



©Shutterstock

L'autonomisation des utilisateurs grâce aux réponses apportées par l'éducation aux médias et à l'information à l'évolution de l'intelligence artificielle générative (IAG)

Messages clés



- L'intelligence artificielle et l'IA générative influent grandement sur les rapports des individus à l'information, aux technologies numériques et aux médias. Mais cela ne va pas sans inquiétudes quant à leur **contrôle**, au pouvoir d'action et à l'autonomie dont disposent les utilisateurs vis-à-vis de l'information, à la prise de décisions, à l'égalité des genres et aux libertés en général. .



- Pour constituer une réponse à l'IAG, technologie qui en est à ses balbutiements, l'**autonomisation des utilisateurs** grâce à l'éducation aux médias et à l'information (EMI) doit être pleinement déployée, et les autorités publiques doivent s'efforcer de la promouvoir dès le début.



- L'EMI est indispensable pour garantir une **utilisation éthique des médias synthétiques**, c'est-à-dire des vidéos, textes, images et sons totalement ou partiellement créés à l'aide de systèmes d'IA.



- L'IAG offre de nombreuses opportunités sociales dans différents domaines, dont l'accès à l'information, **la participation, l'employabilité, la créativité**, l'apprentissage tout au long de la vie et les industries créatives, entre autres.



- Il existe cependant des risques sociaux potentiels aggravés par l'IA générative, dont : **la désinformation**, la perte de contrôle sur la **confidentialité des données**, les menaces pour l'intégrité des élections, **la surveillance**, le manque de **fiabilité des sources**, **la discrimination notamment fondée sur le genre et les stéréotypes raciaux** et les violations des droits d'auteur.



- Pour mettre à profit la **familiarité** face à l'urgence, la maîtrise de l'IA pourrait être intégrée dans les programmes d'EMI afin d'éduquer et de former des communautés très diverses (éducateurs, bibliothécaires, animateurs pour la jeunesse, réseaux de femmes, etc.).



- Pour bien concevoir non seulement les programmes d'EMI mais aussi la gouvernance de l'IAG et les politiques connexes, **l'IA explicable** joue un rôle essentiel.



- Pour renforcer la confiance dans l'information et l'éducation, **la fiabilité des sources** doit être réexaminée pour englober tous les différents types de « preuves » fournies par l'IAG.



- L'EMI peut former des acteurs éclairés n'appartenant pas au secteur de la technologie afin qu'ils contribuent à **la conception, à la mise en œuvre et à la réglementation de l'IA** d'une manière qui reste centrée sur l'humain, sensible au genre et soucieuse de l'intérêt public.



- La formation dans le cadre de l'EMI relève **des gouvernements et des établissements d'enseignement supérieur**, qui doivent veiller à ce que les actions politiques en la matière soient soutenues et renforcées dans la durée, **pour s'adapter aux évolutions constantes de l'IA/IAG**.



Introduction

L'objectif de cette note politique est de montrer comment les parties prenantes peuvent traiter l'intelligence artificielle/l'intelligence artificielle générative (IA/IAG) sous l'angle de l'éducation aux médias et à l'information pour tous afin de limiter les risques et saisir les opportunités qu'elles représentent dans les écosystèmes de l'information et du numérique. L'accent est mis sur l'autonomisation des jeunes et l'importance qui doit être accordée à la diversité culturelle et des genres. Elle propose aussi des recommandations fondées sur les compétences et les principes de l'éducation aux médias et à l'information (EMI) qui pourraient déboucher sur des mesures politiques concrètes. Elle entend également être utile aux organisations de la société civile et à une grande variété de parties prenantes, y compris au sein des médias, des plateformes numériques et des gouvernements.

De nos jours, la plupart des systèmes d'IA reposent sur l'intelligence artificielle générative (IAG). Les systèmes d'IAG peuvent réorganiser statistiquement les vastes quantités de données sur lesquelles ils ont été formatés (textes, images, vidéos, etc.) et ils peuvent produire des résultats conformes aux instructions (requêtes) des utilisateurs. Ils s'appuient sur des modèles de fondation (Jones, 2023) et sont actuellement utilisés par des millions d'individus et d'institutions à travers le monde, sous la forme d'outils comme ChatGPT, qui a gagné 100 millions d'utilisateurs actifs dans le monde en deux mois.

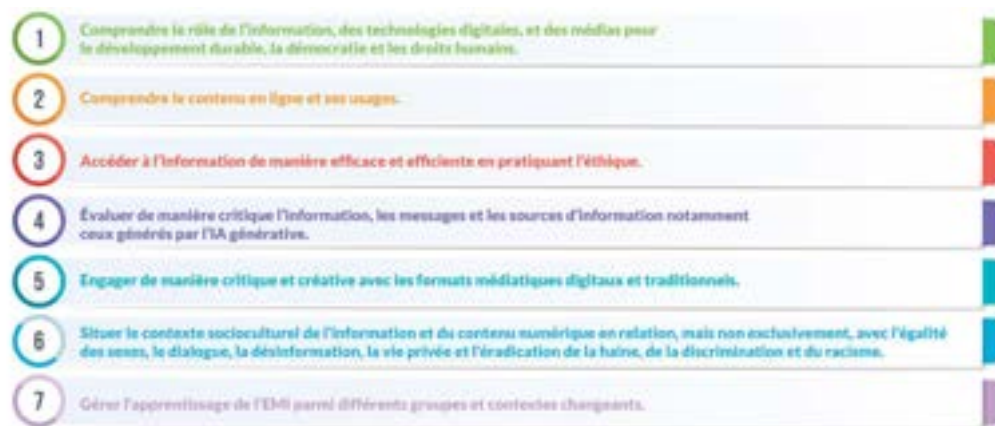
Les systèmes d'IA offrent de nombreux avantages dans presque tous les aspects de notre vie, tels que la santé et l'éducation (Trust et al., 2023), mais ils soulèvent également un certain nombre de préoccupations liées notamment à leur supervision, à leur réglementation et à leur éthique. Les considérations éthiques incitent à surveiller les atteintes aux principes des droits humains, à garantir une utilisation responsable et impartiale, à s'assurer du respect des droits de propriété intellectuelle et à tenir compte des répercussions sociales potentielles (UNESCO, 2022a). Des recherches ont par exemple montré que des outils d'IA et des hypertrucages, comme ceux produits à l'aide de l'IAG, ont généré de la désinformation et des discours de haine en les présentant comme crédibles aux utilisateurs (Ngo et al., 2023). De tels usages pourraient avoir un impact profond sur des moments cruciaux de prise de décision collective, comme lors d'élections ou de consultations publiques.

Depuis le lancement de systèmes d'IAG destinés au grand public, plusieurs dirigeants de grands laboratoires numériques ont appelé à faire « une pause » (Future of Life Institute, 2023) et ont réclamé « une gouvernance de l'IA » (Open AI, 2023), arguant que l'IA représentait un « risque existentiel » pour l'humanité car une super-IA pourrait surpasser les humains et se programmer d'une manière qui irait à l'encontre des intérêts et des valeurs de l'humanité. Cette demande de supervision conduit les régulateurs et de nombreux autres groupes de parties prenantes, mais aussi des particuliers, à entamer un dialogue direct avec les plateformes numériques.

Plusieurs modes de régulation sont envisagés, mais la question de l'autonomisation des utilisateurs est trop peu abordée et sous-estimée. Il est possible d'y remédier en déployant des stratégies et des politiques d'EMI, jugées de plus en plus indispensables pour donner aux individus les connaissances, les aptitudes, les attitudes et les valeurs dont ils ont besoin non seulement pour se protéger des risques, mais aussi pour tirer pleinement parti des opportunités offertes par l'IA. Tout autant préventive que restauratrice, la promotion de l'EMI est une réponse à la nécessité de permettre aux citoyens de participer aux débats actuels sur la place de l'IA dans leurs vies, et donc d'influencer les développements de cette technologie et de demander des comptes sur ses conséquences, qu'elles soient souhaitées ou non.

L'EMI englobe les nombreuses informations et compétences numériques et médiatiques utiles pour maîtriser l'environnement actuel de la communication, de plus en plus complexe, en s'adaptant à son évolution constante. L'EMI permet aux individus de développer leurs capacités de réflexion critique et d'autres compétences indispensables (Figure 1).

Figure 1 : 7 compétences générales et non exhaustives de l'éducation aux médias et à l'information



Source : UNESCO, Citoyens éduqués aux médias et à l'information : penser de manière critique, cliquer à bon escient ! (Programme)

I. L'urgence de l'EMI pour tous face à l'IA générative et aux médias synthétiques



I.1 Pourquoi est-ce important ?

ML'EMI a toute sa place dans le traitement de l'IAG car cette évolution rapide – portée par un ensemble d'applications variées, conviviales, relativement peu coûteuses et faciles d'accès – dans les secteurs de l'information et de la communication met en évidence la nécessité d'éduquer les individus et de leur apprendre à bien utiliser les médias synthétiques et interagir avec des agents non humains dans leurs activités quotidiennes.

Les utilisateurs ont progressivement découvert l'IA étroite conçue pour exécuter des tâches spécifiques (Schlegel et Uenal, 2021), certains de ces systèmes étant d'ailleurs directement liés aux réseaux sociaux ou aux plateformes de diffusion en continu. Ils sont aussi de plus en plus exposés à l'IA générale. L'IAG, forme d'IA générale, s'appuie sur des modèles de fondation tels que les grands modèles de langage (LLM) qui sont pré-entraînés à l'aide d'immenses bases de données regroupant des millions de documents, y compris des contenus provenant des réseaux sociaux (Bryant, 2023). Les utilisateurs doivent également apprendre à gérer les médias synthétiques, c'est-à-dire les vidéos, textes, images ou sons totalement ou partiellement générés par une manipulation informatique ou la modification de données (Stieglitz et al., 2022). En observant la chronologie des évolutions des médias et de l'IA, on constate que la convergence progressive entre ces deux domaines laisse entrevoir un nouveau bouleversement majeur dans l'écosystème de la communication numérique. (Figure 2. Chronologie des médias et de l'IA).



Figure 2: Media and AI @evolutions timeline

Cette convergence aboutit à la naissance des réseaux neuronaux artificiels, qui peuvent modifier leur structure interne selon leur objectif fonctionnel (Grossi et al., 2008) et de l'apprentissage profond, défini comme une structure multicouche qui essaye de reproduire la structure du cerveau humain (Möeller, 2023). Ces technologies sont utilisées pour apporter des améliorations sensibles dans des domaines tels que la reconnaissance et la création d'images, la classification de textes et d'images, l'identification d'objets, l'organisation de données ou encore la recommandation et la prédiction algorithmiques. Dans cette note politique, les termes IA et IAG sont utilisés indifféremment.

Par conséquent, les écosystèmes des médias – et donc l'EMI – sont actuellement influencés par l'émergence de cinq domaines principaux, avec pour chacun un certain nombre d'écueils à éviter : l'interaction avec les consommateurs (assistants virtuels), la prise de décisions (mécanismes de recommandation), l'analyse (analyse des sentiments, apprentissage personnalisé), la prédiction (détection du genre, de l'origine ethnique ou de l'âge des utilisateurs, entre autres caractéristiques) et la communication (médias synthétiques, réalité virtuelle). En ce qui concerne la communication, nous pouvons par exemple faire référence aux systèmes populaires d'IAG présentés comme des « transformateurs génératifs pré-entraînés » (GPT). Le marché comprend désormais plusieurs marques reconnues, dont Google (Bard), Open AI (ChatGPT), Microsoft Copilot (anciennement Bing Chat, qui utilise GPT-4), mais aussi de nouveaux entrants comme MidJourney, Anthropic (Claude 2), Open AI (DALL·E 2) et Stability AI (Stable Diffusion). Ces « assistants virtuels intelligents », conversationnels et répondant à des requêtes écrites sont aujourd'hui intégrés à de nombreux objets du quotidien et à toutes sortes de services d'information. Ils peuvent par exemple aider les utilisateurs à préparer des synthèses, à écrire des scénarios ou à créer des images ressemblant à s'y méprendre à des photographies de scènes qui n'existent pas toujours réellement, et cela les rend difficiles à repérer.

Cette démocratisation pourrait révolutionner l'idée même de l'information comme bien commun – soit en l'amplifiant pour accroître les avantages liés à la connaissance, soit en la privatisant au moyen de « boîtes noires » (UNESCO, 2021a). Ce constat appelle à renforcer d'urgence la littératie de l'IA, pour permettre à chacun de mieux appréhender ses utilisations et interactions quotidiennes non techniques avec des médias de masse, des réseaux sociaux et des médias synthétiques qui influent sur l'information, l'éducation et la culture. Compte tenu de cet objectif, la littératie de l'IA relève de l'éducation aux médias et à l'information. En effet, l'EMI peut faire comprendre les interdépendances entre les médias et les données et s'appuyer sur des méthodes familières pour faciliter l'acquisition de connaissances sur ce nouveau sujet qui évolue rapidement. Cette familiarité pourrait rendre la maîtrise de l'IA moins intimidante pour les éducateurs comme pour les apprenants, en l'inscrivant dans une continuité d'apprentissage et d'enseignement sans nécessiter un grand bond en avant en matière de formation et de perfectionnement des compétences.



1.2 Quelle vision commune adopter ?

Dans le cadre de l'EMI, la réponse à l'IAG consiste à considérer la littératie de l'IA comme un sous-ensemble de littératies imbriquées (couvrant les données, les algorithmes et l'IA elle-même) pour que le processus soit plus familier pour les éducateurs comme pour les apprenants.

Aujourd'hui, la littératie de l'IA en est encore à ses débuts et fait suite à la « littératie des données » et de la « littératie des algorithmes ». La littératie des données vise avant tout à comprendre les données et tend à se concentrer sur la confidentialité et la protection des consommateurs (Nguyen et Beijnon, 2023). La littératie des algorithmes consiste à mettre en lumière ce concept ainsi que la sélection, l'organisation et la présentation des contenus (Droguet, Masur et Joeckel, 2022), en partant du postulat que « la littératie des algorithmes – sous-ensemble de l'éducation à l'information – est l'aptitude à appréhender de manière critique ce que sont les algorithmes, leurs interactions avec les données comportementales humaines dans les systèmes d'information et les enjeux sociaux et éthiques liés à leur utilisation (Head, Fister et MacMillan, 2020).

Un corpus encore restreint porte aussi sur les débuts de la littératie de l'IA en dehors des approches centrées sur l'informatique et réservées aux spécialistes des STIM. Certains chercheurs s'intéressent aux compétences dont les utilisateurs ont besoin pour interagir avec l'IA et pour mettre au point des technologies et des méthodes d'IA axées sur les apprenants (Long et Magerko, 2020). Selon la définition qui se rapproche le plus du concept de l'EMI, la littératie de l'IA renvoie à « l'ensemble des compétences qui permettent aux individus de bien évaluer les systèmes d'IA, mais aussi de communiquer et collaborer efficacement avec ceux-ci. » (Hargittai et al., 2020). En l'occurrence, les quelques exemples récents de cours ayant pour objet la littératie de l'IA sont très similaires, par leur contenu, à la multitude de cours et ressources d'EMI qui existent déjà à travers le monde.

Les chercheurs dont les travaux portent sur ces littératies imbriquées confirment qu'il est urgent de développer l'enseignement et la formation, en mettant en évidence d'importantes lacunes. Ils appellent à multiplier les outils et ressources pour favoriser la littératie des algorithmes et ainsi aider les jeunes et les adultes en général à acquérir les connaissances dont ils ont besoin pour se protéger et préserver les informations qui les concernent

dans l'environnement numérique. D'autres difficultés sont signalées dans le domaine de l'éducation : « (1) les enseignants n'ont ni les connaissances et compétences suffisantes en matière d'IA, ni la confiance nécessaire pour traiter ce sujet ; (2) il n'existe pas de programmes d'enseignement dédiés ; et (3) les directives pédagogiques sont insuffisantes. » (Su et al., 2023).

Ces constats tendent à confirmer les liens forts qui existent entre les données, les algorithmes et l'IA, comme autant de littératies imbriquées à maîtriser. Ils confirment également que ces éléments peuvent s'inscrire dans le paradigme de l'EMI car ils encouragent la réflexion critique sur les données et favorisent une utilisation éthique et sociale de l'information et des systèmes d'IA. Préférable à un traitement indépendant de chaque littératie, leur inclusion dans l'EMI permet de couvrir l'intégralité de la chaîne d'information/communication, du système de production à la consommation par les utilisateurs. Cette approche globale est l'une des caractéristiques de l'EMI en tant que translittératie (Frau-Meigs, 2012). Elle répond à ce que les individus ressentent lorsqu'ils sont confrontés aux cultures de l'information (médias, documents, données), acquièrent des connaissances, forment leurs identités et font leurs choix (Figure 3 : La littératie de l'IA dans le cadre de l'EMI).

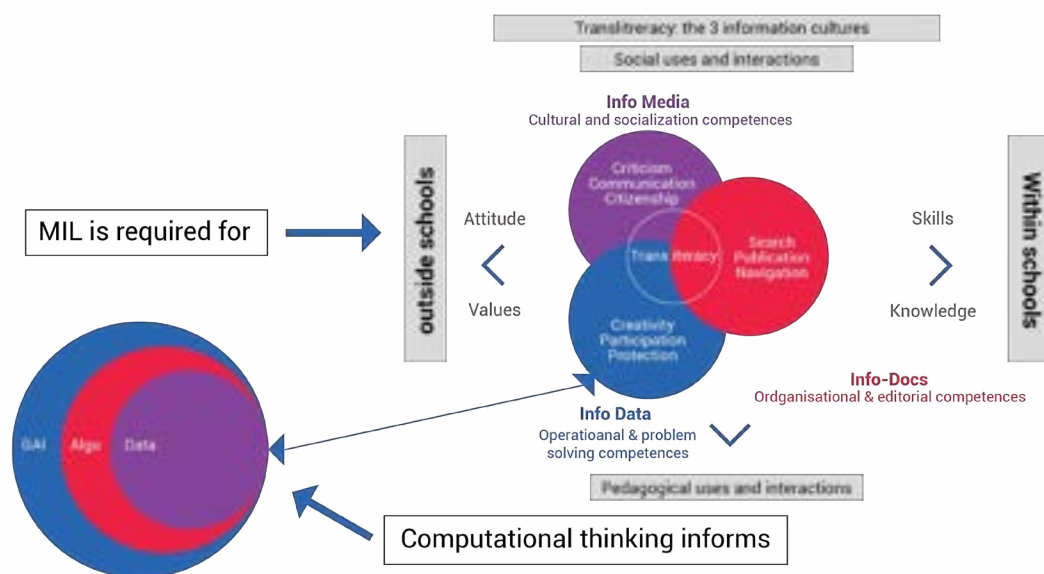


Figure 3 : La littératie de l'IA dans le cadre de l'EMI

Pour favoriser la résilience des utilisateurs, l'EMI peut aider à déterminer ce qu'il faut enseigner et apprendre, et pourquoi, non seulement pour les enfants et les jeunes mais aussi pour les adultes dont les usages quotidiens, les attitudes et les valeurs sont influencés par l'évolution rapide des systèmes d'IA. Cette approche globale incite à adopter une vision commune pour l'élaboration et la mise en œuvre de politiques publiques visant à allouer les ressources appropriées pour permettre l'autonomisation des individus grâce à l'IA/AG.

II. Opportunités de l'EMI liées à l'IA/AG

Dans le cadre de l'EMI, l'expansion de l'IA/AG stimule l'autonomisation des utilisateurs en renforçant leur pouvoir d'action civique (accès à l'information, liberté d'expression, utilisation éthique des médias synthétiques) et en favorisant leur employabilité (apprentissage tout au long de la vie, industries créatives, etc.).

L'éducation aux médias et à l'information a souvent tiré sa légitimité de son rôle en faveur de la citoyenneté. En effet, elle augmente le pouvoir d'action civique et favorise l'utilisation éthique des médias et de l'information, pour une meilleure participation aux affaires sociales. Parallèlement, la nature productive et participative des médias rend l'EMI de plus en plus pertinente pour une autre raison : elle contribue à l'employabilité sur un marché du travail qui évolue rapidement et dans lequel la créativité et les interactions homme-média-machine occupent une grande place. L'employabilité, définie comme « l'aptitude (perçue) de chacun à trouver et conserver un emploi tout au long de sa carrière » (Römgnès, Scoupe et Beusaert, 2020) dépend de la préparation éducative, qui peut conduire à entreprendre toutes sortes d'activités. Ces deux volets, le pouvoir d'action civique et l'employabilité, sont profondément modifiés par l'IA/AG, et modifient à leur tour notre manière d'appréhender et de pratiquer l'EMI.

La généralisation rapide du recours aux systèmes d'AG, leur commercialisation et leur prétendue démocratisation peuvent être vues comme des atouts en faveur de la citoyenneté et de l'employabilité dans le cadre de l'EMI qui fait appel à la réflexion critique et créative. Ces systèmes peuvent stimuler la créativité

lorsqu'il s'agit de produire des contenus médiatiques, en passant du copier-coller au copier-cr  er, sans obstacle majeur    la cr  ation et    la diffusion de contenus. Ils peuvent accro  tre les interactions avec des robots, des capteurs, des agents conversationnels et des assistants virtuels.

Cette d  mocratisation a aussi des r  percussions sur des domaines li  s    l'EMI, dont les sciences de l'information et de la communication, la biblioth  conom   et les sciences de la documentation, le journalisme, la science des donn  es et les industries cr  atives. Elle a des cons  quences profondes sur les individus et la soci  t   dans son ensemble, car des m  dias num  riques de plus en plus personnalis  s se d  veloppent et changent la mani  re dont les utilisateurs communiquent entre eux, pour apprendre, travailler ou se divertir.



II.1 Pouvoir d'action civique

Les r  ponses apport  es par l'EMI face aux syst  mes et services d'IA peuvent renforcer l'autonomie des citoyens et le pouvoir d'action des utilisateurs dans diff  rents domaines, tels que la recherche d'information, le rapport aux r  seaux sociaux et aux m  dias synth  tiques, la v  rification et l'annotation collaboratives, ou encore la reconnaissance vocale et la traduction (Tableau 1 : Progr  s de l'IAG et r  ponses de l'EMI en faveur du pouvoir d'action civique), mais aussi en ce qui concerne les politiques associ  es.

Progr��s	Domaines politiques	R��ponses de l'EMI
Recherche d'informations et acc��s �� celles-ci	Contr��le de la qualit�� des donn��es	Donner aux utilisateurs les moyens d'obtenir des r��sultats de recherche qui refl��tent la diversit��, notamment en ce qui concerne la prise en compte du genre
Participation sur les r��seaux sociaux	Libert�� d'expression	Favoriser l'utilisation ��thique de ces outils et la capacit�� �� y promouvoir et prot��ger les droits humains, notamment la libert�� d'expression et l'��galit�� des genres
Cr��ation de m��dias synth��tiques	Cr��ativit��, innovation	Encourager l'esprit critique et l'utilisation ��thique
Reconnaissance vocale, vision artificielle	Accessibilit�� des personnes en situation de handicap, lutte contre les discriminations	Permet de comprendre les lacunes – les pertes et les gains lors des interactions avec les technologies num��riques
Transcription, cr��ation d'images �� partir d'un texte, traduction automatique	Inclusion, dialogue interculturel, diversit�� culturelle	Souligner l'int��r��t du passage de textes d'un support �� l'autre (ou d'une langue �� l'autre), rep��rer les dangers li��s au changement de support et �� la modification du message
Recoupement de multiples sources, v��rification collaborative, annotation	Participation, contribution	Mettre en valeur le r��le que peuvent jouer les technologies num��riques pour la recherche, donner aux citoyens les moyens d'��viter les probl��mes caus��s par l'IA pendant les ��lections et d'autres ��v��nements publics importants

Tableau 1 : Progr  s de l'IAG et r  ponses de l'EMI en faveur du pouvoir d'action civique

Plusieurs acteurs de l'EMI esp  rent tirer parti du d  veloppement et de la d  mocratisation de l'IA, au premier rang desquels les diffuseurs (publics), les journalistes, les   ducateurs et les biblioth  caires confront  s    une surcharge d'informations (sous la forme d'actualit  s, documents et donn  es). Les   diteurs, les diffuseurs et les plateformes de m  dias en ligne peuvent h  berger et distribuer des contenus synth  tiques cr   s par les utilisateurs. Ils peuvent aussi diffuser leur propre production et utiliser l'IA dans diff  rents domaines qui prennent de l'ampleur tels que la recherche, l'aide    l'indexation, la transcription, etc., ainsi que pour la gestion des contenus multicanaux qui se multiplient.

Les journalistes effectuent un travail complexe et indispensable pour la soci  t  , que l'IA peut faciliter de multiples mani  res. Par exemple, elle peut leur   viter d'avoir    effectuer des t  ches r  p  titives qui n  cessitent de manipuler de vastes quantit  s de donn  es. Les bulletins m  t  orologiques localis  s pourraient ainsi   tre pr  sent  s par des avatars. Les journalistes peuvent mieux r  pondre aux attentes de leurs lecteurs en menant des investigations longitudinales plus complexes ou en s'appuyant sur des bases de donn  es pour mettre au jour des tendances de fond telle que l'  vasion fiscale dans des paradis fiscaux. C'est ce qu'a r  v  l   l'enqu  te dite des « Pandora Papers » dirig  e par le Consortium international des journalistes d'investigation. Les journalistes peuvent proposer aux citoyens des reportages critiques au sujet de l'IA,

qui explorent par exemple les rapports de force entre les plateformes numériques et les autorités ou bien traitent des risques d'exclusion numérique ou des atteintes aux droits humains en lien avec l'utilisation de l'IA (UNESCO 2023b).

Les éducateurs (enseignants, bibliothécaires, animateurs pour la jeunesse) constituent une autre communauté de l'EMI qui peut grandement tirer parti de la possibilité de recouper les sources, de traduire automatiquement ou d'examiner et annoter collaborativement des documents. Les outils reposant sur l'IA qu'utilisent les journalistes et les vérificateurs d'informations s'ouvrent désormais aux enseignants, aux chercheurs et aux militants de la société civile. Ces acteurs savent de mieux en mieux lutter contre la désinformation et promouvoir la citoyenneté, comme le montre le projet Crossover (<https://crossover.social>). Il apparaît donc que l'intégration de l'EMI dans les politiques et la programmation des médias pourrait contribuer à rétablir la confiance dans les médias et le journalisme, et renforcerait ainsi leur viabilité et leur pérennité. Ce serait aussi un moyen d'encourager les influenceurs à diffuser sur les réseaux sociaux des messages positifs et éthiques au sujet des évolutions de l'IAG (UNESCO, 2019).

II.2 Employabilité

Parallèlement, l'intégration de l'EMI dans les formations des entrepreneurs numériques pourrait leur être utile en ligne. Les générations dont les emplois vont être transformés plus rapidement que prévu doivent adapter leurs compétences, sous peine d'être marginalisées. Les plus exposées à ces risques sont les femmes, suivies des membres des minorités et des personnes ayant des revenus inférieurs (OCDE, 2023). Les compétences que l'EMI permet d'acquérir – aussi bien en ce qui concerne la gestion de l'information et la réflexion critique que la créativité – peuvent donc combler le fossé actuel entre la formation et l'emploi, qui nuit aux conditions de travail dans un milieu professionnel de plus en plus fondé sur les médias et les TIC. Dans le monde du travail, elles viennent s'ajouter aux compétences en TIC car de nombreuses entreprises et organisations effectuent la transition vers la littératie numérique. Elles peuvent favoriser l'inclusion et lutter contre la pauvreté, le dénuement et la marginalisation, si elles sont assorties de politiques adaptées.

Les réponses de l'EMI aux systèmes et services d'IAG peuvent renforcer l'employabilité des individus dans différents domaines et améliorer les politiques associées, notamment en ce qui concerne l'enseignement supérieur, l'apprentissage en ligne et les industries créatives (UNESCO, 2023b). L'EMI est elle-même modifiée par les médias synthétiques car les systèmes d'IA permettent de « médiatiser » les apprentissages, notamment en ligne, par exemple avec des plateformes de création de vidéos au moyen de l'IA telles que Synthesia ou Pictory (Tableau 2 : Progrès de l'IAG et réponses de l'EMI en faveur de l'employabilité).

Progrès	Domaines politiques	Réponses de l'EMI
Analyse de l'apprentissage	Apprentissage tout au long de la vie, éducation continue	Faire connaître les droits relatifs aux données, les mécanismes de protection et de recours ; détecter précocement les problèmes d'apprentissage
Technologies pédagogiques : visualisation des données, avatars, ludification	Conception de programmes et de cursus visant à soutenir les enseignants et les apprenants	Favoriser l'apprentissage logique grâce à de nombreux médias fondés sur le raisonnement, et le rapport entre causes et effets, apprendre en s'appuyant sur les interactions entre les contextes formels et informels
Adaptation aux différents types d'apprenants	Inclusion, diversité	Mettre en lumière l'ensemble des compétences numériques
Retour d'information renforcé	Participation, contribution	Encourager l'analyse critique des contributions du numérique (gains et pertes) ; attirer l'attention sur l'utilité de la cocréation de contenus pour l'EMI et de contenus en général
Grand nombre de pédagogies personnalisées	Éducation, apprentissage tout au long de la vie	Favoriser la prise en compte des interactions entre le milieu et les expériences des individus d'une part et les messages et supports utilisés d'autre part ; promouvoir la coconception de projets d'apprentissage (y compris dans le cadre de l'EMI)
Plateformes de création de contenus (musique et vidéo)	Industries créatives, culture	Stimuler les interventions et la participation des utilisateurs dans le domaine culturel, l'engagement du public et notamment des femmes

Tableau 2 : Progrès de l'IAG et réponses de l'EMI en faveur de l'employabilité

Plusieurs acteurs de l'EMI peuvent tirer parti de ces progrès de l'IA, dans l'éducation comme dans la culture. Les formateurs en EMI qui interviennent dans les écoles et les universités peuvent aider les apprenants à reconnaître les médias synthétiques, et cela peut supposer de leur apprendre aussi à maîtriser les requêtes, autrement dit à formuler des demandes qui donneront des résultats utiles. Une telle exposition aux contenus produits à l'aide de l'IA peut se faire par le biais de pédagogies actives qui reposent sur une approche pratique des outils mis à disposition pour l'enseignement et l'apprentissage, comme ceux utilisés dans le cadre de la conception de programmes d'EMI. Ils peuvent renforcer l'employabilité en aidant les élèves et les éducateurs à prendre conscience de la nécessité de constamment perfectionner leurs compétences dans le monde du travail numérique.

Les artistes, les designers et les professionnels de la création qui travaillent dans les secteurs de la télévision, du cinéma et des jeux voient leurs possibilités créatives s'élargir, car les systèmes d'IA leur donnent une plus grande visibilité et lisibilité de l'activité cérébrale. Ils peuvent explorer de nouveaux rapports à la mémoire ou à l'imagination, et disposent de moyens d'interagir avec les utilisateurs de manière très inclusive. Les générateurs de textes qui font appel à l'IA proposent des services de personnalisation pour aider les auteurs à améliorer la qualité des contenus qu'ils publient en ligne en analysant des échantillons de leurs propres écrits. Des œuvres de fiction et d'art synthétiques peuvent être créées en remplaçant les palettes par des algorithmes et des écrans, et les nouveaux arrivants – y compris les amateurs – peuvent ainsi accéder à la production et à la consommation culturelles.

Ces différentes réponses offertes par l'EMI indiquent que les professionnels comme les utilisateurs en général doivent s'adapter aux différents médias et modifier leurs interactions face à des situations, expériences et outils variés. Les institutions culturelles et les établissements d'enseignement doivent évoluer en conséquence, pour éviter les pénuries de personnel qualifié et rendre accessibles au plus grand nombre différentes formes de contribution créative et de création de contenus. C'est l'occasion, pour les programmes d'EMI déployés dans les écoles, les universités, les bibliothèques et les musées comme dans les milieux professionnels, de combler une fracture numérique spécifique : celle qui existe entre les STIM et les autres domaines, notamment ceux des lettres et des arts, où la notion d'humanité tient une place centrale..

III. Les défis et risques de l'IA/IAG pour l'EMI

Dans le cadre de l'EMI, l'IAG peut aussi nuire à l'autonomisation des utilisateurs en affaiblissant la citoyenneté (enjeux liés à désinformation, la surveillance, au respect de la vie privée ou encore à l'éthique) et limiter leur employabilité (enjeux liés aux pseudo-sciences, à la fiabilité des sources, aux droits d'auteur ou encore à la propriété intellectuelle).

III.1 Pouvoir d'action civique



L'une des principales préoccupations liées à l'IAG découle du fait qu'il s'agit d'une technologie qui fait appel à des expérimentations algorithmiques sans réelle compréhension, c'est-à-dire qui reposent uniquement sur la puissance de calculs informatiques. Cela s'oppose aux expérimentations de l'intelligence humaine, principalement motivées par la recherche de la compréhension. Autre inquiétude concernant les médias synthétiques, l'acte de création est déconnecté du processus de réflexion et des émotions que suscitent l'écriture et la visualisation. Les efforts nécessaires à la consolidation et au transfert des connaissances humaines qui étaient – et sont encore – mobilisées lors de cet acte de création ne doivent pas être sous-estimés, et encore moins disparaître. Ces valeurs devraient être prises en compte dès la conception des systèmes d'IA, au même titre que le respect des droits humains universels.

Le principal risque lié à l'information est celui de la désinformation produite et diffusée par le biais de l'IAG, qui peut être automatisée grâce à des hypertrucages intraquables, en quelques clics seulement. Cela peut provoquer d'autres problèmes, dont la production de contenus automatisés qui ne sont pas sensibles au genre et s'appuient sur des préjugés raciaux et qui ont donc pour effet de renforcer et de perpétuer les stéréotypes liés au genre et le profilage racial (Ngo et al., 2023). Les facteurs qui alimentent la désinformation genrée et ses effets négatifs (Nations Unies, 2023), la personnalisation des actualités (Van Drunen et al., 2022), les discours de haine (Henderson et al., 2023), la diffusion non consensuelle d'images intimes (Garon, 2023) ou le vol de données (Wach et al., 2023), mais aussi les risques collectifs tels que la fraude électorale (Srivastava et al., 2023) et la surveillance de masse font également partie des enjeux à traiter, et les réponses apportées par l'EMI relèvent alors de la protection des droits humains.

Les réponses de l'EMI aux risques liés à l'IAG peuvent renforcer l'autonomie des utilisateurs dans différents domaines et améliorer les politiques associés, notamment en ce qui concerne la liberté d'expression, la confidentialité des données, la concentration commerciale et la surveillance (Tableau 3 : Risques liés à l'IAG et réponses de l'EMI). Ces risques existaient déjà avant l'émergence de l'IA, mais l'IAG les a exacerbés et rendus plus difficiles à traiter (UNESCO, 2023c).

Risques	Domaines politiques	Réponses de l'EMI
Risque existentiel/prolifération	Éthique de la concurrence géopolitique	Mettre en avant l'autogestion et la cogestion ; prendre conscience du déséquilibre des valeurs, notamment en ce qui concerne l'égalité des genres
Concentration/monopole	Inégalité des chances, absence de concurrence loyale et donc de pluralisme dans l'écosystème numérique	Améliorer la compréhension de l'importance de la diversité et du pluralisme dans les secteurs de l'information et des technologies numériques
Spams, publicité intrusive	Manque de données et de contenus de qualité	Expliquer aux utilisateurs comment exercer des recours et demander des comptes aux acteurs
Personnalisation des actualités	Manque de diversité	Alerter sur les dangers des bulles de filtres et des chambres d'écho
Désinformation/hypertrucage	Érosion de la confiance, atteinte à la liberté d'expression et réduction de l'accès à des informations fiables	Apprendre à employer diverses stratégies de vérification des faits
Biais algorithmique	Racisme, manipulation sociale, inégalités, discrimination	Aider les utilisateurs à repérer les biais, dont certains sont intrinsèques, et à identifier des techniques d'atténuation
Discours de haine et cyberharcèlement	Diffamation, atteinte à la liberté d'expression	Aider à démasquer les idées fausses, proposer des contre-discours et défendre la liberté d'expression, le dialogue, la tolérance et l'éthique
Diffusion non consensuelle de contenus à caractère sexuel	Diffamation par l'image et maltraitance, atteinte à la vie privée	Mettre en avant l'autogestion et la cogestion, ainsi que les recours que la réglementation offre aux utilisateurs
Fraude électorale	Atteinte à la démocratie et à l'intégrité des électeurs	Sensibiliser les utilisateurs, leur donner les moyens de s'impliquer, d'alerter, de vérifier des faits et de se protéger
Vol de données	Violation de données à caractère personnel, perte d'emploi en raison de l'automatisation	Faire connaître les droits relatifs aux données, doter des compétences nécessaires à la protection de la vie privée et expliquer comment demander réparation
Surveillance de masse	Atteinte à la vie privée et à la sécurité, censure	Favoriser la prise de conscience, l'indépendance et la transparence
Empreinte environnementale	Développement durable	Faire prendre conscience de la pollution causée par les éléments matériels de l'IA (minerais, puces, serveurs de données, etc.)

Tableau 3 : Risques de l'IAG et réponses de l'EMI en ce qui concerne le pouvoir d'action civique

Il n'existe aucun dispositif infallible permettant d'entraîner un système d'IA pour qu'il refuse systématiquement de nuire ou de produire des contenus relevant de la désinformation. Des chercheurs ont montré qu'il était possible de débriquer ces systèmes – des requêtes spéciales peuvent donner lieu à des réponses inattendues – et sont parvenus à contourner les mécanismes de protection mis en place par les fournisseurs de LLM (Zou et al., 2023). Ce manque de garanties expose à des risques de nombreux acteurs de l'EMI, dans la mesure où les journalistes, les éducateurs et les contributeurs aux cultures médiatiques en général sont confrontés à des incertitudes quant aux contenus auxquels ils accèdent et qu'ils utilisent pour prendre des décisions. L'IA peut donc saper la cohésion sociale et entraver l'exercice des droits collectifs. L'engagement et le pouvoir d'action des utilisateurs sont aussi remis en question, parce que la confiance s'érode et que la polarisation des publics et des communautés pourrait s'amplifier dans des chambres d'écho synthétiques.

Les acteurs de l'EMI déploient des moyens variés pour donner aux utilisateurs les moyens de contrer ces phénomènes. Ils peuvent compter sur des dispositifs de surveillance des médias tels que NewsGuard, qui contrôle les médias synthétiques grâce à sa plateforme de suivi de l'IA. En septembre 2023, il avait déjà identifié 450 sites d'information et d'actualité générés par l'IA et gérés avec peu voire pas de supervision humaine (<https://www.newsguardtech.com/fr/>). NewsGuard met aussi ses outils à la disposition des écoles, des bibliothèques publiques et des universités pour favoriser l'EMI.

Les utilisateurs peuvent également choisir des médias qui font preuve de transparence en ce qui concerne leur utilisation de l'IA. Certains groupes de presse publient leurs directives en la matière ou adoptent des labels

permettant d'identifier le degré de recours à l'IA (Shane et al., 2021). Il est nécessaire de mettre en place des normes éthiques, et ce constat a donné lieu à plusieurs initiatives dont Partnership on AI (PAI), qui a travaillé avec des représentants de plateformes technologiques, des membres de la société civile, des journalistes et des universitaires pour publier un ensemble de recommandations, intitulé Responsible Practices for Synthetic Media. L'objectif est d'encourager le développement et le déploiement responsables des médias synthétiques (Partnership on AI, 2023).

Les utilisateurs peuvent également faire appel à des solutions de rétro-ingénierie, c'est-à-dire à des systèmes d'IA conçus pour détecter l'IA. Par exemple, GPTZero repère les utilisations de ChatGPT (<https://gptzero.me/>). Ils peuvent aussi compter sur des équipes de journalistes, de vérificateurs des faits et d'analystes de données qui ont mis au point des outils reposant sur l'IA pour lutter contre la désinformation et les « comportements inauthentiques coordonnés » (campagnes de désinformation lancées de manière coordonnée sur les réseaux sociaux). Par exemple, Vera.ai est un outil qui s'appuie sur l'IA pour détecter les hypertrucages et proposer différents services, y compris un mécanisme d'analyse de fichiers audio et un répertoire des contenus déjà démasqués (<https://www.veraai.eu/>).

Les utilisateurs peuvent également suivre des formations en ligne, comme celle proposée par l'Université d'Helsinki (Finlande), Elements of AI, qui a été consultée par des participants venus de 170 pays (elementsofai.com). L'UNESCO a lancé la formation Défendre les droits humains à l'ère de l'intelligence artificielle (<https://www.edapp.com/course/defendre-les-droits-humains-l-ere-de-l-intelligence-artificielle/>), disponible en 25 langues ; un module dédié à l'EMI et l'IA figure également dans sa publication Citoyens éduqués aux médias et à l'information : penser de manière critique, cliquer à bon escient !, programme d'EMI pour les éducateurs et les apprenants (UNESCO, 2021c).



III.2 Employabilité

Les risques de l'IAG pour l'employabilité sont plus difficiles à déterminer. Ils peuvent être directs, liés à la qualité des informations utilisées pour l'enseignement et la formation (pseudo-sciences, plagiat remettant en question la validité des diplômes et des qualifications, etc.), mais aussi découler des effets de l'économie de l'attention sur les utilisateurs. Ils peuvent être indirects, au travail, dans les industries créatives, en raison du manque de compétences et de l'insuffisance du perfectionnement, associés à une certaine dévalorisation du domaine des lettres par rapport aux STIM. Les conditions de travail peuvent également être affectées par la collecte de données, un moindre respect de la vie privée ou la surveillance au travail, sans oublier les risques de perte d'emploi ou de réorganisation des fonctions en raison de l'automatisation. En outre, la fracture entre les genres pourrait se creuser. C'est ce qu'indique le rapport de l'UNESCO intitulé Les effets de l'IA sur la vie professionnelle des femmes (UNESCO, 2022b), en soulignant par exemple que les assistants vocaux virtuels sont encore très souvent féminisés.

Les évolutions qu'entraîne l'IAG obligent les travailleurs à acquérir de nouvelles compétences numériques à travers la formation continue et l'apprentissage tout au long de la vie. L'insuffisance des compétences en matière de réflexion critique, civiques et créatives peut être un frein susceptible de maintenir des individus dans la pauvreté ou de les marginaliser. Par ailleurs, l'utilisation abusive de l'information, y compris la pseudoscience, les « hallucinations » (Wach et al., 2023), la désinformation (Lo, 2023), le plagiat et la violation du droit d'auteur (Vincent, 2022), peut entraver l'acquisition d'attitudes éthiques à l'égard des nouvelles, des données et des documents (Figure 4 : Risques de l'IAG et réponses de l'EMI en ce qui concerne l'employabilité).

Risques	Domaines politiques	Réponses de l'EMI
Robots et automatisation	Perte d'emploi, réorganisation des fonctions	Faire connaître les avantages et les limites, d'où la nécessité d'un contrôle et d'une évaluation
Préjugés dans les processus de recherche d'emploi et de recrutement	Inégalités entre les genres, stéréotypes	Doter les individus des compétences nécessaires pour repérer la discrimination liée au genre et les représentations stéréotypées dans les informations et les contenus
Surveillance des performances au travail	Inégalités entre les genres, atteinte à la vie privée	Permettre de repérer la fracture entre les genres et la perpétuation des stéréotypes, alerter au sujet du respect de la vie privée
Plagiat	Atteinte à l'intégrité académique	Promouvoir des utilisations éthiques de l'information et des technologies numériques
Hallucinations et erreurs de l'IA	Recherche et éducation	Mettre en avant l'autogestion et la cogestion, lutter contre les mythes et les imaginaires de l'IA

Pseudo-sciences	Recherche et éducation	Raffermir la confiance dans les recherches scientifiques en établissant des liens avec les connaissances scientifiques fondamentales
Fiabilité des sources	Inégalité des chances, absence de concurrence loyale	Souligner la possibilité d'évaluer le cycle de vie de l'information et d'y contribuer
Propriété intellectuelle et droits d'auteur	Usurpation de la qualité d'auteur, perte d'une rémunération équitable	Favoriser l'utilisation éthique et promouvoir le respect de la qualité d'auteur et des droits de propriété intellectuelle

Tableau 4 : Risques de l'IA et réponses de l'EMI en ce qui concerne l'employabilité

Pour contrer ces risques, les acteurs de l'EMI déploient des moyens variés. Fondamentale pour l'accès à des informations de qualité et le maintien d'un environnement en ligne sain, la fiabilité des sources est devenue un enjeu de l'EMI, associée à l'écriture critique, la vérification des faits et l'indexation fiable. La fiabilité des sources est essentielle pour renforcer la confiance dans l'apprentissage, l'étude et la recherche, dans le milieu universitaire mais aussi dans d'autres domaines comme le journalisme. Elle peut avoir des répercussions directes sur l'employabilité si les travailleurs ne sont pas formés à produire des contenus fiables, vérifiés et fondés sur des preuves.

Les enseignants, les bibliothécaires, les rédacteurs et les éditeurs regrettent l'absence de systèmes ou de normes fiables qui pourraient les aider, ainsi que leurs élèves, à indiquer clairement leurs usages des outils d'IA et à citer correctement leurs sources. Pourtant, l'IA crée des sources d'un nouveau genre, qui s'ajoutent aux sources primaires et secondaires et pourraient être appelées « sources tertiaires », pour les distinguer des deux premiers moyens traditionnels d'apporter des preuves et pour les identifier comme étant partiellement ou totalement générées par l'IA. Certaines solutions, essentiellement proposées par les plateformes numériques, sont destinées à avertir les utilisateurs. Elles prennent par exemple la forme de mentions jointes aux contenus, de filigranes ou de notes contextuelles au sujet des contenus synthétiques dont la création et/ou le partage sont autorisés, mais il n'existe pas de règle commune à toutes les plateformes.

La question de la fiabilité et de l'intégrité des sources soulève un autre problème, lié aux violations des droits d'auteur. En effet, les contenus qui servent à entraîner les systèmes d'IA proviennent de bases de données alimentées sans le consentement explicite des créateurs de ces données, qui ne sont pas rémunérés pour le temps et l'expérience qu'ils y consacrent. Des inquiétudes quant au respect des droits d'auteur et de la propriété intellectuelle sont émises par les professionnels de la création, qui s'opposent à l'utilisation de l'IA. À titre d'exemple, on peut citer la campagne « No to AI generated images » ou encore la longue grève menée par la Writers Guild of America, syndicat de scénaristes américains, pour protester contre les effets effrayants de ces pratiques sur leurs emplois et leurs carrières. Les membres de cette profession craignent de voir leur travail détourné et surexploité sans juste rémunération.

Ces enjeux éthiques sont traités par l'EMI car les jeunes utilisateurs de médias doivent comprendre que le travail créatif dans les sociétés de l'information et du savoir repose sur une rémunération équitable, en particulier au regard de leur future employabilité. La protection des droits de propriété intellectuelle et la fiabilité des sources sont indispensables pour la mise en place d'un système de communication sain et pour la reconnaissance de l'information comme bien commun.

IV.. Éléments à prendre en compte pour une EMI qui donne aux utilisateurs les moyens d'appréhender l'IA/IAG : sensibilisation et compétences

L'EMI peut sensibiliser les utilisateurs et renforcer leur réflexion critique et créative autour de différents enjeux : la fracture numérique, y compris entre les genres, la concentration et la propriété de l'IA, les imaginaires de l'IA ou encore les interactions avec des médias synthétiques, des agents non humains et des robots.

Les stratégies de l'EMI doivent viser à faire prendre conscience des risques posés par les IA entraînées, volontairement ou non, à produire des contenus nocifs ou à nuire d'une toute autre façon, c'est-à-dire en agissant de manière à porter atteinte aux valeurs et aux droits humains). Elles respectent les principes des droits humains et renforcent le pouvoir d'action civique des utilisateurs pour éviter le plus grand des dangers, la remise en question des sociétés démocratiques et de l'autonomie humaine. Pour que l'EMI soit prête à traiter les défis actuels et futurs liés à l'IA, de vastes campagnes de sensibilisation du grand public doivent être lancées, peut-être au moyen d'outils et de plateformes faisant appel à l'IA pour faciliter la traduction et la localisation des contenus. Le déploiement de programmes d'EMI pour tous à grande échelle, par exemple sous forme de cours en ligne, peut aussi être utile à de nombreux individus de tous âges et de toutes nationalités. En définitive, l'EMI axée sur l'IA devrait être considérée comme un savoir fondamental, que les jeunes doivent acquérir du début du primaire à la fin du secondaire pour être en mesure de vivre, d'apprendre, de travailler et de créer dans un environnement numérique.

Par ailleurs, d'importantes lacunes persistent en ce qui concerne la conception et la mise en œuvre de programmes d'EMI tenant spécifiquement compte de l'égalité des genres, à même de favoriser des interventions transformatrices du genre (en particulier la création de supports de sensibilisation visant à lutter contre les préjugés de genre perpétués par l'IA). Ce constat s'applique aussi aux formations portant sur les compétences numériques, comme cela est démontré dans la publication *Je rougirais si je pouvais* (UNESCO, 2019), et aux nouveaux cursus axés sur la littératie de l'IA. Les utilisateurs doivent être en mesure d'analyser et de remettre en question les écarts entre les hommes et les femmes en ce qui concerne leur utilisation de l'IA et leur contribution au développement de cette technologie (UNESCO, 2020). L'EMI peut faciliter la prise de conscience quant aux répercussions genrées de l'IA, et donner aux citoyens les moyens de plaider en faveur des applications positives et des avantages de l'IA pour l'autonomisation des femmes. La conception d'un cadre de compétences sur l'IA qui serait compatible avec les principes de l'EMI suppose d'énumérer les compétences en EMI nécessaires pour maîtriser l'IA/IAG et de les répartir en quatre catégories : connaissances, aptitudes, attitudes et valeurs. L'objectif est de permettre aux éducateurs et aux décideurs d'évaluer l'impact de ces interventions et d'établir des références en la matière, grandement nécessaires (Tableau 5 : Cadre de compétences sur l'IA, compatible avec les principes de l'EMI).

1/ connaissances : identification de l'IA (potentiel et risques), reconnaissance de la géopolitique de l'IA (acteurs, propriété, motivations), évaluation des conséquences et proposition de réponses à apporter aux avantages pour l'information (LLM, agents conversationnels et assistants virtuels) et aux problèmes (discours de haine, désinformation, etc.), surveillance de l'empreinte écologique des systèmes d'IA.

2/ aptitudes : utilisation responsable des outils reposant sur l'IA, navigation sur des sites et des bases de données de qualité, comparaison des moteurs de recherche, distinction entre les médias de masse, les réseaux sociaux et les médias synthétiques, interaction avec des objets et des agents non humains, organisation et vérification des sources, maîtrise des requêtes et des outils de création.

3/ attitudes : interprétation critique et créative des données, de leur utilisation et des effets des algorithmes, identification et signalement des contenus nocifs (violence, discours de haine, pornographie, désinformation, etc.), lutte contre des risques systémiques (manipulation, viralité, automaticité, atteinte à la confidentialité des données, vol de données, etc.), reconnaissance des idées reçues, des mythes, des représentations et des imaginaires personnels, élaboration de contre-discours face à la désinformation et communication.

4/ valeurs : défense de la liberté d'expression et d'opinion, promotion de l'égalité des genres, engagement en faveur du respect de la vie privée, de la participation et de la propriété intellectuelle, promotion des informations et des bases de données de qualité, opposition à la surveillance de masse et recherche de la sécurité et du bien-être.

Tableau 5 : Cadre de compétences sur l'IA, compatible avec les principes de l'EMI

V. Promotion et protection de l'EMI dans le cadre de la gouvernance de l'IA/IAG : réponse des politiques publiques

L'EMI peut contribuer aux politiques publiques et aux principes qui les sous-tendent en donnant aux utilisateurs les moyens de participer de manière proactive à la surveillance et à la réglementation de l'IA explicable, mais aussi en les encourageant à dialoguer avec d'autres parties prenantes pour améliorer la gouvernance de l'IA et bâtir des sociétés du savoir saines.

Les plateformes numériques ont certes fait d'immenses progrès dans le domaine de la sécurité de l'IA, mais les garanties mises en place peuvent toujours être contournées, volontairement ou non, et permettre un usage contraire à l'éthique. L'insuffisance de la supervision humaine et des contrôles peut faire naître une méfiance néfaste vis-à-vis de l'informatique, parmi les institutions et les acteurs qui en font la promotion, mais aussi dans les médias professionnels et d'autres organismes de gouvernance établis (gouvernements, entités étatiques, organisations intergouvernementales, etc.). Les réponses actuelles des autres parties prenantes montrent les cadres émergents de la politique de gouvernance de l'IA aux niveaux mondial et national, car les plateformes d'IA et les entreprises technologiques résistent ou font appel à la corégulation.

Plusieurs cadres sont en cours d'élaboration, et adoptent une perspective multipartite plus large. Les Nations Unies sont en train de constituer un Organe consultatif multipartite sur la coopération mondiale en matière d'intelligence artificielle. Cela s'inscrit dans le Plan d'action de coopération numérique établi par le Secrétaire général (A/74/821). Il fournira des recommandations pour optimiser la gouvernance internatio-

nale de l'intelligence artificielle. De son côté, l'UNESCO a mis au point les Principes pour la gouvernance des plateformes numériques : préserver la liberté d'expression et l'accès à l'information - une approche multipartite (UNESCO, 2023d). Le volet de ces principes consacré à l'autonomisation des utilisateurs donne lieu à l'élaboration d'un Plan d'action multipartite pour l'intégration de l'éducation aux médias et à l'information dans les plans stratégiques, politiques et produits des plateformes numériques (UNESCO, 2023e, non publié).

La plupart des cadres, excepté celui de l'UNESCO, mettent l'accent sur la gestion des risques et la transparence, en négligeant les enjeux liés à l'EMI ou en mentionnant très vaguement et généralement l'éducation comme moyen centré sur l'humain pour intégrer l'IA dans la société, sans détailler les questions de mise en œuvre ou de financement. Dans ce cadre, les principes qui renvoient le plus à l'autonomisation des utilisateurs et à l'EMI sont ceux de la responsabilité et de l'explicabilité. Ils sont les plus à même de fournir les contrôles et contre-poids nécessaires pour contrer les risques liés à la gouvernance de l'IA : la dégradation de cette gouvernance en raison de résultats non désirés, l'instabilité interne et les menaces existentielles pour l'humanité, y compris le transfert des pouvoirs décisionnels des gouvernements à des robots et aux humains qui les contrôlent.

Les exigences en matière d'explicabilité se multiplient, pour le bien des utilisateurs. L'IA explicable (ou XAI) fait tomber le mythe de la boîte noire – et des imaginaires anthropomorphiques de « l'intelligence », des « hallucinations » – en présentant des informations techniques compréhensibles pour certaines catégories d'utilisateurs, qui leur permet de s'assurer qu'elle correspond bien à leurs besoins (Sajid et al., 2023). L'IA explicable fournit des explications précises permettant d'ouvrir la voie à des audits et des métriques d'évaluation. Ces informations aideraient certains membres du grand public (utilisateurs, chercheurs, représentants d'organes de gouvernance ou auditeurs) à garantir le respect des principes d'équité, d'égalité, de diversité, d'égalité et de confidentialité et à prendre des mesures pour corriger d'éventuels manquements, par exemple en introduisant des réclamations (judiciaires et non judiciaires).

L'EMI peut ajouter l'explicabilité à son répertoire de notions, car cela pourrait éclairer l'examen des enjeux de transparence et de responsabilité, qui sont souvent difficiles à appréhender pour les utilisateurs (Ananny et Crawford, 2018). L'explicabilité peut aussi être prise en compte dans la conception des programmes d'EMI, au même titre que les interactions avec des outils et des objets non humains (Long et Magerko, 2020). Les principes de la gouvernance de l'IA/IG devraient également être diffusés dans les différents secteurs industriels et sociaux sous la forme de directives et de normes. À titre d'exemple, c'est ce qu'a fait le Bureau du Commissaire à l'information britannique (Information Commissioner's Office, ICO) en publiant sa directive sur l'explication de décisions prises à l'aide de l'IA. Ce document aborde spécifiquement les questions relatives à l'utilisation des données, à la conception équitable, à la sécurité et à l'impact.

Toutefois, sans appropriation active par des citoyens, ces cadres n'auront que des effets limités et ne seront pas pérennes. Si les citoyens ne sont pas informés de leur existence et des avantages recherchés, ils ne peuvent pas dûment assurer la surveillance de l'IA en s'appuyant sur leurs propres tiers fiables. Les représentants des utilisateurs, comme ceux qui existent dans la communauté des acteurs de l'EMI, devraient contribuer aux mécanismes de régulation envisagés pour l'IG. Ils peuvent participer aux discussions menées aux niveaux local, national et mondial au sujet de la gouvernance de l'IA et devenir des atouts pour la mise en œuvre, le déploiement et la surveillance de la gouvernance de l'IA.

À l'heure actuelle, les organes gouvernementaux de régulation sont terriblement mal préparés à assumer leur rôle d'administrateur et de surveillant de l'IA afin d'établir une réglementation adéquate. En outre, il existe des risques d'ordre réglementaire, au premier rang desquels les risques d'inaction ou de mise en œuvre insuffisante. En ce qui concerne l'action publique, les décideurs devraient s'engager davantage à se former aux dernières évolutions de l'IA, afin d'être aptes à appliquer les garanties et les outils de surveillance appropriés. Cette prise de conscience a conduit 32 pays d'Afrique à demander l'assistance de l'UNESCO pour établir des normes politiques et renforcer les capacités. Cette démarche a donné lieu à la création d'une trousse à outils destinée aux décideurs (UNESCO, 2023e). Les fonctionnaires et les décideurs politiques peuvent également tirer parti des formations et des ressources en EMI (UNESCO, 2021c). L'UNESCO a également lancé le MOOC « Plongée dans l'éducation aux médias et à l'information pour les décideurs politiques » pour stimuler la formation à la matière, grandement nécessaire (UNESCO, 2023f).

Conclusion

Il est attendu des décideurs qu'ils exercent pleinement leur rôle d'administration de l'IA, que même ses créateurs appellent de leurs vœux. Ce rôle peut être favorisé en veillant à ces quatre points :

1. des garanties et directives pour une IA fiables doivent être créées et contrôlées,
2. les utilisateurs doivent être dotés des compétences en EMI qui ont été élargies pour englober l'IA,
3. la qualité de l'information est préservée, grâce à l'explicabilité et à des ensembles de données pertinents,
4. une partie des contrôles doit toujours être assurée par des humains, pour que les objectifs de l'IA soient conformes aux droits humains et que les systèmes d'IA traitent les enjeux urgents pour l'humanité, dont la crise environnementale et la fracture numérique.

Faire porter la responsabilité de la formation et de la sensibilisation à des individus en particulier ne suffira pas. Il faut parallèlement apporter des changements systémiques et structurels au marché de l'IA/AG. Les gouvernements devraient soutenir les plaidoyers des utilisateurs afin de promouvoir une concurrence loyale, d'éviter de potentiels troubles géopolitiques et de veiller à la bonne application des lois relatives à la confidentialité des données. Les gouvernements devraient aussi s'assurer que les développeurs d'AG, lorsqu'ils conçoivent leurs systèmes, considèrent comme prioritaires les questions de confidentialité, de responsabilité et de sécurité. Ils doivent s'efforcer de rétablir la confiance et la bienveillance parmi les utilisateurs en ce qui concerne les informations et les données, en coupant court à la manipulation et à la surveillance. Ils doivent également veiller au bien-être des publics et des communautés, toujours plus nombreux, qui utilisent des médias synthétiques.

Traiter l'IA sous l'angle de l'EMI peut ouvrir de nouvelles perspectives sur la gouvernance de l'IA et le rétablissement de la confiance dans les médias et l'information. Il est indispensable de renforcer le pouvoir d'action et la résilience des citoyens pour que les sociétés démocratiques soient en mesure de tirer parti des données, des algorithmes et de l'IA, tout en réduisant leurs répercussions négatives sur la qualité de l'information et la transparence. Relier ces formations à des activités de résolution de problèmes, au quotidien, peut conduire à l'adoption de pratiques plus efficaces et plus judicieuses qui englobent les concepts des algorithmes et de l'IA. L'EMI pour tous les citoyens peut contribuer à ce renforcement nécessaire des compétences et à l'établissement de coalitions, sans partir de zéro et sans renforcer les cloisonnements. L'EMI peut combler la fracture numérique en fournissant des solutions qui établissent des liens entre les secteurs STIM et les autres et en aidant les personnes à maîtriser les concepts fondamentaux nécessaires pour utiliser l'IA de manière efficace, sûre et éthique, quel que soit leur profil (technique ou non).

En fin de compte, dans les systèmes d'IA comme au-delà, l'objectif pour l'humanité est d'établir des sociétés du savoir viable et durable en s'appuyant sur l'intelligence collective – un environnement dans lequel l'éducation des citoyens aux médias et à l'information est un impératif.

10 recommandations clés pour l'EMI, la littératie de l'IA et la gouvernance de l'IA

1. Constituer un groupe de haut niveau qui établira les principes généraux d'un programme fondamental de littératie de l'IA, applicable du début du primaire à la fin du secondaire. Ce programme, inspiré de la perspective globale propre à l'EMI, qui est centrée sur l'humain, insistera sur l'égalité des genres et la lutte contre la discrimination raciale dans le contexte de la désinformation et des discours de haine.
2. Établir des priorités stratégiques pour l'EMI/IA dans les ministères et autorités chargés de l'éducation, de la culture, de l'égalité des genres et de la jeunesse, dans l'optique de gérer et d'évaluer les programmes de littératie de l'IA, en mettant en place les garde-fous appropriés pour les éducateurs et les apprenants.
3. Adopter et adapter des normes et directives sur l'IA explicables dans les politiques et programmes d'EMI ; les appliquer au secteur des technologies éducatives et à tous les autres acteurs proposant des solutions en faveur de la littératie de l'IA.
4. Mettre en place un portail mondial proposant des ressources de formation en EMI aux éducateurs (enseignants, bibliothécaires, animateurs pour la jeunesse, réseaux de femmes, etc.) pour favoriser le renforcement des capacités et l'expansion de l'EMI grâce à des outils en ligne faisant appel aux ressources éducatives libres sur l'IA.
5. Combler la fracture numérique dans le domaine de l'IA, et notamment les écarts entre les genres, en utilisant l'IA pour localiser davantage les programmes d'EMI afin de les proposer à des communautés variées (femmes, jeunes, etc.).
6. Encourager la diversité et l'inclusion en incluant des mesures de l'impact sur la diversité et l'inclusion dans les processus de conception des ressources en EMI et des programmes utilisés pour accroître la littératie de l'IA.
7. Créer un observatoire mondial de l'EMI pour financer la recherche et le développement de manière durable, afin de produire des données factuelles et de référence sur les apprentissages au sujet de l'IA et leurs effets sur le bien-être, l'éducation et la société.
8. Soutenir l'Organe mondial de contrôle des Nations Unies sur l'information et l'IA, avec l'ensemble des parties prenantes, dont les acteurs de l'EMI, afin de favoriser la surveillance du secteur et la production de rapports réguliers.
9. Diffuser une recommandation officielle de l'UNESCO - qui sera lancée à l'occasion de la Semaine mondiale de l'EMI sur l'autonomisation des utilisateurs à l'aide de l'EMI et de l'IA, pour stimuler un dialogue public mondial sur la confiance et la gouvernance d'une fiable et explicable.
10. Organiser un événement intersectoriel à l'UNESCO pour faire le point sur les questions relatives à la fiabilité des sources avec les principaux acteurs de différents secteurs, dont ceux de la société civile et des ONG (IFLA, AIERI, MILID, WAN-IFRA, etc.), pour garantir à l'avenir l'intégrité des sociétés du savoir.

Bibliographie

- Bryant, A. (2023). AI Chatbots: Threat or Opportunity? *Informatics*, 10, 49. <https://doi.org/10.3390/informatics10020049>
- Dogruel, L., Masur, P. and Joeckel, S. (2022). Development and Validation of an Algorithm Literacy Scale for Internet Users, *Communication Methods and Measures*, 16:2, 115-133, DOI: 10.1080/19312458.2021.1968361
- Frau-Meigs, D. (2012). Transliteracy as the new research horizon for media and information literacy. *Media Studies*, 3(6) <https://hrcak.srce.hr/ojs/index.php/medijske-studije/article/view/6064>
- Future of life (2023). Pause Giant AI Experiments: An Open Letter <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>
- Garon, J.M., (2023). An AI's Picture Paints a Thousand Lies: Designating Responsibility for Visual Libel, *Journal of Free Speech Law* 13:425 | 2023, <https://ssrn.com/abstract=4543822>
- Grossi, E. and Buscema, M. (2008). Introduction to artificial neural networks. *European journal of gastroenterology & hepatology*. 19. 1046-54. 10.1097/MEG.0b013e-3282f198a0.
- Hargittai, E., Gruber, J., Djukaric, T., Fuchs, J. and Brombach, L. (2020). Black box measures? How to study people's algorithm skills, *Information, Communication & Society*, 23:5, 764-775, DOI: 10.1080/1369118X.2020.1713846.
- Head, A.J., Fister, B. and MacMillan M. (2020.) Information literacy in the age of algorithms, https://projectinfilit.org/pubs/algorithm-study/pil_algorithm-study_2020-01-15.pdf
- Hendersen, P., Hashimoto, T. and Lemley, M. (2023). Where's the Liability in Harmful AI Speech? *Journal of Free Speech Law*, arXiv:2308.04635v2, doi.org/10.48550/arXiv.2308.04635
- Information Commissioner's Office (ICO) (2020). Guidance on explaining decisions made with AI, developed with the Alan Turing Institute, <https://ico.org.uk/media/for-organisations/documents/2620329/explaining-decisions-made-with-ai-summary.pdf>
- Jones, E. (2023). Explainer: What is a foundation model? Ada Lovelace Institute <https://www.adalovelaceinstitute.org/resource/foundation-models-explainer/>
- Lo, C K. (2023). What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature, *Educ. Sci.* 13, 410. <https://doi.org/10.3390/educsci13040410>
- Long, D. and Magerko, B. (2020). What is AI Literacy? Competencies and Design Considerations, CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems April 2020 Pages 1–16 <https://doi.org/10.1145/3313831.3376727>
- Möller, D. (2023). *Machine Learning and Deep Learning*, Springer DOI: 10.1007/978-3-031-26845-8_8.
- Ngo, R., Chan, L. and Mindermann, S. (2023). The alignment problem from a deep learning perspective, *ArXiv*, <https://arxiv.org/abs/2209.00626>
- Nguyen, D., and Beijon, B. (2023). The data subject and the myth of the 'black box' data communication and critical data literacy as a resistant practice to platform exploitation. *Information, Communication & Society*, DOI: 10.1080/1369118X.2023.2205504
- OECD Employment Outlook (2023). Artificial Intelligence and the Labour Market https://www.oecd-ilibrary.org/fr/employment/oecd-employment-outlook-2023_08785bba-en
- Open AI (2023). Governance of super intelligence, <https://openai.com/blog/governance-of-superintelligence>
- Römgens, I., Scoupe, R. and Beausaert, S. (2020). Unraveling the concept of employability, bringing together research on employability in higher education and the workplace, *Studies in Higher Education*, 45:12, 2588-2603, DOI: 10.1080/03075079.2019.1623770
- Sajid A., Tamer A., Shaker El-Sappagh, Khan M., Alonso-Moral, J-M., Confalonieri, R., Guidotti, R., Del Ser J., Díaz-Rodríguez, N., Herrera F. (2023). Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence, *Information Fusion*, Volume 99, 2023, <https://doi.org/10.1016/j.inffus.2023.101805>.
- Schlegel, D. and Uenal, Y. (2021). A Perceived Risk Perspective on Narrow Artificial Intelligence, *PACIS 2021 Proceedings*. 44. <https://aisel.aisnet.org/pacis2021/44>
- Shane, T., Saltz, E. and Leibowicz, C. (2021). There are lots of ways to label AI content. But what are the risks? <https://firstdraftnews.org/long-form-article/there-are-lots-of-ways-to-label-ai-content-but-what-are-the-risks/>
- Srivastava, B., Nikolich, A. and Koppel, T. (eds). (2023). AI and elections: An introduction to the special issue, *AI Magazine*, <https://doi.org/10.1002/aaai.12110>
- Stieglitz, S., Zeffass, A., Ziegele, D., Clausen, S., and Berger, K. (2022). *Communications Trend Radar 2022*. Language awareness, closed communication, gigification, synthetic media & cybersecurity (Communication Insights, Issue 14). Academic Society for Management & Communication. www.academic-society.net
- Trust, T., Whalen, J. and Mouza, C. (2023). Editorial: ChatGPT: Challenges, Opportunities, and Implications for Teacher Education. *Contemporary Issues in Technology and Teacher Education*, 23(1), 1-23. <https://www.learntechlib.org/primary/p/222408/>
- UNESCO (2023a). Handbook on Reporting about Artificial intelligence <https://unesdoc.unesco.org/ark:/48223/pf0000384551.locale=en>
- UNESCO (2023b). Generative AI and the future of education <https://unesdoc.unesco.org/ark:/48223/pf0000385877>
- UNESCO (2023c). Foundation Models such as ChatGPT through the prism of the UNESCO recommendation on Ethics of AI <https://unesdoc.unesco.org/ark:/48223/pf0000385629>
- UNESCO (2023d). Guidelines for Safeguarding Freedom of Expression and Access to Information in the Governance of Digital Platforms: A multistakeholder approach" https://www.unesco.org/sites/default/files/medias/fichiers/2023/04/draft2_guidelines_for_regulating_digital_platforms_en.pdf
- UNESCO (2023e). Building institutional capacity in public policy development in the field- A decision maker's toolkit of AI <https://en.unesco.org/artificial-intelligence/decision-makers-toolkit>
- UNESCO (2023f). "Deep Dive for Policymakers in Media and Information Literacy" MOOC <https://stageweb.unesco.org/en/articles/newly-launched-mooc-deep-dive-policymakers-media-and-information-literacy>
- UNESCO (2022b). The effects of AI on working lives of women <https://unesdoc.unesco.org/ark:/48223/pf0000380861>
- UNESCO (2022a). Recommendation on the Ethics of Artificial Intelligence <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
- UNESCO (2021a). Windhoek + 30 Declaration: information as a public good, World Press Freedom Day 2021. <https://unesdoc.unesco.org/ark:/48223/pf0000378158>
- UNESCO (2021b). Artificial Intelligence. Media and Information Literacy, Human Rights and Freedom of Expression https://iite.unesco.org/wp-content/uploads/2021/03/AI_MIL_HRs_FoE_2020.pdf
- UNESCO (2021c). Media and Information Literate Citizens: Think Critically, Click Wisely" <https://unesdoc.unesco.org/ark:/48223/pf0000377068>
- UNESCO (2020) Artificial Intelligence and Gender Equality: Key findings of UNESCO's global dialogue. <https://unesdoc.unesco.org/ark:/48223/pf0000374174/PD-F/374174eng.pdf.multi>
- UNESCO (2019). I'd blush if I Could: Closing Gender divide in digital skills through education, <https://en.unesco.org/id-blush-if-i-could>
- United Nations (2023) Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Irene Khan <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N23/233/65/PDF/N2323365.pdf?OpenElement>
- Van Drunen, M., Zarouali, B and Helberger, N. (2023). Recommenders you can rely on, https://www.jipitec.eu/issues/jipitec-13-3-2022/5562/van_drunen_13_3_2022.pdf
- Vincent, J. (2022). The scary truth about AI copyright <https://www.theverge.com/23444685/Generative-ai-copyright-infringement-legal-fair-use-training-data>
- Wach, K., Duong, C., Ejdy, J., Kazlauskaitė, R., Korzynski, P., Mazurek, G., Paliszewicz, J., and Ziemba, E. (2023). The dark side of Generative artificial intelligence: A critical analysis of controversies and risks of ChatGPT. *Entrepreneurial Business and Economics Review*, 11(2), 7-30. <https://doi.org/10.15678/EBER.2023.110201>
- Zou, A., Wang, Z., Zico Kolter, J., Fredrikson, M. (2023). Universal and Transferable Adversarial Attacks on Aligned Language Models, <https://llm-attacks.org/>



©Shutterstock

A PROPOS DE L'AUTEUR

Mme Divina Frau-Meigs est Professeur de sociologie des médias et des TIC à l'Université Sorbonne Nouvelle. Elle est réputée pour ses contributions à la recherche qui éclairent les politiques et les pratiques en matière d'éducation aux médias et à l'information. Reconnue par le « Global MIL Award » de l'UNESCO, l'expertise de recherche de Mme Frau-Meigs couvre la diversité culturelle, la gouvernance de l'Internet et l'étude des pratiques médiatiques des jeunes.

Publié en 2024 par l'Organisation des Nations Unies pour l'éducation, la science et la culture, 7, place de Fontenoy, 75352 Paris 07 SP, France et Bureau hors Siège de l'UNESCO/Institut de l'UNESCO.

© UNESCO 2024

Image d'illustration: ©Shutterstock

CI/FMD/MIL/2024/3



This publication is available in Open Access under the Attribution-ShareAlike 3.0 IGO (CC-BY-SA 3.0 IGO) license (<http://creativecommons.org/licenses/by-sa/3.0/igo/>). By using the content of this publication, the users accept to be bound by the terms of use of the UNESCO Open Access Repository (<https://www.unesco.org/en/open-access/cc-sa>).

The designations employed and the presentation of material throughout this publication do not imply the expression of any opinion whatsoever on the part of UNESCO concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

The ideas and opinions expressed in this publication are those of the authors; they are not necessarily those of UNESCO and do not commit the Organization.