

Intelligence artificielle

Opportunités et risques pour les enfants et les jeunes



Introduction

D'aucuns prétendent que l'intelligence artificielle (IA) existe depuis longtemps, même si elle n'est sur toutes les lèvres que depuis les dernières avancées technologiques. Avec ChatGPT, un générateur de textes ou chatbot (agent conversationnel) développé par OpenAI, le sujet de l'intelligence artificielle revient au centre des discussions sociétales. Ce qui relevait autrefois de la pure science-fiction est devenu réalité. La technologie de l'IA se développe à grande vitesse et influence notre vie de différentes manières. Certes, elle est un outil utile, mais elle recèle aussi un risque de dérives.

Cette fiche thématique a pour vocation d'aider à préparer les enfants et les jeunes à utiliser l'IA de manière plus responsable et positive.

En premier lieu, la notion d'intelligence artificielle sera expliquée et différentes formes en seront présentées. Ensuite, les risques seront mis en lumière et nous illustrerons comment favoriser une utilisation plus sûre de l'IA.

Table des matières

1. IA et algorithmes
2. Les différents types d'IA et leurs applications
3. Utilisation de l'IA par les enfants et les jeunes
4. Quels sont les risques à prendre en compte ?
5. Quelques conseils pour limiter les risques
6. Législation
7. Liens utiles

1. IA et algorithmes

Développement de l'IA

L'« intelligence artificielle » désigne des systèmes (informatiques) capables d'exécuter des tâches qui nécessitent normalement l'intelligence humaine. Parmi ces tâches figurent la compréhension du langage naturel, la reconnaissance de modèles, la prise de décisions et l'apprentissage à partir d'expériences, qui se basent tous sur des algorithmes.



Un algorithme

À l'instar d'une recette de cuisine, un algorithme se compose d'une suite précise et logique d'instructions. Son objectif est de résoudre un problème ou d'exécuter une tâche. Les programmes informatiques s'appuient sur des algorithmes. Alors que dans le cas des algorithmes classiques (déterministes), une entrée donnée produit toujours une même sortie, les algorithmes de l'IA sont plus flexibles : ils sont conçus pour que l'IA s'adapte et « apprenne ». Ces algorithmes sont donc capables de venir à bout de tâches plus complexes avec moins d'efforts.

L'histoire de l'intelligence artificielle a débuté dans les années 1950 avec le « test de Turing ». En 1957, le Perceptron a jeté les bases des réseaux de neurones complexes développés 30 ans plus tard, qui sont présentés en détail dans la section suivante. Le premier chatbot « Eliza » a été développé en 1966. Dès 1972, un système dit « expert » a été utilisé avec succès en médecine pour aider à diagnostiquer certaines maladies. En 1997, un ordinateur a battu un champion du monde d'échecs en titre.

En 2012, l'émergence du « *deep learning* » (apprentissage profond), que nous aborderons plus en détail plus tard, a considérablement accéléré le développement de l'IA. Ces progrès ont été possibles grâce aux ressources techniques toujours plus performantes. Depuis les années 2020, la présence de l'IA dans les médias s'est accentuée,

d'abord avec les systèmes d'IA générative comme GPT, ensuite avec les modèles de générateurs d'images comme DALL·E et, enfin, avec les chatbots comme ChatGPT. Avec les correcteurs automatiques sur les smartphones, les algorithmes de recommandation sur Netflix et les assistants vocaux comme Alexa, l'IA est entrée dans notre quotidien.

Principes de base de l'IA

En quoi consiste la technologie qui se cache derrière ?

- **Réseaux de neurones**

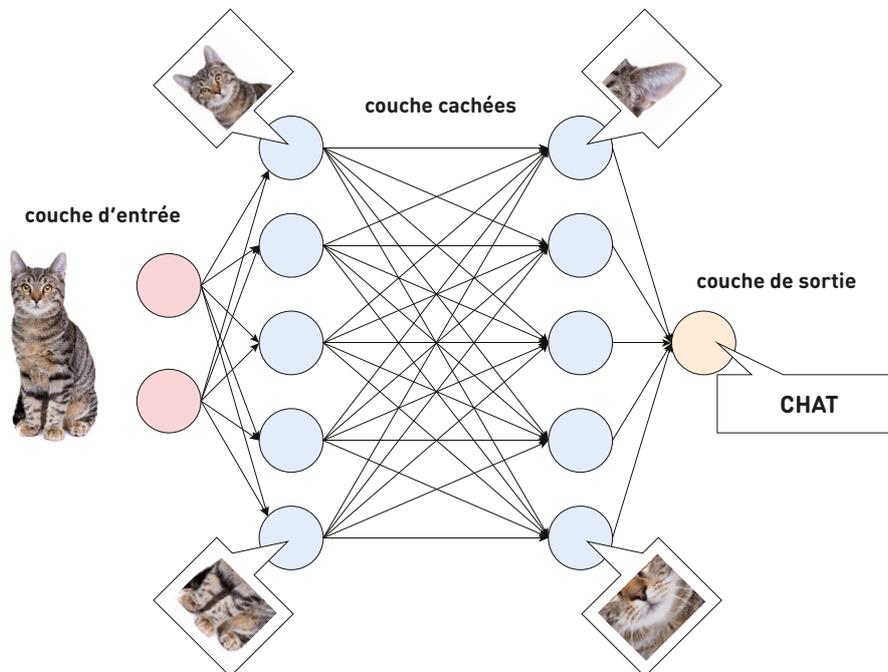
De nombreux systèmes d'IA actuels se basent sur des réseaux de neurones. Les réseaux neuronaux sont des algorithmes composés de nombreux neurones interconnectés et organisés en couches (appelés « layers » en anglais). Un neurone pourrait être décrit comme une « machine à prendre des décisions ». Il contient plusieurs entrées (« inputs »), chacune étant pondérée en fonction de son importance pour une tâche spécifique. En se basant sur la somme de ces données pondérées, une « fonction d'activation » décide quelle est la sortie. Selon le résultat, cette dernière est transférée à d'autres neurones dans la couche suivante. ChatGPT compte quelque 100 milliards de ces neurones organisés en 96 couches.

Quand un réseau neuronal doit reconnaître si un chien ou un chat est représenté sur une photo, il procède comme suit :

Donnée : une image est saisie dans la « couche d'entrée ».

Traitement : l'image traverse plusieurs « couches cachées » (« hidden layers ») de neurones. Chaque neurone examine des caractéristiques spécifiques (la fourrure, la forme des oreilles, etc.). La sortie résulte de la pondération, la somme et la fonction d'activation, et est relayée.

Sortie : la « couche de sortie » reçoit les résultats sur la base desquels le système décide s'il s'agit d'un chien ou d'un chat sur l'image.



• L'entraînement de l'IA – Apprentissage automatique

Comment l'IA apprend-elle à différencier un chien et un chat ou à nous parler ?

En entraînant les réseaux neuronaux ou les « modèles d'IA » avec les données correspondantes. C'est ce que l'on appelle le « *machine learning* » ou l'« apprentissage automatique » en français. Pour les derniers modèles d'IA, on parle également de « *deep learning* » (« apprentissage profond »), puisque ces réseaux neuronaux disposent d'une multitude de couches cachées. Dans ce contexte, le terme « apprentissage » désigne le fait que le réseau neuronal s'adapte aux données. Cette méthode est généralement plus rapide et moins coûteuse que le développement manuel d'un programme avec les mêmes fonctions.

Pour catégoriser les images de chiens et de chats, on a recours à une vaste collection d'images, c'est-à-dire à un ensemble de données (« *dataset* »). Ces images sont dotées d'une étiquette qui en identifie le contenu (c'est-à-dire « chien » ou « chat ») et sont présentées au modèle d'IA en plusieurs passages. À l'instar d'un humain, le modèle d'IA apprend à associer différentes caractéristiques à chaque animal. Une fois l'entraînement terminé avec succès, le réseau neuronal est capable de catégoriser des photos qu'il n'a jamais vues auparavant. Quand on détermine à l'avance des catégories lors de l'entraînement, on parle de « apprentissage supervisé ».

Pour les IA génératives qui produisent du texte (comme GPT), on utilise essentiellement l'« apprentissage non supervisé ». Pour ce faire, des ensembles de données contenant une grande quantité de textes issus de discussions en ligne, de commentaires sur Internet ou de textes littéraires sans catégorisation sont utilisés. Au cours des passages d'entraînement, le modèle d'IA apprend de manière autonome la façon dont la langue est construite, de la même manière qu'un élève apprend le vocabulaire et la grammaire à l'école.

L'« apprentissage par renforcement » (« *reinforcement learning* ») est une autre forme d'apprentissage automatique qui consiste en des récompenses positives ou négatives. Un modèle d'IA qui maîtrise déjà ses tâches, par exemple la génération de textes, produit des tâches qui seront évaluées par des humains ou des programmes. Ces évaluations permettent au modèle d'IA d'apprendre quelles sorties sont pertinentes ou non, et de les adapter en conséquence. ChatGPT a été entraîné de cette manière afin d'améliorer les réponses et d'exclure les sorties non souhaitées, dont les contenus illégaux.

Les modèles d'IA entièrement entraînés donnent souvent l'impression d'être dotés d'une conscience, mais toutes les sorties se basent sur un algorithme et apprennent des modèles !

2. Les différents types d'IA et leurs applications

Les intelligences artificielles ne sont pas toutes pareilles – différents modèles d'IA disposent de différentes capacités qui peuvent être appliquées dans divers domaines.

Traitement du langage naturel (Natural Language Processing)

Le Natural Language Processing (NLP) désigne le traitement et l'analyse du langage naturel (c'est-à-dire humain). Il englobe aussi bien le langage parlé que les textes écrits et se décline en différentes facettes :

- **Modèles génératifs de langage** : ils produisent du texte sur la base des entrées de l'utilisateur. Ils fonctionnent notamment sous la forme de chatbots (comme ChatGPT) et de modèles d'IA générative (tels que GPT-3). Les correcteurs automatiques sur les smartphones font également appel à cette technologie.
- **Reconnaissance et synthèse vocale** : les systèmes d'IA disposant de ces capacités transforment le langage oral en textes écrits et inversement. Les assistants vocaux comme Alexa d'Amazon ou Siri d'Apple peuvent exécuter d'autres tâches en s'appuyant sur cette fonction. Cette technologie peut également être utilisée pour imiter des voix. Au quotidien, elle s'avère utile pour transcrire ou dicter du texte.
- **Analyse de texte** : cette fonction sert à filtrer les informations pertinentes du. Par exemple, grâce à cette analyse, les chatbots et les assistants vocaux « comprennent » l'énoncé des utilisateurs, les courriers indésirables peuvent être détectés et les évaluations de clients peuvent être exploitées afin de proposer des produits adaptés.
- **Traduction automatique** : cette technologie permet aux modèles d'IA de traduire des textes dans d'autres langues.

Traitement des images

Le traitement des images par IA englobe à la fois la génération d'images et la reconnaissance d'objets qui y apparaissent. Les applications vidéo de nombreux smartphones modernes intègrent cette technologie pour optimiser automatiquement les photos, supprimer des objets ou reconnaître des visages.

- **La génération d'images** s'opère, par exemple, à l'aide de générateurs d'image par IA comme DALL·E, qui crée des images à partir d'entrées textuelles. Le transfert de style (« *style transfer* ») fait aussi partie de ces techniques. Il permet de transposer le style d'une image à une autre. Les filtres Instagram et Snapchat utilisent un procédé similaire.
- **La reconnaissance d'images et d'objets** sert à identifier et à classifier automatiquement les objets sur les images, notamment lors de la connexion à un appareil par reconnaissance faciale. Dans le domaine de la médecine, cette technologie est utilisée pour détecter des maladies.

Traitement de vidéos

Tout comme pour les images, des applications d'IA spécialisées peuvent produire des vidéos et y reconnaître automatiquement des objets. La reconnaissance s'étend toutefois aussi aux activités que l'on voit sur les vidéos.

- Au-delà de l'amélioration automatique des vidéos sur les smartphones et les caméras professionnelles, la génération de vidéos à l'aide de l'IA permet aussi de déployer la technologie du deepfake.



Deepfake

La technologie du deepfake permet de réaliser des vidéos contrefaites très réalistes en utilisant les visages de personnes réelles (« face swap ») alors que ces dernières mènent des activités ou prononcent des déclarations. Cette technique est utile dans la production de films, mais elle est également utilisée de façon malintentionnée pour nuire à autrui.

- **Analyse vidéo** : grâce à la reconnaissance automatique d'objets sur les vidéos, les caméras de sécurité sont en mesure d'identifier plus rapidement les intrus. Les plateformes comme Facebook ou YouTube utilisent cette technologie pour identifier automatiquement des personnes (« tagging ») ou pour examiner des contenus inadaptés. En outre, les systèmes d'IA peuvent reconnaître et catégoriser des activités humaines, une fonction qui est notamment intéressante pour les applications de fitness.
- **Réalité augmentée (« Augmented Reality » (AR)) et réalité virtuelle (« Virtual Reality » (VR))** : ces technologies misent sur l'IA dans la mesure où elle permet de superposer des vidéos réelles et des contenus virtuels.

Systèmes de recommandation

Les systèmes de recommandation, souvent appelés « **algorithmes de recommandation** », analysent le comportement de l'utilisateur pour proposer des contenus adaptés. Bien souvent, Netflix, Spotify et Amazon présentent aux utilisateurs des produits et contenus en rapport avec leurs recherches précédentes ou les contenus visualisés. Les publicités que les utilisateurs reçoivent sur les différents sites Internet reposent sur le même principe. Elles sont gérées par les systèmes de recommandation. Sur les réseaux sociaux, les utilisateurs voient essentiellement des contributions correspondant à leurs préférences. Ce système les répertorie en fonction des mentions « j'aime », des commentaires, des recherches et des vues de contenus. Ainsi, pour chaque utilisateur, une « bulle de filtres » (« *bubble* ») est créée. Dans le pire des cas, elle empêche l'utilisateur de se forger sa propre opinion, car elle force un point de vue unilatéral sur le sujet.

Parmi les autres domaines d'applications figurent :

- **la conduite autonome** : les données environnantes d'un véhicule sont évaluées automatiquement pour adapter la conduite.
- **la robotique** : certains robots industriels et appareils domestiques (p. ex. les robots aspirateurs) utilisent, entre autres, la reconnaissance d'objets afin d'exécuter leurs tâches de manière autonome ;
- **les pronostics** : que ce soit dans le secteur financier, dans la recherche ou dans le domaine de la météorologie, l'IA est capable de se prononcer en se basant sur des modèles et des comportements précis.

Une enquête mondiale menée par Statista qui porte sur le sujet « Quels outils d'IA générative utilisez-vous en plus de ChatGPT ? » révèle que parmi ces nombreuses possibilités d'usage, l'IA est principalement utilisée au quotidien dans le cadre des chatbots (15 % des personnes interrogées), des applications de jeux et de divertissement (11 %), des applications vidéo (6 %) et des applications de texte (5 %).

Selon l'enquête Statista « Savez-vous dans quelles applications l'IA est utilisée ? », nombre d'utilisateurs (âgés de 16 à 75 ans et vivant en Allemagne) ne sont pas conscients qu'ils ont recours à l'intelligence artificielle.

En effet, beaucoup de personnes interrogées ignorent que le filtre antispam de leur boîte mail fait appel à la technologie d'IA (50 % des participants) et que les moteurs de recherche et les boutiques en ligne personnalisent les résultats de recherche à l'aide de l'IA (39 et 38 % des personnes interrogées).

3. Utilisation de l'IA par les enfants et les jeunes

L'étude de la fondation Vodafone « Comment les élèves souhaitent-ils utiliser l'IA dans l'enseignement ? » menée auprès de quelque 1 600 jeunes âgés de 14 à 20 ans montre que, dans ce groupe d'âge, l'IA occupe déjà une place importante : 74 % des personnes interrogées l'utilisent de manière ciblée – dans le contexte scolaire essentiellement pour rechercher des informations (58 %), pour comprendre des concepts et des sujets ou pour obtenir des suggestions de solutions (32 à 50 %). Pour 46 % des participants à l'étude, ChatGPT se révèle être une aide dans l'apprentissage.

L'IA accompagne donc les enfants et les jeunes dans leur quotidien scolaire et peut leur apporter de multiples bénéfices.

Apprentissage

Dans le domaine de l'éducation, la capacité d'adaptation de l'IA, en particulier, offre de grands avantages : elle permet de générer des contenus et des stratégies d'apprentissage personnalisés qui tiennent compte des besoins des élèves. Pour les élèves à besoins spécifiques, elle offre également des avantages. En effet, les assistants vocaux et les autres outils d'IA peuvent les aider à surmonter leurs contraintes spécifiques d'apprentissage et favorisent l'inclusion. Le personnel enseignant, quant à lui, en tire des bénéfices, car l'IA suit plus facilement les progrès d'apprentissage des élèves et fournit un feedback.

Daniela Hau, spécialiste de l'IA dans le domaine de l'éducation, explique dans l'article « Impact de l'intelligence artificielle sur le monde de l'éducation » paru sur science. lu, à quoi l'utilisation de l'IA dans l'apprentissage peut ressembler à travers le cadre du projet Erasmus+ « AI4T – AI for and by teachers ». Dans ce contexte, la fonction « Speaker Coach » dans PowerPoint a notamment été utilisée. Elle donne aux élèves un retour objectif sur l'expression et la vitesse d'élocution dans leurs exposés. Les élèves peuvent ainsi améliorer leurs capacités de manière autonome et se servir de l'outil aussi souvent que nécessaire, une disponibilité que le corps enseignant ne peut assurer au quotidien.

ChatGPT est un outil utile pour les élèves et les enseignants : il permet aux enseignants d'élaborer rapidement des exemples de texte, ce qui libère du temps pour d'autres tâches. Les élèves peuvent obtenir rapidement des réponses que les parents et les enseignants ne peuvent peut-être pas fournir immédiatement.

Attention aux informations erronées !

Il est vrai que les informations produites par ChatGPT et d'autres IA semblent généralement crédibles, mais elles sont souvent incorrectes. La raison en est qu'elles sont générées à partir de probabilités statistiques et qu'elles ne sont pas vérifiées. Les réponses peuvent servir de point de départ et expliquer des concepts généraux, mais sans vérification humaine, elles ne peuvent pas être utilisées pour les devoirs ou être considérées comme des sources fiables.

Les « systèmes tutoriels intelligents » (STI) (« *intelligent tutoring systems* » (ITS)) adaptent le niveau de difficulté des tâches au niveau de performance de chaque élève pendant l'apprentissage. Certaines de ces applications, qui peuvent être utilisées en classe, peuvent par exemple aider les élèves à apprendre des langues ou faciliter l'initiation à la programmation et la pensée logique.

Des programmes de formation continue, comme le projet AI4T ou « Elements of AI » de l'université de Helsinki, forment les enseignants aux principes de l'intelligence artificielle et à son utilisation comme outil complémentaire en vue de l'intégrer dans les cours. Les élèves doivent apprendre à utiliser correctement les outils d'IA pour les manipuler en toute sécurité, avec discernement et de manière autonome. Près de 70 % des jeunes interrogés dans l'étude Vodafone considèrent qu'il est important de se familiariser avec l'IA pour pouvoir atteindre des objectifs professionnels. Ce point de vue correspond à l'évolution actuelle des offres d'emploi des entreprises en Allemagne, qui misent de plus en plus sur les emplois liés à l'IA.

Jouets

En faisant appel à l'intelligence artificielle, les jouets intelligents (« *smart toys* ») sont capables de réagir au comportement et à l'environnement des enfants. Les poupées, les peluches et les livres répondent aux questions posées par les enfants, les divertissent et encouragent l'apprentissage par le jeu (« gamification »). L'IA permet

aux jouets intelligents d'apprendre, ce qui devrait rendre l'interaction plus personnelle et plus naturelle.

Créativité

Les générateurs d'images par IA comme Midjourney, DALL·E ou AI Greenscreen de TikTok laissent libre cours à la créativité des utilisateurs, car ils produisent des images en fonction de consignes textuelles (« *prompts* »). Ils visualisent ainsi des idées, génèrent des images dans le style d'artistes connus et modifient de manière fantaisiste les photos que les utilisateurs leur soumettent. Les générateurs de portraits par IA vieillissent nos selfies ou les transforment en illustrations de bande dessinée. Les jeunes créateurs de musique peuvent produire leur propre son en renseignant un style ou en utilisant leurs propres paroles à l'aide d'applications d'IA dédiées.

Divertissement numérique et médias

Les « adversaires virtuels » existent depuis bien longtemps dans l'univers des jeux numériques. Grâce à l'intelligence artificielle, ces adversaires s'adaptent au comportement des joueurs, évitant ainsi l'ennui causé par des modèles récurrents.

Mais l'intelligence artificielle peut imiter un interlocuteur humain dans d'autres contextes que les jeux vidéo. Ainsi, le chatbot My AI de Snapchat figure comme un partenaire de discussion dans la liste d'amis. Les jeunes utilisateurs peuvent se servir de ce chatbot non seulement pour faire leurs devoirs et réaliser des traductions, mais aussi pour discuter, pour jouer aux devinettes ou pour créer de nouveaux « *Snaps* ».

Sur Instagram, on trouve également des personnages virtuels : des influenceuses IA sont à l'origine de canaux qui publient des images et des vidéos faisant appel à des personnes à l'apparence réelle qui sont en fait générées par des ordinateurs. C'est plus rentable pour les entreprises que de coopérer avec de vrais influenceurs pour le placement de leurs produits.

Comme l'illustrent ces différents exemples, l'intelligence artificielle occupe une place de plus en plus importante dans le quotidien des jeunes. Tandis qu'elle offre un immense soutien dans le domaine de l'apprentissage et au niveau scolaire, il est indispensable d'encourager les enfants, les parents et les enseignants à utiliser l'IA de manière réfléchie, et la compréhension des risques associés en fait partie.

4. Quels sont les risques à prendre en compte ?

Comme beaucoup d'applications numériques, les applications d'IA présentent aussi des risques spécifiques dont les enfants et les parents doivent avoir connaissance. Parmi ces risques figurent les suivants :

Contenus inadaptés

Quand ils utilisent des applications d'IA, les enfants peuvent être confrontés à des informations erronées ou non adaptées à leur âge. Même si des chatbots comme ChatGPT sont équipés de filtres pour bloquer les contenus à caractère discriminant, sexuel et injurieux, il existe de nombreuses façons de contourner ces protections, notamment par des prompts ciblés. En outre, certains outils d'IA génèrent des contenus inappropriés sans vérifier l'âge des utilisateurs. Certains générateurs d'images par IA sont capables de produire des images de nu en utilisant les visages provenant de photos fournies à cet effet.

Déformation de la réalité

Les modèles d'IA peuvent déformer la réalité. Les données avec lesquelles de nombreuses IA sont entraînées ne sont pas toujours filtrées en amont pour écarter les contenus discriminants ou inappropriés. C'est pourquoi l'IA peut émettre des avis douteux après l'entraînement. C'est ce que l'on appelle un « biais » (déformation). Dans bien des cas, les IA génératives d'images, elles non plus, ne reproduisent pas la réalité dans leurs créations.

Désinformation

L'exactitude des informations générées par les chatbots n'étant pas vérifiée, il est nécessaire d'examiner ces données de manière critique. Par exemple, les conseils en santé et nutrition fournis par les IA peuvent avoir des conséquences graves, puisqu'ils ne reposent pas sur des sources fiables ni sur l'expertise de spécialistes.

Mais l'IA est également utilisée délibérément à des fins de désinformation : des vidéos deepfake générées automatiquement, qui utilisent le visage de personnes spécifiques, sont de plus en plus répandues, notamment dans le contexte politique ou pour porter atteinte à des personnes. De plus, les applications d'IA facilitent la création de social bots qui propagent de manière ciblée de la désinformation sur les réseaux sociaux. Il faut donc faire preuve d'esprit critique, non seulement quand on utilise activement l'IA, mais surtout quand on s'informe sur Internet.

Enjeux pour la sécurité et la protection des données

Bien souvent, les enfants ignorent quelles informations personnelles ils peuvent partager en ligne, ce qui peut, entre autres, entraîner une violation de la protection des données et un accès non autorisé à ces données.

Les applications basées sur l'IA qui personnalisent les discussions en ligne et traitent les photos demandent souvent des informations ou des photos personnelles. Ces données sont enregistrées sur les serveurs des opérateurs. De plus, l'usage qui en est fait n'est pas toujours clair. Les jouets intelligents connectés à Internet, eux aussi, collectent des données privées telles que la localisation du jouet et de son utilisateur et des données sur l'entourage, et les transmettent à des tiers, généralement inconnus des utilisateurs.

Conséquences sur les relations sociales et harcèlement

Les tendances en ligne favorisent la moquerie des chatbots, un phénomène qui peut inciter les enfants à considérer le harcèlement comme normal en interagissant avec eux, puis à adopter ce comportement dans les interactions sociales réelles. En outre, l'IA permet une nouvelle forme de harcèlement : les deepfakes ainsi que les photos de nu générées par les « applications pour déshabiller » sont de plus en plus utilisées à des fins de chantage (« sextorsion ») et de (cyber)harcèlement, même chez les jeunes internautes. La production et le traitement de telles vidéos et photos sont répréhensibles, en particulier quand elles impliquent des enfants.

Imitation de la communication humaine

ChatGPT et les outils similaires imitent la communication humaine, ce qui peut induire en erreur les enfants. Ils ne savent pas toujours s'ils interagissent avec de vraies personnes ou avec une IA. Par exemple, My AI de Snapchat apparaît dans la liste d'amis au même titre que les autres contacts, ce qui alimente cette confusion.

Effets sur l'esprit critique et la prise d'initiatives dans l'apprentissage

Une autre crainte souvent formulée est que l'utilisation excessive de l'IA générative altère la capacité de faire preuve d'esprit critique et la prise d'initiatives chez les enfants. Le risque est particulièrement présent lorsqu'ils reprennent du contenu sans le remettre en question et qu'ils n'utilisent pas l'IA comme outil, mais comme générateur de réponses ou de dissertations toutes faites.

Atteinte aux droits d'auteur

Lorsque des images ou des textes sont générés par IA, les sources ne sont que très rarement citées de manière correcte. Au-delà des problèmes à l'école, cette pratique peut engendrer une violation des droits d'auteur, une pratique qui est passible de sanctions.

5. Quelques conseils pour limiter les risques

Malgré ces enjeux, il existe plusieurs façons de rendre l'utilisation de l'IA plus sûre pour les enfants et les jeunes (et bien entendu, pour les adultes aussi) :

- **Protection des données – Expliquez aux enfants le comportement à adopter avec les données personnelles :**

Les enfants (tout comme les adultes) ne doivent renseigner que les informations strictement nécessaires sur ChatGPT et les services similaires. Il ne faut en aucun cas transmettre des données confidentielles à des chatbots. Vérifiez les autorisations des applications que les enfants utilisent et consultez les déclarations de confidentialité pour savoir comment les données sont utilisées, y compris en ce qui concerne les jouets intelligents qui sont connectés à Internet ou à une application.

- **Attitude critique – Encouragez l'esprit critique lors de l'utilisation d'outils d'IA :**

Les chatbots peuvent fournir de fausses informations. Les enfants doivent donc toujours prendre ces informations avec beaucoup de recul et les comparer systématiquement avec d'autres sources. Expliquez aux enfants que les images produites par des outils tels que DALL-E sont générées artificiellement et ne représentent pas la réalité. En outre, expliquez-leur que les chatbots comme ChatGPT ou My AI sont des machines et pas de vraies personnes. Les enfants doivent toutefois toujours faire preuve de politesse dans leurs interactions avec ces outils et éviter toute forme d'injure.

- **Contact social et étiquette – Sensibilisez les enfants à l'utilisation respectueuse et responsable des contenus créés par IA :**

Assurez-vous que les enfants ne causent de tort à personne avec des images ou d'autres contenus générés par IA. Les contenus générés par IA peuvent être amusants, mais jamais au détriment des autres. Lors de la diffusion d'images générées par IA, les enfants doivent toujours mentionner qu'elles ont été créées artificiellement.

- **Utilisation générale de l'IA - Informez-vous sur l'intelligence artificielle et la protection des enfants :**

Restez au fait des nouvelles avancées en matière d'intelligence artificielle pour pouvoir mieux accompagner les enfants dans le monde des médias. BEE SECURE propose également un service de conseil en ligne.

6. Législation

La plupart des lois relatives à l'intelligence artificielle s'adressent aux développeurs et fournisseurs de modèles et de services d'IA.

Le règlement concernant l'IA (AI Act)

En mars 2024, le Parlement européen a adopté l'« AI Act », qui est entré en vigueur le 1er août 2024. Il identifie les risques associés aux différents systèmes d'IA et définit les moyens de les réguler. En résumé, les systèmes permettant la notation sociale et les IA à visée manipulatrice sont interdits. Ces systèmes cherchent à influencer l'opinion des utilisateurs, ce qui peut poser des problèmes dans les contextes politiques et sociaux.

Il est essentiel que l'utilisateur puisse identifier facilement qu'il interagit avec une IA, notamment dans le cas des chatbots ou des personnages deepfake pour que ces derniers ne soient pas pris pour des humains.

Les fournisseurs et développeurs doivent faire preuve d'une transparence accrue en donnant aux organismes de contrôle un accès aux données utilisées pour l'entraînement.

La Convention-cadre sur l'intelligence artificielle, les droits de l'homme, la démocratie et l'État de droit

En outre, le Conseil de l'Europe a voté en mai la Convention-cadre sur l'intelligence artificielle, les droits de l'homme, la démocratie et l'État de droit. Elle met l'accent sur la protection des droits de l'homme et des processus démocratiques. Les fournisseurs et développeurs d'applications d'IA doivent veiller à préserver la dignité humaine et l'égalité. Les applications d'IA ne doivent ni discriminer ni porter atteinte à la démocratie et la libre expression d'opinions. La protection des données personnelles, c'est-à-dire de la vie privée, doit être garantie.

Obligation de signaler les contenus deepfake

Hormis l'article 50, paragraphe 4 de la loi sur l'IA, il n'existe pas encore de dispositions légales spécifiques pour l'utilisateur final de l'intelligence artificielle. Cet article sur les « obligations de transparence pour les fournisseurs et les déployeurs de certains systèmes d'IA » régule l'obligation de signaler les deepfakes : « Les personnes qui déploient un système d'IA qui génère ou manipule un contenu image, audio ou vidéo constituant une contrefaçon profonde doivent indiquer que le contenu a été généré ou manipulé artificiellement. »

Droit d'auteur

Il existe toutefois des principes relevant du droit général qui peuvent s'appliquer à la création de contenu par IA, notamment le droit d'auteur. Ces droits de propriété intellectuelle doivent être observés lors de l'utilisation de services d'IA tels que DALL·E ou ChatGPT :

Si vous utilisez une **œuvre originale existante** (photo, image, texte, musique, réplique de sculpture, etc.) en tout ou en partie et que vous souhaitez la traiter avec l'IA, vous devez en obtenir l'autorisation de l'auteur ou de ses représentants, même quand il s'agit d'une utilisation privée. Si le traitement est autorisé, vous devez mentionner l'auteur comme source.

Une **œuvre entièrement créée par IA** n'est en général pas protégée par le droit d'auteur. Cependant, il faut citer la source, en l'occurrence l'IA utilisée, pour indiquer qu'il s'agit de contenu généré par IA. L'obligation de signaler de tels contenus est prévue au paragraphe 4 mentionné ci-dessus du règlement concernant l'IA.

Il convient aussi de vérifier les conditions d'utilisation spécifiques des services d'IA, car elles peuvent contenir des conditions supplémentaires. Ainsi, TikTok exige explicitement que les contenus générés par IA soient signalés. En général, le droit général prévaut, c'est-à-dire la loi modifiée du 18 avril 2001 sur les droits d'auteur, les droits voisins et les bases de données.

De plus, il est conseillé de vérifier les contenus générés par IA en effectuant une recherche sur Internet pour s'assurer qu'ils ne portent pas atteinte aux droits d'auteur en reproduisant le travail d'un artiste. Cette vérification peut être facilement réalisée moyennant une recherche inversée sur Google Images. En cas de doute, vous pouvez consulter des experts en droit de la propriété intellectuelle.

Ce manuel offre aux personnes intéressées un aperçu complet des textes formels, des lois, des projets de loi et des stratégies existants à ce jour à l'échelle européenne, y compris le règlement concernant l'IA. Toutes ces réglementations font partie de la stratégie européenne pour un Internet meilleur pour les enfants (BIK+ Strategy).

7. Liens utiles

Pour les parents

- BEE SECURE:
 - ChatGPT et autres outils d'IA – comment gérer les risques en tant que parents
www.bee-secure.lu/fr/news/chatgpt-et-autres-outils-dia-comment-gerer-les-risques-en-tant-que-parents/
 - Conférence thématique : désinformation
- Klicksafe.de
 - Was müssen Elteren zum neuen Chatbot My AI bei Snapchat wissen? www.klicksafe.de/news/was-muessen-eltern-zum-neuen-chatbot-my-ai-bei-snapchat-wissen
 - Was sollten Elteren über künstliche Intelligenz wissen? www.klicksafe.de/news/was-sollten-eltern-ueber-kuenstliche-intelligenz-wissen
- Srf.ch
 - Chancen und Risiken von AI: Künstliche Intelligenz einfach erklärt – www.srf.ch/sendungen/school/medien-und-informatik/chancen-und-risiken-von-ki-kuenstliche-intelligenz-einfach-erklart
- European strategy for a better Internet for kids (BIK+)
 - Compendium of legislation – <https://digital-strategy.ec.europa.eu/en/library/european-strategy-better-internet-kids-bik-compendium-legislation>

Pour le personnel enseignant et éducatif

- Script:
[Medienkompass : enseigner et apprendre les compétences médiatiques](#)
- Elements of AI: www.elementsofai.fr
- AI for Teachers: www.ai4t.eu/

Bibliographie

- Welche rechtlichen Rahmen gibt es bisher für die Nutzung von generativer KI?
www.jugendschutz.net/themen/geraete-und-technik/artikel/welchen-rechtlichen-rahmen-gibt-es-bisher-fuer-die-nutzung-von-generativer-ki-1
- Deep Fakes: gefälschte Videos, täuschend echt
www.bee-secure.lu/de/news/deep-fakes-gefaelschte-videos-taueschend-echt/
- Welche generativen KI-Tools nutzen Sie neben Chat GPT?
de.statista.com/statistik/daten/studie/1401946/umfrage/nutzung-generativer-ki-tools/
- Wissen Sie, in welchen Anwendungen KI zum Einsatz kommt?
de.statista.com/statistik/daten/studie/1309432/umfrage/umfrage-zu-ki-anwendungen-im-persoelichen-alltag/
- Intelligence artificielle
www.jeunesetmedias.ch/themes-alt/intelligence-artificielle
- KI für die Jugendarbeit
jugendleiter-blog.de/2023/03/31/kuenstliche-intelligenz-jugendarbeit/
- Impact de l'intelligence artificielle sur le monde de l'éducation
www.science.lu/fr/lia-dans-leducation/impact-lintelligence-artificielle-le-monde-leducation
- Pioniere des Wandels – Vodafone Stiftung
www.vodafone-stiftung.de/wp-content/uploads/2024/03/Pioniere-des-Wandels-wie-Schueler-innen-KI-im-Unterricht-nutzen-wollen-Jugendstudie-der-VS-2024.pdf
- The rising importance of Artificial Intelligence (AI) literacy
www.betterinternetforkids.eu/practice/articles/article?id=7214007
- Welche Risiken sind im Kinder- und Jugendmedienschutz zu erwarten?
www.jugendschutz.net/themen/geraete-und-technik/artikel/welche-risiken-sind-im-kinder-und-jugendmedienschutz-zu-erwarten-1
- My AI: Ein Chatbot erobert die sozialen Medien
www.srf.ch/sendungen/school/kuenstliche-intelligenz-my-ai-ein-chatbot-erobert-die-sozialen-medien
- Le Conseil de l'Europe adopte le premier traité international sur l'intelligence artificielle
www.coe.int/fr/web/portal/-/council-of-europe-adopts-first-international-treaty-on-artificial-intelligence-w
- L'impact de l'intelligence artificielle sur le développement des enfants
<https://elblog.pl/fr/2024/05/31/limpact-de-lintelligence-artificielle-sur-le-developpement-des-enfants-3/>
- Wat stécht hannert de gefälschten Deepfake-Videoe vu Politiker?
100komma7.lu/news/Wat-stecht-hanner-de-gefalschten-Deepfake-Videoe-vu-Politker-innen?pd=radio
- Jeder kann in eine Deepfake-Falle tappen
www.lessentiel.lu/de/story/betrug-in-luxemburg-jeder-kann-in-eine-deepfake-falle-tappen-103119959
- Was ist generative KI (Gen-KI) und wie kann sie sich auf das Wohlbefinden von Kindern auswirken?
www.internetmatters.org/de/hub/research/what-is-generative-ai-gen-ai-and-how-can-it-impact-childrens-wellbeing
- Stepstone Jobmarkt-Analyse
www.thestepstonegroup.com/deutsch/newsroom/pressemitteilungen/jobmarkt-analyse-unternehmen-schreiben-rund-50-prozent-mehr-ki-jobs-aus-als-vor-fuenf-jahren/
- Smart Toys: So nutzen Sie smarte Teddys, Autos & Co. sicher
www.bsi.bund.de/DE/Themen/Verbraucherinnen-und-Verbraucher/Informationen-und-Empfehlungen/Internet-der-Dinge-Smart-leben/Smart-Home/Smart-Toys/smart-toys_node.html
- Bots in sozialen Netzwerken
www.lmz-bw.de/medienbildung/themen-von-f-bis-z/soziale-netzwerke/bots-in-sozialen-netzwerken
- IBM – Was ist künstliche Intelligenz (KI)?
www.ibm.com/de-de/topics/artificial-intelligence
- TikTok – À propos du contenu généré par l'IA
support.tiktok.com/fr/using-tiktok/creating-videos/ai-generated-content#2
- Règlement sur l'IA
<https://artificialintelligenceact.eu/fr/>
- Loi du 18 avril 2001 sur les droits d'auteur, les droits voisins et les bases de données
<https://legilux.public.lu/eli/etat/leg/loi/2001/04/18/n2/jo>



Éditeur : Service national de la jeunesse (SNJ)

Service national de la jeunesse L-2926 Luxembourg

www.snj.lu

www.bee-secure.lu



Consulter la licence Creative Commons de cette publication :
www.creativecommons.org/licenses/by-nc-sa/4.0/deed.fr

Initié par :



LE GOUVERNEMENT
DU GRAND-DUCHÉ DE LUXEMBOURG

Opéré par :



Service national
de la jeunesse



Cofinancé par :



Cofinancé par
l'Union européenne

Fiche thématique
Intelligence artificielle - 10.2024
ISBN 978-2-919828-46-3
Ressource électronique