

Médiations & médiatisations

Revue internationale sur le numérique en éducation et communication

Les défis de l'IA dans l'éducation : de la protection des données aux biais algorithmiques

The Challenges of AI in Education: From Data Protection to Algorithmic Bias

Los retos de la IA en la educación: de la protección de datos a los sesgos algorítmicos

Os desafios da IA na educação: Da proteção de dados ao enviesamento algorítmico

Aïssa Messaoudi

Number 18, October 2024

Gestion, gouvernance et politiques d'intelligence artificielle en éducation et en enseignement supérieur
Management, Governance and Artificial Intelligence Policies in Education and Higher Education
Políticas de gestión, gobernanza e inteligencia artificial en la educación y la enseñanza superior
Gestão, governança e políticas para inteligência artificial na educação e no ensino superior

URI: <https://id.erudit.org/iderudit/1114746ar>

DOI: <https://doi.org/10.52358/mm.vi18.409>

[See table of contents](#)

Publisher(s)

Université TÉLUQ

ISSN

2562-0630 (digital)

[Explore this journal](#)

Cite this document

Messaoudi, A. (2024). Les défis de l'IA dans l'éducation : de la protection des données aux biais algorithmiques. *Médiations & médiatisations*, (18), 148–160.
<https://doi.org/10.52358/mm.vi18.409>

Article abstract

This article examines the impact of artificial intelligence (AI) on the field of education and explores its benefits and challenges. The use of AI in the education sector offers many advantages such as the automation of repetitive administrative tasks and the personalisation of learning paths. However, this raises ethical concerns about the protection of personal data and the risk of creating algorithmic biases. In addition, we address other challenges: those related to the opposition between automated and human assessment as well as the complex implications of facial recognition in an educational context. It is essential that a considered and ethical approach to the deployment of AI in education is thought through, emphasising the need for clear and transparent ethical principles and careful pedagogical reflection. We recommend the use of open-source AI tools to promote transparency and compliance with current regulations.

© Aïssa Messaoudi, 2024



This document is protected by copyright law. Use of the services of Érudit (including reproduction) is subject to its terms and conditions, which can be viewed online.

<https://apropos.erudit.org/en/users/policy-on-use/>



This article is disseminated and preserved by Érudit.

Érudit is a non-profit inter-university consortium of the Université de Montréal, Université Laval, and the Université du Québec à Montréal. Its mission is to promote and disseminate research.

<https://www.erudit.org/en/>

Les défis de l'IA dans l'éducation : de la protection des données aux biais algorithmiques

<https://doi.org/10.52358/mm.vi18.409>

Aïssa Messaoudi, professeur invité
Université des finances et de l'économie du Shandong, Chine
aissa@sdufe.edu.cn

RÉSUMÉ

Cet article examine l'impact de l'intelligence artificielle (IA) sur le domaine de l'éducation, et en explore les avantages et les défis. Le recours à l'IA dans le secteur éducatif offre de nombreux avantages tels que l'automatisation des tâches administratives répétitives et la personnalisation des parcours d'apprentissage. Cependant, cela soulève des préoccupations éthiques quant à la protection des données individuelles et au risque de biais algorithmiques. En outre, nous abordons d'autres défis : ceux liés à l'opposition entre l'évaluation automatisée et l'évaluation humaine ainsi que les implications complexes de la reconnaissance faciale dans un contexte éducatif. Il est essentiel qu'une approche réfléchie et éthique dans le déploiement de l'IA en éducation soit pensée en soulignant la nécessité de principes éthiques précis et transparents, et d'une réflexion pédagogique approfondie. Nous préconisons l'utilisation d'outils IA *open source* pour favoriser la transparence et la conformité aux réglementations en vigueur.

Mots-clés : intelligence artificielle, éducation, éthique, IA open source, reconnaissance faciale, données personnelles, automatisation, personnalisation pédagogique



Introduction

Parmi les disciplines les plus influencées par l'intégration de l'intelligence artificielle (désormais IA) comme l'informatique, l'ingénierie, les affaires, la linguistique, la santé, la sécurité, etc., l'éducation est celle qui occupe une position cruciale dans l'évolution du futur citoyen du monde ainsi que le confirme la définition proposée par l'UNESCO (2024) :

Education is the process of facilitating learning or the acquisition of knowledge, skills, values, beliefs and habits. Quality education specifically entails issues such as appropriate skills development, gender parity, provision of relevant school infrastructure, equipment, educational materials and resources, scholarships or teaching force.

L'éducation voit ses enjeux complètement bouleversés dans les différents établissements et en particulier dans l'enseignement supérieur. Ces dernières s'attèlent non sans mal à la redéfinir malgré un paradigme en mutation constante. Cette technologie de rupture (Zovko et Gudlin, 2019) est en train de « bouleverser » (Ancitil, 2023, p.67) les fondements de l'éducation, offrant des potentialités inédites pour la personnalisation de l'enseignement-apprentissage. Certains vont même jusqu'à prédire au secteur éducatif le même sort de *disruption* (Luckin, 2023; AIED, 2024¹). Cette association entre l'intelligence artificielle et l'éducation (désormais IAED) soulève de nombreuses questions et controverses, tant sur le plan pédagogique qu'éthique. Dans cet article, nous analysons les potentialités et les défis engendrés par l'IAED, en explorant ses différentes applications concrètes sur le terrain.

Cadre théorique

Alors que le terme « éducation » est relativement bien établi, celui de l'IA ne fait pas totalement consensus en raison de la nature évolutive de cette dernière (Sheikh *et al.*, 2023, p. 20). Pour autant, l'Organisation de coopération et de développement économiques (OECD, 2024), après un certain recul et aidée de ses trois experts du domaine, propose sa définition de l'IA :

An AI system is a machine-based system that for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.

Celle, plus récente, produite par la Commission européenne (2024a) dans sa législation *EU AI Act* et qui se rapproche sensiblement de celle de l'OCDE, indique un début de consensus :

An AI system is a machine-based systems designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it received, how to generate output such as content, predictions, recommendations, or decisions, that can influence physical or virtual environments.

¹ <https://aied2024.cesar.school/>



Ces deux définitions nous intéressent tout particulièrement, car en plus de jouir d'une certaine autorité, elles englobent les enjeux que nous présentons dans cet article. Toutefois, notre préférence va pour la dernière étant donné qu'elle met bien en avant cette capacité des IA à inférer la manière d'atteindre un but grâce à des approches basées sur l'apprentissage machine afin de produire des contenus. Et cela amène logiquement à évoquer les IA dites génératives (IAG). Le Conseil supérieur de l'éducation du Québec et la Commission de l'éthique en science et en technologie (2024, p. 8) ont opté pour la définition de Murugesan et Cherukuri (2023) : « Une IA générative est un type de système d'intelligence artificielle (SIA) capable de générer du contenu inédit très rapidement, en réponse à la demande d'une utilisatrice ou d'un utilisateur (ou à sa requête [prompt]) ». Pour autant, ne négligeons pas la proposition de définition faite par l'OCDE qui possède l'avantage d'avoir en son sein plus de pays membres, à savoir 38, qui par défaut s'aligneraient sur cette acception.

Problématique

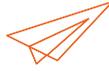
Ce préambule permet de voir toutes les possibilités offertes par l'IA. Elles sont encore à peine exploitées et semblent aussi larges que le permet l'imagination humaine. Or, un rapide survol laisse déjà entrevoir plusieurs questions que soulève l'IAED. Tout d'abord, doit-on personnaliser l'enseignement et l'apprentissage? Si oui, dans quelle mesure l'IAED peut-elle contribuer à la personnalisation de l'enseignement et de l'apprentissage, tout en répondant aux défis éthiques (problèmes liés aux données, inégalités d'accès, transparence, etc.) et aux questions de biais algorithmiques (sous et surreprésentation de groupes, données erronées, discrimination, etc.)?

Méthodologie

Pour tenter de répondre à cette problématique, nous procéderons à une revue narrative en nous appuyant sur une analyse documentaire de la littérature scientifique et des rapports d'experts récents sur l'IAED. Nous répertorions dans un premier temps les applications concrètes de l'IAED, celles qui sont déjà en cours et celles encore en balbutiement. Nous apporterons ensuite une attention particulière aux études évaluant l'impact de l'IA sur l'enseignement-apprentissage et les pratiques pédagogiques.

Résultats

Les possibilités de l'IAED sont majoritairement axées sur la personnalisation de l'apprentissage. Même si quelques chercheurs pointent certains de ses travers : « marketings tactics », « parents and teacher not included », « isn't being used as it could be », etc. (Bulger, 2016, p. 19-20), la majorité démontre qu'elle prend « en considération les caractéristiques de chaque personne en évitant de fonctionner sur le mode de la fragmentation des supports et des savoirs » (Connac, 2022, p. 18). Cette personnalisation peut se faire au niveau macro ou au niveau micro. On parlera alors respectivement de macro-adaptation ou de micro-adaptation. Corno (2008, p. 163) définit la première comme suit : « programs planned for groups of similar students based on formal assessments of qualities such as intellectual ability [...], or sociocultural background [...] »; et la seconde en ces termes : « continually assessing and learning as one teaches—thought and action intertwined ». La personnalisation s'est vue grandement accrue grâce aux technologies numériques et la venue de l'IA lui apporte une tout autre dimension. En effet, cette personnalisation peut s'appliquer, entre autres, sur le contenu, le rythme d'apprentissage, les méthodes d'enseignement et les activités. En nous concentrant sur la littérature scientifique et professionnelle récente (2023 et 2024), tant francophone qu'anglophone, nous avons identifié six moyens récurrents de personnalisation.



L'apprentissage adaptatif

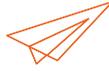
Une partie des spécialistes de l'éducation pense que chaque apprenant est différent sur le plan de l'apprentissage (Shah *et al.*, 2013, p. 212) et qu'il nécessite en conséquence une prise en charge personnalisée. Les experts de la pédagogie différenciée (ou différenciation pédagogique) ont tenté en vain de convaincre de son utilité, mais « les modalités de sa mise en œuvre, l'effectivité des dispositifs produits au fil du temps comme, au final, la cohérence d'ensemble des politiques conduites semblent insuffisantes » (Torres, 2016, p. 164). L'apprentissage adaptatif répond aux attentes de cette pédagogie. C'est un type d'apprentissage qui utilise l'IA pour adapter le contenu, le rythme d'apprentissage et les méthodes d'enseignement aux besoins individuels des apprenants. Cette adaptation en temps réel est possible « grâce à des algorithmes (machine) qui font des inférences à partir des actions de l'apprenant en session d'apprentissage » (Psyché et Ruer, 2019, p. 1). Parmi ses nombreux atouts, on peut citer l'optimisation du processus d'apprentissage grâce à un contenu et un rythme adaptés. Il favorise également l'engagement des apprenants avec des défis et de vraies opportunités de progression. Enfin, il réduit la frustration des élèves qui pouvaient se voir confronter à des exercices trop difficiles ou trop faciles comme dans l'apprentissage traditionnel. La difficulté est dosée en fonction des résultats de l'apprenant pour ne pas le frustrer à l'image des nouveaux jeux vidéo qui utilisent une difficulté dite dynamique (Vang, 2022, p. 6).

Les ressources pédagogiques adaptatives

De manière générale, les ressources pédagogiques sont « les instruments de présentation et de transmission du matériel éducatif prescrit. Cela inclut, entre autres : images, cartes, photographies, croquis, diagrammes, films, matériel écrit tel que des coupures de journaux ou des articles de la littérature scientifique et technique » (Bušljeta, 2013, p. 55). Elles se divisent en deux groupes : traditionnel (ex. : livre imprimé, photo imprimée, magazine, jeu de société, photo argentique, etc.) et numérique (ex. : livre électronique, GIF, site web, logiciel ludo-éducatif, application smartphone, etc.) comme il est précisé par Bourda (2004) qui reprend la définition du LOM (*Learning Object Metadata*) lors d'une communication. Les ressources pédagogiques dites adaptatives prennent logiquement leur place dans les ressources numériques, plus précisément dans celles dotées d'une IA. Par « adaptative », il faut comprendre une ressource qui s'adapte tant sur le fond que sur la forme. Un contenu sera plus ou moins détaillé selon les facilités de l'apprenant et si on le sait porté sur la bande dessinée pourrait se le voir présenter sous cette forme. Alazemi (2022, p. 58) liste les nombreux bienfaits des ressources pédagogiques. Elles rendent l'apprentissage plus engageant étant donné que les apprenants sont confrontés à des activités qui correspondent à leurs intérêts et à leurs besoins. Elles favorisent aussi l'autoapprentissage en les guidant et en les encourageant si nécessaire. Enfin, les apprenants progressent à leur rythme.

Les systèmes tutoriels intelligents

Comme nous l'avons vu plus haut, la pédagogie différenciée est difficilement applicable dans un contexte scolaire traditionnel et c'est bien souvent le tutorat à la maison ou en institut privé (dans des groupes d'apprenants plus réduits) qui correspond au mieux à cette démarche coûteuse. Désormais, l'IA peut y répondre en partie grâce à ce que l'on appelle les systèmes tutoriels intelligents (STI). Ce sont des « computer-based learning systems that use artificial intelligence (AI) to provide personalized and adaptive instruction for students », rappellent Lin, Huang et Lu (2023, p. 2). À l'instar des nombreux bénéfiques du tutorat traditionnel, les STI en possèdent également plusieurs, comme indiqué par Akyuz qui n'en



répertoire pas moins d'une dizaine (2020, p. 955-956). Sa recherche démontre également des progrès de 27 % chez les apprenants qui en font usage. Les apprenants tout comme les enseignants y sont favorables à 90 % (Akyuz, 2020, p. 975).

L'analyse des performances des apprenants

L'analyse des performances des étudiants est le processus de collecte, de mesure et d'interprétation des données sur les progrès des apprenants. Elle peut être utilisée pour suivre leurs progrès au fil du temps, identifier ceux à risque et prendre les décisions pédagogiques adéquates. Un logiciel d'évaluation basé sur l'IA offre la possibilité d'évaluer diverses compétences et performances, le plus souvent sous forme de tableau de bord. Elle permet aux enseignants de suivre les progrès des apprenants de manière plus précise au fil du temps tout en identifiant les points à renforcer pour éviter tout décrochage. Elle guide l'enseignant dans ses prises de décision pédagogiques, notamment sur la composition de différents groupes homogénéisés par niveau (Dhara *et al.*, 2022, p. 162).

Les évaluations automatisées

L'évaluation automatisée est un programme qui évalue les performances des élèves lors de devoirs, tests ou examens. Reiners, Dreher et Dreher (2011, p. 47) soulignent qu'elle est déjà utilisée depuis des décennies. Lorsqu'elle est couplée à une IA, elle offre des commentaires personnalisés et détaillés, adaptés à chaque élève, permettant ainsi une orientation pédagogique plus précise. Dans les devoirs quotidiens, l'IA est capable d'identifier facilement les erreurs courantes. Parmi les avantages de cette approche, on peut citer le gain de temps pour l'enseignant et la possibilité pour l'élève de recevoir un retour presque instantané sur ses performances. De plus, l'évaluation gagne en objectivité et en efficacité (Bablu, 2024) pourvu que l'approche préconisée n'induisse pas de biais (Wiley *et al.*, 2017, p. 784). Enfin, si nous devons la comparer à l'évaluation humaine, elle peut manquer de discernement pour des nuances subtiles dans le langage. Un apprenant qui rendrait une rédaction riche de quelques expressions idiomatiques, pour ne citer qu'un exemple, risquerait de mettre en déroute ces programmes (Tudor, 2022, p. 123).



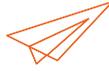
La prédiction de la réussite

Pouvoir prédire l'avenir est un désir que l'Homme tente d'assouvir depuis la nuit des temps : les Étrusques lisaient les entrailles d'animaux, les Babyloniens inventaient l'astrologie et les Grecs consultaient les oracles. Il est donc tout naturel de trouver de nombreux écrits sur la prédiction de la réussite à l'école bien avant l'avènement de l'IA, comme en témoigne l'article de Romainville (1997) ou celui de Corbière (1997), qui lui confèrent une place fondamentale dans l'éducation. Certaines Académies françaises font passer respectivement des tests de suite de logique et des tests d'efficience intellectuelle comme le B53 (créé en 1967) et le BV9 (créé en 1966) qui « constituent un maillon essentiel dans l'identification des élèves potentiellement fragilisés dans l'apprentissage » (Aubry et Bourdin, 2016). Aidé de l'IA, la prédiction de la réussite devient plus précise, particulièrement dans l'analytique de l'apprentissage grâce à la collecte et l'analyse des données des apprenants. Ces données reflètent directement l'expérience de l'apprenant, ses perceptions et ses comportements linguistiques pendant son apprentissage, comme le signalent Yan *et al.* (2024, p. 108). Dès lors, cela permet de prédire les résultats scolaires des apprenants, voire leur réussite ou non dans une filière donnée. Cela peut être utilisé pour identifier les élèves à risque de décrochage ou d'échec scolaire; elle peut soutenir la prise de décision pédagogique. Les enseignants peuvent alors prendre des mesures adaptées au niveau pédagogique et, ce faisant, réduire le taux d'échec scolaire.

Discussion

Sur le plan de l'enseignement-apprentissage, l'IAED se traduit par une ultrapersonnalisation des approches pédagogiques. Cela se matérialise par un apprentissage adaptatif, des ressources adaptatives, un système de tutoriel intelligent, des analyses des performances et des évaluations automatisées ainsi que la possibilité de prédire la réussite. Chacun des champs d'action de la personnalisation rencontre un taux d'efficacité discutable, mais ce n'est qu'une question d'années avant que tous ces outils n'atteignent des résultats plus que satisfaisants. L'autre plan sur lequel l'IAED joue indirectement un rôle important sans réellement constituer le cœur de l'éducation est le fait qu'elle permette une automatisation des tâches, comme la notation des apprenants, la formation de groupes homogènes, la gestion des inscriptions ou la remise de comptes-rendus réguliers aux parents. Faute de place dans cet article, nous ne pouvons malheureusement pas nous attarder sur ce versant administratif de l'IAED, mais il n'est pas à négliger, car il permet à l'enseignant de se concentrer sur l'essentiel, à savoir s'occuper de ses apprenants.

Au-delà des coûts très élevés que peut engendrer le développement de logiciels et de programmes basés sur l'IA (maintenance, mises à jour constantes, etc.), des problèmes techniques (Christodoulou et Angeli, 2022, p. 12) et d'une nécessaire formation de ces outils aux enseignants et aux apprenants, il existe plusieurs défis d'envergure qu'il convient de relever au plus vite. Il faut noter tout d'abord que les sociétés n'ont pas attendu une quelconque loi encadrant l'IA pour proposer leurs services. Et cela donne des situations discutables sur le plan éthique et confidentiel. Par exemple, le logiciel DALIA, un prédicteur de réussite en vigueur dans certains cégeps, ne fait pas l'unanimité à raison. La Fédération étudiante collégiale du Québec, dans sa note sur l'utilisation de l'IA sur les campus, pointe directement ce système en ajoutant qu'« une certaine vigilance quant à l'emploi des données personnelles est tout de même de mise » (2022, p. 6). Toutes ces entreprises mentionnent le point épineux des données personnelles et rassurent leurs futurs clients comme elles le peuvent, mais n'est-il pas plus sensé que les informations sensibles soient entre les mains d'un pouvoir public, qui en théorie n'entend pas réaliser de profits?



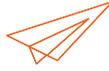
Par ailleurs, rien ne garantit l'exactitude de ces outils de prédictions qui sont dans les faits utilisés depuis plusieurs années. Sans aller jusqu'à faire le parallèle avec les logiciels prédictifs de crimes et leur taux de bonnes prédictions oscillant de 0,6 % à 0,1 % (Sankin et Mattu, 2023), le meilleur algorithme, le « regression-type random forest » est à peine au-dessus des 50 % de bonnes prédictions (Ahajjam, Haidar et Farhaoui, 2021, p. 840). Ces IA sont à notre image, pleines de stéréotypes : « AI perpetuates stereotypes » (Salazar *et al.*, 2024, p. 28). Les IA discriminent négativement aussi bien sur le plan de l'âge (*ibid.*, 2024, p. 14), des handicaps (*ibid.*, 2024, p. 15), des ethnies/races² (*ibid.*, 2024, p. 16), des orientations sexuelles et des identités de genre (*ibid.*, 2024, p. 17). Ces résultats ne doivent donc servir que de simples indicateurs. Il serait dommage qu'un apprenant se retrouve mal guidé dans sa scolarité et par la même occasion dans son avenir, sans compter un effet pygmalion à double tranchant, comme le fait remarquer Gras avec sa question rhétorique : « Est-ce éthique de prendre le risque de démotiver un apprenant? » (2019, p. 20).

Selon les types de données d'apprenants fournis à l'IA (origine sociale, origine ethnique, sexe, âge, dossier médical, etc.), celle-ci peut reproduire des biais, en l'occurrence algorithmiques. Ces dernières pourraient être corrigées en partie, mais, inopportunistement, la majorité des IA les plus en vogue ont un code source qui n'est pas en libre accès. Qui plus est, les IA utilisées sont loin d'être à l'abri des erreurs. La Commission nationale de l'informatique et des libertés (CNIL) met en garde les utilisateurs contre ces erreurs qui sont soit liées à la « conception du système » soit aux « conditions d'utilisation », (CNIL, 2022). Un système mal conçu manquera de représentativité, aura des hypothèses floues ou des critères inappropriés par exemple, sans oublier des données de mauvaise qualité ou des contraintes matérielles.

Bien que les outils d'analyse des performances des apprenants offrent de nombreux avantages, ils peuvent également créer une pression supplémentaire chez les élèves qui se sentent constamment évalués par ces systèmes IA. Il en va de même pour l'évaluation automatisée : elle reste moins précise que l'évaluation humaine en particulier pour les tâches complexes. Une IA n'aura pas l'indulgence d'un enseignant qui distingue une erreur de frappe au lieu d'une erreur de compréhension.

Enfin, dans le cas d'un examen où la reconnaissance faciale serait obligatoire pour éviter la fraude à l'identité, certaines minorités de genre et d'ethnie pourraient rencontrer des difficultés à franchir cette étape (Hassanin, 2023). Pourtant, en analysant les expressions non verbales des apprenants, la reconnaissance faciale pourrait s'avérer précieuse pour l'enseignant en lui indiquant les moments les plus intéressants ou ennuyeux de ses cours (Wang *et al.*, 2020, p. 6), mais les bénéfices semblent être trop peu nombreux comparés aux risques et c'est ce qui a conduit l'État de New York à la bannir de ses écoles en septembre 2023 : « The order prohibits schools in New York State from purchasing or utilizing facial recognition technology » (New York State Education Department, 2023).

² Rappelons que le mot *race* proposée par les auteures américaines est une notion discutable. « La race est une construction sociale. Elle n'a aucune base génétique », comme le rappelle le généticien Drouin (2021).



Les STI ne dérogent pas non plus à la règle. Ils posent de sérieux défis (biais algorithmiques, éthiques, qualité des données, etc.) pouvant « compromettre les objectifs de l'éducation » ainsi que le soulignent Lin, Huang et Lu (2023, p. 18). Il faudrait que les interactions soient majoritairement contrebalancées avec celles de l'enseignant afin de ne pas nuire à la qualité de l'apprentissage. Chan et Tsi (2023, p. 15) disent ceci à ce sujet : « The emotional and interpersonal skills of human teachers are essential for personal growth, resilience, and critical thinking ». De fait, la qualité des dialogues et des retours est loin d'atteindre celle d'un enseignant et d'un élève. Sans l'enseignant, l'apprenant se retrouverait dans une bulle.

Conclusion

L'IA a le potentiel de transformer l'éducation en profondeur, mais son utilisation doit être guidée par des principes éthiques clairs et une réflexion pédagogique approfondie. Des recherches et des expérimentations sont nécessaires pour évaluer l'impact de l'IA sur les apprentissages et les pratiques pédagogiques, et pour développer des solutions inclusives et responsables.

À notre avis, se tourner vers les IA *open source* est l'action la plus raisonnable pour le monde de l'éducation et ses établissements publics. Ces IA, dont le code source est accessible au grand public et qui offrent tous les avantages de la transparence telles des révisions constantes par la communauté, ne manquent pas. Elles prennent d'ailleurs bien souvent la forme d'IA génératives mentionnées plus haut qui peuvent aisément être reliées aux types d'usages énumérés dans cet article pour peu que l'utilisateur sache leur parler avec les *prompts*. On peut citer : Flan-T5-xxl, Mistral AI (en France), Aleph Alpha (Allemagne), Falcon 40B (Abou Dabi), etc. Même la multinationale Meta (anciennement Facebook), connue pour prospérer sur les données de ses utilisateurs (Hutchinson, 2024), a décidé dernièrement en août 2023 de faire basculer son modèle de langage LLaMA 2 vers l'*open source*³. Il ne faut pas y voir là un geste désintéressé de la firme, mais plutôt une stratégie pour prendre de l'avance sur ses concurrents afin de correspondre au cadre législatif en préparation par la Commission européenne dans son *AI Act*. En effet, l'encadrement des IA *open source* sera bien plus souple, comme on peut le comprendre dans l'article 53 :

« The obligations [...] shall not apply to providers of AI models that are released under a free and open-source licence that allows for the access, usage, modification, and distribution of the model, and whose parameters, including the weights, the information on the model architecture, and the information on model usage, are made publicly available. This exception shall not apply to general-purpose AI models with systemic risks » (Commission européenne, 2024b).

Le seul point qui pourrait rebuter ses utilisateurs est une performance moindre lors des requêtes (Wu *et al.*, 2024, p. 4), mais la tendance indique que l'écart se réduit entre les deux IA.

N'oublions pas non plus qu'il est très aisé de les faire fonctionner localement dans des serveurs gérés en interne par le ministère de l'Éducation afin de créer une harmonisation des algorithmes entre tous les établissements. Et ces derniers pourraient héberger eux-mêmes leurs propres serveurs et protéger au mieux les données de leurs apprenants. Enfin, dans l'absolu, un étudiant peut même élaborer une IA sur son propre ordinateur et être moins exposé aux cybercriminels qui préfèrent s'attaquer aux grosses entités.

³ L'inverse s'applique également. Le mot « Open » dans *OpenAi* était une allusion à son code *open source*. L'entreprise américaine était jusqu'en 2019 une organisation à but non lucratif.



Il devient dès lors difficile de faire plus confidentiel.

D'autres pistes gagneraient à être étudiées, comme celle des agents d'IA, « ces systèmes/programmes capables d'effectuer de manière autonome des tâches au nom d'un utilisateur ou d'un autre système en concevant son flux de travail et en utilisant les outils disponibles » (Gutwoska, 2024). Les experts ne tarissent pas d'éloges à son sujet, notamment pour leurs capacités de réactions et de proactions dans leur environnement (Chudleigh, 2024).

Liste de références

- Ahajjam, T., Haidar, A., et Farhaoui, Y. (2021). Artificial Intelligence and Machine Learning to Predict Student Performance during the COVID-19. *Procedia Comput Sci.* (2021)184, 835-840. <https://doi.org/10.1016/j.procs.2021.03.104>
- Akyuz, Y. (2020). Effects of Intelligent Tutoring Systems (ITS) on Personalized Learning (PL). *Creative Education*, 11(6), 953-978. <https://doi.org/10.4236/ce.2020.116069>
- Alazemi, N. (2022). The impact of digital learning resources on developing the educational process for faculty members at the PAAET. *Amazonia Investiga*, 11(59), 54-63. <https://doi.org/10.34069/AI/2022.59.11.5>
- Anctil, D. (2023). L'éducation supérieure à l'ère de l'IA générative. Réflexion pédagogique. *Pédagogie collégiale*, 36(3). <https://educu.info/xmlui/bitstream/handle/11515/38833/Anctil-36-3-23.pdf>
- Aubry, A., et Bourdin, B. (2016). Les tests BV9 et B53 peuvent-ils prédire la réussite scolaire? *L'Orientation scolaire et professionnelle*, 45(3). <https://doi.org/10.4000/osp.5070>
- Bablu, T. A. (2024). Machine Learning in Automated Assessment: Enhancing Objectivity and Efficiency in Educational Evaluations. *Journal of Advanced Computing Systems*, 4(7). <https://scipublication.com/index.php/JACS/article/view/20/9>
- Bourda Y. (2004). Pourquoi indexer les ressources pédagogiques numériques? [compte-rendu 16 novembre 2004 d'une communication orale]. <https://tinyurl.com/4y4ee94t>
- Bulger, M. (2016). *Personalized Learning: The Conversations We're Not Having*. Data & Society. https://datasociety.net/pubs/ecl/PersonalizedLearning_primer_2016.pdf
- Bušljeta, R. (2013). Effective Use of Teaching and Learning Resources. *Czech-Polish Historical and Pedagogical Journal*, 5, 55-69. <https://doi.org/10.2478/cphpj-2013-0014>
- Chan, C.K.Y., et Tsi, L.H.Y. (2023). *The AI Revolution in Education: Will AI Replace or Assist Teachers in Higher Education?* <https://tinyurl.com/wk8mt69m>
- Chudleigh, S. (2024, 31 mai). *Qu'est-ce qu'un agent d'intelligence artificielle?* Botpress. <https://botpress.com/fr/blog/what-is-an-ai-agent>
- Christodoulou, A., et Angeli, C. (2022). Adaptive Learning Techniques for a Personalized Educational Software in Developing Teachers' Technological Pedagogical Content Knowledge. *Frontiers in Education*, 7. <https://doi.org/10.3389/educ.2022.789397>
- Commission européenne (2024a). EU AI Act. Chapter I: General Provisions. Chapter I: General Provisions. Article 3, Definitions (1). <https://www.euaiact.com/article/3>
- Commission européenne. (2024b). EU AI Act. Chapter V: General-Purpose AI Models. Section 2: Obligations for Providers of General Purpose AI Models. Article 53. <https://www.euaiact.com/article/53>
- Commission nationale de l'informatique et des libertés (2022, 25 mars). *Intelligence artificielle, de quoi parle-t-on?* <https://www.cnil.fr/fr/intelligence-artificielle/intelligence-artificielle-de-quoi-parle-t-on>



- Connac, S. (2022). La personnalisation des apprentissages : agir face à l'hétérogénéité à l'école et au collège, *ESF Sciences humaines*. http://media.sodis.fr/cat/9782710145455/15/9782710145455-15-220816_151933.pdf
- Conseil supérieur de l'éducation et Commission de l'éthique en science et en technologie (2024). *Intelligence artificielle générative en enseignement supérieur : enjeux pédagogiques et éthiques*. <https://tinyurl.com/2ksfct4p>
- Corbière, M. (1997). Une approche multidimensionnelle de la prédiction de la réussite scolaire. *L'Orientation scolaire et professionnelle*, 26(1), 109-135. https://www.persee.fr/doc/binop_0249-6739_1997_num_26_1_1178
- Corno, L. (2008). On Teaching Adaptively. *Educational Psychologist*, 43:3, 161-173, DOI: 10.1080/00461520802178466. <https://doi.org/10.1080/00461520802178466>.
- Drouin, G. (2021). L'évolution et le concept de race avec Guy Drouin [balado]. Radio-Canada. <https://tinyurl.com/259x7sxm>
- Dhara, S., Chatterjee, S., Chaudhuri, R., Goswami, P. A. et Ghosh, S. (2022). Chapter 8: Artificial Intelligence in Assessment of Students' Performance. Dans *Artificial Intelligence in Higher Education: A Practical Approach* (1^{re} éd.) (p.153-167). Taylor and Francis. <https://doi.org/10.1201/9781003184157-8>
- Fédération étudiante collégiale du Québec (2022). *Note sur l'utilisation de l'intelligence artificielle sur les campus*. 118^e Congrès ordinaire, Commission des affaires collégiales. https://docs.fecq.org/FECQ/M%C3%A9moires%20et%20avis/2022-2023/Note-IA_118eCo_LaMalbaie.pdf
- Gras, B. (2019). Éthique des Learning Analytics. *Distances et médiations des savoirs*, 26. <https://doi.org/10.4000/dms.3768>
- Gutwoska, A. (2024, 3 juillet). *What are AI Agents?* IBM. 3 juillet 2024. <https://www.ibm.com/think/topics/ai-agents>
- Hassanin, N. (2023, 23 août). *Law professor explores racial bias implications in facial recognition technology*. UCalgary News. University of Calgary. <https://ucalgary.ca/news/law-professor-explores-racial-bias-implications-facial-recognition-technology>
- Hutchinson, A. (2024, 27 janvier). *Meta Loses Data Scraping Case, Highlighting the Need For Clarified Regulation in a Social Media Context*. Social Media Today. <https://www.socialmediatoday.com/news/meta-loses-data-scraping-highlighting-need-clarified-regulation/705814/>
- Lin, C.-C., Huang, A.Y.Q., et Lu, O.H.T. (2023). Artificial intelligence in intelligent tutoring systems toward sustainable education: a systematic review. *Smart Learn. Environ.* 10, 41. <https://doi.org/10.1186/s40561-023-00260-y>
- Luckin, R. (2023, 14 juillet). Yes, AI could profoundly disrupt education. But maybe that's not a bad thing. *The Guardian*. <https://tinyurl.com/wh3rv6ks>
- OECD (2024, 6 mars). *What is AI? Can you make a clear distinction between AI and non-AI systems?* <https://oecd.ai/en/wonk/definition>
- New York State Education Department (2023). *State Education Department Issues Determination on Biometric Identifying Technology in Schools*. <https://tinyurl.com/mr38unr5>
- Parlement européen (2024). *Proposal for a Regulation of the European parliament and of the Council laying down harmonised rules on the artificial intelligence (Artificial Intelligence Act) and amending certain union legislative*. EUR-Lex. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>
- Psyché, V. et Ruer, P. (2019). L'apprentissage adaptatif intelligent. I *Le Tableau*, 8(4). <https://pedagogie.quebec.ca/le-tableau/lapprentissage-adaptatif-intelligent>
- Reiners, T., Dreher, C. et Dreher, H. (2011). Six Key Topics for Automated Assessment Utilisation and Acceptance. *Informatics in Education*, 10(1), 47-64, Vilnius University. <https://files.eric.ed.gov/fulltext/EJ1064287.pdf>
- Romainville, M. (1997). Peut-on prédire la réussite d'une première année universitaire? *Revue française de pédagogie. L'éducation préscolaire*, 119, pp. 81-90. https://www.persee.fr/doc/rfp_0556-7807_1997_num_119_1_1169
- Salazar L. R., Peoples, S. F., Brooks, M. E. (2024). Generative AI Ethical Considerations and Discriminatory Biases on Diverse Students Within the Classroom. Dans S. Elmoudden et J. Wrench (dir.), *The Role of Generative AI in the Communication Classroom* (p. 191-213). IGI Global. <https://doi.org/10.4018/979-8-3693-0831-8.ch010>
- Sankin, A., et Mattu, S. (2023, 2 octobre), Predictive Policing Software Terrible at Predicting Crimes. *Wired*. <https://www.wired.com/story/plainfield-geolitica-crime-predictions/>

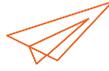


- Shah, K, Ahmed, J., Shenoy, N. et Srikant, N. (2013). How different are students and their learning styles? *International Journal of Research in Medical Sciences*, 1(3), 212–215. <https://doi.org/10.5455/2320-6012.ijrms20130808>
- Sheikh, H., Prins, C., et Schrijvers, E. (2023). Artificial Intelligence: Definition and Background. Dans Sheikh, H., Prins, C., et Schrijvers, E. (dir.), *Mission AI. Research for Policy*. Springer. https://doi.org/10.1007/978-3-031-21448-6_2
- Torres, J.-C. (2016). Les enjeux de la différenciation pédagogique : entre résolutions formelles et indécisions pratiques. *Administration & Éducation*, 2(150), 159-164. <https://doi.org/10.3917/admed.150.0159>
- Tudor, M.-D. (2022). La traduction des expressions idiomatiques à l'aide de moteurs de traduction automatique. *Analele Universității București. Limbi și Literaturi Străine*, 68(2), 109-128. <https://tinyurl.com/22zy55ku>
- UNESCO (2024). *Education for sustainable development. Sustainable development*. <https://en.unesco.org/themes/education/sdgs/material/04>
- Vang, C. (2022). The Impact of Dynamic Difficulty Adjustment on Player Experience in Video Games. *Scholarly Horizons: University of Minnesota, Morris Undergraduate Journal*: Vol. 9: Iss. 1, Article 7. <https://doi.org/10.61366/25762176.1105>
- Yan, L., Martinez-Maldonado, R., et Gasevic, D. (2024). Generative Artificial Intelligence in Learning Analytics: Contextualising Opportunities and Challenges through the Learning Analytics Cycle. Dans *LAK '24: Proceedings of the 14th Learning Analytics and Knowledge Conference* (p. 101-111). ACM. <https://doi.org/10.1145/3636555.3636856>
- Wang, W., Xu, K., Niu, H. et Miao, X. (2020). Emotion Recognition of Students Based on Facial Expressions in Online Education Based on the Perspective of Computer Simulation. *Complexity*. 2020. 1-9. 10. <https://doi.org/10.1155/2020/4065207>
- Wiley, J., Hastings, P., Blaum, D., Jaeger, A. J., Hugues, S., Wallace, P., Griffin, T. D. et Britt, M. A. (2017). Different Approaches to Assessing the Quality of Explanations Following a Multiple-Document Inquiry Activity in *Science*. *International Journal of Artificial Intelligence in Education*, 27, 758-790. <https://doi.org/10.1007/s40593-017-0138-z>
- Wu, S., Koo, M., Blum, L., Black, A., Kao, L. Fei, Z., Scalzo, F. et Kurtz, I. (2024, 17 janvier) Benchmarking Open-Source Large Language Models, GPT-4 and Claude 2 on Multiple-Choice Questions in Nephrology. *NEJM AI*, 1(2) Datasets, Benchmarks, and Protocols. <https://doi.org/10.1056/Aldbp2300092>
- Zovko, V. et Gudlin, M. (2019). Artificial Intelligence as a Disruptive Technology in Education. *The Future of Education*. 9th ed. *International Conference*. <https://tinyurl.com/326uncnr>
-

The Challenges of AI in Education: From Data Protection to Algorithmic Bias

ABSTRACT

This article examines the impact of artificial intelligence (AI) on the field of education and explores its benefits and challenges. The use of AI in the education sector offers many advantages such as the automation of repetitive administrative tasks and the personalisation of learning paths. However, this raises ethical concerns about the protection of personal data and the risk of creating algorithmic biases. In addition, we address other challenges: those related to the opposition between automated and human assessment as well as the complex implications of facial recognition in an educational context. It is essential that a considered and ethical approach to the deployment of AI in education is thought through, emphasising the



need for clear and transparent ethical principles and careful pedagogical reflection. We recommend the use of open-source AI tools to promote transparency and compliance with current regulations.

Keywords: artificial intelligence, education, ethics, open source AI, facial recognition, personal data, automation, personalised teaching

Los retos de la IA en la educación: de la protección de datos a los sesgos algorítmicos

RESUMEN

Este artículo examina el impacto de la inteligencia artificial (IA) en el ámbito de la educación y explora sus ventajas y retos. El uso de la IA en el sector educativo ofrece muchas ventajas, como la automatización de tareas administrativas repetitivas y la personalización de las vías de aprendizaje. Sin embargo, esto plantea problemas éticos sobre la protección de datos personales y el riesgo de sesgo algorítmico. Además, abordamos otros retos: los relacionados con la oposición entre la evaluación automatizada y la humana, así como las complejas implicaciones del reconocimiento facial en un contexto educativo. Es esencial que se diseñe un enfoque reflexivo y ético para el despliegue de la IA en la educación, haciendo hincapié en la necesidad de unos principios éticos claros y transparentes y de una cuidadosa reflexión pedagógica. Recomendamos el uso de herramientas de IA de código abierto para fomentar la transparencia y el cumplimiento de la normativa vigente.

Palabras clave: inteligencia artificial, educación, ética, IA de código abierto, reconocimiento facial, datos personales, automatización, enseñanza personalizada

Os desafios da IA na educação: Da proteção de dados ao enviesamento algorítmico

RESUMO

Este artigo analisa o impacto da inteligência artificial (IA) no domínio da educação e explora os seus benefícios e desafios. A utilização da IA no setor da educação oferece muitas vantagens, como a automatização de tarefas administrativas repetitivas e a personalização dos percursos de aprendizagem. No entanto, suscita preocupações éticas sobre a proteção dos dados pessoais e o risco de enviesamento algorítmico. Além disso, abordamos outros desafios: os relacionados com a oposição entre avaliação automatizada e humana, bem como as implicações complexas do reconhecimento facial num contexto educativo. É essencial



refletir sobre uma abordagem ponderada e ética da utilização da IA na educação, sublinhando a necessidade de princípios éticos claros e transparentes e de uma reflexão pedagógica cuidadosa. Recomendamos a utilização de ferramentas de IA de fonte aberta para promover a transparência e o cumprimento da regulamentação atual.

Palavras-chave: inteligência artificial, educação, ética, IA de fonte aberta, reconhecimento facial, dados pessoais, automatização, ensino personalizado