

Les enjeux éthiques de l'IA en éducation

entre nouveauté et continuité

Simon Collin

Université du Québec à Montréal



RÉSUMÉ. Depuis novembre 2022, ChatGPT a placé l'intelligence artificielle (IA) en tête des préoccupations éducatives. Bien que certains enjeux éthiques soient régulièrement discutés, ils sont souvent dissociés des recherches appliquées sur l'IA en éducation et difficiles à appréhender de manière systématique. Cet article propose une typologie des enjeux éthiques de l'IA en éducation issue d'une revue systématique de la littérature.

MOTS-CLES. Intelligence artificielle – Enjeux éthiques – Tensions

Introduction

Depuis novembre 2022, l'apparition de ChatGPT a propulsé l'intelligence artificielle (IA) au premier plan des préoccupations éducatives, dépassant de loin le seul intérêt des spécialistes (praticiens et chercheurs) du numérique en éducation. Plusieurs enjeux sont régulièrement discutés au gré des usages éducatifs de l'IA. Parmi les principaux, on retrouve, par exemple, la protection des renseignements personnels, la prévention des biais algorithmiques, ou encore l'opacité de l'IA et la complexité d'y consentir de façon libre et éclairée. Dans la mesure où ces enjeux impliquent « une valeur ou un principe moral [qui] est mis en jeu dans une question une situation » (Commission de l'éthique en science et technologie¹), ils se prêtent bien à une démarche éthique (Collin et Marceau, 2022), définie comme « une réflexion sur nos pratiques individuelles et sociales et sur les valeurs qu'elles actualisent menant à leur évaluation à travers l'exercice d'un jugement pratique et à des actions conséquentes » (Marchildon, 2017, 2^e §).

Ces enjeux éthiques sont relativement dissociés des recherches appliquées sur l'IA en éducation, de sorte qu'ils sont peu anticipés lors de la conception, de l'implantation et des usages éducatifs de l'IA. En outre, ces enjeux sont difficiles à appréhender de façon systémique dans la mesure où ils sont abordés par les chercheurs au travers de terrains éducatifs, de techniques computationnelles et d'angles d'analyse hétérogènes. Aussi, l'objectif de cet article est de donner un aperçu des enjeux éthiques de l'IA en éducation en présentant une typologie issue d'une revue systématique de la littérature (voir Collin, Lepage et Nebel, 2023).

Les deux premières sections de ce texte abordent des points qui nous semblent importants pour appréhender les enjeux éthiques de l'IA en éducation : leur continuité par rapport aux enjeux éthiques soulevés par les technologies précédentes, ainsi que leur caractère latent. La troisième section présente une typologie des enjeux éthiques de l'IA en éducation organisée en six tensions entre des dimensions éducatives et des dimensions techniques. Finalement, nous concluons en présentant quelques pistes d'action visant à rendre l'IA la plus éthique possible en éducation.

¹ <https://www.ethique.gouv.qc.ca/fr/ethique/quelques-notions-d-ethique/les-enjeux-ethiques/>

Continuités des enjeux éthiques des technologies en éducation

Les discours sociaux politiques et médiatique sur l'intelligence artificielle (IA), ont tendance à mettre l'emphasis sur les nouveautés que cette dernière apporte par rapport aux technologies précédentes. Ces discours de rupture ne sont pas propres à l'IA. Ils accompagnent toute innovation technique, tant en éducation qu'en société plus généralement. En effet, comme le rappellent Scardigli (1989) et Musso et al. (2007), le processus d'appropriation d'une innovation technique tend à débiter par des discours prophétiques dans lesquels « les fantasmes scientifico-techniques alimentent les espérances et les frayeurs. Les essais prospectifs s'accompagnent d'une surestimation de l'ampleur des changements préparés par les techniques nouvelles » (Musso et al., 2007, p. 48).

Par le fait même, ces discours tendent à négliger les continuités et les filiations qu'il est possible d'établir d'une technologie à une autre. Le même principe vaut pour les enjeux éthiques que suscite actuellement l'IA. À titre d'exemple, Richard Dyer (1997, cité dans Mackenzie et Wajcman (1999)) souligne déjà en 1997, comment la conception des techniques photographiques a été porteuse de biais raciaux.

« Les médias photographiques et, a fortiori, l'éclairage cinématographique supposent, privilègent et construisent la blancheur. L'appareil a été conçu en pensant aux personnes à la peau claire, et les usages et les instructions continuent dans la même veine, à tel point que photographier des personnes non blanches est typiquement considéré comme un problème » (p. 134, traduction libre).

Ces biais raciaux s'exprimaient notamment à travers les photos de classe des écoles américaines, qui offraient peu de nuances des visages des élèves noirs américains.

Dans ce sens, les biais algorithmiques dont est porteuse l'IA en éducation ne sont pas nouveaux. Ils nous rappellent l'intérêt de tirer profit des continuités techniques et des questionnements éthiques précédents pour cerner plus précisément les nouveautés induites par l'IA.

Latence des enjeux éthiques de l'IA en éducation

Un deuxième point qui nous semble important pour situer les enjeux éthiques de l'IA en éducation est leur caractère latent. En effet, les enjeux éthiques sont toujours potentiels par nature et s'actualisent ou non suivant la rencontre singulière d'un système d'IA, et d'une situation éducative eux-mêmes singuliers. Plus précisément, les enjeux éthiques, s'actualisent lorsqu'un système d'IA interfère avec les valeurs, les finalités et les pratiques d'une situation éducative.

L'usage de ChatGPT par les étudiants universitaires en représente un bon exemple. En effet, cette IA générative pose des enjeux considérables dans la mesure où son usage par les étudiants est susceptible d'interférer avec le développement des compétences visées par leur formation. C'est notamment le cas si un étudiant demande à ce système de rédiger à sa place un travail universitaire. À l'inverse, un étudiant peut utiliser ChatGPT pour aller plus loin dans la maîtrise de ses compétences, par exemple, en lui demandant d'identifier des voix de bonification possibles de son travail, en lui demandant de lui poser des questions sur la thématique d'un cours, etc. Ainsi, les enjeux que soulève ChatGPT sont latents et s'actualisent ou non selon chaque situation éducative.

Les enjeux éthiques de l'IA et de l'éducation en six tensions

Une revue systématique de la littérature que nous avons menée sur les enjeux éthiques et critiques de l'IA en éducation (voir Collin, Lepage et Nobel, 2023) nous a amenés à catégoriser ces derniers en six tensions (voir Figure 1). Au même titre que les enjeux éthiques, les tensions entre l'IA et l'éducation sont d'abord latentes. Elles s'actualisent et deviennent problématiques lorsque l'IA interfère avec les valeurs, les finalités ou les pratiques d'une situation éducative.

La première tension concerne la complexité des situations éducatives, d'une part, et la standardisation technique de l'IA, d'autre part. Lorsque la seconde interfère avec la première, elle soulève le risque que les propositions suggérées par l'IA soient moyennes et, par conséquent, peu adaptées aux singularités de chaque situation éducative. Cette tension regroupe des enjeux éthiques liés à la réduction factice des réalités scolaires, ainsi qu'à l'industrialisation de l'éducation.

La deuxième tension porte sur l'agentivité des acteurs scolaires, d'un côté, et l'automatisation technique de l'IA, de l'autre. Les enjeux éthiques concernent alors la limitation du choix, du jugement et des compétences des acteurs scolaires, parfois au profit d'autres acteurs, tels que les gestionnaires scolaires et les acteurs techniques.

La troisième tension met en relation la justice scolaire et la rationalité technique et soulève le risque de sous-estimer les implications scolaires et sociales que peut entraîner l'intégration de l'IA en éducation. Les enjeux éthiques regroupés sous cette section sont, par exemple, la recherche démesurée de l'efficacité aux dépens d'autres valeurs éducatives jugées plus souhaitables, les biais algorithmiques et les discriminations scolaires qui peuvent en découler, ou encore l'hégémonie ethnoculturelle que peut alimenter l'IA.

La quatrième tension porte sur la gouvernance scolaire et la conception technique de l'IA. Elle regroupe des enjeux éthiques tels que le manque de cadre éthique et de responsabilité morale et légale des différentes parties prenantes, ce qui peut résulter en une gouvernance lacunaire de l'IA en éducation. Le risque ultime de cette tension est que la régulation scolaire soit davantage soumise à des logiques techniques et commerciales, plutôt que des logiques proprement éducatives et publiques.

La cinquième tension a trait au besoin d'intelligibilité des acteurs scolaires par rapport à l'opacité technique de l'IA. Elle recouvre des enjeux éthiques liés à un manque de transparence de l'IA en éducation, entraînant une difficulté d'évaluer, d'assumer et d'expliquer les propositions suggérées par l'IA.

Finalement, la sixième tension concerne la dignité, des actrices et des acteurs scolaires et l'exploitation de leurs données par l'IA. Les enjeux éthiques qui en ressortent concernent la collecte, le traitement et l'usage des données scolaires par les acteurs techniques et la question du consentement libre et éclairé par les acteurs scolaires.

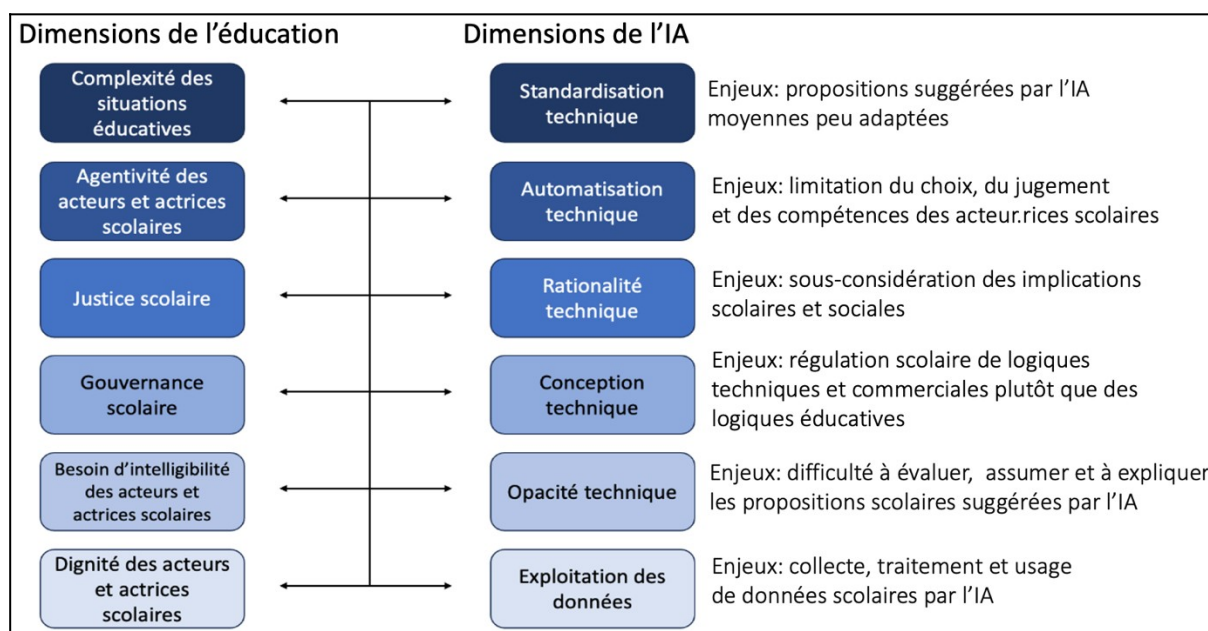


Figure 1 : Six tensions latentes entre l'IA et l'éducation et enjeux éthiques associés (adapté de Collin, Lepage et Nebel, 2023)

Pistes d'action pour une IA la plus éthique possible en éducation

Face aux enjeux éthiques que suscite l'IA en éducation se pose la question de savoir comment la rendre la plus éthique possible. De toute évidence, interdire l'IA en éducation peut apparaître comme une solution à court terme, mais elle génère des contre-effets non négligeables, en premier lieu, celui d'empêcher l'éducation des élèves aux usages éthiques de l'IA. Une solution plus ambitieuse et plus complexe consiste plutôt à monitorer les enjeux éthiques de l'IA en éducation tant au niveau de la conception, que de l'institution scolaire et des pratiques éducatives.

Sur le plan de la conception, il existe différents modèles de conception éthique de l'IA (*ethics by design*, voir p. ex., Iphofen et Kritikos, 2021), dont la caractéristique commune est d'impliquer les acteurs scolaires dès le début du processus de conception et d'articuler ce dernier autour de leurs contextes et de leurs besoins autant que possible. Bien qu'ils ne soient pas sans limite, ces modèles de conception sont susceptibles d'améliorer la qualité et l'acceptabilité éducatives de l'IA, notamment en augmentant la pertinence éducative des technologies d'IA en éducation, ainsi que leur transparence et leur accessibilité, par exemple, en encourageant des solutions libres et ouvertes.

Toutefois, il serait déraisonnable de confier aux seuls acteurs techniques la responsabilité de monitorer l'éthique de leurs produits, aussi bien intentionnés soient-ils. Il est donc nécessaire de compléter les initiatives des acteurs techniques par des initiatives issues des institutions scolaires.

Un premier exemple d'initiative institutionnelle concerne les dispositifs de démocratie technique, telle que la Conférence de consensus sur l'utilisation du numérique en éducation qui a été menée au Québec durant l'année 2022 et qui a donné lieu au rapport *Équité et valeur ajoutée dans les usages du numérique pour l'enseignement et l'apprentissage* (rapport EVA, Beaudoin et al., 2022). Comme le montre la figure ci-dessous (Figure 2), la conférence de consensus a permis de mettre en dialogue un comité composé d'acteurs scolaires non spécialistes du numérique en éducation, des experts

scientifiques et scolaires du numérique en éducation, ainsi que l'ensemble du réseau scolaire. Elle a permis d'aboutir à des recommandations tangibles, concertées et informées.



Figure 2 : Démarche de la Conférence de consensus de consensus sur l'utilisation du numérique en éducation (tiré de Beaudoin et al., 2022)

Bien que la conférence de consensus sur l'utilisation du numérique en éducation n'ait pas porté spécifiquement sur l'IA, elle pourrait facilement être reproduite pour contribuer à orienter l'IA en éducation vers des usages plus éthiques.

Toujours sur le plan institutionnel, certains systèmes scolaires mettent en œuvre des bureaux chargés d'évaluer l'acceptabilité des technologies destinées au milieu scolaire en fonction de standards déterminés et transparents. Les compagnies technologiques qui souhaitent offrir leurs services aux milieux scolaires doivent alors répondre à ces standards et être accréditées. C'est par exemple le cas de l'initiative australienne et néo-zélandaise *Safer technologies for schools* (voir Figure 3). En plus des standards de respect de la vie privée, de sécurité et de robustesse, d'interopérabilité et d'adaptabilité à différents supports valables pour toute technologie, *Safer technologies for schools* a récemment ajouté des critères spécifiques à l'IA, tels que l'utilisation d'une charte éthique de conception de l'IA et le test régulier des modèles d'IA.

Finalement, en plus des initiatives issues des concepteurs techniques et des institutions scolaires, il est aussi possible d'aborder les enjeux éthiques de l'IA en éducation sur le plan pédagogique. Les ateliers de délibération éthique semblent particulièrement intéressants, d'une part pour statuer sur l'acceptabilité de technologies d'IA au sein des écoles, d'autre part pour former les enseignants et les apprenants à la réflexion éthique sur l'IA. Un premier exemple est le modèle techno-éthique développé par Krutka, Heath et Bret Staudt Willet (2019) à l'intention des futurs enseignants. Ce modèle permet d'évaluer l'acceptabilité éducative d'une technologie en considérant ses dimensions éthique, légale, démocratique, économique, technologique et pédagogique. Parce qu'il est exhaustif, il est également complexe à mettre en œuvre auprès d'apprenants de scolarité obligatoire.

Plus pratique, la trousse à outils *Intégrer l'éthique de l'intelligence artificielle en enseignement supérieur*² propose différentes situations pour alimenter des ateliers de délibération éthique. En revanche, on peut regretter que les situations de délibération éthique ne soit pas davantage lié à des problématiques spécifiquement éducatives. Quoiqu'il en soit ces deux exemples peuvent facilement être repris et adaptés à divers profils d'élèves.

² Voir https://poleia.quebec/wp-content/uploads/2020/02/Guide_IA_VF.pdf

Conclusion

Les pistes d'action présentées ci-dessous ne sont assurément pas exhaustives, mais elles suffisent à souligner que les milieux scolaires disposent de plusieurs leviers pour orienter l'IA vers des usages plus éthiques en éducation. Également, elles permettent d'illustrer que le monitoring des dimensions éthiques de l'IA en éducation ne peut être le fait d'un seul acteur ou d'une seule instance.

Au contraire, pour maximiser l'éthique de l'IA en éducation, il convient de mettre en responsabilité plusieurs acteurs (techniques, politiques, scolaires) depuis la conception jusqu'aux usages de l'IA en éducation. Surtout, ces pistes d'action montrent que l'éthique de l'IA en éducation ne peut pas être réduite à des principes ou à des typologies (incluant celle présentée dans ce texte), aussi exhaustifs soient-ils. Au-delà de l'identification des enjeux éthiques et de la sensibilisation des acteurs scolaires, il convient de mettre en œuvre des dispositifs de démocratisation technique comme ceux suggérés ci-haut afin que les acteurs scolaires puissent « pratiquer » l'éthique de l'IA en éducation et contribuer à orienter cette dernière en conséquence.

Références bibliographiques

- Beaudoin, J., Laferrière, T., Collin, S., Ruel, C. et Voyer, S. (2022). Rapport ÉVA : Équité et Valeur Ajoutée dans les usages du numérique pour l'enseignement et l'apprentissage. CTREQ.
- Collin, S., Lepage, A., et Nebel, L. (2023). Enjeux éthiques et critiques de l'intelligence artificielle en éducation: une revue systématique de la littérature. *Revue canadienne de l'apprentissage et de la technologie*, 49(4), 1-29.
- Collin, S. et Marceau, E. (2022). Enjeux éthiques et critiques de l'intelligence artificielle en enseignement supérieur. *Éthique publique*, 24(2). DOI : <https://doi.org/10.4000/ethiquepublique.7619>
- Dyer, R. (1997). Making 'white' people white. In D. MacKenzie et J. Wacjman (dir., 1999), *The Social Shaping of Technology*, second edition (p. 134-137). Open University Press.
- Iphofen, R. et Kritikos, M. (2021). Regulating artificial intelligence and robotics: Ethics by design in a digital society. *Contemporary Social Science*, 16(2), 170-184.
- Krutka, D. G., Heath, M. K. et Willet, K. B. S. (2019). Foregrounding technoethics: Toward critical perspectives in technology and teacher education. *Journal of Technology and Teacher Education*, 27(4), 555-574.
- Marchildon, A. (2017). Le pouvoir de déployer la compétence éthique. *Éthique publique*, 19(1). DOI : [10.4000/ethiquepublique.2920](https://doi.org/10.4000/ethiquepublique.2920)
- Musso, P., Ponthou, L. et Seulliet, É. (2007). *Fabriquer le futur 2. L'imaginaire au service de l'innovation*. Village mondial.
- Scardigli, V. (1989). Nouvelles technologies : l'imaginaire du progrès. Dans A. Gras et S. Poirot-Delpech (dir.), *L'imaginaire des techniques de pointe. Au doigt et à l'œil* (p. 97-114). L'Harmattan.