



Full Length Research Paper

L'évaluation en Ligne à l'ère de l'Intelligence Artificielle Générative : Défis et Opportunités

Amanvon Ferdinand Atta^{1*}, Boris Stéphane Zounémé², Dagou Dangui Augustin Sylvain Legrand Koffi³

¹Université Virtuelle de Côte d'Ivoire, Unité de Recherche et d'Expertise du Numérique – Abidjan, Côte d'Ivoire

²Université Nangui Abrogoua, Laboratoire de Mathématique et d'Informatique – Abidjan, Côte d'Ivoire

³Ecole Supérieure Africaine des Technologies de l'Information et de la Communication, Laboratoire des Sciences et Technologies de l'Information et de la Communication – Abidjan, Côte d'Ivoire

Received October 2024 – Accepted December 2024



*Corresponding author. amanvon.atta@uvci.edu.ci

Author(s) agree that this article remain permanently open access under the terms of the Creative Commons Attribution License 4.0 International License.

Résumé :

L'Intelligence Artificielle Générative ou *Generative Artificial Intelligence* (GenAI) est un sous-ensemble de technologies d'IA capables de générer de nouveaux contenus, qu'il s'agisse de textes, d'images, de musique, de codes ou de données synthétiques. Cet article explore les défis et opportunités de l'intégration de l'intelligence GenAI dans l'évaluation en ligne. À travers une analyse des technologies disponibles, nous examinons comment GenAI peut transformer les processus d'évaluation en termes de feedback immédiat et impartial, tout en mettant en lumière les risques liés à l'intégrité académique, au biais algorithmique et à l'équité. Nous concluons en proposant des recommandations pour une utilisation équilibrée et éthique de GenAI dans le contexte éducatif.

Mots clés: Intelligence Artificielle Générative ; Enseignement supérieur ; GPT ; Large Language Model ; Evaluation en Ligne ; intégrité académique

Cite this article:

Amanvon Ferdinand Atta, Boris Stéphane Zounémé, Dagou Dangui Augustin Sylvain Legrand Koffi. (2024). L'évaluation en Ligne à l'ère de l'Intelligence Artificielle Générative: Défis et Opportunités. *Revue RAMReS – Sci. Appl. & de l'Ing.*, Vol. 6(2), pp. 11-15. ISSN 2630-1164.

1. Introduction

L'évaluation est un élément important de l'éducation et sert à recueillir, analyser et interpréter les progrès d'un apprenant. Ce retour d'information est essentiel pour les étudiants car il leur permet de comprendre leurs points forts et leurs points faibles dans un module. De plus, ce retour d'information est également bénéfique pour les enseignants car il peut leur permettre d'adapter leurs expériences d'enseignement et d'apprentissage à leurs étudiants.

L'évaluation en ligne fait référence à l'évaluation qui utilise la technologie pour évaluer les compétences, les connaissances et les capacités d'un étudiant via des plateformes numériques telles que des applications Web, des appareils mobiles ou des tests informatisés. Il existe divers outils d'évaluation qui peuvent être utilisés pour l'évaluation en ligne. Il s'agit notamment des choix multiples, des réponses courtes, des essais, des portfolios électroniques, de l'auto-évaluation et de l'évaluation par les pairs. Le type d'outil d'évaluation utilisé dépendra des objectifs d'apprentissage et du

niveau du module. L'évaluation en ligne peut être mise en œuvre de plusieurs manières, cela dépendra des objectifs d'apprentissage, des buts d'évaluation et de la technologie disponible. L'évaluation en ligne peut être administrée sur des systèmes de gestion de l'apprentissage tels que Moodle. Des paramètres d'évaluation peuvent également être utilisés, tels que les limites de temps, la randomisation et le nombre de fois que les étudiants peuvent passer l'évaluation.

L'évaluation en ligne présente de nombreux avantages [1]. L'un de ces avantages est la commodité et la flexibilité, qui permettent aux individus de passer les évaluations à leur propre rythme, à leur propre heure et à leur propre endroit, réduisant ainsi les contraintes de planification tout en augmentant l'accessibilité. L'évolutivité est un autre avantage de l'évaluation en ligne, car elle peut accueillir de nombreux étudiants simultanément. Les évaluations en ligne peuvent également fournir un retour d'information immédiat. Cela est utile aux étudiants qui peuvent identifier les domaines dans lesquels ils doivent s'améliorer. Elle est

également rentable car elle réduit le besoin de papier, les coûts d'impression, la notation manuelle et la surveillance. De plus, comme l'évaluation en ligne est administrée à distance, elle peut faire gagner du temps et des frais de déplacement aux étudiants. Les évaluations en ligne peuvent être conçues pour empêcher la tricherie [2] en utilisant des fonctionnalités telles que la randomisation, les limites de temps et les outils de surveillance, ce qui en fait un autre avantage.

L'évaluation en ligne dans l'éducation a augmenté partout dans le monde. Cette accélération est en partie due à la pandémie de COVID-19, qui a forcé les établissements d'enseignement supérieur à passer à l'apprentissage et à l'évaluation à distance. Bien que l'évaluation en ligne ait été mise en œuvre dans de nombreux pays, il n'existe pas encore de consensus clair quant à savoir si l'évaluation en ligne est préférable à l'évaluation traditionnelle en face à face.

L'Intelligence Artificielle Générative ou *Generative Artificial Intelligence* (GenAI) est un type de technologie d'intelligence artificielle qui peut générer des résultats nouveaux et uniques [3]. Parmi ces résultats, on peut citer les images, le texte, les audios, les vidéos et les modèles 3D. Parce qu'elle est capable de produire un contenu sophistiqué et réaliste qui ressemble à l'ingéniosité humaine, GenAI est un outil utile pour un large éventail d'industries, y compris l'éducation, le divertissement et la conception de produits [4]. Plus précisément, ces industries peuvent bénéficier de son utilisation. De récentes avancées technologiques, telles que le transformateur génératif pré-entraîné (GPT), ont permis d'améliorer considérablement ces capacités. Ces nouveaux développements ont ouvert de nouvelles portes pour l'utilisation d'applications GenAI dans l'enseignement supérieur. Parmi ces applications, citons ChatGPT, Midjourney et GitHub Copilot.

De tout ce qui précède, il est possible que les GPTs perturbent les pratiques d'évaluation en ligne dans les établissements d'enseignement; en particulier dans un contexte d'apprentissage à distance et en ligne. C'est pour cette raison que nous nous posons les deux questions suivantes :

- Quelles sont les opportunités offertes par l'intégration de l'intelligence artificielle générative ou *Generative Artificial Intelligence*(GenAI) dans l'évaluation en ligne dans le contexte de l'enseignement supérieur ?
- Quels sont les défis liés à l'utilisation de l'intelligence artificielle générative (GenAI) dans l'évaluation en ligne dans le contexte de l'enseignement supérieur ?

2. Matériels et méthodes

Cette étude repose sur une revue de la littérature des études récentes qui examinent les effets de l'IA générative dans le domaine de l'éducation. Les méthodes d'évaluation employées dans différents contextes éducatifs sont analysées pour identifier les pratiques courantes et les limites de l'utilisation des technologies d'IA.

Cette revue repose sur une méthode de recherche en deux principales phases inspirée de PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses)[5] :

- 1) Phase de sélection des travaux : Nous avons effectué des recherches dans des bases de données universitaires dans les domaines de l'intelligence artificielle générative et de l'éducation, notamment IEEE Xplore, Web of Science (WoS), Scopus et Google Scholar. La chaîne de recherche comprenait des mots-clés soigneusement choisis et des opérateurs booléens, tels que ("online assessment" OR "assessment") AND (ChatGPT OR Claude OR Bard OR Gemini OR "Generative AI"), ("higher education" OR "e-learning") AND (ChatGPT OR "Generative AI"). Nous avons procédé indépendamment à la sélection des résumés à l'aide de Rayyan, une application Web conçue pour faciliter les revues systématiques. Cette phase nous a permis de recenser 31 articles. Rayyan a facilité le processus de sélection en (1) identifiant les doublons, (2) en gérant efficacement de nombreuses références en fournissant des fonctionnalités de marquage et d'organisation des articles, (3) en permettant une révision indépendante des résumés par les trois chercheurs pour inclusion ou exclusion. Ainsi, chaque résumé identifié lors de la recherche initiale est examiné de manière indépendante par les trois chercheurs impliqués. Cette étape est cruciale pour assurer une objectivité maximale et éviter les biais personnels. Nous avons évalué les résumés selon des critères prédéfinis (pertinence par rapport aux objectifs de l'étude, qualité méthodologique, etc.) et avons déterminé s'ils doivent être inclus ou exclus. Cette révision indépendante a permis à chacun de formuler un jugement impartial avant toute discussion de groupe. La dernière étape (4) de la phase de sélection des travaux est l'étape de conservation des résumés et résolution des divergences. Ainsi, Une fois que chaque chercheur a marqué les résumés comme "inclus" ou "exclus", les résultats de leurs évaluations sont comparés. Les fonctionnalités collaboratives de l'outil Rayyan facilitent cette comparaison en mettant en évidence les désaccords entre les évaluateurs. Ces divergences sont ensuite résolues collectivement à travers des discussions, permettant d'aboutir à un consensus sur les articles à conserver. Cette approche garantit que seuls les articles les plus pertinents et conformes aux critères sont retenus pour l'analyse approfondie. Grâce aux fonctionnalités collaboratives de Rayyan, nous avons pu comparer leurs décisions et résoudre les éventuelles divergences par la discussion. Il est important de souligner que Rayyan ne prend pas de décisions de manière autonome et n'influence pas les critères de sélection.

- 2) Phase d'Analyse des travaux : Nous avons lu en profondeur l'article pour voir si ces 31 articles abordent l'une des deux questions de recherche de notre étude. A l'issue de cela nous avons pu identifier 10 articles dont la répartition en fonction des deux questions abordées est présentée par la Figure 1.

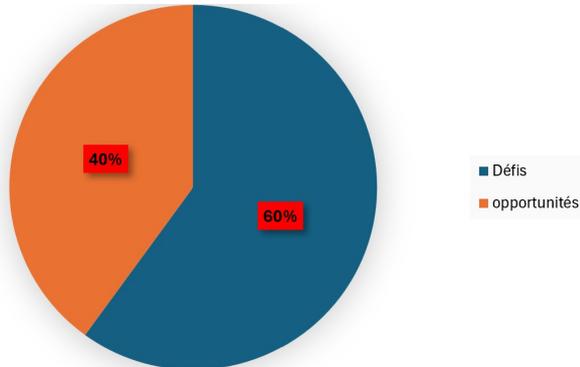


Figure 1: Proportion d'articles sur opportunités vs défis

3. Résultats

3.1. Opportunités

Comme présenté par la Figure 2, GenAI présente de multiples avantages aussi bien que pour les enseignants et les étudiants.

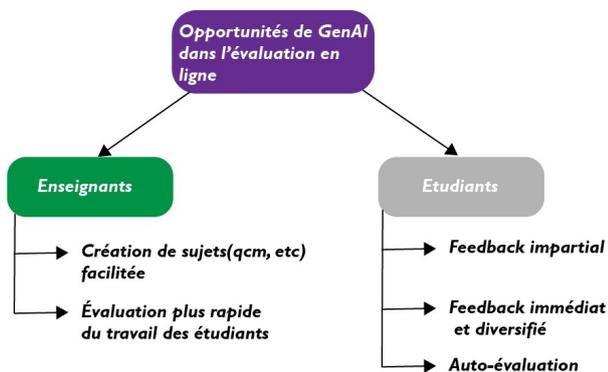


Figure 2: Opportunités de GenAI pour les étudiants vs enseignants

Au niveau des enseignants, GenAI a plusieurs atouts. GenAI facilite la mise en place d'énoncé d'évaluation, avec la possibilité de générer automatiquement des questions variées, telles que des QCM, des questions ouvertes. Cela réduit considérablement le temps nécessaire à la conception des évaluations, tout en assurant une diversité dans les types de questions. Par ailleurs, GenAI facilite la personnalisation des sujets d'évaluation [6]. En effet, GenAI permet une adaptation des examens en fonction de certains critères tels que les niveaux d'apprentissage, les compétences et besoins individuels des étudiants. Un autre avantage pour les enseignants est que GenAI offre également une correction automatique des tests (ex: QCM, QCU). Cela permet aux enseignants de fournir des feedbacks quasi-instantanés aux apprenants et d'effectuer des analyses poussées de leurs réponses.

Concernant les étudiants également il y a plusieurs opportunités qui sont identifiées. L'un des principaux

avantages est la possibilité de recevoir un feedback (ou retour) instantané [7], [8]. Cela permet aux apprenants de s'apercevoir très tôt de leurs manques de connaissances et de les combler rapidement. Ce retour immédiat est essentiel pour favoriser un apprentissage plus réactif, où les étudiants peuvent modifier leur niveau de connaissance de manière autonome. De plus, cette technologie offre une variété de retours. Au fait, en présentant aux apprenants différentes perspectives et méthodes pédagogiques sur un même sujet, cela renforce leur compréhension en favorisant une vision globale des contenus d'apprentissage.

En outre, les étudiants considèrent l'évaluation via GenAI comme une option incontournable, car le retour fourni par un GPT est généralement moins critique et plus objectif que celui fourni par un enseignant. Cette attitude renforce la confiance des étudiants dans les résultats de leurs évaluations et diminue le stress lié à la réception d'un jugement de la part d'un maître. Finalement, GenAI favorise l'évaluation personnelle, en favorisant la réflexion des étudiants [9]. L'auto-évaluation permet aux étudiants de développer des compétences d'apprentissage autonome, qui sont indispensables dans un environnement éducatif où l'autonomie et la capacité à s'améliorer jouent un rôle dans le succès des apprenants.

3.2. Défis

A côté de la pluralité des opportunités offertes par la GenAI dans l'évaluation en ligne, il subsiste également des défis pertinents à relever.

Tout d'abord, sur l'aspect de l'intégrité académique, la possibilité mise en avant par GenAI de créer des réponses même plus complexes et cohérentes augmente la probabilité de tricherie [10], [11]. Les étudiants pourraient remettre des travaux sans une bonne compréhension des sujets traités, ce qui rend difficile la préservation des standards académiques. En outre, les dispositifs conventionnels de détection du plagiat se trouvent également souvent peu opérants à l'égard des productions de GenAI, ce qui rend l'identification des œuvres non originales plus difficile. Dans le même temps, une telle forte dépendance à ces technologies risque de limiter le perfectionnement des compétences analytiques et créatives des étudiants, ce qui les conduira à utiliser des solutions d'IA à la place de leur propre pensée [12].

Un autre défi majeur qui est difficile à ignorer est les biais intégrés dans les modèles d'IA générative, qui sont connus pour s'appuyer sur des bases de données déjà biaisées [13]. Les sources de discrimination peuvent être de différentes natures : raciales, culturelles, de genre. Elles ont la possibilité de fausser les résultats en faveur de groupes particuliers tout en désavantageant d'autres, comme les minorités et les sociétés non occidentales. Cela pose un risque pour l'équité des évaluations, car les jugements des personnes qui seront générés naturellement ne seront pas représentatifs des capacités de tous les étudiants. De plus, il est également probable que l'IA générative assure que la diversité des points de vue sera réprimée

en centrant les réponses sur le contexte et les valeurs des données d'entraînement plutôt qu'en encourageant une véritable diversité d'opinions.

L'évaluation basée sur l'IA générative dans les plateformes d'enseignement en ligne peut également aggraver l'inégalité numérique déjà existante [14]. Tous les étudiants ne sont sur un même pied d'égalité, car certains ont davantage accès à des outils numériques et à une infrastructure que d'autres, créant cette disparité, en particulier dans les contextes socio-économiques moins favorisés. Ainsi, la fracture numérique, qui comprend le manque d'accès à Internet et aux technologies avancées, peut affecter certains étudiants de manière disproportionnée, entraînant des écarts de performance et des contraintes dans leurs opportunités d'apprentissage. Les systèmes éducatifs ont la responsabilité de créer un accès égal à l'évaluation numérique pour tous les étudiants afin d'éviter les écueils liés à l'infrastructure.

En même temps, l'intégration de la GenAI dans l'évaluation exige que les étudiants soient formés correctement afin qu'ils puissent l'utiliser de manière efficace et éthique [15]. La lenteur des outils de GenAI peut entraîner des erreurs d'interprétation ou une exploitation inconsidérée des résultats. Le programme doit inclure le concept de « culture générative de l'IA » pour que les étudiants sachent comment fonctionnent les algorithmes, posent des questions critiques et spécifiques et évaluent les réponses trouvées. Ces compétences des étudiants sont essentielles pour minimiser les risques d'utilisation abusive et pour encourager la réflexion critique sur les produits de l'IA générative.

Enfin, on a les aspects éthiques concernant la confidentialité des données et les questions de propriété intellectuelle [15], [16]. Les étudiants, souvent peu conscients des possibilités de protection des données, peuvent partager des informations personnelles. Les risques liés aux violations de données sont cependant à peine mentionnés. Ainsi, les étudiants n'accordent pas beaucoup d'attention à la sécurité de leurs informations personnelles et les partagent sur les médias sociaux.

3.3. Recommandations

L'évolution du paysage de l'éducation en ligne nécessite de repenser les stratégies d'évaluation pour tenir compte des capacités de l'IA générative. Alors que les formats d'évaluation traditionnels sont confrontés aux défis posés par les outils d'IA, les enseignants sont encouragés à explorer des méthodes alternatives qui peuvent tirer parti des atouts des GPTs tout en atténuant ses risques. Ainsi, sur la base des travaux de [16], nous faisons quelques recommandations adaptées aux évaluations en ligne dans le contexte de l'enseignement supérieur à distance :

- Intégration de méthodes d'évaluations multiples : En ligne, il est possible d'utiliser une variété de formats d'évaluation, tels que des quiz interactifs, des projets collaboratifs, des discussions en forum, et

des présentations vidéo. Cela permet de répondre à différents styles d'apprentissage.

- Promouvoir des évaluations authentiques : Les évaluations en ligne peuvent inclure des études de cas, des simulations et des projets basés sur des problèmes réels, où les étudiants doivent appliquer leurs connaissances dans des contextes pratiques, même à distance.

- Assurer l'intégrité académique : Des questions ouvertes et des évaluations basées sur des projets peuvent aider à maintenir l'authenticité du travail des étudiants, tout en intégrant des éléments de réflexion personnelle.

- Prioriser les compétences interpersonnelles : Les évaluations en ligne peuvent inclure des travaux de groupe, des discussions en ligne et des présentations virtuelles, permettant aux étudiants de développer et de démontrer leurs compétences en communication et en collaboration.

4. Conclusion

L'intégration de l'intelligence artificielle générative dans les évaluations en ligne de l'enseignement à distance présente à la fois des défis importants et des opportunités passionnantes. Si les préoccupations concernant l'intégrité académique et la dépendance des étudiants aux outils d'IA sont légitimes, le potentiel d'apprentissage personnalisé, de processus administratifs simplifiés et de stratégies d'évaluation innovantes ne peut être négligé. Alors que les établissements d'enseignement évoluent dans ce paysage complexe, il est impératif d'aborder l'utilisation des outils de GenAI avec une perspective équilibrée, en reconnaissant à la fois son potentiel de transformation et les responsabilités éthiques qui accompagnent sa mise en œuvre. La recherche et le dialogue continus seront essentiels pour façonner l'avenir des évaluations en ligne d'une manière qui améliore les résultats scolaires tout en respectant les principes d'intégrité académique.

REFERENCES

- [1] A. Al-Maqbali and R. M. Raja Hussain, 'The impact of online assessment challenges on assessment principles during COVID-19 in Oman', *J. Univ. Teach. Learn. Pract.*, vol. 19, no. 2, pp. 73–92, Dec. 2022, doi: 10.53761/1.19.2.6.
- [2] T.-Y. Lee and I. Aslam, 'Policy Review: Academic Cheating in Online Examinations during the COVID-19 Pandemic', *J. Sci. Res. Rep.*, pp. 1–6, Jan. 2023, doi: 10.9734/jsrr/2023/v29i11720.
- [3] T. K. F. Chiu, 'Future research recommendations for transforming higher education with generative AI', *Comput. Educ. Artif. Intell.*, vol. 6, p. 100197, Jun. 2024, doi: 10.1016/j.caeai.2023.100197.
- [4] M. Castelli and L. Manzoni, 'Special Issue: Generative Models in Artificial Intelligence and Their Applications', *Appl. Sci.*, vol. 12, no. 9, p. 4127, Apr. 2022, doi: 10.3390/app12094127.
- [5] S. Mateo, 'Procédure pour conduire avec succès une revue de littérature selon la méthode PRISMA',

- Kinésithérapie Rev.*, vol. 20, no. 226, pp. 29–37, Oct. 2020, doi: 10.1016/j.kine.2020.05.019.
- [6] Q. Xia, X. Weng, F. Ouyang, T. J. Lin, and T. K. F. Chiu, ‘A scoping review on how generative artificial intelligence transforms assessment in higher education’, *Int. J. Educ. Technol. High. Educ.*, vol. 21, no. 1, p. 40, May 2024, doi: 10.1186/s41239-024-00468-z.
- [7] Y. Fu, ‘A research of the impact of ChatGPT on education’, *Appl. Comput. Eng.*, vol. 35, no. 1, pp. 26–31, Feb. 2024, doi: 10.54254/2755-2721/35/20230354.
- [8] E. Kasneci et al., ‘ChatGPT for good? On opportunities and challenges of large language models for education’, *Learn. Individ. Differ.*, vol. 103, p. 102274, Apr. 2023, doi: 10.1016/j.lindif.2023.102274.
- [9] H. Crompton and D. Burke, ‘The Educational Affordances and Challenges of ChatGPT: State of the Field’, *TechTrends*, vol. 68, no. 2, pp. 380–392, Mar. 2024, doi: 10.1007/s11528-024-00939-0.
- [10] B. Guo et al., ‘How Close is ChatGPT to Human Experts? Comparison Corpus, Evaluation, and Detection’, 2023, *arXiv*. doi: 10.48550/ARXIV.2301.07597.
- [11] K. Elkins and J. Chun, ‘Can GPT-3 Pass a Writer’s Turing Test?’, *J. Cult. Anal.*, vol. 5, no. 2, Sep. 2020, doi: 10.22148/001c.17212.
- [12] A. D. Samala, X. Zhai, K. Aoki, L. Bojic, and S. Zikic, ‘An In-Depth Review of ChatGPT’s Pros and Cons for Learning and Teaching in Education’, *Int. J. Interact. Mob. Technol. IJIM*, vol. 18, no. 02, pp. 96–117, Jan. 2024, doi: 10.3991/ijim.v18i02.46509.
- [13] K. Naidu and K. Sevnarayan, ‘ChatGPT: An ever-increasing encroachment of artificial intelligence in online assessment in distance education’, *Online J. Commun. Media Technol.*, vol. 13, no. 3, p. e202336, Jul. 2023, doi: 10.30935/ojcm/13291.
- [14] B. Li, V. L. Lowell, C. Wang, and X. Li, ‘A systematic review of the first year of publications on ChatGPT and language education: Examining research on ChatGPT’s use in language learning and teaching’, *Comput. Educ. Artif. Intell.*, vol. 7, p. 100266, Dec. 2024, doi: 10.1016/j.caeai.2024.100266.
- [15] S. Beckingham, J. Lawrence, S. Powell, and P. Hartley, *Using Generative AI Effectively in Higher Education: Sustainable and Ethical Practices for Learning, Teaching and Assessment*, 1st ed. London: Routledge, 2024. doi: 10.4324/9781003482918.
- [16] C. K. Y. Chan and T. Colloton, *Generative AI in Higher Education: The ChatGPT Effect*, 1st ed. London: Routledge, 2024. doi: 10.4324/9781003459026.