

Revue de littérature (2024–2025)

Les IA génératives en éducation : usages et limites

28 janvier 2026

1 Introduction

Les IA génératives (IAG) fondées sur des grands modèles de langage (*large language models*, LLM) se sont rapidement installées dans les pratiques d'étude et d'enseignement, notamment via des assistants conversationnels. La littérature récente (2024–2025) documente à la fois (i) des usages centrés sur l'aide à l'apprentissage, l'écriture et la production de ressources et (ii) des risques liés à la fiabilité, à l'intégrité académique, à la protection des données et à la dépendance cognitive.

2 Méthodologie de recherche (bornée à 2024–2025)

Cette revue synthétise des travaux publiés en 2024 et 2025, repérés sur des dépôts/archives de recherche (principalement *HAL* et *ACL Anthology*, en complément d'*arXiv*). Les références retenues couvrent des (a) revues/synthèses, (b) études empiriques sur les usages étudiants, et (c) travaux méthodologiques sur la littérature à l'IAG. En complément, des documents institutionnels français (cadres d'usage et rapports d'inspection) permettent de situer les enjeux de gouvernance, de conformité et de déploiement (MINISTÈRE DE L'ÉDUCATION NATIONALE, DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE 2025 ; CORNU et al. 2025 ; PASCAL et al. 2025).

3 Usages documentés des IAG en éducation

Des synthèses hors arXiv apportent des éclairages complémentaires, notamment sur l'école (primaire/secondaire) et la formation des enseignants. Par exemple, une revue systématique (K–12) discute des opportunités (personnalisation, motivation, évaluation) et des défis (risques éthiques, fiabilité, encadrement) (KHAN et al. 2025). Une autre revue de littérature en formation des enseignants souligne les questions d'appropriation, d'alignement avec les objectifs pédagogiques et de gouvernance des usages (MARQUES 2024).

3.1 Aide à l'apprentissage et au travail scolaire

Plusieurs études empiriques décrivent des usages d'assistance (explication, reformulation, exemples, planification) et des variations selon l'âge, la discipline et les compétences numériques. Par exemple, une enquête auprès de 395 élèves/étudiants (France/Italie) met en évidence un usage massif des LLM, mais aussi une évaluation critique parfois limitée (p. ex. relecture et vérification insuffisantes chez les plus jeunes), ainsi que des écarts de genre pouvant annoncer une fracture de littératie IAG (SUBLIME et RENNA 2024).

3.2 Perceptions, acceptabilité et cadrage pédagogique

Les perceptions des apprenants vis-à-vis des IAG, leur acceptabilité et leurs représentations éthiques influencent fortement les modalités d'intégration en classe. Une étude combinant analyses qualitative et quantitative discute notamment les enjeux d'usage « contrôlé » en contexte de classe et les défis d'adoption (formation, engagement enseignant, transfert au-delà des disciplines techniques) (ALTARES-LÓPEZ et al. 2024).

3.3 Conception d'activités et soutien à l'enseignement (enseignement supérieur)

En contexte d'enseignement supérieur, des travaux sur HAL rapportent des retours d'expérience et des recommandations pour « enseigner (avec) » l'IA, en soulignant la nécessité de scénarios explicites (objectifs, consignes d'usage, exigences de traçabilité), ainsi que le besoin d'accompagnement à l'esprit critique (MANY 2024a; MANY 2024b).

3.4 Évaluation et intégrité académique

La littérature 2024–2025 insiste sur la tension entre l'IAG comme outil d'étayage et l'IAG comme substitut au travail cognitif de l'apprenant. Des revues ciblées sur des domaines (p. ex. programmation/informatique) documentent des bénéfices (feedback, gain de temps, aide au débogage) mais aussi des effets indésirables (surconfiance, apprentissage superficiel, contournement des consignes, triche) et appellent à des conceptions d'évaluation davantage centrées sur le processus (justification, traces, oral, itérations) (ELNAFFAR et al. 2025; BERNSTEIN et al. 2025).

En parallèle, des travaux en sciences du langage informatique discutent les limites des détecteurs de texte généré (faux positifs, généralisation, attaques), ce qui est particulièrement critique en contexte éducatif où les erreurs d'attribution ont un coût humain et institutionnel (WU et al. 2025).

4 Limites et risques fréquemment discutés

4.1 Fiabilité, hallucinations et qualité épistémique

Les IAG peuvent produire des erreurs factuelles, des raisonnements fallacieux, ou des références inventées, ce qui impose des pratiques systématiques de vérification et une formation

à l'évaluation de sources. Cela motive notamment des propositions de cadres de « littératie IAG » visant à articuler sélection d'outil, formulation de requêtes, et validation critique des sorties (ZHANG et MAGERKO 2025).

Enfin, l'évaluation des LLM « en tant qu'enseignants » fait l'objet de protocoles dédiés, qui rappellent que la performance générale d'un modèle ne prédit pas nécessairement ses compétences pédagogiques (questionnement, feedback adaptatif) (SHI et al. 2025).

4.2 Biais, équité et effets distributifs

Des travaux empiriques relèvent des disparités d'usage (p. ex. par genre ou niveau) et soulignent le risque d'amplifier des inégalités (accès aux outils, qualité des prompts, capacité à vérifier/éditer) (SUBLIME et RENNA 2024).

4.3 Protection des données et conformité

La littérature insiste sur les risques liés à la saisie de données personnelles/sensibles dans des services tiers et sur la nécessité de politiques institutionnelles, de chartes d'usage et d'alternatives hébergées/maîtrisées quand cela est requis (MINISTÈRE DE L'ÉDUCATION NATIONALE, DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE 2025; CORNU et al. 2025).

4.4 Dépendance cognitive et « délégation » de l'effort

Plusieurs synthèses et revues de domaine mentionnent la sur-délégation de tâches (rédaction, résolution de problèmes) comme un risque central, susceptible de réduire l'effort d'élaboration et de favoriser un apprentissage superficiel si l'activité n'est pas scénarisée (p. ex. obligation d'argumenter, de comparer des solutions, d'expliquer ses choix) (ELNAFFAR et al. 2025; BERNSTEIN et al. 2025).

5 Cadres d'analyse et leviers d'intégration

5.1 Former à la littératie IAG

Au-delà des auto-déclarations, des instruments de mesure apparaissent pour objectiver des compétences (comprendre les limites, évaluer une sortie, collaborer étape par étape). Par exemple, le test GLAT propose une mesure standardisée de littératie IAG et montre des liens avec la performance sur des tâches soutenues par IAG (JIN et al. 2024).

5.2 Concevoir des tâches « résistantes » à la génération automatique

Les travaux convergent vers des principes de conception : (1) tâches contextualisées (données locales, expériences de terrain), (2) évaluation du raisonnement (brouillons, traces, justification), (3) oraux et interactions, (4) rubriques explicites distinguant usage autorisé/interdit, (5) activités d'analyse critique d'une réponse d'IAG (détection d'erreurs, biais, omissions).

6 Conclusion

La littérature 2024–2025 issue de dépôts de recherche montre une adoption rapide des IAG en soutien à l'étude, à l'écriture et à certaines formes de tutorat. Elle met toutefois en évidence des risques robustes (fiabilité, biais, intégrité, inégalités, dépendance), suggérant que la valeur éducative des IAG dépend moins de l'outil que des scénarios pédagogiques, de l'évaluation et de la formation à la littératie IAG.

Références

- ALTARES-LÓPEZ, Sergio et al. (2024). *Qualitative and quantitative analysis of student's perceptions in the use of generative AI in educational environments*. arXiv. arXiv : 2405.13487 [cs.CY].
- BERNSTEIN, Seth et al. (2025). *Beyond the Benefits : A Systematic Review of the Harms and Consequences of Generative AI in Computing Education*. arXiv. arXiv : 2510.04443 [cs.CY].
- CORNU, Bernard et al. (oct. 2025). *Intelligence artificielle et établissements scolaires*. Inspection générale de l'éducation, du sport et de la recherche (IGÉSR). URL : <https://www.enseignementsup-recherche.gouv.fr/sites/default/files/2025-10/intelligence-artificielle-et-etablissements-scolaires-101583.pdf>.
- ELNAFFAR, Said, Farzad RASHIDI et Abedallah Zaid ABUALKISHIK (2025). *Teaching with AI : A Systematic Review of Chatbots, Generative Tools, and Tutoring Systems in Programming Education*. arXiv. arXiv : 2510.03884 [cs.HC].
- JIN, Yueqiao et al. (2024). *GLAT : The Generative AI Literacy Assessment Test*. arXiv. arXiv : 2411.00283 [cs.HC].
- KHAN, Muhammad, Muhammad Asif QURESHI, Mustafa ŞAHIN et al. (juin 2025). « Generative Artificial Intelligence (GAI) in Teaching and Learning Processes at the K-12 Level : A Systematic Review ». In : *Technology, Knowledge and Learning*. DOI : 10.1007/s10758-025-09853-7. URL : <https://link.springer.com/article/10.1007/s10758-025-09853-7>.
- MANY, Holly (2024a). *Enseigner (avec) l'IA dans l'enseignement supérieur*. HAL (communication). Référencé via la page auteur HAL. URL : <https://cv.hal.science/holly-many>.
- (2024b). *Usages raisonnés de l'intelligence artificielle générative pour l'enseignement du FLE/FLS*. HAL (communication). Référencé via la page auteur HAL. URL : <https://cv.hal.science/holly-many>.
- MARQUES, M. (2024). « Generative Artificial Intelligence in Teacher Education : Systematic Literature Review ». In : *EDULEARN24 Proceedings*, p. 2503-2511. DOI : 10.21125/edulearn.2024.0686. URL : <https://library.iated.org/view/MARQUES2024GEN>.
- MINISTÈRE DE L'ÉDUCATION NATIONALE, DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE (2025). *Cadre d'usage de l'intelligence artificielle en éducation*. Document institutionnel (PDF). URL : <https://www.education.gouv.fr/media/221135/download>.

- PASCAL, Christine et al. (oct. 2025). *L'intelligence artificielle dans l'enseignement supérieur et la recherche : enjeux et perspectives*. Inspection générale de l'éducation, du sport et de la recherche (IGÉSR). URL : https://www.enseignementsup-recherche.gouv.fr/sites/default/files/2025-10/rapport_ia_enseignement_sup_et_recherche_igesr_101109.pdf.
- SHI, Yao, Rongkeng LIANG et Yong XU (juill. 2025). « EducationQ : Evaluating LLMs' Teaching Capabilities Through Multi-Agent Dialogue Framework ». In : *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1 : Long Papers)*. Vienna, Austria : Association for Computational Linguistics, p. 32799-32828. DOI : 10.18653/v1/2025.acl-long.1576. URL : <https://aclanthology.org/2025.acl-long.1576/>.
- SUBLIME, Jérémie et Ilaria RENNA (2024). *Is ChatGPT Massively Used by Students Nowadays ? A Survey on the Use of Large Language Models such as ChatGPT in Educational Settings*. arXiv. arXiv : 2412.17486 [cs.CY].
- WU, Junchao et al. (mars 2025). « A Survey on LLM-Generated Text Detection : Necessity, Methods, and Future Directions ». In : *Computational Linguistics* 51.1, p. 275-338. DOI : 10.1162/coli_a_00549. URL : <https://aclanthology.org/2025.cl-1.8/>.
- ZHANG, Chengzhi et Brian MAGERKO (2025). *Generative AI Literacy : A Comprehensive Framework for Literacy and Responsible Use*. arXiv. arXiv : 2504.19038 [cs.HC].