

Intelligence Artificielle en Littérature : Entre innovation créative et risques de standardisation

Artificial Intelligence in Literature: Between Creative Innovation and Risks of Standardization

OULDAMMAR Hassina

Université d'Abbas Laghrour / Khenchela,
Laboratoire ILLAAC (Algérie)

h.ouldammar@univ-khenchela.dz

Résumé:

Cet article propose une revue critique et synthétique des recherches récentes sur l'intégration de l'intelligence artificielle (IA) en littérature, en examinant ses impacts sur la création, l'analyse critique, et les pratiques pédagogiques. Centré sur l'authenticité et la diversité culturelle, il postule une collaboration homme-machine enrichissante sans altérer la richesse littéraire. Il traite les enjeux éthiques et juridiques (standardisation, biais algorithmiques, propriété intellectuelle) et évalue les perspectives d'un champ hybride, prônant vigilance critique et adaptations méthodologiques/législatives.

Abstract :

This article offers a critical and synthetic review of recent research on the integration of artificial intelligence (AI) in literature, examining its impacts on creation, critical analysis, and pedagogical practices. Centered on authenticity and cultural diversity, it posits an enriching human-machine collaboration without altering literary richness. It addresses ethical and legal issues (standardization, algorithmic biases, intellectual property) and evaluates the prospects of a hybrid field, advocating critical vigilance and methodological/legislative adaptations.

informations sur l'article

Reçu

06 Janvier 2026

Acceptation

28 Avril 2026

Mots clés:

- ✓ Intelligence artificielle
- ✓ Création littéraire
- ✓ Diversité culturelle
- ✓ Biais algorithmiques

Article info

Received

January 06.2026

Accepted

April 28.2026

Keywords:

- ✓ Artificial Intelligence
- ✓ Literary Creation
- ✓ Cultural Diversity
- ✓ Algorithmic Biases

1. INTRODUCTION

L'intelligence artificielle (AI) s'inscrit aujourd'hui comme un acteur incontournable dans le domaine littéraire, suscitant autant d'enthousiasme que d'inquiétudes. D'une part, par la génération automatisée de textes, les outils d'IA bouleversent les modes traditionnels d'écriture et questionnent profondément la place de l'humain dans la création. D'autre part, au-delà de la création, l'IA offre des méthodes inédites d'analyse textuelle à grande échelle et influence les pratiques pédagogiques par la personnalisation et l'automatisation. Néanmoins, ces innovations soulèvent des enjeux éthiques, juridiques et culturels majeurs liés aux droits d'auteur, à la diversité culturelle et aux biais algorithmiques.

Dans ce contexte complexe, cette étude s'appuie sur une méthodologie rigoureuse combinant plusieurs approches complémentaires. En premier lieu, nous avons conduit une revue systématique de la littérature scientifique récente, permettant d'identifier les avancées les plus significatives dans le domaine. En second lieu, nous avons analysé des études empiriques menées dans des contextes académiques contrôlés, notamment les expériences menées à l'Université d'Aix-Marseille, telles que l'Atelier IA & EDUcation (IA-ÉDU, 2025), qui constituent un fondement empirique important. Enfin, nous avons procédé à un examen comparatif des cadres juridiques internationaux pour appréhender la diversité des réponses réglementaires. Cette approche plurielle a été complétée par des recherches internationales publiées

dans des revues de premier plan telles que Nature, MIT Press, ACM et PNAS, garantissant ainsi la robustesse de nos conclusions.

Cette recherche interroge une question centrale : comment intégrer l'IA dans la création, l'analyse et la pédagogie littéraires sans compromettre l'authenticité créative, la diversité culturelle et la réflexivité critique qui fondent la pratique littéraire ? Plus spécifiquement, quelles conditions méthodologiques, éthiques et juridiques permettraient une synergie équilibrée entre humains et systèmes artificiels dans le domaine littéraire ?

Nous postulons qu'une collaboration homme-machine réfléchie peut enrichir la littérature à condition de respecter trois principes fondamentaux : le maintien de l'intentionnalité créative humaine comme principe directeur du processus créatif ; l'adoption de protocoles structurés et validés empiriquement pour minimiser les biais algorithmiques et préserver la diversité des expressions culturelles ; et l'établissement de cadres juridiques et pédagogiques adaptés qui reconnaissent la spécificité de la création assistée par l'IA tout en préservant les droits des créateurs.

Cet article apporte trois contributions majeures. Sur le plan théorique, nous proposons un cadre multidimensionnel de la collaboration humain-IA dépassant les modèles binaires simplistes. Sur le plan empirique, nous synthétisons les résultats d'études contrôlées récentes (2024-2025) documentant les effets mesurables de différents modes de collaboration. Sur le plan méthodologique, nous établissons des protocoles opérationnels validés et

identifions les limites critiques des systèmes actuels à travers une typologie rigoureuse des biais algorithmiques.

2. Méthodologie

Notre approche méthodologique repose sur trois piliers complémentaires garantissant la robustesse et la validité de l'analyse présentée dans cet article.

2.1. Revue systématique de la littérature

Nous avons conduit une revue systématique des publications scientifiques de 2024-2025 selon un protocole adapté des lignes directrices PRISMA (Page et al., 2021) (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*, en français : Éléments Rapportés de Préférence pour les Revues Systématiques et les Méta-analyses) pour les sciences humaines et sociales. La recherche documentaire a porté sur quatre bases de données académiques majeures : Web of Science, Scopus, JSTOR et ACM Digital Library. Les descripteurs utilisés incluaient : *AI AND literature* (IA ET littérature), *generative AI AND creativity* (IA générative ET créativité), *large language models AND bias* (grands modèles de langage ET biais), *human-AI collaboration* (collaboration homme-IA), et *AI AND copyright* (IA et droit d'auteur). Après élimination des doublons et application rigoureuse des critères d'inclusion (études empiriques, publications dans des revues à comité de lecture, date de publication postérieure à

2023), un corpus de 47 articles a été retenu pour une analyse approfondie et systématique.

2.2. Analyse d'études empiriques contrôlées

Trois études empiriques majeures structurent le fondement de notre analyse. Premièrement, les expériences menées à l'Université d'Aix-Marseille (IA-ÉDU, 2025) ont comparé 86 textes (43 générés par ChatGPT-4 et 43 produits par des étudiants en première année de licence Lettres et écritures) selon un protocole rigoureux en double aveugle impliquant 12 évaluateurs experts (6 professeurs de littérature et 6 critiques littéraires professionnels). Deuxièmement, le projet CoAuthor développé à l'Université de Stanford (Lee, Liang & Yang, 2022) a documenté la collaboration humain-GPT-3 en conditions contrôlées, mesurant les impacts sur la productivité et la diversité lexicale. Troisièmement, l'étude de Dhillon et ses collaborateurs (2024) a examiné les effets du niveau d'assistance de l'IA sur la qualité créative auprès de 131 participants.

2.3. Analyse comparative des cadres juridiques

Nous avons procédé à une analyse comparative approfondie des cadres juridiques de quatre juridictions majeures représentant différentes traditions légales : l'Union Européenne, les États-Unis, la Chine et le Royaume-Uni. Cette analyse a examiné les textes législatifs récents,

notamment l'AI Act européen (European Union, 2024), les jurisprudences émergentes, et les documents de consultation publique de l'Organisation Mondiale de la Propriété Intellectuelle (WIPO, 2024). Cette approche comparative permet d'identifier les convergences et les divergences substantielles dans le traitement juridique des œuvres assistées par l'IA et d'éclairer les enjeux de l'harmonisation internationale.

3. État de l'art et cadre théorique

3.1. Création littéraire assistée par l'IA : avancées et limitations

L'IA générative de texte représente une avancée technologique majeure dans la création littéraire. À cet égard, les expériences menées à l'Université d'Aix-Marseille (IA-ÉDU, 2025) offrent des données empiriques particulièrement précieuses en comparant les productions de ChatGPT et d'étudiants en première année de licence (L1) Lettres et écritures. Ces travaux révèlent que l'IA excelle dans la génération de textes fluides et syntaxiquement corrects. Cependant, elle rencontre des difficultés majeures à saisir les subtilités émotionnelles, les métaphores complexes et les « non-dits » littéraires caractéristiques des œuvres humaines.

3.1.1. Protocole de recherche et résultats

Le protocole expérimental mis en œuvre s'est articulé autour d'un échantillon de 86

textes analysés, répartis équitablement entre 43 textes générés par ChatGPT-4 et 43 produits par des étudiants en L1 Lettres et écritures. L'évaluation de ces productions s'est fondée sur quatre critères principaux : l'authenticité émotionnelle (mesurée sur une échelle de 1 à 10), la cohérence narrative, la richesse métaphorique, et la présence de « non-dits » littéraires. Pour garantir l'objectivité de l'analyse, un panel de 12 experts a été constitué, comprenant 6 professeurs de littérature et 6 critiques littéraires professionnels, qui ont procédé à une évaluation en double aveugle ayant atteint un coefficient de Cohen kappa de 0,78, attestant d'une fiabilité inter-juges satisfaisante. Par ailleurs, les conditions expérimentales ont été standardisées : tous les textes partaient du même prompt initial, présentaient une longueur standardisée de 800 à 1200 mots, et relevaient du genre narratif court.

Les résultats quantitatifs obtenus sont révélateurs des forces et des faiblesses respectives de l'IA et des écrivains humains. Concernant la cohérence syntaxique, l'IA a obtenu un score de 9,2 sur 10, significativement supérieur aux 7,8 sur 10 des productions humaines ($p < 0,001$), démontrant ainsi sa maîtrise technique du langage. En revanche, l'authenticité émotionnelle a révélé une situation inverse : l'IA n'a atteint que 5,3 sur 10 contre 8,4 sur 10 pour les humains ($p < 0,001$), soulignant la difficulté des systèmes artificiels à reproduire la

profondeur émotionnelle. De même, la richesse métaphorique a favorisé les productions humaines avec 7,9 sur 10 contre seulement 4,7 sur 10 pour l'IA ($p < 0,001$). Enfin, un résultat particulièrement frappant concerne la capacité des experts à identifier les textes générés par l'IA : 87% d'entre eux ont été correctement identifiés dans les 30 premières secondes de lecture, témoignant de patterns stylistiques distinctifs et reconnaissables.

3.1.2. *Étude comparative internationale*

Au-delà de ces résultats français, des recherches complémentaires menées à l'international confirment et nuancent ces observations. Le projet CoAuthor de Lee, Liang et Yang (2022) démontre que la collaboration humain-IA peut augmenter la productivité d'écriture de 30% à 40% et la diversité lexicale de 15% à 20%, tout en maintenant une qualité littéraire comparable aux productions purement humaines. Toutefois, cette efficacité est conditionnée par le maintien d'un rôle de co-créateur actif pour l'écrivain plutôt qu'une simple posture d'éditeur de contenus générés par l'IA. Cette distinction cruciale a été établie par McGuire, De Cremer et Van de Cruys (2024) dans leur étude qui souligne l'importance du positionnement de l'auteur humain dans le processus créatif.

3.1.3. *Limites créatives identifiées*

L'analyse qualitative approfondie révèle plusieurs limitations spécifiques et

systématiques des systèmes d'IA actuels. Premièrement, ces systèmes manifestent une incapacité persistante à saisir l'ironie subtile et l'ambiguïté intentionnelle, éléments pourtant fondamentaux de nombreuses œuvres littéraires. Deuxièmement, ils montrent une tendance marquée à produire des récits génériques ou prévisibles, suivant des patterns narratifs conventionnels sans véritable originalité structurelle. Troisièmement, et peut-être plus fondamentalement, l'absence de véritable expérience émotionnelle ou de conscience existentielle limite profondément leur capacité à produire des œuvres authentiquement humaines. Enfin, ces systèmes tendent à reproduire fidèlement les patterns présents dans leurs corpus d'entraînement sans innover structurellement de manière profonde.

Un exemple récent illustre concrètement ces limites. Un journaliste a exploité les capacités de Claude 4 pour générer un roman de 500 pages en seulement sept jours (Mahier, GPT & Raphael, 2025). Cette expérience démontre que Claude, grâce à sa capacité à traiter jusqu'à 500 pages de contexte, peut effectivement assurer une cohérence narrative sur de longs textes. Néanmoins, l'auteur conclut que l'IA demeure une assistante capable de reformuler, d'adapter et d'explorer des variations stylistiques, mais qu'elle ne parvient pas à créer des univers narratifs profondément originaux sans une direction humaine substantielle et continue.

3.2. Traitement automatique du langage (NLP) et analyse littéraire

Le traitement du langage naturel révolutionne l'analyse littéraire par l'automatisation de tâches auparavant extrêmement chronophages. En effet, les techniques de tokenisation, de lemmatisation et d'analyse morphosyntaxique permettent désormais d'explorer rapidement de vastes corpus pour identifier des patterns invisibles à l'analyse traditionnelle manuelle. Cette capacité transforme radicalement les possibilités de la recherche littéraire à grande échelle.

3.2.1. Outils et méthodologies validés

Une revue systématique récente des outils d'analyse littéraire assistée par l'IA, conduite par Bolaños, Salatino, Osborne et Motta (2024), a identifié 21 plateformes spécialisées présentant des niveaux de validation scientifique variables. Parmi ces outils, plusieurs se distinguent par leur robustesse méthodologique et leur adoption par la communauté scientifique. D'abord, NLTK et spaCy s'imposent comme des standards pour l'analyse morphosyntaxique, l'identification de patterns linguistiques et l'extraction d'entités nommées. Ensuite, les modèles Transformers développés par Hugging Face (Wolf et al., 2020) permettent une analyse sémantique profonde, la détection de cohérence narrative et l'évaluation de la diversité stylistique. Par ailleurs, Semantic

Scholar offre des fonctionnalités avancées de visualisation de réseaux d'influence littéraire et d'identification de patterns intertextuels. Enfin, Elicit se spécialise dans l'extraction automatisée de données issues de milliers de textes et la synthèse thématique à grande échelle.

3.2.2. Applications concrètes

Ces outils NLP permettent d'aborder des questions de recherche auparavant inaccessibles ou excessivement laborieuses. Premièrement, ils facilitent l'identification de l'évolution stylistique d'auteurs sur plusieurs décennies grâce à l'analyse diachronique automatisée, révélant des transformations subtiles dans le vocabulaire, la syntaxe ou les thématiques privilégiées. Deuxièmement, ils rendent possible la cartographie des influences intertextuelles en calculant les similarités entre textes par des méthodes telles que la distance cosinus ou les word embeddings, permettant ainsi de visualiser des réseaux d'influence complexes. Troisièmement, ces technologies permettent l'extraction de thèmes dominants dans des corpus multilingues de plusieurs milliers d'œuvres, une tâche qui aurait été impossible par des moyens traditionnels. Quatrièmement, elles autorisent l'analyse de sentiment à grande échelle pour étudier les évolutions émotionnelles dans la littérature sur de longues périodes historiques.

3.2.3. *Limites méthodologiques*

Cependant, malgré cette puissance analytique indéniable, ces outils présentent des limites importantes qu'il convient de ne pas négliger. Une étude comparative menée par Yadav (2024), s'est penchée sur l'interprétation de Hamlet par des systèmes d'IA. Cette recherche révèle que si l'IA excelle dans l'identification de patterns thématiques récurrents tels que la vengeance, la folie ou la corruption morale, elle manque cruellement de profondeur contextuelle et de compréhension des références culturelles qui donnent tout leur sens à ces thèmes. L'analyse traditionnelle humaine demeure donc indispensable pour apporter les insights culturels, la résonance émotionnelle et la compréhension nuancée nécessaires à une interprétation complète des œuvres littéraires.

De surcroît, un risque majeur accompagne l'utilisation de ces outils. Ces analyses s'effectuent souvent selon des modèles dominants, principalement anglophones et occidentaux, réduisant ainsi la diversité des voix littéraires mondiales et risquant une homogénéisation culturelle préoccupante. Cette dimension critique nécessite une vigilance constante de la part des chercheurs utilisant ces technologies.

4. *Développement*

4.1. *Mutation de la création littéraire et redéfinition de l'auteur*

L'avènement de l'IA transforme radicalement la notion même d'auteur, concept pourtant central dans la théorie littéraire depuis des siècles. Traditionnellement, l'auteur était perçu comme un être conscient, engagé dans une démarche créative porteuse de sens subjectif et d'intentionnalité. Cette définition, héritée des théoriciens du romantisme allemand (Schlegel, 1801) et consolidée par la modernité, s'érode progressivement face à l'émergence d'IA capables de générer des textes porteurs d'émotion apparente et de complexité narrative.

4.1.1. *Évolution du paradigme créatif*

Les outils de génération de texte par l'IA, tels que ChatGPT, Claude 4 et Gemini, ont démontré une capacité remarquable à produire des textes cohérents et stylisés dans une variété impressionnante de genres et de registres. Cependant, cette prouesse technologique n'est pas sans conséquences profondes. Elle pose des questions fondamentales sur la définition même de la créativité, de l'authenticité et de la paternité intellectuelle dans le contexte numérique contemporain. En effet, si une machine peut produire un sonnet formellement parfait ou un récit émotionnellement touchant, qu'advient-il de notre conception traditionnelle de l'acte créatif comme expression d'une subjectivité consciente ?

4.1.2. Cadre multidimensionnel de la collaboration

Face à cette complexité, des recherches récentes proposent des cadres conceptuels plus sophistiqués que la simple opposition binaire entre création humaine et génération automatique. Hutson (2024) a élaboré un framework multidimensionnel qui dépasse le modèle linéaire simpliste « aucune IA ↔ IA complète ». Ce cadre novateur considère quatre axes indépendants de collaboration, chacun pouvant varier en intensité.

Le premier axe concerne la génération de contenu, s'étendant du simple brainstorming initial à la rédaction complète de sections entières. Le deuxième axe porte sur l'assistance structurelle, englobant l'organisation narrative, la cohérence globale et les transitions entre sections. Le troisième axe traite de l'input créatif proprement dit, incluant les suggestions métaphoriques, les variations stylistiques et les perspectives narratives alternatives. Enfin, le quatrième axe se concentre sur la contribution analytique, comprenant l'identification de patterns, la vérification de cohérence interne et la détection de répétitions indésirables.

4.1.3. Études de cas empiriques

Plusieurs études empiriques récentes éclairent concrètement les dynamiques de cette collaboration humain-IA. Le projet CoAuthor (Lee, Liang & Yang, 2022), s'est penché sur la co-écriture avec GPT-3 et a

révélé des résultats encourageants, mais nuancés. Cette recherche documente une augmentation substantielle de 30% à 40% de la productivité d'écriture et de 15% à 20% de la diversité lexicale. Néanmoins, ces bénéfices ne se matérialisent pleinement que si l'écrivain humain maintient son agentivité et conserve un contrôle créatif actif. La délégation excessive du processus créatif à l'IA, au contraire, tend à compromettre l'originalité et l'authenticité du résultat final.

Cette observation trouve une confirmation expérimentale dans l'étude de Dhillon et ses collaborateurs (2024). Cette recherche contrôlée, portant sur 131 participants, a démontré un effet surprenant avec un graphique en forme de U du niveau d'assistance fourni par l'IA. Concrètement, une assistance faible, opérant phrase par phrase, permet de maintenir une qualité et une productivité élevées tout en préservant l'autonomie créative du rédacteur. À l'inverse, une assistance moyenne, opérant au niveau du paragraphe, entraîne paradoxalement une baisse significative tant de la qualité que de l'engagement des participants. Enfin, une assistance élevée, générant des sections complètes, permet un retour à une qualité élevée, mais soulève des inquiétudes légitimes quant au risque de dépendance cognitive et d'atrophie des capacités d'écriture autonome.

Enfin, la recherche de McGuire De Cremer et Van de Cruys (2024) sur la co-création poétique établit l'importance cruciale du

rôle adopté par l'utilisateur humain. Cette étude distingue deux postures fondamentalement différentes : celle du co-créateur, qui travaille avec l'IA dès le début du processus créatif, et celle de l'éditeur, qui se contente de modifier des textes déjà générés par l'IA. Les résultats sont sans équivoque : les participants adoptant une posture de co-créateurs maintiennent leur créativité et leur sentiment d'auto-efficacité créative, tandis que ceux se positionnant comme simples éditeurs manifestent un déficit créatif significatif et éprouvent un sentiment d'aliénation vis-à-vis du processus créatif.

4.2. Nouvelles formes narratives et exploration créative

Paradoxalement, malgré ces défis et ces questionnements, l'IA ouvre également des perspectives créatives inédites qui méritent d'être explorées avec optimisme. En premier lieu, elle rend possibles les récits interactifs et adaptatifs qui se modifient dynamiquement selon les choix du lecteur, créant ainsi une expérience littéraire véritablement personnalisée et non-linéaire. En deuxième lieu, elle facilite l'expérimentation stylistique systématique par la génération rapide de variations multiples d'un même passage, permettant à l'auteur d'explorer efficacement différentes options narratives ou stylistiques. En troisième lieu, elle permet l'accélération de certaines étapes créatives jugées moins cruciales, libérant ainsi l'auteur pour se concentrer sur des tâches de conception

globale et de révision critique où sa contribution est la plus précieuse.

Enfin, et peut-être plus significativement encore, l'IA multimodale contemporaine ouvre la voie à la création de littératures hybrides intégrant de manière organique le texte, l'image et le son générés par l'IA. Ces nouvelles formes littéraires, encore émergentes, suggèrent que l'IA peut effectivement être un catalyseur d'innovation formelle, pourvu que l'intention créative humaine demeure le principe directeur du processus. L'émergence récente de modèles capables de générer et d'analyser simultanément différents médiums ouvre des possibilités encore plus vastes pour des œuvres où les différents modes d'expression sont co-crésés de manière intégrée et cohérente.

5. Analyse littéraire assistée par IA : opportunités et biais

5.1. Biais algorithmiques et représentation culturelle

Si l'IA facilite indéniablement l'exploration de grands corpus textuels pour identifier des patterns complexes, révolutionnant ainsi certains aspects de la critique littéraire, cette révolution s'accompagne néanmoins de défis majeurs. Ces défis concernent notamment la transparence algorithmique, c'est-à-dire la compréhension de comment les systèmes parviennent à leurs conclusions, et surtout la gestion des biais culturels qui peuvent se trouver amplifiés par ces technologies.

5.1.1. Typologie des biais dans les grands modèles de langage (LLMs)

Les recherches les plus récentes, notamment celles de Gallegos et ses collaborateurs (2024), ainsi que les travaux de Zewe (2025), ont permis d'identifier et de catégoriser plusieurs types critiques de biais algorithmiques affectant les grands modèles de langage.

Premièrement, les biais de représentation culturelle et de genre se révèlent particulièrement préoccupants. Une étude conjointe UNESCO/UCL, de Drobnjak, Shawe-Taylor et Perez Ortiz (2024), analysant les comportements de GPT-3.5, GPT-2 et Llama 2, révèle que ces systèmes associent systématiquement les prénoms féminins avec des termes tels que « famille », « enfants » et « mari », reflétant une conformité aux rôles traditionnels de genre. À l'inverse, les prénoms masculins se trouvent liés à des concepts comme « carrière », « cadres », « management » et « entreprise ». Ces biais se traduisent concrètement par la génération de narratives stéréotypées dont le caractère négatif varie selon la culture et l'orientation sexuelle des personnages dépeints.

Deuxièmement, la recherche de Zewe (2025), met en lumière un phénomène jusqu'alors méconnu : le biais de position. Ce biais se manifeste par une tendance des grands modèles de langage à sur-accentuer le début et la fin des documents tout en négligeant systématiquement les sections

centrales. Ce phénomène, causé par les mécanismes d'attention des réseaux de neurones, affecte directement la génération littéraire en privilégiant artificiellement les éléments narratifs des incipits et des dénouements au détriment du développement central, créant ainsi des structures narratives prévisibles et déséquilibrées.

Troisièmement, et de manière particulièrement troublante, l'étude de Laurito et ses collaborateurs (2025), démontre que les grands modèles de langage favorisent systématiquement les contenus générés par d'autres systèmes d'IA. Dans des tests de sélection portant sur des produits et des articles académiques, les modèles choisissaient les contenus générés par l'IA dans 62% des cas contre seulement 38% pour les contenus humains. Ce biais d'auto-préférence, désigné sous le terme d'« AI-AI Bias », pourrait marginaliser progressivement les écritures humaines dans les processus de sélection automatisés et créer des boucles d'auto-renforcement favorisant dangereusement l'homogénéisation stylistique.

5.1.2. Mécanismes techniques des biais

Au-delà de leur manifestation observable, il importe de comprendre que ces biais algorithmiques constituent un problème systémique opérant à trois niveaux techniques distincts (Gallegos et al., 2024). Au niveau des embeddings, c'est-à-dire des représentations vectorielles des mots, les

associations stéréotypées issues des corpus d'entraînement se trouvent encodées dans la structure mathématique même du modèle. Au niveau des probabilités, les distributions de probabilités des tokens reflètent mécaniquement les fréquences biaisées présentes dans les données sources. Enfin, au niveau du texte généré, ces biais encodés dans les couches profondes du modèle se manifestent concrètement à travers des choix lexicaux, syntaxiques et narratifs systématiques et prévisibles.

Cette compréhension technique est corroborée par les travaux de Caliskan, Bryson et Narayanan (2017), qui ont démontré que les modèles de représentation des mots associent automatiquement certaines professions spécifiques à des genres ou à des origines ethniques, renforçant ainsi des biais préexistants dans la société. Ces travaux pionniers ont ouvert la voie aux recherches contemporaines sur l'atténuation de ces biais.

5.2. Vers une IA critique et inclusive

Face à ces constats préoccupants, la communauté scientifique a développé plusieurs approches visant à contrer ces biais, approches désormais soutenues par des données empiriques récentes permettant d'évaluer leur efficacité respective.

5.2.1. Stratégies de mitigation validées

Les approches dites « pré-modèle », intervenant avant l'entraînement proprement dit, se révèlent particulièrement prometteuses. L'augmentation de données contrefactuelles, ou *Counterfactual Data Augmentation* (CDA), consiste à remplacer systématiquement les attributs biaisés dans le corpus d'entraînement. Par exemple, on permute méthodiquement les pronoms « il » et « elle » dans les descriptions professionnelles, forçant ainsi le modèle à apprendre des associations moins stéréotypées. Parallèlement, la diversification volontaire des corpus d'entraînement, par l'intégration délibérée de littératures minoritaires, non-occidentales et expérimentales, permet d'élargir la base de connaissances du modèle, comme l'a démontré Ferrara (2023). Enfin, l'échantillonnage équilibré, qui consiste à sur-échantillonner les textes sous-représentés et sous-échantillonner les sources dominantes, contribue à rééquilibrer les distributions d'entraînement.

Complémentairement, les approches intra-modèle et post-modèle offrent des leviers d'action additionnels. Le fine-tuning ciblé permet d'entraîner les modèles sur des corpus spécifiquement débiaisés, ajustant ainsi leurs paramètres internes. L'audit régulier des systèmes, via l'implémentation de frameworks d'évaluation continue des biais comme ceux proposés par Frontiers in Education (2025), assure une vigilance constante sur les dérives potentielles. Le

développement de modèles d'IA spécialisés et adaptés à des contextes culturels spécifiques permet de mieux respecter les particularités linguistiques et culturelles locales. La mise en place de cadres de gouvernance impliquant des représentants des communautés culturelles affectées garantit que les décisions techniques prennent en compte les perspectives des groupes concernés. Enfin, le post-processing, qui consiste à filtrer les outputs générés pour détecter et corriger les stéréotypes manifestes, constitue une dernière ligne de défense, bien qu'imparfaite.

Cependant, malgré ces avancées théoriques et méthodologiques, il convient de reconnaître que ces solutions restent largement expérimentales et nécessitent des investissements importants et soutenus en recherche et développement. Leur déploiement à grande échelle demeure un défi majeur pour l'industrie et le monde académique.

6. Intégration pédagogique de l'IA : défis et opportunités

6.1. Personnalisation et engagement pédagogique

L'intégration de l'IA dans l'enseignement littéraire offre des avantages significatifs et mesurables, désormais validés par des recherches empiriques récentes menées dans divers contextes éducatifs. Ces bénéfices touchent tant à la personnalisation des parcours

d'apprentissage qu'à l'engagement des étudiants dans les activités littéraires.

6.1.1. Bénéfices documentés

Premièrement, les systèmes d'IA permettent désormais de créer des parcours d'apprentissage véritablement personnalisés. Contrairement aux approches pédagogiques traditionnelles qui tendent vers l'uniformisation, l'IA ajuste dynamiquement les contenus selon le niveau et les besoins spécifiques de chaque apprenant. Cette personnalisation se traduit par une augmentation significative de l'engagement étudiant, comme l'ont démontré Hwang et Lee (2024). Deuxièmement, les outils d'analyse automatisée ouvrent l'accès à l'analyse littéraire avancée à des étudiants qui ne disposent pas encore d'une expertise critique approfondie, facilitant ainsi leur entrée progressive dans l'exploration de corpus complexes. Troisièmement, les systèmes d'IA génèrent des ressources pédagogiques adaptatives telles que des explications sur mesure, des résumés personnalisés et des exercices ajustés au niveau de chaque étudiant, permettant un apprentissage différencié à grande échelle qui serait impossible à mettre en œuvre par des moyens purement humains.

Ces innovations pédagogiques favorisent simultanément l'engagement des étudiants et la rétention des connaissances à moyen et à long terme. De surcroît, elles permettent aux enseignants de se libérer de

tâches répétitives pour se concentrer sur l'accompagnement individualisé et le développement de la pensée critique, compétences qui demeurent fondamentalement humaines.

6.2. Défis cognitifs et éthiques

Néanmoins, en dépit de ces bénéfices indéniables, la dépendance excessive à l'IA pose des risques sérieux et documentés par plusieurs études récentes qui méritent une attention particulière de la part des éducateurs et des décideurs institutionnels.

6.2.1. Risques identifiés

Le premier risque, et sans doute le plus préoccupant, concerne l'affaiblissement potentiel de la pensée critique. Comme l'a souligné Jo (2023), si les étudiants s'appuient excessivement sur l'IA pour l'analyse textuelle, ils risquent de ne jamais développer leurs propres capacités d'interprétation critique, compétence pourtant centrale dans la formation littéraire. Le second risque est lié à ce que la littérature scientifique désigne sous le terme de « dette cognitive » comme le montrent les travaux de Kosmyna et ses collaborateurs (2025) qui suggèrent que l'utilisation prolongée et intensive de ChatGPT et d'outils similaires peut entraîner une atrophie progressive des capacités de structuration autonome de la pensée, les utilisateurs déléguant de plus en plus ces processus cognitifs à la machine. Le troisième risque, d'ordre socio-économique, concerne les inégalités

d'accès. Seuls les étudiants et les institutions disposant d'un accès privilégié aux outils d'IA de pointe peuvent bénéficier pleinement de ces ressources, creusant ainsi dangereusement les inégalités pédagogiques entre établissements favorisés et défavorisés, ainsi qu'entre contextes socio-économiques contrastés.

6.2.2. Cadre pédagogique recommandé

Face à ces défis, les recherches menées en 2025 convergent vers la nécessité d'un encadrement strict et réfléchi de l'utilisation de l'IA en contexte pédagogique. Les recommandations issues de ces travaux s'articulent autour de plusieurs axes complémentaires. En premier lieu, l'enseignement explicite de ce que Alqasham (2023) nomme la « prompt literacy » (ou littératie des prompts) s'avère indispensable. Il s'agit de former les étudiants à comprendre comment ces outils fonctionnent réellement, quelles sont leurs capacités effectives et leurs limites intrinsèques, afin de développer un usage critique et éclairé.

En second lieu, la délimitation claire et explicite des contextes où l'IA est autorisée versus interdite constitue un impératif pédagogique. Les experts recommandent un ratio d'environ 70% d'apprentissage sans assistance de l'IA pour construire les fondations cognitives solides, et 30% d'activités où l'IA peut être mobilisée de manière encadrée (Kosmyna et *al.*, 2025). En troisième lieu, l'encadrement strict de l'usage pédagogique de l'IA doit

s'accompagner d'une documentation obligatoire de son utilisation, permettant aux enseignants de suivre comment les étudiants emploient ces outils et d'intervenir en cas d'usage problématique. Enfin, les modalités d'évaluation doivent être repensées pour se focaliser davantage sur la réflexion critique et l'originalité de la pensée plutôt que sur le produit final, qui peut désormais être partiellement ou totalement généré par l'IA.

7. Enjeux éthiques et perspectives juridiques

7.1. Question de la propriété intellectuelle

Le cadre juridique actuel du droit d'auteur peine considérablement à reconnaître et à protéger les œuvres produites de manière autonome par l'IA. En droit français et européen, la protection conférée par le droit d'auteur repose fondamentalement sur le critère d'originalité, définie comme une « empreinte de la personnalité de l'auteur » (Gautier, 1994). Or, par définition, une création générée de manière autonome par l'IA échappe à cet ancrage subjectif essentiel, car elle ne résulte pas d'une intention créative humaine consciente, critère jusqu'à présent considéré comme indispensable.

7.1.1. Cadres juridiques internationaux comparés

Les jurisprudences récentes et les évolutions législatives à travers le monde offrent des perspectives remarquablement

contrastées, créant une incertitude juridique majeure tant pour les créateurs que pour les entreprises technologiques développant ces systèmes.

Dans l'Union Européenne, l'AI Act adopté en juillet 2024 représente la première tentative de régulation comprehensive de l'IA à l'échelle d'un continent (European Union, 2024). Ce texte impose des obligations substantielles de transparence aux développeurs de modèles d'IA à usage général. Concrètement, ces obligations incluent la production d'une documentation technique détaillée, la publication d'un résumé public des contenus utilisés pour l'entraînement, et le respect scrupuleux du droit d'auteur via un mécanisme d'opt-out prévu par la Directive sur le marché unique numérique (UE) 2019/790 (Parlement européen et Conseil de l'Union européenne, 2019). Toutefois, comme le souligne une étude de Martens (2025), la gestion croissante des demandes d'opt-out crée des défis techniques considérables et risque paradoxalement de biaiser les datasets en faveur des grandes communautés linguistiques anglophones, menaçant ainsi la diversité culturelle que la législation cherchait précisément à protéger.

Aux États-Unis, l'approche demeure plus fragmentée et pragmatique. Le *US Copyright Office* (2024, 2025) maintient fermement sa position selon laquelle les œuvres générées uniquement par l'IA ne peuvent bénéficier de la protection du droit d'auteur en raison de l'absence d'auteur

humain identifiable. En revanche, les œuvres assistées par l'IA peuvent être protégées si l'intervention créative humaine est substantielle et correctement documentée. Le *Generative AI Copyright Disclosure Act* (US Congress, 2024), vise à obliger les entreprises à divulguer publiquement les datasets d'entraînement utilisés, donnant ainsi plus de contrôle aux détenteurs de droits. Parallèlement, plusieurs litiges majeurs sont actuellement en cours devant les tribunaux fédéraux, notamment *Alter v. OpenAI*, *Andersen v. Stability AI* et *Concord Music Group v. Anthropic*, dont les issues façonneront probablement la jurisprudence future.

La Chine adopte quant à elle une approche singulière et potentiellement pionnière. La Cour Internet de Beijing, dans une décision rendue en novembre 2023 qui a fait grand bruit, a reconnu le droit d'auteur sur une image générée par l'IA, attribuant explicitement la paternité à la personne ayant fourni les instructions détaillées au système (Pérot & Bouzayen, 2023). Le tribunal a considéré que cette contribution intellectuelle constituait un apport suffisant pour justifier la protection. Cette décision progressive s'accompagne néanmoins d'une régulation stricte : obligation d'étiquetage clair et transparent du contenu généré par l'IA et responsabilité explicite des développeurs pour les contenus problématiques ou illégaux produits par leurs systèmes.

Enfin, le Royaume-Uni présente un cas juridique particulièrement intéressant. Le *Copyright, Designs and Patents Act* de 1988 (Royaume-Uni, 1988) demeure l'un des rares textes législatifs à prévoir spécifiquement la protection des œuvres générées par ordinateur, attribuant les droits à « la personne qui a pris les dispositions nécessaires à la création de l'œuvre » (Royaume-Uni, 1988). Toutefois, une ambiguïté persistante demeure concernant la tension entre cette disposition ancienne et l'exigence contemporaine d'originalité définie comme la création intellectuelle de l'auteur. Face à cette incertitude, le gouvernement britannique conduit actuellement une consultation publique sur d'éventuels amendements législatifs. Cette disparité juridique internationale crée une incertitude majeure pour tous les acteurs concernés. L'Organisation Mondiale de la Propriété Intellectuelle conduit depuis 2019 un processus consultatif d'envergure ayant reçu plus de 250 contributions de parties prenantes variées. La dixième session tenue en 2024 a débattu intensivement de l'éligibilité au copyright des œuvres générées par l'IA et de l'infrastructure juridique et technique nécessaire pour gérer, licencier et faire respecter efficacement les droits à l'ère de l'IA (WIPO, 2024).

7.2. Responsabilité, transparence et impact sociétal

Au-delà de la propriété intellectuelle stricto sensu, la question de la responsabilité juridique et morale se révèle tout aussi critique et complexe. Lorsqu'une œuvre générée par l'IA reproduit, même partiellement, des éléments d'œuvres protégées par le droit d'auteur, la violation potentielle implique une chaîne complexe de responsabilités. L'utilisateur final qui a formulé le prompt, l'entreprise ayant développé et commercialisé le système d'IA, et les créateurs originaux dont les œuvres ont été utilisées, souvent sans autorisation explicite, pour entraîner le modèle se trouvent tous potentiellement concernés par cette responsabilité juridique. Cette question demeure largement sans réponse juridique claire et fait l'objet de débats intenses tant dans les tribunaux que dans les enceintes législatives.

7.2.1. Transparence de l'entraînement

Par ailleurs, la question de la transparence se trouve au cœur des préoccupations réglementaires contemporaines. L'Union européenne, dans le cadre de l'AI Act (European Union, 2024), impose une transparence substantiellement accrue. Concrètement, les créateurs de contenu et les plateformes de diffusion doivent désormais informer explicitement le public lorsqu'un contenu a été généré par l'IA, particulièrement en cas de diffusion large

ou commerciale. Cette obligation vise un double objectif : d'une part, préserver la confiance publique dans l'information et la culture, d'autre part, permettre aux lecteurs et aux consommateurs de faire des choix véritablement éclairés concernant les contenus qu'ils choisissent de consommer ou d'acquérir.

7.2.2. Représentation des minorités

Enfin, la question cruciale de la représentation des minorités dans les corpus d'entraînement demeure un enjeu éthique et social de premier plan. Si l'IA utilisée pour générer ou analyser la littérature perpétue systématiquement les stéréotypes existants et exclut structurellement les voix minoritaires, elle risque fortement d'aggraver les inégalités culturelles et sociales existantes plutôt que de contribuer à les réduire. Une approche véritablement responsable et éthique nécessite donc une vigilance constante et la mise en place de mécanismes d'audit régulier et transparent des systèmes d'IA utilisés en contexte littéraire, audit impliquant idéalement des représentants des communautés concernées.

8. Discussion

8.1. Vers une collaboration homme-machine réfléchie

Le potentiel de renouvellement de la littérature par l'IA ne doit en aucun cas occulter les risques sérieux et bien documentés de standardisation et

d'appauvrissement culturel. Dans ce contexte complexe, la collaboration homme-machine constitue une voie prometteuse, à condition toutefois que plusieurs principes fondamentaux soient fermement établis et respectés. Ces principes, désormais étayés par les recherches récentes et les protocoles validés empiriquement, doivent guider le développement futur de ce champ hybride.

8.1.1. Protocoles opérationnels recommandés

Le premier protocole recommandé concerne l'établissement d'une communication véritablement bidirectionnelle entre l'humain et la machine. Concrètement, cela implique une clarification explicite des rôles respectifs : l'humain conserve fermement l'intention narrative et les choix créatifs majeurs, tandis que l'IA fournit des variations stylistiques et des suggestions structurelles. Pour maintenir ce contrôle humain, il est préférable de procéder par itérations courtes, typiquement des cycles de 2 à 3 paragraphes avec une révision humaine systématique après chaque génération. De plus, l'écrivain doit fournir un feedback explicite et régulier pour guider les suggestions de l'IA et affiner progressivement le comportement du modèle selon ses préférences créatives.

Le deuxième protocole porte sur l'apprentissage adaptatif mutuel entre l'humain et l'IA. Cela commence par la constitution d'une base de style

personnelle : l'IA analyse typiquement 5 à 10 textes représentatifs de l'auteur pour apprendre son style unique, ses tournures caractéristiques et ses préférences thématiques. Ensuite, une évolution progressive s'opère : les suggestions du système s'affinent continuellement selon les acceptations et les rejets de l'auteur, créant ainsi un processus d'apprentissage par renforcement. Parallèlement, l'écrivain développe progressivement ses compétences dans la formulation de prompts de plus en plus précis et efficaces, apprenant à communiquer avec le système de manière optimale.

Le troisième protocole concerne les principes fondamentaux de confiance et de transparence qui doivent sous-tendre toute collaboration humain-IA. Cela nécessite d'abord une documentation explicite des limites du système : identifier clairement ce que l'IA ne peut pas faire est aussi important que de comprendre ses capacités. Ensuite, la traçabilité s'impose : il faut marquer clairement quels segments de texte proviennent de l'IA versus de l'auteur humain, permettant ainsi une transparence totale du processus créatif. Enfin, une éthique de divulgation doit être établie : adopter une politique claire et assumée sur la mention de l'assistance IA dans les œuvres publiées, respectant ainsi le droit du public à l'information.

8.1.2. Principes éthiques fondamentaux

Au-delà de ces protocoles opérationnels, plusieurs principes éthiques fondamentaux

doivent guider l'intégration de l'IA dans le domaine littéraire. Premièrement, il faut reconnaître et valoriser la créativité humaine comme élément irremplaçable et fondateur de la littérature. L'IA, aussi sophistiquée soit-elle, doit rester un outil puissant au service de la vision créative humaine, et non l'inverse. Cette hiérarchie doit être maintenue pour préserver l'authenticité et la profondeur de l'expression littéraire.

Deuxièmement, il est impératif de mettre en place des garde-fous éthiques et juridiques robustes et cohérents. Cela inclut plusieurs dimensions complémentaires. D'abord, l'établissement d'un régime juridique unifié pour la propriété intellectuelle des œuvres impliquant l'IA, harmonisant les approches actuellement disparates entre juridictions. Ensuite, l'audit régulier et systématique des biais dans les systèmes d'IA utilisés pour l'analyse littéraire, en s'appuyant sur des frameworks standardisés comme ceux proposés par Frontiers in Education (2025). Par ailleurs, la préservation active de la diversité culturelle par l'intégration volontaire et méthodique de littératures minoritaires dans les corpus d'entraînement. Enfin, l'encadrement strict de l'usage pédagogique de l'IA pour éviter l'affaiblissement de la pensée critique chez les apprenants.

Troisièmement, les chercheurs, les créateurs et les critiques doivent rester constamment vigilants face à la tendance

inhérente de la technologie vers l'automatisation et la standardisation. Il faut continuer à célébrer et à valoriser activement les formes littéraires singulières, expérimentales et minoritaires, même et surtout lorsqu'elles ne correspondent pas aux patterns que l'IA peut facilement générer ou recommander. Cette vigilance culturelle constitue un rempart essentiel contre l'homogénéisation.

Enfin, la recherche d'un équilibre durable entre innovation technologique et préservation de la richesse littéraire nécessite une adaptation continue et concertée des cadres méthodologiques, éthiques et juridiques. Un dialogue continu et constructif entre technologues, créateurs, critiques littéraires et législateurs s'avère indispensable pour construire un avenir où l'IA enrichit véritablement la littérature sans en appauvrir la diversité et la profondeur humaine. Ce dialogue doit demeurer ouvert, inclusif et attentif aux évolutions rapides tant technologiques que sociales.

9. Limites et perspectives

Cette recherche, bien qu'elle offre un panorama approfondi de l'impact de l'IA sur la littérature, présente certaines limites qu'il convient de reconnaître explicitement :

Premièrement, l'évolution extrêmement rapide des technologies d'IA générative constitue une limite temporelle majeure. Les observations formulées dans cette

étude reflètent l'état des systèmes d'IA au moment de la rédaction, alors que de nouveaux modèles aux capacités substantiellement accrues émergent régulièrement. Cette obsolescence potentielle nécessite une mise à jour continue des analyses et des cadres théoriques proposés.

Deuxièmement, notre analyse s'appuie principalement sur une revue systématique de la littérature scientifique existante, suivant le protocole PRISMA (Page et *al.*, 2021), qui a permis d'identifier 47 articles empiriques pertinents. Toutefois, cette approche documentaire, bien que rigoureuse, ne capture pas l'ensemble des pratiques émergentes et des usages informels de l'IA dans les communautés littéraires et éducatives. Des études empiriques directes auprès d'écrivains, d'étudiants et d'enseignants utilisant activement ces technologies permettraient d'enrichir considérablement notre compréhension des dynamiques réelles d'appropriation et de résistance.

Troisièmement, le caractère interdisciplinaire de notre objet d'étude, qui mobilise simultanément les sciences de l'éducation, les études littéraires, l'informatique et le droit, rend difficile l'atteinte d'une profondeur analytique équivalente dans chacun de ces domaines. Certains aspects juridiques et techniques mériteraient des investigations spécialisées plus approfondies, notamment concernant les mécanismes d'apprentissage

automatique et les jurisprudences émergentes en matière de propriété intellectuelle.

Quatrièmement, notre analyse se concentre principalement sur le contexte académique et littéraire occidental. Les dynamiques d'adoption et les enjeux culturels liés à l'IA générative peuvent différer substantiellement dans d'autres contextes linguistiques et culturels, limitant ainsi la généralisabilité de certaines de nos conclusions.

Ces limites ouvrent néanmoins des perspectives de recherche prometteuses pour les travaux futurs. Plusieurs pistes apparaissent particulièrement fécondes.

D'une part, des études longitudinales permettraient d'observer l'évolution des pratiques créatives et pédagogiques sur le long terme, au-delà des réactions initiales d'enthousiasme ou de résistance. Comment les usages se stabilisent-ils ou se transforment-ils à mesure que la technologie devient plus familière ? Quels nouveaux genres littéraires ou pratiques d'écriture émergent de cette familiarisation progressive ?

D'autre part, des recherches empiriques qualitatives approfondies, utilisant des méthodologies ethnographiques ou d'analyse de l'activité, permettraient de documenter finement les processus réels de collaboration homme-IA dans des contextes créatifs authentiques. L'observation des stratégies d'utilisation

des prompts, des processus itératifs de révision, et des négociations entre intention humaine et suggestion algorithmique constituerait un apport majeur à la compréhension des nouvelles formes d'agentivité créative.

Sur le plan théorique, l'élaboration de cadres conceptuels intégrant les apports de la philosophie de la technique, de l'esthétique et de la théorie littéraire permettrait de mieux penser les transformations ontologiques de l'œuvre littéraire et du statut d'auteur à l'ère de l'IA. La notion d'intentionnalité créative, héritée du romantisme, nécessite une refondation conceptuelle pour s'adapter aux configurations hybrides actuelles.

Enfin, les dimensions éthiques et juridiques appellent des recherches interdisciplinaires urgentes. L'élaboration de nouveaux cadres réglementaires équilibrant protection de la propriété intellectuelle, accès démocratique aux technologies et préservation de la diversité créative constitue un chantier juridique et politique majeur. Les questions de transparence algorithmique, de traçabilité des sources et d'attribution équitable des droits économiques demeurent largement ouvertes et requièrent l'implication conjointe de juristes, d'informaticiens, de créateurs et de décideurs politiques.

Ces perspectives de recherche soulignent que l'étude de l'impact de l'IA sur la littérature ne constitue pas un champ clos,

mais un domaine en pleine construction, appelant des approches méthodologiques diversifiées et des collaborations disciplinaires renouvelées.

10. Conclusion

L'IA bouleverse profondément et durablement la littérature, transformant simultanément la création, l'analyse critique et les pratiques pédagogiques, tout en posant des défis éthiques et méthodologiques majeurs dont nous commençons seulement à saisir l'ampleur. L'hypothèse centrale d'une synergie équilibrée entre humains et IA offre une piste conceptuelle prometteuse pour penser ces transformations, dans un cadre respectueux de la diversité culturelle et de l'intégrité créative qui définissent la littérature depuis ses origines.

Cependant, cette synergie ne peut véritablement émerger que si plusieurs conditions essentielles sont réunies, conditions désormais étayées par des données empiriques robustes et des protocoles opérationnels validés par la recherche internationale. Premièrement, le maintien absolu de la singularité créative humaine comme fondement inaliénable de la littérature doit s'accompagner d'une documentation systématique des processus de collaboration, garantissant transparence et traçabilité. Deuxièmement, l'encadrement strict des biais algorithmiques pour préserver la diversité culturelle nécessite l'audit régulier et rigoureux des systèmes ainsi que la diversification délibérée des corpus d'entraînement. Troisièmement, l'adaptation des cadres juridiques pour

clarifier définitivement la propriété intellectuelle doit progresser vers une harmonisation internationale progressive, dépassant les disparités actuelles. Quatrièmement, la formation critique et approfondie des utilisateurs (étudiants, chercheurs et créateurs) aux capacités effectives et aux limites intrinsèques de l'IA s'impose, développant ce que nous avons appelé la « prompt literacy » et une pensée critique robuste. Enfin, une vigilance éthique constante et collective demeure indispensable pour éviter la standardisation et l'uniformisation culturelle qui menacent la richesse du patrimoine littéraire mondial.

Les données empiriques récentes confirment de manière convergente que l'IA peut effectivement enrichir la pratique littéraire lorsqu'elle est utilisée comme outil de co-crédation plutôt que comme substitut à la créativité humaine. Les expériences menées à Aix-Marseille, Stanford, et dans d'autres institutions de premier plan à travers le monde démontrent que l'efficacité de cette collaboration dépend crucialement du maintien de l'agentivité humaine et de l'adoption de protocoles structurés et réfléchis. En définitive, l'avenir de la littérature à l'ère de l'IA dépendra de notre capacité collective à naviguer entre innovation technologique et préservation des valeurs humanistes qui ont toujours animé la création littéraire. Si nous parvenons à établir un cadre équilibré (technologique, éthique, pédagogique et juridique) alors l'IA pourra véritablement devenir un partenaire créatif enrichissant la littérature plutôt qu'un concurrent menaçant son essence profonde. Ce défi majeur nécessite une vigilance critique constante, une adaptation continue des cadres méthodologiques et législatifs,

ancrés dans la recherche empirique rigoureuse et le dialogue interdisciplinaire ouvert. L'enjeu dépasse largement la seule sphère littéraire : il touche à notre conception même de la créativité, de l'authenticité et du rôle de l'humain dans un monde de plus en plus augmenté par la technologie.

11. Liste Bibliographique

- Alqasham, F. (2023). ChatGPT in the Saudi EFL classroom: A study of learner usage patterns and possibilities in learning optimization. *Migration Letters*, 20(S7), 1251-1263. <https://doi.org/10.59670/ml.v20iS7.4828>
- Bolaños, F., Salatino, A., Osborne, F., & Motta, E. (2024). Artificial Intelligence for Literature Reviews: Opportunities and Challenges. *Artificial Intelligence Review*, 57, article 259. <https://doi.org/10.1007/s10462-024-10902-3>
- Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics Derived Automatically from Language Corpora Contain Human-like Biases. *Science*, 356(6334), 183-186. <https://doi.org/10.1126/science.aal4230>
- Dhillon, P. S., Molaei, S., Li, J., Golub, M., Zheng, S., & Robert, L. P. (2024). Shaping Human-AI Collaboration: Varied Scaffolding Levels in Co-writing with Language Models. *CHI '24: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 1044, 1-18.
- Drobnjak, I., Shawe-Taylor, J., & Perez Ortiz, M. (2024, 12 April). *Large Language Models Generate Biased Content, Warn Researchers*. UNESCO/UCL Report.
- European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council on Artificial Intelligence (AI Act).

Official Journal of the European Union, July 12.

- Ferrara, E. (2023). Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies. *Sci*, 6(1), 3. <https://doi.org/10.3390/sci6010003>
- Gallegos, I. O., Rossi, R. A., Barrow, J., Tanjim, M. M., Kim, S., Dernoncourt, F., Yu, T., Zhang, R., & Ahmed, N. K. (2024). Bias and Fairness in Large Language Models: A Survey. *Computational Linguistics*, 50(3), 1097-1179. https://doi.org/10.1162/coli_a_00524
- Gautier, P.-Y. (1994). Les critères qualificatifs pour la protection littéraire et artistique en droit français. *Revue internationale de droit comparé*, 46(2), 507-519.
- Hwang, Y., & Lee, J. H. (2024). Exploring Students' Experiences and Perceptions of Human-AI Collaboration in Digital Content Making. *International Journal of Educational Technology in Higher Education*, 22(44), 1-21.
- IA-ÉDU (2025). *Actes de l'atelier Intelligence Artificielle générative et ÉDUcation : Enjeux, Défis et Perspectives de Recherche*. Marseille. https://ia-edu.lis-lab.fr/iaedu/doc/Actes_IA_EDU_2025.pdf (Consulté le 13/12/2025).
- Jo, H. (2023). *Factors Contributing to the Adoption of AI Models in Higher Education: Developing Hybrid Intelligence Systems*. Educational Technology Research and Development.
- Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). *Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task*. Preprint, arXiv:2506.08872.
- Laurito, W., Davis, B., Grietzer, P., Gavenčiak, T., Böhm, A., & Kulveit, J. (2025). AI-AI bias: Large language models favor communications generated by large language models. *PNAS*, 122(31), e2415697122. <https://doi.org/10.1073/pnas.2415697122>
- Lee, M., Liang, P., & Yang, Q. (2022). CoAuthor: Designing a Human-AI Collaborative Writing Dataset for Exploring Language Model Capabilities. *CHI Conference on Human Factors in Computing Systems*. Stanford University.
- Mahier, T., GPT J., & Raphael, B. (2024, 12 mai). *Comment j'ai écrit un roman en 7 jours avec l'IA*. Génération IA. <https://generationia.flint.media/p/comment-ecrire-un-roman-en-7-jours-avec-chat-gpt-ia> (Consulté le 11/12/2025).
- Martens, B. (2025, 10 September). *The European Union is Still Caught in an AI Copyright Bind*. Bruegel Analysis. <https://www.bruegel.org/analysis/european-union-still-caught-ai-copyright-bind> (Consulté le 20/12/2025).
- McGuire, J., De Cremer, D., & Van de Cruys, T. (2024). Establishing the Importance of Co-creation and Self-efficacy in Creative Collaboration with Artificial Intelligence. *Scientific Reports*, 14, 18525. <https://doi.org/10.1038/s41598-024-69423-2>
- Page, M. J. et al. (2021). The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *BMJ*, 372(71), 1-9. <https://doi.org/10.1136/bmj.n71>
- Parlement européen et Conseil de l'Union européenne. (2019). *Directive (UE) 2019/790 du Parlement européen et du Conseil du 17 avril 2019 relative à l'auteur et aux droits voisins dans le marché unique numérique et*

modifiant les directives 2001/29/CE et 2006/116/CE (JO L 130). <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX:32019L0790> (Consulté le 11/12/2025).

- Pérot, P., & Bouzayen, I. (2023, 8 décembre). *Décision historique d'un tribunal chinois : des images générées par IA protégées par le droit d'auteur*. AD Article. <https://www.august-debouzy.com/fr/blog/2032-decision-historique-dun-tribunal-chinois-des-images-generrees-par-ia-protegees-par-le-droit-dauteur> (Consulté le 10/12/2025).
- Royaume-Uni (1988). *Copyright, Designs and Patents Act 1988* (c. 48). <https://www.legislation.gov.uk/ukpga/1988/48/contents> (Consulté le 24/12/2025).
- Schlegel, F. (1801). *Athenaeum: Eine Zeitschrift*. Friedrich Schlegel & August Wilhelm Schlegel.
- US Congress. (2024, 9 April). *Generative AI Copyright Disclosure Act of 2024*. <https://www.congress.gov/bill/118th-congress/house-bill/7913> (Consulté le 20/12/2025).
- US Copyright Office (2024). Copyright and Artificial Intelligence, Part 1: Digital Replicas. A Rapport of the Register of Copyright. <https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-1-Digital-Replicas-Report.pdf> (Consulté le 07/12/2025).
- US Copyright Office (2025). Copyright and Artificial Intelligence, Part 2: Copyrightability. A Rapport of the Register of Copyright. <https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-2-Copyrightability-Report.pdf> (Consulté le 07/12/2025).
- WIPO. (2024). *Artificial Intelligence and Intellectual Property: 10th Session of the WIPO Conversation*. World Intellectual Property Organization.
- Wolf, T. et al. (2020). *Transformers: State-of-the-art natural language processing*. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations (pp. 38-45). <https://doi.org/10.18653/v1/2020.emnlp-demos.6> (Consulté le 09/12/2025).
- Yadav, D. (2024). The Role of Artificial Intelligence in Literary Analysis: A Comparative Study. *International Journal of Education and Knowledge Studies*, 3(9), 558-565.
- Zewe, A. (2025). *Unpacking the Bias of Large Language Models*. MIT News. <https://news.mit.edu/2025/unpacking-large-language-model-bias-0617> (Consulté le 29/12/2025).