

RÉINVENTER L'INTÉGRITÉ ACADÉMIQUE À L'ÈRE DE L'INTELLIGENCE ARTIFICIELLE

Sous la direction de
Michelle Bergadaà et Paulo Peixoto



Collection « [Questions de société](#) » dirigée par Olivier MEIER

Réinventer l'intégrité académique à l'ère de l'intelligence artificielle

Sous la direction de Michelle BERGADAÀ et Paulo PEIXOTO



136 boulevard du Maréchal Leclerc
14000 CAEN

© 2025. EMS Editions

Tous droits réservés.

www.editions-ems.fr

ISBN : 978-2-38630-177-3

(Versions numériques)

REMERCIEMENTS

Un premier témoignage de reconnaissance pour Béatrice Durand, membre du Comité éditorial de l'IRAFPA et relectrice du présent volume. Selon la culture IRAFPA, elle a déployé un dialogue bienveillant avec les auteurs en ne sacrifiant rien de l'exigence du mot au service de l'idée.

Merci à tous les participants au 3^e Colloque International de Recherche et d'Action sur l'Intégrité Académique (20-22 juin 2024) pour leurs contributions, leurs questions et leurs discussions qui ont grandement contribué à l'amélioration des chapitres publiés dans cet ouvrage.

Nous remercions tout particulièrement ceux qui ont participé à nos trois ateliers et à la table ronde du Colloque. Après maints débats passionnés, ils sont à l'origine des trois communiqués présentés dans cet ouvrage : Isabelle Baraille, Boubacar Camara, Sandrine Charles, Jacques Athanase Gilbert, France Houle, Eléna Ishow, Ludovic Jeanne, Nicholas Jobidon, Dominique Leglu, André Le Roux, Tom Lindemann, Alexandra Mendoza Caminade, Hervé Maisonneuve, Christina Mitsopoulou, Marian Popescu, Jean-Yves Puyo, Jacques Py, Johanne Saison, Jean-Baptiste Soufron, Hervé Vezin, Jean Winand, Alexandre Zollinger.

Un très grand merci à Henri Prade et Vincent Gaillaud qui nous ont offert une de leurs œuvres qui fait la couverture de ce volume. Un grand merci aussi à Émilie Prade, graphiste et designer talentueuse, qui a réalisé cette couverture.

Nous sommes également reconnaissants envers notre éditeur, Gaël Letranchant, qui a accepté de publier pour la troisième fois un de nos ouvrages collectifs dans la collection « Questions de société » des Éditions EMS.

SOMMAIRE

REMERCIEMENTS

INTRODUCTION

PARTIE 1. FORMATION, INFORMATION ET INTELLIGENCE ARTIFICIELLE

Communiqué 1. Des valeurs aux normes académiques

Chapitre 1. Concevoir l'intégrité académique à l'ère des systèmes d'IA générative : une approche expérimentale

1. Introduction
2. Exploration empirique des capacités de l'IA générative dans un contexte de recherche académique
3. Méthodologie
4. Discussion
5. Conclusion

Chapitre 2. Les nouvelles représentations mentales du travail universitaire : le cas des générations Y et Z

1. Introduction
2. La perception de la valeur du travail intellectuel pour les générations Y et Z
3. Le même comportement ne reflète pas forcément les mêmes motivations
4. Ils vivent une époque sensationnelle
5. Sommes-nous des modèles ou des illusions ?
6. Qu'attendons-nous d'eux ?
7. Sur quelle toile de fond tissent-ils leur avenir ?
8. Discussion

Chapitre 3. IA générative et intégrité académique : approches pédagogiques pour former à l'éthique de l'IA en enseignement supérieur

1. Introduction
2. Les fondements de notre proposition
3. Activités pédagogiques
4. Bilan et discussion

Chapitre 4. L'utilisation de l'IA générative pour la rédaction académique en français par les étudiants allophones.

L'expérience du CPU Lyon

- [1. Introduction](#)
- [2. Revue de littérature sur l'IA par l'IA](#)
- [3. Résultats de l'enquête exploratoire](#)
- [4. Discussion finale et voies de recherche](#)
- [5. Conclusion](#)

PARTIE 2. PUBLICATIONS ET INTELLIGENCE ARTIFICIELLE

Communiqué 2. Qualité de signataire de publications scientifiques

Chapitre 5. L'écosystème de la surpublication : acteurs, *modus operandi* et impact sur le monde universitaire

- [1. Introduction](#)
- [2. Cadre théorique](#)
- [3. Les acteurs : Le Bon, La Brute et Le Truand](#)
- [4. Motivations de la surpublication : mais pourquoi ?](#)
- [5. *Modus operandi* : comment s'y prennent-ils ?](#)
- [6. Quelles sont les conséquences de ce phénomène ?](#)
- [7. Conclusions et prospective](#)

Chapitre 6. Récit, littérature et pratique littéraire. Délices du métier d'éditeur et les vertiges de l'IA

- [1. Introduction](#)
- [2. Engel, la vérité dans la littérature et *Le Maître et Marguerite* comme vérité et séduction](#)
- [3. Le vertige de l'IA et le rôle des sciences humaines](#)
- [4. ChatGPT, récit, littérature](#)
- [5. Une étude de cas en exemple](#)
- [6. Conclusion](#)

Chapitre 7. Intégrité et conflit d'intérêts : financement de la recherche publiée en libre accès et dans les revues hybrides ?

Le cas de la chirurgie cardiaque

- [1. Introduction](#)
- [2. Publication en libre accès, un modèle loin d'être parfait](#)

- [3. Présentation du champ d'étude](#)
- [4. Méthodologie](#)
- [5. Les résultats de la recherche](#)
- [6. Discussion](#)
- [7. Conclusion](#)

PARTIE 3. LES ORGANISATIONS ET L'INTELLIGENCE ARTIFICIELLE

Communiqué 3. Les dix conséquences du plagiat et de la fraude académiques

Chapitre 8. Promouvoir les usages de l'intelligence artificielle (IA) dans le respect de l'intégrité académique. Le cas du Business Science Institute

- [1. Introduction](#)
- [2. L'apparition des IA génératives faisant suite à la Covid-19](#)
- [3. IA comme Intelligence artificielle](#)
- [4. Le point de vue de l'intégrité académique](#)
- [5. Réconcilier l'intelligence artificielle et l'intégrité académique](#)
- [6. Pratiques pour les différents acteurs du Business Science Institute](#)
- [7. Discussion](#)

Chapitre 9. De l'IAïsation de la gestion des ressources humaines dans le milieu académique. Conséquences en termes de droit à un recours juridictionnel effectif garant du respect des droits fondamentaux en présence

- [1. Introduction](#)
- [2. Le recours en annulation](#)
- [3. Le recours en responsabilité](#)
- [4. Conclusion](#)

Chapitre 10. La culture éthique dans les organisations de recherche face aux (in)certitudes liées à l'intelligence artificielle

- [1. Introduction](#)
- [2. Utilisations de l'IA dans la recherche et directives](#)
- [3. Les valeurs partagées](#)
- [4. Application à l'Institut Portugais de Recherche Biomédicale](#)
- [5. Discussion](#)

NOTICES BIOGRAPHIQUES DES AUTEURS

INTRODUCTION

[Michelle Bergadaà](#) et [Paulo Peixoto](#)

Nous vivons aujourd'hui une période de transformations technologiques d'une ampleur sans précédent, où l'intelligence artificielle (IA) ne se contente plus d'être une simple innovation parmi d'autres, mais s'impose comme un phénomène qui redessine en profondeur les structures mêmes de notre société. Son essor fulgurant dans presque tous les aspects de notre quotidien constitue une révolution technique et économique, mais aussi sociologique, culturelle et éthique. En bouleversant les façons dont nous interagissons, travaillons, communiquons et apprenons, l'IA détermine la posture des individus vis-à-vis de la connaissance, de l'autorité scientifique et, plus largement, des institutions sociales.

Le monde académique, socle premier de la production et de la diffusion du savoir, est particulièrement touché par cette révolution. L'université, longtemps perçue comme le bastion de la rigueur intellectuelle et de l'éthique de la recherche, est régulièrement confrontée à des défis qui ébranlent ses normes traditionnelles. Mais l'arrivée des systèmes d'IA, et plus spécifiquement des IA dites « génératives » va bien au-delà : en produisant des contenus de manière autonome, en introduisant de nouvelles dynamiques qui reconfigurent les pratiques académiques, elle met à l'épreuve les valeurs sur lesquelles repose la légitimité scientifique.

D'un point de vue sociologique, l'introduction massive de ces technologies dans le champ académique entraîne une redéfinition des rapports de pouvoir, de la production des savoirs et des critères de reconnaissance dans les sphères scientifiques et éducatives. Il est désormais possible pour une machine de produire des articles, des résumés, voire des idées entières avec un degré de sophistication qui, dans bien des cas, dépasse la simple imitation et confine à la création. Cette réalité soulève des questions cruciales : qu'est-ce que l'intégrité académique à l'ère de l'automatisation ? Comment garantir la valeur et l'authenticité des productions

intellectuelles dans un contexte où les frontières entre l'humain et la machine s'amenuisent rapidement ? Pour cerner les propositions de réponses qui commencent à poindre, nous avons très classiquement organisé cet ouvrage en trois parties : la formation et l'information, la recherche et la publication, les bouleversements des organisations académiques.

Chacune des trois parties de cet ouvrage est introduite par un communiqué de l'IRAFPA qui prend un sens performatif essentiel à une époque de mutation et d'instabilité. Ce besoin de directives se généralise, car la montée en puissance des systèmes d'intelligence artificielle générative, capables de produire des contenus textuels, visuels et sonores en quelques instants, bouleverse profondément le paysage académique et scientifique. Les communiqués de l'IRAFPA ont pour rôle d'alerter la communauté académique sur des phénomènes ou des tendances mettant en question les fondements et le bon fonctionnement de l'intégrité académique. Tout au long de l'année, des ateliers thématiques – en ligne ou en présentiel – sont animés par des membres de l'IRAFPA. Ils réunissent des personnes qui se mobilisent pour envisager des réponses à la question traitée. Les propositions de leurs débats sont validées lors de l'un de nos colloques avant d'être publiées sur le site web de l'IRAFPA¹. Ces communiqués constituent les bases essentielles d'un dialogue pragmatique entre membres de la communauté académique. Ils sont devenus indispensables, alors même que les organisations faïtières ne dissertent que sur les valeurs (honnêteté, transparence, morale...), bien trop abstraites dans une époque en pleine mutation.

Le [premier de nos communiqués](#) pointe le fait que les interactions au sein du monde académique s'enracinent dans les valeurs que nous partageons tous, quels que soient les disciplines, les pays ou les époques. Mais pour agir au quotidien, ce sont de normes explicites dont nous avons besoin, car elles sont la traduction concrète de nos valeurs. Les normes s'accompagnent de dispositifs de régulation (de règles explicites) et de sanctions. L'IRAFPA distingue la nature de la transgression des normes académiques

selon trois dimensions : l'éthique formelle, la déontologie et la responsabilité individuelle et collective.

Avec [notre deuxième communiqué](#), nous éclairons le dilemme que pose cette évolution technologique. Il est essentiel de préciser ce qui définit le statut d'auteur scientifique avant de nous interroger sur l'impact de l'IA. Dans le cadre académique traditionnel, la notion d'auteur est centrale : être auteur confère non seulement les droits qui découlent de la propriété intellectuelle (sur la forme littéraire ou artistique originale, le cas échéant), mais aussi la responsabilité éthique et scientifique des idées et des découvertes. Il est urgent de rappeler selon quels critères on peut être admis à signer des publications scientifiques. En effet, les systèmes d'IA capables de générer des contenus originaux et pertinents risquent de brouiller la notion d'auctorialité. Si une IA rédige un article ou une analyse complexe à partir de données préexistantes, qui doit en être considéré comme l'auteur légitime ? Il faut une parfaite transparence entre les coauteurs pour décider qui signe la publication et dans quel ordre et qui apparaîtra dans les remerciements. Ceci est d'autant plus essentiel qu'en cas de rétractation de la publication (pour fraude ou plagiat), tous les auteurs – les signataires – seront considérés comme coresponsables.

Notre [troisième communiqué](#) s'inscrit dans la mission même de l'IRAFPA. Nous sommes tous conscients que le plagiat et la fraude académiques ne peuvent plus s'appréhender selon une logique de la cause qui consisterait à établir la gravité des conduites illicites sur des faits avérés mais passés. Si la fraude et le plagiat représentent un risque majeur pour la société de la connaissance, alors c'est une logique de la conséquence qu'il faut adopter. Car, une fois que les responsabilités ont été établies, il faut être en mesure de mettre en place des dispositifs individuels et collectifs pour circonscrire les risques à venir. Dans toutes les médiations réalisées à l'IRAFPA, nous proposons d'évaluer la gravité d'une inconduite académique au regard de ses conséquences individuelles et collectives. Ce communiqué nous invite à envisager les dix conséquences potentielles du plagiat et de la fraude académiques. Ces

conséquences font apparaître que l'objet des sciences de l'intégrité est bien un « fait social total » (au sens de Mauss).

La [première partie](#) de cet ouvrage s'intitule « Formation, information et intelligence artificielle ».

Les outils d'IA, autrefois perçus comme de simples assistants technologiques, sont aujourd'hui suffisamment sophistiqués pour rédiger des articles de recherche, synthétiser des revues de littérature ou même proposer des solutions à des problématiques complexes, avec une intervention humaine minimale ou même absente. Cela soulève des questions cruciales et sans précédent pour l'intégrité académique, pilier fondamental de la recherche et de la production de savoirs. Historiquement, l'intégrité académique repose sur ces principes immuables que sont la rigueur méthodologique, l'honnêteté intellectuelle et la transparence dans le processus de recherche. Ces valeurs guident les chercheurs dans les processus rigoureux de collecte, d'analyse et de communication des données et assurent la crédibilité et la validité des résultats scientifiques. Cependant, l'émergence des systèmes d'IA générative, comme ChatGPT ou les grands modèles de langage (LLM), remet en question ces fondements mêmes. Nous devons nous demander ce qui se passe lorsque la technologie devient non seulement un outil, mais également un agent actif dans la production des connaissances ? quand elle est capable de surpasser certaines capacités humaines, tout en automatisant des tâches intellectuelles autrefois réservées aux chercheurs ?

[Morgan Blangeois](#) a exploré les requêtes et les usages de l'IA générative dans trois activités du chercheur : la recherche bibliographique, le traitement de données, la rédaction de textes. Repenser l'intégrité académique à l'ère de l'IA exige de comprendre comment elle fonctionne. Or, les IA génératives, comme ChatGPT ou d'autres modèles de langage avancés, permettent d'automatiser des tâches qui, autrefois, nécessitaient un travail intellectuel considérable. Par exemple, la génération d'idées, la rédaction d'essais ou la synthèse d'informations, qui étaient traditionnellement le fruit d'un processus laborieux de réflexion, de discussion et de

recherche, peuvent désormais être partiellement confiées à ces systèmes. Si cette automatisation peut sembler prometteuse et efficace, elle pose toutefois des problèmes éthiques majeurs, car elle risque de diluer la responsabilité individuelle dans le processus de production de connaissances. Les résultats des simulations réalisées par l'auteur de ce premier chapitre indiquent non seulement les zones grises dans les textes générés (pertinence, hallucinations, etc.), mais aussi les risques de commettre des manquements à l'intégrité. Aussi est-il nécessaire d'informer/de former chercheurs, étudiants, encadrants, éditeurs et... fraudeurs potentiels. Ceci afin de redessiner les extensions comme les limites des critères d'originalité, de créativité, d'éthique et de déontologie académique menant à une connaissance scientifiquement valide et socialement intègre.

Dans le deuxième chapitre, [Marie-Frédérique Bacqué et Pedro Urbano](#) analysent la relation étroite de l'IA et des nouvelles générations de chercheurs, en particulier ceux des générations Y et Z. D'un point de vue sociologique, ces générations se distinguent par leur rapport singulier aux technologies de l'information et de la communication. Ils sont les premiers à avoir grandi avec l'accès instantané à des ressources infinies *via* Internet et les réseaux sociaux, ce qui a profondément façonné leur manière d'appréhender le savoir et les processus d'apprentissage. Ils abordent l'IA non seulement comme un outil d'innovation, mais aussi comme une extension naturelle de leur quotidien. Pour ces personnes qui ont aujourd'hui entre 15 et 40 ans, l'IA fait partie intégrante de leur manière de penser, d'apprendre et de produire du contenu. Cette familiarité avec la technologie, si elle est un atout indéniable, peut cependant les exposer à la tentation de tricher ou d'utiliser ces outils de manière non éthique. Le fait que l'IA générative puisse aujourd'hui produire des textes apparemment originaux et cohérents en quelques instants à partir d'instructions simples peut encourager certains à contourner les efforts nécessaires à l'acquisition de compétences et de savoirs approfondis. La culture de l'instantanéité et de l'accès rapide à l'information redéfinit la manière dont ces

jeunes générations conçoivent le travail intellectuel. La tentation d'une utilisation non éthique de ces outils est d'autant plus grande pour les générations Y et Z, que les systèmes d'IA générative sont plus accessibles, intuitifs et performants.

Avec une approche résolument pragmatique, [Frédéric Bruneault et Andréane Sabourin Laflamme](#) expliquent comment leur « trousse à outils pédagogique » propose dix activités qui apprennent réellement aux étudiants à utiliser intelligemment les outils de l'IA. Les auteurs proposent leur méthode d'exercices utilisables dans le premier et le deuxième cycle universitaire, et même avant. Nous sommes ici au cœur même du *learning by doing* si cher à nos collègues nord-américains. En tant que formateurs et éducateurs, il est essentiel que nous comprenions que ces générations sont façonnées par des phénomènes sociaux inédits : comme la culture des réseaux sociaux, l'image de soi, la validation sociale et la rapidité de la reconnaissance jouent un rôle crucial dans la construction de l'identité individuelle et collective. De plus, la prolifération de plateformes d'information, souvent non régulées et de qualité variable, contribue à brouiller la distinction entre information fiable et information douteuse, ce qui renforce la nécessité d'une éducation aux médias et à l'information. Ainsi, l'utilisation de l'IA générative dans le cadre académique ne peut être appréhendée uniquement sous l'angle de la technique. Elle doit également être comprise à travers le prisme de l'évolution des valeurs, des attentes sociales et des comportements des étudiants dans un monde hyperconnecté. Cela implique d'intégrer, dès les premiers cycles universitaires, une éducation à l'éthique numérique qui ne se contente pas de prévenir la triche, mais qui aide les étudiants à comprendre l'importance de la rigueur intellectuelle, de l'authenticité et de l'honnêteté dans la production du savoir.

S'il est impératif de repenser les approches pédagogiques pour ces nouvelles générations, il est tout aussi important de sensibiliser les enseignants et les encadrants aux responsabilités qui accompagnent l'usage de ces technologies. Parce qu'ils accompagnent des étudiants en thèse allophones, [Yves Frédéric](#)

[Livian et Robert Laurini](#) livrent une analyse très actuelle des échanges entre accompagnants et étudiants. L'IA ne doit pas être perçue comme un substitut au travail académique, mais comme un outil complémentaire qui, lorsqu'il est utilisé de manière responsable, peut enrichir le processus de recherche et d'apprentissage. Selon les auteurs, il est possible de mettre en place des dispositifs pédagogiques qui encouragent les étudiants à développer une pensée critique face aux contenus générés par l'IA. En considérant l'IA générative comme étant au cœur du dialogue accompagnant-étudiant, en questionnant les pratiques de rédaction, les étudiants allophones doivent être capables de questionner l'origine des informations fournies par ces systèmes. Ils peuvent évaluer la qualité des réponses générées et comprendre que l'IA, toute sophistiquée qu'elle soit, repose sur des algorithmes et des données qui peuvent être biaisés ou incomplets. L'objectif est de former de futurs chercheurs à une utilisation éthique et responsable de l'IA, tout en conservant un sens aigu de leur réflexion critique et de leur contribution intellectuelle personnelle. Dans cette démarche qui appelle une grande maturité, les étudiants ne peuvent ni ne doivent pas être laissés à eux-mêmes. Une conclusion s'impose : le renforcement du lien social entre le pédagogue et l'étudiant, mais plus encore le retour en force d'une culture de l'oralité dans le processus de formation qui devra contrebalancer des décennies de simple évaluation « scolaire » de l'écrit.

La [deuxième partie](#) de cet ouvrage s'intitule « Publications et intelligence artificielle ».

La prolifération des publications académiques est un phénomène qui a connu une explosion exponentielle ces dernières décennies. Le nombre d'articles, de revues et de résultats de recherche n'a cessé d'augmenter depuis vingt ans, en grande partie grâce à la numérisation et à la facilitation des processus de soumission et de diffusion. L'IA générative est capable de produire des contenus de haute qualité en quelques instants, ce qui allège le travail des chercheurs en leur offrant des outils puissants pour analyser des données ou rédiger des articles. Mais elle risque également

d'entraîner des dérives – des fraudes scientifiques et de plagiat, dont la surpublication ne sera que la face visible de l'iceberg. Cette « fièvre de la publication académique » reflète une obsession de la productivité, l'accent est mis davantage sur la quantité des publications que sur leur impact ou leur contribution réelle au savoir. Deux visions de notre métier s'opposent, tant au niveau des chercheurs individuels qu'à celui des responsables d'établissements universitaires et des éditeurs : certains procèdent encore de manière traditionnelle et considèrent que la quête de la vérité est ce qu'il y a de plus sacré en sciences ; d'autres sont tentés d'accélérer la production de publications sans nécessairement respecter les standards rigoureux de qualité qui devraient caractériser le processus de recherche.

[Cinta Gallent Torres et Rubén Comas-Forgas](#) analysent l'écosystème de la surpublication et son impact sur le monde universitaire. D'un point de vue sociologique, la surpublication peut être analysée comme une conséquence directe de l'évolution des normes institutionnelles et professionnelles. Au cours des dernières décennies, les systèmes d'évaluation des chercheurs se sont de plus en plus centrés sur des indicateurs quantitatifs tels que le nombre de publications ou le facteur d'impact des revues. Ces indicateurs, censés mesurer la productivité et l'influence scientifique, sont devenus des outils incontournables pour déterminer les promotions, les financements de recherche ou encore la reconnaissance académique. Cette course à l'échalote a des effets pervers : non seulement elle encourage la multiplication des articles qui ne contiennent aucune véritable innovation, mais elle crée aussi un environnement de compétition exacerbée, dans lequel la publication devient une fin en soi plutôt qu'un moyen de partager des découvertes scientifiques importantes. Le risque de surpublication lié à l'utilisation de l'IA générative n'est pas seulement un problème technique, c'est l'aboutissement d'une crise plus profonde au sein du système académique. Les auteurs de ce chapitre montrent comment les dynamiques actuelles de production de savoir sont davantage régies par des impératifs de rentabilité et de visibilité : la recherche

est souvent perçue comme un produit à vendre sur un marché compétitif. Dans ce cadre, l'IA devient une aubaine pour ces marchands de publications. L'automatisation de la recherche encourage la multiplication d'articles redondants et superficiels, ce qui dilue la valeur réelle du savoir produit. L'optimisation de la chaîne de production se fait au détriment de la qualité et de l'intégrité scientifique.

Cette évolution risque de compromettre la confiance accordée aux publications académiques. [Ignace Haaz](#) analyse la pratique littéraire depuis son poste d'observation d'éditeur. Traditionnellement, la publication scientifique est le fruit d'un processus rigoureux de recherche, de vérification et de validation par les pairs. Cependant, avec la capacité de l'IA à générer de manière autonome des textes cohérents et détaillés, il devient plus difficile de distinguer les travaux véritablement innovants de ceux qui ne sont que des répliques ou des compilations automatisées de travaux existants. Comment garantir que l'utilisation de l'IA ne conduira pas à une érosion de la rigueur linguistique ? Capable de rédiger des articles, de synthétiser des données ou même de proposer des résultats de recherche en quelques instants, l'IA peut devenir un outil hors contrôle. Elle peut être utilisée pour réécrire ou paraphraser des articles déjà publiés, donnant ainsi l'apparence d'une nouvelle recherche, alors qu'il ne s'agit que de la reformulation d'un contenu préexistant. Or, incorporer le savoir-faire pratique d'écrire, repose sur la capacité à former des connaissances et l'intériorisation d'une ambition de vérité. Sans cette dimension cognitive essentielle, ChatGPT peut bien être entraîné à générer un texte ou à le modifier selon des injonctions précises. Mais de telles pratiques peuvent constituer une menace pour l'éthique scientifique, car elles brouillent la ligne entre l'originalité et le plagiat, entre l'innovation et la répétition. Pour l'auteur de ce chapitre, une automatisation de la production scientifique poserait des questions cruciales sur l'authenticité et l'originalité du contenu d'un manuscrit.

Dans ce contexte, il est essentiel de repenser les mécanismes de validation et de régulation. [Delphine Szeceł](#), [Tom Melvin](#) et [Wouter Oosterlinck](#) se sont penchés sur le lien entre deux phénomènes : le

financement de la recherche médicale publiée... et les conflits d'intérêts. L'automatisation de la production scientifique incite à une surpublication dans les disciplines telles que la médecine, où la pression de publier compromet la qualité des recherches. En facilitant la production massive de contenus, l'IA encourage potentiellement des publications malhonnêtes fondées sur du plagiat ou de la fabrication de données. Les auteurs clarifient ce qu'est la « science ouverte » avant de montrer que ces outils présentent des risques considérables, publier dans des revues prédatrices ou dans des revues peu regardantes dans leurs évaluations des articles soumis. En appui de leur démonstration, les auteurs analysent comment les revues médicales gèrent les conflits d'intérêts, situation à laquelle elles sont confrontées de manière récurrente. Les auteurs considèrent que l'intégration de l'IA dans le processus de recherche doit être accompagnée d'une réflexion éthique approfondie pour garantir que ces technologies enrichissent la production scientifique plutôt que de l'appauvrir. Les comités éditoriaux, les institutions de recherche et les agences de financement doivent développer de nouveaux cadres pour évaluer non seulement la quantité, mais surtout la qualité des publications, leur originalité et leur impact réel.

La troisième partie de cet ouvrage s'intitule « Les organisations et l'intelligence artificielle ».

Tant dans leur organisation que dans leur structure, les institutions d'enseignement et de recherche doivent s'adapter aux défis posés par l'IA. Et ce, qu'il s'agisse de gestion des ressources humaines, de promotion d'une culture éthique ou de relations avec la société. Au-delà des questions d'intégrité individuelle, la transformation en cours est le reflet des mutations globales de la société qui se répercutent dans l'organisation du travail académique. L'automatisation croissante facilitée par l'IA exige de repenser les formes de collaboration, le rôle des enseignants et des chercheurs ainsi que l'impact de ces innovations sur les carrières universitaires. La gestion des ressources humaines, traditionnellement fondée sur des interactions humaines, devient plus technocratique. La transformation des pratiques managériales au sein des universités

peut avoir des conséquences sur la motivation et l'engagement des chercheurs, ainsi que sur leur perception de la justice organisationnelle.

[Jean Moscarola et Michel Kalika](#) exposent comment un institut international d'études doctorales a dû très rapidement intégrer les usages de l'intelligence artificielle dans le respect le plus absolu de l'intégrité académique. Parce que le Business Science Institute est une entreprise en réseau, cette adaptation a pu être relativement rapide... du moins pour les doctorants. L'IA générative, avec sa capacité à synthétiser de grandes quantités d'informations en un temps record, semble parfois court-circuiter le lent processus de maturation de la recherche. Comment, dès lors, garantir que les connaissances produites par ces systèmes respectent les critères académiques de rigueur et d'originalité ? La rapidité et l'efficacité de l'IA ne risquent-elles pas de compromettre la qualité des savoirs en privilégiant la quantité ou la surface au détriment de la profondeur et de la réflexion critique ? Ces questions appellent une transformation des pratiques pédagogiques : l'accent doit être mis non seulement sur l'apprentissage technique, mais aussi sur le développement d'une pensée critique capable de naviguer dans un monde de plus en plus automatisé. Car, c'est *in fine* l'acquisition de cette compétence critique que valide l'obtention de ce diplôme d'État qu'est le doctorat.

L'un des enjeux majeurs de l'IA ne se situe pas seulement au niveau de la recherche et de la publication, mais au cœur même de notre système. [Ghislaine Alberton](#) montre que l'impact de l'intelligence artificielle (IA) s'étend également à la gestion des ressources humaines dans les institutions académiques. L'une des particularités du milieu universitaire est la diversité des trajectoires de ses acteurs. Or, l'introduction croissante des systèmes d'IA dans la gestion des carrières – phénomène que l'on pourrait appeler « IAïsation » – transforme profondément ces processus clés que sont le recrutement, la promotion ou l'évaluation des performances des chercheurs et des enseignants. En théorie, les systèmes d'IA sont censés être objectifs et impartiaux. Cependant, les algorithmes

sur lesquels ils reposent sont souvent biaisés par les données avec lesquelles ils ont été entraînés, lesquelles sont le produit de choix humains subjectifs. Dans ce contexte, l'utilisation de l'IA pour des décisions aussi importantes que le recrutement ou la promotion peut introduire des biais systémiques discriminant certains groupes de chercheurs en fonction de critères de genre, d'origine ethnique, ou encore de type d'institution. Le recours à l'IA dans la gestion des carrières académiques soulève donc des questions importantes concernant les droits fondamentaux des chercheurs, notamment en ce qui concerne la transparence des décisions, la non-discrimination et l'équité dans le traitement des candidatures et des promotions.

Dans une approche holistique qui englobe non seulement les chercheurs individuels, mais aussi les institutions et la culture de recherche en général dans la société, [Susana Magalhães](#) fait des propositions pour adapter les directives en matière d'intégrité dans un contexte où l'intelligence artificielle provoque beaucoup d'incertitudes. Le véritable défi auquel sont confrontés les centres de recherche ne réside pas uniquement dans l'adoption de l'IA, mais dans la manière dont cette technologie peut être intégrée de manière éthique. Il faut réformer les systèmes d'évaluation de la recherche de manière à favoriser la qualité et l'originalité plutôt que de se concentrer uniquement sur les métriques quantitatives. Les récits d'expérience et la délibération éthique sont appelés à jouer un rôle crucial pour gérer les incertitudes. Il s'agit enfin d'intégrer l'IA aux pratiques de recherche d'une manière transparente et éthique, en reconnaissant ses potentiels et en évitant ses dérives.

Au fil des chapitres, le livre *Réinventer l'intégrité académique à l'ère de l'intelligence artificielle* jette les bases d'une réflexion approfondie sur la manière dont l'intelligence artificielle, et plus particulièrement l'IA générative bouleverse les paradigmes de la recherche, de l'enseignement et de la gestion des institutions académiques. Cette technologie a fait des avancées spectaculaires en très peu de temps. Elle offre aujourd'hui des opportunités sans précédent pour accélérer la production et la diffusion du savoir : elle permet de traiter des volumes de données massifs, d'automatiser

des tâches complexes et d'optimiser la gestion des ressources humaines. Toutefois, ces succès indéniables s'accompagnent de défis inédits, notamment sur les plans de l'éthique, de l'intégrité académique et de l'organisation institutionnelle.

Cet ouvrage collectif propose d'analyser les implications de cette mutation technologique au prisme de plusieurs questions d'ordre sociologique : l'intégrité académique, les pratiques pédagogiques, la publication scientifique et la gestion des ressources humaines dans les institutions d'enseignement et de recherche. En abordant l'IA comme un phénomène social aux répercussions profondes, nous voulons comprendre comment les normes, les valeurs et les relations de pouvoir dans le monde académique sont en train de se reconfigurer. Les outils d'IA imposent une réflexion sur la cohabitation entre des générations de chercheurs qui, selon leurs socialisations technologiques respectives, n'appréhendent pas de la même manière ces nouvelles capacités de production automatique du savoir. Loin de se limiter à une analyse technologique, cet ouvrage met en lumière les nouvelles configurations sociales, culturelles et politiques que l'IA entraîne au sein des institutions du savoir.

Les auteurs sont issus de multiples disciplines : chirurgie, droit, éducation, gestion, informatique, médecine, philologie, philosophie, psychologie, psychanalyse et sociologie. En explorant ces transformations à partir de leurs différentes perspectives disciplinaires, ils nous invitent à réfléchir aux implications à long terme de l'IA pour le monde académique et pour la société dans son ensemble. Un modèle de production académique « à la chaîne », dans lequel la créativité humaine et l'intuition intellectuelle seraient partiellement confiées à des machines, pourrait bouleverser les fondements mêmes de l'autorité et de la légitimité scientifique. Cela conduit à interroger les structures de reconnaissance académique, souvent basées sur des critères de production quantifiable, et les conséquences d'une telle évolution sur la qualité du savoir produit. En fin de compte, l'IA représente un défi pour les institutions

académiques, non seulement en termes d'adaptation technique, mais aussi dans leur rôle social.

In fine, l'objectif de cet ouvrage est de trouver un équilibre délicat entre l'innovation technologique et la préservation des valeurs académiques. L'IA, si elle est bien encadrée, peut être un atout considérable pour l'enseignement et la recherche. Mais pour cela, elle doit être utilisée avec discernement, dans un cadre de règles pratiques et éthiques claires, qui garantissent que la technologie reste au service de l'humain et de la connaissance, et non l'inverse. Les recommandations concrètes proposées tout au long de ce livre visent ainsi à orienter les universités et les chercheurs dans leur adoption de l'IA, en veillant à ce que cette dernière serve à renforcer l'intégrité académique et non à la mettre en péril.

1. <https://irafpa.org/>

PARTIE 1. FORMATION, INFORMATION ET INTELLIGENCE ARTIFICIELLE

COMMUNIQUÉ 1. DES VALEURS AUX NORMES ACADÉMIQUES

[Michelle Bergadaà](#), Professeure émérite, Université de Genève

Les principes de l'intégrité scientifique sont précisés dans le *Code de conduite européen pour l'intégrité en recherche* ; ils sont donc reconnus dans les différents pays de l'Union européenne. Ces principes guident les individus, les institutions et les organisations dans leur travail. Ils sont au nombre de quatre : la fiabilité de la recherche, l'honnêteté dans le travail, le respect des collègues et collaborateurs et la responsabilité à l'égard des pairs et de la société en général. Hors Union européenne, il existe des dispositifs similaires. Ainsi, les universités du Canada se réfèrent à cinq valeurs fondamentales proposées par l'International Center for Academy Integrity. Elles sont les suivantes : l'honnêteté, la confiance, l'équité, le respect et la responsabilité.

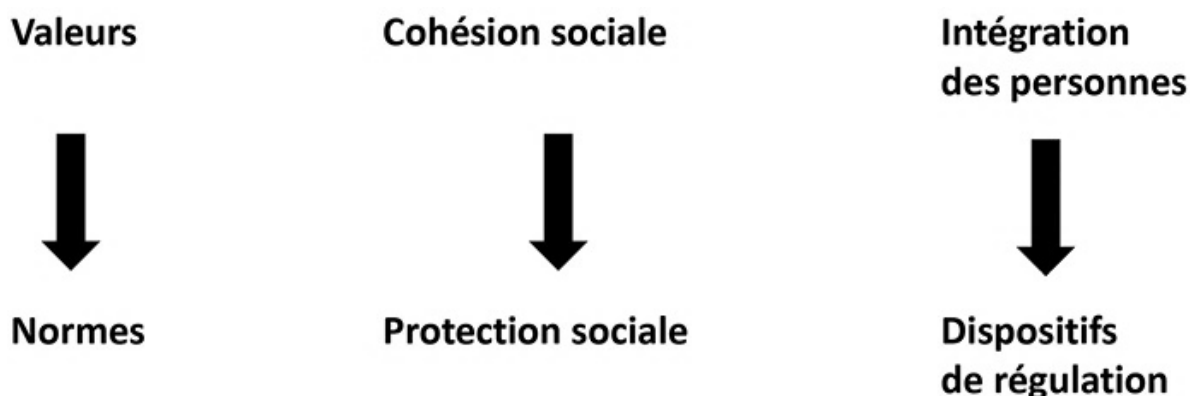
Ces principes sont en réalité des valeurs, non des normes de conduite. Mais, parce que l'intelligence artificielle bouscule nos convictions, l'IRAFPA a renforcé son communiqué concernant la transgression des normes. La zone grise doit en effet être clarifiée si l'on ne souhaite pas voir les inconduites académiques se développer de manière exponentielle. Commençons par ce que sont les valeurs et les normes (Figure 1).

- **Les valeurs** guident, voire conditionnent nos conduites ; elles sont abstraites et communes à tous ceux qui ont choisi d'embrasser une carrière académique. L'énoncé des valeurs génère de la cohésion sociale pour une communauté ou pour une société (la nôtre est le milieu académique). Le but est de pouvoir intégrer des personnes qui vont les partager (par exemple les doctorants). Ainsi des énoncés tels que « la reconnaissance de la dignité de la personne humaine » ou « l'utilisation honnête et rationnelle des fonds publics » font référence à des valeurs.

- **Les normes** sont la traduction concrète de nos valeurs. L'énoncé des normes a pour objet la protection sociale. Les normes s'accompagnent de dispositifs de régulation (de règles explicites) et de sanctions. Il s'agit donc de dispositifs destinés à contrôler les déviations et à les traduire devant les organes d'évaluation et de sanction. L'analogie avec

l'application des lois par la police et le système judiciaire souligne l'importance de ces normes pour maintenir l'ordre et la cohésion sociale (Livet, 2006).

Figure 1. Des valeurs aux normes



Clarifions également ce que nous entendons par « intégrité ». L'intégrité, selon le Centre National de Ressources Textuelles et Lexicales (CNRTL), est *l'état d'une chose, d'un tout, qui est entier, qui a toutes ses parties*. Le concept d'intégrité, selon l'IRAFPA, est donc constitué de quatre dimensions que sont : la morale personnelle (*virtue ethics*), l'éthique normative (lois et règlements régissant notre métier), la déontologie professionnelle et la responsabilité individuelle (Bergadaà, 2023).

Si nous ne réussissons pas à concilier ces quatre éléments de l'intégrité dans notre pratique quotidienne, nous risquons de ne pas satisfaire à la deuxième définition de l'intégrité proposée par le CNRTL : *caractère, qualité d'une personne intègre, incorruptible, dont la conduite et les actes sont irréprochables*.

Ce communiqué de l'IRAFPA émis en juin 2016, a été précisé avec les participants des deux Écoles d'été de juillet 2024 « Implémenter une culture de l'intégrité académique » (11-12 juillet 2024) et « Conseiller en intégrité » (15-17 juillet 2024). Qu'ils soient remerciés de leur implication.

Références bibliographiques

ALLEA (All Europeans Academies) (2018). *Code de conduite européen pour l'intégrité en recherche* (édition révisée). <https://allea.org/code-of-conduct/>

Bergadaà, M. (2023). Le contrepois de la parole performative dans les sciences de l'intégrité. Dans M. Bergadaà (dir.), *Les nouvelles frontières de l'intégrité académique* (pp. 33-48). Éditions EMS.

Centre National de Ressources Textuelles et Lexicales (CNRTL). <https://www.cnrtl.fr/lexicographie/int%C3%A9grit%C3%A9>

International Center for Academic Integrity. <https://academicintegrity.org/>

Livet, P. (2006). *Les normes*. Armand Colin.

Communiqué de l'IRAFPA – Transgression des normes académiques

L'IRAFPA distingue la nature de la transgression des normes académiques selon trois dimensions que sont : l'éthique formelle, la déontologie et la responsabilité individuelle et collective.

En matière de plagiat, sont des atteintes à l'éthique formelle et aux codes de conduite qui en découlent les faits avérés suivants :

- La publication sous son propre nom de résultats de travaux et de découvertes de tiers (plagiat)
- L'omission délibérée des noms de collaborateurs du projet y ayant apporté des contributions essentielles
- L'autoplégat, soit la reprise d'une partie ou de la totalité d'un travail déjà publié en omettant volontairement d'y faire référence
- La publication dans une langue d'un article déjà publié dans une autre langue sans y faire explicitement référence
- L'usage de l'IA sans y faire référence

En matière de fraude scientifique, sont des atteintes à l'éthique formelle et aux codes de conduite qui en découlent les faits avérés suivants :

- L'invention des résultats de recherche
- La falsification intentionnelle de données de base
- La présentation ainsi que le traitement intentionnellement trompeur de résultats de recherche
- L'exclusion de données de base de données, sans la consigner ou sans en donner les raisons
- La dissimulation de données
- Le piratage de données

En matière de plagiat et de fraude scientifique, sont considérés comme des problèmes de déontologie les faits avérés suivants :

- Le fait d'obtenir le statut de coauteur d'une publication sans avoir apporté de contribution essentielle au travail
- Les citations intentionnellement erronées tirées de travaux existants ou supposés de tiers
- La modification de l'ordre des auteurs sans leur accord et sans justification
- La mention volontaire d'une personne en qualité de coauteur alors qu'elle n'a pas contribué au projet

- L'omission délibérée de contributions essentielles d'autres auteurs sur le même sujet (bibliographie incomplète)
- Le fait de passer sciemment sous silence des conflits d'intérêts
- La violation du devoir de discrétion (de l'obligation de réserve)
- La négligence du devoir de dénonciation des fraudes scientifiques observées
- Le refus d'accorder à des tiers dûment autorisés le droit de consulter les données de base
- La présentation non-objective d'opinions divergentes
- Les indications incorrectes sur le stade d'avancement de la publication de ses propres travaux (par exemple « publication en cours d'impression », alors que le manuscrit n'a pas encore été accepté)

En matière de plagiat et de fraude scientifique, sont considérés comme problèmes de responsabilité les faits avérés suivants :

- Le refus de personnes détenant l'autorité, élues ou nommées à des postes hiérarchiques au sein d'établissements, d'écouter les lanceurs d'alerte lorsqu'ils informent d'atteintes à l'intégrité
- Le manque de protection accordé par les personnes détenant l'autorité aux victimes et aux lanceurs d'alerte durant toute la durée de la procédure d'enquête
- Le fait de passer sous silence des conflits d'intérêts en matière de financement de la recherche
- L'absence de réactivité face à une situation de crise organisationnelle due à des faits supposés de manquement à l'intégrité d'une ou de plusieurs personnes

CHAPITRE 1. CONCEVOIR L'INTÉGRITÉ ACADÉMIQUE À L'ÈRE DES SYSTÈMES D'IA GÉNÉRATIVE : UNE APPROCHE EXPÉRIMENTALE

[*Morgan Blangeois*](#), *doctorant à Clermont Recherche Management, Université Clermont
Auvergne (France)*¹

1. INTRODUCTION

L'horizon de la production scientifique se redessine à la vitesse de l'intelligence artificielle (IA). Des articles générés en quelques secondes, des revues de littérature synthétisées automatiquement et des lacunes de recherche identifiées par des algorithmes : ce scénario, autrefois relégué à la science-fiction, devient réalité. Cette mutation technologique remet en question les fondements mêmes de l'intégrité du chercheur, traditionnellement ancrée dans quatre piliers interdépendants : morale, déontologie, éthique formelle et responsabilité (IRAFPA, 2024). La rigueur méthodologique et la transparence, garantes de la validité des résultats scientifiques, se trouvent désormais à la croisée des chemins entre innovation et préservation des valeurs académiques. Dans ce contexte, il devient impératif d'examiner de manière approfondie l'impact potentiel des systèmes d'IA générative sur les principes fondamentaux de l'intégrité académique. Notre étude plonge au cœur de cette révolution pour en explorer les implications profondes sur la recherche et ses pratiques.

Le vocabulaire de l'IA

Avant d'approfondir cette analyse, il convient de clarifier certains termes clés.

- IA générative : paradigme émergent en IA, caractérisé par la capacité des systèmes à produire des contenus originaux à partir de données d'apprentissage (Bommasani *et al.*, 2022).
- Systèmes d'IA générative : ensembles technologiques intégrés, capables de générer du contenu multimodal (texte, image, son, vidéo) présentant un degré

élevé de cohérence et de pertinence contextuelle (Commission de l'intelligence artificielle, 2024).

- Grands modèles de langage (LLM) : architectures neuronales avancées, entraînées sur des corpus de données massifs, conçues pour comprendre et générer du langage naturel avec un niveau de sophistication proche de celui des humains (Vaswani *et al.*, 2017).

L'émergence de ces systèmes soulève de nombreuses questions éthiques et pratiques². Navigli *et al.* (2023) et Barrett *et al.* (2023) soulignent les inquiétudes concernant les pratiques abusives, les menaces pour la vie privée et la sécurité, ainsi que les biais présents dans les modèles d'IA générative. Barrett *et al.* (2023) et Perkins *et al.* (2024) mettent en garde contre la propagation de désinformation due aux « hallucinations » que provoquent ces biais. Dwivedi *et al.* (2023) soulignent l'applicabilité croissante de ces outils dans divers contextes professionnels, de la santé à l'éducation. Face à ces enjeux, diverses instances, telles que le Comité National Pilote d'Éthique du Numérique (2023) et l'Union européenne (2024), se mobilisent pour évaluer l'impact sociétal de ces technologies et formuler des recommandations. À l'heure des LLM, les questions de cocréation et de paternité des œuvres sont au cœur des débats (Bozkurt, 2024 ; Dell'Acqua *et al.*, 2023 ; Stokel-Walker & Van Noorden, 2023). Les LLM sont capables de produire des textes cohérents à partir d'instructions génératives (prompt) sur une variété de sujets, y compris des contenus académiques. Dell'Acqua *et al.* (2023) et Stokel-Walker & Van Noorden (2023) soulignent que les frontières et les possibilités des modèles de langage ne sont pas encore bien établies. Selon Bozkurt (2024) et Dwivedi *et al.* (2023), les LLM remettent en question notre conception traditionnelle de la création intellectuelle dans le domaine scientifique. Cette évolution rend nécessaire une redéfinition des codes déontologiques de la recherche en réponse à la zone grise née de l'apparition de ces nouveaux outils.

Certaines études soutiennent que le contenu créé par des systèmes d'IA peut constituer des imitations (Jeanne, 2023), d'autres

soulignent les dilemmes déontologiques inhérents à l'utilisation d'une telle technologie à des fins académiques (Eke, 2023). Pour anticiper et prévenir efficacement les dérives, il est essentiel d'explorer le point de vue des délinquants de la connaissance et d'analyser les potentialités offertes par l'IA en termes de sophistication et de professionnalisation des pratiques de fraude (Jeanne, 2023). Le concept de « délinquant de la connaissance », issu de l'approche interactionniste, désigne les divers comportements fraudeurs dans l'interaction avec les objets de connaissance et les autres acteurs du monde académique. Les recherches menées à l'IRAFPA ont permis d'identifier un spectre de profils allant de ceux qui commettent de petites négligences quotidiennes à ceux qui s'engagent dans des fraudes scientifiques ou des plagats massifs (IRAFPA, 2024). Cette perspective nous permet d'appréhender les défis posés par l'IA générative à l'intégrité académique.

Notre chapitre adopte une approche expérimentale pour évaluer empiriquement les capacités des LLM dans un contexte académique, en simulant le point de vue d'un acteur cherchant à exploiter ces technologies à des fins potentiellement frauduleuses. Ce chapitre se concentre sur deux propositions principales :

(1) Démontrer expérimentalement la capacité des LLM à produire des contenus académiques substantiels avec une implication humaine minimale et examiner les défis que cela pose aux principes actuels d'intégrité académique.

(2) Analyser les implications de ces capacités sur les critères d'originalité et de création de connaissance conventionnels, ouvrant ainsi une réflexion sur l'évolution nécessaire des politiques éditoriales et des modes de validation des connaissances.

Nous présenterons d'abord une expérimentation explorant les capacités avancées des outils d'IA du point de vue d'un délinquant de la connaissance, et détaillant la méthodologie et en illustrant graphiquement les résultats. Ensuite, nous analyserons les implications de cette démonstration en examinant les critères d'originalité et de création de connaissance à l'ère de l'IA générative.

Enfin, nous proposerons des recommandations pratiques pour l'évolution des politiques éditoriales et l'usage des systèmes d'IA générative dans la recherche académique.

2. EXPLORATION EMPIRIQUE DES CAPACITÉS DE L'IA GÉNÉRATIVE DANS UN CONTEXTE DE RECHERCHE ACADÉMIQUE

L'IA générative est un changement de paradigme dans le domaine de l'IA (Bommasani *et al.*, 2022). Selon la Commission de l'intelligence artificielle (2024), l'IA générative est qualifiée de générative « car elle permet de générer de nouveaux contenus sous la forme de texte, d'image, de son, de vidéo ou de code ». Cette technologie s'appuie sur des techniques d'apprentissage automatique avancées, notamment :

- L'architecture *transformer* : permettant une prise en compte des dépendances à « longue distance » dans les données (Vaswani *et al.*, 2017).
- Les modèles de fondation : des réseaux neuronaux pré-entraînés sur d'immenses corpus de données (Bommasani *et al.*, 2022).

La Commission de l'intelligence artificielle (2024) souligne plusieurs caractéristiques clés des systèmes d'IA générative, notamment leur simplicité d'utilisation, leur rapidité de génération, et le réalisme du contenu produit. Il est crucial de noter que, comme le souligne la Commission, « les modèles d'IA générative n'appréhendent pas la signification des mots, des images ou du son ». La signification perçue provient en réalité de l'interprétation humaine des contenus générés. En d'autres termes, ces modèles fonctionnent sans compréhension intrinsèque du sens ou de la nature fondamentale de ce qu'ils produisent (absence de métaphysique) et sans méthode pour acquérir ou valider une telle compréhension (absence d'épistémologie).

3. MÉTHODOLOGIE

Les caractéristiques que nous venons d'exposer soulèvent des questions cruciales quant à leur utilisation dans le domaine académique. Afin d'explorer ces enjeux de manière empirique, nous avons développé un cadre méthodologique spécifique pour notre étude.

Notre démarche évalue empiriquement l'IA générative dans la recherche académique en adoptant la perspective de l'acteur qui optimise ces technologies pour produire un contenu scientifiquement discutable. Cette approche repose sur trois principes directeurs :

(1) Contextualisation technologique : nous analysons un environnement où l'IA générative est omniprésente et en constante évolution, ce qui permet aux délinquants de la connaissance d'innover continuellement dans leurs méthodes de fraude. Cette réalité rend obsolètes les méthodes de détection proposées par *filigrane invisible*³.

(2) Perspective déontologique : nous explorons l'utilisation maximale de l'IA générative pour produire du contenu académique, afin d'examiner les implications profondes sur l'intégrité scientifique, au-delà de la simple qualité du contenu.

(3) Point de vue du fraudeur : nous cherchons à optimiser la génération de résultats en minimisant l'intervention humaine, pour explorer les limites de ces systèmes et mettre en lumière les défis éthiques et pratiques qu'ils soulèvent.

3.1. Expérimentation

Notre expérimentation utilise des scripts Python, elle s'inspire des méthodologies récentes d'évaluation de l'IA en contextes professionnels et académiques (Dell'Acqua *et al.*, 2023 ; Perkins *et al.*, 2024), adaptées pour explorer les capacités des LLM du point de vue d'un délinquant de la connaissance :

(1) Définition des tâches de recherche à automatiser : analyse de données, synthèse d'articles, identification de lacunes de recherche. L'Union européenne (2024) précise que ce sont spécifiquement ces

usages qui peuvent être problématiques : « Par exemple, utiliser l'IA générative comme un simple outil de support à la rédaction n'est pas un usage substantiel. Cependant, interpréter une analyse de données, effectuer une revue de littérature, identifier des lacunes de recherche, formuler des objectifs de recherche, développer des hypothèses, etc. pourraient avoir un impact substantiel. »

(2) Sélection des LLM les plus adaptés à chaque tâche selon leurs capacités spécifiques.

(3) Élaboration d'instructions génératives détaillées, utilisant la technique « *metaprompt* ».

(4) Itérations et ajustements des modèles.

(5) Intégration d'outils complémentaires pour créer des « agents ».

(6) Validation des agents sur des tâches réelles.

Deux aspects clés de notre approche méritent une attention particulière :

Nous avons employé le « *metaprompt* », une méthode qui utilise un modèle de langage pour générer automatiquement les prompts adaptés à nos besoins (Anthropic, 2024). Cette approche s'est révélée particulièrement pertinente pour notre étude, car elle allie efficacité et rapidité de mise en œuvre – des attributs cruciaux dans le contexte de notre simulation d'un délinquant de la connaissance cherchant à maximiser l'efficacité de ses activités frauduleuses.

Le développement d'agents autonomes est un autre aspect clé de notre approche. Ces agents, contrairement aux LLM isolés, interagissent avec leur environnement en utilisant des outils externes comme des interfaces de programmation pour accéder à des bases de données. Notre expérimentation intègre un interprète Python et des accès aux bases d'articles scientifiques, améliorant ainsi les performances des agents dans les tâches académiques définies.

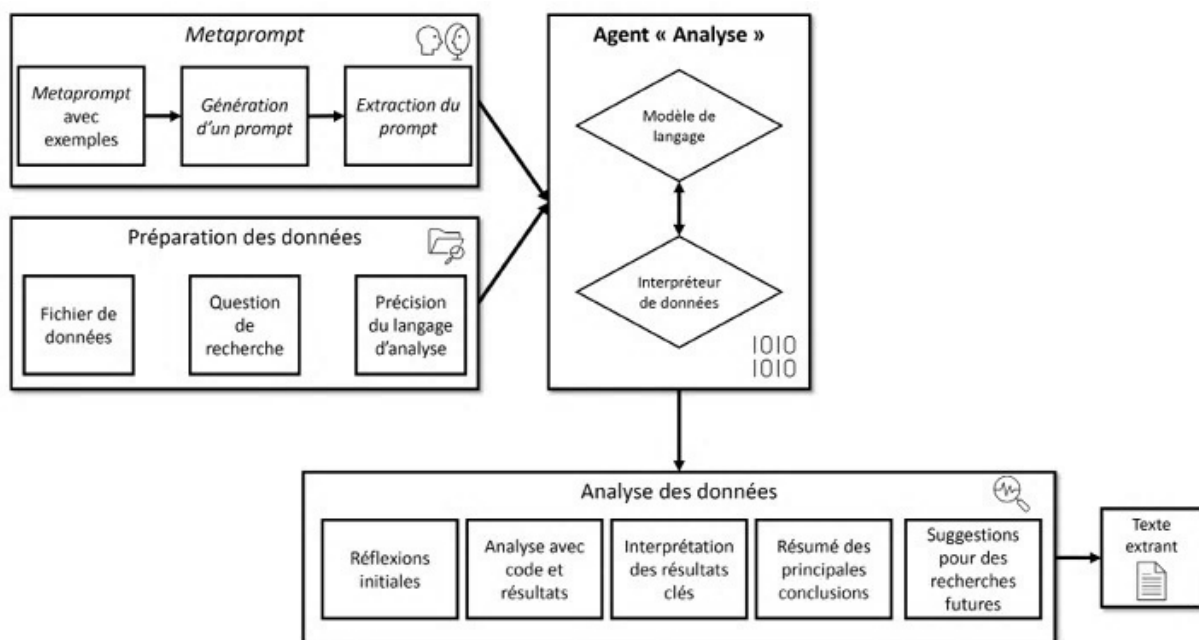
Pour mener à bien cette expérimentation, nous avons fait appel aux LLM les plus performants selon *lmsys.org* au mois d'avril 2024. Spécifiquement, nous avons utilisé GPT-4 (OpenAI, 2023) pour la création et l'opérationnalisation des agents, tandis que les modèles

de la série Claude d'Anthropic (2024) ont été employés pour la génération des *metaprompts*. L'évaluation des sorties des agents est basée sur la cohérence, la pertinence et l'originalité apparente des textes générés par chaque agent.

Nous avons créé trois agents pour différentes tâches académiques. Chaque agent est conçu pour effectuer une fonction spécifique dans le processus de recherche.

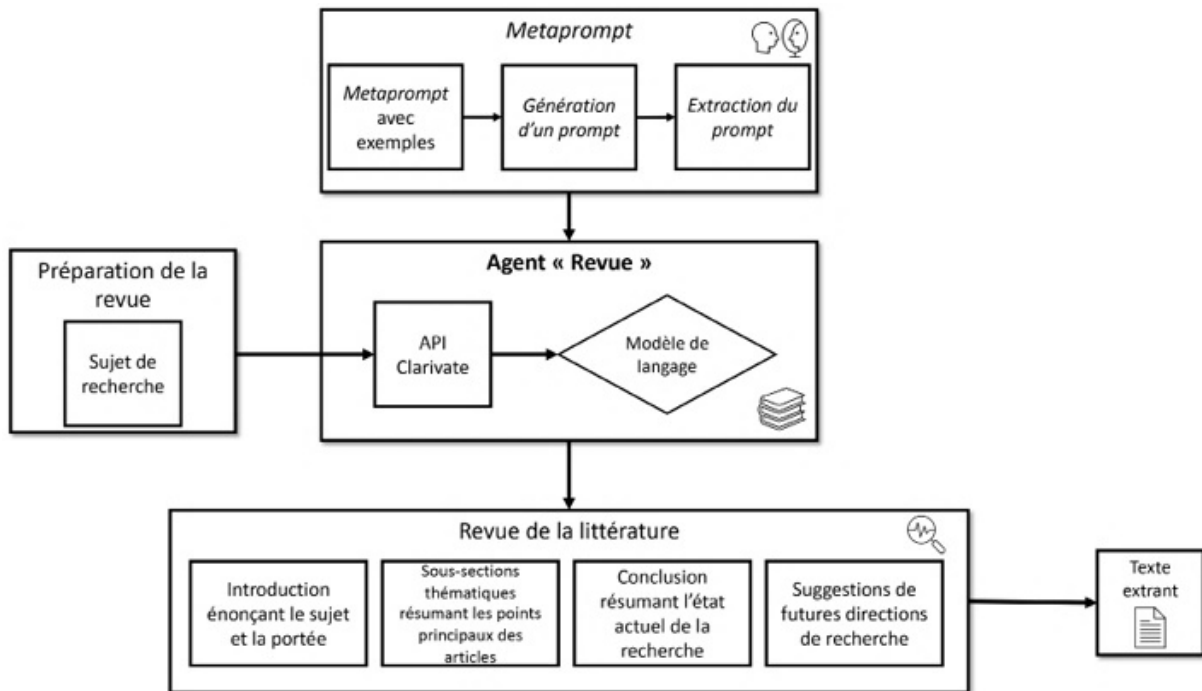
L'agent « Analyste » effectue des analyses de données approfondies en Python pour répondre à une question de recherche. Il prend en entrée un fichier de données et un code d'analyse existant et produit une analyse détaillée comprenant des réflexions initiales, l'analyse avec code et résultats, une interprétation des résultats et un résumé des principales conclusions (Figure 1).

Figure 1. L'agent « Analyste » pour l'analyse de données en Python



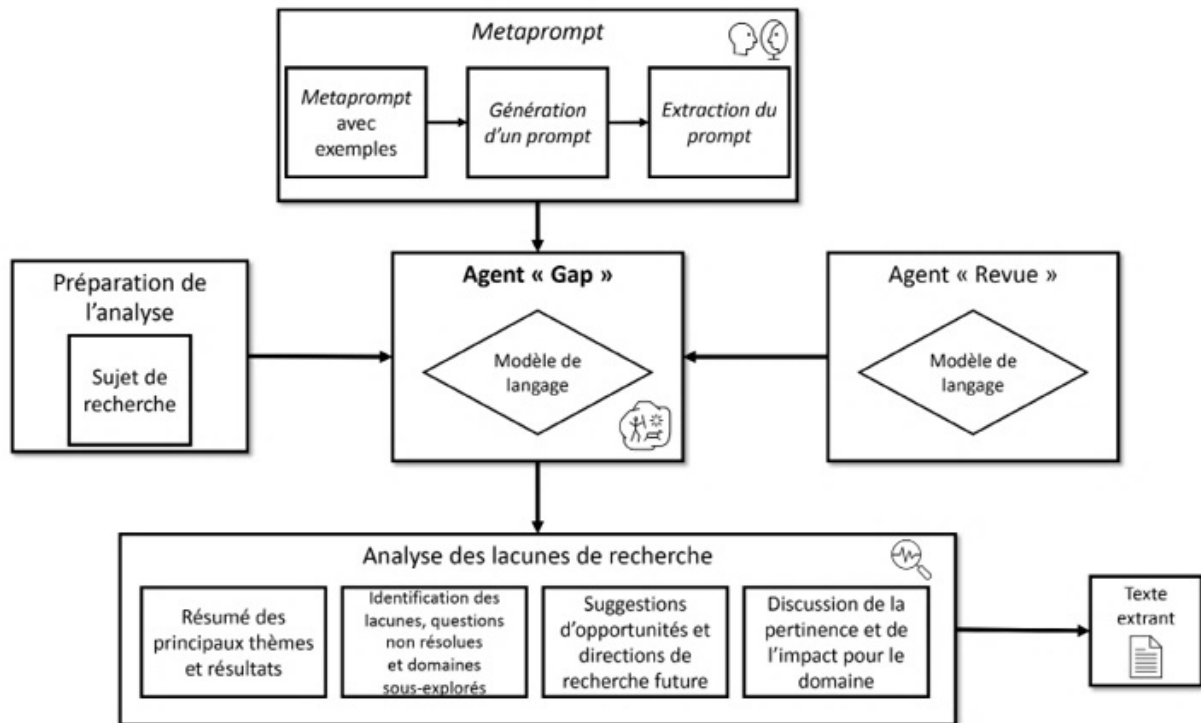
L'agent « Revue » rédige une revue de littérature sur un sujet donné. En utilisant l'interface de programmation *Clarivate* du *Web of Science*, il synthétise un ensemble d'articles pour produire une revue narrative comprenant une introduction, des sous-sections thématiques et une conclusion sur l'état de la recherche (Figure 2).

Figure 2. L'agent « Revue » pour la synthèse d'articles et la rédaction de revues de littérature



L'agent « Gap » identifie les lacunes et opportunités de recherche future dans un domaine, basé sur les articles analysés par l'agent « Revue ». Il produit une analyse des principaux thèmes, identifie les questions non résolues et suggère des pistes de recherche futures (Figure 3).

Figure 3. L'agent « Gap » pour l'identification de lacunes et d'opportunités de recherche



Contrairement aux LLM isolés qui peuvent inventer des citations fictives (Stokel-Walker & Van Noorden, 2023), notre approche par agents permet de produire un contenu sans hallucination ou confabulation. Pour l'expérimentation, nous avons utilisé un ensemble d'articles sur divers sujets scientifiques contemporains.

Les récentes avancées en IA générative soulèvent des questions cruciales sur la détection des contenus générés par des LLM. Perkins *et al.* (2024) mettent en évidence les limites significatives des outils de détection actuels, avec des taux d'exactitude faibles (39,5 %), diminuant à 17,4 % face à du contenu manipulé. Ces résultats soulignent la difficulté croissante de distinguer les textes générés par l'IA de ceux rédigés par des humains : notre étude adopte une approche déontologique, se concentrant sur les implications éthiques et pratiques de l'utilisation de l'IA générative dans la recherche académique, plutôt que sur l'analyse des textes générés. Cette perspective est d'autant plus pertinente que la capacité des systèmes d'IA à produire des textes de qualité est désormais établie, comme le soulignent Dwivedi *et al.* (2023) et Stokel-Walker et Van Noorden (2023).

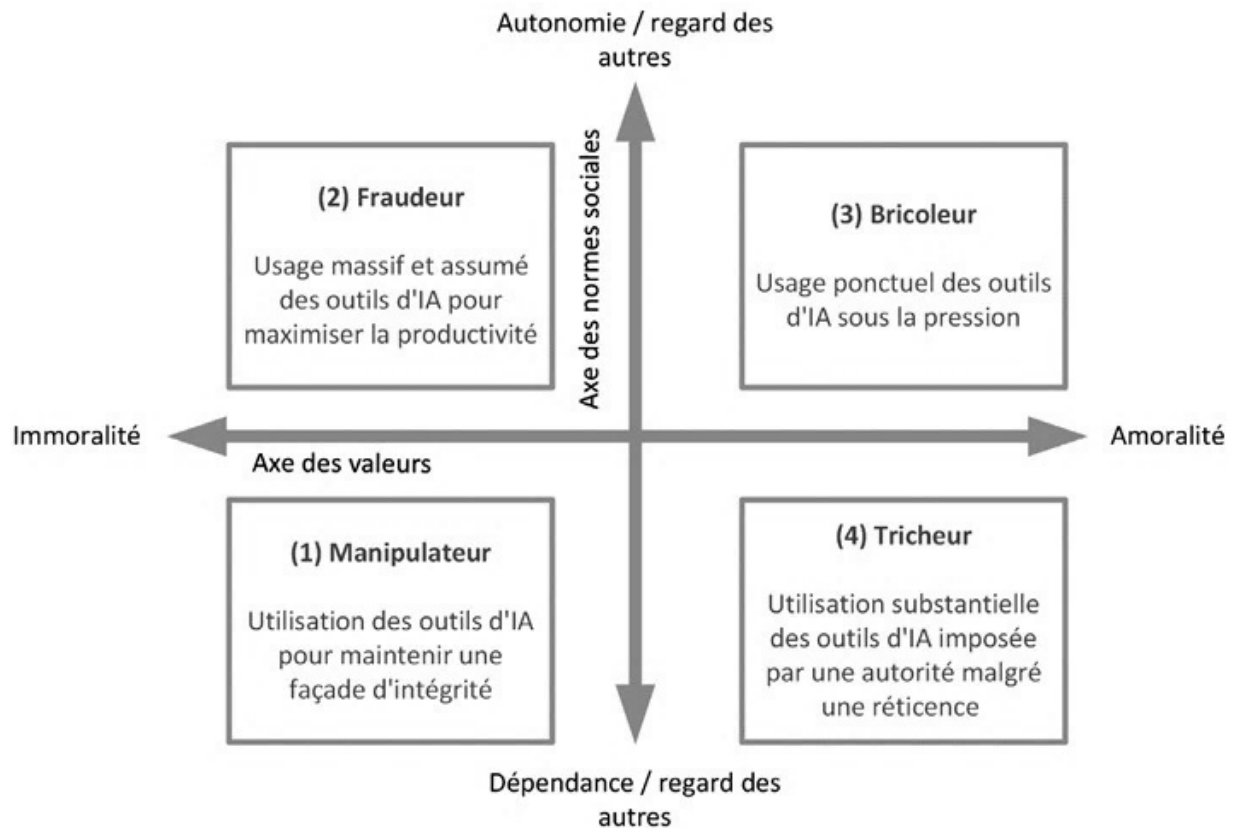
L'interconnexion de nos agents « Revue » et « Gap » démontre empiriquement les possibilités d'automatisation du processus de recherche. Ces résultats suggèrent que la détection et la prévention des usages abusifs de ces technologies émergentes nécessitent une réévaluation des principes fondamentaux de l'intégrité académique (morale, déontologie, éthique formelle et responsabilité) (Bergadaà, 2015 ; IRAFPA, 2024 ; Jeanne, 2023).

3.2. Application

Nous avons montré comment des acteurs pourraient potentiellement contourner les garde-fous existants en omettant des étapes cruciales du processus de recherche tout en créant l'illusion d'une productivité accrue. Pour analyser de manière systématique ces implications, nous appliquons le cadre conceptuel des quatre profils de délinquants de la connaissance proposé par Bergadaà (2015). Ce modèle identifie quatre types distincts de comportements délinquants : le manipulateur, le fraudeur, le bricoleur et le tricheur. Chaque profil est caractérisé par un rapport spécifique aux valeurs académiques et aux normes, ainsi que par des réactions stéréotypées lorsque la fraude est découverte.

Ces quatre profils peuvent être vus selon les deux axes orthogonaux issus du cadre proposé par l'IRAFPA⁴ :

Figure 4. Les quatre profils de délinquants de la connaissance



Source : Bergadaà, 2015

En les appliquant à notre expérimentation, nous pouvons décrire chaque profil comme suit :

(1) Pour le « **manipulateur** », les LLM représentent des outils puissants pour maintenir une image d'intégrité tout en s'adonnant à des pratiques discutables comme les falsifications de données de recherche. Il pourrait les utiliser pour générer des textes en apparence irréprochables et rendre la détection de la fraude encore plus ardue. Démasqué, il pourrait nier « avoir eu recours à l'IA » et affirmer que le contenu provient entièrement de son travail.

(2) Le « **fraudeur** » verrait dans l'IA générative un moyen d'accroître sa productivité apparente et de gravir les échelons plus rapidement. Peu attaché aux valeurs d'intégrité, il n'hésiterait pas à en faire un usage massif et assumé. La facilité d'utilisation de ces outils pourrait l'inciter à multiplier les fraudes dans différents contextes s'il parvient à créer de multiples agents pour automatiser la production d'articles creux quant à leur contenu.

(3) Pour le « **bricoleur** », les LLM constitueraient une tentation dans les moments de surcharge ou de pression. Il pourrait y avoir recours ponctuellement pour pallier un manque de temps ou d'inspiration, tout en culpabilisant de franchir la ligne rouge. La frontière entre usage acceptable et fraude deviendrait encore plus ténue pour ce type de délinquant ambivalent.

(4) Enfin, le « **tricheur** » pourrait se voir imposer l'utilisation d'un système de génération d'articles scientifiques par un supérieur mal intentionné, devenant complice de pratiques frauduleuses à grande échelle. N'ayant pas d'autonomie, il peinerait à s'y opposer et se réfugierait derrière la figure d'autorité. Découvert, il dirait que cela s'est fait « malgré lui ».

Des LLM de plus en plus performants amplifieront les comportements de ces profils. La sophistication de la fraude académique et l'accessibilité accrues de ces outils risquent d'exacerber les tendances existantes, tout en rendant la détection et la prévention de fraudes encore plus complexes. Ces développements risquent de mener à l'émergence de « systèmes de génération automatique de contenu scientifique », capables d'automatiser des pans entiers, voire l'intégralité du processus de recherche. Ces systèmes pourvus d'agents, comme dans notre expérimentation, permettront la production à grande échelle de travaux en apparence rigoureux mais dénués de véritable valeur scientifique. Leur existence soulève des questions sur la nature même de la recherche et de la production de connaissances à l'ère de l'IA générative.

4. DISCUSSION

Les résultats de notre expérimentation font apparaître plusieurs enjeux cruciaux pour l'intégrité académique à l'ère de l'IA générative. Cette démonstration contribue de manière significative à la littérature existante en explorant de façon concrète et plus approfondie les capacités de ces technologies du point de vue d'un délinquant de la connaissance (Bergadaà, 2015 ; Jeanne, 2023). Nos travaux se

distinguent des études précédentes qui se concentraient principalement sur l'émblématique ChatGPT (Dwivedi *et al.*, 2023 ; Emsley, 2023) ; ils montrent que ces enjeux concernent plus largement tous les modèles de langage. De même, alors que les recherches emploient des modèles de langage seuls (Dwivedi *et al.*, 2023 ; Stokel-Walker & Van Noorden, 2023), nous avons créé des agents capables d'interagir avec le monde numérique, révélant ainsi des possibilités d'automatisation encore plus poussées du processus de recherche. Ces agents produisent des résultats textuels d'une qualité telle qu'il devient manifestement impossible de les distinguer de ceux d'un chercheur humain, marquant ainsi un tournant critique dans la détection de leur usage et dans la création des connaissances.

Nous avons vérifié qu'il est faisable de développer des systèmes automatiques de recherche académique efficaces sans recourir à une plateforme tierce hébergeant des modèles génératifs, la seule limite étant la créativité du concepteur d'un tel système⁵. L'élargissement de ces enjeux à l'ensemble des modèles de langage fait apparaître un enjeu crucial : la détection des textes générés par LLM, les détecteurs actuels souffrant de taux de faux positifs trop élevés (Perkins *et al.*, 2023). Contrairement aux technologies antérieures de lutte contre le plagiat ou les fraudes qui sont liées à la puissance de calcul (Jeanne, 2023), l'IA générative présente de nouveaux défis. Les LLM, déjà largement utilisés, ne laissent aucune trace détectable, ce qui rend leur identification particulièrement complexe et pose des problèmes inédits en termes de mésusages potentiels.

À l'ère de l'IA générative, le concept d'auctorialité est profondément remis en question. Traditionnellement, l'auctorialité est définie comme « le statut, la qualité ou l'essence d'un phénomène généré par l'acte d'une personne produisant une œuvre de l'esprit, ou plus généralement une forme d'expression relevant du génie idiotique d'un individu, comme le font par exemple artistes, auteurs et autrices »⁶. Notre expérimentation apporte des éléments concrets

pour réexaminer les critères d'originalité et distinguer « l'originalité textuelle » de la véritable « création » de connaissance. L'émergence des systèmes d'IA pose de nouveaux défis à notre compréhension traditionnelle de la paternité intellectuelle et de la création de savoir.

En effet, nous avons prouvé qu'il est possible de générer des textes en apparence originaux, cohérents et pertinents sur un sujet donné, sans pour autant créer de nouvelles connaissances. Les agents « Revue » et « Gap » que nous avons développés peuvent produire une synthèse de la littérature existante et identifier des pistes de recherche futures de façon automatisée à partir d'un corpus d'articles. Mais ces capacités soulèvent des questions fondamentales : peut-on parler de véritable originalité et de création de connaissance dans ce cas ? Ne s'agit-il pas plutôt d'une forme avancée de compilation et de génération de texte, certes inédite, mais sans apport substantiel à la connaissance ? De même, l'agent « Analyste » peut générer une analyse de données avec interprétation écrite des résultats et conclusion, donnant l'illusion d'une contribution originale. Mais là encore, il s'agit en réalité d'une automatisation de tâches à partir de données et de codes existants, sans nouveauté conceptuelle ou méthodologique.

Ces exemples nous montrent que les LLM nous obligent à dissocier l'originalité apparente d'un texte (sur le plan de la forme, du style, de la combinaison inédite d'éléments) de l'originalité réelle de la contribution scientifique (sur le plan des idées, des concepts, des résultats nouveaux). Un texte généré par un LLM peut présenter une originalité *a priori* sans pour autant apporter une contribution scientifique authentique. Cette distinction souligne la nécessité cruciale de l'expertise humaine dans l'évaluation critique de la valeur réelle des productions scientifiques. Les chercheurs qui délèguent la révision d'articles à des LLM sans y apporter leur expertise risquent de perpétuer une évaluation superficielle, incapable de discerner la véritable originalité et la contribution scientifique d'un travail. Cette pratique, si elle se généralise, pourrait avoir des conséquences graves sur l'avancement des connaissances. Elle risque de créer un environnement où la prolifération de publications apparemment

« originales » masque une stagnation de l'avancement scientifique réel. À terme, cela pourrait compromettre la qualité et la fiabilité de la littérature scientifique, entravant le progrès dans divers domaines de recherche. Il est donc impératif que la communauté scientifique développe des méthodes rigoureuses pour intégrer les LLM dans le processus de révision, tout en préservant l'intégrité et la profondeur de l'évaluation par les pairs.

Cette dissociation entre originalité apparente et réelle soulève des questions fondamentales sur la nature de l'IA générative dans le contexte de la recherche scientifique. Il est impératif d'examiner en quoi ces systèmes diffèrent ontologiquement d'autres outils technologiques utilisés en science, tels que le télescope ou la calculatrice. Contrairement à ces outils qui augmentent nos capacités d'observation ou de calcul dans des domaines spécifiques, l'IA générative présente une dimension transformative plus profonde, qui remet en question les processus mêmes de la création et de l'évaluation du savoir scientifique.

Une différence majeure est que les systèmes d'IA générative ne se contentent pas d'augmenter nos capacités de recherche d'information ou de calcul, mais sont capables de produire des imitations textuelles passant pour des créations humaines originales. Ces systèmes brouillent les frontières entre augmentation, automatisation, assistantat et substitution dans des tâches cognitives complexes. En remettant en question les principes éthiques formels de la recherche, notamment en estompant la distinction entre l'assistance légitime et la substitution du travail intellectuel du chercheur, l'IA générative soulève des questions inédites sur le statut et la valeur des textes produits par ces systèmes. Peut-on les considérer comme des contributions scientifiques valables au même titre que ceux écrits par des humains ? Faut-il les évaluer selon les mêmes critères ? La seule obligation de transparence dans l'usage de ces systèmes et dans la mention des sources suffit-elle à garantir leur légitimité ?

Nos résultats suggèrent une réponse nuancée, voire négative, à ces interrogations. *Stricto sensu*, le texte généré par un système d'IA

est unique, mais cela ne garantit pas *ipso facto* son originalité ni sa valeur épistémique : ce risque est alors de voir se multiplier des publications « originales » en apparence, mais creuses ou redondantes sur le fond, sans réel apport de connaissance. D'où la nécessité d'élaborer des règles spécifiques aux systèmes d'IA générative, au-delà des principes généraux de transparence et de citation. Il faut pouvoir évaluer ce que l'usage de ces systèmes apporte réellement à la démarche et au raisonnement scientifique, au-delà de la seule performance générative. Cette évaluation peut se situer dans des critères qualitatifs comme la pertinence, la cohérence, la rigueur méthodologique, l'articulation aux travaux existants ou la portée épistémique des résultats.

L'adoption d'une posture critique s'avère indispensable pour éviter les pièges de l'illusion d'originalité et résister à la tentation de lui préférer la facilité apparente offerte par ces outils. L'enjeu réside alors dans le développement d'un usage scientifique intègre de l'IA générative, qui exploiterait pleinement son potentiel tout en maîtrisant ses risques. Il est nécessaire de soumettre à l'examen les critères habituels de l'originalité et de la créativité intellectuelles, garantes d'une authentique contribution à l'avancement de la connaissance, à la lumière de ces nouvelles technologies. Dans la section suivante, nous proposons des recommandations concrètes visant à établir un cadre préliminaire pour l'utilisation de l'IA générative dans la recherche.

5. CONCLUSION

Bien que les directives développées par l'UE constituent une base importante pour encadrer l'utilisation de l'IA générative dans la recherche, nous pensons qu'elles risquent de ne pas couvrir adéquatement toutes les possibilités de mésusage que nous avons identifiées. Au vu de notre expérimentation et de l'analyse qui en découle, on peut formuler plusieurs recommandations pratiques :

Pour les revues, il s'agit d'exiger l'explicitation des systèmes d'IA générative utilisés dans le processus de recherche, comme le

préconise l'Union européenne (2024). Comme le souligne Jeanne (2023), les politiques éditoriales ne sont pas adaptées à ces nouveaux défis et leur refonte en profondeur s'impose. En ce sens, les politiques éditoriales des revues scientifiques devraient exiger systématiquement que les auteurs signalent quel usage a été fait des modèles génératifs pour accomplir quelle tâche. En sus, nous recommandons aux revues scientifiques d'exiger spécifiquement des modèles et de la version utilisés.

Cette exigence de transparence permettra de répondre à plusieurs impératifs : d'une part, elle facilitera la reproductibilité des recherches en fournissant des informations détaillées sur les outils d'IA utilisés ; d'autre part, elle permettra une meilleure évaluation de l'impact de ces technologies sur le processus de recherche et les résultats obtenus. En outre, cette pratique contribuera à l'établissement d'une base de connaissances sur l'utilisation de l'IA générative dans la recherche académique, favorisant ainsi une compréhension plus approfondie de ses applications et de ses limites. Il incombe aux revues de demander aux auteurs d'indiquer quels contenus ont été produits à l'aide de quels outils génératifs. La transparence dans l'utilisation des systèmes d'IA générative pour la production de contenu académique est cruciale à plusieurs égards. D'un point de vue éthique, elle permet de maintenir l'intégrité du processus de recherche et la confiance de la communauté scientifique. L'absence de déclaration soulève en effet des questions sur la nature de la contribution intellectuelle du chercheur et sur le respect des normes académiques établies.

Cependant, il convient de nuancer ces remarques. L'utilisation des systèmes d'IA générative ne remet pas nécessairement en cause l'intégrité morale du chercheur, elle invite plutôt à une réflexion sur l'évolution des pratiques de recherche à l'ère de leur émergence. La question centrale est de savoir comment intégrer ces outils de manière éthique et productive dans le processus scientifique. Par ailleurs, la transparence soulève des questions importantes sur la responsabilité du chercheur vis-à-vis du contenu produit. Les modèles d'IA générative, malgré leur sophistication, peuvent générer

des erreurs ou reproduire des biais présents dans leurs données d'entraînement. Il incombe donc au chercheur de développer une approche critique et vigilante dans l'utilisation de ces outils, en vérifiant rigoureusement les *outputs* et en assumant la responsabilité finale du contenu publié.

Aux chercheurs, nous faisons une recommandation complémentaire aux directives de l'Union européenne (2024). Nous pensons qu'il est essentiel pour chaque chercheur de s'interroger sur l'utilisation de l'IA générative dans chacun des rôles qu'il assume. Partant des réflexions de Eke (2023) et de l'IRAFPA (2024), nous identifions d'abord plusieurs mauvaises pratiques dans l'utilisation des systèmes d'IA générative pour la production scientifique :

(1) Générer des données ou résultats non obtenus par une expérimentation ou recherche réelle.

(2) Accepter les informations fournies par l'outil sans vérification indépendante.

(3) Dépendre excessivement de l'outil pour générer du contenu, au détriment de l'effort personnel et de la pensée critique.

Pour éviter ces pratiques, nous proposons d'en user avec les contenus produits par une IA comme on en userait avec un stagiaire. L'analogie entre l'IA et le stagiaire permet d'aborder les systèmes d'IA de façon plus intègre :

(1) Superviser et vérifier activement le contenu généré par les LLM, comme on guiderait et évaluerait le travail d'un stagiaire.

(2) Ne pas surestimer les compétences de l'outil, en maintenant une expertise supérieure et un regard critique.

(3) Vérifier systématiquement les résultats fournis par l'outil en les confrontant à d'autres sources et à sa propre expertise.

(4) Éviter une dépendance excessive envers ces outils pour les tâches cruciales.

En définitive, l'émergence de l'IA générative nous invite à repenser en profondeur nos critères d'originalité, de créativité et de

contribution scientifique.

Notre expérimentation met en lumière la nécessité d'élaborer, au-delà des principes généraux de transparence et des règles de citation, des règles spécifiques à l'usage de l'IA générative. Il s'agit de développer une éthique exigeante pour l'usage intègre de ces technologies. Cela passe par une évolution de la politique éditoriale des revues, par une sensibilisation des chercheurs aux enjeux et par l'adoption de bonnes pratiques dans la supervision de ces « stagiaires » d'un nouveau genre que sont les systèmes d'IA générative. C'est à ce prix qu'au lieu d'une menace nous pourrions en faire de véritables outils au service de l'intégrité et d'une recherche féconde.

Références bibliographiques

- Anthropic (2024). *Metaprompt*. GitHub. <https://github.com/anthropics/anthropic-cookbook/blob/main/misc/metaprompt.ipynb>
- Baldassarre, M. T., Caivano, D., Fernandez Nieto, B., Gigante, D. & Ragone, A. (2023). The Social Impact of Generative AI: An Analysis on ChatGPT. *Proceedings of the 2023 ACM Conference on Information Technology for Social Good*, 363-373.
- Barrett, C., Boyd, B., Bursztein, E., Carlini, N., Chen, B., Choi, J., Chowdhury, A. R., Christodorescu, M., Datta, A., Feizi, S., Fisher, K., Hashimoto, T., Hendrycks, D., Jha, S., Kang, D., Kerschbaum, F., Mitchell, E., Mitchell, J., Ramzan, Z., ... Yang, D. (2023). Identifying and Mitigating the Security Risks of Generative AI. *Foundations and Trends in Privacy and Security*, 6(1), 1-52.
- Bergadaà, M. (2015). *Le plagiat académique. Comprendre pour agir*. L'Harmattan.
- Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., ... Liang, P. (2022). *On the Opportunities and Risks of Foundation models* [Prépublication]. arXiv.
- Bozkurt, A. (2024). GenAI et al. Cocreation, Authorship, Ownership, Academic Ethics and Integrity in a Time of Generative AI. *Open Praxis*, 16(1), 1-10.
- Comité National Pilote d'Éthique du Numérique (2023). *Systèmes d'intelligence artificielle générative : Enjeux d'éthique* (Avis 7 du CNPEN).
- Commission de l'intelligence artificielle (2024). *IA : notre ambition pour la France*.
- Dell'Acqua, F., McFowland, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Kraye, L., Candelon, F. & Lakhani, K. R. (2023). Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality. *SSRN Electronic Journal*, 58.

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., ... Wright, R. (2023). Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, 102642.

Eke, D. O. (2023). ChatGPT and the rise of generative AI: Threat to academic integrity? *Journal of Responsible Technology*, 13(4).

Emsley, R. (2023). ChatGPT: These are not hallucinations – they're fabrications and falsifications. *Schizophrenia*, 9(1), 52.

IRAFPA (2024). *Les repères de l'intégrité académique*. <https://www.iraipa.org/les-reperes-de-lintegrite-academique/>

Jeanne, L. (2022). Intégrité Académique et Intelligence Artificielle. Réflexions prospectives sur la base du cas Speedwrite. *Actes du 2^e Colloque IRAIPA*, 211-229.

Jeanne, L. (2023). Production d'imitations textuelles par une intelligence artificielle et prospective de la fraude académique. Dans M. Bergadaà (dir.), *Les nouvelles frontières de l'intégrité académique* (pp. 113-134). Éditions EMS.

Navigli, R., Conia, S. & Ross, B. (2023). Biases in Large Language Models: Origins, Inventory, and Discussion. *Journal of Data and Information Quality*, 15(2), 1-21.

OpenAI (2023). *GPT-4 Technical Report*. arXiv.

Perkins, M., Roe, J., Vu, B. H., Postma, D., Hickerson, D., McGaughran, J. & Khuat, H. Q. (2024). GenAI Detection Tools, *Adversarial Techniques and Implications for Inclusivity in Higher Education*. arXiv.

Stokel-Walker, C. & Van Noorden, R. (2023). What ChatGPT and generative AI mean for science. *Nature*, 614(7947), 5.

Union européenne (2024). *Responsible use of generative AI in research*.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. & Polosukhin, I. (2017). *Attention Is All You Need*. arXiv.

1. Note de l'auteur : ce document a été relu et affiné avec l'aide du modèle Claude-3-opus-20240229 d'Anthropic, en complément de l'édition humaine. L'auteur a validé le contenu de manière critique, il a évalué les biais potentiels et assume l'entière responsabilité de la version finale.

2. Par exemple, ChatGPT est l'outil d'IA générative qui emploie les grands modèles de langage propriétaires d'OpenAI, tels que GPT-4.

3. Ceci restant possible pour les modèles propriétaires (<https://arxiv.org/abs/2301.10226>), mais impossible pour les modèles publiés en *open source*.

4. Disponible à l'adresse suivante : <https://iraipa.org/a-propos/profils-delinquants/>.

5. Déjà même, à l'issue de la rédaction de ce chapitre, l'outil *open source* Flowise permet de réaliser tout ce qui a été présenté ici sans l'écriture d'une ligne de code : <https://github.com/FlowiseAI/Flowise>.

[6.](#) Wiktionnaire, 2024.

CHAPITRE 2. LES NOUVELLES REPRÉSENTATIONS MENTALES DU TRAVAIL UNIVERSITAIRE : LE CAS DES GÉNÉRATIONS Y ET Z

[Marie-Frédérique Bacqué](#), Professeure des Universités, Université de Strasbourg (France)

[Pedro Urbano](#), Professeur auxiliaire, Université de Coimbra (Portugal)

1. INTRODUCTION

Le fossé entre les générations n'est pas un phénomène récent, ni même strictement humain. Plus marqué dans les périodes de grands changements comme celle que nous vivons, ce phénomène peut se manifester dans divers aspects de la vie quotidienne, des valeurs morales aux choix de style de vie. Nos étudiants en master et en doctorat ainsi que les jeunes chercheurs font partie des générations dites Y et Z¹, qui ont grandi avec la technologie numérique et les réseaux sociaux. En tant que psychologues, éducateurs, mais aussi psychanalystes, nous les avons observés et considérons dans ce chapitre comment ils réagissent à la brusque émergence de l'intelligence artificielle générative dans leur vie quotidienne. Mais, rappelons au préalable les fondements de notre métier de chercheur dans un contexte fortement perturbé.

La recherche de la vérité est un aspect tellement fondamental de la science que, sans lui, la pratique de la science et la recherche scientifique seraient dénuées de sens. Bien entendu, la recherche de la vérité ne se fait pas n'importe comment, elle n'est pas une fin qui justifie n'importe quel moyen. Au contraire, comme l'ont montré les expériences innommables de Mengele à Auschwitz, elle doit absolument obéir à des principes éthiques et déontologiques.

Aujourd'hui, plus que jamais dans l'histoire de la science, on attend et on exige de ceux qui mènent des recherches scientifiques qu'ils le fassent avec intégrité, honnêteté et dans le respect des normes éthiques. Dans ce sens, il va presque sans dire que les

comportements répréhensibles, tels que le plagiat et la fabrication ou la falsification de données, compromettent la recherche de la vérité et la crédibilité des connaissances scientifiques. Toutes ces exigences et considérations s'appliquent *a fortiori* à ceux² qui apprennent à faire de la science et qui font leurs premiers pas dans la recherche, aux jeunes générations de chercheurs, y compris aux étudiants en master et en doctorat.

Toutefois, si l'exigence d'intégrité, d'honnêteté et de respect des normes éthiques dans l'exercice de la science existe depuis longtemps, le panorama a changé ces dernières années, avec la disponibilité et même la démocratisation de nouveaux outils dits d'intelligence artificielle (IA), facilement accessibles au grand public. Cette disponibilité a créé de nouveaux défis pour la communauté scientifique et surtout pour les universités, qui forment les nouvelles générations de scientifiques. C'est notamment le cas de ChatGPT, un agent conversationnel utilisant l'intelligence artificielle générative, capable de générer des textes cohérents et pertinents sur le plan contextuel, ressemblant à ceux d'un humain. Et tout porte à croire que ces outils seront encore perfectionnés.

Hélas, cette prouesse technologique peut être utilisée à mauvais escient, comme cela s'est produit dans le passé avec de nombreuses technologies, créées pour améliorer la vie humaine, mais qui ont souvent produit des résultats inattendus et parfois nuisibles (Tenner, 1997). Il existe un risque non négligeable que l'adoption de ces outils se généralise de manière incontrôlée. Par exemple, l'IA peut être utilisée pour rédiger des essais ou d'autres textes académiques sur mesure ; elle produit un contenu qui semble original, mais qui n'est pas le fruit du travail de ses auteurs ; ces textes ne représentent aucun effort véritable, ni compréhension ou connaissance authentique du sujet. L'IA peut même être utilisée, non sans ironie, pour paraphraser ou réécrire automatiquement un texte, afin de contourner les systèmes de détection du plagiat.

Il est évident que les outils d'IA ne sont pas en eux-mêmes à l'origine des mauvaises conduites dans la recherche. Certes, ils

peuvent les faciliter, les rendre plus fréquents ou plus difficiles à détecter, mais le problème de ces mauvaises pratiques, en particulier le plagiat, existait déjà avant même que les ordinateurs personnels n'existent. Érasme faisait déjà, dans son *Éloge de la folie*, la satire des plagiaires au XVI^e siècle. En effet, les causes de ces mauvais comportements sont anciennes et nous obligent à une réflexion plus approfondie. En même temps, il est de la responsabilité de tous – établissements d'enseignement, éducateurs, chercheurs et étudiants – de veiller à l'utilisation éthique des technologies d'IA dans la recherche et de favoriser une culture de l'intégrité académique. La sensibilisation, l'instruction civique et les mesures proactives peuvent jouer un rôle crucial pour minimiser l'impact négatif potentiel de l'IA sur les mauvaises conduites.

Il convient de noter que cette réflexion, limitée dans sa portée, ne traitera que du cas des nouvelles générations de chercheurs et d'étudiants universitaires, en master et surtout en doctorat. Mais cela ne veut pas dire que les autres acteurs impliqués ne doivent pas être pris en considération dans le cadre d'un débat plus étendu. Le problème des mauvaises conduites en milieu universitaire (où se déroule une part importante de la recherche scientifique) touche tout le monde.

Il est également essentiel de noter que ces problèmes ne sont pas abstraits, historiques (c'est-à-dire concernant uniquement le passé) ou purement techniques, et qu'ils ne se produisent pas dans le vide. Au contraire, ils font clairement partie de la vie et des expériences de vie des personnes qui étudient et travaillent dans les universités et les centres de recherche scientifique. Ces problèmes relèvent donc des domaines de la sociologie et de la psychologie sociale. Ils peuvent et doivent être traités en collaboration avec des disciplines plus philosophiques, notamment l'éthique et la déontologie.

Enfin, l'idée la plus erronée qui puisse émerger d'une lecture hâtive de ce texte serait que nous cherchons à montrer quelqu'un du doigt, à commencer par les deux générations identifiées, Y et Z. Nous prenons soin de nos étudiants et notre mission est de leur

transmettre nos connaissances, mais aussi nos valeurs éthiques. C'est pourquoi nous pouvons et devons aborder les problèmes concrets et spécifiques qui affectent leur travail aujourd'hui et qui, pour diverses raisons, n'ont pas affecté celui des générations précédentes. C'est aussi pourquoi nous parlons de fossé entre les générations, dans ce domaine spécifique, en sachant qu'il existe aussi dans d'autres domaines. En sachant aussi que nous utilisons une figure de rhétorique : il ne s'agit pas d'un véritable fossé, mais plutôt des marches sur l'échelle qui symbolise la longue histoire de l'évolution culturelle de l'humanité.

La question qui nous occupe est celle des différences qui peuvent exister dans la manière dont les jeunes générations perçoivent la recherche scientifique et tout ce qui a trait à l'intégrité dans les universités et les centres de recherche. Et même si les différences sont limitées, elles peuvent être significatives. C'est le cas, comme nous le verrons plus loin, de la perception de la valeur du travail intellectuel, qui a connu une transformation majeure au cours des dernières décennies. La génération Y, en particulier, a été en partie formée par des phénomènes sociaux inconnus auparavant, tels que le service Napster. C'est peut-être aussi le cas de la très connectée et technophile génération Z, qui peut rêver une manière idiosyncrasique de percevoir le travail scientifique, inspirée par l'expérience des jeux vidéo pendant l'enfance, l'adolescence et le début de l'âge adulte. En effet, les jeux en ligne offrent des plateformes sur lesquelles ces jeunes interagissent avec des pairs du monde entier, ce qui renforce chez eux un type d'interaction sociale inconnu auparavant. Or, la question n'est pas de savoir si ces interactions ont un effet sur les compétences sociales ou créent des sentiments d'appartenance ou d'identité. Il est évident et logique qu'elles le font. La question est de savoir comment elles fonctionnent et de comprendre dans quelle mesure elles auront un effet sur la pratique quotidienne de la recherche.

Tout cela ne signifie en aucun cas que notre attitude doive être vindicative ou répressive. Bien au contraire, la clé principale est le dialogue, la confrontation des idées, le partage des responsabilités

et la responsabilité mutuelle. Après ces remarques introductives, nous allons maintenant présenter notre réflexion, divisée en cinq points principaux.

2. LA PERCEPTION DE LA VALEUR DU TRAVAIL INTELLECTUEL POUR LES GÉNÉRATIONS Y ET Z

Commençons par un problème de fond, à l'origine de plusieurs mauvaises pratiques : la perception de la valeur du travail intellectuel. Sans pointer qui que ce soit du doigt, il faut reconnaître qu'il existe un fossé générationnel dans l'appréhension de la valeur du travail intellectuel et le respect des droits d'auteur.

La croyance que tout est gratuit dans le domaine numérique est une tendance observée chez les jeunes générations. Cette manière de se représenter la réalité est souvent influencée par l'omniprésence de contenus, de services et d'applications en ligne gratuits, apparemment gratuits ou partiellement gratuits (voir le modèle dit « freemium ») : cela renforce l'idée que les services de base sont gratuits. En effet, Internet offre un accès libre à de vastes quantités d'informations, y compris à des ressources éducatives, des articles et des contenus médiatiques. Cet accès facile à l'information sans coûts directs (évidents) peut contribuer à l'idée que tout est disponible gratuitement. En ce sens, le plagiat tend à ne pas être perçu par les jeunes générations comme un problème ou, encore moins, comme un délit. Dans certains cas, ils ne savent peut-être même pas ce qui constitue un plagiat. Dans d'autres cas, c'est la dynamique sociale entre pairs qui peut contribuer à la transgression : les étudiants peuvent se sentir obligés de se conformer au comportement des autres, en particulier si la fraude est courante et perçue comme acceptable au sein de leurs cercles sociaux.

Dans ces cas, la transgression peut même être considérée (tout comme le piratage informatique) sous un jour romantique, comme s'il s'agissait d'un acte de rébellion, perçu comme courageux, de la part de ceux qui n'acceptent pas passivement l'autorité des institutions, qu'elles soient académiques ou gouvernementales. Cette croyance,

selon laquelle tout est gratuit sur Internet, a une histoire qui remonte à une vingtaine d'années (au moins), aux formes les plus populaires de « piratage informatique », notamment par le biais des logiciels d'échange « *peer-to-peer...* » (P2P). L'exemple le plus connu est peut-être le phénomène Napster, qui a rapidement attiré des millions d'utilisateurs, notamment des étudiants et des jeunes adultes, séduits par la possibilité d'accéder à une quantité de musique illimitée sans frais. Le phénomène a eu un impact culturel profond. En raison de l'immense popularité du modèle P2P et de la façon dont il a été exploité pour le partage non autorisé de matériel protégé par des droits d'auteur, le piratage sur Internet a cessé (en règle générale) d'être perçu comme un problème grave ou comme un délit par les nouvelles générations. Cette représentation sociale perdure encore aujourd'hui : les chiffres de téléchargement illégal de musique, films, séries, logiciels et autres contenus numériques se chiffrent en milliards par an, sans que les membres de ces générations en conçoivent la moindre gêne.

Il en va de même pour le téléchargement illégal d'articles et de livres scientifiques, qui s'avère coûteux. Il existe des sites web bien connus, fréquentés par de nombreux universitaires du monde entier, qui sont spécialisés dans ce type de produits intellectuels. Le plus connu est peut-être « Sci-Hub », qui peut être comparé à un marché noir et qui propose plus de 80 millions d'articles scientifiques en accès libre sans respecter le droit d'auteur : on les obtient par *web scraping* (moissonnage) en contournant les *paywalls* (péages) des éditeurs académiques. On pourrait nommer plusieurs autres sites, comme la Library Genesis. Dans tous ces cas, le piratage informatique est perçu au mieux comme un crime sans victimes. Ou bien il est minimisé par le récit autojustificateur d'une justice directe ou « poétique » : voler aux riches, notamment aux éditeurs avides d'argent, pour distribuer aux pauvres. C'est le cas d'Alexandra Elbakyan, qui a fondé Sci-Hub en 2011, ce qui lui a valu le surnom de « Robin des Bois de la science ».

Enfin, nous ne devons pas oublier que l'essor de l'enseignement en ligne et de l'apprentissage à distance, notamment en réponse à la

pandémie de Covid-19, a modifié l'environnement social traditionnel de l'apprentissage. La socialisation par les pairs et la pression de ces derniers, y compris la pénalisation des mauvais comportements, ont largement disparu au cours de cette période.

Mais comment parler aux jeunes générations du travail intellectuel ? Pour les raisons évoquées précédemment, un premier pas vers la résolution du problème des mauvaises conduites pourrait consister à faire appel aux contextes d'éducation et de développement, en particulier la famille, l'école et la société, dans leur ensemble, pour promouvoir la culture numérique et un comportement éthique : il faut souligner l'importance du respect des droits de propriété intellectuelle, du soutien aux créateurs de contenu et de la divulgation des coûts réels associés au développement et à la maintenance des produits et services numériques.

3. LE MÊME COMPORTEMENT NE REFLÈTE PAS FORCÉMENT LES MÊMES MOTIVATIONS

Un autre point fondamental, qu'il nous semble essentiel de clarifier au plus vite, est de ne pas considérer les mauvaises conduites en milieu universitaire comme un ensemble, dans lequel rentreraient toutes les raisons, toutes les motivations, toutes les catégories ou tous les types d'utilisation de données. Faute de quoi, on risque de confondre l'arbre et la forêt et de voir plus de malhonnêteté qu'il n'y en a en réalité.

Si on parle spécifiquement des cas réels (*versus* des cas surestimés) de malhonnêteté intellectuelle des étudiants, on doit reconnaître que leurs causes sont multiples, comme ce serait le cas dans tout autre groupe de personnes. De même, il est nécessaire de prendre en compte le type et la gravité du comportement malhonnête. Dans certains cas, il peut s'agir de situations ponctuelles ou sporadiques, résultant (par exemple) d'une mauvaise gestion du temps ou du fait que les étudiants considèrent que le travail n'est pas important. Comme Internet a rendu de vastes quantités d'informations facilement accessibles tout en *permettant*

l'anonymat, il est inévitable que la tentation d'utiliser des sources en ligne sans les citer se présente, ce qui se traduit par des actes de plagiat par copier-coller. Il en va de même pour l'utilisation de documents pré-écrits et d'autres ressources en ligne. Bien qu'il s'agisse de comportements malhonnêtes, ils ne sont pas nécessairement systématiques, mais plutôt opportunistes et occasionnels. Pour lutter contre ces formes occasionnelles de plagiat, il faudra peut-être encourager l'investissement dans le travail et expliquer ce qu'est le plagiat et l'attribution correcte des sources.

Dans d'autres cas, cependant, les causes sont pour ainsi dire endémiques ou plus profondes. Les pressions académiques, notamment la pression de la réussite et le désir d'obtenir de bonnes notes, font partie de ce groupe. Des pressions qui affectent non seulement les étudiants, mais aussi l'ensemble de la communauté universitaire. Ces pressions résultent souvent du stress et des exigences imposées aux individus dans un cadre éducatif. Elles sont liées aux attentes, à la compétition, à la charge de travail et à la recherche de la réussite. Ces pressions, et bien d'autres encore, peuvent contribuer à la malhonnêteté, parce qu'elles encouragent certains étudiants à recourir non seulement aux outils d'IA, mais aussi à d'autres ressources ou services glanés sur la toile, par exemple les « usines à dissertations » ou les « services de rédaction de devoirs », qui se transmettent entre étudiants, pour faire face aux exigences de la charge de travail ou pour répondre aux attentes perçues. Ces cas constituent un problème complexe, la solution devra donc tenir compte de cette complexité, mais elle devra passer, une fois de plus et comme dans tous les cas mentionnés ci-dessus, par la promotion d'une culture de recherche intègre. Les différentes chartes européennes de déontologie et d'intégrité scientifique convergent sur ce point, parce qu'elles définissent des stratégies pour une telle promotion³.

Enfin, il existe des cas de malhonnêteté prémédités y compris la fraude aux thèses. Ces cas, dont le nombre est difficile à estimer, mais qui devraient varier considérablement en fonction de la tradition

de recherche scientifique de chaque pays, dépassent le cadre de cet exposé.

4. ILS VIVENT UNE ÉPOQUE SENSATIONNELLE

Un troisième point, qui peut passer inaperçu en raison de la tendance des médias et de l'opinion publique à rechercher le sensationnel, est la connaissance rigoureuse de la réalité. Bien qu'il y ait occasionnellement des articles de presse suggérant une augmentation des mauvaises pratiques, en particulier le plagiat, dans les milieux universitaires, il est essentiel de s'appuyer sur des études approfondies et rigoureuses, ou sur des données fiables (provenant notamment d'établissements d'enseignement), avant de tirer des conclusions sur cette tendance supposée. Par ailleurs, afin d'obtenir une image plus complète de ce qui est sans aucun doute un problème complexe, il est essentiel de considérer aussi les facteurs psychosociaux. Il ne faut surtout pas minimiser l'importance des études psychologiques et sociologiques approfondies pour comprendre les raisons de ces éventuelles tendances.

En d'autres termes : l'existence d'outils d'IA, leur accès facile et gratuit, l'anonymat permis par l'Internet, la culture dominante parmi les jeunes générations, forment une véritable *anthropologie du savoir* nouvelle selon laquelle tout ce qui se trouve sur Internet est (ou devrait être) gratuit. La pression sociale exercée par les pairs pour que le monde soit perçu d'une certaine manière, etc., peut faciliter la fraude, la malhonnêteté ou le manque d'intégrité, mais elle ne crée pas chez chaque personne la volonté de les commettre. De la même manière, on n'imagine pas que la plupart des (jeunes) gens soient malhonnêtes ou choisissent de tricher chaque fois que cela est possible. Au contraire, il est essentiel de reconnaître que tous les étudiants ne trichent pas et que nombre d'entre eux accordent une grande importance à l'intégrité académique. Il est également essentiel de reconnaître que la société et les établissements d'enseignement jouent un rôle crucial dans la promotion d'une culture de l'honnêteté et d'un comportement éthique ; mais il faut

également tenir compte des raisons socio-économiques en soutenant, le cas échéant, les étudiants confrontés à des difficultés scolaires. Enfin, il est essentiel de reconnaître que le fait même de s'attaquer aux causes profondes de la malhonnêteté académique peut contribuer à créer un environnement dans lequel les étudiants sont motivés pour apprendre et réussir par des moyens éthiques.

5. SOMMES-NOUS DES MODÈLES OU DES ILLUSIONS ?

Le quatrième point, peut-être le moins évident de tous, concerne les exemples inconscients et involontaires que nous pouvons donner à nos étudiants. En d'autres termes, même si le problème des mauvaises conduites finit par se manifester chez nos étudiants, il se peut qu'il provienne en partie de certaines de nos pratiques.

Il s'agit d'un sujet complexe, qui relève avant tout de la psychologie ou de la psychanalyse⁴ et qui doit être étudié en profondeur. On peut néanmoins reconnaître, en quelques mots, que le tutorat d'un doctorant, ne se résume pas à la transmission de connaissances ou à la simple supervision de son travail. Tout cela est bien sûr important, mais ne correspond qu'à la partie visible de « l'iceberg ». Cette métaphore est souvent utilisée pour expliquer des problèmes dans lesquels on risque de prendre la partie visible pour le tout comme dans la notion de « zone aveugle de la science » proposée par Edgar Morin (1982). Il s'agit aussi de transmettre des valeurs, d'enseigner par l'exemple, etc.

Il est essentiel de souligner à ce stade que nous avons la conviction profonde que la plupart des chercheurs scientifiques ne tolèrent pas les mauvaises pratiques et qu'en outre, la situation varie considérablement d'un pays à l'autre, les scientifiques qui effectuent des recherches dans des économies émergentes ou dans lesquelles il n'existe pas de solide tradition de recherche scientifique étant plus vulnérables. Cela ne signifie pas pour autant que nous devons fermer les yeux et ignorer ou oublier une contrainte primordiale, mais qui tend à être reléguée au second plan, étant même devenue, pour utiliser l'expression à la mode, « l'éléphant dans la pièce » : le

« *publish or perish* », c'est-à-dire cette réalité bien connue et omniprésente depuis plusieurs décennies de la pression intense (voire tyrannique) exercée sur le monde universitaire par l'obligation de publier le plus régulièrement possible des articles dans des revues scientifiques. Connue superficiellement par tout le monde, cette question est presque devenue un sujet tabou dans la pratique, probablement en raison de plusieurs facteurs : l'inconfort et les conflits qu'elle peut créer, le déni et la peur des conséquences ou encore le conformisme d'une situation connue depuis des décennies, mais à laquelle il n'y a, manifestement, pas beaucoup de volonté de trouver une solution.

Des dizaines de milliers de pages pourraient être écrites, et l'ont déjà été, sur cette question et ses diverses conséquences négatives, voire néfastes : elle tend notamment à générer de la quantité au détriment de la qualité ; elle a aussi de nombreuses implications éthiques qui affectent l'intégrité scientifique. On peut signaler à titre d'exemple que le stress considérable entraîné par ce système encourage la publication de nombreuses études peu significatives, au détriment de travaux plus importants et approfondis.

La pression à la publication peut également conduire à d'autres pratiques contraires à l'éthique : la soumission de travaux non originaux ou le fractionnement artificiel des recherches en plusieurs articles (« *salami slicing* »), afin d'essayer de manipuler les systèmes de récompense académiques ou simplement d'augmenter le nombre total de publications, ce qui peut être avantageux pour une évaluation de carrière, une demande de financement, etc. Mais il y a un deuxième « éléphant dans la pièce » : le fait que les mauvaises pratiques sont parfois tacitement acceptées par les encadrants eux-mêmes. Sans là encore pointer du doigt qui que ce soit, nous pouvons et devons nous poser quelques questions gênantes. Par exemple, si nos étudiants en master et en doctorat, qui font leurs premiers pas dans la recherche scientifique, voient leurs maîtres ou tuteurs diviser artificiellement les résultats d'un projet de recherche, contre leur volonté, mais en raison des exigences et de la pression imposées par le système, quelles valeurs leur apprend-on ? Ne

sommes-nous pas tenus, au contraire, d'aborder ce genre de situations avec nos étudiants, de manière tout à fait claire, malgré la gêne qu'elles peuvent occasionner ? Faire autrement revient à ignorer « les éléphants dans la pièce » et à refuser d'accepter la réalité, parce qu'elle est trop perturbante ou difficile à gérer. Or, dans la pratique, cela se traduit presque toujours par une aggravation de la situation. Les reconnaître et les aborder est crucial pour résoudre les problèmes sous-jacents et avancer.

Bien entendu, nous n'épuisons pas la situation : à ces problèmes bien connus, nous pourrions ajouter plusieurs autres problèmes connexes qui ont également un impact sur la formation de nos étudiants. Un exemple relativement récent et très pertinent est celui de la soi-disant « crise de la réplication... » (ou de la « répliquabilité »), sans aucun doute un problème majeur dans les sciences expérimentales : les biais de publication et le biais de confirmation, la manipulation (consciente ou pas) des analyses pour obtenir des résultats statistiquement significatifs (« *p-hacking* »), le « *data dredging* » ou encore l'ajustement des variables, la sélection des données ou la modification des hypothèses après coup (« HARKing »). Depuis au moins vingt ans, plusieurs auteurs (par exemple Ioannidis, 2005) s'inquiètent régulièrement du fait que la plupart des résultats de recherche publiés actuellement sont faux ou que la publication sélective des seuls résultats « positifs » pour corroborer les hypothèses testées peut fausser les preuves disponibles (Scheel *et al.*, 2021). Toutes ces pratiques doivent être discutées et dénoncées avec nos doctorants de manière ouverte et exhaustive. Prétendre qu'elles n'existent pas est une invitation à les utiliser.

6. QU'ATTENDONS-NOUS D'EUX ?

Un point qui tend à passer encore plus inaperçu que les précédents et qui ne sera qu'effleuré ici, concerne la société et ses attentes actuelles à l'égard des jeunes générations, des étudiants

universitaires en général et plus précisément des doctorants et des jeunes chercheurs.

En effet, les attentes élevées – de la part des parents, des enseignants, des pairs et des étudiants et aspirants chercheurs eux-mêmes – est un phénomène actuel remarquable. Ces attentes, qui se combinent à l'anxiété de la performance et à la peur de l'échec, créent une pression intense. Nous pouvons également nous interroger sur le rôle d'une culture axée sur les notes, qui, elle aussi, engendre la compétition et le désir d'excellence académique.

Enfin, en ce qui concerne encore et spécifiquement les jeunes générations, il est impossible de ne pas prendre en compte des phénomènes tels que la comparaison sociale constante, le « syndrome FOMO » (« *fear of missing out* »), l'anxiété sociale, la peur constante de manquer quelque chose, très favorisée par les médias sociaux ; ou encore l'équilibre difficile entre des sollicitations trop nombreuses et des engagements multiples. Tous ces phénomènes, tant sociaux qu'académiques, engendrent une pression très forte et prédisposent aux mauvais comportements, le manque d'intégrité, le recours au plagiat ou d'autres manières d'agir inappropriées. Les attitudes sociétales à l'égard de la réussite scolaire peuvent, par exemple, influencer les comportements : si la société accorde trop d'importance aux notes ou aux résultats scolaires, les étudiants peuvent être davantage tentés de plagier pour répondre à ces attentes.

7. SUR QUELLE TOILE DE FOND TISSENT-ILS LEUR AVENIR ?

Ces différents problèmes de nature psychosociale ne se produisent évidemment pas dans le vide. Il est donc de notre devoir de considérer leur contexte général, la toile de fond sur laquelle ils s'inscrivent. Et il nous semble impossible de caractériser ce contexte sans évoquer un phénomène culturel majeur aux États-Unis et dans les pays de leur sphère d'influence.

Il s'agit de ce que l'écrivain et journaliste Tom Wolfe (1976) a mis en évidence dans un article devenu emblématique : il a qualifié les

années 1970 aux États-Unis de « décennie du moi ». L'idée sera développée quelques années plus tard par Christopher Lasch (1979, 1985), qui défend longuement la thèse selon laquelle le pays cultive le narcissisme, un amour de soi exagéré. Dans les pathologies du narcissisme, l'amour de soi est survalorisé au détriment de l'intérêt pour autrui et le monde. Selon Freud (1914), puis Kernberg (1975), les personnalités narcissiques projettent sur autrui leur idéal du moi en clivant la réalité. Elles rejettent alors ce qui ne s'intègre pas dans leur souhait inconscient, tout en se repliant sur leur autoappréciation d'un moi infatué. Quelques décennies plus tard, Jean Twenge a même postulé l'existence d'une épidémie (métaphorique) de narcissisme : conséquence de la culture dénoncée par Lasch, cette épidémie se manifeste dans la psychopathologie et la santé mentale (Twenge & Campbell, 2010) ; cette dernière hypothèse est cependant relativement controversée.

Ce phénomène est difficile à cerner et doit donc faire l'objet d'une analyse plus approfondie. En effet, indépendamment de l'existence ou non d'une épidémie de narcissisme, un phénomène culturel et social a certainement émergé aux États-Unis à partir des années 1970. Selon Lasch (1979), ce phénomène correspond à une normalisation de la pathologie narcissique dans la culture populaire. À partir d'une évolution signalée par la psychiatrie à cette époque – la multiplication des cas-limites (troubles « *borderline* » ou troubles de la personnalité-limite), Lasch signalait l'émergence d'une personnalité narcissique non plus à l'échelle individuelle (psychopathologique), mais en tant que phénomène de société. De là sont nés des épiphénomènes tels que le culte de la jeunesse – en fait, un culte du moi déguisé – ou une culture de l'instant présent, centrée sur la satisfaction des besoins immédiats.

D'une manière ou d'une autre, le narcissisme a donc pris de l'ampleur dans sa double dimension psychologique et sociologique, comme dans un processus de coévolution : il est sorti du cabinet de consultation pour coloniser la place publique, passant de rare ou résiduel à fréquent ou normal. Par ailleurs, il a été exporté vers l'Europe et les autres pays dans la sphère d'influence des États-Unis

comme élément du bagage culturel populaire américain bénéficiant d'une large diffusion, au même titre que le cinéma hollywoodien et les soi-disant médias sociaux. L'*American way of life*, qui met l'accent sur l'individualisme ou l'éthique de l'autosuffisance, n'a jamais perdu son pouvoir de séduction.

Les jeunes générations n'ont pas créé le monde dans lequel elles ont été formées et vivent, ce sont les générations de leurs parents et grands-parents qui ont créé ce monde organisé par des valeurs propres, dont la focalisation sur l'individu (le Moi) plutôt que sur le collectif. On en trouve un exemple, caricatural, mais non négligeable, chez De Wall *et al.* (2011), qui ont analysé les paroles de près de 300 chansons de musique pop parmi les plus populaires aux États-Unis entre 1980 et 2007 : « Au fil du temps, l'utilisation de mots liés à la focalisation sur soi et au comportement antisocial a augmenté, tandis que les mots liés à la focalisation sur les autres, aux relations sociales et aux émotions positives ont diminué. » (p. 200)

Les nouvelles générations n'ont pas été les réceptacles passifs de ces formes culturelles, elles ont aussi créé leur propre culture. Mais il est plus ou moins établi que le développement des tendances narcissiques (en termes de traits de personnalité, mais aussi au niveau pathologique) est influencé, entre autres causes, par les contextes éducatifs, les influences culturelles et divers agents environnementaux.

8. DISCUSSION

L'arrivée de l'IA et des incroyables exploits qu'elle va réaliser pour des jeunes extrêmement prompts à s'approprier toutes les ressources technologiques gratuites et à les transformer en extensions d'eux-mêmes nous interpelle. Et c'est très bien ainsi, car nous ne saurions comprendre ce qui est en train de se jouer sans commencer par déchiffrer les comportements de ceux qui adoptent déjà ces outils.

C'est un sujet vaste et complexe, avec de nombreuses subtilités et une grande variabilité selon les pays et les cultures, en particulier les

cultures davantage axées sur la consommation et le consommateur et saturées de médias (Tyler, 2022). Il n'est pas possible de tirer des conclusions directes de la supposée culture du narcissisme ni pour la population générale (dans les pays occidentaux), ni pour les jeunes générations.

Néanmoins, il est possible de poser des questions et de tenir compte de ces hypothèses, en particulier de celles de Lasch, dans nos futures recherches. Nous nous demandons ainsi quels sont les effets concrets, au jour le jour, d'une vision du monde caractérisée par des rêves narcissiques de célébrité et de gloire (hypertrophiés par les réseaux sociaux), rendant difficile l'acceptation d'une vie banale, créant et renforçant le solipsisme et diverses formes de dépendance, notamment à la validation de l'estime de soi par des tiers. Plus spécifiquement, nous nous demandons si nos doctorants réalisent que faire de la science n'est pas l'apanage de génies (comme Galilée, Darwin ou Einstein), mais un travail collectif, un humble et patient travail de fourmi dans la fourmilière. Seront-ils prêts à faire les sacrifices nécessaires ?

Nous nous demandons si nos doctorants, pris au piège d'une culture de l'instant présent centrée sur la satisfaction des besoins immédiats, seront capables de tolérer la frustration implicite dans la pratique de la recherche. Comme le soulignait si justement Einstein (1950/2006), la construction de la connaissance en science est le résultat d'un processus d'adaptation extrêmement laborieux, hypothétique (par essence) et toujours soumis à la remise en question et au doute. Une carrière scientifique dure en moyenne toute la vie, sans grandes récompenses matérielles, sans grande (voire aucune) notoriété médiatique, parfois même sans congés ni vacances. Nos doctorants en sont-ils conscients ?

Pour conclure, nous nous demandons également si certains cas de mauvaises conduites ne reflètent pas finalement une vision fondamentalement erronée – ou trop idéalisée – de la science et de la pratique de la recherche. Les représentations stéréotypées de la science et des scientifiques, véhiculées notamment par la télévision mondiale, les films ou les médias sociaux, ne seraient-elles pas à

l'origine d'importants malentendus sur la nature laborieuse, lente et frustrante de la construction du savoir scientifique ? Après tout, la confusion sur l'essence même de la science est peut-être à l'origine de certains mauvais choix, dans l'université ou la recherche ?

Il y a encore du chemin à faire.

Références bibliographiques

Baumeister, R., Campbell, J., Krueger, J. & Vohs, K. (2005). Exploding the self-esteem myth. *Scientific American Mind*, 16(4), 50-57.

DeWall, C., Pond, R., Campbell, W. & Twenge, J. (2011). Tuning in to psychological change: Linguistic markers of psychological traits and emotions over time in popular U.S. song lyrics. *Psychology of Aesthetics, Creativity, and the Arts*, 5(3), 200-207.

Einstein, A. (1950/2006). *The Einstein Reader*. Citadel Press. (Réédité de *The Einstein Reader*, 1950, Citadel Press).

Freud, S. (1914). Zur einföhrung des narzissmus. *Jahrbuch der Psychoanalyse*, 6, 1-24.

Ioannidis, J. (2005). Why most published research findings are false. *PLOS Medicine*, 2(8), e124. <https://doi.org/10.1371/journal.pmed.0020124>

Kernberg, O. F. (1975). *Borderline conditions and pathological narcissism*. Jason Aronson.

Lasch, C. (1979). *The culture of narcissism*. W.W. Norton.

Lasch, C. (1985). *The minimal self*. W.W. Norton.

Morin, E. (1982). *Science avec conscience*. Fayard.

Scheel, A. M., Schijen, M. R. M. J. & Lakens, D. (2021). An excess of positive results: Comparing the standard psychology literature with registered reports. *Advances in Methods and Practices in Psychological Science*, 4(2).

Tenner, E. (1997). *Why things bite back*. Vintage Books.

Twenge, J. & Campbell, W. (2010). *The narcissism epidemic*. Simon & Schuster.

Tyler, I. (2022). De la « décennie du moi » au « millénium du moi ». Une histoire culturelle du narcissisme aux États-Unis. *Le Temps des médias*, 38, 53-72.

Wolfe, T. (23 August 1976). *The "Me" decade and the third great awakening*. New York.

1. Les membres de la génération Y ou « milléniaux » ont aujourd'hui entre 30 et 45 ans et ont fait leurs études avec un Internet omniprésent. Les membres de la génération Z ont aujourd'hui 15 à 30 ans et sont ultra connectés.

2. Par souci de brièveté, les accords grammaticaux seront mis au masculin dans un sens générique.

3. Voir par exemple la charte de l'Agence Nationale de la Recherche française (ANR).

[4.](#) Voir par exemple les concepts psychanalytiques d'identification, de projection ou de l'idéal du Moi.

CHAPITRE 3. IA GÉNÉRATIVE ET INTÉGRITÉ ACADÉMIQUE : APPROCHES PÉDAGOGIQUES POUR FORMER À L'ÉTHIQUE DE L'IA EN ENSEIGNEMENT SUPÉRIEUR

[*Frédéric Bruneault*](#), École des médias, UQAM et philosophie, Collège André-Laurendeau,
Montréal (Canada)

[*Andréane Sabourin Laflamme*](#), philosophie, Collège André-Laurendeau, Montréal (Canada)

1. INTRODUCTION

Depuis l'arrivée de ChatGPT en 2022, suivi par de nombreux autres modèles d'IA générative, le paysage technologique a profondément évolué. La démocratisation du recours aux grands modèles de langage (LLM) dans la population a suscité une transformation marquante dans l'enseignement supérieur, où l'utilisation croissante de l'IA générative par les étudiants, déjà très habiles pour se saisir des innovations technologiques, soulève de nombreuses questions. Si l'IA générative offre des possibilités inédites en matière d'éducation, elle a aussi un potentiel disruptif pour de nombreuses facettes de l'enseignement et pour l'évaluation des compétences et des connaissances (Farrelly & Baker, 2023).

Une étude réalisée en 2023 révèle que plus de la moitié des étudiants canadiens de 18 ans et plus ont recours à l'IA générative pour réaliser des travaux académiques. Ils l'utilisent notamment pour générer des idées, rédiger des essais et synthétiser des informations. Mais 60 % d'entre eux considèrent cette pratique comme une forme de tricherie (KPMG, 2023). L'IA générative offre également des possibilités inédites aux professeurs, qui peuvent l'utiliser pour les aider dans une très grande variété de tâches : élaboration des plans de cours, d'exercices personnalisés et de questions d'évaluation, tutorat intelligent pour la correction des évaluations (Baidoo-Anu & Owusu Ansah, 2023).

Cependant, bien que ces technologies aient élargi le champ des possibles tant pour les étudiants que pour leurs professeurs, leur utilisation croissante en contexte académique soulève des questions complexes en matière d'éthique et d'intégrité académique. Deux tendances se manifestent aujourd'hui. Certains enseignants, parfois des établissements, choisissent d'interdire l'usage de l'IA générative, assimilant son utilisation au plagiat. Mais cette interdiction est difficile à mettre en œuvre de manière rigoureuse (Bretag & Mahmud, 2009). Plusieurs établissements éducatifs préfèrent encourager une culture de l'utilisation responsable plutôt que d'imposer des interdictions unilatérales.

La thèse que nous défendons est que pour favoriser l'utilisation éthique, responsable et intègre de l'IA générative dans l'enseignement supérieur, il faut préférer une approche préventive large à une approche répressive étroite. Plus précisément, nous pensons qu'avec la démocratisation des outils d'IA générative, le développement d'une culture de l'intégrité académique renouvelée doit nécessairement s'appuyer sur le développement d'une compétence éthique en matière d'IA. Cette compétence concerne l'ensemble de la communauté académique et passe par une formation adéquate. Il faut donc penser les modalités et les composantes d'une telle formation.

C'est dans cet objectif que nous avons conçu une « trousse pédagogique » pour l'enseignement de l'éthique de l'IA (Bruneault *et al.*, 2024). Cette trousse, qui s'adresse aux étudiants du premier cycle universitaire, a notamment fait l'objet d'expérimentations dans le cadre de journées de la citoyenneté numérique et dans des cours complémentaires dans les collèges québécois. Elle peut servir dans des contextes disciplinaires variés et dans des cadres tant académiques qu'extra-curriculaires. Les activités peuvent être combinées ou utilisées de manière indépendante. Partant du cadre théorique développé dans *Comment former à l'éthique de l'IA en enseignement supérieur. Référentiel de compétence* (Bruneault *et al.*, 2022), la trousse s'inscrit dans l'approche de l'éthique développée par Dewey (Pappas, 2008) et propose de développer la

compétence en éthique de l'IA dans une perspective pragmatiste (Aiguier *et al.*, 2012) par le biais d'approches telles que l'apprentissage contextuel, l'expérimentation active, l'apprentissage réflexif, le jeu de rôle et l'apprentissage collaboratif. Nous exposerons dans ce texte des moyens pédagogiques concrets qui permettent de développer la compétence en éthique de l'IA dans l'enseignement supérieur.

2. LES FONDEMENTS DE NOTRE PROPOSITION

L'approche pragmatiste de la formation à l'éthique part du constat suivant : les approches purement cognitives (fondées sur la connaissance), comportementales (centrées sur l'application de règles) ou fonctionnelles (basées sur la mise en œuvre d'obligations professionnelles) sont insuffisantes pour rendre compte de la complexité des dilemmes éthiques auxquels nous sommes confrontés en pratique (Boudreau, 2019). Dans cette optique, une approche dynamique, réflexive et contextuelle de la compétence éthique est préférable à une approche théorique, abstraite et statique. Cette approche reconnaît la nécessité d'intégrer et d'articuler ces composantes ; elle permet aussi de synthétiser les éléments cognitifs, comportementaux et fonctionnels pour appréhender les situations éthiques de manière plus globale.

Notre approche du développement de la compétence en éthique de l'IA en vue de la promotion d'une culture de l'intégrité académique à l'ère des outils d'IA générative s'inspire grandement du travail de Lacroix *et al.* (2017), qui identifient trois composantes essentielles de la compétence éthique, qui sont interconnectées et nécessaires pour agir de manière appropriée dans des situations éthiques complexes : la sensibilité éthique, les capacités réflexives et les capacités dialogiques.

- **La sensibilité éthique** désigne la capacité de l'étudiant à identifier et à percevoir une situation éthique. Il s'agit d'une sensibilité particulière qui permet de détecter un déséquilibre ou une rupture dans une situation donnée. Par exemple, dans un contexte

professionnel, il ne suffit pas de suivre et d'appliquer aveuglément les règles, mais il faut également percevoir des tensions éthiques qui pourraient ne pas être évidentes à première vue. Cette sensibilité éthique, souvent sous-estimée, est cruciale pour agir de manière éthique.

- **Les capacités réflexives** impliquent de savoir réfléchir sur les situations éthiques et de mobiliser des ressources internes (comme la connaissance de soi, la maîtrise des concepts éthiques) et externes (comme les avis d'experts ou les règles en vigueur) pour analyser et comprendre une situation problématique. Il s'agit d'une forme d'intelligence pratique, qui ne se contente pas d'appliquer des cadres théoriques, mais permet de trouver des solutions adaptées à des contextes spécifiques, à condition de faire preuve d'imagination morale.

- **Les capacités dialogiques** mettent l'accent sur l'importance de l'interaction avec les autres pour résoudre des situations éthiques. Elles ne concernent pas seulement l'écoute des autres, mais aussi la capacité à participer activement à un dialogue constructif autour des enjeux éthiques. Cela implique d'être conscient des significations sociales partagées et des normes qui influencent les relations humaines, tout en ayant la volonté de répondre de ses propres actions devant les autres.

Nous avons identifié des champs de compétence spécifiques en éthique de l'IA (Bruneault *et al.*, 2022) pour structurer cette réflexion.

- Avant de développer la **compétence éthique** dans le domaine de l'IA, il faut avoir une compréhension minimale des processus techniques qui sous-tendent les systèmes d'IA. La personne compétente doit être en mesure de suivre l'évolution rapide de la technologie pour anticiper les applications et les impacts potentiels des nouveaux outils.

- Les systèmes d'IA créent **des dilemmes moraux complexes**, comme les conflits entre autonomie et bien-être, ou entre confidentialité et transparence. La personne compétente en éthique de l'IA doit non seulement être en mesure de mobiliser les théories

morales classiques dans le traitement de ces conflits, mais aussi être capable d'appliquer ces cadres à des situations nouvelles ou inédites.

- Il est aussi indispensable de situer les enjeux éthiques de l'IA dans un **contexte socio-économique** plus large. L'IA, comme d'autres technologies, peut exacerber des inégalités existantes, notamment par la discrimination algorithmique ou le renforcement des préjugés sociaux.

Nous proposons dans les pages qui suivent une synthèse des activités qui pourraient être intégrées dans une formation des étudiants sur l'utilisation responsable de l'IA à l'ère de la démocratisation des outils d'IA générative. Ces activités permettent de développer les trois composantes de la compétence éthique, à savoir la sensibilité éthique, les capacités réflexives et les capacités dialogiques. On propose aux participants de les déployer notamment dans des jeux de rôle éthiques sous la forme de scénarios fictifs : ils incarnent différentes parties prenantes lors d'une négociation autour de l'utilisation de l'IA ou doivent prendre une décision qui suppose des choix éthiques.

3. ACTIVITÉS PÉDAGOGIQUES

Dans *Former à l'éthique de l'IA en enseignement supérieur : trousse pédagogique* (Bruneault et al., 2024), nous proposons 10 activités pédagogiques qui visent le développement de la compétence en éthique de l'IA essentielle au développement des valeurs associées à l'intégrité académique. Nous présenterons ici brièvement chacune de ces activités.

3.1. Activité « Qui se cache derrière ce fil d'actualité ? »

Cette activité aide les participants à mieux comprendre le fonctionnement des algorithmes de recommandation utilisés par les plateformes de médias sociaux, comme YouTube, TikTok ou Facebook. Ces systèmes, en analysant les interactions des utilisateurs, filtrent et sélectionnent les contenus qu'ils jugent les plus

pertinents pour maximiser leur engagement (Tolentino, 2019). L'objectif est d'analyser de manière critique ces recommandations et leur impact (Seaver, 2019).

Les participants visionnent des enregistrements vidéo qui montrent les suggestions de contenus proposées à un « profil mystère ». Ils doivent alors, sur la base des contenus recommandés et des comportements observés (vidéos liées, commentaires, etc.), deviner les caractéristiques du profil utilisateur (âge, sexe, intérêts, etc.). Après avoir émis leurs hypothèses, ils les comparent avec les informations du profil réel.

Cette analyse permet de comprendre la manière dont les plateformes créent des profils détaillés à partir de simples interactions en ligne, qu'elles soient actives (commentaires, partages) ou passives (temps passé à visionner une vidéo). Une fois les résultats comparés, une discussion s'engage sur les questions éthiques liées à l'utilisation des données personnelles dans la construction de ces profils ainsi que sur les biais qui peuvent influencer les recommandations.

Le débat porte également sur l'effet des bulles de filtres, ces phénomènes qui limitent l'exposition des utilisateurs à des points de vue divers et renforcent ainsi leurs croyances préexistantes (Pariser, 2011). Cette activité vise à sensibiliser à l'impact des algorithmes sur la consommation d'information et sur la manipulation subtile des comportements en ligne (Cardon, 2015).

3.2. Activité « Profiler pour vendre : les dessous de la publicité ciblée »

Cet atelier immerge les participants dans le monde de la publicité ciblée, où des algorithmes de profilage sont utilisés pour personnaliser les annonces publicitaires en fonction des comportements en ligne des utilisateurs (Zuboff, 2020). L'activité commence par une observation des publicités affichées sur les réseaux sociaux des participants. Ils sont invités à noter les produits

ou services promus, la marque et l'adéquation de la publicité avec leurs intérêts ou leurs activités récentes.

Après cette observation, les participants se regroupent pour analyser les liens entre leur comportement en ligne et les publicités qui leur sont proposées. Cette analyse amène à questionner la transparence des processus de profilage et à réfléchir aux stratégies publicitaires déployées pour influencer les comportements de consommation (Durand-Folco & Martineau, 2023 ; Piguët, 2023).

Les discussions permettent aussi d'aborder les risques liés aux biais algorithmiques dans les publicités ciblées, par exemple dans des domaines comme l'emploi ou l'accès au logement. En fin d'atelier, les participants proposent des solutions pour rendre plus transparent le processus de publicité ciblée, notamment en améliorant la compréhension des utilisateurs sur la manière dont leurs données sont utilisées.

3.3. Activité « Sécurité informatique 101 : comment déjouer les escrocs du cyberspace ? »

Cet atelier initie les participants aux bases de la cybersécurité, un enjeu crucial à l'ère de l'IA et de la numérisation généralisée (Billois, 2022). L'objectif est d'aider les participants à identifier les risques associés à l'utilisation des outils numériques et à adopter des pratiques responsables et sécuritaires.

L'activité commence par une réflexion sur les données personnelles partagées en ligne et la vulnérabilité de ces informations face aux cyberattaques. Les participants sont ensuite invités à créer des mots de passe pour des comptes fictifs, puis à les tester pour évaluer leur robustesse. Cette partie de l'atelier aborde les règles de création de mots de passe sécurisés et l'importance des gestionnaires de mots de passe.

La session se poursuit par une introduction aux tentatives de *phishing* (hameçonnage), avec des exemples concrets de messages frauduleux qui cherchent à obtenir des informations confidentielles (Gouvernement du Canada, 2022). Les participants apprennent à

repérer les indices de *phishing* (adresse e-mail suspecte, fautes de grammaire, demandes urgentes) et à appliquer des stratégies pour éviter d'être piégés.

Enfin, les participants discutent de l'utilisation de l'IA pour amplifier les cyberattaques (Begou *et al.*, 2023). Par exemple, l'IA peut générer des *deepfakes* (hypertrucages) utilisés pour des arnaques sophistiquées. La discussion s'étend à la responsabilité des individus et des entreprises dans la mise en place de mesures de cybersécurité adéquates, tant sur le plan personnel que professionnel.

3.4. Activité « Études de cas en éthique de l'IA : quelles pistes d'action pour mitiger les risques ? »

Cette activité permet aux participants de plonger dans des scénarios concrets où des systèmes d'IA sont utilisés. Chaque groupe de participants reçoit une étude de cas centrée sur un domaine particulier, comme la santé, l'éducation ou le recrutement (Stahl *et al.*, 2023).

Une première étape consiste à analyser les avantages et les risques pour les parties prenantes. Les participants doivent lister les acteurs directement ou indirectement impliqués (utilisateurs, entreprises, communautés) et anticiper leurs intérêts et préoccupations.

Ensuite, les groupes procèdent à une évaluation des risques et bénéfices : quels sont les effets positifs de l'IA dans ce contexte ? Quels dangers peut-elle poser, notamment en matière de biais, de confidentialité ou d'équité ? Une fois cette analyse faite, les groupes formulent des recommandations concrètes pour atténuer ces risques, en se basant sur des principes éthiques comme la justice, la transparence et le respect de la vie privée.

3.5. Activité « Déclaration de Montréal pour un développement responsable de l'IA »

L'objectif de cette activité est de familiariser les participants avec la *Déclaration de Montréal pour un développement responsable de l'IA* (Abrassart *et al.*, 2018) et la grille de réflexivité éthique développée par l'OBVIA pour analyser les systèmes d'IA (Marchildon *et al.*, 2022). En utilisant ces outils, les participants doivent évaluer un système d'IA fictif ou réel en s'appuyant sur les dix principes de la *Déclaration*, tels que l'autonomie, la protection de la vie privée, la transparence ou l'équité.

Les participants commencent par une présentation des principes de la *Déclaration de Montréal*. Ces principes sont ensuite appliqués à un cas d'IA dans un domaine spécifique. Chaque groupe est chargé d'identifier les enjeux éthiques spécifiques à ce système d'IA, en répondant à une série de sous-questions issues de la grille de réflexivité.

Cette activité permet d'anticiper les impacts sociaux, moraux et environnementaux des systèmes d'IA (Mittelstadt *et al.*, 2016), en aidant les participants à proposer des solutions concrètes pour minimiser les risques (Morley *et al.*, 2019).

3.6. Activité « Algorithme de recrutement biaisé : quand l'IA fait de la discrimination »

Dans cet atelier, les participants explorent comment les algorithmes de recrutement peuvent présenter des biais (Gentelet & Mathieu, 2021). Ils sont invités à analyser le cas fictif d'une entreprise qui utilise un algorithme pour présélectionner des candidats à un poste d'ingénieur informatique. Leur mission est de comparer leurs propres choix de candidats à ceux que sélectionne l'algorithme.

L'activité montre comment les algorithmes entraînés à partir de données historiques peuvent perpétuer des inégalités. Par exemple, si les données d'entraînement reflètent des biais de genre ou d'origine ethnique, l'algorithme risque de privilégier certains groupes démographiques au détriment d'autres.

La discussion porte sur les moyens de détecter et corriger ces biais dans les systèmes automatisés ainsi que sur les conséquences de ces biais en termes de discrimination à l'embauche (Black & van Esch, 2020). Les participants sont également sensibilisés aux questions de transparence, cruciales pour garantir l'équité des processus de recrutement et prévenir la discrimination systémique (Lacroux & Martin-Lacroux, 2021).

3.7. Activité « L'IA créatrice, vraiment ? »

Cette activité invite les participants à examiner la capacité créative de l'IA (Wu, 2023). En comparant des œuvres créées par des IA (textes, images, vidéos) à des créations humaines, ils sont encouragés à réfléchir à la différence entre la créativité humaine et les processus génératifs automatisés.

Les participants utilisent des outils d'IA générative pour réaliser eux-mêmes des créations textuelles ou visuelles. Cette expérience pratique permet de mettre en lumière les possibilités et les limites des systèmes IA : par exemple, les résultats générés ne possèdent pas la profondeur émotionnelle ou l'intention artistique qui caractérisent les œuvres humaines.

La discussion porte aussi sur les enjeux éthiques et juridiques de l'IA générative, notamment en matière de droits d'auteur (Azzaria, 2023). Les œuvres générées par l'IA sont-elles réellement originales ou simplement des amalgames de données existantes ? Quelles sont les implications pour les artistes humains dont les œuvres sont utilisées pour entraîner ces systèmes (Chayka, 2023) ?

3.8. Activité « L'IA, plus matérielle qu'on ne le pense et moins verte qu'on ne le dit »

Cet atelier sensibilise les participants à l'impact environnemental des technologies d'IA, souvent perçues à tort comme immatérielles. Les participants découvrent que le développement et le fonctionnement des systèmes d'IA nécessitent d'importantes ressources matérielles, telles que des métaux rares (lithium, cobalt)

et de grandes quantités d'énergie pour alimenter les centres de données (Ferreboeuf *et al.*, 2021).

L'activité consiste à cartographier la chaîne de production des systèmes d'IA, de l'extraction des matières premières à la gestion des déchets électroniques, en passant par l'assemblage des composants et le fonctionnement des infrastructures technologiques. Ce processus permet de mieux comprendre les enjeux écologiques et géopolitiques liés à l'industrie de l'IA (Pitron, 2021). Plus spécifiquement, les participants sont invités à placer des cartes ressources sur une carte du monde.

Les participants discutent ensuite des solutions pour minimiser l'empreinte environnementale des technologies d'IA (Cowls *et al.*, 2023) et ils réfléchissent également à la responsabilité des utilisateurs dans cette transition écologique ainsi qu'à l'impact du consumérisme technologique sur l'environnement.

3.9. Activité « La chasse aux illusions : comment détecter les fausses nouvelles et les hypertrucages »

Les fausses nouvelles et les *deepfakes* (hypertrucages) sont des phénomènes amplifiés par les récents progrès en IA générative. Dans cet atelier, les participants apprennent à identifier ces contenus trompeurs en utilisant un guide de détection critique. Ils analysent des images, vidéos ou articles publiés en ligne pour déterminer lesquels sont authentiques et lesquels relèvent de la désinformation.

L'activité aide à développer une pensée critique face aux contenus numériques et à comprendre comment l'IA peut être utilisée pour manipuler l'opinion publique ou semer la confusion (Naffi *et al.*, 2022). Les *deepfakes* posent des problèmes éthiques graves, notamment en matière de consentement et de réputation.

Après l'analyse, les participants discutent des conséquences sociales et politiques de la désinformation, notamment l'impact sur la démocratie et les processus électoraux (CEST, 2023). Ils réfléchissent également aux moyens de contrer la propagation des

fausses nouvelles, en renforçant la réglementation et en responsabilisant les plateformes numériques.

3.10. Activité « Déconnexion créative : dessiner pour réfléchir aux automatismes numériques »

Cet atelier invite les participants à réfléchir à leur utilisation compulsive des téléphones intelligents et des gestes automatiques qui y sont associés, comme le *scroll* ou le *pull-to-refresh*. Ces petits gestes sont au cœur des stratégies des plateformes numériques pour capter et retenir l'attention des utilisateurs (Citton, 2014).

Au moyen de deux exercices de dessin (une ligne et un gribouillis), les participants explorent leur état intérieur et comparent cette expérience à l'usage quotidien du téléphone intelligent. L'objectif est de prendre conscience de la manière dont les gestes numériques répétitifs influencent leur comportement et leurs émotions (Mondoux *et al.*, 2014).

Le dessin de la ligne, réalisé les yeux fermés, incite les participants à se recentrer sur leur état émotionnel et à s'interroger sur l'intention derrière chaque geste. Le gribouillage, quant à lui, sert de point de départ à une réflexion sur la créativité et la communication interpersonnelle, contrastant avec les gestes répétitifs et automatisés induits par les technologies numériques.

Cette activité permet d'aborder l'impact des plateformes numériques sur le bien-être, en incitant les participants à réfléchir à la manière dont la technologie influe sur leur relation à eux-mêmes et aux autres, ainsi qu'aux moyens de réduire la dépendance aux contenus proposés par des algorithmes (Marcus & Davis, 2019).

4. BILAN ET DISCUSSION

Notre démarche, tributaire de l'expertise des membres de l'équipe ayant participé à l'élaboration de la trousse pédagogique, est également liée aux particularités des champs d'expertise que réunit cette équipe. Nous en avons validé la pertinence en la présentant

lors de conférences multidisciplinaires et internationales, comme celle de l'IRAFPA en juin 2024. L'accueil positif fait à notre travail et les débats suscités montrent que notre proposition pragmatique peut fournir un tremplin pour aborder les nombreuses questions éthiques soulevées par l'IA. Les débats ont porté notamment sur sa contribution à l'insertion de ces questions à différents niveaux des cursus d'enseignement supérieur.

Développons ces propositions d'action.

Une approche pragmatiste en enseignement de l'éthique (Lacroix *et al.*, 2017) constitue en effet une avenue particulièrement prometteuse pour le développement de la capacité à utiliser les outils d'IA générative de manière responsable en enseignement supérieur. Selon nous, une telle approche, inspirée du cadre conceptuel établi par John Dewey (1920/2004), permet de penser la formation à l'éthique de l'IA comme un exercice d'enquête autonome, réflexive et contextuelle. Cette formation vise à outiller les apprenants pour qu'ils puissent interagir de manière critique avec les outils d'IA générative.

Il paraît aujourd'hui nécessaire non seulement de mettre en place les conditions institutionnelles et pédagogiques favorisant l'intégrité intellectuelle et la prévention du plagiat (Eaton, 2021), mais aussi de développer la compétence en éthique de l'IA chez les membres de la communauté académique. Il importe en effet que tous soient formés à une utilisation éthique et responsable de ces outils, tout en étant informés de leurs possibilités et de leurs limites.

Les activités proposées dans notre trousse pédagogique visent à renforcer la compréhension des enjeux éthiques de l'IA, compréhension qui nous semble nécessaire pour favoriser le renouvellement d'une culture de l'intégrité académique. En mettant l'accent sur la réflexion critique, la sensibilité éthique et les capacités dialogiques, ces approches permettent d'intégrer l'IA de manière responsable dans les pratiques académiques.

Nous pensons que ces activités pourraient être intégrées dans une formation à l'éthique de l'IA en contexte académique et qu'elles sont

susceptibles de favoriser l'utilisation des outils d'IA générative de manière responsable et intégrée, tout en tenant compte des défis éthiques, sociaux, politiques, juridiques et environnementaux que posent ces technologies. Une approche proactive et préventive plus large est également nécessaire en raison des limites des approches prohibitives dues aux difficultés de détection de la fraude liée à l'IA générative.

Bien qu'il soit absolument nécessaire de se doter de règles claires dans les institutions académiques pour l'utilisation de l'IA, ces règles seules ne peuvent anticiper toutes les situations possibles. L'approche éthique vient en complément des règles. Il faut donc que la compétence en éthique de l'IA soit elle aussi développée pour que la communauté académique puisse agir de manière responsable dans leur utilisation de l'IA.

Références bibliographiques

Abrassart, C., Bengio, Y., Chicoisne, G., de Marcellis-Warin, N., Dilhac, M-A., Gambs, S., Gautrais, V., Gibert, M., Langlois, L., Laviolette, F., Lehoux, P., Maclure, J., Martel, M., Pineau, J., Railton, P., Régis, C., Tappolet, C. & Voarino, N. (2018). *La déclaration de Montréal pour un développement responsable de l'intelligence artificielle*. Université de Montréal. <https://declarationmontreal-iaresponsable.com/>

Aiguier, G., Le Berre G., Vanpee, R. & Cobbaut, J.-P. (2012). Du pragmatique au pragmatisme : quels enjeux pour la formation à l'éthique ? *Journal international de bioéthique*, 23(4), 123-148.

Azzaria, G. (2023). Droit d'auteur et intelligence artificielle : un aperçu canadien. Dans C. Castets-Renard & J. Eynard (dirs.), *Un droit de l'intelligence artificielle : entre règles sectorielles et régime général. Perspectives comparées* (pp. 553-587). Éditions Bruylant.

Baidoo-Anu, D. & Owusu Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4337484>

Begou, N., Vinoy, J., Duda, A. & Korczynski, M. (2023). *Exploring the dark side of AI: Advanced phishing attack design and deployment using Chat GPT*. Consulté en ligne le 28 février 2023 sur <https://arxiv.org/abs/2309.10463>

Billois, G. (2022). *Cyber-attaques : les dessous d'une menace mondiale*. Hachette.

Black, J. S. & van Esch, P. (2020). AI-enabled recruiting: What it is and how should manager use it? *Business Horizons*, 63(2), 215-226.

Boudreau, M.-C. (2019). *La compétence éthique en milieu de travail : Une perspective pragmatiste pour sa conceptualisation et son opérationnalisation*. Thèse de doctorat,

Université de Sherbrooke.

Bretag, T. & Mahmud, S. (2009). A model for determining student plagiarism: Electronic detection and academic judgment. *Journal of University Teaching & Learning Practice*, 6(1), 57-69. <https://doi.org/10.53761/1.6.1.6>

Bruneault, F., Sabourin Laflamme, A., Boivin, B., Grondin-Robillard, L. & Le Calvez, E. (2024). *Former à l'éthique de l'IA en enseignement supérieur. Trousse pédagogique*. <https://osf.io/preprints/socarxiv/z8m42>

Bruneault, F., Sabourin Laflamme, A. & Mondoux, A. (2022). *Former à l'éthique de l'IA en enseignement supérieur : Référentiel de compétence*. <https://doi.org/10.31235/osf.io/38tfv>

Bussièrès McNicoll, F. (2023). IA à l'université : le passage de la réflexion à l'action tarde. *Ici Radio-Canada*. <https://ici.radio-canada.ca/nouvelle/2061040/intelligence-artificielle-ia-universite-plagiat-etudiantes>

Cardon, D. (2015). *À quoi rêvent les algorithmes ? Nos vies à l'heure des big data*. Seuil.

Chayka, K. (2023, february 10). Is AI art stealing from artists? *The New Yorker*. <https://www.newyorker.com/culture/infinite-scroll/is-ai-art-stealing-from-artists>

Citton, Y. (2014). *L'économie de l'attention : Nouvel horizon du capitalisme ?* La Découverte.

Commission de l'éthique en science et en technologie (CEST) (2023). *Désinformation, démocratie et intelligence artificielle*. <https://www.ethique.gouv.qc.ca/fr/actualites/ethique-hebdo/desinformation-democratie-et-intelligence-artificielle/>

Cowls, J., Tsamados, A., Taddeo, M. & Floridi, L. (2023). The AI gambit: leveraging artificial intelligence to combat climate change. Opportunities, challenges and recommendations. *AI & Society*, 38, 283-307. <https://doi.org/10.1007/s00146-021-01294-x>

Dewey, J. (1920/2004). *Reconstruction in philosophy*. Holt and Co.

Durand-Folco, J. & Martineau, J. (2024). *Le capital algorithmique*. Écosociété.

Eaton, S. E. (2021). *Plagiarism and Higher Education*. Tackling Tough Topics in Academic Integrity, Bloomsbury Publishing.

Farrelly, T. & Baker, N. (2023). Generative artificial intelligence: Implications and considerations for higher education Practice. *Education Sciences*, 13(11), 1109. <https://doi.org/10.3390/educsci13111109>

Ferreboeuf, H., Efoui-Hess, M. & Verne, X. (2021). *Environmental impacts of digital technology: 5-year trends and 5G governance*. The Shift project. <https://theshiftproject.org/en/article/environmental-impacts-of-digital-technology-5-year-trends-and-5g-governance/>

Gouvernement du Canada (2022). *Ne mordez pas à l'hameçon. Reconnaître et prévenir les attaques par hameçonnage*. Consulté le 28 février 2023 sur <https://www.cyber.gc.ca/fr/orientation/ne-mordez-pas-lhamecon-reconnaître-et-prévenir-les-attaques-par-hameconnage>.

KPMG Canada (2023). Répertoire sur l'adoption de l'IA générative. <https://assets.kpmg.com/content/dam/kpmg/ca/pdf/2023/06/kpmg-in-canadas-generative->

[ai-adoption-index-2023-fr.pdf](#)

Lacroix A., Marchildon, A. & Bégin, L. (2017). *Former à l'éthique en organisation*. Presses de l'Université du Québec.

Lacroux, A. & Martin-Lacroux, C. (2021). L'intelligence artificielle au service de la lutte contre les discriminations dans le recrutement : nouvelles promesses et nouveaux risques. *Management & Avenir*, 122(2), 121-142.

Marchildon, A., Jacob, S. & Marcoux, M. A. (2021). *Grille de réflexivité portant sur les enjeux éthiques des systèmes d'intelligence artificielle*, Observatoire international sur les impacts sociaux de l'IA et du numérique (OBVIA). <https://observatoire-ia.ulaval.ca/grille-de-reflexivite-portant-sur-les-enjeux-ethiques-des-systemes-dintelligence-artificielle-sia/>

Marcus, G. & Davis, E. (2019). *Rebooting AI: building artificial intelligence we can trust*. Pantheon Books.

Mittelstadt, B. R., Allo, P., Taddeo, M., Wachter, S. & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 1-21. <https://doi.org/10.1177/2053951716679679>

Mondoux, A., Ménard, M. & Bonenfant, M. (2014). Quand le Pathos devient Ethos. Esquisse de la dépendance psychosociale contemporaine. *Drogues, santé et société*, 13(1), 1-18.

Morley, J., Floridi, L. Kinsey, L. & Elhalal, A. (2019). From what to how: An initial review of publicly available AI ethics tools, methods and research to translate principles into practices. *Science and Engineering Ethics*. <https://doi.org/10.1007/s11948-019-00165-5>

Naffi, N., Davidson, A.-L., Barma, S., Bernard, M.-C., Brault, N., Berger, F. & Gagnon-Tremblay, A. (2022). Pour une éducation aux hypertrucages malveillants et un développement de l'agentivité dans les contextes numériques. *Érudit*, 49(2). <https://www.erudit.org/fr/revues/ef/2021-v49-n2-ef06680/1085307ar/>

Pappas, G. F. (2008). *John Dewey's Ethics. Democracy as experience*. Indiana University Press.

Pariser, E. (2011). *The filter bubble: What the internet is hiding from you*. Penguin Press.

Piguet, J.-G. (2023). Nudge, désinformation et autonomie citoyenne. Une critique de Sunstein. *Éthique Publique*, 24(2). <http://journals.openedition.org/ethiquepublique/7583> <https://doi.org/10.4000/ethiquepublique.7583>

Pitron, G. (2021). *L'enfer numérique. Voyage au bout d'un like*. Les Liens qui Libèrent.

Sadasivan, V. S., Kumar, A., Balasubramanian, S., Wang, W. & Feizi, S. (2023). Can AI-Generated text be reliably detected? *arXiv preprint arXiv:2303.11156*. <https://arxiv.org/abs/2303.11156>

Seaver, N. (2019). Captivating algorithms : Recommender systems as traps. *Journal of Material Culture*, 24(4), 1-16. <https://doi.org/10.1177/1359183518820366>

Stahl, B. C., Schroeder, D. & Rodrigues, R. (2023). *Ethics of artificial intelligence: Case studies and options for addressing ethical challenges*. Springer Briefs in Research and Innovation Governance. Springer International Publishing. <https://doi.org/10.1007/978-3-031-17040-9>

Tolentino, J. (2019, 30 septembre). How TikTok holds our attention. *The New Yorker Magazine*. Brave New World Dept.
<https://www.newyorker.com/magazine/2019/09/30/how-tiktok-holds-our-attention>

United Nations Educational, Scientific and Cultural Organization (UNESCO) (2023). UNESCO survey: Less than 10% of schools and universities have formal guidance on AI.
<https://www.unesco.org/en/articles/unesco-survey-less-10-schools-and-universities-have-formal-guidance-ai>

Wu, Y. (2023, January 17). AI is outrageous – and wonderful. It's also prompting a new art form. *Washington Post*. <https://www.washingtonpost.com/opinions/interactive/2024/ai-image-generation-art-innovation-issue/>

Zuboff, S. (2020). *L'âge du capitalisme de surveillance*. Zulma.

CHAPITRE 4. L'UTILISATION DE L'IA GÉNÉRATIVE POUR LA RÉDACTION ACADÉMIQUE EN FRANÇAIS PAR LES ÉTUDIANTS ALLOPHONES. L'EXPÉRIENCE DU CPU LYON

[Yves Frédéric Livian](#), Professeur émérite, Université Lyon 3

[Robert Laurini](#), Professeur émérite, INSA de Lyon, Ancien président d'Universitaires sans Frontières

1. INTRODUCTION

L'introduction de l'IA dans les pratiques éducatives n'est pas nouvelle et a fait l'objet de nombreuses publications spécialisées. Mais la large diffusion de ChatGPT en 2022 et des nombreux modèles d'IA générative ont ajouté une dimension nouvelle aux préoccupations des enseignants comme aux ressources des étudiants¹. Plus largement, les questions éthiques soulevées par l'utilisation de l'IA en français remplissent les pages de magazines et de blogs. L'émotion créée est grande et les réactions pas toujours en rapport avec les usages réels et les limites des outils disponibles actuellement.

À l'intérieur de ce vaste champ, nous voudrions traiter un aspect particulier, qui touche de nombreux enseignants en général, et les enseignants de français langue étrangère (FLE) en particulier, celui des effets de l'usage de l'IA dans la rédaction et la correction de textes académiques par des étudiants allophones. Nous sommes partis du constat de la large disponibilité (gratuite pour partie) de ces outils pour les étudiants (dès l'enseignement secondaire), de leur diffusion irrépressible et des progrès constants sur le plan des données traitées et des formes d'accessibilité (évolution de ChatGPT4 par rapport à la version 3 par exemple, articulation de ChatGPT4 avec Bing pour Copilot). Les questions soulevées par la rédaction assistée par IA sont nombreuses : qui est réellement

l'auteur du texte généré ? Quelles sources utilise-t-il ? Quel rôle ce texte généré peut-il jouer pour rédiger un mémoire, voire une thèse ? Quelle attitude doit avoir l'enseignant (ou l'accompagnateur) face à cet usage ? La tentation d'utiliser ces outils sans discernement est-elle plus grande chez les étudiants allophones ? Comment les accompagner dans leur cheminement ?

Nous avons observé la rencontre des étudiants allophones avec l'IA au sein d'une association² créée en 2011 et localisée à Lyon. Elle a pour objet d'accueillir et d'aider les étudiants internationaux inscrits dans un établissement à Lyon. Moyennant une adhésion modique, l'étudiant est suivi individuellement par un accompagnateur bénévole³ qui l'aide dans ses études et l'amélioration de son français. Cet accompagnement comporte des groupes de conversation, des ateliers d'expression écrite, de grammaire et de méthodologie (Livian & Laurini, 2018). Pour ceux qui sont en master ou thèse, l'accompagnateur travaille avec l'étudiant pour améliorer le texte qu'il rédige, typiquement une fois par semaine ; l'étudiant le mentionnera dans son mémoire. Les demandeurs d'asile de niveau universitaire, eux suivent des cours de français dispensés par des spécialistes de FLE. En 2023, il y avait 310 étudiants inscrits à l'association, 142 bénévoles et 2 salariés. La majorité des étudiants (hors demandeurs d'asile) sont inscrits dans des écoles ou des facultés de lettres, sciences humaines, droit et gestion. Les pays représentés sont principalement l'Asie du Sud-Est (37 %), le Moyen-Orient (16 %), l'Afrique (14 %), l'Europe (16 %) et l'Amérique latine (14 %).

Notre principale question de recherche est de savoir si l'IA générative ne constitue pas un effet d'aubaine pour les étudiants internationaux⁴ ?

Sans chercher à répondre à toutes les questions posées par l'arrivée impromptue de l'IA, nous avons interrogé nos collègues accompagnateurs et quelques étudiants du CPU sur les usages de l'IA dans la préparation de leurs écrits en contexte universitaire.

2. REVUE DE LITTÉRATURE SUR L'IA PAR L'IA

Notre première question de recherche a porté sur la possibilité de réaliser des revues de littérature. La requête (formulée en anglais) a été de chercher des références sur « l'usage de l'IA par les étudiants de l'enseignement supérieur » datant de moins de 5 ans.

Nous avons réalisé un exercice d'utilisation de l'IA pour la recherche documentaire en nous mettant au diapason des étudiants que nous interrogeons. Nous avons déjà repéré des références par nos sources habituelles (lectures, Google Scholar, Cairn, etc.) et avons exploré les apports de sept systèmes d'IA à finalité documentaire dans leur version gratuite (donc facilement accessible aux étudiants).

- **Consensus** nous a fourni une quinzaine de références, toutes pertinentes, avec le nombre de citations de chaque article : une liste limitée, mais Consensus AI prévient qu'il ne délivre pas *a final truth but a reflection of relevant research*. Il y a donc une certaine sélection, aux critères peu clairs.

- **Perplexity** (*where knowledge begins*) va plus loin, puisqu'il fournit une synthèse de la littérature trouvée en une quinzaine de lignes. Parmi les articles proposés, il en a trouvé cinq en français : le texte est pertinent, mais très condensé.

- **Elicit** sélectionne quatre références *top* dont il fait un résumé, mais on aurait plus, dit-il, grâce à la version Premium payante. Il propose une ligne de résumé sur les 55 articles trouvés (tous en anglais, même quand le prompt est rédigé en français).

- **Scispace** (*Discover and analyze research with no limits*) trouve 120 articles, propose une reformulation des prompts pour relancer la recherche. Il fournit les PDF quand ils sont disponibles⁵. Le résultat est donc abondant et utile, mais la pertinence des références est à vérifier.

- **You.com** n'est pas un outil documentaire scientifique et est plutôt complémentaire. Il propose un chatbot permettant de questionner sur un texte (ou une vidéo) dont on fournit l'URL. Sa réponse est

habituellement générale et assez « plate » (sans être fausse) et il fournit plusieurs sources à sa réponse. Il reformule aussi la demande en la transformant en sujet sur lequel il peut répondre.

- **Paper Digest** fournit de courtes revues de littérature (avec des filtres par disciplines, types de document recherché et dates). Sur l'usage de l'IA par les étudiants (qui était notre prompt), il propose une dizaine de références comprenant aussi des documents qui ne sont pas des articles scientifiques. Relancé sur le sujet « problèmes éthiques de la rédaction d'essais par les étudiants » (toujours en formulé en anglais), il a fourni quelques sources utiles, notamment le rapport sur l'IA du MLA-CCC Joint Task Force (2023), un organisme regroupant des enseignants nord-américains de communication, et le livret de recommandations de l'Université Carleton (Canada).

- Le prompt *Chat GPT and plagiarism* donne plus d'une centaine de réponses, mais dont beaucoup sont hors sujet.

- Enfin, Semantic Scholar (basé sur plus de 217 millions d'articles) se voulant « *a free, AI – powered research tool for scientific literature* » a eu du mal à répondre au même prompt. Il a trouvé en revanche sur *writing assignment by AI* 98 articles, pour lesquels il donne parfois le super-résumé (appelé TLDR, *too long, don't read*, surtout appliqué aux sciences médicales et à l'informatique). Pour chaque article, il propose un résumé et suggère trois questions-types (buts, méthodes, principaux résultats). Il signale aussi les PDF disponibles. Un survol des résumés donne une vingtaine d'articles pertinents pour le sujet.

En résumé, il apparaît que les modèles utilisés pour la recherche documentaire sont fonctionnels. Ils sont la continuation des bases de données déjà existantes, mais ils brassent d'énormes quantités de textes (125 millions selon l'un d'eux) et sont capables de rédiger des synthèses en répondant quasi immédiatement. La majorité des articles est écrite par des chercheurs non européens (très nombreux articles écrits par des Indiens, Indonésiens, Chinois...) ce qui montre une diffusion rapide de l'IA dans l'enseignement supérieur de ces pays.

Chacun des « outils » a une présentation différente et une certaine familiarisation est nécessaire pour en tirer parti. Il semble indispensable d'utiliser plusieurs prompts différents pour être sûr d'avoir des résultats. Il est clair également que l'utilisateur a intérêt à comparer les résultats et bien sûr à sélectionner les sources pertinentes, qu'il devra vérifier⁶. Les *literature surveys* fournis par les versions gratuites sont succincts et parfois simplistes, mais ils peuvent servir de point de départ et certains étudiants peu ambitieux les utiliseront !

Un point semble important pour l'avenir de l'utilisation de l'IA à des fins de bibliographie. La littérature récente en anglais depuis trois ans est déjà abondante, bien que les outils génératifs ne soient diffusés que depuis deux ans. La tonalité générale des articles sélectionnés par l'IA est positive, en raison sans doute des revues les plus souvent citées (revues d'informatique ou d'éducation) : l'IA peut apporter dans l'enseignement supérieur de l'efficacité, de la productivité et de l'innovation. L'utilisation de l'IA est fréquente dans tous les échantillons étudiés. Rares sont en revanche les études observant concrètement les pratiques des étudiants et abordant leur perception des pratiques illicites. Rares aussi sont les références aux règles précises qui pourraient être émises par les établissements.

Or les revues de littérature et les résumés générés avec ces outils ont leurs limites : ils sont plats et consensuels. Nous n'y avons pas trouvé d'idées tranchantes ni stimulantes, mais seulement beaucoup de descriptions. Une partie non négligeable des articles compilés cherche simplement à mesurer les avantages de l'IA dans l'enseignement : ils proposent de comparer les résultats obtenus par les apprenants avant/après le recours à des dispositifs expérimentaux et concluent généralement à l'utilité de l'IA.

Si certains articles soulignent les risques éthiques, c'est toujours en termes généraux. Les articles traitant des questions éthiques posées par l'IA dans l'enseignement supérieur étaient rares : une revue de littérature (Zawacju-Richter *et al.*, 2019, cité par Collin & Marceau 2022) donnait deux articles sur 146). Leur nombre a

cependant dû augmenter considérablement en 2023 (ce qui n'était pas notre requête).

De nombreux articles concluent à la nécessité pour les établissements d'enseignement supérieur d'être clairs sur les règles d'utilisation par les étudiants et sur l'urgence de former les enseignants à l'usage de l'IA. Il n'y a là rien de très original et cela relève du lieu commun.

3. RÉSULTATS DE L'ENQUÊTE EXPLORATOIRE

En janvier-mars 2024, nous avons procédé à une enquête auprès d'étudiants et de collègues accompagnateurs sur l'usage qu'ils font de l'IA générative⁷. Nous avons posé une série de questions à un groupe de 21 accompagnateurs et avons approfondi l'entretien avec deux d'entre eux ; nous avons également réalisé douze entretiens individuels en face à face (en français) avec des étudiants étrangers et un entretien de groupe avec sept autres étudiants. Ce matériau a fait l'objet d'une analyse thématique classique.

3.1. Le recours à l'IA : une découverte aléatoire

Parmi les étudiants interrogés, le recours à l'IA n'est pas généralisé et il est moins fréquent que dans la population totale des étudiants fréquentant le CPU. Certains, notamment les demandeurs d'asile, sont en apprentissage du français basique, suivent des cours en présentiel adaptés à leur niveau et, le plus souvent, ne possèdent pas d'ordinateur. Parmi les autres, un doctorant se déclare « réfractaire » à l'usage de l'IA et a le soutien de sa directrice de recherche, qui lui présente la thèse comme un « travail artisanal » ne nécessitant pas le recours à des outils sophistiqués. Une autre (chinoise) dit ne pas en avoir besoin, elle est considérée comme ayant un niveau de français excellent et a déjà à sa disposition les livres des auteurs sur lesquels porte sa thèse de littérature. Deux étudiants en géographie et anthropologie n'ont aucune expérience de l'IA et ne pensent pas s'en servir dans l'immédiat.

Pour les autres étudiants, les cas d'utilisation sont nombreux, mais ils paraissent y être venus par curiosité ou par hasard : « *Pendant un cours de français, il fallait recréer les dialogues d'une séquence d'un film muet projeté par le professeur... certaines étudiantes n'ont rien dit, mais ont soumis une réponse qui était bonne. Elles l'avaient demandée à ChatGPT. C'est comme cela que je l'ai découvert.* »

3.2. Traduire et écrire en bon français

Les étudiants allophones que nous avons interrogés utilisent déjà les logiciels de traduction (Reverso, DeepL, Google Translate...). L'IA ne fait que compléter les outils disponibles. Les usages d'approfondissement linguistique sont nombreux : regarder le sens du mot, en chercher la définition, relier grammaire et syntaxe, perfectionner un texte...

Ces systèmes de traduction permettent de jongler entre plusieurs langues. Par exemple, une de nos étudiantes sud-américaines doit rédiger en anglais, elle demande des reformulations en anglais, et vérifie avec une traduction en espagnol. Pour le français, elle utilise ChatGPT pour la définition de certains mots. Autre exemple, ce juriste qui rédige en arabe et en anglais et fait traduire en français par Google ou DeepL.

Mais l'utilisation est relative aux langues concernées : un doctorant en littérature arabe considère que les traductions en français sont fiables. Une étudiante africaine regrette de ne pas pouvoir utiliser le fulani (peul), sa langue maternelle.

Évidemment, avec l'émergence de l'IA, le travail de correction des accompagnateurs de CPU a changé, puisque le texte de l'étudiant, s'il a utilisé l'IA, est déjà correct du point de vue grammatical et orthographique. Mais comment s'assurer de la maîtrise de la langue par l'étudiant ?

Après une période confuse, le passage par la traduction ou la génération de texte par l'IA est aujourd'hui le plus souvent transparent dans le rapport étudiant-accompagnateur. Une vingtaine de bénévoles volontaires ont reçu une sensibilisation et une

formation à l'IA générative et sont donc à même de gérer l'usage que leurs étudiants font de l'IA.

Nous constatons alors l'importance fondamentale de la communication orale en appui à la simple production d'un écrit. Ainsi, une accompagnatrice déclare : « *La première fois, le texte d'une étudiante taïwanaise était hors sujet... Je lui demande pourquoi... Elle me répond que c'était Chat qui l'avait écrit. On refait le travail sur le fond, elle me propose un second texte, bien écrit et répondant au sujet.* »

Nous constatons aussi l'importance des notions de confiance et de respect. Dans la relation étudiant/correcteur, l'usage de l'IA peut être difficilement dissimulé, car le niveau linguistique de départ de l'étudiant est connu. Selon les accompagnateurs interrogés, pour ceux qui ont déjà un bon niveau linguistique, on repère facilement le recours à ChatGPT au caractère standardisé du texte et à son adéquation parfois maladroite au thème à traiter malgré l'absence d'incorrections purement linguistiques.

3.3. Explorer un sujet et produire un texte

L'usage de l'IA comme outil documentaire dépend du stade de recherche où en sont les étudiants. Ceux qui n'ont pas de gros besoins documentaires ne mentionnent pas l'usage de l'outil documentaire. Un seul dit avoir trouvé trop d'erreurs dans les références pour faire confiance à l'IA sur ce point. Mais la majorité cite spontanément l'aide à la découverte d'un champ : « *Quand je ne connais pas, j'interroge* » ou « *C'est quoi le concept de X ? Je vais chercher, il [Chat] me donne quelques sources.* » Mais aussi : « *Il me propose parfois des choses fausses.* » (doctorant asiatique)

Pour une recherche initiale, l'IA paraît très utile, mais très peu des étudiants interrogés ont pu citer les différents modèles de recherche documentaire et les comparer entre eux. Il nous semblerait donc indispensable de les former à l'usage de ces outils pour approfondir leur champ de recherche, au lieu de laisser cet apprentissage au simple hasard.

La production d'un écrit pose d'autres problèmes. Pour ceux qui ont un bon niveau linguistique (B2 réel ou C), la stratégie la plus courante n'est plus tellement de produire un texte, mais de l'adapter finement. Dans la correspondance courante d'abord. Les étudiants allophones sont nombreux à faire réécrire leurs mails par l'IA. Mais il y a une limite à cette pratique comme le constate une doctorante sud-américaine : « *Il ne faut pas que ça sonne faux. Donc j'enlève certaines formules, trop pompeuses, que l'IA propose.* » Encore faut-il avoir un niveau suffisant non seulement dans la langue, mais dans la culture des interlocuteurs pour s'en apercevoir.

Mais l'IA peut communiquer et, par exemple, une étudiante lui demande : « Réponds de manière positive » et elle constate que « *ça marche, il s'adapte !* ». Une autre s'en sert aussi dans la correspondance courante : « *Il fait quelque chose de plus poli, de plus amical.* » Une doctorante déclare cependant que, de toute façon, pour un message plus important, elle fera quand même relire par un Français.

Pour la rédaction du mémoire ou de la thèse, les avis sont plus variés. Une limite observée par les accompagnateurs est la fragmentation des textes du fait des limites quantitatives du Chat gratuit. Certains étudiants hésitent à s'en servir de peur de fournir un texte qui ne soit pas réellement le leur. La question de l'appropriation est posée, plus que celle du plagiat. Les avis sont très partagés : s'en servir, oui, mais rester maîtres de son œuvre : « *Je ne veux pas que Bard entre dans l'avancement de ma thèse.* » (doctorant asiatique). D'autres l'utilisent couramment pour la rédaction : « *Il y a des phrases à reformuler, mais je m'en sers... et ma directrice de recherche me corrige ensuite de toute façon.* » (doctorant africain)

Un constat nous a interpellés : contrairement aux professeurs, les étudiants interrogés ne réfléchissent pas à l'usage de l'IA en termes d'intégrité académique. Ils sont simplement persuadés qu'il y a une différence entre la forme et le fond et que les lecteurs-évaluateurs seront capables de distinguer s'il y a des passages non originaux.

Mais le doute existe sur le fait qu'il faut le dire ou non : « *Je ne l'utilise pas pour les idées, donc c'est inutile de le dire.* »

D'ailleurs de l'avis général tout le monde utilise l'IA et il n'y a aucune raison qu'eux s'en vantent et aucune personne interrogée ne parle de consignes précises de la part des établissements. Deux étudiants seulement, parmi les plus avancés en âge et en connaissance du français, expriment des craintes générales pour l'avenir, mais de manière assez floue. On voit bien qu'aucun débat n'a encore eu lieu pour discuter de cette nouvelle manière de chercher et de produire de la connaissance.

3.4. Des pratiques pédagogiques en cours de transformation

Plusieurs accompagnateurs interrogés font état des pratiques auxquelles ils et elles ont recours depuis 2023, date de l'irruption de ChatGPT, comme critiquer et ajuster un texte généré par l'IA (syntaxe, clarté, caractère convaincant, original...) ou encore comparer deux textes générés par l'IA.

Les accompagnateurs demandent aux étudiants d'ajouter le cas échéant la mention « texte élaboré à l'aide de ChatGPT » (ou Gemini...). Au-delà de cette recommandation, un échange entre les accompagnateurs serait souhaitable. Pour le moment, chacun innove de son côté. Ainsi, un accompagnateur demande à un doctorant (de faible niveau linguistique) de faire une lecture à haute voix du texte traduit et/ou généré afin de perfectionner la phonétique. Une autre propose de vérifier la compréhension réelle de certains mots, expressions ou notions, puis de faire une recherche complémentaire si nécessaire. Un autre incite à vérifier les sources et demande de modifier la présentation des références. De nombreux collègues s'interrogent sur l'avenir tant de leur rôle d'accompagnateur que de l'évaluation des travaux.

3.5. Des risques pour tous

Une partie des étudiants et la totalité des accompagnateurs se posent des questions d'ordre général sur l'IA et son impact sur

l'enseignement supérieur : les formes de contrôle devront évoluer, l'oral prendra une plus grande place, il faudra être vigilant sur les textes, la pédagogie devra changer... : « *Les enseignants devront sortir de leur petit système.* » (une accompagnatrice, ex-professeure)

Seules deux personnes (une étudiante afghane et un étudiant russe) se soucient de la confidentialité des données personnelles et décident de ne pas se servir de l'IA pour certains textes (par exemple leur CV) de peur d'une utilisation extérieure. Le doctorant africain évoque le décrochage numérique de l'Afrique, en raison de ressources électriques et informatiques insuffisantes et du coût des abonnements. Un doctorant asiatique est le seul à évoquer le contenu des données : « *Il y a asymétrie au niveau mondial, nous n'avons pas accès aux corpus chinois, il y aura des gens qui vont avoir accès aux deux !* » Il souligne les effets possibles du développement de corpus différents à travers le monde.

En revanche, pour produire des idées, explorer ou documenter un sujet, l'utilisation d'une IA est assez rare dans notre échantillon. Mais cet état de choses est appelé à évoluer avec la diffusion toujours plus grande des outils. L'IA étant considérée comme un outil de perfectionnement linguistique, et non comme concepteur de textes nouveaux, la question du plagiat ou de la fraude est peu abordée. Les étudiants n'ont pas le sentiment de frauder, ce qui s'explique par le fait que la plupart des textes faisant l'objet d'accompagnement au CPU sont des mémoires et des thèses, et non des travaux sur tables lors d'un examen.

On est frappé par l'absence (là encore peut-être provisoire) de recommandations et d'interdictions de la part des établissements, qui ne donnent aucune consigne claire. De même, ni les universités, ni les écoles doctorales n'ont encore proposé de formation⁸.

4. DISCUSSION FINALE ET VOIES DE RECHERCHE

À une exception près, le bilan de l'IA fait par les étudiants est positif ; ils ont conscience de ses limites, mais pensent qu'elle va

continuer à s'améliorer. Leur attitude est pragmatique ; comme on pouvait s'y attendre, ils utilisent énormément l'IA pour la traduction et le perfectionnement des textes. L'utilisation de plusieurs langues (maternelle, intermédiaire, française), la correction et l'affinement de textes sont grandement facilitées par les outils d'IA. On peut penser que l'enseignement du FLE va être fortement impacté par l'IA, comme en témoignent déjà certaines publications dans les milieux spécialisés.

L'usage de l'IA dans l'enseignement supérieur soulève des questions spécifiques ; dans le cas particulier des étudiants allophones, cet usage accentue des problèmes préexistants. Nombreux sont les étudiants ne disposant pas en réalité d'une maîtrise suffisante du français pour faire des études supérieures sans difficulté. Le succès aux tests linguistiques (B1, B2, C1) est en principe nécessaire pour s'inscrire dans une formation, mais ces tests ne sont pas toujours passés dans des conditions sérieuses, les attachés culturels à l'étranger et Campus France jouant surtout leur rôle de développeur de l'attractivité française.

La rédaction d'un mémoire ou d'une thèse en français est donc pour beaucoup une épreuve considérable, qui poussera à utiliser tous les moyens disponibles, dont l'IA. La rédaction autorisée en anglais dans certains établissements facilite la tâche pour certains (les anglophones, surtout), mais elle n'évite pas toujours le recours à des corrections de textes.

À cette difficulté linguistique s'ajoute le faible taux d'encadrement des travaux universitaires en SHS dans de nombreuses universités françaises. Or un accompagnement méthodologique est d'autant plus nécessaire pour les étudiants allophones, surtout quand ils viennent de cultures dont les traditions académiques sont différentes (styles de rédaction, formulation du plan, présentation de références, attitude face au plagiat...). De ce point de vue, l'IA peut fournir des solutions dont l'étudiant peut manquer par ailleurs.

Contrairement à ce qu'on pourrait croire, l'utilisation de l'IA pourrait aboutir à rendre les interventions de l'enseignant encore plus

nécessaires : ce dernier devra d'une part lire avec attention les écrits produits pour y détecter l'éventuel usage de l'IA (quand celui-ci est interdit) et d'autre part la vérification des connaissances acquises passera par des situations de communication orales.

Ces situations modifieront considérablement la relation pédagogique, comme elles modifient déjà le rôle des accompagnateurs dans notre association. Même sur un texte déjà corrigé, l'enseignant ou l'accompagnateur devra utiliser les outils de correction comme des outils d'apprentissage et vérifier que l'étudiant est capable de lire correctement et de comprendre le nouveau texte français ainsi produit. La logique du raisonnement et la clarté des arguments ne sont pas prises en charge par l'IA et dépendront surtout d'un dialogue humain.

L'absence de recommandations éthiques de la part des établissements – pour l'heure – et la faible sensibilité des étudiants internationaux à ce sujet ne dispense pas les accompagnateurs de délivrer un message de prudence et de vigilance aux étudiants.

• Pour les étudiants :

Outre la recommandation générale de vérifier les résultats obtenus par l'IA, il faut, selon nos expériences, insister sur quelques points :

- Il faut demander d'indiquer systématiquement les outils utilisés.
- Les références bibliographiques proposées par une IA de bibliographie pouvant être fausses bien que plausibles (et donc trompeuses), elles doivent être vérifiées.
- Les résumés sur des sujets précis peuvent être des paraphrases d'un texte original. Il y aurait donc plagiat si le texte généré est repris par l'étudiant tel quel sans mention de la source. D'une manière générale, les paraphrases doivent être utilisées avec prudence et systématiquement attribuées à l'auteur du texte original (par l'usage d'expressions comme « selon l'auteur », « X déclare que », « écrit que », « selon lui... », etc.).

Il faut en outre inciter les étudiants à suivre les formations méthodologiques proposées par leurs établissements. En définitive,

il faut établir l'obligation de signaler sur chaque texte, ou à chaque séance, l'utilisation éventuelle d'un modèle d'IA.

- **Pour les accompagnateurs :**

Sur la base de cette enquête et d'échanges avec les bénévoles de l'association, nous proposons les mesures suivantes :

- Sensibiliser des encadrants à l'IA, à ses possibilités et limites de manière à ce qu'ils puissent dialoguer avec les étudiants utilisateurs.
- Diffuser des recommandations sur les limites et dangers de l'IA pour la rédaction et la traduction de textes.
- Recommander de mentionner dans les travaux académiques le nom des outils d'IA utilisés.
- Élaborer des méthodes pédagogiques pour la rédaction et la correction de textes intégrant l'IA (mesure également valable pour les enseignants de FLE).
- Faire signer par l'étudiant en début de cycle un court document fixant les « règles du jeu » du CPU quant à l'usage de l'IA.
- Inciter les étudiants à suivre les formations dispensées à l'université, qu'il s'agisse de formations pratiques (rédaction de prompts, usage des chatbots, usages pour la recherche documentaire...) ou de formations à l'éthique (chartes éthiques des établissements).

5. CONCLUSION

L'IA générative n'est pas (encore) utilisée totalement et systématiquement par les étudiants allophones de notre échantillon, mais il faut se souvenir qu'ils travaillent principalement sur des sujets littéraires et de SHS (disciplines dans lesquelles l'usage systématique de l'IA semble moins développé qu'en médecine ou en sciences de la nature et en technologie) et qu'ils appartiennent à une génération découvrant l'IA dans leur cursus supérieur. L'IA est vraisemblablement déjà en train de se généraliser dans

l'enseignement secondaire, sous toutes ses applications (textes, images, vidéos...).

Il semblerait que l'enseignement supérieur français soit encore dans une zone floue : l'usage se répand, on a conscience des risques qu'il comporte et les établissements proposent quelques séminaires, mais peu de consignes précises en dehors des chartes d'éthique existantes, antérieures à l'apparition de l'IA.

Il est donc nécessaire de sensibiliser tous les acteurs à un usage raisonné de ces outils. Au-delà des grands débats éthiques, il serait souhaitable de proposer des formations pour montrer concrètement leurs apports et leurs limites.

La question générale posée par l'IA rejoint celle qui a déjà été posée par les anthropologues de la technique : l'outil extériorise toujours un savoir par rapport à l'individu, comme la machine avait remplacé l'artisan et créé l'ouvrier spécialisé. Le savoir-écrire (en français pour nos étudiants) risque d'être délégué à la machine, le savoir-faire restant à l'individu n'étant plus qu'un résidu, celui de bien formuler les prompts. La rédaction en tant qu'expression personnelle d'une pensée propre risque de reculer au profit d'un texte correct, mais standardisé – même s'il faut reconnaître que la « moyennisation » était déjà une tendance observable avant l'apparition des IA génératives. Les pratiques pédagogiques et les pratiques d'évaluation doivent donc être transformées, si l'on veut éviter « l'industrialisation de l'insignifiance » (Alombert, 2023). À un autre niveau, la position de l'institution doit être également clarifiée dans chaque établissement par des consignes et des recommandations⁹.

En ce qui concerne notre association, il nous semble que la relation personnelle entre l'accompagnateur et l'étudiant allophone ou la relation au sein d'un petit groupe d'étudiants reste indispensable. Mais une modification des pratiques pédagogiques et une redéfinition du rôle que peut jouer cette relation durable (quelques mois, voire quelques années pour les thésards) s'imposent. Plus généralement, à l'échelle des universités,

l'apparition de l'IA obligera à développer la communication orale. Or les structures universitaires traditionnelles – notamment dans les écoles doctorales – ne sont pas préparées.

Notre expérience confirme celle des didacticiens : quels que soient les outils utilisés, c'est par une relation humaine que se crée le climat le plus favorable à l'apprentissage.

Références bibliographiques

- Alombert, A. (2023). Panser la bêtise artificielle. *Appareil*, (26). <https://doi.org/10.4000/appareil.6979>
- Chan, C. & Hu, W. (2023). Student's voice on Gen AI: Perceptions, benefits and challenges in higher education. *International Journal of Education. Technology in Higher Education*, 20(43). <https://doi.org/10.1186/s41239-023-00411-8>
- Collin, S. & Marceau, E. (2022). Enjeux ethniques et critiques de l'intelligence artificielle en enseignement supérieur. *Éthique Publique*, 24(2). Consulté le 05/09/2024 sur <https://doi.org/10.4000/ethiquepublique.7619>.
- Day, T. (2023). A Preliminary Investigation of fake peer reviewed citations and references generated by ChatGPT. *The Professional Geographer*, 75(6), 1024-1027.
- De la Higuera, C. & Lyer, J. (2024). IA for teachers: an open text book. <https://pressbooks.pub/aiforteachers/>
- Dussarps, C. & Vaugier, E. (2023). L'usage de Chat GPT comme assistant du chercheur en SHS : apports, limites et risques. Dans *Actes de la 18^e Conférence Internationale EUTIC 2023 (Humanisme numérique et durabilité sociale)*, Octobre 2023, Bordeaux, France. <https://hal.science/hal-04248025v1/document>
- Firat, M. (2023). What Chat GPT means for universities: perception of scholars and students. *Journal of Applied Learning and Teaching*, 6(1), 57-63.
- Grinbaum, A., Chatila, R., Devillers, L., Martin, C., Kirchner, C., Perrin, J. & Tessier, C. (2023). Systèmes d'IA générative : enjeux d'éthique. Comité national pilote d'éthique du numérique. 2023, Avis 7 du CNPEN. <https://cea.hal.science/cea-04153216/document>
- Livian, Y. F. & Laurini, R. (2018). *Réussir son mémoire de master ou sa thèse. Guide pour les étudiants étrangers*. Éditions Campus Ouvert.
- Lobet, M., Honet, A. & Wathelet, W. (2023). Chat GPT : l'usage par les étudiants de première année à l'université. HAL-SHS-04190226
- MLA-CCC Joint Task force on writing and AI writing paper (2023). Modern Language Association of America. Voir <https://hcommons.org/app/uploads/sites/1003160/2023/07/MLA-CCCC-Joint-Task-Force-on-Writing-and-AI-Working-Paper-1.pdf>.
- Rudolph, J., Tan, S. & Tan, S. (2023). Chat GPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied learning and teaching*, 6(1), 342-363.

Silvestri, S. (2023). ChatGPT for scientific writing: Navigation potentials and challenges. *The Fifteenth International Conference on Future Computational Technologies and Applications, Future Computing 2023*, IARIA. <https://www.iaria.org/conferences2023/ProgramFUTURECOMPUTING23.html>

Song, C. & Song, Y. (2023). Enhancing academic writing skills and motivation assessing the efficacy of Chat GPT in AI assisted language learning for EFL students. *Frontiers in Psychology*, 14, 1260843. <https://doi.org/10.3389/fpsyg.2023.1260843>

UNESCO (2023). Guidance for generative AI in education and research. Voir <https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>.

Wang, T., Lund, B.D., Marengo, A., Pagano, A., Mannuru, N.R., Teel, Z.A. & Pange, J. (2023). Exploring the potential impact of artificial intelligence (AI) on international students in higher education: Generative AI, chatbots, analytics, and international student success. *Appl. Sci.*, 13(11), 6716. <https://doi.org/10.3390/app13116716>

Yilmag, H., Maxutov, S., Baitekoy, A. & Balta, N. (2023). Students attitudes towards Chat GPT: a technology acceptance model survey. *International Educational Review*, 1(1), 57-83.

1. Par souci de simplicité, nous avons adopté le masculin générique, pour indiquer les étudiants et étudiantes.
2. « Coup de Pouce Université ». <https://www.cpu-lyon.org/>
3. C'est le terme que nous utiliserons dans ce chapitre. L'accompagnateur ne vient pas se substituer au directeur de thèse ou de mémoire, mais le compléter dans une relation linguistique, méthodologique et parfois humaine.
4. Rappelons qu'il y a en France 400 000 étudiants étrangers, un chiffre en hausse de 20 % depuis 5 ans. Ils constituent la majorité des étudiants dans certaines écoles doctorales.
5. On trouve notamment l'article « Students voices », inclus dans notre bibliographie, que Google Scholar avait aussi découvert.
6. Plusieurs papiers de recherche notent cependant la fréquence des fausses références, dont certaines très trompeuses (revue existante, auteur existant... mais référence fausse). Voir Day (2023).
7. Nous comprenons dans cette rubrique les modèles généralistes (ChatGPT, Gemini, Copilot, Mistral...), les outils de traduction (DeepL, Google Translate...) et les outils de recherche documentaire (You.com, Paper Digest, Perplexity Semantic Scholar...). Nous n'avons pas intégré les outils de création d'images (Midjourney, Dall-E...).
8. Pour étayer ce constat, on peut relever que l'URFIST de Lyon (organisme de formation inter-universitaire des documentalistes) a proposé trois séances de deux heures sur l'IA et son utilisation académique pendant le premier semestre 2024. Ces séances affichent « complet » jusqu'à l'été, avec une quinzaine d'inscrits par séance. Il y a sans doute eu aussi d'autres initiatives locales. Mais il y a à Lyon 120 000 étudiants, ce qui signifie que l'offre de formation est pour l'instant faible.
9. On en trouve certains exemples en Amérique du Nord, par exemple les *Guidelines* de l'Université Carleton (Canada), version d'octobre 2023.

PARTIE 2. PUBLICATIONS ET INTELLIGENCE ARTIFICIELLE

COMMUNIQUE 2. QUALITÉ DE SIGNATAIRE DE PUBLICATIONS SCIENTIFIQUES

[Alexandre Zollinger](#), Maître de conférences HDR, Université de Poitiers (France)

Le nombre croissant d'articles rétractés pose le problème de la responsabilité des auteurs de ces publications. Si le nombre des coauteurs varie significativement selon les disciplines, le fait de s'octroyer le titre de coauteur répond à des dispositifs génériques.

Le communiqué qui suit a été rédigé par les participants à l'atelier « Auctorialité : articulation entre Droit et Éthique académique » lors du 3^e Colloque International de l'IRAFPA sur l'intégrité académique (Figueira da Foz, 20-22 juin 2024). L'atelier, que nous avons animé, a permis de dégager des voies d'action spécifiques. Les participants de l'atelier sont nommés dans les remerciements de cet ouvrage. Nous avons retenu la formulation figurant dans l'encadré suivant après de nombreux débats, dans le souci de rester synthétique et de dégager ce qui a été collectivement considéré comme « essentiel ». Le propos devait n'être ni trop précis, ni trop général. La diversité des contextes de recherche conduit notamment à écarter la préconisation de critères d'auctorialité scientifique trop précis qui risqueraient de méconnaître certains usages disciplinaires légitimes. Les renvois aux textes de référence en matière d'intégrité scientifique permettent un approfondissement.

L'action concertée des trois principes que sont la responsabilité, la distinction entre signature et remerciements et la transparence offre les moyens d'identifier les pratiques non intègres, des moyens applicables à toutes les disciplines pour identifier : auteurs « oubliés », signatures indues (pour des raisons de complaisance, de rapports d'autorité...). L'exigence de réserver la signature de publications aux « personnes ayant réalisé une contribution qualitativement ou quantitativement substantielle aux travaux » ne résout évidemment pas l'ensemble des difficultés : comment déterminer le caractère suffisamment « substantiel » ou non d'une

contribution ? C'est avant tout aux chercheurs concernés qu'il revient de se poser cette question et d'y répondre au regard des spécificités de leur recherche. Le principe de transparence permet quant à lui de relativiser les inévitables problèmes de « frontière » dans la détermination des signataires et d'appliquer concrètement le précepte *Suum cuique tribuere*.

Communiqué de l'IRAFPA – Qualité de signataire de publications scientifiques

La [Déclaration de Singapour sur l'intégrité en recherche \(2010\)](#) a le mérite de distinguer, aux points 6 et 7, les qualités d'auteur et de personne remerciée dans le cadre d'une publication scientifique :

« 6. *Publication* : Les auteurs doivent assumer la responsabilité de leur contribution à la rédaction d'articles scientifiques, de demandes de contrat, de rapports de recherche et de toutes autres formes de publication concernant leurs travaux de recherche. La liste des auteurs doit inclure ceux et seulement ceux qui remplissent les critères de la qualité d'auteur.

« 7. *Les remerciements* : Les auteurs doivent faire figurer dans leurs publications le nom et le rôle des personnes qui ont contribué à la recherche, mais qui ne remplissent pas les conditions pour être auteur : aide à la rédaction, sponsors, organismes financeurs ».

Dans le respect des différences de pratiques d'ordre disciplinaire, l'IRAFPA rappelle fermement son attachement aux trois considérations suivantes :

Responsabilité

Les cosignataires d'une publication scientifique sont solidairement responsables de son contenu ([Code de conduite européen pour l'intégrité en recherche, 2023](#)).

Distinction entre signature et remerciements

Une approche intégrée de la recherche implique que la qualité de signataire/cosignataire d'une publication scientifique soit reconnue à toutes les personnes ayant réalisé une contribution qualitativement ou quantitativement substantielle aux travaux, mais qu'elle leur soit exclusivement réservée. D'expérience, déterminer à l'avance (dans une recherche de consensus entre contributeurs et non de manière imposée) qui est signataire d'une publication et dans quel ordre sont nommés les cosignataires facilite le processus et permet d'éviter la survenance de conflits.

Transparence

Afin de respecter la distinction susmentionnée entre signature et remerciements tout en prenant en considération les contextes différents des projets de recherche (le champ disciplinaire, le caractère éventuellement pluridisciplinaire des travaux...) pouvant expliquer des pratiques de signature variables, une attention particulière doit être portée au principe de transparence. La nature (voir par exemple la [taxonomie CRediT](#)) et l'objet des contributions des différents signataires des publications doivent être précisés. Les éditeurs scientifiques intègrent d'ailleurs de plus en plus

fréquemment cette exigence à leur charte éditoriale. Dans le même souci de transparence, le recours à des outils d'intelligence artificielle générative doit être explicité ([Commission européenne, *Living guidelines on the responsible use of generative AI in research*](#), mars 2024).

CHAPITRE 5. L'ÉCOSYSTÈME DE LA SURPUBLICATION : ACTEURS, *MODUS OPERANDI* ET IMPACT SUR LE MONDE UNIVERSITAIRE¹

[Cinta Gallent Torres](#), Professeure, Université de Valence

[Rubén Comas-Forgas](#), Maître de conférences, Université des îles Baléares (Espagne)

1. INTRODUCTION

Dans le travail académique, la diffusion des résultats de recherche est essentielle pour le progrès et le développement des connaissances. Les chercheurs et universitaires investissent un temps et des efforts considérables dans la réalisation de travaux et dans le partage de leurs résultats avec la communauté scientifique au sens large du terme. Cependant, cet engagement est confronté à des défis complexes en raison de l'expansion rapide des publications au cours des dernières années (Johan *et al.*, 2024). L'augmentation substantielle du nombre de publications a suscité des préoccupations quant à l'exactitude, la pertinence et la durabilité de ces productions. L'accent mis sur la quantité plutôt que sur la qualité a conduit à une prolifération de publications qui apportent peu de valeur ajoutée au corpus des connaissances existant.

Cette production massive d'écrits a incité les acteurs de l'écosystème de la recherche (universitaires, universités, éditeurs et agences de financement) à évaluer de manière critique les conséquences de cette obsession, parfois appelée « fièvre de la publication académique » (Guion, 1988), et son impact sur l'intégrité scientifique. Cependant, aucune solution efficace n'a encore été trouvée.

Pour illustrer ces problèmes à l'aide de données, nous nous référerons au *Rapport sur la Science* publié par l'UNESCO en 2015, qui fournit des preuves rigoureuses et complètes. Ce document

révèle qu'en 2013, il y avait 7,8 millions de chercheurs dans le monde, un chiffre en hausse de 21 % depuis 2007. Des études ultérieures donnent des chiffres similaires et aucune donnée définitive n'a été trouvée pour actualiser ce chiffre ; il est donc raisonnable de supposer qu'en raison de la croissance continue de l'activité scientifique mondiale, le nombre de chercheurs a augmenté depuis de manière significative. Cette augmentation apparaît clairement dans le volume des publications scientifiques. En effet, le même rapport indique qu'en 2014, il y avait environ 1,27 million de publications scientifiques, ce qui représente une croissance de 23,4 % depuis 2008.

Pour fournir des données bibliométriques plus récentes, nous pouvons nous tourner vers l'étude menée par Thelwall et Sud (2022), qui examine l'évolution de la base de données bibliographiques Scopus de 1900 à 2020. Cette étude révèle un schéma de croissance exponentielle constante du nombre d'articles indexés, avec une accélération significative à partir de 2004. Elle note également que, tandis que le nombre moyen d'articles par revue est resté stable jusqu'en 1980, il a ensuite triplé en raison de l'émergence des méga-revues et de la publication en ligne. De plus, la portée de la base de données s'est considérablement élargie, le nombre de domaines spécialisés passant d'un seul en 1945 à 308 en 2020.

Dans la même perspective, Cabanac (2022) souligne qu'environ huit millions de chercheurs produisent environ cinq millions d'articles par an, toutes disciplines et langues confondues. Il note également que, selon *Dimensions*, la plus grande base de données sur la recherche au monde, cela se traduit par environ 16 000 articles publiés chaque jour. L'ampleur de cette production est difficile à appréhender, d'autant plus que chaque article doit passer par un processus d'évaluation rigoureux avant publication. Il devient rapidement évident que la tâche de lire et d'évaluer correctement un volume aussi immense d'informations dépasse les limites du possible.

Sur la base de ce qui précède, la présente étude vise à analyser la dynamique de la surpublication dans le milieu académique, en se concentrant sur les acteurs clés (qui ?), leurs motivations (pourquoi ?) et les méthodes qu'ils utilisent (comment ?). Notre objectif est d'explorer l'impact de la surpublication sur la qualité et la fiabilité de la recherche académique, tout en abordant les questions de la faute professionnelle et de la malhonnêteté dans la recherche. Nous examinerons comment ces pratiques affectent la communication scientifique et quelles sont leurs implications pour l'avancement de la connaissance.

2. CADRE THÉORIQUE

Ce travail s'appuie sur les conclusions présentées par les auteurs lors du 3^e Colloque International IRAFPA qui s'est tenu à Coimbra en juin 2024. Les discussions du colloque rejoignent les préoccupations plus vastes de la scientométrie, discipline qui englobe l'étude quantitative de la science, l'évolution des disciplines, les interactions entre la science et la technologie ainsi que la productivité des chercheurs (Spinak, 1996). Au fil du temps, de nombreux chercheurs ont montré que la croissance exponentielle des publications scientifiques conduit à un système de récompenses académiques pervers (Grimes *et al.*, 2016). Ce système pousse les chercheurs à produire constamment de nouveaux résultats pour progresser dans leur carrière. La conséquence de cette pression est que les publications ne contiennent plus toujours de contribution à la connaissance ; la production académique s'est transformée en course aux citations et à la reconnaissance, mécanisme qui altère l'intégrité de la recherche.

Parmi les facteurs menaçant la science, on trouve : l'augmentation exponentielle du nombre de chercheurs (Cabanac, 2022), le flot d'informations écrasant (Akbashev & Kalinin, 2023) et la pression pour produire toujours plus et avoir un impact de publication important. De plus, le développement des revues prédatrices (Gallent Torres & Comas-Forgas, 2022) et des usines à papiers

(*papermills*), ainsi que l'environnement compétitif incitent les chercheurs à recourir à des pratiques douteuses telles que la manipulation de données, le *salami slicing*, le plagiat, la paternité fantôme et la fausse évaluation par les pairs. Tous ces facteurs sont complexes et interconnectés et ils ont une conséquence commune : non seulement ils sapent la fiabilité et la crédibilité des publications scientifiques, mais ils créent un cercle vicieux qui érode la confiance du public dans la science.

À mesure que la recherche de mauvaise qualité se répand, il devient difficile pour la communauté académique comme pour la société de distinguer les études solides des travaux trompeurs ou triviaux. D'où la nécessité urgente d'un changement dans la manière dont la recherche scientifique est publiée et évaluée. Cependant, sans une reconnaissance claire du problème et sans une volonté de le dénoncer, modifier les pratiques de publication actuelles et réduire la productivité scientifique sera une tâche difficile. Aborder ces questions nécessite non seulement des changements systémiques, mais aussi un engagement collectif pour établir des normes de qualité et d'intégrité plus élevées.

On a observé à plusieurs reprises que la prolifération de publications émanant d'un seul auteur ou d'un petit groupe de chercheurs viole nécessairement les principes fondamentaux de l'intégrité académique. Il est en effet difficile d'imaginer qu'on puisse publier une nouvelle étude tous les trois jours (Maisonneuve, 2024), être éditeur associé dans de nombreuses revues internationales prestigieuses, faire office d'éditeur invité pour 40 numéros spéciaux et être mentionné comme auteur principal de centaines d'articles. C'est pourtant le cas du Pr Dr Vo Xuan Vinh, un des grands chercheurs en économie du Vietnam, l'un des quatre Vietnamiens figurant dans le « top 1 % des chercheurs les plus cités au monde » selon Clarivate (Library Learning Space, 2024). Malgré sa carrière illustre, l'un de ses articles a récemment été retiré par la revue *Environmental Science and Pollution Research*, éditée par Springer. L'article intitulé « Trade Openness and CO2 Emanations: a Heterogeneous Analysis on the Developing Eight (D8) Countries »

contient des références inexactes et des phrases torturées ; le processus d'évaluation par les pairs a été manipulé (Else, 2021) ; de plus, son contenu n'était pas en adéquation avec le champ de recherche de la revue. L'article a pourtant été consulté 1 054 fois et a cité 54 fois.

Il est également difficile de croire qu'une personne puisse avoir déjà 37 795 citations et un indice i10 de 544 selon les métriques de Google Scholar au moment de la rédaction sans une certaine forme de manipulation des citations. C'est apparemment le cas de Juan Manuel Corchado, le nouveau recteur de l'Université de Salamanque en Espagne, qui a été accusé de « gonfler ses métriques par des auto-citations » et de « booster artificiellement son CV pour se classer au-dessus des chercheurs de premier plan dans le monde » (Ansedo, 2024). De telles actions ont conduit Corchado à être mentionné dans un article de *Nature* portant sur la manipulation des citations (Chawla, 2024), ainsi qu'à ce que la même revue rétracte 75 articles de lui et de ses collaborateurs.

Une autre pratique étrange est l'attribution de la paternité à des individus décédés, un phénomène de plus en plus fréquent (University of Melbourne, 2022). Le cas de Jiří Jaromír Klemeš, ingénieur scientifique de l'Université de Technologie de Brno, est particulièrement remarquable, puisqu'il a coécrit au moins 49 articles publiés après sa mort en janvier 2023 (Youmshajekian, 2024). Plus étonnants encore sont les résultats d'une étude menée par Jung *et al.* (2022), qui explore cette tendance dans le domaine biomédical. Les auteurs ont analysé 2 601 457 articles de 1990 à 2020 et identifié 1 439 auteurs décédés, qui ont collectivement publié 38 907 articles, dont 5 477 après leur décès. Cela pourrait « intensifier la croissance du nombre moyen d'auteurs par publication, ce qui pose également un problème éthique pour la paternité en général » (Jung *et al.*, 2022, p. 12).

La fièvre de publication freine le progrès scientifique et conduit souvent à des pratiques non éthiques. Le système est perverti par la priorité accordée aux indicateurs quantitatifs plutôt qu'au véritable

mérite scientifique, à la malhonnêteté plutôt qu'à l'intégrité. Comme l'ont montré Perceval et Fornieles Alcaraz (2008), un système qui récompense la quantité plutôt que la qualité scientifique risque de passer à côté des véritables découvertes. On encourage ainsi la routine et le conservatisme intellectuel, ce qui, combiné avec la fragmentation des institutions en disciplines et départements finira par transformer le champ de la recherche en un désert stérile (*ibid.*, p. 221).

En résumé, les causes de cette fièvre de publications doivent être analysées, à commencer par les acteurs-clés à l'origine de ce problème, leurs motivations et leur *modus operandi*. Une compréhension plus claire de ces trois éléments nous permettra d'évaluer plus précisément l'impact du phénomène, de maintenir des normes éthiques élevées et de mettre en œuvre des mesures efficaces pour garantir la transparence et la responsabilité dans la publication académique.

3. LES ACTEURS : LE BON, LA BRUTE ET LE TRUAND

L'écosystème académique est composé de multiples acteurs, chacun jouant un rôle distinct dans la dynamique de la surpublication. Ces acteurs-clés sont :

- **Les chercheurs** : Au cœur du processus de publication académique, les chercheurs subissent une immense pression pour publier. Qu'il s'agisse de jeunes chercheurs en début de carrière ou d'universitaires seniors, ils sont jugés en fonction de leur dossier de publications, qui influence directement leur carrière, la sécurité de leur emploi et leur réputation. Ainsi, ils se laissent souvent prendre dans le cercle vicieux de la surpublication pour maintenir leur position. Les vrais chercheurs, « Le[s] Bon[s] », sont ceux qui parviennent à sortir du cercle vicieux tout en produisant des recherches authentiques. Les autres, les surpublieurs, au contraire, recourent à des pratiques moins rigoureuses pour répondre aux exigences du milieu académique : publications en double, *ghostwriting*, autoplagiat et manipulation des données. Certains vont

même jusqu'à compromettre leur intégrité pour le prestige, comme on le voit avec les chercheurs hautement cités (*Highly Cited Researchers*, HCR) affiliés aux universités saoudiennes pour assurer leur inclusion dans le prestigieux classement de Shanghai (Ansede, 2023).

- **Les institutions** : Les universités et les organismes de recherche encouragent la productivité de leur personnel, car elle impacte directement leur classement, leurs opportunités de financement et leur perception par le public. Les institutions établissent ainsi des indicateurs de performance qui privilégient la quantité de publications plutôt que leur impact réel ou leur caractère innovant. Cette pression se répercute sur les chercheurs, qui doivent répondre aux exigences de l'institution pour obtenir une titularisation ou une promotion. Dans ce groupe, nous incluons également les instances d'évaluation et les financeurs de la recherche, qui évaluent les chercheurs en fonction du nombre de leurs publications. Les indicateurs de performance, la cupidité, le désir de pouvoir, et l'opportunisme sont les caractéristiques de « la Brute ».

- **Les revues académiques et les éditeurs** : Dans le film de Sergio Leone *Il Buono, il Brutto, il Cattivo* (1966), « le Truand » (Eli Wallach) est un personnage chaotique, opportuniste et malhonnête, prêt à exploiter n'importe quelle faiblesse pour un gain personnel. Dans le contexte de la surpublication, le Truand recourt à des pratiques toxiques et prédatrices, notamment à la publication dans des revues fictives qui publient des recherches de faible qualité sans véritable évaluation par les pairs. Ces revues privilégient leurs intérêts propres au détriment de la recherche académique ; souvent elles diffusent des informations fausses ou trompeuses, recrutent auteurs, évaluateurs et membres de leurs comités scientifiques d'une manière agressive et douteuse, font preuve d'un manque flagrant de transparence dans le processus de révision et d'édition.

La publication prédatrice, qui se combine à la surpublication, risque encore d'augmenter avec le développement des grands modèles de langage basés sur l'intelligence artificielle (Durand &

Roux-Steinkühler, 2023 ; Gallent-Torres *et al.*, 2024). Il faudrait donc lutter contre la pression excessive qui s'exerce sur les chercheurs et renforcer les mécanismes d'évaluation des résultats scientifiques. Car les pratiques prédatrices sont à la fois une cause et une conséquence de la surpublication. D'une part, elles offrent une voie aux chercheurs moins scrupuleux pour continuer à diffuser des travaux de mauvaise qualité. D'autre part, elles piègent des chercheurs bien intentionnés, mais moins familiers avec le système de la publication académique.

Le rôle des « *gatekeepers* » (évaluateurs par les pairs et rédacteurs de revues) devient alors crucial. Malheureusement, leur efficacité est souvent compromise par la surcharge de travail, la pression pour accepter des articles et les conflits d'intérêts potentiels. Ces problèmes affectent non seulement la qualité des publications, mais contribuent également à la surpublication, perpétuant un cycle infernal qui nuit à l'intégrité de la recherche. L'augmentation du nombre de rétractations d'articles en raison de fraudes lors de l'évaluation par les pairs est un symptôme des vulnérabilités du système. Selon McCook (2015), les fausses évaluations par les pairs « deviennent une tendance » (en ligne). De plus, Haug (2015) mentionne le cas d'un chercheur sud-coréen, Hyung-in Moon, qui utilisait de fausses adresses e-mail pour évaluer ses propres manuscrits. Plus de 250 de ses articles ont dû être rétractés. Haug (2015) mentionne également le Comité international pour l'Éthique des Publications (COPE), selon lequel certaines activités frauduleuses étaient orchestrées par des agences qui aidaient les auteurs à améliorer leurs articles et leur vendaient ensuite des évaluations par les pairs favorables.

La sophistication de telles activités frauduleuses n'a fait qu'augmenter depuis cette date. Au-delà des questions éthiques et des risques de corruption, un autre problème crucial persiste : le manque d'experts qualifiés pour examiner en profondeur les articles et identifier les erreurs avant publication. Le volume d'articles à évaluer nécessite un nombre extraordinaire d'experts prêts à

consacrer leur temps à cette tâche, et tout le monde n'est pas disposé à le faire (Cabanac, 2022).

La surpublication académique est dans une situation comparable à celle des personnages du film : leurs motivations sont diverses, ils évoluent dans un univers chaotique. Pour que le phénomène prospère à ce point, il faut la complicité des chercheurs, des institutions et des éditeurs. La surpublication n'est pas un problème isolé, mais un symptôme de défaillances plus larges qui nuisent à l'intégrité académique : conflits d'intérêts, auteurs invités, concurrence pour les financements, motivations guidées par le profit, partenariats commerciaux, etc.

4. MOTIVATIONS DE LA SURPUBLICATION : MAIS POURQUOI ?

Pour comprendre ce qui motive la surpublication, il est essentiel d'examiner les incitations au niveau individuel et institutionnel. Au niveau individuel, les chercheurs sont poussés par des facteurs tels que l'avancement de carrière, la reconnaissance et la recherche de financements (Horbach & Halfman, 2019), qui sont souvent directement liés au nombre de publications et de citations qu'elles génèrent.

Au niveau institutionnel, les universités et les centres de recherche ont tout autant intérêt à augmenter leur production de publications. Les institutions académiques rivalisent à l'échelle mondiale pour les classements, les financements et le prestige et les publications sont le principal indicateur d'excellence. La dépendance croissante aux indicateurs bibliométriques (nombre d'articles et facteur d'impact des revues dans lesquelles ils paraissent) incite les institutions à pousser leurs chercheurs à produire davantage (Dani, 2018). Cette pression institutionnelle encourage la quantité plutôt que l'excellence.

Les médias ont relayé des cas de mauvaise conduite institutionnelle. En Arabie saoudite, certaines universités de premier plan ont eu recours à des fraudes – achat de chercheurs hautement cités à l'étranger, manipulation des affiliations, déclaration de faux indicateurs de recherche, etc. – pour améliorer leur production et leur

crédibilité sur la scène mondiale (Johan *et al.*, 2024). La présence d'un chercheur hautement cité est perçue par l'université comme un gage de qualité et peut améliorer le classement de plus de 100 places (SIRIS Academic, 2023).

De plus, les offres financières faites aux chercheurs intensifient les pratiques contraires à l'éthique. Par exemple, Blanca Landa, chercheuse à l'Institut de l'Agriculture Durable du CSIC à Cordoue (Espagne), a reçu de la part d'un professeur de l'Université King Saud une offre de 1 500 dollars par étude publiée ; en échange, elle devait accepter d'être mentionnée comme professeure invitée tous frais payés – elle a immédiatement refusé (Ansede, 2023). De même, Luis Martínez, Professeur à l'Université de Jaén, a accepté une offre annuelle de 66 000 dollars de l'Université King Saud pour être mentionné comme universitaire affilié. Cette offre est venue après qu'il n'avait pas réussi à obtenir des financements du gouvernement espagnol pour ses projets (Ansede, 2023). Un rapport intitulé *The Affiliation Game of Saudi Arabian Higher Education & Research Institutions*, publié par SIRIS Academic (2023), a également révélé ces pratiques manipulatrices, dévoilant même le cas d'une agence intermédiaire qui percevait des commissions pour persuader des chercheurs hautement cités de changer d'affiliation, ouvrant ainsi la voie à un marché de niche potentiellement frauduleux.

Les chercheurs individuels et les institutions sont influencés par les récompenses financières et les pressions compétitives, ce qui a un impact significatif sur l'intégrité académique. Or, l'influence des pairs, la recherche d'un statut plus élevé ou d'une plus grande renommée ne justifient en aucun cas des pratiques contraires à l'éthique.

5. MODUS OPERANDI : COMMENT S'Y PRENNENT-ILS ?

Dans cette section, nous allons explorer certaines des « pratiques de recherche discutables » (Steneck, 2006) des hyperpublicateurs. Ces pratiques varient selon les disciplines. La plus courante est le plagiat : des chercheurs présentent le travail d'autrui comme étant le

leur, sans attribution correcte (Eaton, 2021). Le plagiat peut prendre plusieurs formes : la copie servile, l'autoplégat de travaux déjà publiés et la paraphrase d'idées sans mention correcte de la source.

La fabrication d'études, autre méconduite grave, trompe les lecteurs et les parties prenantes sur les véritables conclusions de la recherche. Les études fabriquées contribuent à une perception faussée de la productivité et de la qualité du travail. Babbage (1980, cité par Fanelli, 2009, p. 1) a défini ce processus comme la « cuisson des données », « un art sous diverses formes, dont l'objectif est de donner à des observations ordinaires l'apparence et le caractère d'un degré de précision plus élevé ». De plus, la falsification et la manipulation des données aggravent encore les choses. Certains chercheurs peuvent présenter sélectivement des données ou les modifier pour qu'elles correspondent aux résultats souhaités.

Une pratique particulièrement insidieuse consiste à acheter des citations. Certains chercheurs exploitent des services qui vendent des citations en gros pour leurs articles. Des rapports indiquent que certaines entreprises proposent des forfaits comme 50 citations pour 300 \$ ou 100 citations pour 500 \$ (Chawla, 2024). Ces dernières sont parfois insérées dans des bases de données académiques comme Scopus, augmentant ainsi la visibilité et l'impact des articles. Selon Labbé (s.d., cité par Chawla, 2024), elles sont même vendues par des usines à articles. Les chercheurs qui les achètent améliorent leur dossier de publication sans effectuer une recherche légitime. La manipulation des indicateurs de citation sape la crédibilité des mesures bibliométriques utilisées pour évaluer la performance académique (Ansedè, 2023).

Il existe un marché noir pour des références truquées. Des études, telles que celles de Ibrahim *et al.* (2024), ont documenté comment des chercheurs participent à ces marchés illicites pour améliorer artificiellement leurs indicateurs de publication. Elles ont également mis en lumière la création de faux profils académiques pour gonfler des réalisations scientifiques. En fait, en juillet 2024, certains chercheurs ont créé un profil pour un chercheur fictif, un chat nommé

Larry, crédité d'une douzaine de faux articles (Richardson, 2024). Ce profil a d'abord été publié sur des plateformes comme ResearchGate et Google Scholar ; une fois la supercherie découverte, les articles ont été supprimés. Au-delà de l'anecdote, ce cas montre à quel point les profils académiques peuvent être facilement manipulés pour gonfler le nombre de publications.

En résumé, la liste des tactiques contraires à l'éthique semble interminable, incluant l'IA-plagiat, la manipulation algorithmique, le vol d'identité, les réseaux de citations, la triche contractuelle et les publications de remplissage (« *meat extender publications* »), entre autres. En tant que chercheurs, il est crucial de rester vigilants face à d'éventuels signaux d'alerte, tels que des schémas de citations suspects (Chawla, 2024), des collaborations inhabituelles, le *salami slicing*, etc. Les observateurs de l'intégrité dans la recherche jouent donc un rôle essentiel dans l'identification et le traitement de ces problèmes : ils aident à préserver la crédibilité des travaux académiques et à garantir que l'intégrité scientifique prévaut.

6. QUELLES SONT LES CONSÉQUENCES DE CE PHÉNOMÈNE ?

Les actions mentionnées précédemment ont des conséquences dramatiques pour le monde académique. Bien que le gaspillage de ressources financières et humaines soit un sujet de préoccupation majeur, la surpublication est encore plus grave, car elle menace de déstabiliser l'ensemble du système universitaire.

Une deuxième conséquence est la baisse de la qualité de la recherche. À l'instar de l'hyperinflation économique, où un excès de monnaie dévalue l'argent, la prolifération incontrôlée des publications scientifiques dévalue les contributions individuelles. À mesure que le volume de publications augmente, il devient difficile pour les chercheurs de discerner quelles études sont réellement pertinentes et lesquelles sont simplement incrémentales, ce qui conduit à une saturation excessive de la littérature (Akbashev & Kalinin, 2023). Une étude menée par Wong *et al.* (2022) sur les auteurs prolifiques en

orthodontie a révélé qu'un groupe de 49 auteurs était « collectivement prolifique pendant 1 à 8 ans au cours de la décennie évaluée (2011-2020) » (p. 416). Ces auteurs ont publié un total de 2 025 articles, avec une production médiane de 18 articles par an. En particulier, un auteur a publié 200 articles en une seule année, 2019, dans la même revue (non liée à la recherche dentaire). Malgré cette production élevée, de nombreuses publications manquaient de rigueur scientifique, se composant principalement d'études cliniques non prospectives, ou des comptes-rendus. Ce schéma montre une tendance plus forte dans d'autres domaines, où un nombre croissant d'articles de faible qualité encombrant le paysage académique, réduisant ainsi la fiabilité globale des recherches publiées (Wong *et al.*, 2022).

Une autre conséquence de la surpublication est l'augmentation des rétractations d'articles en raison d'erreurs ou de données frauduleuses. Pour Kun (2018) la fréquence des rétractations est l'indicateur clair d'une baisse de qualité, due aux diverses pressions qui pèsent sur la publication. Plus grave encore, on continue à citer des articles rétractés, de sorte que ces études défectueuses continuent à influencer la recherche ultérieure et la prise de décision (Zhang *et al.*, 2020).

La surpublication contribue également à la fragmentation des communautés de recherche. À mesure que les chercheurs se concentrent sur la publication dans des sous-domaines restreints, la communauté scientifique dans son ensemble se divise en petits groupes isolés. Cette fragmentation réduit les opportunités de collaboration interdisciplinaire et limite le potentiel de découvertes, qui sont souvent le fruit d'approches pluridisciplinaires (Akbashev & Kalinin, 2023). Dans des domaines comme l'apprentissage automatique et les photovoltaïques, cette division est particulièrement évidente, car les chercheurs peinent à suivre le volume accablant d'études publiées (Akbashev & Kalinin, 2023).

La surproduction de littérature scientifique a des conséquences pour la société au-delà du seul milieu universitaire. Lorsque la qualité de la recherche est compromise, la confiance du public dans la

science diminue, ce qui rend plus difficile l'application des résultats de la recherche. Dans des domaines comme la médecine, par exemple, des études inexactes peuvent conduire à des traitements inefficaces mettant en danger la santé publique (Schulz *et al.*, 2016). La surpublication complique l'identification de recherches fiables par les praticiens et entraîne des retards dans leur mise en œuvre. Enfin, à mesure que la confiance dans la recherche scientifique décline, il devient plus difficile de mobiliser un soutien pour des initiatives cruciales, telles que la lutte contre le changement climatique ou l'innovation technologique (Leiserowitz *et al.*, 2013).

7. CONCLUSIONS ET PROSPECTIVE

Il est regrettable que quelques « brebis galeuses » nuisent à la crédibilité de toute la communauté académique et provoquent une crise de la reproductibilité de la science. Cette préoccupation n'est pas nouvelle, comme le montrent des recherches antérieures (Kun, 2018). Mais le problème s'est aggravé avec le temps. Par exemple, Augustyn et Cillié (2008) se réfèrent à une étude de Tulving et Madigan (1970) qui a évalué 540 articles publiés dans des revues scientifiques et a conclu que deux tiers pouvaient être considérés comme « totalement sans importance ». Les mêmes auteurs mentionnent une autre étude de Garvey et Griffith (1963), qui estimait que 50 % des articles publiés dans les revues de l'American Psychological Association étaient lus par moins de 200 personnes et que deux tiers n'étaient jamais cités par d'autres auteurs. Malheureusement, la situation ne s'est pas améliorée au XXI^e siècle.

De nombreux cas de méconduite dans les publications académiques passent inaperçus en raison de mécanismes inadéquats pour les identifier. Par conséquent, il est impératif que toutes les parties prenantes surveillent attentivement les comportements malhonnêtes dans leurs domaines respectifs. La surpublication représente un risque collectif qui doit être pris en compte et combattu.

Références bibliographiques

- Akbashev, A. R. & Kalinin, S. V. (2023). Tackling overpublishing by moving to open-ended papers. *Nature Materials*, 22(3), 270-271. <https://doi.org/10.1038/s41563-023-01489-1>
- Ansede, M. (2023, mars 3). Arabia Saoudi paga a científicos españoles para hacer trampas en el 'ranking' de las mejores universidades del mundo. *El País*. <https://tinyurl.com/yymdc5ck>
- Ansede, M. (2024, janvier 15). El aspirante a rector que escribió cuatro párrafos y se citó a sí mismo 100 veces. *El País*. <https://tinyurl.com/47vawxnd>
- Augustyn, J. C.D. & Cillié, G.G. (2008). Theory and practice in industrial psychology: *quo vadis?* Theoretical research. *SA Journal of Industrial Psychology*, 34(1), 70-75. <https://hdl.handle.net/10520/EJC89142>
- Cabanac, G. (2022). *Pollution de la littérature scientifique : détection participative d'expressions torturées révélatrices d'articles frauduleux* [Conférence]. Université Toulouse III. <https://www.youtube.com/watch?v=MtGMSfubD8Y>
- Chawla, D. S. (2024). The citation black market: schemes selling fake references alarm scientists. *Nature*, (632), 966. <https://doi.org/10.1038/d41586-024-01672-7>
- Dani, E. (2018). How "Publish or Perish" Can Become "Publish and Perish" in the Age of Objective Assessment of Scientific Quality. *Systemics, Cybernetics and Informatics*, 16(4), 20-25.
- Durand, B. & Roux-Steinkühler, M.A. (2023). Pour une écriture scientifique respectueuse des principes de la paternité intellectuelle. Droit et déontologie. *Les Cahiers de IRAPFA*, 1(1). Cahiers Méthodologiques. Institut de Recherche et Action sur la Fraude et le Plagiat Académiques. <https://doi.org/10.56240/irafpa.cm.v1n1/dur>
- Eaton, S. E. (2021). *Plagiarism in higher education: Tackling tough topics in academic integrity*. Libraries Unlimited.
- Else, H. (2021). 'Tortured phrases' give away fabricated research papers. *Nature*, 596, 328-329. <https://doi.org/10.1038/d41586-021-02134-0>
- Fanelli, D. (2009). How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PLoS One*, 4(5), e5738. <https://doi.org/10.1371/journal.pone.0005738>
- Gallent-Torres, C., Arenas Romero, B., Vallespir Adillón, M. & Foltýnek, T. (2024). Inteligencia artificial en educación: entre riesgos y potencialidades. *Práxis Educativa*, 19, 1-29. <https://tinyurl.com/4w3n7dkx>
- Gallent Torres, C. & Comas-Forgas, R. (2022). Revues prédatrices et mauvaises pratiques éditoriales : une menace croissante pour l'intégrité académique. *Synergies Espagne*, (15). https://gerflint.fr/Base/Espagne15/gallent_comas.pdf
- Garvey, W. D. & Griffith, B. C. (1963). Research frontier: The APA project on scientific information exchange in psychology. *Journal of Counseling Psychology*, 10(3), 297-302. <https://doi.org/10.1037/h0045147>
- Grimes, D. R., Bauch, C. T. & Ioannidis, J. P. (2018). Modelling science trustworthiness under publish or perish pressure. *Royal Society open science*, 5(1), 171511.

Guion, R. M. (1988). *Pitfalls in the march of science. Annual meeting of the American Psychological Association.* APA.

Haug, C. J. (2015). Peer-review fraud – Hacking the scientific publication process. *The New England Journal of Medicine*, 373, 2393-2395. <https://www.nejm.org/doi/full/10.1056/NEJMp1512330>

Horbach, S. P. J. M. & Halffman, W. (2019). The extent and causes of academic text recycling or 'self-plagiarism.' *Research Policy*, 48(2), 492-502. <https://doi.org/10.1016/j.respol.2017.09.004>

Ibrahim, H., Liu, F., Zaki, Y. & Rahwan, T. (2024). Google Scholar is manipulatable. *Computational Engineering, Finance, and Science*. <https://doi.org/10.48550/arXiv.2402.04607>

Johan, D., Neufeld, J., Thomas, K., Rathmann, J. & Rauhut, H. (2024). The impact of researchers' perceived pressure on their publication strategies. *Research Evaluation*, 00, 1-16. <https://doi.org/10.1093/reseval/rvae011>

Jung, C.H., Boutros, P.C., Park, D. J., Corcoran, N.M., Pope, B. J. & Hovens, C.M. (2022). Perish and publish: Dynamics of biomedical publications by deceased authors. *PLoS ONE*, 17(9), e0273783. <https://doi.org/10.1371/journal.pone.0273783>

Kun, A. (2018). Publish and who should perish: You or science? *Publications*, 6(2), 18. <https://doi.org/10.3390/publications6020018>

Leiserowitz, A. A., Maibach, E. W., Roser-Renouf, C., Smith, N. & Dawson, E. (2013). Climategate, public opinion, and the loss of trust. *American behavioral scientist*, 57(6), 818-837. <https://doi.org/10.1177/000276421245>

Library Learning Space (2024). *Vietnam's top 'world's most influential' scientist had his international article removed.* [Blog]. <https://tinyurl.com/mvyd6ckd>

Maisonneuve, H. (2024). Ces guest editors de numéros spéciaux permettent toutes les manipulations ; Mr. Vo Nguyen Dai Viet publie un article tous les trois jours. *Revue et intégrité* [Blog]. <https://tinyurl.com/yj4n38ae>

McCook, A. (2015). 64 More papers retracted for fake reviews, this time from Springer journals. *Retraction Watch*. <https://tinyurl.com/5n6t8xua>

Perceval, J. M. & Fornieles Alcaraz, J. (2008). Confucio contra Sócrates: la perversa relación entre la investigación y la acreditación. *Anàlisi: quaderns de comunicació i cultura*, 36, 213-224. <https://raco.cat/index.php/Analisi/article/view/94533>

Richardson, R. (July 18, 2024). Engineering the world's highest cited cat, Larry [Blog]. <https://tinyurl.com/2yrmnr4y>

Schulz, J. B., Cookson, M. R. & Hausmann, L. (2016). The impact of fraudulent and irreproducible data to the translational research crisis—solutions and implementation. *Journal of Neurochemistry*, 139, 253-270. <https://doi.org/10.1111/jnc.13844>

SIRIS Academic (2023). *The affiliation game of Saudi Arabian higher education & research institutions. Analysing development of affiliation practices in the Highly Cited Researchers list from Clarivate between 2014-2022.* <https://tinyurl.com/nuw4ccdc>

Spinak, E. (1996). *Diccionario enciclopédico de bibliometría, cienciometría e informetría.* UNESCO.

- Stefan, A.M. & Schönbrodt, F.D. (2023). Big little lies: a compendium and simulation of p-hacking strategies. *Royal Society Open Science*, 10, 220346. <https://doi.org/10.1098/rsos.220346>
- Steneck, N.H. (2006). Fostering integrity in research: definitions, current knowledge, and future directions. *Science and Engineering Ethics*, 12, 53-74. <https://pubmed.ncbi.nlm.nih.gov/16501647/>
- Thelwall, M. & Sud, P. (2022). Scopus 1900–2020 : Growth in articles, abstracts, countries, fields, and journals. *Quantitative Science Studies*, 3(1), 37-50. https://doi.org/10.1162/qss_a_00177
- Tulving, E. & Madigan, G. (1970). Memory and verbal learning. *Annual Review of Psychology*, 29, 437-484.
- UNESCO (2015). *UNESCO Science Report. Towards 2030*. UNESCO Publishing. <https://tinyurl.com/zzhpbn8k>
- University of Melbourne (2022). *Researchers perishing and publishing: Journal publishing by deceased authors is increasing*. <https://tinyurl.com/4t29u3dy>
- Wong, C., Papageorgiou, S. N., Seehra, J. & Cobourne, M. T. (2022). Prolific authorship in orthodontic scientific publishing. *Orthodontics & craniofacial research*, 25(3), 416-428. <https://doi.org/10.1111/ocr.12551>
- Younshajekian, L. (2024). Highly cited scientist published dozens of papers after his death. *Retraction Watch*. <https://tinyurl.com/5n8kfhbt>
- Zhang, Q., Abraham, J. & Fu, H. Z. (2020). Collaboration and its influence on retraction based on retracted publications during 1978–2017. *Scientometrics*, 125, 213-232. <https://doi.org/10.1007/s11192-020-03636-w>

1. Remerciements : ce travail s'inscrit dans le cadre du doctorat de Cinta Gallent Torres à l'Université des îles Baléares, ainsi que du projet de recherche intitulé « Congrès, revues et éditeurs prédateurs : conception d'un programme de formation destiné aux étudiants de troisième cycle », avec le numéro de référence HUMA20232024-03, financé par l'Université Internationale de Valence, et du projet de recherche PID2022-141031NB-I00, financé par le MICIU/AEI /10.13039/501100011033/, et par le FEDER « manière de faire l'Europe ». Il fait également partie des activités du Réseau ibéro-américain de recherche sur l'intégrité académique.

CHAPITRE 6. RÉCIT, LITTÉRATURE ET PRATIQUE LITTÉRAIRE. DÉLICES DU MÉTIER D'ÉDITEUR ET LES VERTIGES DE L'IA

[Ignace Haaz](#), Éditeur en chef, Globethics Publications

1. INTRODUCTION

L'intelligence artificielle comme système conversationnel utile ne fait que s'étendre dans toutes les directions ; elle est devenue un instrument indispensable à bon nombre de nos tâches. Parmi ces applications, les chatbots conversationnels tels que ChatGPT n'ont que peu de traits conversationnels véritablement humains et ne favorisent intrinsèquement ni les progrès de la connaissance ni les qualités éthiques, mais ils séduisent quand même. Lorsque lecture et écriture constituent des pratiques bien ancrées chez l'individu, l'IA peut devenir un outil complémentaire utile. Mais Internet et ces outils restent insuffisants pour des personnes qui s'engagent dans la pratique de l'écriture sans se recentrer sur sa fonction de représentation, qui, dans son principe, n'est pas un jeu de rôle, le faire-semblant d'une approche cognitiviste de la littérature.

Le savoir-écrire pratiqué comme un art de la représentation repose sur des conditions préalables : avoir la capacité d'élaborer des connaissances ; avoir intériorisé une ambition de vérité comme raison purement théorique ; et avoir la capacité de former un jugement, ce qui associe d'une certaine manière la sphère de la théorie et celle de l'action pratique. Former un jugement est toujours une pratique orientée vers une finalité (éthique), un engagement dans une posture éthique de vertu (ou de vice), tout comme former un jugement impartial sur un objet revient à penser et organiser le modèle de l'objet (ou au contraire à négliger par paresse ou vanité de le faire)¹.

ChatGPT peut être entraîné à générer un texte ou à le modifier selon des injonctions précises, mais sans cette dimension cognitive du jugement, distincte de la simple imitation d'un modèle sans finalité pensée par nous-mêmes, il manque une raison fondamentale, en quelque sorte de principe, de reproduire d'une manière ou d'une autre dans l'écriture son objet et ainsi de contrer les conditions qui favorisent une forme de vice intellectuel (voire de délinquance académique). Ce vice est en grande partie le produit d'une soumission passive à un jeu de pouvoir, sans pensée ni organisation d'un modèle de connaissance qui suivrait un langage et une méthode définissable. Dans les métiers de la connaissance et de l'information, on justifie et on renforce la complaisance à la formidable pression sociale et économique à publier par la vanité de préférer consolider l'appartenance à un groupe social désirable, au détriment d'une meilleure hygiène dans la formation de ses croyances. On délègue la formation de ses croyances au pouvoir de groupes auxquels on se soumet passivement par commodité sociale et économique. Les raisons de type théorique – en définitive faire ou ne pas faire un travail soigné d'analyse de l'objet de connaissance – sont distinctes des conditions nécessaires qui appartiennent aux motifs prudents basés sur les conséquences d'une manière d'agir vertueuse – ou moins vertueuse par crainte de perdre des avantages sociaux.

Notre réponse à ce dilemme est que l'intégrité académique est radicalement une question à la fois théorique et pratique, elle exige une distinction sans laquelle le problème de l'intégrité académique se dissout dans un simple relativisme des forces sociales et économiques en présence. Prise à la lettre, comme science de l'homme, l'intégrité académique est une science appliquée, qui travaille à partir d'observations empiriques ; elle est guidée par les besoins, des conditions sociales de la production et de la communication du savoir. L'édition scientifique est la dernière étape du processus de recherche. La question de l'intégrité se pose dans les mêmes termes pour l'éditeur et les chercheurs. Éditer des thèses de doctorat portant sur l'éthique, proposer des codes de conduite

aux professionnels, mesurer l'enjeu des technologies d'information et de communication, tout cela engage la responsabilité de l'éditeur, qui doit, lui aussi anticiper les risques de fraude académique.

2. ENGEL, LA VÉRITÉ DANS LA LITTÉRATURE ET *LE MAÎTRE ET MARGUERITE* COMME VÉRITÉ ET SÉDUCTION

Pascal Engel nous éclaire sur la nature de l'œuvre littéraire et sur le caractère à la fois théorique et pratique des modes de définition et des méthodes d'appréhension de ce qui fait d'un récit, d'une fiction un travail à prétention littéraire, par contraste par exemple avec un roman de quai de gare. Comment définir cette qualité, puisque ce n'est pas une activité totalement pratique comme calculer ou faire du vélo ? Pour Engel, l'œuvre présente des qualités littéraires, essentiellement des « savoir-comment » et non des « savoir-quoi » (« *know how* », *not* « *know what* »), pour reprendre les termes de Ryle (1949) et de Benda (1945). Engel nie que le style puisse suffire à donner une définition de l'œuvre littéraire. Il veut concilier un critère robuste de *vérité* dans l'œuvre littéraire avec un mode de vérification de la nature de l'objet d'art littéraire dans la pratique. On parle pour cette raison de *practicalisme*, sans pour autant abandonner l'intellectualisme propre à la littérature, qui lui conserve sa valeur de gardienne de l'idéal d'humanisme des Modernes.

La valeur d'une histoire est de nous faire appréhender du vrai, Shakespeare dans son œuvre « dit le vrai », dit Engel². Certaines œuvres communiquent une vérité sur le monde, d'autres y sont indifférentes. Les premières ont là une valeur cognitive : elles produisent une connaissance et la transmettent, en s'arrêtant à la plus vraie. L'essentiel, pour l'œuvre littéraire par opposition à l'œuvre scientifique, est qu'une valeur cognitive, lorsqu'elle existe, peut contribuer aussi à la valeur esthétique de l'œuvre et donc faire doublement un bon office, pour ainsi dire. L'inverse semble pourtant aussi plausible, quand une souche nous suggère la présence d'un ours au détour d'un sentier ou lorsque le bonhomme de neige est si réussi qu'on le prend pour une vraie personne à la tombée de la nuit.

Dans une perspective précise, un artefact ou une forme donnée dans la nature peuvent exprimer une apparence de vérité troublante, qui peut être aussi agréable, mais réalisée lorsqu'ils provoquent une réaction chez le témoin, malgré la fausseté de la représentation ou de la référence. Des formes, comme un genre ou un style, qui provoquent une réaction, ne font pas une œuvre, elles peuvent inciter à imaginer rapidement un objet comme l'ours ou la présence humaine grâce à la silhouette d'apparence humaine dans la neige.

Enfin, une œuvre qui contient une large proportion de réflexions vraies, mais aussi une spéculation franchement séduisante qui va nous attirer vers l'erreur ou le vice, combinera deux niveaux de valeurs : esthétiques et éthiques. Ainsi, l'imagination nous fait pénétrer dans un territoire peu familier : *Le Maître et Marguerite* de Mikhaïl A. Boulgakov (2004), une œuvre littéraire russe majeure du XX^e siècle, est un roman fantastique qui nous invite aussi à considérer que Satan se manifeste à Moscou autour de 1930 sous l'identité d'un mystérieux magicien nommé Woland. Si l'œuvre fantastique se marie bien avec le genre littéraire du roman, malgré la distance que celui-ci prend vis-à-vis du critère de vérité, c'est pour un ensemble de raisons distinctes de ses propriétés stylistiques ou sémiologiques – des connotations associées au sens littéral des moyens de communication ou aux qualités de la dénotation des choses dans le monde. La vérité dans cette œuvre, si elle est présente, repose d'une part sur l'agencement des différents aspects d'une histoire d'amour, d'une critique politique et sociale et d'une comédie burlesque, décrits avec réalisme comme quelque chose d'observable. Même la personnification du vice et de l'erreur représentés par Woland est quelque chose de réaliste, non une licorne, et surtout aussi quelque chose de séduisant, et non un repoussoir normatif ou une anti-valeur. D'autre part, il y a des références historiques et littéraires importantes à la vie de Jésus, à Ponce Pilate, Faust, Tibère, Staline, etc. Le roman inclut l'autobiographie d'un auteur de roman et les péripéties de la création littéraire. Enfin, bien que ces éléments contribuent à poser un cadre de référence précieux pour une forme d'autocompréhension du

lecteur, ces références mobilisent le caractère proprement humain de l'expérience de lecture.

Le narrateur offre au lecteur son propre commentaire du texte, notons en passant que ce roman a été écrit à quatre mains. Il nous semble que cette mise en abîme du narrateur constitue la partie la plus consistante de l'œuvre littéraire. À côté de la mise en scène soigneuse des actions, cette expérience est comme un palimpseste, elle apporte un jugement d'une qualité exceptionnelle, elle renforce le geste pratique précis de l'écriture. Ensemble, ces trois dimensions produisent une densité de références qui, paradoxalement, éloigne le texte de son caractère de fiction pour lui donner une vérité littéraire.

En conclusion, bien qu'il s'agisse d'un roman fantastique ménageant l'intrusion du surnaturel ou de l'imaginaire – ce qui constitue un peu autre chose que de la littérature au sens cognitiviste –, il semble indiscutable que ce roman a une valeur littéraire comme système décrivant le vrai. Il n'y a pas que les grandes œuvres historiques qui ont une valeur cognitive par la précision de peintures dont le caractère vérifiable est souvent remarquable à travers le temps. Le mot « littérature » est ici essentiellement la dénotation d'une chose normative, comme le pointe bien Engel (2013) : on parle presque automatiquement de *bonne littérature* ou de littérature digne de ce nom, mais on n'emploie pas le mot pour des fictions auxquelles manque cet horizon normatif. Aussi remarquable que soit ChatGPT, l'IA ne produira pas de la littérature, elle génèrera quelque chose qui ressemble à une fiction... à condition d'être programmée soigneusement !

Pourquoi les textes comme ceux que produit ChatGPT ou encore les livres du kiosque de la gare, ne répondent-elles pas pleinement à notre compréhension du mot « littérature » ? Ainsi que nous l'avons vu, la littérature devrait nous apprendre quelque chose sur l'humain, une propriété centrale qui doit déroger aussi peu que possible à une catégorisation précise et exacte de son objet, selon la théorie de la vérité comme correspondance.

3. LE VERTIGE DE L'IA ET LE RÔLE DES SCIENCES HUMAINES

« ChatGPT » ou l'IA comme un « système conversationnel » ne sont pas des outils de conversation³, à moins d'entendre la conversation comme un jeu de questions-réponses⁴. L'outil est conçu « pour produire des réponses et générer du texte de qualité comparable à celle d'un humain » à partir de prompts, – d'instructions très précises, sous forme de liste d'injonctions spécifiques. De la qualité de la liste d'injonctions, c'est-à-dire des prompts, va dépendre la qualité des réponses. On trouve des modèles de listes de prompts déjà préparés pour différents types de tâches sur Internet. Par exemple : « Je veux que tu te comportes comme un rédacteur spécialisé [en littérature et en philosophie] » ; « Mon objectif est de rédiger un article de type [académique], [...] donne-moi une liste de thèmes généraux autour de [la création littéraire, la production de connaissances vraies et le risque de voir substituer le baratin à celles-ci]⁵ ».

À partir de cette liste d'injonctions, la machine produira des réponses ; elle « poursuit la conversation » ou « développe le sujet en question ». La nature mimétique de l'outil est basée sur des modèles de langage produisant des chaînes de pensées, l'outil a été formé à partir de l'ingénierie de prompts et le résultat ressemble à un raisonnement humain. Mais ce résultat n'est pas le produit d'un raisonnement. L'IA n'a pas appris à raisonner⁶. La réponse de l'IA procède par simplifications et schématisations, ce qui explique pourquoi le texte de ChatGPT n'est pas littéraire, il est toujours un cran en-dessous de la complexité humaine qui caractérise le texte littéraire.

Un sens littéraire ou une manière universitaire philosophique de disserter et d'élaborer une pensée reste à définir. À la fois dialectique dans sa manière d'équilibrer thèse et antithèse et soucieuse de former de bons raisonnements exempts de trivialité ou de prêt-à-penser fallacieux, la pensée philosophique est très exigeante. Il est indéniable que seules la lecture et l'écriture comme

pratiques permettent de reconstituer une représentation réaliste de l'activité philosophique comme forme de littérature. La pensée philosophique fait et continuera à produire des œuvres infiniment supérieures et inimitables par les chatbots. Pour penser le futur de nos pratiques, nous devrions admettre que l'intelligence artificielle aide à identifier l'objet de notre investigation, ou facilite certaines de nos tâches. Le néologisme « d'éducation augmentée » décrit un état de fait ; dans les domaines connexes à l'éthique et la philosophie, c'est aussi un horizon technique de réappropriation de l'art de l'écriture (Bellomo, 2023). L'espoir de pouvoir un jour « jouer comme sur un clavier sur le dos des choses »⁷.

Améliorer, générer, traduire, réduire en résumé... sont là des fonctions dont seule la seconde, celle de générer, entendue comme un inventeur littéraire ou scientifique, mérite une attention spéciale, puisque, pour les autres fonctions, il suffit qu'elles soient réalisées sous contrôle attentif, avec corrections de l'auteur et en transparence vis-à-vis du lecteur, afin de l'informer de ce qu'il lit. L'engagement créatif est très différent. Nous avons vu plus haut avec le roman de Boulgakov qu'il y a une concentration extraordinaire de qualités dans l'œuvre littéraire, qualités que le prompt, aussi suggestif soit-il, ne permet pas de rendre avec précision.

Si l'œuvre littéraire est définie, avec Engel, par la vérité, la connaissance et la beauté comme un *continuum*, celle-ci peut souligner le rôle de l'empathie, afin d'incorporer par son intermédiaire un savoir-faire pratique humain comme éthique du caractère (Engel, 2013). Largement basée sur *la mimesis*, on peut aussi entendre une œuvre vraie comme celle qui peint la réalité humaine comme action vertueuse, ou formation du caractère ; elle est alors inspirée des écrits des grands moralistes (de La Rochefoucauld, Montaigne).

Dans le genre de la comédie, l'œuvre peut vouloir porter une critique sociale ou une critique des manques de vertu ou des vices. Pensons au pastiche *Bouvard et Pécuchet* (1881) de Flaubert, où deux personnages passionnés par l'aventure de la connaissance,

manquent à l'application du savoir, mais émerveillent les lecteurs par le comique de répétition et par leur manque de jugement. Saisir l'intuition précise des actions à conduire requiert une habileté. Son manque forme une espèce de bêtise assez innocente, liée à notre propre complaisance dans l'ignorance pratique. Elle est distincte de la sottise, qui est élaborée par des gens réputés plus connaisseurs, mais auxquels manque une juste balance des choses. Engel écrit : « C'est la raison elle-même qui est conne. » La légèreté égoïste est manipulatrice et bricoleuse, elle peut faire souffrir autrui, ainsi que le montrent les atteintes à l'intégrité les plus graves, à la limite d'un crime contre la connaissance, documentées dans les trois précédents ouvrages de l'IRAFPA⁸.

Fraude et malhonnêteté ont leurs zones d'ombre qui préoccupent les praticiens de l'intégrité académique. Avant de devenir blâmable comme délit ou punissable comme crime, le vice de pensée et d'information peut faire rire, – et mérite la sanction sociale que constitue le rire ou cette stratégie d'éthique sociale et politique artistique qu'est la dénonciation théâtrale. De façon cocasse, une pratique sotte, qui semble stupide au plus grand nombre, n'empêche pas son auteur d'être fier. Cela a pour effet d'en propager le contenu, parce qu'aucun responsable ne semble désigné ; or les gens aiment désigner un coupable (Engel, 2014). L'information erronée, qui apparaît stupide parce qu'elle ne véhicule ni une croyance fondée sur une raison ou un raisonnement ni une application concrète ou éventuellement un principe, est un simple buzz nourri par le fait que les gens sont friands de nouvelles informations. Le vice intellectuel constitue un buzz parce que ce qui est stupide paraît aussi souvent nouveau, ce qui à son tour rend l'information contagieuse, puisque les gens ne s'y intéresseraient pas autrement ; c'est aussi parce qu'on adopte avec empressement ce qui distrait des choses sérieuses. Lorsque l'information se rapporte à un vice intellectuel, ceux qui la propagent, par une vanité friande de publicité, savent que la chose véhiculée n'est pas fondée, mais entreprennent quand même d'amplifier cette chose de manière ironique, sachant que la sottise reste un excellent véhicule de contagion de l'information.

4. CHATGPT, RÉCIT, LITTÉRATURE

S'il existe une valeur cognitive essentielle à toute œuvre littéraire, dans la mesure même où elle est de la littérature au sens de P. Engel, alors, par extension, une œuvre littéraire produite en partie par une machine comme ChatGPT doit aussi avoir une valeur cognitive démontrable, sans quoi elle n'engendrerait qu'une pure contre-valeur. Comme en grande partie l'intelligence artificielle n'est pas utile à autre chose qu'à améliorer le style de l'auteur, mais que précisément pour un cognitiviste littéraire est essentiel ce qui se manifeste par les qualités esthétiques de l'œuvre dans la valeur cognitive elle-même, et non un résidu de connotation qui viendrait en sus modifier une signification initiale, cette amélioration automatisée prendra, au mieux, une importance marginale. Former une connaissance de l'être humain dépend d'un ensemble de textes écrits fondateurs, même dans un monde où des acteurs ou des activités peuvent ne plus dépendre directement d'une exploration intellectuelle seulement humaine, mais sont en passe de devenir le produit de services automatisés, indépendants de l'être humain.

La littérature permet de rendre présente une activité qui est à la fois pratique, celle du producteur d'un texte écrit, et théorique, puisque l'accès à certains textes fondateurs, nous amène à une forme de vérité. Comme telle, cette vérité convoque la distinction entre la pratique et la théorie dans les activités liées à l'éducation. Des corpus de doctrine ou d'enseignement qui ne sont pas entièrement constitués de connaissances, ou de propositions qui ont une claire valeur de vérité, peuvent néanmoins avoir une valeur épistémique, comprenant la créativité, fécondité des interprétations ou *valeur herméneutique* (Engel, 2019, p. 351). Enfin ces textes peuvent viser offrir une valeur de sagesse ou *valeur eudémonique*, associée à la vie bonne.

Un auteur peut bien par son activité d'écriture développer l'estime de soi, de toutes les manières dont on définit cette activité, même avec force images poétiques. Lorsque le but recherché est la formation de connaissances avec une référence explicite à la vérité,

c'est-à-dire lorsqu'on privilégie les connaissances qui totalisent davantage de critères de vérité, on peut se dire que l'auteur rejoint le grand projet des premiers penseurs de la Modernité, celui qui conduit au développement des sciences humaines, aux théories sociales, économiques et psychologiques de l'humain. La portée métaphysique de ce projet est bien connue depuis Descartes et cette quête peut finalement se refermer sur une soif postmoderne de se considérer comme une totalité différente d'une autre personne, projet commencé comme métaphysique du sujet. Si, au terme d'une vie, il n'est pas possible d'être en adéquation parfaite avec la vérité comme telle, seule compte peut-être notre propre contingence. Richard Rorty (1986), inspiré par le livre de Nehamas, *Nietzsche, Life as Literature* (1985), l'a pensé. Pourquoi cette fixation sur la vérité dans la littérature ? Après tout, même si on veut sauver l'héritage de la Modernité, le constat s'impose : peu de textes de littérature font des références très explicites à la vérité.

Les limites de conditions d'expérience du caractère « littéraire », de la valeur de littérature d'un récit, lorsque l'IA est partiellement utilisée pour composer ce récit, rajoutent une difficulté non négligeable à un métier déjà très aérien, selon la manière dont on veut entendre le mot « littérature ». Le risque de plagiat existe, quelle que soit la forme du texte soumis. Celui-ci augmente chez de nombreux éditeurs, dont le travail se fait le plus souvent sans rencontrer l'auteur. La valeur cognitive de la littérature nous semble un critère qui présente l'avantage d'être transparent et clair, du moins dans son principe ; la valeur cognitive devrait nous permettre de dégonfler un peu les prétentions associées au mot littérature, pour le prendre vraiment au sérieux, comme science humaine rigoureuse.

N'est-il pas contradictoire de dire que certains récits ne peuvent être que des produits de l'esprit humain et de soutenir par ailleurs des expérimentations innovantes au moyens de ChatGPT ? À notre avis, si les activités de l'esprit humain ne sont pas « des sciences » dans le vieux sens du terme, d'une branche « du savoir-de » [*know-what*], mais des « disciplines », selon la proposition de Ryle, alors

les avancées des connaissances ne consistent pas en une accumulation de connaissances, en des « corps d'information », mais des « branches de savoir-comment », de [*knowledge-how*] pouvant comprendre l'usage d'outils d'IA (Ryle, 1945 ; Engel, 2015).

Examinons trois objections à la ligne argumentative qui prône l'inclusion de ChatGPT dans le travail d'écriture, alors qu'une charte de conduite traditionnelle, limiterait tout apport conçu « pour produire des réponses et générer du texte de qualité comparable à celle d'un humain » en plus de l'apport humain.

La première objection est celle de l'accommodation, fonction de survie sociale. Si l'être humain utilise des technologies de l'information en grand nombre (databases, smartphone, courriels, etc.), pourquoi ne pas utiliser largement des outils de génération de texte du même type que ChatGPT ? Ne pas le faire reviendrait à se murer derrière l'idée fausse que les innovations technologiques ne correspondent pas à la réalité économique humaine qui les soutient et qui, dans de larges mesures, commandent les transformations du mode de vie et des comportements humains. L'intégrité dans la génération de texte repose d'une part sur l'intention de l'utilisateur, non sur l'outil ; d'autre part elle repose sur la prise en compte des risques ou des dommages possibles à autrui. Selon Ryle, « reconnaître les maximes d'une pratique présuppose de savoir comment l'exercer en pratique ».

En effet, les œuvres publiées à caractère littéraire sont des vérités, des connaissances, des règles, qui « comme les oiseaux » devraient « vivre avant de pouvoir être empaillées ». Comme « savoir-comment », s'engager avec ChatGPT dans l'écriture d'un récit littéraire repose sur une maîtrise pondérée de la pratique en question, avant qu'on puisse en tirer de manière explicite des règles ou des principes qui régissent ces usages.

La deuxième objection est celle du biais des professionnels de l'information et de transmission de la connaissance – auteur, éditeur, bibliothécaire numérique ou formateur et enseignant : ces métiers sont, à leur manière, directement responsables (éventuellement

coupables) d'encourager à changer les pratiques, car comment les exercer autrement que de manière professionnelle ? Il est banal de dire que les professionnels de l'information, qui y ont un intérêt existentiel, encouragent et éventuellement favorisent ces usages (techniques). Chacun réalise individuellement l'impact négatif de ces technologies sur l'économie de l'attention, mais vivre avec son temps revient à s'accommoder et à comprendre. Le credo technosolutionniste oublie les efforts qu'il convient d'investir pour former à la connaissance de l'utilité et aussi des limites de ces techniques.

La troisième objection est celle de ChatGPT autour des métiers de la création de la connaissance humaine et littéraire. Travailler dans un métier de la connaissance demande un exercice quotidien de qualités humaines, que cette technologie ChatGPT n'offre pas toute seule, car le connaître-quoi du ChatGPT, « implique » toujours « dans un sandwich », aussi « un connaître-comment » du ChatGPT, et non l'inverse. Le connaître-quoi de la littérature implique toujours un connaître-comment.

5. UNE ÉTUDE DE CAS EN EXEMPLE

Un cas d'édition hybride avec ChatGPT a été réalisé dans un projet d'autobiographie, *Sharing Values* (2024) où l'auteur (Stückelberger) et l'éditeur (Green) adaptent le récit autobiographique, en utilisant une IA (ChatGPT 3.5) pour composer un récit autobiographique « amélioré » à partir d'une interview. La question centrale pour un éditeur est de définir si l'auctorialité est intégralement préservée, et si la proposition d'intention est partagée, en sachant qu'il y a deux personnes à l'origine de l'œuvre.

Admettre que l'auteur utilise de manière responsable des moyens de transformer le matériau initial peut faire partie du projet éditorial qui veut montrer la richesse créative, éthique et organisationnelle d'une œuvre : dans ce cas, il faut mentionner clairement le recours à ChatGPT sur la page de titre, à côté du nom des auteur et éditeur, ce qui ici a été réalisé consciencieusement.

Si nous pensons, selon la ligne argumentative du cognitivisme littéraire développée ci-dessus, que le style d'un auteur est essentiellement un faire qui ne dépend pas d'une collaboration avec ChatGPT et si, par conséquent, le style de l'auteur, son *savoir-comment écrire* est absent de son œuvre, où on découvre essentiellement le produit d'un texte généré par intelligence artificielle, alors nous ne trouvons pas la démonstration pratique du savoir-faire de l'auteur.

Une argumentation intermédiaire pourrait soutenir que, si la contribution de ChatGPT est vérifiée par l'auteur et que l'auteur a bénéficié de la rapidité, de l'économie de moyens et de l'efficacité fonctionnelle incontestable de l'outil, alors le texte généré par ChatGPT a une certaine dignité en tant que produit coorganisé par l'esprit humain.

Le choix d'un auteur ou d'un producteur de livre peut délibérément consister à anticiper une série d'expérimentations avec la technologie. En effet, plus qu'une question de gestion de l'information et de la connaissance, nous sommes au défi de préserver la confiance en une société ouverte autour d'un principe de dialogue et de partage des valeurs, au défi de vivre humainement en résistant à la crainte et aux conflits qui séparent les communautés humaines et renforcent les inégalités existantes.

Par ailleurs, « le prête-plume » ou écrivain fantôme, [...] est l'auteur « sous-traitant » et anonyme d'un texte signé par une autre personne. On trouve déjà ce procédé dans la Bible, les personnalités d'État, les People et certains écrivains à succès y ont recours. Au contraire, l'intégrité de l'œuvre littéraire et l'intégrité du binôme guide expérimenté/apprenant, mentor/mentoré, suppose des limites précises, une direction attentive et sage, pour que l'action d'imitation fondée sur des raisons ou sur un raisonnement, voire sur le projet d'inspirer une œuvre créatrice poétique vraie, fonctionne véritablement entre plusieurs personnes. Le prête-plume utilisé n'est pas toujours un rêveur éveillé, qui embrouille les frontières de la production littéraire honnête, à la façon du sot de Molière ou de Flaubert, qui est toujours une personne éduquée, mais qui

n'applique pas convenablement son savoir dans son rapport à autrui. Ce dernier cas de figure est celui qui nous concerne, lorsque nous pensons à la délinquance académique, qui n'est pas le fait seulement de jeunes apprentis-chercheurs oubliant de citer leurs sources.

Il existe des détecteurs de contenus générés par ChatGPT. Au mieux, ces détecteurs montrent qu'on a utilisé ChatGPT, mais ils n'indiquent pas avec précision de quelle manière quelle partie du texte a été générée, ils offrent seulement un pourcentage approximatif du total de texte généré artificiellement. Selon R. Larsen, le moyen (utilisable en cas de litige juridique, car très précis) de détecter ChatGPT de manière fiable est d'opérer comme on fait avec un prête-plume (*ghostwriter*)⁹. Dans un entretien (ICDELS, Genève, juin 2024), R. Larsen dévoile le principe de deux tests (il s'agit pour l'heure d'un logiciel en phase de développement). Pour le premier test, on demande à l'auteur d'écrire une courte composition (une page et demie) sans Internet et dans un environnement surveillé ou, si le test se déroule à la maison, avec la caméra du laptop qui filme tout ce que fait la personne pendant la durée de l'épreuve. Le second test consiste à soumettre à titre d'objet d'analyse principal un mémoire d'une longueur qui varie entre huit et deux cents pages. Les deux documents sont comparés au moyen d'outils élaborés par des linguistes à partir d'un corpus de textes fourni par la police scientifique des textes numériques par le moyen de calcul de prédictions informatiques élaborés par l'entreprise Nor.Education : ces outils mesurent les temps de réponse pour définir des repères temporels dans les réponses, jalons spécifiques à l'auctorialité de chaque individu. Pour décider ce qui est de l'auteur et ce qui a été produit par la machine, ces différents tests doivent offrir un résultat cohérent.

6. CONCLUSION

Dans sa *Philosophie du comme si* (1924), H. Vaihinger écrivait que « l'apparence, le consciemment-faux joue un rôle fondamental dans

la science, dans la philosophie et dans la vie ». C'est dans un sens différent que nous interprétons l'usage mimétique du *faire comme si*, de l'imitation par une sorte d'empathie des comportements et des actions d'autrui, afin d'incorporer par une discipline pratique un savoir-faire pratique ; le *faire comme si* est un jeu d'apparence qui ne doit être ni « consciemment faux », ni consciemment vrai dans un premier temps... Avant de devenir l'auteur acceptable d'un court essai littéraire de quatre pages, il est indéniable qu'il faut mille fois sur le métier remettre l'ouvrage ! Aucune liste exhaustive de conditions théoriques ne suffit à former à l'habileté nécessaire à l'écriture d'un texte.

Afin d'ériger la vérité et la connaissance comme critères de l'œuvre littéraire et de garder l'idée qu'il y a un essentiel savoir-comment, nous devons remplir certaines conditions. À la suite d'Engel et de Landy (2014, 2012), on retiendra deux pistes principales : la première veut que les fictions vraies soient celles qui nous exercent à une pensée rationnelle et favorisent la paix intérieure. La seconde répond à la question « qu'est-ce qui fait d'une fiction une œuvre littéraire ? » par un ensemble d'exercices destinés à *fortifier les capacités mentales* du lecteur (Landy, 2012 ; notre italique). Il s'agit de constituer un « terrain de jeu » pour le développement de nos facultés, et pas nécessairement de nous informer et de nous améliorer moralement, ce qui ne veut pas encore dire qu'on consolide sciemment le faux. On distingue ici deux choses : apprendre à se conduire dans la pratique grâce à la littérature, ou bien il s'agit de maîtriser des pratiques littéraires, l'une comme l'autre rejoint un concept de la psychologie du flux : il ne s'agit pas de perdre l'apprenant dans un labyrinthe, mais de développer l'intersection logique entre un élément de jeu, la simulation et l'apprentissage. L'intersection logique entre apprentissage et simulation se situe sur le terrain de l'exercice des disciplines, celle entre jeu et apprentissage, forme des applications éducatives non formelles et divertissantes. Enfin comme simulation et comme élément de jeu, ce sont des jeux de simulation, y compris pour nous, littéraires ouverts à la technologie¹⁰.

Nous espérons avoir montré qu'il n'est pas contradictoire de dire que certains récits ne peuvent être que des produits de l'esprit humain, et de soutenir par ailleurs des expérimentations ChatGPT qui ajoutent un ressort supplémentaire à l'exercice favorisant la lecture et l'écriture littéraire.

Références bibliographiques

- Bellomo, S. T. (2023). *Educación aumentada Desafíos de la educación ne la era de la inteligencia artificial*. Globethics Publications.
- Benda, J. (1945). *De la Littérature pure. Essai d'une psychologie originelle du littéraire*. Gallimard.
- Boulgakov, M. (2004). *Le Maître et Marguerite*, dans *Œuvres II*. Gallimard, Coll. « Bibliothèque de la Pléiade ».
- Csikszentmihalyi, M. (1996). *Creativity: Flow and the psychology of discovery and invention*. Harper Perennial.
- Engel, P. (2013). Trois conceptions de la connaissance littéraire : cognitive, affective, pratique. *Philosophiques*, 40(1), 121-138. <https://doi.org/10.7202/1018380ar>
- Engel, P. (2014). L'avenir du crétinisme. Dans A. Reboul (dir.), *Mind, Values, and Metaphysics* (pp. 135-148). Springer.
- Engel, P. (25 mars 2015). Littérature et connaissance pratique (conférence au Collège de France). [Vidéo] <https://www.youtube.com/watch?v=fOAt6OMDXqA> (URL vérifié 19 oct. 2024).
- Engel, P. (2019). *Les vices du savoir*. Agone.
- Inclusive AI for a Better Future. Policy Dialogue Report* (2024). Globethics Publications.
- Landy, J. (2012). *How to do things with fictions*. Oxford University Press.
- Larsen, R. (2024). Nor.Education. <https://nor.education>
- Mulligan, K. (2014). Foolishness, stupidity, and cognitive values. *The Monist*, 97(1), 66-85.
- Navigating the future of higher education with AI: 2024 ICDE Leadership Summit Policy Dialogue Report* (2024). Geneva. Globethics Publications, <https://doi.org/10.58863/20.500.12424/4316083>
- Nehamas, A. (1985). *Nietzsche, life as literature*. Harvard University Press.
- Nietzsche, F. (1873). *De la vérité et du mensonge au sens extra-moral* (nouvelle trad. par P. Wotling, appareil pédag. B. Benoit, 2023). GF-Flammarion.
- Roland, J. (1958). On "Knowing How" and "Knowing That." *The Philosophical Review*, 67(3), 379-388. Rorty, R. (8 May 1986). The contingency of selfhood. *London Review of Books*, XXIII(8), 11-15.

Ryle, G. (1945). Knowing how and knowing that: The presidential address. *Proceedings of the Aristotelian Society*, 46, 1-16.

Ryle, G. (1949). *The concept of mind*. Hutchinson.

Stückelberger, C. (2024). *Sharing values across borders. My story with Globethics* (édité avec Erin Green). Globethics Publications.

Vaihinger, H. (1924). *The philosophy of 'As if', A system of the theoretical, practical and religious fictions of mankind* (traduit par C. K. Ogden). Harcourt, Brace & Co.

1. Immanuel Kant, dans sa *Critique de la Faculté de jugement* (1790) distingue un jugement déterminant, qu'il met de côté, d'un jugement réfléchissant, propre à l'expérience du beau dans l'œuvre. Ce jugement n'est pas le produit d'une loi logique, mais une spéculation de l'esprit ; ce n'est ni un phénomène seulement psychologique, ni une référence à des Idées comme essences dans un monde idéal, mais une tentative cognitive de combler l'écart vers l'Idéal par les ressources (en grande partie innées) propres à l'esprit.
2. Nous sommes très redevables aux concepts forgés par Pascal Engel dans *Littérature et connaissance pratique* (2015).
3. Une condition nécessaire à la conversation que ne remplit pas ChatGPT est le manque de protocoles de reconnaissance non-verbale.
4. Un jeu de messages écrits ne rendra jamais les micro-messages positifs non-verbaux d'un dialogue les allô, hein, sans lesquels il est très difficile d'échanger de manière fonctionnelle et économique.
5. Nous avons adapté ici « Le guide et la liste ultime de prompts ChatGPT » trouvé sur nexus.fr.
6. Dans une conversation qui s'appuie sur un raisonnement, il ne suffit pas que la machine réponde, mais il faut aussi qu'elle articule son raisonnement : $2 \times 2 + 1 = 5$, car *on opère en priorité la multiplication sur l'addition*.
7. Nietzsche, F. « Ns Schreibmaschine », 1882, *Werke, Kritische Studienausgabe* 1, p. 876.
8. Voir Research Ethics No. 1, 2 et 3 publiés par Globethics en collaboration avec IRAFPA.
9. Conversation très intéressante avec M. R. Larsen, lors de la conférence *ICDE Leadership Summit*, codirigée par Globethics et UNESCOICHEI : Ethical Leadership in the Age of AI: Rethinking Futures of Education. Genève, juin 2024. Nor.Education, Oslo, <https://nor.education>.
10. On pensera ici au travail de psychologie de M. Csikszentmihalyi (1996).

CHAPITRE 7. INTÉGRITÉ ET CONFLIT D'INTÉRÊTS : FINANCEMENT DE LA RECHERCHE PUBLIÉE EN LIBRE ACCÈS ET DANS LES REVUES HYBRIDES ? LE CAS DE LA CHIRURGIE CARDIAQUE

[Delphine Szeceł](#), CHU de Liège, Université de Liège (Belgique)

[Tom Melvin](#), Trinity College, Université de Dublin (Irlande)

[Wouter Oosterlinck](#), University Hospitals Leuven (Belgique)

1. INTRODUCTION

Dans le modèle de publication traditionnel, les coûts de publication dans une revue sont généralement supportés par le lecteur *via* un abonnement, souvent par l'intermédiaire des bibliothèques institutionnelles ou par le biais de réimpressions. Tandis que dans un système de publication en libre accès, l'accès à l'article est gratuit et la publication est financée par l'auteur, qui paye les frais de traitement de l'article (*article processing charges*, désormais abrégé en APC). Plus rarement, ces coûts sont pris en charge par une société scientifique agissant en tant qu'éditeur. Dans un pays comme l'Irlande, un accord financier (IReL) a été négocié entre les universités et les éditeurs.

En outre, certains éditeurs offrent aux auteurs la possibilité de publier leurs articles en accès libre dans des revues traditionnelles moyennant des frais supplémentaires. Il s'agit d'un système hybride. En 2023, ce modèle hybride était proposé par 64,5 % des revues de chirurgie cardiothoracique (données non publiées).

Notre objectif est de donner un aperçu de cette évolution afin de soulever une question centrale puisqu'elle a trait à l'intégrité des chercheurs et de leurs publications : si les recherches publiées dans des revues en libre accès et dans des revues payantes n'ont pas les mêmes sources de financement, les liens d'intérêts et les conflits

d'intérêts sont-ils pris en compte de la même manière ? Cette question sera explorée par l'analyse des publications de deux revues majeures dans notre domaine de recherche.

2. PUBLICATION EN LIBRE ACCÈS, UN MODÈLE LOIN D'ÊTRE PARFAIT

Le passage d'un modèle de financement par les lecteurs à un modèle de financement par les auteurs a plusieurs effets négatifs qui ont été décrits avec éloquence dans le dernier rapport du CNRS sur cette question. Alain Schuhl, directeur général délégué à la recherche au CNRS, décrit les inconvénients du système de publication hybride : l'inégalité de publication remplace l'inégalité d'accès ; le fait que les chercheurs doivent payer pour être publiés jette éventuellement le discrédit sur leur recherche ; on encourage les publications inutiles, le développement de revues prédatrices et l'inflation des coûts, les institutions, dans le cas des revues hybrides, payant à la fois pour la lecture et la publication (Schuhl, 2022). Nous pouvons ajouter les nouveaux risques créés par l'IA générative : manque d'originalité du contenu, auctorialité sujette à caution, révision par l'IA plutôt que par les pairs, vitesse exponentielle de la génération de texte, etc.

Définitions de la science ouverte

De nombreuses définitions de la science ouverte coexistent :

- La science ouverte est définie par le ministère français de l'Enseignement supérieur et de la Recherche comme la diffusion sans restriction des publications et des données de recherche. Elle tire parti des possibilités offertes par la transformation numérique pour favoriser le libre accès aux publications et, dans la mesure du possible, aux données de recherche. Cette définition est plus restrictive que celle de l'OCDE, pour qui la diffusion n'est qu'un aspect (Paic, 2021). Selon l'OCDE, les trois principaux aspects de la science ouverte sont l'accès ouvert, les données de recherche ouvertes et la collaboration ouverte rendue possible par les technologies de l'information et de la communication.
- Système de science ouverte (*open science*) : l'OCDE va au-delà de ces trois aspects avec « l'examen par les pairs après publication, les carnets de recherche ouverts, l'accès libre aux matériaux de recherche, les logiciels libres, la science citoyenne et le financement participatif de la recherche » (Paic, 2021).

- Le libre accès (*open access*) a été défini comme la disponibilité en ligne libre et gratuite de la documentation scientifique par l'Open Access Initiative de Budapest en 2002 (Open Society Institute, 2002). Il s'oppose au modèle traditionnel de lecture payante.
- La voie verte désigne la mise à disposition des travaux publiés dans une archive ouverte ou dans un dépôt thématique ou institutionnel.
- La voie dorée est la publication dans des revues en libre accès impliquant parfois le paiement de frais de traitement des articles (APC).
- Dans la voie diamant ou platine, les revues en libre accès sont définies par Science Europe comme des « initiatives de publication menées par la communauté, dirigées par les universitaires et appartenant aux universitaires ». Néanmoins, on peut en donner une définition plus large, c'est « un sous-ensemble de la voie dorée, dans lequel les revues ne font pas payer l'auteur parce qu'elles reçoivent d'autres sources de financement » sans mention d'un caractère universitaire (*Open access publication routes*, University of Oxford).

La voie verte n'implique aucun coût supplémentaire pour le chercheur. Néanmoins, la publication ouverte dans un registre peut potentiellement entrer en conflit avec le transfert des droits d'auteur à l'éditeur au cours du processus de publication. Malgré un cadre juridique régional ou national¹, il existe une incertitude quant à la juridiction compétente en cas de litige, la plupart des éditeurs étant des sociétés basées à l'étranger.

De façon générale, ces exigences augmentent la complexité administrative à laquelle les chercheurs doivent faire face. Elles nécessitent un bagage juridique parfois très éloigné de leur domaine de recherche. Non moins importante est l'impossibilité de comptabiliser correctement le montant total des fonds consacrés au processus de publication, différentes sources de financement étant couramment utilisées à cette fin dans les laboratoires. Enfin, le terme « libre accès » en lui-même implique une sorte d'opposition à un ancien modèle qui aurait été obscur et fermé.

Au-delà de ces obstacles, peu de données sont disponibles sur l'évolution du contenu ou de la qualité des publications en fonction des différents modèles de publication, en particulier dans un souci d'intégrité académique. Nous avons voulu explorer le lien entre financement de la recherche et modèle de publication. Notre objectif

est de recueillir des données sur l'écosystème de la publication et son évolution. En d'autres termes, les sources de financement de la recherche publiée dans des revues dotées de systèmes de publication hybrides ou ouverts sont-elles sensiblement différentes ?

Au sein de certaines sociétés savantes, il existe un système de soumission automatique en cascade en cas de rejet du manuscrit (Paltani-Sargologos, 2022). Nous avons observé que dans cette cascade, lorsque les auteurs choisissent une revue traditionnelle (ou hybride) et que le manuscrit est rejeté, la proposition de soumission à une autre revue éditée par la même société scientifique implique généralement la soumission à une revue en accès libre, qui a un facteur d'impact plus faible et facture des frais de publication. Le mécanisme inverse, quand un manuscrit soumis par des auteurs à une revue en accès libre est repris par les éditeurs pour être proposé à la revue traditionnelle correspondante, pourrait exister, mais nous ne l'avons pas rencontré.

La publication scientifique est un axe central de la communication scientifique, elle fait le lien entre la recherche menée par les scientifiques, la communauté scientifique à laquelle ils appartiennent et éventuellement le grand public, grâce aux efforts de vulgarisation. Dans certains cas, le monde politique peut s'appuyer sur ces publications, pour élaborer des politiques de santé publique par exemple. Les publications scientifiques sont en outre prises en compte dans le recrutement et l'évolution de la carrière des chercheurs. Il est donc crucial que cette étape de la recherche soit aussi intègre que possible. Dans cette optique, le risque de partialité doit être rigoureusement examiné à chaque étape.

De nombreux chercheurs en milieu universitaire sont désormais soumis à des réglementations exigeant le dépôt de toutes leurs publications dans des archives ouvertes (voie verte ou dorée), parfois avec des possibilités d'embargo inférieures à celles qu'imposent les éditeurs et avec des APC éventuellement proportionnels au facteur d'impact.

3. PRÉSENTATION DU CHAMP D'ÉTUDE

Suite à plusieurs scandales (celui des implants mammaires PIP notamment), un *Règlement européen relatif aux dispositifs médicaux* (*Regulation on medical devices*, MDR) a été publié en 2017 et mis en œuvre en 2021. Pour les dispositifs médicaux à haut risque (classe IIa, IIb et III), ce règlement impose une évaluation par un « organisme notifié » (*notified body*) des données fournies par le fabricant avant que le dispositif puisse être mis sur le marché. Dans le cas des dispositifs médicaux à haut risque, catégorie qui englobe de nombreux dispositifs utilisés en chirurgie cardiaque, l'organisme notifié peut consulter un panel d'experts pour évaluer les données cliniques fournies par le fabricant. L'organisme notifié et le groupe d'experts fondent leur évaluation de la sécurité et du bénéfice clinique sur la documentation technique soumise, ce qui désigne la littérature scientifique dans le rapport d'évaluation clinique. Bien qu'il existe une hiérarchie réglementaire des preuves cliniques pour les dispositifs précédemment approuvés (MDR, art.55, §1), elle n'est pas alignée sur les niveaux de preuves cliniques publiés. En outre, l'avis du groupe d'experts est consultatif.

La récente mise en œuvre du MDR vise à renforcer les exigences en matière de preuves cliniques. En vertu de ce règlement, les fabricants de dispositifs médicaux sont tenus de produire des données démontrant que le dispositif est sûr et efficace. Les fabricants doivent rassembler des « données cliniques » qui peuvent provenir d'études cliniques prospectives, mais aussi de la « littérature scientifique » ou de la « littérature scientifique examinée par les pairs » concernant soit le dispositif évalué, soit un autre dispositif dont l'« équivalence » au sens réglementaire peut être démontrée (art.2, §48, du RIM²). En tant que telle, la littérature publiée est donc une source majeure de données utilisée par les fabricants de dispositifs médicaux pour répondre à ces exigences.

La tendance actuelle favorise ou impose parfois la publication de résultats en libre accès. Cela signifie qu'il existe deux types de relations financières concernant l'utilisation des preuves cliniques

pour les dispositifs médicaux : soit l'organisme notifié est payé par le fabricant pour étudier la sécurité et l'efficacité de son produit ; soit les coûts des publications scientifiques afférentes sont payés par le fabricant. Nos deux questions de recherche concernant les risques pour l'intégrité académique sont les suivantes :

- Question 1 : Y a-t-il un risque à faciliter l'accès à la publication des travaux de recherche financés par ces fabricants ou rédigés par des personnes directement ou indirectement financées par ces mêmes fabricants ?

- Question 2 : Le modèle commercial (libre accès ou hybride) des deux revues de chirurgie cardiothoracique examinées est-il lié aux sources de financement de la recherche et au taux de conflit d'intérêts dans la recherche qu'elles publient ?

4. MÉTHODOLOGIE

Notre recherche a comparé les sources de financement et les conflits d'intérêts déclarés dans les articles publiés dans deux revues phares de l'Association Européenne de Chirurgie Cardio-Thoracique (EACTS). L'EACTS est la principale société professionnelle de chirurgie cardiaque en Europe et elle est impliquée dans la formation et la certification des chirurgiens ainsi que dans la rédaction des recommandations de pratique clinique. La revue *European Journal of Cardio-Thoracic Surgery* (EJCTS) fonctionne selon un modèle hybride (modèle d'abonnement traditionnel, bien que l'accès libre soit possible sur demande de l'auteur moyennant des frais supplémentaires), tandis que la revue *Interactive Cardiovascular and Thoracic Surgery* (ICVTS) fonctionne sur un modèle de libre accès payé par l'auteur. Nous avons examiné les articles publiés entre janvier et novembre 2022. Novembre 2022 a été choisi comme date de fin de la collecte des données, la revue ayant changé de nom à cette date. Les deux revues sont publiées par Oxford University Press.

Nous avons effectué une analyse exhaustive de tous les articles publiés au cours de cette période. Ont été inclus dans l'étude les

articles comprenant des recherches originales, des méta-analyses ou des revues de littérature. Les articles relatifs à la chirurgie congénitale ont été exclus, car il s'agit d'une activité de niche représentant moins de 6,4 % de toutes les procédures de chirurgie cardiaque (Vervoort *et al.*, 2024), et les dispositifs utilisés pour ces indications présentent en effet des défis spécifiques (Melvin *et al.*, 2023). En outre, les éditoriaux, les articles d'opinion et les rapports de cas ont été exclus, car ne correspondant pas à des recherches originales.

Conflit d'intérêts : définition et mécanisme en jeu³

Le conflit d'intérêts (CI) est défini comme « une situation dans laquelle les intérêts (financiers ou non) d'un individu peuvent affecter négativement ses performances professionnelles en biaisant son jugement ». Kafaei a théorisé les effets des CI et montré qu'ils provoquent des biais cognitifs. Il interprète les CI comme la possibilité que des conflits ou des alignements d'intérêts affectent inconsciemment la prise de décision (Kafaei *et al.*, 2022). En tant que tel, le CI est une situation, tandis que le biais induit est une conséquence potentielle, qui peut être présente ou non.

Les sources de financement de l'industrie ont été classées comme présentes ou absentes. Les liens d'intérêts potentiels et les conflits d'intérêts ont été repris tels qu'ils ont été déclarés par les auteurs conformément à la politique de la revue. La politique de la revue stipule que... « Au moment de la soumission, chaque auteur doit révéler tout intérêt ou lien financier, direct ou indirect, ou toute autre situation qui pourrait soulever la question de la partialité dans le travail rapporté ou les conclusions, implications ou opinions énoncées – y compris les sources de financement commerciales ou autres pertinentes pour l'auteur individuel ou pour le(s) département(s) ou organisation(s) associé(s), les relations personnelles ou la concurrence académique directe. Lorsque vous vous demandez si vous devez déclarer un conflit d'intérêts ou un lien, veuillez tenir compte du test de conflit d'intérêts : Existe-t-il un arrangement qui vous mettrait dans l'embarras, vous ou l'un de vos coauteurs, s'il apparaissait après la publication et que vous ne l'aviez pas déclaré ? » (*Journal policies*, EACTS)

IBM SPSS Statistics for Windows, version 29.0.2.0 (IBM Corp., Armonk, N.Y., USA) a été utilisé pour l'analyse statistique. Une valeur $p < 0,05$ a été considérée comme statistiquement significative. Les données catégorielles sont exprimées sous forme de nombres et de proportions. Le test du chi carré et le test exact de Fisher ont été utilisés pour les analyser.

L'intelligence artificielle générative, sous la forme de ChatGPT-4 (OpenAI)⁴, a été utilisée uniquement comme assistant linguistique pour réviser ou améliorer des phrases ou des fragments de phrases écrits par les auteurs. Aucune donnée n'a été fournie en entrée à ChatGPT. La traduction française du texte a été réalisé à l'aide de DeepL Translate⁵, puis relue et amendée par les auteurs.

5. LES RÉSULTATS DE LA RECHERCHE

Au cours de la période d'étude de 10 mois, 1 195 articles ont été publiés, dont 746 dans l'EJCTS et 449 dans l'ICVTS. Sur ces 1 195 articles, nous avons analysé les travaux de recherche originaux (selon la définition ci-dessus), qui représentaient un total de 561 articles (297 dans l'EJCTS et 264 dans l'ICVTS). La proportion de publications de recherche financées par l'industrie était significativement plus élevée dans l'EJCTS (9,4 %) que dans l'ICVTS (3,8 %) ($p = 0,008$). De même, la déclaration d'un CI potentiel était plus élevée dans l'EJCTS (30,1 %) que dans l'ICVTS (17 %) ($p < 0,001$). La déclaration des CI manquait dans un article (0,2 %).

Tableau 1. Taux de recherche financée par l'industrie et taux de conflits d'intérêts déclarés par revue

| | | | EJCTS (n = 297) | ICVTS (n = 264) | p |
|------------------------------------|--|--|-----------------|-----------------|---------|
| Recherche financée par l'industrie | | | 28 (9,4 %) | 10 (3,8 %) | 0,008 |
| CI déclarés | | | 89 (30,1 %) | 45 (17 %) | < 0,001 |

La disponibilité des déclarations de conflits d'intérêts était particulièrement élevée, une seule déclaration manquait. Le taux de conflits d'intérêts déclarés est conforme aux données publiées précédemment, selon lesquelles un quart des chercheurs ont un lien avec l'industrie (Bekelman *et al.*, 2003). Nos résultats montrent que les publications des recherches financées par l'industrie et avec des CI déclarés sont plus nombreuses dans l'EJCTS, publié selon un modèle traditionnel, que dans l'ICVTS, publié selon un modèle en accès libre. Il convient de noter la nature purement déclarative des sources de financement et les CI. Les CI ont été analysés par un seul auteur sans processus de vérification. Par ailleurs, notre étude se concentre uniquement sur l'aspect du financement des recherches publiées et ce, dans seulement deux revues d'une discipline spécifique. Le design de l'étude ne permet pas d'évaluer la qualité des publications. Enfin, les conflits d'intérêts décrits dans la politique de la revue sont vagues quant à l'inclusion des conflits d'intérêts non financiers. Ces limites offrent également des pistes pour de futures recherches.

6. DISCUSSION

Comme nous l'avons indiqué dans notre méthodologie, nous avons choisi de comparer deux revues de la même société savante (l'EACTS). Cette décision a été prise dans le contexte de ce que l'on appelle le « transfert de manuscrits au sein des publications de l'EACTS ». Selon la politique de la revue, les éditeurs peuvent offrir aux auteurs des manuscrits sélectionnés qui sont rejetés par l'EJCTS l'option de soumettre une version révisée de leur travail à la revue ICVTS. L'éditeur justifie cette politique par le nombre élevé de manuscrits soumis pour un nombre limité d'acceptations possibles dans l'EJCTS. Les facteurs d'impact des deux revues sont différents (3,4 en 2022 pour l'EJCTS contre 1,978 pour l'ICVTS). Il est intéressant de noter que le site web bibliographique Scimago identifie l'ICVTS comme la revue la plus « similaire » à l'EJCTS, avec une similarité de 95 %. Cette similitude est basée sur le pourcentage de publications référencées qui sont communes aux deux revues.

Cela conforte notre décision de les comparer. Bien que les deux revues soient similaires dans le type de travaux qu'elles publient, nous avons identifié des différences dans le financement de la recherche et dans la proportion d'articles dont l'auteur a déclaré un CI. Les raisons de ces différences ne sont pas évidentes. On peut faire l'hypothèse d'une méthodologie de meilleure qualité dans les études bénéficiant d'un financement plus important, ce qui faciliterait la publication dans l'EJCTS lors de l'évaluation par les pairs. La proportion plus élevée de CI déclarés pourrait, quant à elle, être expliquée par l'implication de chercheurs plus expérimentés : publiant dans une revue à facteur d'impact plus élevé, ils ont eu plus d'occasions d'établir des liens avec l'industrie.

6.1. Impact des sources de financement et des CI déclarés dans la littérature médicale

Il a été démontré que les études financées par l'industrie comparent plus fréquemment les nouveaux médicaments à des placebos plutôt qu'au meilleur traitement médical, ce qui conduit à une probabilité de résultat positif plus élevée (Djulbegovic *et al.*, 2000). Il a également été constaté que la déclaration d'un CI était associée de manière significative à des résultats favorables dans les études tant sur les médicaments (Cho, 1996) qu'en chirurgie (Crossley *et al.*, 2021 ; Dubin *et al.*, 2024). Toutefois, Cho n'a pas constaté de différences en termes de qualité et de pertinence clinique dans les études de médicaments (Cho, 1996). En outre, le lien entre CI déclarés et niveau de preuve n'est toujours pas concluant. Au-delà du biais potentiel, l'existence de conflits d'intérêts – en particulier lorsque ces derniers ne sont pas divulgués en temps utile – peut contribuer à la méfiance du public à l'égard de la science.

6.2. Innovations et conflits d'intérêts

Les CI sont inhérents aux relations entre les cliniciens, les universitaires et l'industrie. De nombreuses innovations médicales reposent sur la recherche et le développement menés par l'industrie.

Pour les médicaments, le test de nouveaux dispositifs et les essais de phase III nécessitent la collaboration entre fabricants et cliniciens pour l'implantation expérimentale ou le recrutement de patients. D'un point de vue financier, en 2012, le financement de la recherche médicale aux États-Unis provenait à 59 % du secteur privé, à 38 % du secteur public et à 4 % de fondations et d'organisations caritatives (Moses *et al.*, 2015).

Une analyse de 2016 du Tufts Center sur les médicaments novateurs des 25 dernières années a révélé que, si 54 % des étapes clés de la recherche fondamentale ont été franchies par le secteur public, 58 % des étapes menant à la découverte d'un nouveau médicament ont été franchies par le secteur privé. En outre, 81 % et 73 % des étapes liées à la production et au développement de médicaments ont été franchies par des entreprises privées (Chakravarthy *et al.*, 2016). Ces données montrent que, pour des raisons pratiques et en raison de la structure du financement de la recherche biomédicale, l'interdépendance entre les secteurs est inévitable. La gestion de ces conflits est donc essentielle pour minimiser les biais.

6.3. Lutte contre la non-déclaration des CI

En 2020, une étude transversale a montré que 92 % des revues de chirurgie avaient des politiques de déclaration des financements et 99 % des politiques de divulgation des conflits d'intérêts, mais qu'aucune ne proposait de procédures pour remédier à la non-déclaration ou à la sous-déclaration des financements. Il a été montré à maintes reprises que les conflits d'intérêts étaient sous-déclarés dans diverses spécialités chirurgicales (El Moheb *et al.*, 2021). Alors que les CI non divulgués peuvent être traqués aux États-Unis grâce à la base de données Open Payments, en Europe où ils dépendent des législations et bases de données nationales, les mécanismes de contrôle sont limités. Lors des conférences scientifiques, le manque de considération pour ces déclarations est également problématique, un tiers des présentateurs ne signalant pas leur CI, alors qu'il y a des enjeux liés à la lisibilité de ces

informations (Crawford *et al.*, 2020). Les plateformes de médias sociaux sont également touchées par la sous-déclaration (Helou *et al.*, 2023). Ceci témoigne du manque d'attention de la communauté scientifique pour cette question.

Au-delà des articles de recherche, l'élaboration des recommandations de pratique clinique, qui s'appuie sur ces publications, est également sujette à des conflits d'intérêts financiers (James, 2020 ; Moore *et al.*, 2020). À cet égard, même si nous n'avons pas pu vérifier l'exactitude des informations, les taux de déclaration sont plus que satisfaisants dans notre échantillon.

6.4. Éviter la partialité due aux conflits d'intérêts

Comme l'indique la littérature, la simple déclaration des conflits d'intérêts ne suffit pas à neutraliser les biais potentiels qu'ils peuvent entraîner (James, 2020 ; Cain & Mohin, 2020). La limite de ce mécanisme est la difficulté pour les lecteurs de quantifier l'importance de l'ajustement nécessaire, en particulier pour les lecteurs non-initiés (James, 2020). Par conséquent, des stratégies actives doivent être développées afin de minimiser les biais potentiels qui en résultent. Plusieurs démarches visant l'amélioration de la compréhension des CI et l'élaboration de réglementations externes ont été développées.

Contrairement à la croyance populaire, il a été démontré que le professionnalisme augmentait potentiellement le risque de partialité en raison d'une confiance excessive de l'individu dans son impartialité. Pour résoudre ce problème, on a développé le concept de « professionnalisme profond », qui inclut la nécessité de « comprendre les limites de l'autorégulation » et de « pratiquer de manière cohérente des comportements qui incarnent cette compréhension » (Sah, 2024).

En outre, les leçons tirées par les intervenants du secteur sont précieuses. À cet égard, le retour d'expérience des rédacteurs d'*Arteriosclerosis, Thrombosis et vascular biology* est intéressant (Schmidt *et al.*, 2023). En informant les auteurs invités à écrire un

article sur leurs préoccupations relatives aux conflits d'intérêts, les rédacteurs ont apporté plusieurs modifications enrichissant au cours du processus de révision, l'article initialement soumis. Ceci souligne l'importance d'une guidance explicite. Il faut également aller au-delà de l'aspect comptable des informations sur les CI et souligner l'importance pour un chercheur d'expliquer comment il est parvenu à sa conclusion.

Concernant la rédaction des recommandations scientifiques, la participation d'au moins un méthodologiste au comité de rédaction et l'attribution de responsabilités différenciées en fonction de l'importance des conflits d'intérêts pourraient améliorer le choix des preuves (Guyatt, 2010). D'une façon plus générale, il en est de même pour les relecteurs externes (Samsa & Solomon, 2019).

La science participative est l'un des piliers de la science ouverte. Inviter des patients à cosigner un article est désormais reconnu comme un moyen d'enrichir les perspectives d'une publication médicale. Toutefois, l'existence de CI concerne également les patients et les organisations de patients (Bruno & Rose, 2019).

7. CONCLUSION

Le changement de modèle de financement des plateformes de publication et d'édition est à l'origine de profondes modifications dans l'écosystème de la publication. Cet écosystème est actuellement en pleine transition. À ce stade, l'impact de cette évolution sur les choix faits par les auteurs et les éditeurs ainsi que sur la qualité de la production scientifique a été peu étudié. Nous montrons que les articles publiés dans deux revues différentes de la même société scientifique et publiés sous deux systèmes de publication (hybride et ouvert) ont des sources de financement significativement différentes. Ceci n'implique aucun jugement sur la qualité attendue.

Par ailleurs, la science ouverte ne se limite pas à la publication en libre accès. D'autres méthodologies doivent être développées pour partager les données et les informations pertinentes. Dans le

contexte de la certification des dispositifs médicaux, des outils supplémentaires, tels que des registres financés par les autorités publiques et complétés par les utilisateurs (professionnels de la santé et patients), peuvent être utiles. Enfin, dans le domaine de la chirurgie cardiaque, très peu de revues indexées par la National Library of Medicine (la bibliothèque nationale de médecine des États-Unis) sont publiées selon la voie diamant. Dans un contexte universitaire où, malgré l'accord de la CoARA (Coalition pour l'Avancement de l'Évaluation de la Recherche), la valeur d'un chercheur est encore souvent déterminée par la métrique des citations, ce qui place les chercheurs face à un dilemme entre publications en accès ouvert et facteur d'impact. Malgré les initiatives émergentes, les alternatives sont actuellement sous-développées. Toutes les parties prenantes doivent unir leurs forces pour créer l'environnement dans lequel le libre accès garantira l'intégrité scientifique et la recherche innovante promue et mise à la disposition des décideurs réglementaires et politiques. La science ouverte n'en est qu'à ses débuts et il reste encore du chemin à parcourir pour qu'elle soit également garantie d'une amélioration de la qualité de la recherche.

Pour la recherche médicale, les données relatives à l'impact de la publication en accès libre sur la qualité des travaux sont rares, ce qui doit faire l'objet d'une étude plus approfondie. En effet, nous avons montré que le financement et le taux de CI étaient différents dans notre recherche exploratoire, mais notre méthodologie ne nous a pas permis d'interpréter davantage nos résultats. Comparer les études à résultats positifs publiées dans les revues en libre accès et dans les revues traditionnelles ou hybrides permettrait tout d'abord de vérifier si l'accès libre contribue effectivement à réduire les biais de publication. Une deuxième tâche consisterait à répliquer la méthodologie de Gardener *et al.* (2022) pour examiner dans quelle mesure les deux revues respectent les procédures de déclaration de conflit d'intérêts. Une troisième tâche consisterait à analyser le niveau de preuve atteint par chacune des deux revues. Cela permettrait une compréhension plus qualitative du fonctionnement

des deux revues. À une échelle plus globale, une comparaison des politiques des différentes sociétés scientifiques en matière de conflits d'intérêts permettrait de mieux comprendre ce qu'on attend de l'obligation de les déclarer.

L'évolution vers la publication en libre accès est une tendance profonde et nous assistons à un véritable changement de paradigme dans la publication académique. Les nombreuses conséquences d'un tel changement n'ont pas encore été étudiées en profondeur. Notre objectif était de contribuer à une meilleure compréhension de ce changement et en particulier du lien entre les modes de financement de la recherche et les modes de publication.

Références bibliographiques

Bekelman, J. E., Li, Y. & Gross, C. P. (2003). Scope and impact of financial conflicts of interest in biomedical research: A systematic review. *JAMA*, 289(4), 454. <https://doi.org/10.1001/jama.289.4.454>

Bruno, B. & Rose, S. (2019). Patient organizations and conflict of interest. *BMJ*, l129. <https://doi.org/10.1136/bmj.l129>

Cain, D. & Mohin, B. (2020). Do conflict of interest disclosures facilitate public trust? *AMA Journal of Ethics*, 22(3), 232-238. 10.1001/amajethics.2020.232

Chakravarthy, R., Cotter, K., DiMasi, J., Milne, C.-P. & Wendel, N. (2016). Public – and private-sector contributions to the research and development of the most transformational drugs in the past 25 years: From theory to therapy. *Therapeutic Innovation & Regulatory Science*, 50(6), 759-768. <https://doi.org/10.1177/2168479016648730>

Cho, M. K. (1996). The quality of drug studies. Published in Symposium Proceedings. *Annals of Internal Medicine*, 124(5), 485. <https://doi.org/10.7326/0003-4819-124-5-199603010-00004>

Commission européenne – Groupe de coordination des dispositifs médicaux (2020). *Preuves cliniques nécessaires pour les dispositifs médicaux précédemment marqués CE en vertu des directives 93/42/CEE ou 90/385/CEE Guide pour les fabricants et les organismes notifiés.* (Règlement (UE) 2017/745). https://health.ec.europa.eu/system/files/2020-09/md_mdcg_2020_6_guidance_sufficient_clinical_evidence_en_0.pdf

Crawford, W., Camm, C. F., Prachee, I., Olivarius-McAllister, J., Ginks, M. R. & Nicol, E. D. (2020). Are conflict of interest declarations appropriate to allow sufficient consideration of potential bias in presentations? *Future Healthcare Journal*, 7(3), 226-229. <https://doi.org/10.7861/fhj.2020-0018>

Crossley, J. R., Wallerius, K., Hoa, M., Davidson, B. & Giurintano, J. P. (2021). Association between conflict of interest and published position on hypoglossal nerve

stimulation for sleep apnea. *Otolaryngology-Head and Neck Surgery*, 165(2), 375-380. <https://doi.org/10.1177/0194599820982914>

Djulbegovic, B., Lacey, M., Cantor, A., Fields, K. K., Bennett, C. L., Adams, J. R., Kuderer, N. M. & Lyman, G. H. (2000). The uncertainty principle and industry-sponsored research. *The Lancet*, 356(9230), 635-638. [https://doi.org/10.1016/S0140-6736\(00\)02605-2](https://doi.org/10.1016/S0140-6736(00)02605-2)

Dubin, J. A., Hameed, D., Baksh, N., Bains, S. S., Mont, M. A., Nace, J., Delanois, R. E. & Golladay, G. (2024). Impact of reporting bias, conflict of interest, and funding sources on quality of orthopaedic research. *The Journal of Arthroplasty*, 39(5), 1348-1352. <https://doi.org/10.1016/j.arth.2023.11.017>

EACTS (n.d.-a). *Highly Cited Articles from ICVTS*. Consulté le 8 septembre 2024 sur https://academic.oup.com/icvts/pages/highly_cited

EACTS (n.d.-b). *EACTS – Journal Policies*. Consulté le 8 septembre 2024 sur https://academic.oup.com/ejcts/pages/Journal_Policies

El Moheb, M., Karam, B. S., Assi, L., Armache, M., Khamis, A. M. & Akl, E. A. (2021). The policies for the disclosure of funding and conflict of interest in surgery journals: A cross-sectional survey. *World Journal of Surgery*, 45(1), 97-108. <https://doi.org/10.1007/s00268-020-05771-0>

Gardener, A. D., Hick, E. J., Jacklin, C., Tan, G., Cashin, A. G., Lee, H., Nunan, D., Toomey, E. C., & Richards, G. C. (2022). Open science and conflict of interest policies of medical and health sciences journals before and during the COVID-19 pandemic: A repeat cross-sectional study: Open science policies of medical journals. *JRSM Open*, 13(11), 205427042211321. <https://doi.org/10.1177/20542704221132139>

Guyatt, G. (2010). The vexing problem of guidelines and conflict of interest : A potential solution. *Annals of Internal Medicine*, 152(11), 738. <https://doi.org/10.7326/0003-4819-152-11-201006010-00254>

Helou, V., Mouzahem, F., Makarem, A., Noureldine, H. A., El-Khoury, R., Al Oweini, D., Halak, R., Hneiny, L., Khabisa, J. & Akl, E. A. (2023). Conflict of interest and funding in health communication on social media: A systematic review. *BMJ Open*, 13(8), e072258. <https://doi.org/10.1136/bmjopen-2023-072258>

James, J. E. (2020). Disclosing conflict of interest does not mitigate healthcare bias and harm: It is time to sever industry ties. *European Journal of Clinical Investigation*, 50(12), e13344. <https://doi.org/10.1111/eci.13344>

Kafaei, M., Kheirkhah, M. T., Balali, R. & Gharibzadeh, S. (2022). Conflict of interest as a cognitive bias. *Accountability in Research*, 29(6), 379-396. <https://doi.org/10.1080/08989621.2021.1938556>

Melvin, T., Kenny, D., Gewillig, M. & Fraser, A. (2023). Orphan medical devices and pediatric cardiology. What interventionists in Europe need to know, and what needs to be done. *Pediatric Cardiology*, 44(2), 271-279. <https://doi.org/10.1007/s00246-022-03029-1>

Moore, A., Straus, S., Lexchin, J. & Thoms, B. D. (2020). Financial conflict of interest among clinical practice guideline-producing organisations. *British Journal of General Practice*, 70(700), 530-531. <https://doi.org/10.3399/bjgp20X713177>

Moses, H., Matheson, D. H. M., Cairns-Smith, S., George, B. P., Palisch, C. & Dorsey, E. R. (2015). The anatomy of medical research: US and international comparisons. *JAMA*, 313(2), 174. <https://doi.org/10.1001/jama.2014.15939>

Open Society Institute (2022, 14 février). *Déclaration de Budapest, BOAI (Budapest Open Access Initiative)*. <https://www.budapestopenaccessinitiative.org/read/>

Oxford University (n.d.). *Open access publication routes*. Consulté le 5 avril 2024 sur <https://openaccess.ox.ac.uk/oa-publication>

Paic, A. (2021). *Science ouverte. Favoriser la découverte à l'ère numérique*. OCDE. <https://doi.org/10.1787/81a9dcf0-en>

Paltani-Sargologos, I. (2022). *Paysage de l'édition scientifique, pratiques des scientifiques et coûts associés*. Direction des Données Ouvertes de la Recherche. CNRS. <https://www.scienceouverte.cnrs.fr/wp-content/uploads/2022/09/Rapport-DDOR-CNRS-Paysage-de-ledition-scientifique-au-CNRS-Printemps-2022.pdf>

Règlement EUM 45/2017. *Règlement sur les dispositifs médicaux*. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017R0745>

Sah, S. (2024). Deep professionalism: Charting a path for effective conflict-of-interest management in medicine. *Journal of General Internal Medicine*, 39(8), 1503-1505. <https://doi.org/10.1007/s11606-024-08668-z>

Samsa, G. & Solomon, A. (2019). Managing the science in the presence of financial conflict of interest. *Accountability in Research*, 26(6), 397-403. <https://doi.org/10.1080/08989621.2019.1636652>

Schmidt, A. M., Sorci-Thomas, M. G., Chen, Y. & Hegele, R. A. (2023). Academic-industry partnerships: Transparency, potential conflict of interest, and communicating state-of-the-art technologies. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 43(7), 1093-1095. <https://doi.org/10.1161/ATVBAHA.123.319633>

Schuhl, A. (2022, 7 avril). *Le CNRS encourage ses scientifiques à ne plus payer pour être publiés*. CNRS. <https://www.cnrs.fr/fr/actualite/le-cnrs-encourage-ses-scientifiques-ne-plus-payer-pour-etre-publies>

Vervoort, D., Lee, G., Ghandour, H., Guetter, C. R., Adrean, N., Till, B. M. & Lin, Y. (2024). Global Cardiac Surgical Volume and Gaps: Trends, Targets, and Way Forward. *Annals of Thoracic Surgery Short Reports*, 2(2), 320-324. <https://doi.org/10.1016/j.atssr.2023.11.019>

1. En France la « Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique » ; dans la Communauté française de la Belgique le Décret du 3 mai 2018 visant à l'établissement d'une politique de libre accès aux publications scientifiques (*open access*).

2. RIM : Regulatory Information Management.

3. Contrairement aux sciences sociales, la plupart des auteurs et des rédacteurs en sciences médicales ne font pas la différence entre un lien d'intérêts et un conflit d'intérêts. C'est pourquoi nous utilisons le terme *conflit d'intérêts* dans son sens plus large.

4. <https://chatgpt.com/>

5. www.deepl.com

PARTIE 3. LES ORGANISATIONS ET L'INTELLIGENCE ARTIFICIELLE

COMMUNIQUÉ 3. LES DIX CONSÉQUENCES DU PLAGIAT ET DE LA FRAUDE ACADÉMIQUES

[Michelle Bergadaà](#), Professeure émérite, Université de Genève (Suisse)

Avec l'émergence des outils d'intelligence artificielle, bien malin qui saura établir le plagiat au regard des simples termes de la législation « droits d'auteur » et « contrefaçon ». Qui est le véritable auteur ? Et qui saura si les données d'une recherche n'ont pas été exploitées par une intelligence artificielle ou des codes informatiques développés par un chatbot ? C'est donc dans une perspective davantage holiste que nous devons maintenant considérer les défis qui se présentent à nous en matière d'intégrité académique.

Lorsque nous représentons les victimes d'inconduites académiques auprès des instances académiques ou des éditeurs, nous vérifions avant toute chose que la déviance peut être déterminée selon la méthodologie de l'IRAFPA. Nous constituons le dossier des preuves formelles des actes délictueux au regard des règles et normes académiques (cf. communiqué 1).

Puis nous évaluons le plus précisément possible la gravité de l'inconduite académique dénoncée en examinant les dix conséquences formulées dans l'encadré ci-après. Si nous mesurons ces conséquences ce n'est pas pour les déplorer, mais pour établir les termes des réparations à réclamer en relation avec les dommages causés par le comportement irrégulier.

Ainsi, dans tous les cas de médiation que nous conduisons à la demande de victimes ou de témoins de méconduites, nous passons au crible la grille d'analyse ci-après, conséquence après conséquence. Toutes, bien entendu ne sont pas impliquées simultanément. Chaque cas d'inconduite doit être considéré comme unique et donc générant des conséquences spécifiques. Cet inventaire précis nous permet de proposer des mesures de réparation adaptées à chacune des conséquences identifiées. Il peut s'agir, par exemple, de la mise au pilon d'un livre plagié, de la rédaction d'une lettre d'excuse par le responsable hiérarchique négligent, de la demande d'un communiqué expliquant le retrait d'un article...

Cette grille nous est également utile pour dialoguer avec les journalistes ou les dirigeants d'établissements qui nous sollicitent, car elle leur permet d'évaluer rapidement la gravité relative des comportements identifiés comme déviants et de savoir si une enquête approfondie s'impose ou non. Et elle prend aussi tout son sens pédagogique quand il s'agit de dialoguer avec des « délinquants de la connaissance » parfois inconscients des conséquences de leurs actes, à l'instar des chauffards de la route qui ne réalisent la dangerosité de leur conduite qu'une fois qu'un accident s'est produit.

In fine, l'IRAFPA considère que son action serait bien limitée s'il ne s'agissait que de constater des faits s'étant produits dans le passé. Ce qui est essentiel est de comprendre pourquoi les actes délictueux se sont produits, puis de mettre en place des dispositifs individuels et collectifs pour prévenir les risques à venir. Et c'est en cela que les responsabilités des acteurs individuels et collectifs sont décrites dans les dix conséquences ci-après. Ils peuvent ainsi prendre conscience qu'ils peuvent et qu'ils doivent agir.

Références bibliographiques

Bergadaà, M. (2013). Tu ne plagieras, point. *Cultures & Sociétés*, 28, 18-22.

Bergadaà, M. (2015). Appréhender la gravité du plagiat. Dans M. Bergadaà, *Le plagiat académique. Comprendre pour agir* (pp. 27-66). L'Harmattan.

Communiqué de l'IRAFPA – Les dix conséquences du plagiat et de la fraude académiques¹

- Conséquence fondamentale 1 – Fraude et plagiat nuisent à l'avancement de la connaissance

Le droit fondamental des chercheurs est l'accès à l'origine des sources de la connaissance. Nous *devons* citer nos sources dans nos publications afin de permettre à tout nouveau chercheur de se pencher à son tour sur nos données de base (quelle qu'en soit la nature), de conduire sa propre analyse et de proposer de nouvelles interprétations. Est un délinquant de la connaissance celui qui rompt ou mystifie le lien liant ceux qui l'ont précédé à ceux qui le suivent.

- Conséquence 2 – Le plagiat est un vol de la création originale

La souffrance des victimes doit être comprise de et accepté par tous. Plagier revient à faire disparaître l'auteur d'origine aux yeux de sa communauté. Se sentir ainsi symboliquement tué peut être dévastateur. Il ne s'agit pas seulement d'une appropriation de l'œuvre, mais de la paternité de l'œuvre. Il s'agit du vol d'une œuvre de l'esprit, donc unique : c'est une atteinte grave aux droits et intérêts moraux de l'auteur, inaliénables et consacrés en tant que droits fondamentaux.

- Conséquence 3 – Plagiat et fraudes académiques vident une production de sens

La raison même de la recherche est la quête de vérité. La technique du patchwork ou du collage à partir d'un ou plusieurs textes sources peut rendre le plagiat difficile à déceler et donner l'illusion au lecteur que ce texte a un sens savant. La technique des faux positifs relève du même comportement et le résultat aboutit à une illusion de véracité scientifique.

- Conséquence 4 – Plagiat et fraudes académiques incitent à la recherche bâclée

Ce qu'il y a de sacré dans la recherche est la liaison fusionnelle qui existe entre le chercheur et son œuvre. Il advient parfois qu'un comportement paresseux d'un auteur qui se contente de plagier ou de frauder ici et là bouleverse les conventions et les règles académiques. Références bibliographiques diluées, auteurs fantômes ou invités, données tronquées, etc., peuvent alors devenir des us et coutumes de son entourage professionnel.

- Conséquence 5 – Le plagiat et la fraude scientifique constituent une fraude du système académique

Dans un système sous tension, l'égalité de traitement doit être une préoccupation majeure. Le délinquant introduit une grande iniquité puisqu'il multiplie à bon compte le nombre de publications figurant sur son curriculum vitae. Il sera en mesure de bénéficier d'avantages professionnels au détriment de la personne honnête qui aura déclaré les écrits relevant de son seul mérite. L'autoplégat, qui consiste à utiliser le même écrit en le modifiant à la marge ou en le traduisant simplement, sans prévenir, pour démultiplier sciemment le nombre de ses publications sur un curriculum vitae, est aussi une fraude.

- Conséquence 6 – Le comportement délictueux peut inhiber des chercheurs compétents

Collaborateurs et coauteurs sont éclaboussés par la révélation d'un scandale. Ce dernier donne naissance à des rumeurs qui blessent des personnes innocentes. Or, certaines avaient une confiance aveugle envers les auteurs des fautes, d'autres ne voulaient simplement rien voir. En jetant l'opprobre sur un groupe de chercheurs, le phénomène du plagiat et de la fraude académiques donne envie à certains chercheurs intègres d'abandonner le métier.

- Conséquence 7 – Le comportement délictueux provoque un dysfonctionnement des revues et des maisons d'édition scientifiques

Ces revues scientifiques et ces éditeurs sont indispensables à l'ordre académique. Ils déterminent la distinction des chercheurs aptes (ou non) à être engagés par un type donné d'institution, ou d'accéder ensuite à des fonctions de direction. La délinquance académique empêche les revues et les éditeurs scientifiques de remplir pleinement ce rôle régulateur. En outre, elle induit de nombreuses turbulences, économiques pour ces mêmes éditeurs lorsqu'ils doivent retirer des articles.

- Conséquence 8 – La lutte contre l'inconduite académique coûte cher

Le coût de prévention sera toujours moins élevé que celui de la répression des comportements délictueux. La découverte d'un cas et la mise en examen de son auteur entraînent des commissions d'enquête longues et coûteuses. Nul ne chiffre, et c'est un tort, les salaires et honoraires des avocats et enquêteurs associés, mais aussi le temps que les membres des commissions d'enquête auraient pu consacrer à la recherche et à l'écriture.

- Conséquence 9 – Plagiat et fraudes académiques portent atteinte à l'image de nos établissements

La délinquance académique génère de la suspicion dans la société civile à l'égard des scientifiques. En jetant l'opprobre sur ses mentors – responsables de laboratoires, directeurs

de thèses, rédacteurs en chef de revues – qui n’ont pas su recadrer les délinquants, elle nuit à l’image de l’institution. La tentation de l’omerta n’est nullement une solution, car tout finit toujours par se savoir. Les établissements ont la responsabilité de développer une culture de l’éthique et de l’intégrité, c’est une composante à part entière de leur mission sociale.

- Conséquence 10 – Le comportement délictueux entretient l’ambiguïté systémique

La responsabilité des dirigeants d’établissements est engagée quand des étudiants obtiennent un diplôme alors qu’ils ont triché, tout comme elle l’est si des agents ou des collaborateurs divulguent de fausses connaissances par le biais de journaux scientifiques ou par celui des médias. Le devoir des dirigeants est de mettre en place les dispositifs qui préserveront et encourageront l’intégrité académique.

1. <https://irafpa.org/methodes/instruire-un-plagiat/les-dix-consequences-du-plagiat/>

CHAPITRE 8. PROMOUVOIR LES USAGES DE L'INTELLIGENCE ARTIFICIELLE (IA) DANS LE RESPECT DE L'INTÉGRITÉ ACADÉMIQUE. LE CAS DU BUSINESS SCIENCE INSTITUTE

[Jean Moscarola](#), Professeur émérite, fondateur et président d'honneur de la société Sphinx Développement (France)

[Michel Kalika](#), Professeur honoraire à iaelyon, Université Jean Moulin (France), fondateur et président du Business Science Institute (Luxembourg)

1. INTRODUCTION

L'observation et la mesure marquent la première révolution scientifique dite copernicienne. Elle est suivie de la révolution numérique. Nous sommes peut-être à l'aube d'une 3^e révolution : celle de l'IA générative. Toutes ces révolutions sont le fruit d'innovations technologiques qui complètent les capacités d'observation (télescope, microscope, rayon x...), puis les capacités de calcul (ordinateur, informatique). Jusqu'ici produire des connaissances issues de l'observation, du calcul ou du raisonnement était le privilège de l'auteur observateur, expérimentateur ou théoricien. En donnant aux machines informatiques la capacité de rédiger, les grands modèles de langage sont-ils en mesure de remplacer ces capacités de l'auteur scientifique que sont l'observation, le calcul et le raisonnement ?

Sommes-nous à l'aube d'une 3^e révolution scientifique, qui, comme les précédentes, contribuera à la croissance du savoir scientifique ? Ou bien nous acheminons-nous vers une nouvelle forme d'humanité dans laquelle une machine égalerait les capacités humaines au point de figurer comme auteur légitime au sein de la recherche ? Les IA génératives seraient-elles en passe de bouleverser la notion même d'intégrité académique ? Toutes ces questions sont posées et provoquent des réactions parfois extrêmes, qui relancent avec

encore plus d'acuité les débats autour de l'impact des technologies sur les différentes formes d'atteinte à l'intégrité académique.

Nous répondrons à ces questions à partir de l'expérience acquise au Business Science Institute. Nous présenterons d'abord les spécificités de notre institut. Puis nous évoquerons les interrogations en matière d'intégrité académique tant des doctorants que des enseignants de l'Institut quand ils utilisent des outils d'IA. Enfin nous présenterons les bonnes pratiques que nous avons mises en place.

Le Business Science Institute : une organisation en réseau pour des doctorants-managers

Il s'agit d'une institution académique internationale en réseau, innovante à plus d'un titre : depuis plus de 10 ans, sans locaux ni corps professoral exclusif, elle est spécialisée dans la préparation du Doctorate of Business Administration (DBA) et dans la production de connaissances par les managers-chercheurs. Les doctorants sont tous des professionnels expérimentés. Ils ont quitté le monde académique depuis longtemps (l'âge moyen est de 45 ans).

Le Business Science Institute organise son DBA en face à face et en ligne en quatre langues. Il compte plus de 200 doctorants-managers issus de 60 pays et plus de 150 professeurs en réseau. 200 thèses de DBA ont été soutenues et plus de 55 ouvrages publiés. Les séminaires dispensés au cours la première année ont pour but d'immerger les participants dans l'actualité des questions de recherche et des méthodologies. À cet effet, des outils de recherche bibliographique, de collecte et d'analyse de données sont mis à leur disposition.

La singularité de cette organisation va de pair avec le respect des traditions académiques. Ainsi, le suivi de la relation doctorant-directeur de thèse, le fonctionnement de la faculté, dont les membres proviennent de différentes institutions, et les procédures de délivrance de la thèse ont été reconnus et accrédités par AMBA (organisme international d'accréditation des formations en management).

Ce terrain permet l'observation des comportements de personnes confrontées à des exigences différentes de celles du monde professionnel dont ils sont issus et soumises à la pression d'une charge de travail importante imposée par le programme.

Depuis sa création, le Business Science Institute a mis en place un système d'accompagnement des doctorants-managers. Il consiste à mettre à leur disposition des instruments pour collecter et analyser leurs données. Comme ils n'ont ni la connaissance technique de ces outils ni le temps de s'y former, ils sont aidés pour cela. À charge pour eux de réaliser les entretiens et d'interpréter les résultats du traitement des données. Cette sous-traitance technique facultative est payante et soumise à l'aval du directeur de thèse. Reste au chercheur le travail de fond : choix du sujet, examen de la littérature, formulation d'une problématique, interprétation des résultats, rédaction de la thèse et formulation des implications managériales. Ce système, développé depuis 10 ans, concerne approximativement une thèse sur six et

s'est étendu à l'occasion de la Covid-19 à des recherches collectives menées entre collègues.

2. L'APPARITION DES IA GÉNÉRATIVES FAISANT SUITE À LA COVID-19

Les bénéficiaires de notre accompagnement gagnent du temps et accomplissent des collectes plus riches et/ou des analyses plus approfondies grâce à des méthodes qu'ils n'auraient pas pu mettre en œuvre seuls : diffusion de questionnaires en ligne, collecte en grand nombre de sources documentaires et recueil de transcriptions d'entretiens. Cette aide technologique leur permet de découvrir des méthodes nouvelles pour eux. Bien que proposée par l'organisation, cette expérience provoque cependant des questions : ne risque-t-elle pas d'être considérée comme abusive dans la mesure où le chercheur aurait répondu positivement aux offres de soutien de l'organisation ?

La crise de la Covid-19 a été vécue au Business Science Institute comme l'opportunité de lancer des recherches collectives impliquant plusieurs professeurs et directeurs de thèse qui ont bénéficié de notre dispositif d'assistance. Ainsi, des collègues appartenant à différents champs disciplinaires et habitués aux approches quantitatives ou qualitatives ont pu coopérer sur un large projet sans se préoccuper de ce qu'ils ne maîtrisaient pas complètement. Cette expérience a non seulement permis de publier deux ouvrages collectifs (Kalika, 2020 ; Kalika & Beaulieu, 2021), mais également de proposer l'idée de CodataLab défendant la mise en commun de données et de résultats produits et analysés avec l'assistance d'un dispositif technologique et d'un facilitateur. Les participants ont pu ainsi découvrir que l'accompagnement dont ils ont bénéficié leur a permis de collecter plus de données et de les analyser différemment (Moscarola *et al.*, 2022).

L'apparition des IA génératives a conduit à compléter cette assistance. Cela s'est manifesté de trois manières.

- **Une aide à la formulation du projet de recherche**

Le professeur en charge du séminaire de méthodologie utilise ChatGPT pour prendre connaissance des projets des participants. Ainsi, depuis 2022, ils reçoivent, au début du séminaire, la copie d'un dialogue dans lequel les réponses de ChatGPT donnent une vision de leur projet résultant des prompts du professeur. Ce retour est bien rédigé, dépourvu d'affect et bienveillant. Il fournit à chacun un commentaire sur sa question de recherche, les théories de référence pertinentes et les méthodes envisageables pour observer le terrain. Ces textes permettent aussi au professeur d'illustrer son propos au cours du séminaire.

- **Le couplage avec les outils traditionnels**

En aval, pour interpréter les résultats et faire parler les données, les outils classiques d'analyse statistique ou d'analyse de données textuelles sont couplés avec ChatGPT. L'assistance porte sur un ensemble de méthodes et d'outils qui « augmentent » les capacités du doctorant (Moscarola, 2024). Il a ainsi l'occasion de découvrir comment utiliser Sphinx sur ses propres données en le couplant avec ChatGPT pour orienter sa recherche et être plus productif. Le temps ainsi gagné peut servir à approfondir les observations et améliorer la qualité des données.

L'originalité de ce cas tient au fait qu'il concerne une population de doctorants et d'encadrants pour lesquels l'assistance peut compenser l'absence du support habituellement offert par un laboratoire traditionnel. Mais ces pratiques modifient les équilibres et posent la question des rapports entre IA et intégrité académique.

- **Le défi de IA à l'intégrité académique**

L'IA constitue un défi et fait se télescoper technologie et éthique ; ce qui nous rappelle que la survenue des IA génératives est concomitante avec le renforcement de l'exigence académique exprimé dans de très nombreux commentaires. Cette opposition, ou cette convergence, repose sur la profonde ambiguïté de ces notions qui empêche de bien considérer les modalités de leur synergie car toutes deux reposent sur un abus de langage. Cet abus s'est imposé par facilité car chacune des deux IA donne l'illusion d'une grande

efficacité pour la production de la recherche d'une part, pour la réputation du chercheur d'autre part, et finalement pour le succès du doctorant.

3. IA COMME INTELLIGENCE ARTIFICIELLE

Nous nous intéressons ici essentiellement aux IA génératives qui, depuis leur apparition, ont entretenu la confusion entre l'IA au sens large et les grands modèles de langage ou Pretrained Transformer entraînés à produire des textes qui créent l'illusion d'une intelligence capable d'écrire de manière autonome.

L'IA est encore une technologie obscure et bien mal nommée : la technologie comme artefact ne modélise pas la conscience et les connaissances qui font l'intelligence humaine, mais son efficacité fait illusion.

Les grands modèles de langage mettent en œuvre des procédés très simples consistant à établir des connexions avec les textes sur lesquels ils sont périodiquement entraînés. Andler (2023) à la suite de Floridi (2014) évoque cette masse gigantesque de données sous l'appellation de « numérisphère » pour indiquer qu'elle contient toutes les connaissances numérisées avec lesquelles les IA sont entraînées. Plus la base d'entraînement est grande, meilleur est le résultat (Le Cun, 2019). Il est non seulement intelligible, mais il fait sens pour le lecteur, car les textes produits portent la trace des connaissances incorporées dans la numérisphère.

Mais si la technologie permet d'écrire correctement, elle n'a pas conscience de ce qu'elle écrit. Ainsi, dans les premières applications, les « hallucinations » et les énoncés inappropriés étaient fréquents (Bender *et al.*, 2021). La plupart de ces aberrations ont été corrigées par différentes formes de supervision. Il demeure que le doctorant est seul responsable du crédit qu'il donne à ce que l'IA écrit. La lecture qu'il en fait dépend de ce qu'il a lui-même formulé dans la question posée. Lui seul est intelligent et aucune intelligence, au sens de l'intelligence humaine, ne peut être attribuée à ces technologies dont le fonctionnement est totalement opaque.

Pourtant l'illusion fonctionne, alors que l'IA ne produit que des énoncés vraisemblables qui doivent toujours être examinés avec un regard critique et confrontés aux faits ou à la logique. *In fine*, tout dépend des usages et de l'intelligence de l'utilisateur. L'IA peut être efficace si la question est très simple et la réponse déjà présente dans la numérisphère, comme si on utilisait un moteur de recherche élaboré. Elle est féconde quand, par une succession de questions, on l'utilise pour fouiller, pour chercher des idées ou des références auxquelles on n'aurait pas pensé.

Certes, l'IA augmente les capacités de réflexion et peut être très utile pour la production de texte relativement courts, une ou deux pages et quelques paragraphes. Mais elle a ses limites : la longueur des textes et leur cohérence argumentative ne reposent que sur les connexions établies à partir des mots de la question dont dépend la qualité des résultats ; plus la question (le prompt) est détaillée, plus longue est la réponse ; plus la question est argumentée ou construite sur un raisonnement, plus le texte produit s'éloignera de la simple énumération de thèmes et sous-thèmes.

Par conséquent, le bon usage des IA génératives dépend avant tout de l'intelligence du doctorant capable de proposer à l'outil le raisonnement qu'il souhaite développer. C'est la grande différence avec les autres formes d'IA, conçues pour résoudre des problèmes bien spécifiés en apportant des solutions directement actionnables (voitures autonomes, attribution de crédit...).

L'art du prompt devient alors essentiel. Les IA génératives sont des systèmes complexes et opaques dont il faut maîtriser l'usage : savoir poser des questions à l'outil en l'informant sur le contexte, en lui donnant les informations spécifiques dont on dispose et en précisant ce qu'on attend. Ces trois éléments entrent dans l'élaboration du prompt, qui désigne autant la question posée que le texte reçu en retour. C'est la fenêtre de l'interaction entre l'utilisateur et l'outil.

En substance : « Bien utiliser ChatGPT, c'est un peu comme maîtriser l'art de dialoguer avec un prompteur. Il s'agit de savoir poser les bonnes questions de manière précise et claire, et ensuite

de lire attentivement et comprendre les réponses pour continuer la conversation de manière efficace. Cela implique souvent d'affiner les questions suivantes en fonction des réponses reçues pour se rapprocher progressivement des informations désirées. C'est un processus interactif qui peut être très enrichissant quand on maîtrise bien ces étapes. » (ChatGPT) ou encore : « *The essence of prompt engineering is about crafting the right questions or instructions to guide AI models, especially Large Language Models (LLMs), to produce desired outcomes.* » (Crabtree, 2023)

Ce bon usage appelé aussi « prompt ingénierie » est l'objet de nombreuses publications et propositions de formation (Frimousse & Besseyre des Horts, 2023). C'est à cela qu'il convient de former autant les doctorants que leur encadrant, car la facilité d'usage et l'apparente efficacité peuvent très facilement conduire à des usages maladroits ou abusifs de la part des doctorants.

4. LE POINT DE VUE DE L'INTÉGRITÉ ACADÉMIQUE

L'intégrité académique comme injonction éthique a ceci de paradoxal qu'elle s'adresse aux chercheurs comme individus, tout en les situant dans le contexte collectif d'une institution. Au-delà de la tension entre individu et institution, le paradoxe tient à la finalité de la recherche – innover, créer de nouvelles connaissances, et à celle de l'institution – conserver et établir les connaissances communes à une discipline. D'une part, l'individu cherche à se distinguer en apportant des connaissances nouvelles, de l'autre l'organisation pose un cadre de référence à ceux qui veulent y entrer. Ainsi, le chercheur authentique cherchera à tout prix à innover et à exercer sa liberté en prenant le risque de se heurter aux règles imposées par sa discipline pour ne pas renoncer à l'exercice de l'esprit critique ou à celui de la créativité.

Mais si certaines règles conventionnelles propres à chaque paradigme peuvent effectivement entraver cette liberté il en est d'autres plus universelles qui s'imposent. Il s'agit de l'éthique de la transparence (citer ses sources, exposer ses données) ou de la

simple honnêteté intellectuelle consistant à ne pas s'attribuer les idées d'un autre ou à tricher sur les données. C'est ce que défend l'IFRAPA (Uhaldeborde, 2023 ; Bergadaà, 2021 ; Bergadaà & Peixoto, 2023). Ainsi, la tension que nous évoquons plus haut disparaît au niveau des valeurs plus universelles de la morale ou de l'éthique qui fondent l'intégrité académique.

L'apparition des IA réactive le débat sur le plagiat : elle permet au chercheur de s'attribuer une capacité d'écriture quand il recopie un texte produit par une machine, sans qu'il soit possible de le prouver factuellement. Plusieurs articles dénoncent ainsi des usages abusifs de l'IA induits par la pression pour publier (Diallo, 2023) ou par l'abandon de l'esprit critique face aux biais et possibles « hallucinations » (Godé *et al.*, 2023). Ainsi, l'émergence des IA génératives peut transformer le paysage de la recherche en faisant émerger des pseudo-textes sans auteur identifiable, qui ne peuvent pas être reconnus comme tels car les logiciels peuvent aider à la détection d'usages abusifs mais n'apportent pas de preuve comme en cas de plagiat textuel.

Gain de temps et source d'inspiration, l'usage des IA génératives augmente les capacités des chercheurs, mais il introduit également une complexité accrue nécessitant un nouveau type d'exigence et de vigilance éthique. Bref, l'IA est un défi pour l'intégrité académique.

En ce qui concerne les encadrants et les professeurs, le risque n'est en fait pas celui d'un abus, mais d'un rejet trop rapide, provoqué par la crainte d'être remplacé ou abusé dans son rôle d'auteur responsable, par la crainte des hallucinations et des biais d'apprentissage, par la question non encore résolue des droits d'auteur et l'hégémonie culturelle des GAFA. De leur côté, les doctorants craignent une perception négative de l'IA par leurs encadrants. Récemment un doctorant déclare qu'il utilise ChatGPT, mais qu'il en a honte et préfère ne pas le dire : il s'en sert pour les recherches bibliographiques, pour lire le résumé d'un article avant de décider de le lire en entier, pour approfondir une notion ou un concept qu'il maîtrise mal ou pour vérifier une interprétation suggérée par l'analyse de ses données. Ce type de réaction est

injustifié : tant qu'on ne recopie pas à l'identique les réponses de ChatGPT sans le dire, il n'y a pas de honte à utiliser un outil qui aide à réfléchir. Ne cite-t-on pas les logiciels utilisés pour analyser des interviews ?

Afin d'analyser cette réaction, nous l'avons soumise à ChatGPT¹. Il en ressort la peur du jugement fondé sur une présomption de tricherie, comme si le simple fait d'utiliser ce type d'outil était suspect. Cette réaction est assez répandue, comme en témoigne une recherche sur la perception de l'IA (Ocal, 2023). On craint que la seule mention de ChatGPT ne fasse douter de l'originalité du propos. De plus, le manque de compréhension d'une technologie opaque peut provoquer un sentiment de perte de contrôle et un manque d'assurance pour ceux qui ne sont pas suffisamment familiarisés avec leur usage (Lebovitz *et al.*, 2022). Ces réactions individuelles témoignent d'un manque de maîtrise qui provoque une contradiction entre l'efficacité perçue du procédé et sa légitimité.

Elles sont aussi imputables au contexte des institutions académiques. En effet, depuis leur apparition, les intelligences artificielles génératives provoquent des réactions contrastées. Les réticences se sont d'abord manifestées quand il s'agissait d'évaluer les travaux des étudiants : comment évaluer des essais dont on ne peut savoir s'ils sont écrits par des robots ou s'ils révèlent vraiment la compétence des personnes évaluées ? (Moscarola & Chauhan, 2024). Cette interrogation trouve une réponse institutionnelle dans la mise en place de règles (Bieliauskaite, 2021 ; Macleod & Eaton, 2020) qui concernent aussi bien les publications des chercheurs que les travaux des étudiants. Ces règles visent à garantir l'intégrité académique selon deux types de stratégies : le contournement ou l'intégration (Anctil, 2023).

Le contournement consiste à interdire l'usage de ces outils en privilégiant les examens sur table et surveillés ou en excluant de publication les chercheurs qui les utilisent. L'intégration au contraire conduit à légitimer l'usage des IA génératives au même titre que de

toutes les autres technologies de la connaissance avec obligation de les citer et de préciser l'usage qui en a été fait (Bergadaà, 2023).

5. RÉCONCILIER L'INTELLIGENCE ARTIFICIELLE ET L'INTÉGRITÉ ACADÉMIQUE

À condition de ramener ces deux notions à ce qu'elles recouvrent vraiment : des outils et des comportements humains, il devrait être possible de penser leur complémentarité dans un souci d'efficacité d'une part et de l'éthique d'autre part.

Les IA génératives sont très efficaces pour faciliter le travail du chercheur, notamment pour mobiliser les références académiques, situer les questions de recherche, explorer le monde des idées et des connaissances, faire parler les données et communiquer (Moscarola, 2024). Cependant, elles ne connaissent pas le monde des phénomènes et sont constitutivement en retard sur l'actualité. Perroquets stochastiques, elles sont circonscrites aux données d'apprentissage et ne contiennent donc aucune connaissance à proprement parler nouvelle (Bender *et al.*, 2021).

Or la découverte et la création de nouvelles connaissances est bien la raison d'être du chercheur, ce qui fonde son identité.

Le temps gagné grâce aux IA génératives pour mobiliser les connaissances existantes permet de se concentrer sur la réflexion génératrice de nouvelles connaissances. Cet usage vertueux de l'IA peut préserver de l'effondrement d'une recherche académique bouclée sur elle-même (Peterson, 2024). En effet un usage généralisé des IA pourrait conduire à ce que leur entraînement ne permette pas de régénérer les connaissances qu'elles contribuent à diffuser. C'est alors l'intégrité des connaissances elle-même qui serait la conséquence du non-respect de l'intégrité par le chercheur. C'est ainsi que la productivité apportée par les outils de l'IA peut contribuer au progrès effectif des connaissances garanti par le respect de l'intégrité académique.

6. PRATIQUES POUR LES DIFFÉRENTS ACTEURS DU BUSINESS SCIENCE INSTITUTE

Le Business Science Institute s'adresse à des professionnels encadrés par un corps professoral issu de nombreuses universités. Il offre aux doctorants un accompagnement technologique d'une qualité pour eux inhabituelle et leur permet de bénéficier de la compétence de professeurs précurseurs en matière d'IA. Ces conditions particulières jointes à la flexibilité de l'organisation ont favorisé l'adoption de règles (ou de pratiques) permettant de concilier les possibilités offertes par l'IA et l'intégrité académique pour l'accompagnement des chercheurs.

Pour les doctorants qui, bien souvent, ne connaissent pas les outils couramment utilisés par les chercheurs, c'est l'occasion de découvrir ces « technologies de la connaissance » (Stiegler, 2016). Pour les chercheurs, ces technologies représentent aussi une rupture par rapport à la tradition qualitative dans laquelle le chercheur lit, observe, réfléchit et écrit sans recourir à aucun dispositif technologique. Cette tradition a précédé l'époque actuelle dans laquelle les bases de données bibliographiques, les moteurs de recherche, les logiciels de collecte et d'analyse de données (Excel, SPSS, NVivo, Sphinx) dotent les chercheurs d'instruments qui permettent de recueillir davantage de données et de les traiter plus vite et de manière plus objective. Ces outils « signent » les papiers autant que leurs auteurs par les méthodes qu'ils mettent en œuvre et les résultats qu'ils produisent. Ils sont *a priori* légitimes. Il faut simplement en faciliter l'usage. Il est moins légitime, en revanche, d'assister les doctorants en leur indiquant la meilleure méthode et en la mettant en œuvre pour eux. Cette aide optionnelle remplace l'entraide de la vie de laboratoire reconstituée à distance par le Business Science Institute Lab.

Il faut apprendre à connaître les qualités et les défauts des différents outils d'IA générative. ChatGPT n'a pas très bonne presse dans le milieu de la recherche. Il est cependant facile d'emploi et peut se révéler très utile. En revanche ce qu'il écrit n'est que

vraisemblable : il n'apporte aucune preuve, aucune garantie objective de la vérité des énoncés qu'il produit. C'est à l'utilisateur de les valider pour se les approprier comme auteur. « ChatGPT a écrit... » n'a pas la valeur de vérité d'une référence, d'un tableau, d'un graphique ou d'un verbatim produit par Scopus, Excel ou Sphinx. Mais il suffit d'expliquer aux doctorants comment mettre en œuvre ChatGPT : en formulant une question et en interprétant les réponses pour aller plus loin. Comme nous le répond ChatGPT lui-même, « l'apprentissage de la formulation d'un prompt, comme la formulation d'une question, est un exercice de réflexion. Utiliser ChatGPT nécessite de réfléchir précisément à ce que l'on cherche à savoir. Cela constitue déjà une partie importante du travail. Ensuite, il est crucial de lire les réponses pour identifier ce qui a été omis ou les points de désaccord. Construire un prompt et continuer la conversation avec ChatGPT est un processus qui permet de développer et d'affiner sa pensée ». Ainsi, citer ChatGPT respecte l'éthique propre à l'écriture de recherche. Il se peut cependant que les règles de publication de certains journaux scientifiques en refusent l'usage.

7. DISCUSSION

Pour une organisation comme le Business Science Institute, dont la seule activité est l'accompagnement de doctorants-managers et la validation de leurs travaux, la question de l'intégrité académique et de l'utilisation de IA est essentielle. C'est pourquoi nous avons mis en place des règles qui devraient permettre d'éviter que ne se reproduise la mésaventure d'une étudiante dont la soutenance a été annulée parce que l'un des membres du jury avait remarqué que certaines références bibliographiques relevaient de « l'hallucination », qu'elles n'existaient pas ! Cet incident a également mis en évidence une faille du dispositif de contrôle qui n'avait été appliqué qu'au texte de la thèse et non à la bibliographie. Depuis ce jour, les références qui ne font l'objet d'aucune occurrence sont examinées de très près...

Nos dispositifs pour concilier intelligence artificielle et intégrité académique

Ces exigences s'inscrivent dans la logique de la certification « Responsible Doctoral School » délivrée par l'IRAFPA au Business Science Institute en septembre 2023.

1/ Nous sensibilisons les étudiants à l'intégrité académique dès le début du cursus. Nous leur donnons accès à un logiciel anti-plagiat (Turnitin) avant le premier séminaire. Nous les informons également que tous les documents remis par les étudiants sont vérifiés à l'aide de ce logiciel. Cette vérification conduit à leur signaler tout problème de citation des sources.

2/ Plus récemment, nous avons réalisé une série de six vidéos de formation à l'usage de ChatGPT à toutes les étapes de la recherche ; elles sont disponibles sur l'intranet Moodle du Business Science Institute. Leur objectif est de faire prendre conscience aux doctorants à la fois de l'apport potentiel, mais aussi des limites de l'IA.

3/ Nous exigeons également que les étudiants précisent au début de leur thèse quelles aides ils ont reçues et quel usage ils ont fait des outils d'IA ; nous leur demandons notamment de faire figurer en annexe les prompts adressés aux outils de l'IA. Ces informations permettent d'identifier les apports et les approfondissements imputables au chercheur lui-même.

4/ Nous tenons compte de l'hétérogénéité des attitudes et des comportements : si l'intégrité académique est une valeur partagée par les professeurs, ils sont moins unanimes à valoriser les apports de l'IA. C'est en se formant aux exigences de l'intégrité académique qu'ils dépasseront ces appréhensions.

Le temps libéré par l'utilisation de l'IA pour chercher et mobiliser des références permet d'en consacrer davantage à l'observation scientifique proprement dite. L'IA permet aussi d'éviter les dérives d'une recherche académique exclusivement centrée sur elle-même. Un usage vertueux des outils d'IA est possible ; ils doivent donc entrer dans la formation des managers.

Cela indique le chemin de la synergie à trouver entre intelligence artificielle et intégrité académique : il passe par les I, de l'intelligence et de l'intégrité et par les A de l'artificiel et de l'académique. Pour cela il faut simplement que la raison pratique des technologies s'accorde avec la sagesse des institutions, comme pourrait l'écrire ChatGPT !

Le chemin est encore long, car la révolution que nous évoquions en introduction n'en est qu'à ses débuts. L'expérimentation menée au sein du Business Science Institute s'inscrit dans ce mouvement

dont le succès dépend notamment de l'engagement individuel des directeurs de thèse.

Références bibliographiques

- Anctil, D. (2023). Higher education in the era of generative AI. *Pédagogie collégiale*, 36(3).
- Andler, D. (2023). *Intelligence artificielle, intelligence humaine : la double énigme*. Gallimard.
- Bender, E. M., Gebru, T., McMillan-Major, A. & Shmitchell, S. (2021, March). On the dangers of stochastic parrots: Can language models be too big? Dans *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610-623).
- Bergadaà, M. (2021). *Academic plagiarism. Understanding it to take responsible action*. Research ethics 1, Globethics.net.
- Bergadaà, M. (2023). *Les nouvelles frontières de l'intégrité académique*. Éditions EMS, Coll. « Questions de société ».
- Bergadaà, M. & Peixoto, P. (2023). *Academic integrity. A call to research and action*. Research Ethics 2, Globethics.net.
- Bieliauskaitė, J. (2021). Solidarity in academia and its relationship to academic integrity. *Journal of academic ethics*, 19(3), 309-322. <https://doi.org/10.1007/s10805-021-09420-6>
- Carvalho, S. (2019). L'éthique de la recherche entre réglementation et réflexivité. *Revue d'anthropologie des connaissances*, 13(2), 299-326.
- Crabtree, M. (2023). The future of data literacy. DataCamp.com.
- Diallo, M. F. (2023). Ce que ChatGPT fait à l'enseignement, à la recherche et aux organisations. *Revue française de gestion*, 312(5), 9-14.
- Floridi, L. (2014). *The fourth revolution: How the infosphere is reshaping human reality*. Oxford University Press.
- Frimousse, S. & Besseyre des Horts, C.-H. (2023). L'art du prompt : Une compétence à maîtriser pour les futurs chercheurs en sciences de gestion. *Management & Datascience*, 7(3). <https://doi.org/10.36863/mds.a.24326>
- Godé, C. Bidan, M. & Lebraty, J. F. (2023). ChatGPT, étudiants et enseignants-chercheurs : sont-ils vraiment félins pour l'autre ? *Management & Datascience*, 7(2).
- Kalika, M. (2020). *L'impact de la crise sur le management*. Éditions EMS, Coll. « Business Science Institute ».
- Kalika, M. & Beaulieu, P. (2021). *Les impacts durables de la crise sur le management*. Éditions EMS, Coll. « Business Science Institute ».
- Le Cun, Y. (2019). *Quand la machine apprend : la révolution des neurones artificiels et de l'apprentissage profond*. Odile Jacob.
- Lebovitz, S., Lifshitz-Assaf, H. & Levina, N. (2022). To engage or not to engage with AI for critical judgments: How professionals deal with opacity when using AI for medical

- diagnosis. *Organization science*, 33(1), 126-148. <https://doi.org/10.1287/orsc.2021.1549>
- MacLeod, P. D. & Eaton, S. E. (2020). The paradox of faculty attitudes toward student violations of academic integrity. *Journal of Academic Ethics*, 18(4), 347-362. <https://doi.org/10.1007/s10805-020-09363-4>
- Moscarola, J. (2024). Coupler analyse de données textuelles et IA génératives pour « augmenter » le processus de recherche qualitative. *JADT 2024 Mots comptés, textes déchiffrés* (pp. 653-662).
- Moscarola, J. & Chauhan, Z. (2024). Humain contre robot : Différencier l'expression des IA génératives de celle d'auteurs humains en utilisant l'analyse de données textuelles. *JADT 2024 Mots comptés, textes déchiffrés* (pp. 663-673).
- Moscarola, J., Dudézert, A., Fuhrer, C., Laval, F., Kalika, M. & Lavastre, O. (2022). Recherche qualitative pluridisciplinaire. CO-DATA-LAB. Dans F. Chevalier, L. M. Cloutier & N. Mitev (dirs.), *La recherche qualitative* (pp. 251-276). Éditions EMS, Coll. « Questions de société ».
- Ocal, A. (2023, mai). Perceptions of AI Ethics on Social Media. *2023 IEEE International Symposium on Ethics in Engineering, Science, and Technology (ETHICS)*, West Lafayette (pp. 1-10). <https://doi.org/10.1109/ETHICS57328.2023.10155069>
- Peterson, A. J. (2024). AI and the problem of knowledge collapse. *arXiv preprint arXiv:2404.03502*.
- Stiegler, B. (2016). *Digital studies organologie des savoirs et technologies de la connaissance*. FYP éditions.
- Uhaldeborde, J.-M. (2023) (dir.). *Les chemins de traverse de l'éthique académique*. Presses Universitaires de Pau et des Pays de l'Adour (PUPPA).
- Zhang, C., Zhang, C., Li, C., Qiao, Y., Zheng, S., Dam, S. K., Zhang, M., Kim, J. U., Kim, S. T., Choi, J., Park, G.-M., Bae, S.-H., Lee, L.-H., Hui, P., Kweon, I. S. & Hong, C. S. (2023). One small step for generative AI, one giant leap for AGI: A complete survey on ChatGPT. *AIGC Era Association of Computing Machinery*.

1. <https://chat.openai.com/share/c2624a20-d6fd-4c4b-8368-6bf4ec1a78a5>

CHAPITRE 9. DE L'IAÏSATION¹ DE LA GESTION DES RESSOURCES HUMAINES DANS LE MILIEU ACADÉMIQUE. CONSÉQUENCES EN TERMES DE DROIT À UN RECOURS JURIDICTIONNEL EFFECTIF GARANT DU RESPECT DES DROITS FONDAMENTAUX EN PRÉSENCE

*[Ghislaine Alberton](#), Professeure agrégée, Référente intégrité scientifique, Université de Pau
et des Pays de l'Adour (France)*

1. INTRODUCTION

« Il n'est guère de sujet plus en vogue que celui de l'intelligence artificielle. L'effervescence scientifique dont elle fait l'objet [...] aliment[e] [en effet] une production vertigineuse d'articles [et] d'actes de colloques [...] » affirme le Conseil d'État (2022, p. 15). Cependant cette effervescence se focalise quasi exclusivement sur les conséquences de l'utilisation, le plus souvent confidentielle, de l'IA par les chercheurs (questions des droits d'auteur, des éventuels plagats ou encore de l'hyperpublication...). Comparativement, la question de son utilisation administrative dans la fonction publique et plus précisément dans le secteur académique public demeure peu étudiée. L'utilisation de l'IA par l'administration concerne pourtant l'ensemble des enseignants-chercheurs. Elle pourrait s'avérer préjudiciable au respect de leurs droits fondamentaux, d'autant qu'elle les concerne à chaque étape de leur carrière. C'est pourquoi il est aujourd'hui primordial de s'interroger sur ce que pourraient être les conséquences d'une potentielle « IAïsation » de la gestion des ressources humaines dans le milieu académique.

Dans la sphère publique, « le défi prioritaire de l'acceptabilité sociétale des systèmes d'intelligence artificielle (ci-après SIA) repose en effet sur la mise en place d'une IA publique fondée sur des

principes structurants et dignes de confiance ». Selon le Conseil d'État, le premier de ces principes résiderait dans « la primauté humaine », principe en vertu duquel « derrière la machine, il y a toujours l'homme ». C'est que « l'idée selon laquelle un SIA ne saurait évincer l'humain et fonctionner en roue libre est en effet l'une des plus évidentes et des plus consensuelles dans la réflexion éthique et déontologique de ces systèmes » (2022, p. 93-102)². Ce principe, dit de « supervision humaine » ou de « contrôle humain », impliquerait donc que, dans notre domaine d'analyse (la gestion des ressources humaines en milieu académique), les SIA, lorsqu'ils existent, assureraient exclusivement une fonction d'aide à la prise de décision humaine (par exemple en formulant des suggestions). Cette première impression semble d'autant plus fondée qu'elle est confirmée par les dispositions de l'arsenal juridique en présence tant au niveau national qu'europpéen.

Au niveau français, ce principe se trouve en effet consacré par l'art. 47 de la loi du 6 janvier 1978 « relative à l'informatique, aux fichiers et aux libertés »³ dans les termes suivants : « **Aucune décision produisant des effets juridiques à l'égard d'une personne** ou l'affectant de manière significative **ne peut être prise sur le seul fondement d'un traitement automatisé** de données à caractère personnel. » Le principe paraît donc être entendu.

Oui, mais précisément, ce même article lui apporte dans son 2^e alinéa une notable **exception** concernant les « **décisions administratives individuelles** ». *A priori* et plus simplement, cela revient à dire que l'administration (et elle seule) se verrait réserver un traitement particulier (pour ne pas dire privilégié) en ce qu'elle ne serait en réalité aucunement liée par ce principe de supervision humaine et qu'elle pourrait donc, en toute légalité, prendre des décisions en matière de recrutement et de promotion académiques sous une forme entièrement automatisée *via* une IA générative. De là à dire que cet art. 47 serait en réalité le « cavalier de l'apocalypse numérique » (Delisle, 2024), il n'y a évidemment qu'un pas... que l'art. 47 se garde précisément de franchir, puisque, conformément

aux dispositions du RGPD et plus particulièrement de son art. 22, il subordonne l'application de cette « exception administrative » à la stricte observance de plusieurs conditions sur lesquelles, bien évidemment, nous reviendrons.

C'est ce même principe de « supervision humaine » que le récent règlement du 13 juin 2024 « établissant des règles harmonisées concernant l'IA » (règlement plus communément dénommé « IA Act »)⁴ vient promouvoir au niveau de l'Union européenne (ci-après UE). L'objet central de ce texte réside dans la protection des droits fondamentaux tels qu'ils sont consacrés par la Charte des droits fondamentaux de l'UE. Il hiérarchise les SIA de manière graduée en fonction des risques d'atteintes prévisibles à ces droits qu'ils présentent (Sénéchal, 2024). Ainsi, au sommet de l'échelle des risques, se trouvent (art. 5) des « SIA à risques inacceptables » et donc prohibés (manipulation mentale, abus de faiblesse, catégorisation biométrique notamment à des fins répressives...), alors qu'au bas de cette même échelle (art. 4) se trouvent les « SIA à usage général », soumis à une réglementation « allégée ».

En fait, c'est la catégorie intermédiaire qui nous intéresse ici, celle des « SIA à hauts risques » (art. 6). Cet article dresse en effet une liste de secteurs d'activité considérés comme tels, parmi lesquels figurent précisément « l'accès à l'emploi et la gestion de la main-d'œuvre », c'est-à-dire, notamment, « le recrutement et la sélection des personnes physiques (diffusion des offres d'emploi, évaluation et filtrage des candidatures), la promotion, le licenciement, le suivi et l'évaluation des performances des travailleurs ». Cette liste ne doit rien au hasard. Elle a été établie sur la base de nombreuses études démontrant les risques de discriminations envers les individus générés par les SIA développés dans ces domaines (Pons, 2022 ; Compani, 2023 ; Huttner, 2024).

Parce qu'elle est considérée comme un secteur « à hauts risques », la gestion des ressources humaines *via* un SIA (en particulier dans le secteur public) se voit donc soumise au respect d'un ensemble d'obligations destiné à garantir sa fiabilité et son

innocuité en termes de respect des droits fondamentaux des agents concernés. Et, parmi l'ensemble de ces obligations, se trouve précisément celle (art. 14) d'un « contrôle humain ».

Après avoir rappelé ce contexte normatif, nous proposons d'examiner un scénario dans lequel la prise de décision administrative en matière de recrutement et de promotion académiques serait entièrement automatisée *via* une IA générative. Selon ce scénario, il appartiendrait à la seule IA de déterminer par exemple qui simplifierait en disant qui peut être recruté ou promu : l'outil d'IA déciderait à partir d'un faisceau de critères préétablis⁵ tels que la durée de rédaction de la thèse, le nombre de publications, la renommée des revues éditrices (pour les critères scientifiques) ou la performance des étudiants aux examens standards (sur le plan pédagogique).

Se poserait alors immanquablement la question du respect des droits fondamentaux des agents publics destinataires de telles décisions : quid du droit « à la transparence », si l'ensemble des textes précités exige que ces agents soient expressément informés de l'utilisation d'un SIA pour la prise de décision et qu'ils bénéficient en conséquence d'un droit d'accès à la documentation du SIA concernant son mode de fonctionnement ? Quid du respect du principe d'égalité, en l'occurrence de non-discrimination en raison du sexe, de la résidence géographique ou de l'université d'origine par exemple⁶ ? Le respect dû aux droits fondamentaux poserait alors nécessairement la question du droit à un recours juridictionnel effectif contre les décisions générées par le SIA, recours sans lequel le respect des droits fondamentaux ne serait ni garanti, ni assuré⁷.

Ces questions sont d'autant plus essentielles que, parallèlement à l'action de l'UE (règlement « IA Act » précité), le Conseil de l'Europe a adopté, le 17 mai 2024, une convention-cadre « sur l'intelligence artificielle et les droits de l'homme, la démocratie et l'état de droit ». Cette convention ouverte à la signature de ses 46 États membres (ce qui inclut individuellement les 27 États membres de l'UE) ainsi

que 11 États tiers⁸ a été signée le 5 septembre 2024 par la Commission européenne au nom de l'UE⁹. Il s'agit là du tout premier texte d'ampleur internationale dans le domaine de l'IA qui soit juridiquement contraignant. Cette convention établit en effet un cadre juridique qui impose des exigences de transparence pour tous les SIA *via* notamment l'information des personnes amenées à interagir avec ces systèmes.

Les États signataires devront également adopter des mesures adéquates pour que les SIA respectent l'égalité, interdisent les discriminations et protègent la vie privée. Et ils devront aussi non seulement imposer l'obligation de rendre des comptes et de déterminer les responsabilités qui en découlent en cas d'impacts négatifs de ces systèmes, mais surtout, veiller à ce que les victimes de violations des droits fondamentaux du fait de l'utilisation de SIA disposent de voies de recours adéquates et bénéficient de garanties procédurales effectives (Rédaction JCP, 2024).

La cause est donc entendue. Reste dès lors à déterminer quels pourraient être les recours qui, exercés sur le plan national devant le juge administratif¹⁰, permettraient précisément d'assurer l'entier respect de ces nouvelles exigences européennes.

En fait, les voies de droit existant aujourd'hui en droit administratif sont d'ores et déjà suffisamment fournies pour offrir un panel de possibilités contentieuses. Leur régime juridique sera très certainement amené à évoluer sous l'impulsion d'un juge qui se montrera de toute évidence d'autant plus garant du respect de nos droits et libertés fondamentaux que la problématique de l'IA est spécifique. Le droit à un recours juridictionnel effectif devrait alors pouvoir porter (conformément à la dichotomie propre au droit administratif) aussi bien sur la légalité des décisions prises par le SIA (recours en annulation, cf. *infra* 1.) que sur la mise en œuvre de la responsabilité de l'administration amenée à réparer financièrement les préjudices qu'il aurait causés (recours en responsabilité, cf. *infra* 2.).

2. LE RECOURS EN ANNULATION

Ce recours, officiellement dénommé recours pour excès de pouvoir (REP), se définit comme celui par lequel on demande au juge d'annuler (donc de supprimer de manière rétroactive) une décision administrative à raison de l'illégalité dont elle serait entachée. Ainsi entendu, ce recours devrait donc pouvoir être exercé à deux stades : en amont (2.1.), ou en aval (2.2.) des décisions individuelles prises par le SIA et dont les requérants seraient les destinataires. Reste évidemment à déterminer s'il en est vraiment ainsi...

2.1. Un recours effectif en amont ?

Ce recours serait exercé à l'encontre de la décision même de recourir à un SIA afin d'obtenir l'arrêt et/ou la suppression immédiate du système d'algorithmes concerné. En vertu d'un principe général du droit consacré dès 1950, un REP peut en effet être exercé à l'encontre de toute décision administrative¹¹, y compris donc celle d'utiliser un SIA pour gérer le parcours professionnel des agents publics.

Cela étant, un obstacle juridique pourrait apparaître et faire échec à l'exercice d'un tel recours : la décision même de recourir à un SIA, plutôt qu'à une intervention humaine exclusive, pourrait en effet être considérée par le juge administratif comme relevant de la catégorie dite des « mesures d'ordre intérieur » et comme telle, soustraite à tout contrôle juridictionnel, car insusceptible d'appréciation juridique, sauf pour le juge à se substituer à l'administration en se transformant littéralement en administrateur. La question est évidemment fondamentale puisque qu'il en va du respect même des droits et libertés des agents publics...

Osons, en conséquence, supposer que le juge administratif accepte de contrôler une telle décision ; quels seraient alors les termes de son contrôle ?

Appliquant sur ce point les prescriptions de l'art. 22 du RGPD, l'art. 47 de la loi du 6 janvier 1978, subordonne, nous l'avons dit, la

possibilité pour l'administration de recourir à titre « exclusif » à un SIA à la stricte observance de plusieurs conditions.

Ces conditions sont les suivantes : mention explicite figurant dans la décision et informant son destinataire qu'elle est fondée sur un traitement algorithmique, interdiction de prendre en compte des données dites « sensibles » (état de santé, données biométriques et/ou génétiques, orientation sexuelle, origine raciale ou ethnique, opinions religieuses et politiques...), maîtrise par l'administration du traitement algorithmique et de ses évolutions afin de pouvoir expliquer, en détail et sous une forme intelligible, à la personne concernée la manière dont le traitement a été mis en œuvre à son égard et obligation, le cas échéant, de les lui communiquer sur demande. Sans qu'il soit besoin, pour la présente étude, d'analyser l'ensemble de ces exigences fondées sur le principe de transparence (et notamment le problème posé par le nécessaire respect du secret des affaires¹²), on notera que ces conditions sont exactement celles qu'on retrouve dans le règlement européen « IA Act » concernant les secteurs à hauts risques, au nombre desquels figure précisément la gestion des ressources humaines.

On notera surtout que, dans un arrêt de principe¹³ rendu le 7 décembre 2023, la Cour de justice de l'UE affirme (§ 65) qu'il « ressort du libellé même de l'art. 22 du RGPD que le droit national qui autorise l'adoption d'une décision individuelle automatisée doit prévoir des mesures appropriées pour la sauvegarde des droits et des libertés ainsi que des intérêts légitimes de la personne concernée ».

C'est dire que le contrôle auquel pourrait alors procéder le juge administratif serait déterminant pour le respect de nos libertés fondamentales. Mais à une condition *sine qua non* : que le juge ne voie pas dans la décision administrative de recourir à un SIA une simple « mesure d'ordre intérieur », auquel cas il n'y aurait d'autre alternative que celle d'attendre qu'une (première) décision individuelle soit prise *via* le SIA, afin que son destinataire exerce, en aval, un REP à son encontre.

2.2. Un recours efficient en aval ?

Le REP se focaliserait alors sur la conformité de cette décision aux règles de droit existantes¹⁴.

Riches, à cet égard, sont les enseignements tirés de la décision rendue par le Conseil constitutionnel le 12 juin 2018 à propos de la loi « relative à la protection des données personnelles »¹⁵, loi « réceptionnant » en effet les exigences communautaires nées du RGPD.

Les parlementaires requérants estimaient alors (§ 66) que « la loi contestée méconnaissait la garantie des droits en ce qu'elle autorisait l'utilisation d'algorithmes autoapprenants susceptibles de réviser eux-mêmes les règles qu'ils appliquent, empêchant, de ce fait, l'administration de connaître les règles sur le fondement desquelles la décision administrative a été effectivement prise » et, selon eux, la loi méconnaissait d'autant plus cette garantie qu'« il n'était absolument pas assuré que les règles appliquées par les algorithmes soient conformes au droit »...

En réponse, le Conseil constitutionnel s'est évertué à expliquer (§ 70) que « si les dispositions de cette loi autorisent effectivement l'administration à procéder à l'appréciation individuelle de la situation de l'administré par le seul truchement d'un algorithme [...], elles n'ont ni pour objet ni pour effet d'autoriser l'administration à adopter des décisions sans base légale ou à appliquer d'autres règles que celles du droit en vigueur ». C'était donc rappeler l'administration au nécessaire respect de la légalité. Et, pour lever toute ambiguïté à cet égard, le Conseil d'État a bien pris soin de préciser que toute décision administrative individuelle (prise *via* un SIA) « doit pouvoir faire l'objet de recours contentieux ».

Les termes de cette décision sont suffisamment clairs et explicites et ne nécessitent aucun commentaire supplémentaire... Ils posent cependant une question : ces recours contentieux pourraient-ils être exercés directement ou devraient-ils être obligatoirement précédés de ce qu'il est convenu d'appeler un recours administratif

obligatoire ? Dans certains cas, en effet, une décision administrative ne peut être déférée immédiatement au juge, parce qu'une disposition (le plus souvent législative) impose aux requérants potentiels de faire précéder leur recours juridictionnel de la saisine de l'administration à l'origine de la décision (recours administratif dit « gracieux »). La finalité de ce recours est d'ouvrir un espace de conciliation propre à régler la situation et donc d'éviter une voie contentieuse qui ne doit intervenir qu'en ultime recours. Une telle option est, du reste, appelée de ses vœux par le Conseil d'État lui-même... Selon lui, en effet, « l'administration doit rester en première ligne dans le contrôle de ses propres décisions. Il pourrait donc être envisagé qu'un tel recours administratif constitue un préalable obligatoire à l'introduction d'un recours contentieux. Ce faisant, l'administration corrigerait ses propres erreurs et, le cas échéant, le modèle algorithmique qui en est à l'origine, afin que soient portés devant le juge des litiges résultant de désaccords de fond et non de dysfonctionnements techniques ou d'une erreur de système » (2022, p. 147).

Reste dès lors à savoir si le législateur se fera un jour l'écho de cette proposition...

Si l'exercice même d'un recours contentieux, que ce soit en amont ou en aval de la décision administrative de recourir à un SIA, soulève de nombreuses questions d'ordre juridique conduisant à douter de l'effectivité d'une telle voie de droit, il convient toutefois de ne pas oublier qu'elle n'est, en réalité, pas la seule à pouvoir être utilisée.

Il est en effet une seconde voie qui tend à la reconnaissance au profit du requérant (ici l'agent public) d'un droit à des dommages-intérêts en réparation du préjudice que lui aurait causé l'administration et dont elle serait tenue pour responsable.

Cette reconnaissance peut évidemment faire suite à la reconnaissance préalable d'une illégalité *via* un REP (toute illégalité est en effet juridiquement considérée comme fautive). Mais elle peut également se faire indépendamment de tout REP dans le cadre d'un recours dit « de plein contentieux » ou « de pleine juridiction ». Eu

égard aux problèmes énoncés relativement au REP, cette dernière voie de droit mérite évidemment une attention toute particulière.

3. LE RECOURS EN RESPONSABILITÉ

Conformément au « mode d'emploi » jurisprudentiel traditionnellement appliqué en la matière, l'exercice de ce recours suppose que soit tranchées deux questions, celle de l'imputation de la responsabilité d'abord (3.1.), celle du fondement de cette responsabilité ensuite (3.2.).

3.1. De l'imputation de la responsabilité

Cette question est en réalité redoutable, car elle soulève des difficultés particulières quand l'utilisation d'un SIA est en jeu. Cela tient plus précisément à la multiplicité des acteurs et à leur inévitable imbrication dans le cycle de vie d'un SIA (Eynard, 2022).

Le dommage subi par l'agent public (l'atteinte à ses droits et libertés) peut en effet trouver sa source dans de multiples causes, que l'on peut schématiquement regrouper en deux catégories : un défaut de conception tout d'abord (mauvais choix d'algorithmes, défauts de programmation notamment), un défaut d'utilisation ensuite (utilisation d'un SIA « sain » à des fins pour lesquelles il n'avait pas été conçu, utilisation d'un SIA que l'on sait « vicié » en raison d'un dysfonctionnement constaté). Dans le premier cas, le responsable serait donc le fournisseur du SIA et dans le second son utilisateur, c'est-à-dire (dans notre cas d'espèce) l'administration. Les choses paraissent d'une simplicité désarmante¹⁶... sauf précisément pour l'agent public destinataire d'une décision administrative qui lui serait préjudiciable. À qui, en effet, peut-il et doit-il en imputer la responsabilité ? Comment peut-il, à son niveau, identifier l'origine même du dommage qui lui est causé ?

À cet égard, le régime de la responsabilité administrative se distingue de celui de la responsabilité civile. En effet, selon une jurisprudence constante, il n'appartient pas à l'administré (ici l'agent

public) de déterminer qui est le « réel » responsable de son préjudice. Il exerce « tout simplement » un recours en dommages-intérêts contre l'administration auteur de cette décision, laquelle est donc susceptible d'être condamnée « pour le tout », à charge pour elle de rechercher ensuite le véritable responsable : elle-même en sa qualité d'utilisateur ou bien le fournisseur du SIA qu'elle a utilisé.

Dans ce dernier cas, il lui appartiendrait alors d'exercer une « action récursoire » à l'encontre de celui-ci pour obtenir le remboursement de la somme versée à l'agent au titre de sa réparation.

L'identification du « responsable » étant faite (l'administration), il nous reste à déterminer le fondement juridique de sa responsabilité. C'est que celle-ci s'avère strictement conditionnée.

3.2. Du fondement de la responsabilité

En principe, et à l'instar du droit privé, la responsabilité de l'administration est une **responsabilité pour faute**. Dans notre cas d'espèce, la faute pourrait tout d'abord résulter de l'illégalité de la décision prise sur le fondement partiel ou exclusif d'un SIA (violation du principe d'égalité par exemple). Une décision reconnue illégale est en effet toujours fautive et, en conséquence, de nature à engager la responsabilité de l'administration dont elle émane.

La faute pourrait également résulter des agissements matériels de l'administration (utilisation à mauvais escient d'un logiciel « sain » ou d'un logiciel qu'elle sait « vicié », non-respect de l'obligation de transparence à laquelle elle est astreinte). La faute enfin, pourrait résulter du dysfonctionnement même du SIA (erreur de conception ou de programmation), alors même, nous l'avons souligné, que ce dysfonctionnement ne serait aucunement imputable à l'administration. Même non coupable, en sa qualité de « gardien de la chose », elle n'en est pas moins responsable de l'utilisation d'un SIA (Conseil d'État, 2022, p. 150). Dans tous les cas, il appartiendra à l'agent public de démontrer que ces fautes sont la « cause directe » du préjudice qu'il subit. Or, pour les raisons déjà énoncées

(multiplicité d'acteurs et imbrication des rôles), ce lien de causalité pourrait, en pratique, s'avérer très difficile à établir...

Mais justement... une **responsabilité sans faute** de l'administration devrait pouvoir être également retenue, même en l'absence de toute illégalité, erreur, mauvaise utilisation ou dysfonctionnement du SIA. C'est là une spécificité de la responsabilité administrative, qui est une « responsabilité de plein droit ». Il en existe en réalité deux types et tous deux devraient précisément avoir vocation à s'exercer dans notre cas d'espèce.

D'une part, la responsabilité de l'administration pourrait être recherchée **pour rupture de l'égalité devant les charges publiques**. Il en serait ainsi lorsque l'utilisation même « non fautive » d'un SIA entraînerait des conséquences gravement préjudiciables pour une catégorie de personnes bien déterminée (exigence dite de « la spécificité du préjudice » propre au régime de cette responsabilité et sans laquelle il n'y aurait pas, en effet, rupture de l'égalité) : par exemple, dans notre cas d'espèce, les femmes, les personnes de nationalité étrangère, les personnes résidant en outre-mer... Nous savons que le risque « systémique » de discrimination est propre à tout SIA et qu'au niveau européen, la gestion des ressources humaines est précisément un secteur classé « à hauts risques » pour cette raison. Encore faudra-t-il prouver « l'anormalité » du préjudice subi, c'est-à-dire, au sens de la jurisprudence, « un préjudice atteignant un certain degré d'importance » et « excédant [...] les gênes et inconvénients ordinaires de la vie en société ». Et précisément, eu égard aux risques systémiques de l'IA, où placer le curseur de la normalité ?

Le deuxième type de responsabilité administrative sans faute est dit justement « **pour risque** ». Il se rapporte aux dommages qui ont leur origine dans des « choses » (arme, ouvrage, engin public...), des « situations » (affrontement, manifestation, épidémie...) ou bien des « méthodes » (traitement médical novateur notamment) considérées comme dangereuses par le juge. Or, ne serait-il pas envisageable que celui-ci étende à notre matière la jurisprudence qui est la sienne concernant cette responsabilité pour risque, soit parce

qu'il considère que l'utilisation d'un SIA dans la gestion des ressources humaines consiste en une « méthode » potentiellement dangereuse (pour le respect de nos droits et libertés fondamentaux) ou plus simplement encore parce que l'IA générative constituerait en soi « une chose » dangereuse.

Mais alors, se poserait inmanquablement une dernière question...

4. CONCLUSION

... celle de savoir si juridiquement un système d'IA générative pourrait un jour être analysé non plus comme une « chose », mais bien comme un « sujet de droit » autonome, disposant d'une personnalité juridique et dont la responsabilité pourrait alors être recherchée comme telle par ses victimes. Pour sa part, le Conseil d'État estime qu'« il n'est aujourd'hui ni utile ni pertinent de conférer la personnalité morale au SIA lui-même » (2022, p. 11), car « en application du principe de supervision humaine, l'erreur de la machine n'est en réalité, rien d'autre juridiquement qu'une erreur humaine » (p. 108). Soit ! Il n'en reste pas moins que cette question est désarmante, car susceptible d'alimenter la peur d'une forme d'intelligence artificielle comparable à l'intelligence humaine, susceptible de rivaliser avec elle et même de se substituer à elle (Méheut, 2024). MAIS... que l'on se rassure ! Contrairement à certains de ses voisins européens¹⁷ et alors même que nombreux sont ceux qui appellent de leurs vœux « l'introduction d'une culture de l'IA générative dans le secteur public, y compris dans le domaine régalien » (Rapport parlementaire, 2024, p. 60 ; Rapport de la Commission de l'IA, 2024). « La France ne vit pas une révolution de l'IA publique..., les SIA étant encore assez peu exploités dans l'assistance à la gestion des ressources humaines (recrutements, gestion des affectations et des promotions) et n'assurant le plus souvent qu'une fonction d'aide à la décision » dit encore le Conseil d'État (2022, p. 7).

Alors oui, on peut être rassuré... mais pour combien de temps encore ?

Références bibliographiques

Rapports et avis officiels

Commission des lois de l'Assemblée nationale (2024). *Rapport d'information sur les défis de l'intelligence artificielle générative en matière de protection des données personnelles et d'utilisation du contenu généré* (N°2207 du 14 février 2024). https://www.assemblee-nationale.fr/dyn/16/dossiers/defis_ia_protections_donnees

Commission de l'IA (2024). *IA : notre ambition pour la France*. Rapport du 13 mars 2024. <https://www.economie.gouv.fr/cge/commission-ia>

Commission Nationale Consultative des Droits de l'Homme (CNCDH). (2022). *Avis A-2022-6 relatif à l'impact de l'intelligence artificielle sur les droits fondamentaux*. 7 avril 2022.

Conseil d'État (2014). *Le numérique et les droits fondamentaux* (Étude annuelle 2014). La Documentation Française.

Conseil d'État (2022). *Intelligence artificielle et action publique : construire la confiance, servir la performance* (Étude annuelle 2022). La Documentation Française.

Défenseur des droits (2020). *Algorithmes : prévenir l'automatisation des discriminations* (Avis en partenariat avec la CNIL). 31 mai 2020. <https://www.defenseurdesdroits.fr/rapport-algorithmes-prevenir-lautomatisation-des-discriminations-283>

Défenseur des droits (2024). *Rapport annuel d'activité*. 26 mars 2024. <https://www.defenseurdesdroits.fr/rapport-annuel-dactivite-2023-la-banalisation-des-atteintes-aux-droits-et-libertes-inquiete-la-597>

Journal Officiel de la République française (2022). N° 91 du 17 avril 2022, texte n° 99.

Ministère de la Transformation et de la Fonction publique (2023). *S. Guérini expérimente l'IA générative dans les services publics*. Communiqué de presse du 5 octobre 2023. <https://www.transformation.gouv.fr/files/presse/cp-experimentation-intelligence-artificielle-generative-services-publics.pdf>

Doctrine

Bensamoun, A. (2023). Maîtriser les risques de l'IA : entre éthique, responsabilisation et responsabilité. *La semaine juridique JCP*, Édit. Générale, 6 février 2023, (5), p. 181.

Compani, S. (2023). Éthique des IA au travail : vaste programme ! *Revue Internationale de la Compliance et de l'Éthique des Affaires*, 18 oct. 2023, (5), p. 236.

Delisle, E. (2024). Réflexions sur la conciliation entre intelligence artificielle et protection des données dans le monde du travail. *Actualité Juridique de la Fonction Publique* (AJPF), mai 2024, (5), p. 257.

Eynard, J. (2022). L'identification des acteurs dans le cycle de vie des SIA » *Dalloz Droit de la propriété intellectuelle et du numérique IP/IT*, février 2022, (2), p. 71.

Godefroy, L. (2020). Éthique et droit de l'intelligence artificielle : osmose ou symbiose ? *Recueil Dalloz*, 6 février 2020, (4), p. 231.

Huttner, L. (2024). Les femmes, l'IA et les systèmes algorithmiques. *La semaine juridique JCP. Édit. Générale*, 11 mars 2024, (10), p. 311.

Méheut, D. (2024). Notre arsenal juridique est-il suffisant face aux enjeux de l'IA ? *Les Petites Affiches*, février 2024, (2), p. 5.

Pons, R. (2022). Biais et discriminations dans les SIA. *Droit de la propriété intellectuelle et du numérique IP/IT*, février 2022, (2), p. 75.

Samman, T. & Vanssay, B. de (2024). Que retenir de la loi européenne sur l'Intelligence artificielle ? *Fondation Robert Schuman*. Paper n°757, 15 juillet 2024. <https://www.robert-schuman.eu/fr/questions-d-europe/757-que-retenir-de-la-loi-europeenne-sur-l-intelligence-artificielle>

Sénéchal, J. (2024). L'AI Act dans sa version finale – provisoire –, une hydre à trois têtes ». *Dalloz actualités*, 11 mars 2024.

(s.a.) (2024). Le Conseil de l'Europe adopte le premier traité international sur l'IA. *La semaine juridique JCP. Édition. générale*, 21(27 mai 2024), p. 669.

1. Ce néologisme est emprunté au Conseil d'État (2022, p. 17).
2. Sur le rôle respectif de l'éthique et du droit dans la régulation de l'IA : Godefroy, 2020.
3. Loi n°78-17 dans sa rédaction issue de la loi n° 2018-493 du 20 juin 2018 « relative à la protection des données personnelles ». Cette loi « réceptionne » les exigences communautaires nées du RGPD (Règlement Général sur la Protection des Données) n° 2016/679 du 27 avril 2016.
4. Ce règlement n° 2024/1689, publié au JOUE du 12 juillet 2024, est entré en vigueur le 1^{er} août 2024. Pour une présentation d'ensemble de ce texte, voir notamment Samman et de Vanssay (2024).
5. Selon la CNIL (avis du 25 avril 2024), tout recrutement doit être gouverné par le principe dit « de minimisation » selon lequel « le recruteur ne doit collecter que des données présentant un lien direct et nécessaire avec l'emploi proposé ou l'évaluation des compétences professionnelles du candidat » (*La Semaine juridique J.C.P. édit. Social* 30 avril 2024 n°17 p. 303).
6. Sur cette question, voir notamment l'Étude (2014) du Conseil d'État, le *Rapport* (2024) du Défenseur des Droits, l'Avis (2020) du Défenseur des droits et l'Avis (2022) de la CNCDH.
7. Ce droit a été consacré par le Conseil constitutionnel le 9 avril 1996 (96-373 DC) sur le fondement de l'article 16 de la Déclaration des Droits de l'Homme et du Citoyen (DDHC) aux termes duquel « Toute société dans laquelle la garantie des droits n'est pas assurée [...] n'a point de constitution ». On notera par ailleurs que ce droit est proclamé par l'art. 47 de la Charte des droits fondamentaux de l'UE et garanti par l'art. 78 § 1 du RGPD.
8. L'Argentine, l'Australie, le Canada, le Costa Rica, les États-Unis d'Amérique, Israël, le Japon, le Mexique, le Pérou, l'Uruguay et... le Saint-Siège.
9. Concernant l'UE, la convention sera mise en œuvre au moyen de l'*IA Act*, car jugée par la Commission « pleinement conforme » à celui-ci.
10. Nous nous limitons ici aux seuls recours exercés devant ce juge puisqu'il est le juge « naturel » de l'administration.

[11.](#) CE Ass. 17 février 1950, n° 86949 « Dame Lamotte ».

[12.](#) Pris « à la lettre », le respect dû au secret des affaires pourrait en effet constituer un rempart absolu à l'obligation pourtant imposée par la loi française et le RGPD de donner toutes informations utiles concernant les caractéristiques du SIA utilisé. Cela étant, l'article 15 DDHC aux termes duquel « la société a le droit de demander compte à tout agent public de son administration » devrait alors pouvoir être interprété par le juge comme imposant à l'administration l'obligation d'explicitier le fonctionnement des outils algorithmiques qu'elle utilise, cela sans pouvoir se retrancher derrière le (faux) prétexte du respect dû au secret des affaires. Telle est en tout cas la conviction du CE (2022, p. 118 sq.).

[13.](#) CJUE, 7 décembre 2023, aff. C-634/21 « Schufa Holding ».

[14.](#) Ne sont ici envisagés que les moyens de légalité interne tirés du respect des droits et libertés fondamentaux (notamment du principe d'égalité) à l'exclusion de tous les autres vices susceptibles d'entacher une telle décision (incompétence, vice de procédure ou de forme pour l'illégalité externe ; erreur de fait ou détournement de pouvoir pour l'illégalité interne).

[15.](#) Décision 2018-765 DC.

[16.](#) C'est du reste cette dichotomie que retient le règlement européen « IA Act ». Sur ce point, voir notamment Bensamoun, 2023.

[17.](#) Le Conseil d'État rapporte ainsi qu'en Suède, une municipalité expérimente depuis 2019 un « robot-recruteur qui se présente comme un boîtier surmonté d'une tête humanoïde ». Il a pour fonction d'analyser le comportement du candidat à l'embauche en lui soumettant des problèmes à résoudre ou des mises en situation. Il est capable de réagir aux réponses du candidat en formulant de nouvelles questions. L'objectif recherché est de supprimer tous les risques de discrimination inhérents à l'intervention humaine et tenant notamment à la prise en compte de l'âge, du sexe, de la tenue vestimentaire ou encore du physique du candidat (2022 p. 117).

CHAPITRE 10. LA CULTURE ÉTHIQUE DANS LES ORGANISATIONS DE RECHERCHE FACE AUX (IN)CERTITUDES LIÉES À L'INTELLIGENCE ARTIFICIELLE¹

[Susana Magalhães](#), *Institute of Research and Innovation in Health, i3S, University of Porto*
(Portugal)

A machine can be considered “intelligent” if a person interacting with it cannot tell whether it is human or computer. (Dignum, 2019, p. 120)

1. INTRODUCTION

En adoptant une approche moins centrée sur l'humain, Dignum (2019) définit l'intelligence artificielle comme « la discipline qui étudie et développe des artefacts computationnels qui présentent certaines facettes d'un comportement intelligent » (p. 10). Si on définit les agents intelligents comme ceux qui peuvent percevoir le contexte et réagir en trouvant les meilleures stratégies d'adaptation, qui sont proactifs dans la recherche des bons chemins pour atteindre leurs objectifs et capables d'établir des interactions avec d'autres êtres humains ou d'autres agents non humains, il est clair que les systèmes d'IA ne possèdent pas tous ces trois caractéristiques. Cependant, le progrès scientifique et technique rapide soulève la possibilité de systèmes plus sophistiqués dans un avenir prévisible. L'incertitude persiste quant aux avantages et aux risques que l'IA représente dans nos vies en général et dans la recherche en particulier.

La loi européenne sur l'IA (*Règlement 2024/1689*) souligne les risques potentiels liés à la manipulation des êtres humains, elle met tout particulièrement l'accent sur les groupes vulnérables, la discrimination engendrée par des algorithmes biaisés et la généralisation des systèmes d'identification biométrique, qui pourraient éventuellement mener à un contrôle à distance des individus et des populations. Le cadre de principes proposé par

l'Union européenne pour répondre à ces risques rencontre un écho important dans le monde scientifique : respect de l'autonomie humaine (en gardant les humains impliqués), c'est-à-dire respect de l'autonomie, de la dignité et de la liberté ; droit à la vie privée et à la protection des données ; équité ; bien-être individuel social et environnemental ; transparence ; responsabilité et surveillance. À ces principes, nous aimerions ajouter le rôle clé de la narration et des récits dans la gestion des incertitudes.

Notre époque est immergée dans une quantité d'informations qui tendent à discipliner nos communautés, à fermer l'accès à d'autres visions du monde et à fournir des recettes pour des individus sains, performants et productifs, au sein d'une société anesthésiée/palliative (Han, 2021, 2024). Réfléchir avec et sur l'éthique nécessite le cadre normatif de principes et de valeurs proposé par l'Union européenne, mais cela exige également d'écouter les récits de tous les acteurs concernés, qui vivent dans ces temps d'incertitudes masquées par la certitude offerte par les « conteurs ». Ainsi, nous faisons face à une époque particulièrement exigeante. Réfléchir éthiquement à l'intelligence artificielle n'a de sens que si nous saisissons l'opportunité de réfléchir sur la société qui nourrit ces systèmes d'IA. Dans l'écosystème scientifique, cela prend un sens tout particulier, car les critères utilisés pour évaluer la recherche reflètent en réalité des (contre-)valeurs pour la vie académique et la recherche : métriques, quantité et une carrière ascendante sans véritable projet de vie ou de carrière significatif.

La réforme du système d'évaluation des chercheurs, promue par la Commission européenne à travers la création de la Coalition pour l'Avancement de l'Évaluation de la Recherche (CoARA), peut faire une réelle différence si elle est effectivement mise en œuvre par les institutions académiques. Le texte de l'accord final, daté de juin 2022, stipule que les critères d'évaluation de la recherche doivent être centrés sur la qualité et valoriser l'originalité des idées, le professionnalisme et des résultats qui vont au-delà de l'état de l'art. La qualité implique que la recherche soit menée à travers des processus et des méthodologies transparentes ainsi qu'une gestion

de la recherche permettant la réutilisation systématique des résultats précédents. L'ouverture de la recherche ainsi que, le cas échéant, des résultats vérifiables et reproductibles contribuent fortement à la qualité. La réforme de l'évaluation de la recherche vise à mettre l'accent sur l'inclusion, la diversité, la qualité, l'ouverture et la collaboration. L'IA peut contribuer à cette réforme et façonner la culture et le climat des organisations de recherche, à la fois comme outil à utiliser dans l'évaluation et comme sujet de préoccupation concernant la conduite responsable de la recherche utilisant l'IA.

L'intelligence artificielle peut jouer un rôle très pertinent pour dévoiler la science bâclée et/ou frauduleuse, si on se rappelle que la quantité ne signifie pas la qualité. Il en va de même pour la réforme de l'évaluation des chercheurs, qui pourrait éventuellement devenir un scénario dystopique, dans lequel, même les quelques résultats pris en compte pourraient manquer de qualité ainsi que le souligne Chawla (2021) : « SciScore analyse un article en recherchant environ 20 éléments d'information différents, tels que la clarté du langage, la transparence des données et la méthodologie, et attribue une note sur 10 pour refléter sa rigueur scientifique et sa reproductibilité. Les études qui précisent les anticorps, les logiciels, les lignées cellulaires et les organismes transgéniques utilisés dans les expériences, par exemple, ont tendance à obtenir une meilleure note que celles qui fournissent moins de détails de ce type. »

Dans un article intéressant récemment publié sur la recherche biomédicale (González-Márquez *et al.*, 2024), l'IA a été utilisée pour cartographier diverses dimensions de ce domaine spécifique, à savoir : l'émergence de la littérature sur la Covid-19, l'évolution de la discipline des neurosciences, l'adoption de l'apprentissage automatique, le déséquilibre entre les genres ainsi que la concentration d'articles frauduleux rétractés. Nous utilisons ces propositions sans les modifier, dans le texte en encadré dans le cadre de notre communication interne pour indiquer la pertinence de cette approche.

Extraits de *The landscape of biomedical research*, González-Márquez, R., Schmidt, L., Schmidt, B. M., Berens, P., Kobak D. (2024). *Patterns*.

L'équipe de recherche indique avoir procédé à la recherche suivante :

Nous avons développé un atlas 2D de la littérature biomédicale basé sur la collection PubMed de 21 millions de résumés d'articles en utilisant un modèle de langage basé sur un transformateur (PubMedBERT) et une visualisation d'encastrement de voisinage (t-SNE) adaptée à la gestion de grandes bibliothèques de documents. [...] Les visualisations 2D des corpus textuels peuvent aider à dévoiler des aspects des données que d'autres méthodes d'analyse pourraient ne pas révéler.

Les auteurs font également référence au déséquilibre entre les genres dans l'attribution des auteurs, ainsi qu'à l'augmentation du nombre d'articles rétractés et à ceux qui pourraient également devoir être rétractés.

Nous avons identifié 11 756 articles signalés comme rétractés par PubMed et ayant encore des résumés intacts (ne contenant pas de termes tels que « rétracté » ou « retiré »). Ces articles n'étaient pas répartis uniformément sur la carte 2D, mais se concentraient plutôt dans plusieurs zones spécifiques, en particulier dans la partie supérieure de la carte, couvrant la recherche sur les médicaments liés au cancer, les gènes marqueurs et les microARN. Ces domaines sont connus pour être des cibles des « usines à articles », des organisations qui produisent des articles de recherche frauduleux à vendre. Notre carte est basée uniquement sur la similarité textuelle entre les résumés. Cela suggère que les articles non rétractés des régions à forte concentration d'articles rétractés pourraient nécessiter une enquête, car leurs résumés sont similaires à ceux des produits d'usines à articles. (p. 5)

L'équipe de recherche a conclu que :

Il y avait 3 572 articles supplémentaires dans leur carte qui n'étaient pas marqués comme rétractés dans PubMed, mais qui l'étaient en réalité. [...] Ils se trouvaient principalement dans les mêmes zones de la carte que nous avons identifiées comme suspectes précédemment, validant ainsi nos conclusions. Cela ne garantit pas que tous les articles dans ces zones soient frauduleux, mais confirme que notre carte 2D peut être utilisée pour mettre en évidence les articles nécessitant une enquête éditoriale plus approfondie. Si des usines à articles supplémentaires sont découvertes à l'avenir, notre carte aidera à mettre en lumière les clusters de littérature nécessitant un examen plus minutieux. (p. 5)

2. UTILISATIONS DE L'IA DANS LA RECHERCHE ET DIRECTIVES

Nous pouvons considérer deux utilisations différentes de l'IA dans la recherche : procédurale et méta-réflexive. En ce qui concerne la première :

- L'IA peut être utilisée tout au long des étapes du processus scientifique (génération d'hypothèses, conception d'expériences, suivi et simulation) jusqu'à la publication.

- La robotique de laboratoire peut automatiser des tâches répétitives telles que la manipulation de tubes à essai et de cultures cellulaires, entre autres, avec l'avantage d'éviter l'exposition humaine à des substances nocives ou à des radiations.

- Certains modèles d'IA sont capables de résumer des articles scientifiques ; ils mettent en évidence les éléments d'un brouillon de papier de recherche qui contribuent à une meilleure ou à une moins bonne compréhension de son contenu (Huang, 2018).

En ce qui concerne l'approche méta-réflexive, l'IA peut être très pertinente, car en utilisant l'IA pour accomplir certaines des tâches mentionnées ci-dessus, on peut économiser du temps pour la réflexion éthique. En effet, le temps nécessaire à la délibération éthique n'est pas le temps accéléré du développement des programmes d'intelligence artificielle, du *machine learning* et du *deep learning*.

En gardant à l'esprit que d'autres types d'organisations sont tenues de se conformer aux Objectifs de Développement Durable (ODD) et aux critères Environnementaux, Sociaux et de Gouvernance (ESG), nous proposons que l'utilisation d'une IA digne de confiance et responsable soit également alignée avec ces objectifs et ces pratiques. Les critères ESG concernent la responsabilité des organisations pour...

1. la protection de l'environnement ;
2. les interactions avec tous les groupes sociaux appartenant à son écosystème ;
3. le leadership, les audits et le contrôle interne.

Dans le rapport de l'atelier FORUM « Favoriser une recherche biomédicale plus verte » [*Enabling greener biomedical research*], organisé conjointement par l'Académie des sciences médicales [Academy of Medical Sciences], le Conseil de la recherche médicale

[Medical Research Council] (MRC) et l'Institut national pour la recherche en santé et en soins [National Institute for Health and Care Research] (NIHR), on souligne que « en étant prêtes à financer des options plus coûteuses, mais plus écologiques, les agences de financement peuvent envoyer un signal résolu sur l'importance qu'elles accordent à la durabilité environnementale » (2023, p. 26).

Nous pensons que le concept de durabilité dans la recherche devrait être réinterprété pour inclure non seulement les pratiques de recherche respectueuses de l'environnement, mais aussi la qualité de la culture et du climat de recherche au sein des ORP (organisations de recherche publique), ce qui constitue en réalité un pilier de la science durable.

L'intégrité de la recherche implique une conduite responsable de la part des individus et des institutions et doit être complétée par une dimension éthique qui englobe les relations entre toutes les parties prenantes. En intégrant les ESG dans l'intégrité de la recherche, comme il a été mentionné ci-dessus, on peut promouvoir un lien entre l'éthique et l'intégrité dans une approche holistique, en tenant compte non seulement de la responsabilité individuelle, mais aussi de celle de l'organisation. Compte tenu du rôle important de toutes les parties prenantes dans la recherche, l'édition révisée du *Code de conduite européen pour l'intégrité en recherche* (2023) rappelle à la communauté académique et à l'environnement de recherche que les organisations jouent un rôle clé dans la prévention et la détection des comportements contraires à l'éthique. L'objectif est de sensibiliser les scientifiques et de promouvoir leur responsabilité quant à l'impact fort et large de leur travail, qui doit respecter les normes élevées d'intégrité scientifique que les chercheurs eux-mêmes ont établies. La définition de l'éthique selon Paul Ricœur (1994), « visée de la vie bonne, avec et pour les autres, dans des institutions justes », souligne précisément la dimension relationnelle de l'approche éthique. En instaurant un dialogue entre les normes procédurales (conformité) et la délibération éthique, les chercheurs peuvent non seulement réfléchir à ce qui doit être fait et comment le faire, mais aussi aux valeurs qui sous-tendent ces normes et à la

manière dont ils y adhèrent eux-mêmes, non pas parce qu'ils doivent le faire, mais parce qu'ils le veulent (Magalhães, 2024). Citant Berling *et al.* (2019), Peuter (2021, p. 10) souligne que la conformité implique l'obéissance à des règles qui, par exemple, empêchent les pratiques frauduleuses, tandis que l'approche fondée sur la vertu « perçoit l'intégrité de la recherche comme une partie intégrante des bonnes pratiques scientifiques : ainsi, l'idée même de fabriquer des données viole ce que signifie être un scientifique ».

L'objectif est donc d'intégrer dans l'évaluation de la recherche une approche dialogique de son écosystème, qui inclut non seulement les chercheurs individuels et les organisations de recherche, mais aussi la culture et le climat dans lesquels ces deux agents agissent et interagissent. Cela est particulièrement pertinent pour prévenir les pratiques inacceptables et frauduleuses. Peuter (2021) souligne l'impact des facteurs systémiques, tels que la pression à la publication, les mécanismes d'évaluation pervers et le climat organisationnel, sur les inconduites en recherche et les pratiques de recherche douteuses. Il cite les contributions importantes de plusieurs auteurs sur cette question (Bouter, 2020 ; Gopalakrishna *et al.*, 2021).

Par « climat de recherche » nous entendons les comportements observables au sein des organisations de recherche (RPO), tandis que le terme « culture » désigne les valeurs et la mission qui confèrent à chaque RPO son identité propre. Culture et climat se combinent dans une approche dialogique entre les normes et lignes directrices et leur interprétation par les individus et les organisations dans leur ensemble : les questions à aborder sont donc : comment les valeurs se rapportent-elles à la façon dont les gens connaissent les choses dans la pratique ? comment les valorisent-ils ? comment les réalisent-ils ? comment attribuent-ils la responsabilité ? (Valkenburg *et al.*, 2021).

L'utilisation responsable de l'IA dans la recherche aura donc un impact sur ces deux dimensions des institutions de recherche. Les codes d'éthique et d'intégrité, à l'exemple du *Code de conduite*

européen pour l'intégrité en recherche (2023), devraient inclure l'utilisation de l'IA dans leurs lignes directrices, en fournissant des exemples pratiques de ce qu'il faut faire et éviter et en l'alignant sur la mission et les valeurs de chaque organisation. De plus, les exigences en matière de Société, Environnement et Gouvernance (SEG) pourraient fournir un cadre pour intégrer l'utilisation responsable de l'IA en tant que contributrice importante à l'évaluation et à la promotion de pratiques de recherche respectueuses de l'environnement et responsables envers différents groupes sociaux ainsi qu'à une gouvernance transparente et éthique.

La figure 1 illustre l'évaluation de la recherche basée sur la diversité et la qualité implique également une approche dialogique avec les communautés dans lesquelles l'organisation de recherche (RPO) opère.

Figure 1. Évaluation de la recherche en contexte



3. LES VALEURS PARTAGÉES

Sur la base de l'expérience de l'auteur en tant que responsable de l'intégrité dans un institut portugais de recherche biomédicale, cette section présente quelques points clés sur l'approche dialogique de la réflexion et de la délibération éthiques ainsi que sur son impact sur le climat éthique d'une organisation de recherche (RPO), y compris l'interaction avec d'autres parties prenantes de la science. Nous avons établi des lignes directrices fondées sur des valeurs pour une conduite responsable et garantissant la pérennité de ces valeurs, même lorsque les principes et les normes évoluent. Les valeurs qui ont sous-tendu l'élaboration des documents de politique (Gouvernance) concernant les dimensions éthiques et d'intégrité de la recherche sont l'honnêteté, le soin, le respect et la responsabilité.

Elles sont adaptées des recommandations du *Code de conduite européen pour l'intégrité en recherche* de l'ALLEA².

- **Honnêteté** : La transparence et la rigueur sont essentielles, en accord avec les pratiques de la science ouverte et en réponse aux préoccupations liées à la reproductibilité. La [checklist de l'intégrité de la recherche](#), la politique générale d'éthique/intégrité ainsi que les lignes directrices spécifiques concernant la [gestion des données](#) sensibilisent à la nécessité de mettre en œuvre la valeur de l'honnêteté à toutes les étapes du processus de recherche.

- **Soin** : Le respect de l'environnement et des collègues a été mis en avant dans l'[Initiative GreenLab](#), qui vise à promouvoir une recherche de plus en plus durable dans trois domaines principaux : l'économie d'énergie, la réduction des déchets et la durabilité sociale. En outre, le [Plan pour l'égalité des genres](#), adopté en 2022, reflète cet engagement. Cette organisation de recherche (RPO) a toujours eu une politique de non-discrimination et d'égalité d'accès, en faisant de l'équilibre entre les genres une priorité. Cependant, il était pertinent de formaliser ce Plan pour l'égalité des genres et de le rendre accessible à l'ensemble de la communauté. Les sessions d'accueil, organisées conjointement par l'Unité de conduite responsable et l'Unité de développement de carrière, offrent aux nouveaux membres l'opportunité de se familiariser avec les lignes directrices éthiques et d'intégrité de cette organisation. Elles soulignent également l'importance de parler ouvertement lorsqu'ils sont témoins d'inconduites et de pratiques inacceptables, telles que le harcèlement moral ou sexuel. On a également porté une attention particulière à la sensibilisation à la protection légale contre les représailles. En juin 2022, cette RPO a mis en place un canal de signalement, définissant et communiquant à toute la communauté sa Politique concernant les lanceurs d'alerte, qui assure leur protection conformément à la loi portugaise n°93/2021 du 20 décembre 2021 (Loi sur les lanceurs d'alerte).

- **Respect** : [Les directives pour gérer les conflits d'intérêts dans la recherche, ainsi qu'en ce qui concerne le recrutement et les](#)

[promotions](#), favorisent le respect envers les collègues et les candidats potentiels aux postes ouverts. Par ailleurs, les [directives concernant la politique de découvertes fortuites](#) sensibilisent à l'importance d'incarner la valeur du respect dans les consentements éclairés, en fournissant non seulement les informations requises sur l'étude (objectifs, procédures, risques et bénéfices, méthodes de recrutement, critères d'inclusion et d'exclusion, résultats de l'étude et partage des résultats), mais aussi en précisant ce qui est autorisé par les participants dans le cas de découvertes fortuites.

Compte tenu des conflits fréquents liés à la paternité des travaux, deux documents ont été publiés pour sensibiliser à l'obligation de respecter le travail de toutes les personnes ayant contribué à un article particulier, en précisant clairement que la paternité fantôme, la paternité invitée et la paternité offerte sont interdites : **le** Plansur la paternité intellectuelle **et les** Directives associées.

- **Responsabilité** : Mettre en place un cadre éthique qui distingue clairement entre éthique, morale et loi constitue la base de la formation à l'intégrité de la recherche dans cette organisation de recherche (RPO). Il est rappelé aux chercheurs que respecter la loi ne signifie pas nécessairement agir de manière éthique ; que l'éthique précède et suit la loi ; et que l'éthique et la morale nécessitent une approche différente des valeurs, des principes et des normes. Cela est particulièrement pertinent dans le domaine de l'IA. Les systèmes d'IA utilisés exclusivement à des fins de recherche scientifique ont été exclus du *Règlement sur l'IA* de l'UE en vertu de son article 2,6. Des préoccupations ont été soulevées quant aux éventuels écarts normatifs et à la manière d'appliquer ces exclusions. D'un côté l'innovation est essentielle, mais de l'autre des questions de sécurité et d'éthique doivent être prises en compte, ce qui nécessite une approche équilibrée de la réglementation de l'IA. Cet équilibre ne peut être atteint que si les organisations de recherche créent des espaces sûrs pour discuter de l'éthique et de l'intégrité tout au long du développement des systèmes d'IA : de la conception à la mise en œuvre et à l'utilisation, en sensibilisant

toutes les parties prenantes à la complémentarité entre conformité et éthique.

4. APPLICATION À L'INSTITUT PORTUGAIS DE RECHERCHE BIOMÉDICALE

Les chercheurs sont activement impliqués car nous les invitons à réfléchir sur divers documents normatifs et sur la manière dont les lignes directrices peuvent être mises en pratique dans leur vie professionnelle quotidienne. On les encourage également à partager les bonnes pratiques liées à une conduite responsable, plutôt que de se concentrer uniquement sur les inconduites et les pratiques inacceptables. Une série de sessions sur les *bonnes pratiques de reproductibilité* a été lancée en 2024 et se poursuivra en 2025.

Le responsable de l'intégrité organise régulièrement des réunions avec différents groupes de recherche sur divers thèmes, notamment l'inconduite (falsification, fabrication et plagiat), les pratiques inacceptables, la paternité intellectuelle, l'éthique de la publication, la gestion des données, la science ouverte, l'utilisation éthique de l'IA, le mentorat, la confidentialité et la protection de la vie privée, les questions éthiques dans les expériences humaines et l'autoévaluation éthique. L'objectif de cette approche est collaboratif, elle combine une approche ascendante (*bottom-up*) avec une application normative descendante (*top-down*) en tenant compte des piliers de la politique de l'UE sur la conduite responsable. Outre ces réunions réflexives et argumentatives, le responsable de l'intégrité fournit également un soutien aux chercheurs concernant la rédaction de documents d'autoévaluation éthique et de plans de gestion des données. Ce soutien dépasse les simples besoins de conformité, il est une occasion de revisiter les fondements des lignes directrices éthiques mises en œuvre et de combler les éventuels écarts entre les dimensions normatives et pratiques.

Comme nous l'avons mentionné précédemment, une checklist éthique générale a été publiée pour permettre aux nouveaux arrivants de se familiariser avec les principales thématiques et les

documents relatifs à la conduite responsable. Une section sur l'utilisation éthique de l'IA dans la recherche est actuellement en cours de rédaction et sera discutée lors de réunions de groupe avant d'atteindre sa version finale. Les principaux points de ces directives collaboratives pour une utilisation éthique et intégrée de l'IA peuvent être résumés ainsi :

1) La responsabilité est partagée entre les chercheurs et les développeurs de logiciels en ce qui concerne l'identification et le contrôle des biais liés à l'IA.

2) Les systèmes d'IA doivent toujours être explicables et il faut éviter le jargon difficile à comprendre pour les non-experts.

3) Chaque fois que nécessaire, nous encourageons l'engagement avec les communautés impactées et les autres parties prenantes et les incitons à prendre en compte leurs besoins et leurs préoccupations.

4) L'utilisation non éthique de l'IA et contraire à l'intégrité doit être considérée comme une faute passible de sanctions préalablement définies.

5) La paternité intellectuelle est réservée aux agents humains, car elle implique une responsabilité et une obligation de rendre des comptes, caractéristique que ne possèdent pas les systèmes d'IA eux-mêmes. Toute contribution de ces systèmes doit être reconnue et précisément décrite.

6) L'utilisation responsable des systèmes d'IA impose de respecter la protection des données personnelles. Des mesures de précaution doivent être prises pour éviter tout manquement aux principes fondamentaux de la confidentialité et du respect de la vie privée.

7) Les institutions soucieuses de recherche responsable devraient créer des espaces sûrs pour discuter de l'utilisation éthique de l'IA.

Enfin, une de nos mesures phares est un cours de formation en éthique et intégrité de la recherche qui est régulièrement dispensé depuis son lancement en 2021. Les quatre précédentes éditions ont été conçues pour s'adresser aux chercheurs à tous les stades de

carrière, tant de cette organisation de recherche (RPO) que de la communauté externe. L'objectif a été d'offrir une opportunité de discussion sur les questions éthiques et d'intégrité entre chercheurs issus de différentes institutions.

La qualité interdisciplinaire des sessions de formation, qui découle des divers champs scientifiques des participants, offre un environnement propice à la conception de nouvelles voies pour l'évaluation de la recherche. En écoutant les récits des uns et des autres, les chercheurs peuvent élargir leur perspective sur la diversité des parcours professionnels et des critères de qualité. Les questions de savoir ce qu'est le talent et ce qu'est une bonne recherche suscitent des réponses variées qui devraient être prises en compte dans la conception des systèmes d'IA destinés à évaluer la recherche.

De plus, partager différents récits peut sensibiliser à des problèmes qui n'ont pas encore été identifiés comme tels, en raison d'une injustice herméneutique (manque de familiarité avec certains concepts d'éthique et d'intégrité). Un avertissement clair se dégage concernant l'utilisation des systèmes d'IA pour évaluer la recherche : s'ils sont utilisés seuls, sans la boucle humaine, ces lacunes ne seront pas identifiées et le manque de compréhension peut même être perpétué par les algorithmes eux-mêmes. L'accent unilatéral mis sur les indicateurs quantitatifs doit être équilibré par des indicateurs qualitatifs, notamment en ce qui concerne l'éducation, l'impact sociétal et le leadership. Les systèmes d'IA peuvent être utilisés pour intégrer les critères qualitatifs dans l'analyse des articles de recherche et des CV (narratifs) des chercheurs, tout en étant adaptés aux caractéristiques spécifiques des différentes organisations et domaines scientifiques. C'est précisément cette diversité que met en lumière Lindemann et Häberlein (2023, p. 5), en prônant une approche contextuellement sensible des pratiques de science ouverte qui tient compte des *imperfections de nombreux environnements de recherche et du besoin, ressenti par les chercheurs, de concilier la science ouverte avec d'autres exigences*. Ainsi, au lieu d'une approche standardisée insensible aux contextes,

ce sont les lignes directrices de la science ouverte qui s'adaptent à la diversité du système scientifique dans le respect des principes d'une recherche responsable. Cette perspective devrait réconcilier les chercheurs les plus sceptiques avec la science ouverte et améliorer l'adhésion à ces pratiques au sein de la communauté scientifique.

5. DISCUSSION

Étant donné que les chercheurs peuvent être évalués sur l'utilisation responsable de l'IA et que leur travail peut être analysé par l'IA pour détecter des pratiques inacceptables, nous proposons que les organisations incluent l'utilisation responsable de l'IA dans leurs documents politiques et leurs codes éthiques.

L'IA peut être utilisée pour analyser la conformité de ces documents avec les critères Environnement, Société et Gouvernance (ESG) dans la recherche. Cela mettrait en avant le rôle de la culture de la recherche, au lieu de se concentrer exclusivement sur l'individu : « la responsabilité de réaliser des valeurs spécifiques constituant l'intégrité de la recherche (IR) est souvent attribuée aux chercheurs individuels par les codes de conduite de la recherche, bien que des valeurs telles que la transparence, le respect et la responsabilité puissent tout aussi bien être considérées comme faisant partie de la responsabilité d'une institution décente, ces attributions sont en effet souvent ambiguës » (Valkenburg *et al.*, 2021).

Nous acceptons que l'IA soit utilisée aux différentes étapes de la recherche ou dans la diffusion des connaissances, mais son utilisation doit être transparente et justifiée. Nous nous référons à *Living Guidelines on the Responsible Use of Generative AI in Research* (2024), récemment publiées par l'ERA FORUM, la responsabilité, la pensée critique, l'autonomie, l'honnêteté et l'équité sont considérées comme fondamentales.

Quelques exemples de recommandations des lignes directrices de ERA

FORUM³

Pour les chercheurs qui doivent rester responsables de leurs résultats scientifiques :

- Les chercheurs sont responsables de l'intégrité du contenu généré par ou avec le soutien des outils d'IA.
- Les chercheurs adoptent une approche critique vis-à-vis des résultats produits par l'IA générative et sont conscients des limites de ces outils, telles que les biais, les hallucinations et les inexactitudes.
- Les systèmes d'IA ne sont ni auteurs ni coauteurs. La paternité implique une responsabilité et un rôle actif qui sont réservés aux chercheurs humains.
- Les chercheurs n'utilisent pas de matériel fabriqué par l'IA générative dans le processus scientifique, par exemple en falsifiant, modifiant ou manipulant les données de recherche originales (p. 6).

Pour les organisations de financement de la recherche il faut exiger des candidats la transparence concernant leur utilisation de l'IA générative :

- Les candidats déclarent s'ils ont utilisé de manière significative des outils d'IA générative pour préparer leur candidature.
- Les candidats fournissent des informations sur le rôle de l'IA générative dans les activités de recherche proposées et réalisées (p. 9).

De plus, il doit être clairement indiqué comment ont été abordés les dangers et risques liés à l'utilisation de l'IA dans la recherche, tels que la probabilité d'interpréter une corrélation comme une causalité ou l'introduction de biais dans le système, qui risqueraient d'influencer les résultats. Selon l'OCDE (2023, p. 110), « les modèles d'IA apprennent simplement des corrélations dans les données, et non les relations causales impliquées. Des modèles causaux sont nécessaires pour distinguer la corrélation de la causalité ».

Un autre aspect de l'utilisation responsable de l'IA est l'approche interdisciplinaire, dans la mesure où « les innovations en IA pour la science sont souvent faciles à transférer d'un domaine scientifique à un autre » (OCDE, 2023, p. 111). Relvas et Ribeiro (2017, p. 231, notre traduction) nous rappellent que « la réalité est interdisciplinaire et sa perception est faite par un organe qui est structurellement et fonctionnellement interdisciplinaire, le cerveau ». L'IA appliquée à la science peut être un outil précieux pour une approche interdisciplinaire, tant en ce qui concerne les méthodes que les

contenus. L'évaluation responsable par les chercheurs devrait inclure la pensée interdisciplinaire et critique comme un critère à valoriser. En précisant comment l'IA a été utilisée dans une étude, un projet de recherche ou la rédaction d'un article, les chercheurs pourraient également inclure le processus de réflexion critique employé pour gérer les risques potentiels, les menaces et les biais des outils d'IA.

La nécessité d'une utilisation transparente de l'IA en général, y compris dans la recherche, peut rappeler les valeurs qui sous-tendent le *Code de conduite européen pour l'intégrité en recherche* (2023). De plus, les principales déclarations sur l'IA, telles que les *Principes* d'Asilomar et la *Déclaration de Barcelone*, peuvent également servir de phare pour souligner le rôle crucial de l'hétéro-régulation de la science et de la technologie : « La responsabilité en matière d'IA commence par un discours adéquat sur l'IA, qui démystifie les possibilités et les processus des technologies de l'IA et permet à tous de participer à la discussion sur le rôle de l'IA dans la société. » (Dignum, 2019, p. 101) Cela signifie que toute réflexion éthique entourant l'IA doit se placer dans le cadre plus général de l'éthique de la recherche et prendre en compte le contexte (parfois implicite) des valeurs humaines qui sous-tendent l'écosystème de production de connaissances. L'évaluation de la recherche faisant partie intégrante de cet écosystème, il s'ensuit que penser la responsabilité dans l'utilisation de l'IA implique également de penser la responsabilité de toutes les parties prenantes. En définitive, la réflexion sur l'utilisation responsable de l'IA *est une excellente occasion de réfléchir à la signification de l'éthique et de l'intégrité dans des environnements imparfaits* (Lindemann & Häberlein, 2023, p. 4).

Intégrer le développement des systèmes d'IA dans le cadre de valeurs humaines plus générales impose de repenser la notion de responsabilité en matière de transparence dans la recherche, plutôt que de se concentrer uniquement sur les résultats, l'efficacité et la performance. De plus, il ne faut pas oublier que l'évaluation du système scientifique doit primer sur celle des chercheurs individuels. Il faut mettre en place des incitations à adopter certains

comportements plutôt que de se focaliser uniquement sur les résultats individuels (Von Shomberg, 2024). L'incitation à publier des résultats nuls, non concluants, négatifs ou « non standards », à fournir des données ouvertes, à préenregistrer des études et à s'engager avec la communauté pourrait par exemple changer les comportements de recherche. Une recherche orientée vers les missions et défis sociétaux, comme celle que promeut Horizon Europe 2021-2027, s'aligne sur les valeurs de réactivité mutuelle, de réciprocité et d'ouverture. Les systèmes d'IA peuvent être utilisés pour simuler et tester des futurs souhaitables basés sur ces valeurs avant que les acteurs humains prennent des décisions finales. Une relation interactive et collaborative entre la science et la société (Von Shomberg, 2024, p. 10-11) est souhaitable et devrait constituer la base de la conception d'une IA responsable, en incluant ainsi les parties prenantes et les participants à la recherche. En résumé, nous proposons d'utiliser les directives pour un usage responsable de l'IA dans la recherche comme un outil permettant une meilleure évaluation et encourageant une réflexion plus approfondie sur la culture et le climat dans la recherche.

Références bibliographiques

- Berling, E., McLeskey, C., O'Rourke, M. & Pennock, R. T. (2019). A new method for a virtue-based responsible conduct of research curriculum: Pilot test results. *Sci Eng Ethics*, 25, 899-910. <https://doi.org/10.1007/s11948-017-9991-2>
- Bouter, L. (2020). What research institutions can do to foster research integrity. *Sci Eng Ethics*, 26, 2363-2369. <https://doi.org/10.1007/s11948-020-00178-5>
- Chawla, D. (2021). AI tools measure rigour of COVID-19 preprints : Automated reviews are picking up gaps in papers prior to peer review. *Nature Index*. Consulté le 9 mai 2024 sur <https://www.nature.com/nature-index/news/artificial-intelligence-tools-measure-scientific-rigour-covid-nineteen-research-preprints>.
- CoARa (2022). Agreement on reforming research assessment. Consulté le 6 mai 2024 sur https://coara.eu/app/uploads/2022/09/2022_07_19_rra_agreement_final.pdf.
- Code de conduite européen pour l'intégrité en recherche. https://www.allea.org/wp-content/uploads/2018/01/FR_ALLEA_Code_de_conduite_europeen_pour_lintegrite_en_recherche.pdf
- Dignum, V. (2019). *Responsible artificial intelligence: How to develop and use AI in a responsible way*. Springer. https://doi.org/10.1007/978-3-030-30371-6_2.

Directorate-General for Research and Innovation. European Commission (2024). Living guidelines on the responsible use of generative AI in research. https://research-and-innovation.ec.europa.eu/document/2b6cf7e5-36ac-41cb-aab5-0d32050143dc_en

González-Márquez, R., Schmidt, L., Schmidt, B. M., Berens, P. & Kobak, D. (2024). The landscape of biomedical research. *Patterns*. <https://doi.org/10.1016/j.patter.2024.100968>

Gopalakrishna, G., ter Riet, G., Cruyff, M. J., Vink, G., Stoop, I., Wicherts, J. M. & Bouter, L. (2021). Prevalence of questionable research practices, research misconduct and their potential explanatory factors: a survey among academic researchers in The Netherlands. Utrecht University Repository (Preprint). [https://doi.org/ 10.1371/journal.pone.0263023](https://doi.org/10.1371/journal.pone.0263023)

Han, B. (2021). *The Palliative Society*. John Wiley & Sons.

Han, B. (2024). *The Crisis of Narration*. John Wiley & Sons.

Huang, J.B. (2018). Deep paper gestalt. *arXiv:1812.08775*. <https://arxiv.org/abs/1812.08775>

Lindemann, T. & Häberlein, L. (2023). Contours of a research ethics and integrity perspective on open science. *Front. Res. Metr. Anal.*, 8, 1052353. <https://doi.org/10.3389/frma.2023.1052353>

Magalhães, S. (2024). Ethics and integrity in research: Why bridging the gap between ethics and integrity matters. *Journal of Academic Ethics* 22, 137-147. <https://doi.org/10.1007/s10805-024-09504-z>

OECD (2023). *Artificial intelligence in science: Challenges, opportunities and the future of research*. OECD Publishing. <https://doi.org/10.1787/a8d820bd-en>.

Peuter, S. (2021). Fostering a research integrity culture: Actionable advice for institutions [preprint]. ResearchGate. https://www.researchgate.net/publication/355264654_Fostering_a_Research_Integrity_Culture_Actionable_Advice_for_Institutions. Accessed May 02 2024.

Règlement européen sur l'IA (2024/1689) (2024). https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=OJ%3AL_202401689

Relvas, J. & Ribeiro, F. (2017). Manifesto pela interdisciplinaridade – Por uma educação transformadora. *Cadernos de Literatura Comparada*, 37, 229-250. <http://dx.doi.org/10.21747/21832242/litcomp37a10>.

Ricoeur, P. (1994). *Oneself as another*. University of Chicago Press.

The Academy of Medical Sciences (2023). Forum workshop report *Enabling greener biomedical research*. Consulté le 15 mars 2023 sur <https://acmedsci.ac.uk/file-download/61695123>.

Valkenburg, G., Dix, G., Tjldink, J. & de Rijcke, S. (2021). Expanding research integrity: A cultural-practice perspective. *Sci Eng Ethics*, 27(10). <https://doi.org/10.1007/s11948-021-00291-z>

Von Schomberg, R. (2024). Towards a new ethos of science or a reform of the institution of science? Merton revisited and the prospects of institutionalizing the research values of openness and mutual responsiveness (July 05, 2024). *Novation*, sixth special edition 2024, consulté sur SSRN : <https://ssrn.com/abstract=4892603>.

1. Remerciements : le travail développé par la chercheuse est financé par le FCT : CEECINST/00091/2018/CP1500/CT000.

<https://doi.org/10.54499/CEECINST/00091/2018/CP1500/CT0008>

2. <https://allea.org/code-of-conduct/>

3. https://research-and-innovation.ec.europa.eu/news/all-research-and-innovation-news/guidelines-responsible-use-generative-ai-research-developed-european-research-area-forum-2024-03-20_en

NOTICES BIOGRAPHIQUES DES AUTEURS

Ghislaine Alberton, ORCID : 0000-0002-3736-6271

Professeure agrégée, Université de Pau et des Pays de l'Adour, France, Ghislaine Alberton s'est spécialisée dans l'analyse des rapports normatifs et du processus d'interactions normatives entre le droit de l'Union européenne et l'ordre juridique national sur les plans constitutionnel, administratif et judiciaire. Les recherches qu'elle mène dans ce domaine s'inscrivent dans le droit fil des travaux conduits par le Centre de Documentation et de Recherches Européennes (CDRE. E.A n° 3004. UPPA Collège 2EI Bayonne) et font régulièrement l'objet de publications dans le cadre de revues juridiques spécialisées. Elles sont notamment consacrées à la fonction publique (sujet de sa thèse de doctorat obtenu à l'Université de Montpellier I en 1995). En janvier 2022, Ghislaine Alberton a été nommée Référente intégrité scientifique de l'UPPA.

Marie-Frédérique Bacqué, ORCID : 0000-0001-6223-4409

Marie-Frédérique Bacqué est psychologue, docteure en psychologie, Professeure de psychopathologie à l'Université de Strasbourg. Sa formation en anthropologie sociale à l'École des Hautes Études en Sciences Sociales (EHESS) lui permet de comparer les rites de passage après les grands maîtres français Arnold Van Gennep et Louis-Vincent Thomas. Rédactrice en chef des revues *Études sur la Mort* et *Psycho-oncologie* (jusqu'en 2022), elle dirige le CIEM (Centre International des Études sur la Mort) qui prend la suite de la Société de Thanatologie, fondée en 1966 pour étudier la mort et le mourir. Ses thèmes de recherche : le deuil, les traumatismes, la mort, les cancers, la psychanalyse. Elle a dirigé ou écrit 16 ouvrages scientifiques, 52 chapitres d'ouvrage, 131 articles dans des revues spécialisées nationales et internationales, 90 éditoriaux de revues scientifiques.

Michelle Bergadaà, ORCID : 0000-0001-6272-4061

Michelle Bergadaà est présidente de l'Institut International de Recherche et d'Action sur la Fraude et le Plagiat Académiques (IRAFPA). Elle a obtenu son Ph.D. à Montréal. Elle a publié plus de 150 articles et 10 livres dans son domaine avant de se consacrer, dès 2004, au développement des sciences de l'intégrité. Son livre *Le plagiat académique. Comprendre pour agir* (L'Harmattan), en 2015, suscite l'adhésion d'une communauté académique en quête de méthodes et de processus adaptés aux défis du XXI^e siècle. Elle fonde l'IRAFPA en 2016. Elle a publié, en 2020 *Le temps entre science et création* (Éditions EMS, Coll. « Les grands auteurs francophones »). Elle préside également la Fondation pour une Éducation Responsable et Équitable (FERE) et le Comité d'éthique de l'Université de Pau et des Pays de l'Adour (UPPA). Professeur émérite, Université de Genève, elle prend grand plaisir à diriger des thèses de doctorants de diverses nationalités. Elle donne des conférences au niveau international afin de renforcer le mouvement des sciences de l'intégrité.

Morgan Blangeois, ORCID : 0009-0006-9699-037X

Doctorant en sciences de gestion et du management à l'Université Clermont Auvergne, Morgan Blangeois mène ses recherches au sein du laboratoire Clermont Recherche Management (CleRMa). Titulaire d'un Master en Gestion de projet, Innovation et Transformation Numérique, il poursuit actuellement une thèse sur les enjeux de l'intelligence artificielle générative pour les entreprises de services du numérique. Ses recherches s'inscrivent à l'intersection du management stratégique, des systèmes d'information et des sciences de l'intégrité. Il s'intéresse particulièrement aux implications de l'IA générative pour la gestion stratégique et pour l'intégrité académique. Au sein de CleRMa, Morgan Blangeois est représentant des doctorants et membre du conseil scientifique. Il enseigne également la Data Science et l'IA générative en Master à l'IAE Clermont Auvergne.

Frédéric Bruneault, ORCID : 0000-0002-6963-6467

Frédéric Bruneault, Ph.D. (philosophie) est Professeur de philosophie au Collège André-Laurendeau et Professeur associé à l'École des médias de l'UQAM. Il est cofondateur du Laboratoire d'éthique du numérique et de l'intelligence artificielle (LEN.IA), chercheur associé au Groupe de recherche sur la surveillance et l'information au quotidien (GRISQ), à l'Observatoire international sur les impacts sociétaux de l'IA et du numérique (OBVIA) et à l'Observatoire du numérique en éducation (ONE). Il codirige des projets de recherche qui portent notamment sur l'enseignement de l'éthique de l'IA, sur la citoyenneté numérique et sur les outils d'évaluation en éthique de l'IA. Il agit aussi à titre de formateur et consultant en éthique de l'IA auprès de différentes organisations.

Rubén Comas-Forgas, ORCID : 0000-0002-8885-753X

Maître de conférences, Université des îles Baléares (Espagne), Rubén Comas-Forgas est titulaire d'un doctorat européen en Sciences de l'Éducation. Il occupe son poste actuel au Département de Pédagogie Appliquée et de Psychopédagogie de l'UIB. Il a été chercheur invité dans plusieurs centres internationaux, notamment aux universités d'East Anglia, John Moores de Liverpool, Panteion des Sciences Sociales et Politiques, à l'Université Autonome du Yucatan et à l'Université de Stockholm. Ses intérêts de recherche portent sur l'intégrité académique, la documentation et communication scientifiques, l'utilisation sociale et éducative des TIC, l'éducation à l'environnement et le décrochage scolaire précoce. Il participe à plusieurs projets de recherche d'envergure nationale en Espagne sur l'intégrité académique et le plagiat académique. Il coordonne le Réseau de Recherche Ibéro-américain sur l'Intégrité académique (Red-IA).

Cinta Gallent Torres, ORCID : 0000-0002-4260-7594

Professeure à l'Université de Valence (Espagne), Cinta Gallent Torres enseigne au Département de Philologie française et italienne. Titulaire d'un doctorat en intervention éducative et d'un double diplôme traduction/interprétation et français, elle a été assistante d'enseignement et chercheuse au Département des

Langues Modernes et Classiques de l'Université du Connecticut (États-Unis) où elle a obtenu une maîtrise en espagnol. Sa recherche se développe dans le domaine des technologies émergentes, de l'éthique de la publication scientifique, de l'enseignement des langues secondes et du tutorat des étudiants universitaires. Elle est membre de l'Institut de Recherche et d'Action sur la Fraude et le Plagiat Académiques (IRAFPA), du Réseau Ibéro-Américain de Recherche sur l'Intégrité académique (Red-IA) et de l'Alliance Forthem.

Ignace Haaz, ORCID : 0000-0003-1064-7994

Après son doctorat ès Lettres, philosophie, Université de Genève (Suisse), Ignace Haaz a été Assistant-docteur Université de Fribourg (Suisse). Aujourd'hui responsable des Éditions Globethics (Globethics Publications) à Genève, il a assumé la supervision et la production de trois cents livres sur l'éthique, à travers cinq continents, dont quarante-cinq thèses. Il est éditeur avec la professeure Amélé Ékué et Jakob Bühlmann Quero de la revue *Journal of Ethics in Higher Education* (JEHE), dont 4 numéros sont parus. Il a été nommé membre du Global Pool of Ethics Experts en mars 2023, par le Comité académique Globethics, présidé par professeur Rudolf von Sinner. Il est à l'origine d'une réalisation notable pour les sciences de l'intégrité : la création de la série Globethics Higher Education Series – Research Ethics.

Michel Kalika, ORCID : 0009-0004-2587-8650

Professeur émérite, Michel Kalika est Président du Business Science Institute, un réseau académique international proposant un Executive Doctorate in Business Administration (accrédité AMBA) en présentiel (10 sites) et en ligne dans quatre langues : français, anglais, allemand et espagnol. Auteur ou coauteur de plus de 25 livres (dont *Stratégie*, le livre francophone le plus réédité dans le domaine) et d'une centaine d'autres publications diverses en stratégie ou technologies de l'information. Ancien Professeur à l'iaelyon School of Management, Université Jean Moulin et à l'Université Paris Dauphine, où il a créé des programmes de DBA et

MBA, il fut aussi le premier Directeur de l'EM Strasbourg Business School. Il est aussi le fondateur de la méthodologie BSIS (Business School Impact System-EFMD-FNEGE), un processus qui a été utilisé à ce jour dans plus de 60 *business schools* à travers plus de 18 pays. 40 collègues lui ont consacré un livre en 2019 : *Entrepreneur à l'Université*, Éditions EMS.

Robert Laurini, ORCID : 0000-0003-0426-4030

Actuellement Professeur émérite en informatique, Robert Laurini a pratiquement fait toute sa carrière à l'Institut National des Sciences Appliquées (INSA de Lyon) avec des séjours d'un an à l'Université de Cambridge (UK) et à celle de l'Université du Maryland (USA), et pendant 10 ans à temps partiel à l'Université IUAV de Venise (Italie). Il parle français, anglais, italien et espagnol. Il a rédigé ou corédigé plus de 200 publications et 10 livres. Il a encadré 44 doctorants et a été membre du jury de thèse dans 19 pays. Il a créé l'ONG « Universitaires sans Frontières » dont le rôle est d'aider les universités des pays en voie de développement à se moderniser en liaison avec l'emploi local et en améliorant l'organisation de la recherche. Actuellement, il travaille sur l'intelligence artificielle dans le cadre des Smart Cities et il est bénévole au sein de l'association Coup de Pouce Université où il accompagne les doctorants.

Yves Frédéric Livian, ORCID : 0000-0001-7323-1254

Yves Frédéric Livian est Professeur honoraire à l'IAE, Université Jean-Moulin Lyon 3. Il a écrit seul ou en collaboration une trentaine d'articles et une quarantaine d'ouvrages dans le domaine de la sociologie des organisations, du management des ressources humaines et du management international. Étant passionné par la pédagogie, il a assuré également de nombreuses formations de méthodologie de la recherche. Il a été président du comité scientifique d'une agence nationale française. Il a enseigné en Europe centrale, Chine, Pérou et Afrique de l'Ouest. Il accompagne aujourd'hui des doctorants internationaux dans le cadre de l'association Coup de Pouce Université Lyon, où il assure les

enseignements de méthodologie intégrant les apports de l'IA générative.

Susana Magalhães, ORCID : [0000-0002-3440-4244](https://orcid.org/0000-0002-3440-4244)

Susana Magalhães est la coordinatrice de l'Unité de Conduite Responsable dans la Recherche à l'Institut de Recherche et d'Innovation en Santé (i3S) de l'Université de Porto. Elle est titulaire d'un doctorat en Bioéthique de l'Université Catholique Portugaise et est chercheuse dans les domaines de la médecine narrative et de l'éthique et de l'intégrité de la recherche. Elle est Professeure adjointe à l'Université Fernando Pessoa et collaboratrice du Forum d'Éthique de Católica Porto Business School de l'Université Catholique Portugaise. Membre du groupe fondateur du Réseau Portugais pour la Reproductibilité Scientifique ([PTRN](#)) ; membre de la Société Européenne de Médecine Narrative ([EUNAMES](#)) et formatrice certifiée du programme européen de formation des formateurs [VIRT2UE](#).

Tom Melvin, ORCID : 0000-0001-6370-4177

Tom Melvin est Professeur associé en Affaires Réglementaires des dispositifs médicaux au Trinity College de Dublin. Avant cela, il a été médecin-chef pendant 7 ans à l'Autorité de réglementation des produits de santé (HPRA). Il a coprésidé le groupe de travail sur l'investigation et l'évaluation cliniques de la Commission européenne pendant 5 ans, et a dirigé divers groupes de travail nationaux et internationaux sur des enjeux réglementaires et scientifiques liés aux dispositifs médicaux. Il a dirigé plusieurs groupes de travail et projets au niveau national et international sur des sujets réglementaires et scientifiques liés à la réglementation des dispositifs médicaux.

Jean Moscarola, ORCID : 0009-0008-5862-2132

Jean Moscarola a fondé la société Sphinx Développement dont il est aujourd'hui président d'honneur. Professeur des Universités, il a obtenu son diplôme d'HEC (France) avant de se diriger vers la recherche. Au cours de son doctorat en Économie Appliquée de l'Université Dauphine, il se spécialise dans les méthodes quantitatives. Très tôt cependant, son attirance pour le monde

concret et son attitude pragmatique le conduisent sur la voie de la recherche appliquée. Il tire parti des nouvelles technologies de l'information et de la communication pour renouveler les méthodes de recherches en sciences sociales et en marketing. Son expérience de chef d'entreprise et ses contributions au développement d'outils et méthode de recherche le distingue au sein de la communauté académique. Depuis plus de 10 ans, il accompagne les chercheurs du Business Science Institute en animant le BSI Lab. Expert dans l'usage des IA génératives il continue de renouveler les méthodes de recherche.

Wouter Oosterlinck, ORCID : 0000-0002-8313-3559

Wouter Oosterlinck, médecin et Professeur de chirurgie cardiaque à l'Université de Leuven (KULeuven), est un spécialiste reconnu dans la prise en charge des maladies coronariennes, tant sur le plan clinique que scientifique. Fort d'une expérience approfondie en chirurgie cardiaque, il maîtrise aussi bien les approches traditionnelles que les techniques minimalement invasives. Ses travaux de recherche actuels se concentrent sur des innovations de pointe, notamment la chirurgie cardiaque robotisée. Auteur et coauteur de nombreuses publications scientifiques, il a été distingué pour ses contributions à la recherche, en particulier sur la protection cardiaque des patients atteints de diabète de type II. En tant qu'expert pour l'Agence Européenne des Médicaments, il a développé une expertise sur la nouvelle législation MDR. Il est régulièrement sollicité pour évaluer des dispositifs innovants en chirurgie cardiaque, alliant ses compétences en réglementation et innovation technologique.

Paulo Peixoto, ORCID : [0000-0003-2275-667X](https://orcid.org/0000-0003-2275-667X)

Paulo Peixoto est Professeur associé avec habilitation à l'Université de Coimbra et chercheur au Centre d'Études Sociales. Il est membre de la direction de l'Institut International de Recherche et d'Action sur la Fraude et le Plagiat Académiques. Il a été l'Ombudsman des étudiants de l'Université de Coimbra (2019-2023). Depuis 2023, il est pro-recteur pour l'innovation pédagogique, les

politiques inclusives, le volontariat et les « repúblicas ». Il est président de la Direction de l'Association portugaise de sociologie et président du Council of National Associations de l'Association Européenne de Sociologie. Docteur en sociologie de l'Université de Coimbra (2007), Paulo Peixoto est titulaire d'une licence (1994, avec une partie de la formation effectuée à l'Université Catholique de Louvain) et d'un master (1997) en sociologie de l'Université de Coimbra.

Andréane Sabourin Laflamme, ORCID : 0009-0008-9663-5700

Professeure de philosophie, Collège André-Laurendeau, Andréane Sabourin Laflamme est candidate au doctorat, Faculté de Droit, Université de Sherbrooke. Elle est également titulaire d'une maîtrise en philosophie. Son projet de recherche doctoral, financé par le Conseil de recherche en sciences humaines du Canada, porte sur le rôle de l'éthique et du droit dans l'encadrement normatif de l'IA. Boursière de la Canadian Foundation for Governance Research. Chercheuse collégiale à l'OBVIA (Observatoire international sur les impacts sociétaux de l'IA et du numérique), elle est cofondatrice du LEN.IA. (Laboratoire d'éthique du numérique et de l'IA) où elle codirige des projets de recherche qui portent sur l'enseignement de l'éthique de l'IA, sur la citoyenneté numérique et sur les outils d'évaluation éthique des systèmes d'IA. En 2024, elle a reçu le Prix Relève Étoile du Fonds de recherche du Québec – Société et Culture (FRQSC) pour son article « Droit et soft ethics dans l'encadrement normatif de l'IA ».

Delphine Szeceł, ORCID : 0000-0001-5728-7338

Delphine Szeceł est chirurgienne cardiaque et vasculaire au Centre Hospitalier Universitaire de Liège, tout en exerçant également en tant que consultante aux Hôpitaux universitaires de Leuven. Dans le cadre de sa thèse de doctorat, elle mène des recherches innovantes sur de nouveaux modèles de valves cardiaques sans fixation. Sa curiosité scientifique, alliée à une ouverture d'esprit envers les approches interdisciplinaires, l'a conduite à s'intéresser aux systèmes de publication scientifique ainsi qu'aux enjeux

éthiques inhérents à la recherche. En parallèle de ses activités cliniques et de recherche, elle joue un rôle actif dans la formation des internes en chirurgie. Elle enseigne la chirurgie cardiaque dans le cadre du programme de master en Santé Publique, participant ainsi au développement de la prochaine génération de professionnels de la santé.

Pedro Urbano, ORCID : 0000-0002-7547-3035

Docteur en Psychologie (spécialité : Théorie et Histoire de la Psychologie), Professeur et chercheur à la Faculté de Psychologie et des Sciences de l'Éducation, de l'Université de Coimbra (FPCEUC), depuis 1997. Il enseigne notamment la psychologie évolutionniste, l'épistémologie et l'éthique de la psychologie, qui sont aussi les domaines de recherche dans lesquels il s'est spécialisé. Il est membre intégré du Centre de Recherches Interdisciplinaires (CEIS20) de l'Université de Coimbra et membre fondateur du Centre International des Études sur la Mort/International Center for Death Studies (CIEM/ICDS). Il a été membre de la commission d'éthique de la FPCEUC et il a fait partie du comité de rédaction de la revue *Psychologica*. Depuis 2019, il fait partie du comité de rédaction de la revue *Études sur la mort* (publiée par Cairn).

Alexandre Zollinger, ORCID : 0000-0002-0617-9254

Alexandre Zollinger est Maître de conférences HDR, Université de Poitiers, France. Après un doctorat de Droit privé consacré aux relations entre droit d'auteur et droits de l'Homme, Alexandre Zollinger a été recruté à l'Université de Poitiers. Il y codirige le Magistère en Droit des T.I.C./Master mention Droit du numérique. Spécialiste de propriété intellectuelle, il a rédigé deux ouvrages en cette matière, dirigé ou codirigé deux ouvrages collectifs et est responsable de deux chroniques de jurisprudence (au sein de *La Semaine Juridique – Entreprise et Affaires* et de la *Revue Semestrielle de Droit Animalier*). Il intervient mensuellement au sein de la revue *L'ESSENTIEL Droit de la propriété intellectuelle*. Chargé de la sensibilisation à la question du plagiat dans le cadre de la formation assurée à l'École Doctorale Pierre Couvrat, il a consacré

certaines de ses publications à ce sujet ou à des thématiques connexes.