

## Actes du 3e Colloque International de Recherche et Action sur l'Intégrité Académique, IRAFPA, Université de Coimbra, 20-22 juin 2024

“Défis et incertitudes de l'intégrité académique à l'ère de l'intelligence artificielle”

---

### **Interprétation et littéralité : l'« intégrité » du texte remise en cause par les IA génératives ?**

Jacques Athanase Gilbert, Professeur des universités, Nantes Université (France)

**ORCID :** [0000-0003-4808-3119](https://orcid.org/0000-0003-4808-3119)

**DOI :** <https://doi.org/10.56240/dfart05> (provisoire)

**Mots-clés :** texte, literacy, format, auctorialité

**Keywords:** text, literacy, format, auctoriality

**Résumé :** Il s'agit d'examiner dans le domaine des lettres et en SHS, les conséquences des modifications du statut du texte avec l'utilisation des IA. Le point de départ de cette présentation est une expérience désagréable de remise par un étudiant d'un manuscrit très largement plagié dont je découvre qu'il est fabriqué comme un centon de bribes de livres et d'articles assemblés selon une méthode toutefois aberrante. Il arrive que dans une page soient entrelacés des fragments de phrases et des expressions empruntés à des ouvrages tout à fait hétérogènes : des articles critiques, des comptes rendus de discours, des quatrièmes de couverture comme une sorte de patchwork. Je me doute alors que le travail a probablement été écrit avec une intelligence artificielle. Les emprunts sont incohérents sur le plan scientifique mais ils composent un texte à la fois lisible et improbable sur le plan scientifique. Je préviens mon établissement, l'université de Nantes et refuse que la thèse soit soutenue. C'est donc bien l'intégrité de même du texte qui était affecté par l'auteur. L'œuvre et son interprétation forment un tout problématique. Le « plagiat » est par conséquent une question des études littéraires qui interroge au plus profond la matérialité des œuvres et leur signification.

**Abstract:** The aim is to examine, in the field of humanities and social sciences, the consequences of changes in the status of the text with the use of AIs. The starting point for this presentation is the unpleasant experience of a student handing in a largely plagiarized manuscript, which I discovered to be made up as a centon of snippets from books and articles, assembled using an aberrant method. On a single page, sentence fragments and expressions borrowed from completely heterogeneous works are intertwined: critical articles, reports of speeches, back covers in a sort of patchwork. I then suspect that the work was probably written with artificial intelligence. The borrowings are scientifically incoherent, but they make up a text that is both readable and scientifically improbable. I warned my institution, the University of Nantes, and refused to allow the thesis to be defended. It was the very integrity of the text that was affected by the author. The work and its interpretation form a problematic whole. "Plagiarism" is therefore a question of literary studies that raises profound questions about the materiality of works and their meaning.

## 1. Introduction

Je vais commencer par le récit d'une anecdote personnelle assez peu significative sans doute en tant que telle à ceci près qu'elle a pris dans mon travail de recherche une place paradigmatique. Cela fait un moment que je m'intéresse avec plusieurs de mes collègues, dont Madame Vignon à l'impact des technologies digitales sur les disciplines en Lettres et Sciences Humaines. La proposition de cette communication est de tirer d'une expérience effective un motif de réflexion sur l'« intégrité » du texte à l'ère des IA générative, mais aussi peut-être de dessiner quelques pistes pour le futur tel qu'il se présente aujourd'hui.

Le point de départ de cette présentation est une expérience désagréable de remise par un étudiant d'un manuscrit très largement plagié dont je découvre qu'il est fabriqué comme un centon de bribes de livres et d'articles assemblés selon une méthode toutefois aberrante. Il arrive que dans une page se trouvent entrelacés des fragments de phrases et des expressions empruntés à des ouvrages tout à fait hétérogènes : des articles critiques, des comptes-rendus de discours, des quatrièmes de couverture comme une sorte de patchwork. Toutefois il ne s'agit pas exactement d'un plagiat au sens que je peux m'imaginer, celui d'un emprunt de passages excellents ou significatifs. Il s'agit d'un plagiat « inspiré » comme on dit, ce qui signifie que les guillemets ont été « oubliés » selon la litote généralement utilisée dans ce genre de circonstances. La situation de plagiat dépend largement du statut des textes. Et même du fait qu'ils soient des textes éditorialisés. Une étudiante en thèse que je rencontre régulièrement et qui travaille sur des sujets proches des miens me demande un jour si elle doit citer ce que je lui dis parfois. Je lui réponds que le travail que nous faisons constitue de fait une communauté de travail et qu'il nous appartient non pas de manière forcément égale mais certainement commune. On ne confond pas Braque et Picasso mais certaines de leurs productions sont à peu près indiscernables. Le reste relève de la courtoisie et même de l'amitié qui préside à toute pensée. Un jour, dans un livre consacré à un sujet que j'avais travaillé, je retrouve des expressions que j'avais utilisées dans ma thèse presque mot pour mot. L'emprunt est évident. L'auteur ne met pas de guillemets et il ne me mentionne pas dans sa bibliographie. Je suis agacé. Sans doute un peu content quand même: il est plus flatteur d'être le plagié que le plagiaire. Surtout l'emprunteur n'est pas allé au bout de ma démonstration. Il n'a pas compris où je voulais en venir et s'est comporté en emprunteur pressé aux mauvaises manières. Si l'objet du plagiat est la doxa commune, on s'expose à recycler indéfiniment les mêmes lieux communs. L'art du plagiaire, comme celui du faussaire, demande quand même un peu de talent. Il devient alors un art, une manière, un « style » au sens où l'entend Marielle Macé qui porte avec lui sa manière de vivre. Une imitation.

## 2. Évolution

Il faut reconnaître que les usages ont largement changé. Les ouvrages du XIXe siècle et même ceux de la première partie du XXe sont moins chargés de notes et de référence qu'aujourd'hui. Cela tient à plusieurs raisons. La plus évidente est certainement le caractère plus large de la diffusion et le fait que la société de « réception » est plus large et par conséquent plus indéterminée. Il est intéressant de constater que le contemporain de Platon, Isocrate dans un discours consacré aux sophistes (*Aux sophistes*) semble partager la réticence de Socrate envers l'écriture mais l'argumente d'une façon plus explicite. Que signifiaient dans *Le Phèdre* ces paroles mortes et dépourvues de leur paternité ? Isocrate donne une forme de réponse de nature politique et quasiment « sociologique ». La connaissance d'un maître permet l'établissement d'un dialogue comme l'illustre la forme dialogique pratiquée par Platon. Le dialogue est une forme « traversante » et interagissante. L'écriture fige un discours qui ne peut désormais plus « répondre », du moins d'une manière aussi immédiate. Il y a toutefois une nuance sémiotique que la distance historique ne nous permet peut-être pas d'appréhender. C'est le maître qu'on imite plus que ses seules paroles, surtout si elles sont transcrites. On trouve chez Derrida, et ce dès la *Grammatologie*, une critique du phonocentrisme au profit d'une « archécriture » fondamentale. Les débats qui ont opposés Jacques Derrida à Jack Goody sont

de ce points de vue très intéressants. Pour le premier, toute différence sémiologique est déjà « écriture » de par son pouvoir différenciant. Pour Jack Goody, au contraire, il existe une distinction beaucoup plus marquée entre un monde de l'oralité qui ignore la pratique de l'écriture et celui de la scription au sein duquel, les « outils de l'intellect » façonnent à la fois l'oral et l'écrit.

Ce qu'on pressent à la lecture du Phèdre, mais aussi du Discours aux sophistes tient à cette distinction. Platon a certes beaucoup écrit mais il semble, comme Isocrate, éprouver cette nostalgie d'une parole qui ne peut être séparée d'elle-même parce qu'elle tient toujours à la situation spécifique de celui qui l'énonce. Socrate n'a rien laissé d'écrit. Il devient une sorte de « personnage » platonicien qu'on retrouve aussi chez Aristophane et Xenophon. Et de manière indirecte chez Aristote. Après lui semble s'instituer de manière irréversible un certain monde de l'écrit qui fait éprouver des regrets à Platon et Isocrate. Pour quelles raisons ? Isocrate donne quelques réponses : les paroles que peut dispenser un maître ne peuvent être comprises sans qu'il soit là. Le risque, en son absence, est que ses propos ne soient pas compris et leur sens dévoyé. Ce qui compte est moins la teneur de la parole elle-même que les conditions de son énonciation. Le maître interagit avec le disciple et le charisme disparaît si on se contente de la transcription des paroles. Il manquera toujours la modulation de la voix, le rythme du discours et son adaptation aux circonstances. En un sens, la tâche de l'herméneutique, telle qu'elle se trouve envisagée à partir du XIXe siècle en Allemagne, mais peut-être plus tôt encore en France avec les « lectures » bibliques, je pense au Platonisme dévoilé du protestant Jacques Souverain, vise moins à saisir ce qui est impossible à saisir que de susciter une sorte de vie propre. « La lettre tue, mais l'esprit vivifie » (Corinthiens, 3 : 6), écrit l'apôtre Paul dans un tout autre contexte que celui d'Isocrate. L'enjeu est différent de celui d'Isocrate. Le « maître » est le Christ ressuscité dont il faut maintenir la parole vive, la « bonne nouvelle ». Le lien du texte sacré à son lecteur est désormais personnel et orienté. Le moins que l'on puisse faire est de respecter la « bonne nouvelle » sans l'ensevelir sous les strates d'une tradition exogène. D'inaugurer ainsi une sorte d'archéologie du sens qu'on doit, comme Luther et Calvin l'ont fait pour les reliques, débarrasser de sens ajouts. Dans le premier christianisme, l'imitation n'est pas textuelle ou figura le, elle est de type sacramentel, comme le déclare Ignace d'Antioche quand il se déclare « pain du Christ ».

### 3. Le plagiat

Nous étions partis d'un cas problématique de plagiat. Il semble bien que nous ayons opéré un détour. Pas vraiment. Le manuscrit plagié que j'avais lu m'avait surtout paru dépourvu de toute saillance. Il est fréquent, me semble-t-il que les textes plagiés vous « tombent des mains » littéralement. La raison en est simple à comprendre : ils sont dépourvus d'une organisation propre. Le critère est toutefois trop subjectif et bien moins probant que l'identification de chaînes de caractères identiques. Il faut ajouter que dans les travaux en littérature, on cite beaucoup et on se trouve amené à paraphraser parfois. En réalité, je n'ai pas décelé le plagiat. C'est ma collègue Daphné Vignon qui l'a identifié. J'ai d'abord cru qu'il n'était effectivement qu'un « oubli » de guillemets et j'ai demandé à l'étudiant de replacer les guillemets. C'est alors seulement que Madame Vignon et moi-même nous sommes rendus compte que le plagiat était généralisé. Et pourtant très peu visible pour deux raisons. La première tenait à un rewriting soigné qui remplaçait les mots expressions originales par d'autres à peu près équivalentes. Le logiciel de comparaison de texte avait identifié certaines ressemblances de chaînes de caractères mais rien de vraiment significatif. La thèse était en co tutelle mais aucun des membres du jury proposé par la partie étrangère, francophone, je tiens à la préciser, n'avait levé le lièvre.

La seconde raison tient à une spécificité que je n'aurais jamais soupçonnée si je ne l'avais eue sous les yeux. Le texte était littéralement un centon comme l'est par exemple le *Christus patiens*

de Grégoire de Nazianze qui est réalisé à partir de fragments de tragédies d'Euripide. Il s'agissait, au sens strict, d'un « bricolage » au sens ou l'entend Claude Lévi-Strauss. Les éléments du système original étaient conservés mais leur fonction dans la structure du nouveau système se trouvait entièrement modifiée. Jamais la métaphore « textile » du texte ne me parut mieux appropriée. Chaque page s'avérait un patchwork de pièces issues de corpus différents, assez courtes pour ne pas permettre facilement une identification par chaîne de caractères. La plupart du temps les morceaux n'avaient pas de lien entre eux. Une page pouvait ainsi mettre bout à bout des extraits d'un discours de réception à l'Académie française, d'un quatrième de couverture d'un ouvrage de Paul Ricoeur, un morceau de thèse soutenue à Montréal, quelques lignes arrachées à un article de tourisme issu du Magazine Géo. Il y avait un texte, on peut même dire « hypertexturé » mais pas de thèse ! En un sens, je me sentis soulagé, un peu à la manière de quelqu'un qui souffre d'un syndrome inexplicable. Le fait d'avoir été trompé ne me consolait pas de m'être trompé moi-même. La thèse ne fut pas soutenue mais je demandai cependant à mon université d'entreprendre une démarche disciplinaire. Je n'ai eu depuis près d'un an et demi aucune nouvelle ! C'est à ce moment que je pensai que l'étudiante avait eu accès à un robot conversationnel du genre de ChatGPT. À ce moment, j'entrepris pour mes étudiants de L2 un cours intitulé « Vivre avec l'intelligence artificielle » et je leur ai proposé cette année de pouvoir utiliser une IA pour l'examen à condition d'expliquer dans le détail le protocole de leur usage et de développer une petite partie critique. Cette année j'ai reçu environ 400 copies de L2 réalisée selon cette méthode. Cette évaluation un peu expérimentale s'est avérée riche d'enseignements.

Je me doute que le travail a probablement été écrit avec une intelligence artificielle. Les emprunts sont incohérents sur le plan scientifique mais ils composent un texte à la fois lisible et improbable sur le plan scientifique. Je préviens mon établissement, l'université de Nantes et refuse que la thèse soit soutenue. Le taux de plagiat est détecté par les logiciels anti-plagiat mais il n'apparaît pas forcément à la lecture. Les membres du jury franco-marocain ne l'avaient pas détecté au premier abord. Pour nous littéraires, il s'agit d'un « objet » d'un type nouveau dont le possible développement pose problème s'il se généralise.

La situation en Lettres pose certaines difficultés particulières dans la mesure où il s'agit d'un domaine où la notion d'auteur et celle de la spécificité d'une œuvre n'ont été admises que tardivement. L'emploi et le réemploi de motifs ont longtemps été la norme aussi bien dans les « œuvres » (sources premières) que dans l'interprétation dont ils ont fait l'objet (sources secondaires). Les « idées » sont rarement exclusives, contrairement aux brevets scientifiques et techniques. Il ne reste à protéger qu'un « droit d'auteur » par un dispositif de références qui tient essentiellement à la présence de guillemets et de notes. Mais l'écosystème est fragile et les potentialités de réénonciation produites par des IA génératives qui ne citent pas leurs sources remettent en cause toute protection du travail intellectuel dans ces domaines.

#### **4. Le cas particulier des études en Lettres et Sciences Humaines et Sociales**

Sur le plan pédagogique, les conséquences sont également assez consistantes. Une partie du travail s'effectue de manière plutôt « qualitative » si bien qu'il n'existe pas de norme unique et mesurable qui permette de déterminer un résultat de manière absolue. Le savoir ne tient pas tant à une mesure qu'à une pratique d'écriture et une méthode qui ne peuvent jamais donner lieu à des certitudes. En littérature par exemple, la connaissance des modalités du récit, qui constitue un des objets d'étude, ne peut être appréhendé comme une science exacte et cela même si les approches linguistiques et structuralistes ont pu développer une tentative d'objectivation de leur objet. Tout jugement en matière de littérature doit être rapporté à la société qui l'émet et à la connaissance du contexte de la production des objets d'étude. Cela n'implique pas qu'il n'y ait aucun « savoir » mais il faut admettre que les conditions de ce savoir

ne peuvent être établies quantitativement. Le terme « scientifique » ne doit pas alors, comme cela se fait facilement en français, être réservé aux sciences dites « dures ». Il faut l'entendre au sens allemand de *wissenschaft* comme objet de connaissance. Est scientifique ce qui peut être connu et su. Le problème est qu'un savoir de type interprétatif répond difficilement au critère établi par Thomas Kühn de « falsifiabilité ». Ce qui ne peut être démontré comme faux ne peut non plus être démontré vrai. Les sciences dites « dures » fondent largement leur effectivité sur l'exactitude ou l'établissement d'une mesure. Certaines sciences humaines, comme la sociologie, recourent aussi à la mesure mais, de manière plus générale, ces savoirs recourent plutôt à une méthode ou à des principes pour s'établir.

Le discours de la méthode de René Descartes, Les règles de la méthode sociologique d'Émile Durkheim ou l'exigence de neutralité axiologique chez Max Weber en sont des exemples illustres. Il n'est pas possible d'entrer ici dans le détail mais la question de l'intégrité scientifique peut être établie comme une éthique scientifique, comme le montre Pascal Engel dans *Les vices du savoir, Essai d'éthique intellectuelle*, Agone, 2019. Une « éthique » répond à un éthos, c'est-à-dire à un comportement au sein d'une société. Pour les Anciens, il ne peut être question d'une « morale personnelle ». Aristote l'explique nettement dans son *Éthique à Nicomaque* : le bien doit toujours être envisagé selon une perspective commune. Les sociétés anciennes définissent leur comportement éthique sous le signe de la réciprocité dans la communauté. De ce point de vue, l'invention de l'écriture, accompagnée de tous les outils de l'intellect tels que les a définis Jack Goody : la formule, le tableau, la liste, constituent indéniablement une rupture. L'écrit devra lui aussi définir et codifier ses usages. Le principe de non-contradiction, pierre de touche de la logique aristotélicienne, présuppose une dimension écrite. Il est fréquent que l'être humain dise une chose et son contraire à des moments différents. La présentation simultanée, sur un même plan, d'une formule ou d'un syllogisme rend inacceptable la contradiction en raison du tiers exclu. Il n'est plus alors nécessaire de recueillir un assentiment social. L'éthique du savoir peut être exposée et comprise. Mais il faut garder à l'esprit que la norme est toujours en partie autotélique. Dans la mesure où elle se définit par ses outils même. Comme le montre Jack Goody, un tableau à double entrée suscite naturellement, et par suite façonne le contenu des cases par le seul fait de les produire à la vue.

Littérature et Sciences Humaines n'échappent pas à cette détermination de la literacy. La « sécurisation » scientifique des publications et des travaux dans le contexte spécifique des lettres et des sciences humaines s'établit moins par la mesurabilité que par une suite de procédures qui en garantissent la qualité. L'herméneutique est aussi un savoir constitué comme l'exposait Georges Gadamer dans *Vérité et méthode*. On pourrait aussi citer *Les limites de l'interprétation* de Umberto Eco. Pour les études littéraires par exemple, le travail scientifique concerne l'histoire des textes et des œuvres mais pour une large part, il s'agit d'un travail d'interprétation. Dans le cas des études littéraires, la référence est d'une importance première. Elle indique où et quand un auteur a produit un texte spécifique. À la différence des historiens qui privilégient les sources premières, les littéraires travaillent la plupart du temps sur des textes déjà publiés. Cela signifie que le corpus se trouve déjà médiatisé par la culture.

La « fidélité » au texte original, l'Urtext, occupe une place particulière qui peut aller jusqu'à la sacralisation du texte et le refus de sa traduction. La littéralité déplace alors la fidélité de « la lettre » vers un « esprit » « qui vivifie » son sens, ouvrant la porte aux interprétations « charismatiques » plus ou moins garanties par les institutions. L'herméneutique des textes littéraires est en partie l'héritière des traditions interprétatives des textes sacrés. L'œuvre et son interprétation forment dès lors un tout problématique. Jorge Luis Borges pousse au maximum,

dans une de ses nouvelles, Pierre Ménard auteur du *Quichotte*<sup>1</sup>, l'hypothèse de deux textes « différents », car produits dans des conditions différentes, et de sens différents, mais dont la matérialité littérale serait strictement identique. Le « plagiat » est par conséquent une question des études littéraires qui interroge au plus profond la matérialité des œuvres et leur signification. Umberto Eco dans *Les limites de l'interprétation*<sup>2</sup> s'est interrogé sur les possibles effets de cette « dérive sémiotique ».

La production de textes par intelligence artificielle grâce à l'utilisation de robots conversationnels touche désormais la fonction même du langage humain dans sa capacité à rendre compte du monde environnant. C'est l'objet du dernier ouvrage de Daniel Andler (Gallimard, 2023) consacré à l'intelligence humaine et l'intelligence artificielle. Se trouve remise en question l'« intégrité » même de l'objet d'étude des Lettres.

---

<sup>1</sup> Wikipédia article Pierre Ménard, auteur du *Quichotte* : « Elle a été publiée initialement dans la revue argentine *Sur* en mai 1939, puis dans les recueils *El Jardín de senderos que se bifurcan* (1941) et *Ficciones* (1944). »

<sup>2</sup> Umberto Eco, *Les limites de l'interprétation*, Grasset, 1987.