

# **Pour un usage intègre et responsable de l'IA dans les études universitaires**

**Guide à l'intention des étudiantes et étudiants de la Sorbonne Nouvelle**

Version 1 (février 2025)

## Table des matières

<b>Première partie. L'IA dans le contexte universitaire .....</b>	<b>3</b>
Anatomie de l'IA .....	3
L'IA dans le processus de préparation d'un devoir.....	10
Risques liés à l'usage de l'IA .....	14
Pour aller plus loin.....	16
<b>Deuxième partie. Foire aux questions .....</b>	<b>18</b>
<b>Troisième partie. Check list pour un usage responsable de l'IA dans mes travaux universitaires.....</b>	<b>19</b>

# Première partie. L'IA dans le contexte universitaire

## Anatomie de l'IA

### 1° Une brève introduction à l'IA

Le terme “intelligence artificielle” n’est pas nouveau, ayant été forgé dès 1956 lors de la « conférence de Dartmouth » par John McCarthy, dans le prolongement des travaux mathématiques développés pendant la seconde guerre mondiale et immédiatement après, et plongeant ses racines dans des systèmes philosophiques esquissés depuis le Moyen Âge, qui tentent de formaliser le raisonnement humain. L’actuel usage courant désigne comme “Intelligence artificielle” les systèmes de génération automatique de texte, de son ou d’images, illustrés notamment par GPT et son robot conversationnel ChatGPT, lancé en octobre 2022 par la société OpenAI. Les performances de ce transformer<sup>1</sup>, tout comme son accessibilité, ont mis dans l’ombre le fait qu’il prend le relais de techniques plus anciennes, celles de l’apprentissage machine (*machine learning*), qui continuent à être utilisées dans de nombreux domaines.

Comme le *machine learning*, l’IA générative se base sur l’utilisation de jeux de données qui peuvent avoir été étiquetés par des humains, ou non. Dans le cas de l’utilisation de données étiquetées on parle d’apprentissage supervisé, qui permet à la machine d’inférer des informations à partir d’un grand nombre d’exemples déjà catégorisés. Dans les approches non supervisées, l’ordinateur est confronté à une grande quantité de textes, non-étiquetés, entre lesquels il lui est demandé d’identifier des similarités (en termes de vocabulaire, de structure, voire de graphie ou de disposition etc.). Enfin, une troisième technique utilisable en *machine learning* est l’apprentissage par renforcement : le système est récompensé pour les réponses correctes qu’il donne et pénalisé pour les mauvaises réponses. Ces récompenses et ces pénalités sont basées sur le calcul de la distance entre les réponses de la machine et les rétroactions des humains.

La réussite de GPT et des systèmes qui ont suivi se fonde sur trois facteurs. Le premier est l’utilisation d’énormes quantités de données (ex. tout le contenu de tous les sites web dans une langue donnée). Le deuxième est la mobilisation d’une architecture informatique complexe, appelée “réseau de neurones”, elle aussi développée depuis les années 1950, mais ayant connu des avancées significatives à partir des années 1990. Le produit de cette combinaison de réseaux de neurones et de grandes quantités de texte est désigné comme un “grand modèle de langue” (*Large Language Model*, LLM\*), soit une représentation statistique de la distribution des éléments (phonèmes, syllabes, mots, syntagmes, phrases...) d’une langue naturelle. Enfin,

---

<sup>1</sup> Un Transformer est un type d'architecture informatique qui a révolutionné le traitement du langage naturel.

la puissance des LLMs est basée sur leur ajustement en fonction d'une très grande quantité de rétroactions humaines.

Il est important de souligner que l'IA ne sait pas, à proprement parler, raisonner ; elle se fonde sur des calculs statistiques effectués à très grande vitesse pour livrer une suite de mots qui a le plus de chances de se trouver dans le même "espace de discours" que la question posée. Par ailleurs, en dépit de leur capacité à reproduire les codes de la conversation humaine, les systèmes d'IA générative ne sont dotés ni de compréhension, ni d'empathie et encore moins d'intentionnalité. Représentables comme des "perroquets stochastiques", ils produisent toujours des réponses : une IA ne sera jamais silencieuse et ne dira jamais "je ne sais pas". Par ailleurs, elle ne donnera jamais des réponses fondées sur une connaissance (même rudimentaire) de la réalité : celle-ci reste, pour l'IA, intégralement médiée par les données ayant servi à son entraînement. Enfin, elle ne peut pas être considérée *stricto sensu* comme créative : s'il ne fait aucun doute qu'elle peut proposer des poèmes inédits, des scénarios nouveaux, des images qui n'ont pas été captées ou peintes par la main de l'homme, les IA génératives ne produisent pas de tels objets de leur propre chef, étant toujours saisies et guidées par un humain.

Les caractéristiques des LLMs en font des agents mobilisables pour une grande variété des tâches : réponses à des questions, résumés, génération de texte ou de code informatique, correction d'erreurs etc. Leur caractère généraliste signifie également que leurs performances sont moins bonnes pour des tâches hautement spécialisées. L'apport des LLMs dans des domaines de pointe (lecture de radiographies, optimisation d'itinéraires de transport, aide à l'automatisation des chaînes de production, apprentissages personnalisés etc.) se base toujours sur des ré-entraînements spécifiques (*fine-tuning*) et sur une coopération avec des opérateurs humains. La quantité et la propreté des données utilisées pour ce *fine-tuning* sont, de nouveau, cruciales pour la réussite des modèles.

## 2° Glossaire

### **Algorithme**

Un algorithme est une série d'instructions logiques et ordonnées pour accomplir une tâche ou résoudre un problème. Des données en entrée (images, textes, sons, code informatique, vidéos, etc.) sont traitées selon une série d'instructions, ce qui produit des données en sortie (images, textes, sons, code informatique, vidéos etc.).

Exemples d'algorithmes : trier une liste de nom par ordre alphabétique, calculer une moyenne de notes dans une classe.

### **Apprentissage automatique**

L'apprentissage automatique (*machine learning*) est une branche de l'intelligence artificielle qui développe des machines capables d'exécuter des tâches sans avoir été spécifiquement programmées pour celles-ci, grâce à un entraînement sur des données. À partir d'une série d'exemples (« le 13 mars », « avril 2022 », « vendredi passé » etc.), une machine peut inférer qu'il s'agit de dates et reconnaître comme appartenant à la même catégorie une suite de mots qu'il n'aura jamais rencontrée pendant la phase d'apprentissage (ex. « novembre de l'année précédente »).

Il existe différents types d'entraînement : supervisé (la machine apprend à partir de données étiquetées), non supervisé (la machine apprend sans données étiquetées en faisant des regroupements), par renforcement (la machine apprend en recevant des récompenses ou des pénalités)

### **Apprentissage profond**

L'apprentissage profond est une sous-branche de l'apprentissage automatique (*machine learning*) où on utilise des réseaux de neurones artificiels, composés de plusieurs couches, pour permettre à l'intelligence artificielle de construire des représentations complexes à partir de grandes quantités de données et des algorithmes d'entraînement et d'optimisation. L'apprentissage profond (*deep learning*) peut traiter des tâches complexes comme la reconnaissance faciale ou la traduction automatique.

La plupart des IAG récentes comme ChatGPT reposent sur l'apprentissage profond.

### **Chatbot (agent conversationnel)**

Un chatbot est un outil permettant de simuler des conversations avec des humains. Les chatbots peuvent utiliser différentes technologies, allant de simples scripts de programmation à des algorithmes de *machine learning*, de *deep learning* ou des modèles de langues.

### **Deep learning**

Voir « Apprentissage profond »

### **Fine-tuning**

Le *fine-tuning* est une opération qui consiste à adapter un LLM à des tâches spécifiques, en l'entraînant de nouveau sur des données propres à l'utilisateur.

### **Hallucinations**

Les hallucinations dans le contexte des modèles de langage et des IA désignent la génération d'affirmations qui ne sont pas fondées sur la réalité. Cela peut inclure la fabrication de faits, de chiffres, de citations ou de réponses qui semblent plausibles mais qui sont complètement incorrectes. Ce phénomène est particulièrement courant dans les modèles génératifs comme les LLMs (*Large Language Models*), qui peuvent parfois inventer des réponses "convaincantes" mais fausses.

### **Intelligence artificielle (IA)**

D'après le dictionnaire Larousse, l'intelligence artificielle est « [l'ensemble des théories et des techniques mises en œuvre pour réaliser des machines dont le fonctionnement s'apparente à celui du cerveau humain » (*Dictionnaire Larousse*, n.d.).

### **Intelligence artificielle générative (IAG)**

Une intelligence artificielle générative est une intelligence artificielle qui génère un contenu (texte, image, vidéo, musique, code informatique, etc.) à partir de données d'entraînement et de contenus donnés par l'utilisateur.

On parle parfois d'IAG *text-to-text*, *image-to-text*, *text-to-video* : le premier terme désigne le contenu fourni par l'utilisateur, tandis que le second terme désigne le contenu généré par l'IAG.

## Exemples

- Vous rédigez un prompt et l'intelligence artificielle génère du texte en réponse, il s'agit d'une IAG *text-to-text*.
- Vous fournissez une image et l'IAG vous génère du texte en réponse, il s'agit d'une IAG *image-to-text*.

## **Intelligence artificielle générative multimodale (IAG multimodale)**

Il s'agit d'une intelligence artificielle générative qui est capable de prendre plusieurs types de contenu en entrée et qui génère plusieurs types de contenus en sortie.

Exemple d'intelligence artificielle multimodale : chatGPT 4o car peut interpréter des textes comme des images.

## ***Large Language Models (LLM)***

Un Large Language Model (LLM) est un modèle de traitement du langage naturel entraîné sur une énorme quantité de texte provenant de sources diverses. Ces modèles, comme GPT (*Generative Pretrained Transformer*), sont capables de comprendre, générer et manipuler du langage humain à un niveau très avancé. Ils sont utilisés pour des tâches comme la traduction automatique, la génération de texte, la réponse à des questions, et la classification de texte.

## **Modèle de fondation**

On appelle modèle de fondation un modèle pré-entraîné sur un vaste corpus de données (textes, images, vidéos, etc.). Ce modèle généraliste et non spécifique peut être adapté en différents scénarios. Pour cela il faut entraîner de nouveau le modèle de fondation, lors d'une opération appelée le *fine-tuning* (nouvel entraînement), pour qu'il apprenne à traiter des tâches spécifiques, pour en faire un modèle spécialisé, performant sur une tâche spécifique.

Par exemple, on peut prendre un modèle de fondation GPT-4 pour en faire un modèle spécialisé capable de reconnaître dans un texte les différentes figures de style.

## ***Machine learning***

Voir « Apprentissage automatique »

## **Paramètres**

Les paramètres sont des variables ajustables d'un modèle LLM, qui sont modifiés pendant l'entraînement, afin d'optimiser les performances du modèle sur une tâche donnée. Les deux types de paramètres les plus importants dans une architecture Transformer sont le « poids » et le « biais ». Plus l'architecture du réseau de neurones est complexe, plus le nombre de paramètres ajustables sera important. À titre d'exemple, ChatGPT se caractérise par plusieurs trillions de paramètres.

## **Prompt**

Le prompt est le message envoyé par l'utilisateur à un modèle d'intelligence artificielle dans le but d'obtenir une réponse.

## **Réseau de neurones (*Neural Network*)**

Un réseau de neurones est une architecture informatique inspirée du cerveau humain. Il est constitué de nœuds (ou neurones) interconnectés organisés en couches. Un « neurone », dans ce contexte, est une unité de calcul, destinée à traiter certains signaux en entrée selon un protocole défini. Les résultats d'un neurone peuvent être envoyés aux autres neurones, qui les traitent à leur tour. Le neurone ajuste également ses connexions en fonction de l'apprentissage.

### **Transformer**

Un Transformer est un type de modèle d'architecture de réseau de neurones qui a révolutionné le traitement du langage naturel. Introduit en 2017 dans l'article "Attention is All You Need" par Vaswani et al., il se distingue par sa capacité à traiter des séquences de données (comme du texte) en parallèle, grâce à un mécanisme d'attention. Ce mécanisme permet au modèle de se concentrer sur des parties spécifiques de l'entrée, facilitant ainsi l'apprentissage sur de grandes quantités de données et la modélisation des relations à longue portée. Les modèles comme GPT et BERT sont basés sur cette architecture.

### **3° Exemples**

GPT, le modèle à la base de ChatGPT, a été suivi par la mise sur le marché d'une multitude de LLMs. Certains sont fermés : on ne connaît pas les données sur lesquelles ils ont été entraînés et, plus gênant encore, les paramètres qui ont été utilisés pour leur entraînement (c'est-à-dire, les poids et les biais des différents neurones participant à l'apprentissage). C'est le cas des modèles GPT-1, 2, 3, 4 avec leurs variantes, ainsi que de PaLM ou Gemini. D'autres sont ouverts, accessibles via des plateformes bien connues des informaticiens (Hugging Face, par exemple), et même téléchargeables sur une machine personnelle : Llama, Bloom, certaines versions de Claude etc. Attention : l'ouverture ou la fermeture d'un LLM n'est pas à confondre avec le fait que vous puissiez utiliser des chatbots associés de façon gratuite ou non. Cette différence est importante, parce qu'elle est également significative de ce qui se passera avec les conversations et les informations que vous aurez avec les robots vous permettant d'interagir avec ces modèles : dans le cas des LLMs fermés, vous risquez de perdre la propriété de vos données. Dans le cas des LLMs ouverts, vous pouvez garder la maîtrise, surtout si vous les faites fonctionner sur votre propre machine, ou dans un environnement sécurisé (via un robot et un serveur de l'université, par exemple). Toutefois, dans ces deux derniers cas vous aurez besoin de certaines compétences informatiques pour faire fonctionner les modèles.

Voici une liste des IAG les plus connues existant à la date de rédaction de ce guide :

<b>Modèle</b>	<b>Version</b>	<b>Organisation</b>	<b>Fonctionnalités principales</b>
GPT	GPT 4o	OpenAI	Capacités multimodales étendues, interactions vocales en temps réel, traduction, analyse de contenus téléchargés (photos, documents)
GPT	GPT 4o mini	OpenAI	Modèle plus léger et moins gourmand en ressources que GPT 4o

Claude	Claude 3.5 sonnet	Anthropic	Analyse de textes, d'images, de photos de documents, création de code informatique, compétences linguistiques, traduction, création de contenu textuel.
Dall·e	Dall·e 3	OpenAI	Génération d'image
Gemini	Gemini 1.5 Flash	Google	Compréhension du langage, traduction, génération de texte, analyse d'images, création de code informatique

Voici également quelques exemples d'applications utiles, qui recourent à de l'IAG.

Application	Organisation	Rôle principal	Fonctionnalités
Perplexity	Perplexity AI	Moteur de recherche conversationnel	Moteur de recherche conversationnel, génération de résumés, transparence des sources, mémoire contextuelle, questions liées, bibliothèque de discussion, page découvrir, propose des images et des vidéos
NotebookLM	Google	Analyse de ressources basée sur le LLM Gemini 2.0	Importation de contenu (PDF, Word, liens, YouTube, image etc.) Analyse de contenu Génération de contenu Interaction en langage naturel Génération d'un podcast audio en anglais basé sur le contenu apporté

Dans le contexte des études universitaires, ce sont probablement les assistants de recherche académique qui peuvent également vous intéresser :

Outil	Fonction principale	Origine des articles	Organisation
Semantic Scholar	Recherche d'articles académiques	Plus de 220 000 000 de publications scientifiques Archives ouvertes dont HAL Éditeurs académiques Sociétés savantes Presses universitaires	Allen Institute for AI
Consensus	Réponse à des questions basées sur des articles scientifiques	Semantic Scholar Open Research Corpus	Consensus inc.



Research rabbit	Élabore une carte de connexions interactives entre les articles pour suggérer d'autres articles à étudier	Semantic Scholar Open Research Corpus	Research Rabbit inc.
Connected Papers	Créer un graphe pour présenter les relations pertinentes entre les articles scientifiques	Semantic Scholar Open Research Corpus	Connected Papers inc.
Litmaps	Créer un graphe présentant des connexions entre les articles, à partir des citations entre articles	Semantic Scholar Open Research Corpus	Litmaps
Elicit	Assistant de recherche permettant d'aider à faire une revue de littérature	Semantic Scholar Open Research Corpus	Ought inc

Enfin, notez qu'il existe des sites qui vous permettent soit d'utiliser plusieurs LLMs à la fois (<https://poe.com>), soit d'identifier l'application la plus appropriée à votre objectif (<https://theresanaiforthat.com/>).

## **L'IA dans le processus de préparation d'un devoir**

La formation universitaire consiste à transmettre un certain nombre de compétences qui ont trait à l'exercice et à l'expression de la pensée. L'IA ne saurait se substituer à l'expression d'une pensée humaine autonome, construite et argumentée. Néanmoins, l'IA possède des fonctionnalités qui peuvent être sollicitées pour effectuer certaines tâches, notamment dans le cadre d'un processus de rédaction.

Tout processus de rédaction peut être schématisé en trois grandes étapes : la planification, la mise en texte et la révision. L'IA peut intervenir à chacune de ces étapes.

### **1° Planifier**

#### a) Cerner le sujet

On parle également d'idéation. L'IA peut être une aide pour dépasser le blocage de la page blanche. Elle peut également aider à organiser des idées ou à les prolonger vers des concepts annexes dont l'exploration peut être utile, par exemple en faisant une recherche lexicale pour mieux comprendre votre sujet.

#### b) Rechercher des informations

- Équations de recherche

L'intelligence artificielle générative peut vous aider à formuler des équations de recherche, que vous pourrez ensuite utiliser dans des bases de données académiques, des catalogues de bibliothèque ou des moteurs de recherche.

Exemple de prompt pour obtenir une équation de recherche :

« Tu es un expert en recherche académique et en élaboration d'équations de recherche. Je fais une recherche académique sur [...] Peux-tu rédiger une équation de recherche adaptée qui inclut des synonymes ou termes alternatifs pour ces concepts ? »

- Recommandations intelligentes

Vous pouvez vous servir des outils d'intelligence artificielle spécialisés dans la recherche académique comme Semantic Scholar pour obtenir des recommandations intelligentes basées sur des documents que vous avez sauvegardés dans des fils de recherche intelligents. L'intelligence artificielle vous suggèrera alors quotidiennement de nouvelles recommandations, à l'intérieur de ces fils de recherche intelligente.

D'autres outils d'intelligence artificielle spécialisée dans la recherche académique comme Research Rabbit peuvent vous recommander des articles à lire sous forme de graphe intelligent. Ces graphes mettent en évidence des connexions avec des articles que vous avez sauvegardés dans des collections d'articles scientifiques.

- Questions de recherche ciblées

L'IA peut répondre à des questions, plus ou moins focalisées, de façon plus ou moins correcte. Il convient de se rappeler que la plupart des modèles de langage n'actualisent pas leurs informations et qu'ils n'ont pas été entraînés sur de la littérature scientifique (validée par les pairs, certifiée) ou, du moins, pas principalement.

c) Synthétiser des connaissances

Cet aspect est assez difficile à distinguer des deux précédents dans les productions de l'IA générative. Tout en esquissant des pistes ou en fournissant des informations, l'IAG propose des synthèses. Il convient de retenir que ces synthèses ne se fondent pas sur une connaissance du domaine concerné, mais de l'espace de discours présent dans les données d'entraînement de l'IAG. Pour les connaissances courantes, le produit de l'IAG va être probablement assez correct, tout en introduisant des erreurs et des biais ; pour des connaissances plus spécialisées, la synthèse proposée va rester probablement très générale, peu utile en fin de compte. Sur ce point, la plus grande précaution reste de mise lorsque vous utilisez une IAG.

## **2° Mettre en texte**

a) Rédiger

À partir d'un sujet donné, l'IAG peut rédiger un texte d'une longueur spécifiée, qui sera organisé selon les règles (introduction, développements, conclusions) et qui proposera une synthèse des connaissances sur le sujet. Cet usage est toutefois celui qui pose les plus grands problèmes en contexte universitaire : si vous donnez le produit de l'IA à votre enseignante ou enseignant au lieu de votre propre texte, vous commettez une fraude. Il est également probable que le texte produit par l'IA reste très passe-partout et ne réponde finalement pas aux attendus de l'enseignant ou de l'enseignante, même s'il a l'air correct et même élégant. Vous vous privez également, dans le processus, de l'apprentissage de l'écriture académique, qui constitue une compétence dont vous aurez besoin tout au long de votre vie professionnelle.

Dans certains cas, le texte de l'IAG est utilisé comme une base de départ, à partir de laquelle les étudiants mobilisent leurs connaissances, ajoutent des exemples, corrigent des affirmations fausses ou nuancent d'autres. Cet usage semble plus licite, surtout quand le texte de base de l'IAG a été beaucoup retravaillé. Il n'en reste pas moins que vous risquez de produire un devoir qui ne répond pas tout à fait à ce qui est attendu, car l'IAG n'a pas été entraînée sur les matériaux de votre cours. Il faudra également être très attentif/ attentive à signaler systématiquement, au moyen de barres obliques, toutes les parties de texte que vous n'aurez pas écrites par vous-même.

En bref, l'utilisation de l'IAG comme base de départ rédactionnelle reste peu recommandée.

b) Reformuler

L'IAG peut reprendre vos phrases ou vos paragraphes et les réécrire, y compris selon une consigne de style donnée. Cet usage semble plus acceptable que la rédaction pure et simple d'un devoir ou d'un mémoire par l'IAG. Dans ce cas aussi, vous vous exposez

toutefois à différents risques : modification du sens de la phrase ; perte de cohérence de l'ensemble ; dépersonnalisation. Tout comme lors de l'utilisation de l'IAG pour rédiger à votre place, vous n'apprenez pas à améliorer vous-même vos productions. La reformulation par l'IAG reste donc à utiliser avec parcimonie, seulement quand vous n'arrivez pas à faire mieux, ou dans le cas d'un objectif précis (ex. passer d'un texte complexe à un texte facile à lire, pour une population spécifique).

### **3° Réviser**

#### a) Corriger l'orthographe

L'IAG peut vous aider à corriger vos fautes, surtout dans ses versions spécialisées (Quilbot, Scribens, Grammarly...) mais ce type d'aide est fournie par n'importe quel éditeur de texte (Word, Pages, LibreOffice, OpenOffice, Google Doc etc.). Il semble peu approprié de recourir à l'IA pour une telle tâche, compte tenu de la consommation électrique qu'elle suppose, largement supérieure à un éditeur de texte.

Alternativement, il est possible de demander à l'intelligence artificielle générative de relever vos fautes d'orthographe, de grammaire et de ponctuation dans un tableau sans reformuler votre texte et de mettre en gras toutes vos fautes. De cette manière, vous pourrez corriger vos fautes d'orthographe, sans que l'intelligence artificielle reformule votre devoir.

#### b) Revoir la syntaxe

La révision de la syntaxe, effectuée par les mêmes outils énumérés supra, est également une tâche sur laquelle l'IAG est efficace, mais aussi concurrencée par d'autres solutions numériques, moins gourmandes en énergie.

Vous pouvez également demander à l'intelligence artificielle de mettre en évidence les fautes de syntaxe dans votre rédaction, sans reformuler votre texte. Ce sera à vous d'apporter les corrections.

#### c) Résumer

L'IAG peut aider à retenir l'essentiel d'un texte en un paragraphe ou un autre texte d'une longueur donnée. Toutefois, des études existent qui montrent des différences importantes entre le résumé fait par un humain et par l'IAG, cette dernière réalisant plutôt une opération de raccourcissement que de résumé.

#### d) Transformer un devoir en une présentation autre

L'IAG peut vous aider à gagner du temps pour réaliser un diaporama à partir d'un texte écrit, ou pour transcrire une présentation orale, ou pour d'autres types de conversion d'un format à un autre. Il vous appartient toutefois de vérifier la correction de la conversion : est-ce que dans l'opération des éléments importants n'ont pas été omis, décontextualisés ou modifiés ? Comme dans toutes les tâches décrites plus haut, savoir faire par soi-même ce qu'on demande à l'IA est la meilleure garantie d'obtenir un résultat de qualité, grâce au recul critique et au travail supplémentaire que vous investirez dans l'amélioration du produit de l'IA.

e) Mettre en forme des références bibliographiques

L'intelligence artificielle générative peut vous aider à mettre en forme des références bibliographiques dans un style bibliographique spécifique (APA, MLA, Chicago, Harvard, etc.).

## Risques liés à l'usage de l'IA

Comme tous les outils, l'IA demande de prendre un certain nombre de précautions et d'être conscient de ce que son utilisation implique.

### Aggravation de la dette écologique

Toute IA est énergivore. Des travaux sont en cours pour améliorer la frugalité de l'outil, mais il est nécessaire d'être conscient de sa consommation énergétique. À titre d'exemple, l'IA requiert dix fois plus d'énergie que Google pour répondre à une question du type "quelle est la date de la Révolution française ?". À noter également que les encyclopédies en ligne sont plus résilientes face aux erreurs et aux fausses informations introduites volontairement : l'intervention d'un grand nombre de rédacteurs permet la correction de telles erreurs et fausses informations dans des délais parfois inférieurs à une minute. Les "hallucinations" de l'IA sont, elles, bien plus persistantes.

### Non-confidentialité

Vos conversations avec un outil IA ne sont pas protégées par des règles de confidentialité. Ne communiquez pas des données intimes ou confidentielles (par exemple, recueillies sur le terrain) et ne tombez pas dans le piège de croire qu'il y a de l'empathie dans l'outil. La machine n'a pas de compréhension psychologique car ce n'est pas un humain à l'écoute. Comme on l'a vu plus haut, elle fonctionne à partir de probabilités statistiques de rencontre de lettres et de mots les uns après les autres.

### Fausseté ou inexactitude des réponses

Les réponses générées par l'IA ne sont pas vraies en soi. Ce sont des suites de mots et de phrases qui ont une grande probabilité de se trouver dans le même "espace linguistique" que la question que vous avez posée. Leur degré de vérité ne peut être attesté que par quelqu'un qui a une connaissance préalable du domaine : votre enseignante ou enseignant, ou bien l'auteur d'un livre ou d'un article qui a été évalué par les pairs (donc estampillé "vrai" selon les connaissances les plus avancées à un moment donné).

L'utilisation d'une IA sans recul et analyse critique vous expose à prendre pour vraies des "hallucinations". Elles sont plus nombreuses que vous ne le pensez : les IA, par exemple, ne lisent pas des romans pour les résumer, ni ne visionnent des films ; elles recombinaient des discours en lien avec les titres et quand ces éléments manquent, elles vont inventer à partir de résumés d'œuvres qui n'ont rien à voir avec celles qui vous intéressent. De ce fait, la fraude est facilement détectable par votre regard critique ou par votre enseignant ou enseignante.

## Fausseté ou inexactitude des sources mentionnées

Comme pour les résumés ou les synthèses, les sources citées par l'IA peuvent être inexistantes, inappropriées ou mal citées. L'IA peut tout à fait attribuer à un auteur des articles qu'il n'a jamais écrits ou des idées qu'il n'a jamais énoncées. Même quand ces idées ou articles sont exacts, l'IA les extrait de leur contexte : la citation peut donc ne pas correspondre à l'intention de l'auteur ; elle s'inscrit dans un agencement statistique de mots qui peut en modifier complètement le sens. À l'inverse, l'IA peut reprendre des idées et des phrases sans les signaler comme des citations : elle se rend donc coupable de plagiat. Cependant, l'IA n'est pas une personne qui puisse être poursuivie : c'est vous qui devenez responsable si vous reprenez ses propos sans vérification.

N'oubliez pas non plus que l'IA présente des biais liés à la nature des données sur laquelle elle a été entraînée. L'anglais étant la langue dominante sur Internet, c'est également la vision anglo-saxonne du monde qui domine dans les productions de l'IA. Des biais de genre ou raciaux ont également été constatés dans les productions de l'IA. Tout ceci invite donc à la plus grande prudence lors de l'usage de cet outil.

**En bref : un usage de l'IA non éclairé et non critique peut avoir de sérieuses conséquences sur vos apprentissages. Un travail au long cours, basé sur de multiples lectures et la réalisation de nombreux exercices, reste la meilleure façon de produire de bons devoirs et donc de réussir vos études.**

## Pour aller plus loin

### Bibliographie

[collectif] *L'IA peut-elle nous remplacer ?*, Montrouge, Bayard, 2024.

Bersini Hugues, Hasselman Ken, *L'Intelligence artificielle en pratique avec Python : recherche, optimisation, apprentissage*, Paris Éditions Eyrolles, 2023 (seconde édition).

Le Cun Yann, *Quand la machine apprend*, Paris, Éditions Odile Jacob, 2023.

Mermillod, Martial, *De l'intelligence humaine à l'intelligence artificielle*, Louvain la Neuve, De Boeck supérieur, 2024.

Peters M., « Note éditoriale : Intelligence artificielle et intégrité académique peuvent-elles faire bon ménage ? », *Revue des sciences de l'éducation*, 49(1), 2023, [<https://doi.org/10.7202/1107846ar>]

### MOOCs

- [INRIA] *C'est quoi l'IA ?*

[<https://epoc.inria.fr/epocs/E009LL/>]

- [FUN] *L'Intelligence artificielle avec intelligence*

[<https://www.fun-mooc.fr/fr/cours/lintelligence-artificielle-avec-intelligence/>]

- Radio France. (n.d.) *Série : L'intelligence artificielle, utopie dystopique* [Podcast]. France Culture.

[<https://www.radiofrance.fr/franceculture/podcasts/serie-l-intelligence-artificielle-utopie-dystopique>]

- L'Université numérique, *L'IA dans l'enseignement supérieur*

[<https://moodle.luniversitenumérique.fr/course/view.php?id=1083>]

### Sites utiles

- [ChatGPT] *Student's Writing Guide*

[<https://openai.com/chatgpt/use-cases/student-writing-guide/>]

- Consensus

[<https://chatgpt.com/g/g-bo0FiWLY7-consensus>]

- CulturIA

[<https://cultureia.hypotheses.org/>]

- [MESR], *L'IA : de quoi parle-t-on ?*

[<https://www.enseignementsup-recherche.gouv.fr/fr/intelligence-artificielle-de-quoi-parle-t-91190>]

- [Savoir Devenir] *Glossaire EMI. Data, Algorithmes, Intelligence artificielle*

[<https://savoirdevenir.net/2024/11/glossaire-emi-data-algorithmes-intelligence-artificielle/>]

- Scholar AI

[<https://chatgpt.com/g/g-L2HknCZTC-scholar-ai>]

- ScholarGPT

[<https://chatgpt.com/g/g-kZ0eYXlJe-scholar-gpt>]

- SciSpace

[<https://chatgpt.com/g/g-NgAcklHd8-scispace>]

- Perret Arthur, *Guide de l'étudiant pour ne pas écrire avec ChatGPT*



[\[https://www.arthurperret.fr/blog/2024-11-15-guide-etudiant-ne-pas-ecrire-avec-chatgpt.html\]](https://www.arthurperret.fr/blog/2024-11-15-guide-etudiant-ne-pas-ecrire-avec-chatgpt.html)

- Ligozat, A.-L. et De Vries, A., *IA générative : la consommation énergétique explose.*

Polytechnique Insights

[\[https://www.polytechnique-insights.com/tribunes/energie/ia-generative-la-consommation-energetique-explose/\]](https://www.polytechnique-insights.com/tribunes/energie/ia-generative-la-consommation-energetique-explose/)

## Deuxième partie. Foire aux questions

Q: Puis-je utiliser l'IA pour mes devoirs ?

R: Cela dépend. L'usage de l'IA n'est pas interdit par principe à la Sorbonne Nouvelle. Toutefois, il existe des enseignements dont les objectifs pédagogiques ne pourraient être atteints si l'IA était utilisée. Pour une langue étrangère, par exemple, l'IA aura sans doute de meilleures performances que vous : il peut être tentant de l'utiliser pour traduire un texte difficile à comprendre. Cependant, ce faisant, ce n'est pas vous qui consultez les dictionnaires, analysez la syntaxe, comprenez la construction du texte ; ce n'est donc pas vous qui accumulez des connaissances ou améliorez vos compétences. C'est la raison pour laquelle l'usage de l'IA peut être parfois interdit, ou du moins fortement déconseillé par votre enseignant ou enseignante.

Q: Comment puis-je connaître les devoirs pour lesquels l'usage de l'IA est permis ?

R: Reportez-vous au descriptif de chaque enseignement. Il devrait comporter une "Fiche d'information sur l'usage de l'IA" qui vous éclairera sur ce qui est permis ou interdit. En absence d'un tel descriptif, considérez que l'usage de l'IA **n'est pas permis par votre enseignant**. D'une manière générale, par défaut, il est interdit de présenter comme sien un devoir qui aurait été intégralement réalisé avec l'IA ou qui aurait donné lieu à quelques modifications de votre part (comme un changement de mots par ci, par-là, la réorganisation de phrases, l'ajout de courtes portions de texte etc.). Le travail doit rester, dans sa plus grande partie, le vôtre.

Q: Comment dois-je indiquer que j'ai eu recours à l'IA ?

R: La "Fiche d'information sur l'usage de l'IA", remise par votre enseignant ou enseignante, donne des consignes de référencement et de citation.

Q: Est-ce que je serai noté plus ou moins bien si j'utilise l'IA pour mes devoirs ?

R: Cela dépend de l'étendue de l'usage et de la qualité du résultat. En fonction des disciplines et des exercices, certains usages peuvent être valorisés. Une rédaction correcte effectuée par vos propres forces peut valoir plus de points qu'une rédaction corrigée avec l'aide de l'IA. À l'inverse, une recherche d'informations qui aura intelligemment utilisé l'IA peut vous apporter un bonus. Dans tous les cas, demandez à vos enseignantes ou enseignants leurs critères de notation.

## **Troisième partie. Check list pour un usage responsable de l'IA dans mes travaux universitaires**

1. J'ai vérifié que l'usage de l'IA est permis pour le travail que je suis en train de faire.  
OUI/ NON
2. J'ai bien compris quelles sont les activités que je peux accomplir avec l'IA et quels sont les interdits.  
OUI/ NON
3. J'ai analysé de façon critique les résultats, en croisant les réponses avec celles que j'obtiens de sources traditionnelles (livres, articles, notes de cours).  
OUI/ NON
4. J'ai vérifié toutes les sources citées par l'IA (la responsabilité des citations et de la bibliographie me revient).  
OUI/ NON
5. Je n'ai pas dévoilé des informations intimes, confidentielles ou de terrain dans mes échanges avec l'IA.  
OUI, j'ai fait attention./ NON, j'ai communiqué des informations intimes, confidentielles ou de terrain dans mes échanges avec l'IA.
6. J'ai indiqué tout ce qui vient de l'IA dans mes rédactions, et j'ai documenté mes usages dans la section appropriée de la bibliographie. De la même manière, j'ai indiqué explicitement tout ce qui vient de la bibliographie parcourue ou d'autres sources.  
OUI/ NON
7. Je n'ai pas demandé à un tiers de produire le travail, avec ou sans IA.  
OUI, j'ai fait mon travail moi-même./ NON, je me suis fait aider.

Si vous n'êtes pas en mesure d'assumer une de ces affirmations (réponse « non »), il convient de reprendre votre travail et de corriger ce qui doit l'être jusqu'à pouvoir le communiquer en confiance à votre enseignante ou enseignant.