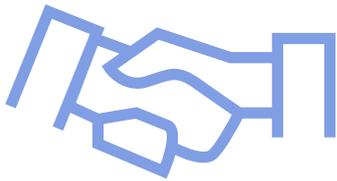


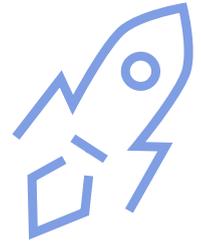
NUMÉRIQUE
DANS LES GRANDES ÉCOLES,
UNE APPROCHE À 360°

SOMMAIRE

INSPIRER
et renouveler
les pédagogies



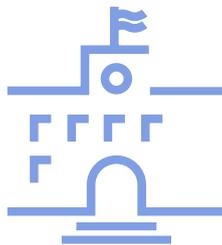
LABELLISER
et promouvoir



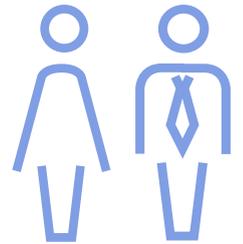
PILOTER



RATIONALISER
et sécuriser les
environnements
techniques



ADAPTER
les locaux et
les process



ACCOMPAGNER
les acteurs

HORS SÉRIE

Intelligence Artificielle
*Pour un usage raisonné et critique de
l'IA dans les Grandes écoles*

SOMMAIRE

Intelligence artificielle : pour un usage raisonné et critique de l'IA dans les Grandes écoles

INTRODUCTION	p.1
1. INTELLIGENCE(S) ARTIFICIELLE(S) : DE QUOI PARLE-T-ON ?	p.2
2. LES SPÉCIFICITÉS DES IA DITES GÉNÉRATIVES	p.4
3. LES BIAIS	p.5
3.1. Les biais amont, liés à la conception des LLMs	
3.2. Les biais aval, liés à l'utilisation des résultats des LLMs	
4. LES COÛTS	p.7
4.1. Les coûts environnementaux	
4.2. L'IA peut-elle être écologiquement frugale ?	
4.3. Les coûts humains	
5. LE SUJET ÉPINEUX DES ÉVALUATIONS	p.9
5.1. Les LLMs, des outils puissants, mais pas omnipotents	
5.2. Détecter l'usage de l'IA générative dans les travaux étudiants ?	
5.3. Le besoin de revoir les modalités d'évaluation	
6. LES APPORTS POSSIBLES DE L'IA : EXEMPLES ET ENRICHISSEMENT DES FONCTIONNALITÉS	p.11
6.1. Différents types d'usages de l'IA	
6.2. RAGs et Fine-Tuning : deux moyens complémentaires pour enrichir l'IA générative	
7. COMMENT SE LANCER ?	p.17
7.1. « Unlearn & Relearn » : le nouveau monde de l'IA générative	
7.2. Une vision collaborative à 360° de l'IA générative selon un principe test & learn	
7.3. Accompagner et former le professionnel et le citoyen	
7.4. L'IA générative : des conseils pour se projeter dans un futur mouvant	
CONCLUSION	p.21

INTRODUCTION

Symbolisées par des outils comme ChatGPT, Midjourney, le chat de MISTRAL AI, Deepseek ou encore Dall-E, **les technologies d'intelligence artificielle (IA) générative sont en train de transformer des activités** aussi diverses que la création de contenus (écrit, audio, vidéo, graphique, code, etc.), la conception de produits, ou encore l'éducation pour n'en citer que quelques-unes.

S'appuyant sur des algorithmes de traitement automatique du langage et des réseaux de neurones pour produire automatiquement des contenus, **le potentiel des IA génératives semble considérable**, alors même qu'elles ne furent révélées au plus grand nombre que le 30 novembre 2022 lorsqu'OpenAI ouvrit à tous l'accès à ChatGPT 3.0.

Alors que de nombreux outils d'IA sont utilisés quotidiennement depuis de nombreuses années, **l'émergence de modèles conversationnels a permis au grand public de découvrir le potentiel des outils d'IA dites génératives**. Alors derrière les effets d'annonce, nous nous proposons de faire un petit rappel de ce qu'est, ou plutôt **ce que sont les IA telles que nous les connaissons actuellement**.

Nous ne rentrerons pas dans les débats philosophiques (une intelligence peut-elle être artificielle ?) mais **nous nous concentrerons sur l'apport d'éléments concrets permettant à n'importe quel acteur de l'ESR de se faire une idée des mécanismes sous-jacents et ainsi mieux comprendre les possibilités, mais aussi les limites actuelles d'outils basés sur ces algorithmes**.

Ainsi, **ce chapitre traitera des spécificités des IA génératives ainsi que des biais qui leur sont associés**. Nous en évoquerons ensuite les coûts environnementaux et humains (passant volontairement sous silence leur coût financier, variable selon les usages et les prestataires). Puis nous aborderons la question spécifique (et épineuse s'il en est) des évaluations dans un monde post-IA générative. Nous concluons ce chapitre en donnant des exemples concrets de l'IA dans les différentes activités de nos établissements, et des conseils pour se lancer avec ces outils.

1. INTELLIGENCE(S) ARTIFICIELLE(S) : DE QUOI PARLE-T-ON ?

Comme nous l'avons souligné en introduction, il n'existe pas une mais plusieurs IA avec différents courants de pensée, algorithmes et applications.

Pour simplifier, nous pouvons séparer ces techniques en 2 grandes familles : l'IA symbolique et l'IA numérique.

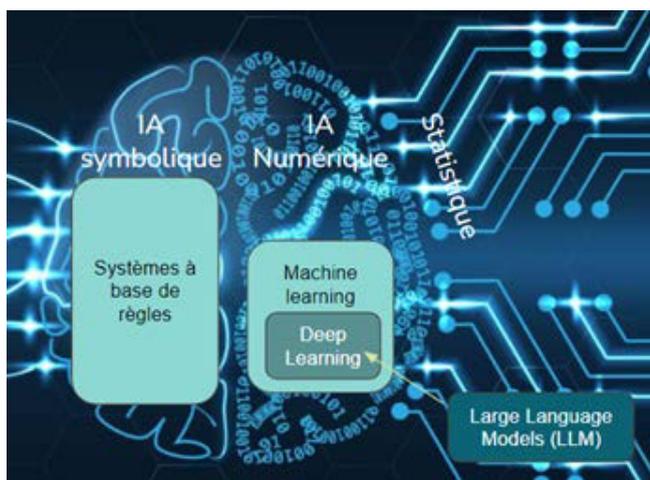


Figure 1 : IA symbolique et IA numérique. Source : *IA et ESR, entre mythes et réalités, quelles limites aux outils existants et quels apports pour l'ESR*, 2023, A. Duffoux.

Les puristes diront que c'est une vision réductrice, et ils auront raison, mais cela permet de fixer le cadre de ce chapitre. Si les techniques basées sur l'IA symbolique ont été plébiscitées dans les années 70 à 90, les techniques basées sur l'IA numérique sont les techniques à l'origine des « buzz modèles » que nous connaissons aujourd'hui et notamment celles basées sur les réseaux de neurones.

Un réseau de neurones est un modèle statistique entraîné, sur la base d'éléments qu'on lui aura appris à reconnaître, à prédire une réponse la plus probable – ou celle qui a été la plus mise en avant lors de son apprentissage (c'est d'ailleurs ce qui mènera à un des biais que nous verrons plus loin dans ce chapitre).

Ce concept n'est pas récent, le premier neurone artificiel datant de 1943, et trouve de nombreuses applications dans notre quotidien

depuis de nombreuses années (notamment dans la reconnaissance d'images, par [exemple](#)). Vous avez vous-même certainement utilisé sans le savoir un des premiers réseaux de neurones déployés à grande échelle par... la [Poste française](#) pour reconnaître automatiquement les codes postaux dans les années 1990 (LeNet, Le Cun 1989).

Un réseau de neurones, c'est un peu une chambre chinoise, nom du **concept proposé par Searle en 1980** (J. R. Searle, "Minds, Brains and programs", *The Behavioral and Brain Sciences*, vol. 3, Cambridge University Press, 1980) et très bien expliqué dans cette [vidéo de Defakator](#).

Pour résumer ce principe, disons que si l'on vous enfermait dans une bibliothèque contenant tous les écrits existants en mandarin et que vous aviez le temps nécessaire pour apprendre par cœur tous ces ouvrages, vous seriez capable de compléter une suite de sinogrammes avec une bonne probabilité que cela semble cohérent pour un éventuel interlocuteur, lui donnant l'illusion que vous parlez couramment mandarin alors qu'en fait vous ne comprenez pas du tout la signification de ce que vous avez écrit. Vous avez juste retenu quel était le sinogramme le plus souvent utilisé à la suite des précédents. **Les outils d'IA actuels fonctionnent ainsi : ils calculent la réponse la plus probable à une question sans en comprendre le sens ni le contexte.** Les réseaux de neurones font ainsi partie des techniques dites d'IA « faible » en opposition aux techniques qualifiées d'IA « fortes » qui, pour l'heure, n'existent pas encore mais alimentent tous les fantasmes et font couler beaucoup d'encre.

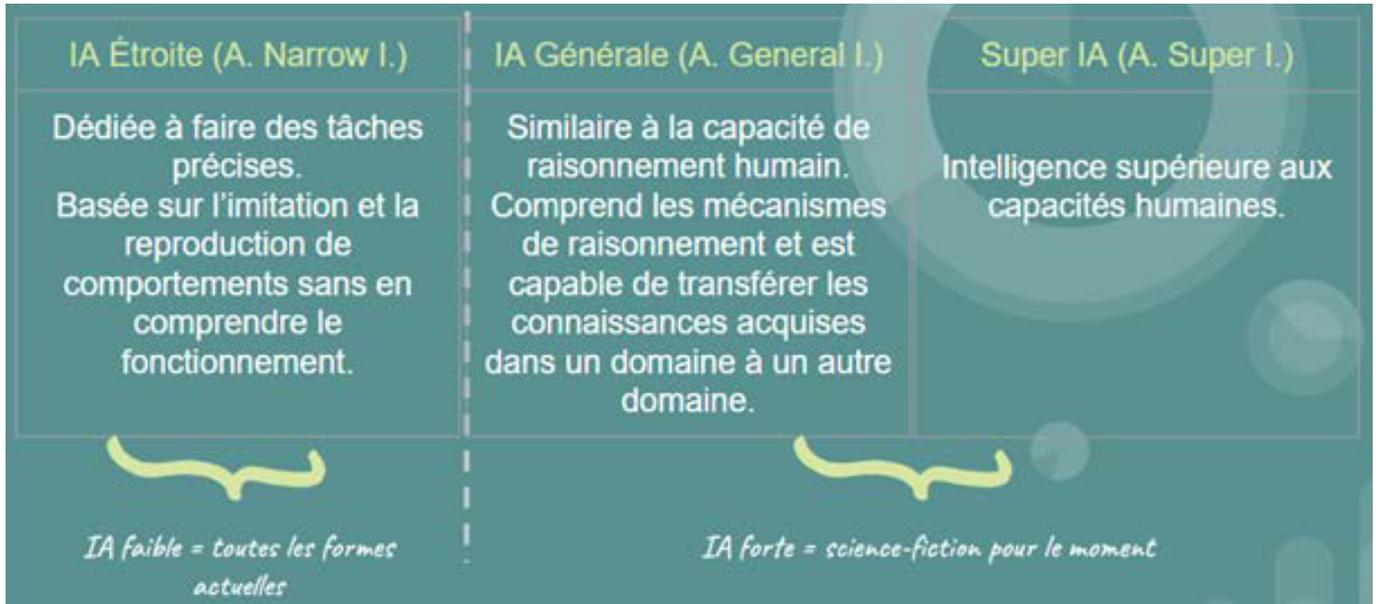
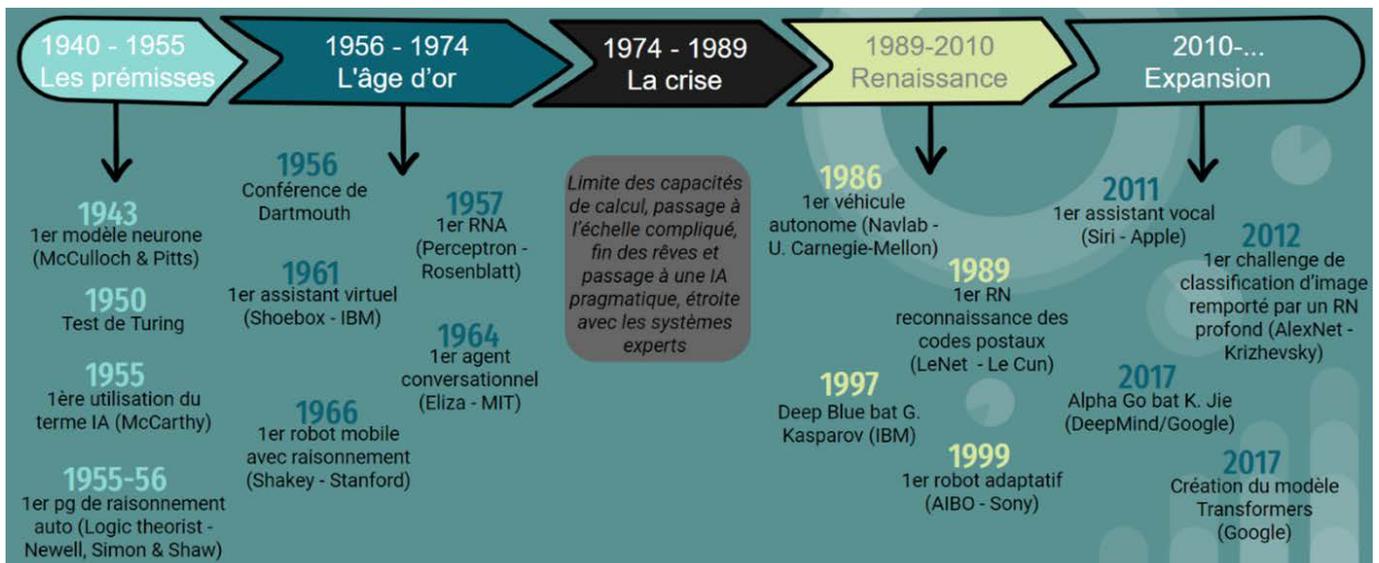


Figure 2 : De l'IA étroite à la Super IA. Source : *IA et ESR, entre mythes et réalités, quelles limites aux outils existants et quels apports pour l'ESR*, 2023, A. Duffoux

Beaucoup d'applications existent déjà et les principes utilisés aujourd'hui ne sont pas récents.

La dernière grande avancée, qui est d'ailleurs plus dans **le domaine du Traitement Automatique du Langage (TAL)**, est la création du modèle Transformers par les chercheurs de Google, modèle présenté dans [une publication](#) qui est devenue une référence et qui a constitué le terreau pour le développement des Large Language Models (LLMs) sur lesquels s'appuient les outils d'IA dites génératives connus aujourd'hui.

Figure 3 : Développement chronologique de l'IA. Source : *IA et ESR, entre mythes et réalités, quelles limites aux outils existants et quels apports pour l'ESR*, 2023, A. Duffoux



2. LES SPÉCIFICITÉS DES IA DITES GÉNÉRATIVES

Les IA dites génératives sont des IA que l'on a entraînées pour générer un contenu (texte, image, son, vidéo, code ...). Dans cette partie, nous nous concentrerons sur les IA génératives textuelles mais le procédé peut être transposé pour la génération d'autres contenus.

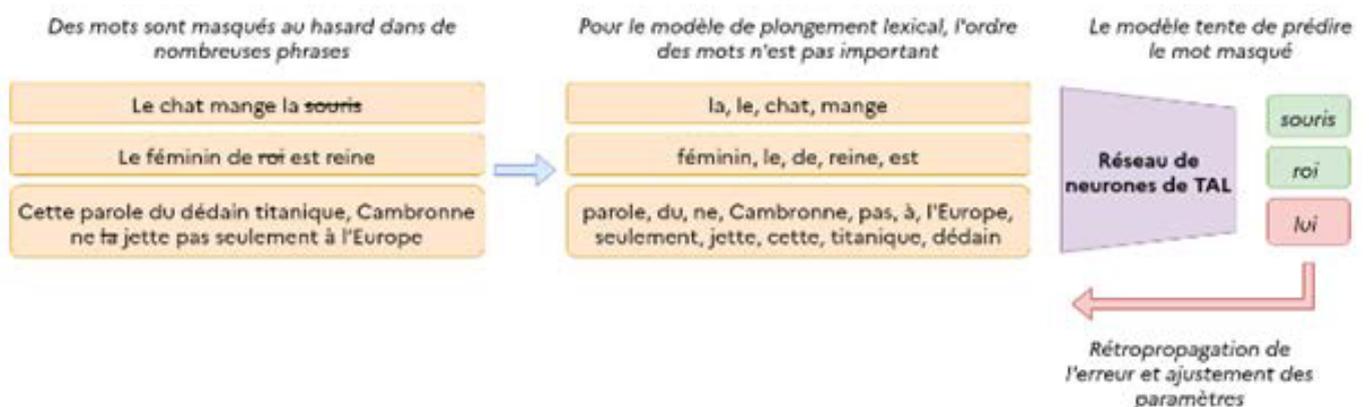
Les IA génératives textuelles actuelles sont entraînées pour prédire le mot le plus probable à la suite d'autres mots en fonction de données d'entrées. Les techniques ont bien sûr évoluées dans ce domaine depuis le plongement lexical (1957) jusqu'au modèle Transformers (2017). Ce dernier permet de traiter d'énormes quantités de données en entrée et utilise le mécanisme d'attention pour reconnaître l'importance des mots dans le contexte des autres mots de la phrase. Ce faisant, il donne l'illusion d'une compréhension et d'une « mémoire ».

On voit déjà que, dès cette première étape, la capacité du réseau de neurones à prédire un mot dépend très fortement du corpus de documents utilisés pour l'entraînement. On devine donc les premiers biais de ce type d'outil.

La plupart des Large Language Models (LLM) conversationnels sont, de plus, affinés en utilisant le principe d'apprentissage par renforcement à l'aide de retours humains (ou RLHF pour Reinforcement Learning from Human Feedback) présenté par [DeepMind](#) en 2017.

Dès lors, le but de ces outils ne sera plus exactement de prédire le mot (ou la suite de mots) le plus probable, mais la suite de mots qui permettra d'obtenir le score de feedback le plus important. Voilà pourquoi une interface de conversation avec ce type de LLM conversationnel renforcé par retour humain tendra à ne pas vous contredire, quitte à dire tout et son contraire dans une même conversation, ou campera obstinément sur ses positions selon son entraînement. Cela implique de garder un recul nécessaire sur les réponses apportées et les biais, amont ou aval, inhérents à ces techniques.

Figure 4 : Exemple d'entraînement de réseau de neurones par plongement lexical. *Source* : ChatGPT où la percée des modèles d'IA conversationnels, Pôle d'expertise de la régulation numérique, 2023



3. LES BIAIS

Les techniques utilisées afin de construire les LLMs utilisés aujourd'hui sont **à l'origine de nombreux biais qu'il faut comprendre pour utiliser de manière pertinente ces nouveaux outils**, dont les apports pour l'enseignement supérieur et la recherche sont cependant réels. Nous distinguerons ici **les biais amont**, liés à conception des LLMs, **des biais aval**, liés à l'utilisation des résultats obtenus. Ces deux catégories de biais, étroitement interreliées (un biais amont pouvant déclencher un biais aval) représentent autant de risques liés à l'usage des LLMs dont il convient d'être conscient pour en avoir un usage raisonné.

3.1. Les biais amont, liés à la conception des LLMs

Ces biais dits amont sont **issus des données d'entraînement et des choix d'architecture** des modèles utilisés pour la conception des LLMs.

Tout d'abord, Les LLMs sont entraînés sur d'immenses corpus de textes issus du web, de livres, d'articles, qui contiennent inévitablement certains biais sociétaux (racisme, sexisme, stéréotypes, etc).

Les modèles peuvent ainsi "absorber" et reproduire ces biais, les perpétuant et risquant de les inscrire d'autant plus dans la société. **On parlera alors de biais de représentation.** En 2018, Amazon a dû arrêter d'utiliser un [outil](#) de recrutement basé sur l'IA qui favorisait les candidats masculins pour des postes techniques (biais aval), car il avait été entraîné à partir de CV reçus au cours des 10 années précédentes, période pendant laquelle la majorité des candidats étaient des hommes. Autre exemple inquiétant : lorsque l'on demande à l'outil de création d'images Stable Diffusion de générer des photos avec le mot clé « inmate » (détenu), plus de 80% des [résultats](#) sont des personnes avec une peau sombre (et ce alors même que les personnes de couleur représentent moins de 50% de la population incarcérée aux USA).

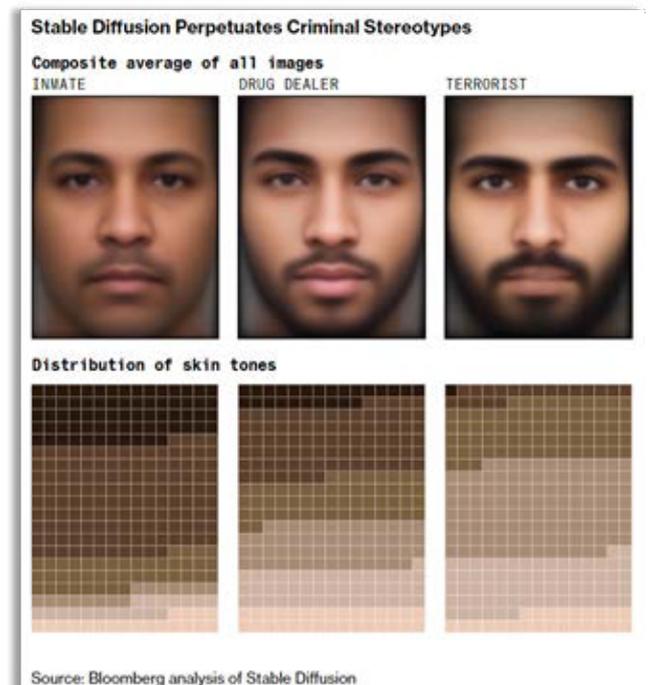


Figure 5 : Perpétuation des stéréotypes criminels par Stable Diffusion. [Source](#)

Ensuite, **les décisions prises lors de la configuration des paramètres du modèle**, telles que le choix des poids ou la définition des objectifs d'apprentissage, **peuvent également introduire des biais.** Si un modèle est optimisé pour maximiser la précision sans prendre en compte l'équité, **il peut ignorer ou mal représenter des groupes minoritaires**, générant des biais de représentation culturelle ou géopolitique. Par exemple, [une étude](#) sur des tâches de classification de texte a montré des résultats inférieurs pour les langues comme le swahili, le télougou ou le ourdou par rapport à l'anglais (Wu et Dredze, 2020).

Conscientes de ces biais et des risques qui y sont associés, **les entreprises tentent de les limiter ou de les rectifier**, développant par exemple des [partenariats](#) avec des producteurs de contenus fiables comme le Financial Times ou le Monde pour entraîner leurs modèles. **Leurs efforts peuvent cependant déboucher sur des résultats pour le moins inattendus**, sinon posant plus de problèmes encore que les biais initiaux. Ainsi, en février 2024, Google fut obligé de suspendre en

urgence la création de visuels à travers son IA générative Gemini car celle-ci proposait, lorsqu'on lui demandait des images de sénateurs américains du 19ème siècle ou de soldats allemands en 1943, des résultats historiquement faux.



3.2. Les biais aval, liés à l'utilisation des résultats des LLMs

Les biais aval émergent de l'utilisation des résultats des LLMs. **Ils peuvent survenir même si le modèle en lui-même est relativement neutre**, mais sont exacerbés par l'interprétation et l'application des résultats du modèle dans des environnements biaisés (les biais amont pouvant donc générer plusieurs biais aval). Parmi les biais aval, **on retiendra notamment les biais d'interprétation** (façon dont les humains perçoivent et interprètent les capacités des LLMs, leur prêtant parfois à tort des capacités de compréhension alors qu'il s'agit de modèles statistiques) **ou les biais de confiance excessive** (fait d'accorder trop de confiance aux résultats obtenus par un LLM).

Comme expliqué précédemment, **ces outils cherchent avant tout à « plaire à leur utilisateur »**. Ceci les amène à générer ce que l'on appelle des hallucinations, c'est-à-dire des résultats créés de toutes pièces, plausibles, mais entièrement faux. Le modèle calcule la suite de mots la plus probable et la plus susceptible de plaire à l'interlocuteur, **indépendamment d'une quelconque connaissance de la véracité de la suite de mots** (qui ne sont que des vecteurs pour le modèle) fournis. Ainsi, ChatGPT a pu vanter les valeurs nutritionnelles des œufs de vache ou indiquer que Simone Veil a été la première femme présidente de la République française.

Utiliser les résultats des IA génératives sans les questionner, les vérifier ou les contextualiser peut mener à des situations incongrues et problématiques. Ce fut le [cas pour un avocat new-yorkais](#) qui, ayant interrogé ChatGPT pour préparer une plaidoirie, n'a cité que des arrêts n'ayant jamais existé ! **Il est donc crucial que, dans nos missions d'enseignement et de recherche, nous soyons conscients de ces risques et formions nos étudiants pour qu'ils prennent toute la distance nécessaire dans leur usage de l'IA générative.**

Figure 6 : Résultats historiquement erronés issus de Gemini. [Source](#)

4. LES COÛTS

4.1. Les coûts environnementaux

L'impact environnemental des usages liés aux outils d'IA générative est loin d'être négligeable et malheureusement, il est encore compliqué de trouver des études complètes sur le sujet. Les grands acteurs du domaine sont réticents à communiquer des éléments à ce sujet et généralement, on ne parle que des impacts liés à l'entraînement et l'utilisation des modèles, mais les impacts liés à la fabrication du matériel nécessaire (serveurs, puces, cartes graphiques...) sont généralement passés sous silence alors qu'ils représentent la majeure partie des impacts. L'impact environnemental dépend aussi grandement du nombre de paramètres des modèles ; les modèles « généralistes » comme GPT4 ou Gemini sont beaucoup plus consommateurs que des modèles spécialisés.

L'utilisation d'eau est un des sujets les plus parlant en termes d'impact environnemental, une [étude](#) intitulée « Making AI Less « Thirsty »: Uncovering and Addressing the Secret Water Footprint of AI Models” (Li et al, 2023) a été publiée à ce sujet en octobre 2023 et tente d'apporter un certain nombre d'éléments factuels à ce sujet. **Les conclusions donnent le vertige**, sans rentrer dans les détails techniques, voici quelques chiffres clés :

- **Fabrication** : très peu de données existent, mais on estime qu'une grande usine de semi-conducteurs peut prélever plusieurs millions de litres d'eau chaque jour. L'eau rejetée peut, de plus, contenir des produits chimiques toxiques et/ou des déchets dangereux, nécessitant un traitement supplémentaire avant d'être réutilisée à d'autres fins. Malgré les efforts des industriels, le taux de recyclage de cette eau est très faible, de l'ordre de 30% en moyenne.
- **Entraînement** : l'entraînement des modèles d'IA nécessite des quantités considérables d'eau, principalement pour le refroidissement des centres de données. L'entraînement d'un modèle de grande envergure comme GPT3 peut

consommer environ 700 000 litres d'eau (une quantité comparable à celle utilisée par la tour de refroidissement d'un réacteur nucléaire). Il s'agit bien de la quantité d'eau consommée, c'est-à-dire évaporée lors du refroidissement des supercalculateurs, et donc non réutilisable.

- **Utilisation** : Une estimation admise aujourd'hui est une évaporation de l'ordre de 1/2 litre d'eau pour 10 à 50 réponses apportées par un modèle comme GPT-3.x. Cette estimation correspond à une utilisation pour de la génération de texte, il n'existe malheureusement pas, à notre connaissance, d'étude comparative d'analyse pour la génération d'autres tâches (génération d'image, vidéo, analyse de texte, ...) mais celles-ci seront plus consommatrices, car elles demandent plus de capacité de calcul. De même les derniers modèles intégrant plus de paramètres se retrouveront également plus consommateurs.

L'énergie nécessaire et l'impact équivalent carbone est un autre moyen de mesurer l'impact environnemental des outils d'IA générative. De ce côté-là aussi, il est compliqué de trouver des études fiables et complètes bien que de nombreux chiffres soient avancés. L'entraînement du modèle GPT-3.5 aurait entraîné une consommation de près de 1 287 MWh, soit l'équivalent de la consommation de 275 foyers français durant une année.

En termes d'[impact équivalent carbone](#), l'entraînement de ce même modèle aurait occasionné 552 tonnes équivalent carbone. Par comparaison, l'entraînement de BLOOM, modèle français, a occasionné un impact équivalent carbone de 25 tonnes, la différence étant due au mix énergétique différent entre USA et France. Il est intéressant de noter que **l'impact n'est pas le même selon les usages**, la génération d'images étant la tâche la plus impactante et consommatrice de ressources.

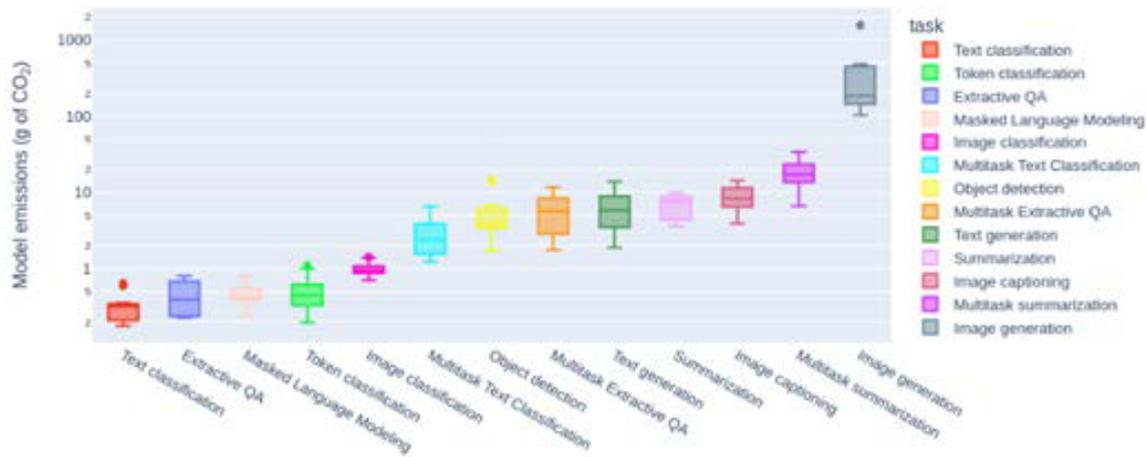


Figure 7 : Impact carbone par type de tâche. [Source](#)

4.2. L'IA peut-elle être écologiquement frugale ?

En réponse aux préoccupations croissantes concernant l'empreinte carbone et énergétique des systèmes d'IA, **l'AFNOR a publié un nouveau référentiel**, initié par Ecolab et fruit de la collaboration d'une centaine d'experts, chercheurs, industriels et administrations. Le référentiel AFNOR Spec 2314 a été publié le 28 juin 2024 en France et offre aux entreprises **un outil concret pour mesurer l'empreinte carbone de leurs modèles d'IA** tout au long de leur cycle de vie.

Un premier exemple de LLM **conçu au sein d'une Grande école française** avec un souci de l'impact environnemental de celui-ci est le [LLM Croissant](#) (CentraleSupélec).

Par ailleurs, la société chinoise DeepSeek a communiqué fortement en janvier 2025 sur le fait que leurs modèles étaient beaucoup plus économes en énergie et en temps (ainsi que financièrement) que les modèles américains, pour des résultats globalement similaires. Cependant, les informations communiquées ne permettent pas encore d'assurer avec certitude qu'il s'agit d'une martingale pour un usage plus écologiquement raisonnable de l'IA.

4.3. Les coûts humains

En complément des coûts environnementaux, l'entraînement des LLMs soulève également **des questions éthiques importantes**, notamment en ce qui concerne les impacts psychologiques sur les travailleurs chargés d'étiqueter les données sources.

Une grande partie du travail nécessaire pour entraîner et maintenir ces systèmes repose sur des millions de travailleurs du clic, souvent basés dans des pays à revenu faible ou intermédiaire. Ces travailleurs, le plus souvent localisés dans des pays en développement ou à très faible coût de main-d'œuvre (par exemple, au Kenya ou aux Philippines), effectuent des tâches répétitives et peu rémunérées, telles que l'étiquetage de données, la modération de contenu et d'autres formes de travail manuel en ligne. Ils sont souvent exposés à des contenus explicites, violents ou perturbants et **peuvent souffrir de stress post-traumatique, d'anxiété et de dépression**. Les [conditions de travail](#) précaires et la pression pour traiter rapidement de grandes quantités de données aggravent ces [effets](#).

5. LE SUJET ÉPINEUX DES ÉVALUATIONS

La sortie de ChatGPT, puis des autres LLMs, a **immédiatement généré des fantasmes plus ou moins réalistes concernant leur utilisation à des fins peu avouables par les étudiants**. S'il est compliqué d'avoir des chiffres fiables sur ce sujet, plusieurs études montrent que les étudiants utilisent effectivement des LLMs (notamment ChatGPT), sans forcément toujours savoir si leur usage est licite ou non. **Les institutions d'enseignement supérieur n'ont en effet pas encore toutes établi de règles claires sur le sujet**. Dans tous les cas, même si les limites des LLMs font qu'ils ne sont pas encore tous en mesure de faire tout le travail à la place des étudiants, **la question des évaluations se pose de manière cruciale dans ce nouvel environnement** : peut-on détecter leur usage ? Comment revoir les évaluations ?

5.1. Les LLMs, des outils puissants, mais pas omnipotents

Petit exemple de limites existant sur les premiers modèles grand public : des enseignants de Terminale ont eu l'idée de donner à ChatGPT le sujet de mathématiques du bac 2023. Une des questions était de faire un calcul de probabilité. Et **si ChatGPT a su utiliser la bonne formule, il a échoué dans le calcul correspondant**, amenant l'enseignant à faire ce commentaire : « La formule utilisée est correcte mais pas le résultat ! Le résultat devrait être 0,32 mais ChatGPT trouve bizarrement 0,18. L'outil est donc sujet, comme les élèves, aux erreurs de calculs ! Erreur dont il est difficile de saisir l'origine ici. ».

Cette anecdote résume parfaitement les capacités et limites des LLM conversationnels, **ils calculent le mot ou le caractère statistiquement le plus adapté sans forcément comprendre le calcul sous-jacent**. Dans ce domaine, l'utilisation de différents modèles d'IA spécialisés capables de communiquer entre eux pour effectuer une tâche semble beaucoup plus efficace et réaliste que l'avènement d'une IA générale.

5.2. Détecter l'usage de l'IA générative dans les travaux étudiants ?

Le premier réflexe de nombreux établissements fut de **procéder avec l'IA générative comme avec le plagiat traditionnel**, c'est-à-dire d'essayer d'en détecter l'usage non autorisé avec des outils informatiques. Parmi ceux qui indiquent pouvoir détecter l'usage de l'IA générative dans les travaux rédigés par les étudiants, citons Turnitin, Compilatio, OpenAi text Classifier, GPT Zero ou encore Draftgoal.

Mais **ces outils sont incapables de repérer parfaitement l'utilisation de l'IA générative pour tricher** (il en va de même si l'on demande à ChatGPT s'il a écrit un [texte](#)). Par exemple, le document d'un étudiant peut renvoyer des résultats positifs alors qu'il n'a fait appel à l'IA générative que pour réviser l'anglais. **Ces faux positifs peuvent entraîner des situations désastreuses pour les étudiants**.

Cela a conduit plusieurs institutions à cesser d'utiliser les détecteurs GPT, comme l'Université Montclair (USA) en novembre 2023 qui a cessé d'utiliser AI Detection via Turnitin (pour information, en février ou mars 2023, Turnitin mettait en avant un taux de fiabilité de 97% de leur détecteur GPT). **Une étude a également révélé que ces « détecteurs » avaient tendance à discriminer les personnes dont l'anglais n'est pas la langue maternelle**, du fait de leurs biais amont.

Attention : ces limites sont valables dans le cadre de la [détection](#) de texte généré par l'IA générative. À l'inverse, **il existe de nombreuses techniques pour détecter des images, de la vidéo ou du son** ainsi créé, rendant leur utilisation illicite plus complexe.

Turnitin. As of November 2023, in response to concerns about false positives and the inability to guarantee the validity of AI detection, the **Advisory Board of the Office for Faculty Excellence** recommended to discontinue the use of AI detection via Turnitin for the assessment of student submissions. The Academic Affairs Council supported that recommendation.

The full text of the Provost's Memorandum on AI Detection Service within Turnitin / Canvas, 11/14/2023.

The AI detection functionality within Turnitin, originally available to Montclair Canvas users, was officially discontinued on November 20, 2023.

This decision was made following similar ones by many other universities, including Vanderbilt, Michigan State, Northwestern, and the University of Texas at Austin. Turnitin has **acknowledged** that its product may have a higher error rate than was initially suggested.

Figure 8 : Publication officielle de l'université de Montclair concernant l'arrêt de son utilisation du module de détection de l'IA générative de Turnitin. [Source](#)

5.3. Le besoin de revoir les modalités d'évaluation

Devant la vague de fond que représentent les LLMs, les établissements de formation n'ont d'autre choix que de **repenser leurs modalités d'évaluation en profondur**. À ce stade, le tâtonnement est général – ce qui est certainement une bonne chose, car il faudra un peu de temps pour voir ce qui fonctionne et ne fonctionne pas.

Dans l'ensemble (en France et à l'étranger), le réflexe principal pour faire face à ces changements a été de **revenir à des examens traditionnels** avec papier et crayon. En parallèle, la mise en place **d'évaluations orales** a également le vent en poupe.

Cependant, ces deux modalités ne permettent pas d'évaluer la même chose qu'un travail qui aurait auparavant été produit pendant plusieurs semaines par des étudiants qui auraient eu le temps de mûrir leur réflexion et d'approfondir le sujet sur lequel ils travaillaient. Il semble donc compliqué de se limiter à ces formats d'évaluation ou à leurs déclinaisons.

De **nombreuses expérimentations voient ainsi le jour** : réalisation d'examens « ouverts » en autorisant l'utilisation de l'IA générative, analyse par les étudiants de la qualité des rendus de LLMs, de leurs limites, et de leurs propres prompts pour en expliquer la construction.

Le foisonnement des expérimentations actuelles aboutira certainement à **un renouvellement en profondeur des évaluations tout en les rendant plus résistantes à la triche** via des LLMs, mais il s'agit d'un processus de (très) longue haleine. D'autant plus que réviser les modalités d'évaluation ne représente qu'une partie de la problématique. L'IA générative bousculant les métiers, la nature même de ce à quoi nos étudiants doivent être formés, évolue rapidement et considérablement, impliquant de **transformer aussi le contenu de ces évaluations**.

6. LES APPORTS POSSIBLES DE L'IA : EXEMPLES ET ENRICHISSEMENTS DES FONCTIONNALITÉS

Les apports de l'IA pour les différentes parties prenantes de l'enseignement supérieur sont extrêmement divers, et dépendent des contextes d'utilisation, du niveau de familiarité des acteurs avec ces outils, des process locaux, ou encore de l'imagination des utilisateurs. Dès lors, nous n'avons pas l'ambition de proposer une liste exhaustive desdits apports, mais plutôt **un ensemble de possibilités que chacun pourra s'approprier et approfondir selon ses propres besoins et contraintes**. Nous concluons la section en présentant rapidement deux modalités complémentaires d'enrichissement des LLMs contribuant à l'amélioration de leurs résultats.

6.1. Différents types d'usages de l'IA

Afin de guider au mieux enseignants, étudiants, équipes administratives et gouvernances des Écoles dans leurs usages de l'IA générative, nous avons tout d'abord **identifié quatre types d'activités « génériques »** que sont l'enseignement, l'apprentissage, la recherche et les missions administratives (elles-mêmes caractérisées par une profonde diversité). Nous proposons, pour chacune d'elles, des moyens de mobiliser l'IA qui ne prétendent bien sûr aucunement à l'exhaustivité.

6.1.1. Activités d'enseignement

Le **tableau 1** présente plusieurs exemples d'utilisation de l'IA générative par des enseignants pour gagner du temps sur des activités chronophages telles que la préparation de documents administratifs ou de cours, la préparation de nouvelles activités, de nouvelles évaluations, etc. Il va sans dire que, comme nous l'avons déjà indiqué, **l'ensemble des contenus ainsi préparés doivent faire l'objet d'une vérification de la part des professeurs**, afin d'éviter les hallucinations susceptibles d'en affecter la qualité. Ce gain de temps peut être redéployé en faveur des étudiants (ex : temps passé à leur faire du feedback, coaching plus individualisé, etc.).

Si les outils « traditionnels » d'IA générative (ChatGPT, Claude, Gemini, Copilot et bien d'autres) peuvent être très utiles pour ces différents usages, **un nombre croissant d'outils répondant à des problématiques spécifiques arrive sur le marché**. En France, certains de ces outils peuvent être identifiés facilement via l'association [EdTech France](#), mais seuls les adhérents y sont répertoriés.

À titre d'exemple, l'**IESEG School of Management** a ainsi déployé la solution de NOLEJ, EdTech française multiprimée. Celle-ci permet aux professeurs de créer rapidement, à partir de leurs supports de cours, des vidéos interactives, des quizz, des résumés de cours, des flashcards ou des mini-jeux permettant aux étudiants de s'entraîner, etc.

L'ensemble de ces outils, s'ils ne sont pas dénués de limites comme nous l'avons vu, offrent des perspectives considérables pour **supporter les enseignants et les apprenants**. C'est ce que montre par exemple une [vidéo](#) de Salman Khan, fondateur de la Khan Academy qui s'est appuyé sur ChatGPT pour mettre au point Khanmigo, outil qui accompagne les étudiants dans leur apprentissage et les professeurs dans leur activité d'enseignement.



Figure 9 : Salman Khan présentant Khanmigo lors d'un TedTalk en avril 2023.

Tableau 1 : Cas d'usage de l'IA générative pour l'enseignement supérieur

Nature de l'usage	Exemples spécifiques
Préparation de syllabus	<ul style="list-style-type: none"> • Définition de Learning Outcomes (taxonomie de Bloom) • Proposition de plan de cours • Glossaire
Préparation de supports de cours	<ul style="list-style-type: none"> • Génération de supports • Génération d'illustrations • Génération de modèles
Préparation d'évaluations	<ul style="list-style-type: none"> • Aide à la génération de QCM/Quiz • Proposition d'exercices d'évaluation • Test de réponses à une évaluation • Aide à la création de grilles d'évaluation (avec barème et rubriques détaillées)
Aide à la correction d'évaluations	<ul style="list-style-type: none"> • Feedback détaillé • Orthographe et grammaire
Aide à l'analyse et à la décision	<ul style="list-style-type: none"> • Génération de tableaux de bord analytiques • Proposition d'analyse des résultats

Tableau 2 : Cas d'usage de l'IA générative pour les apprenants

Nature de l'usage	Exemples spécifiques
Organisation	<ul style="list-style-type: none"> • Plan d'apprentissage • Plan de révisions
Apprentissages	<ul style="list-style-type: none"> • Traduction • Simplification • Reformulation • Explication
Contrôles de connaissances et évaluations	<ul style="list-style-type: none"> • Génération de quiz • Échanges sur un sujet avec l'outil conversationnel • Feedback
Soutien à l'adaptive learning	<ul style="list-style-type: none"> • Évolution dans le cours adaptée aux connaissances et compétences de chacun • Adaptation aux handicaps à l'apprentissage • Rendre les contenus éducatifs plus accessibles

6.1.2. Apprentissage

Le **tableau 2** présente plusieurs exemples d'utilisation de l'IA générative par les apprenants pour **faciliter leurs apprentissages**. Les outils d'IA générative peuvent être utilisés comme tuteur d'apprentissage, proposant un cadre méthodologique, des échanges adaptés aux problématiques des apprenants (adaptation des propositions en fonction du contexte de l'apprenant, comme dans le cas de dyslexie), des reformulations et traductions, des espaces d'entraînement et de confrontation d'idées ... Ces différents cas d'usage s'avèrent **d'autant plus efficaces qu'ils s'insèrent dans un schéma d'apprentissage validé par les enseignants** afin de limiter les éventuels biais inhérents aux outils grand public.

Les apprenants peuvent également s'approprier les outils « traditionnels » d'IA générative pour améliorer leur apprentissage. L'UNESCO a par exemple [listé](#) 10 rôles que ChatGPT peut endosser dans ce cadre.

Les écoles peuvent là encore travailler avec des EdTech, voire développer leurs propres outils.

L'EXEMPLE DU "SKEMA AI TUTOR" POUR AUGMENTER L'EXPÉRIENCE APPRENANTE

Contributeurs : Denys GIORDANO, Irina OTMANINE & Vaitea JACQUIER (SKEMA Business School)

Une réponse à l'évolution des pratiques d'apprentissage :

Compte tenu de **la transformation des usages induite par l'essor des outils d'intelligence artificielle générative**, et pour optimiser le temps des professeurs par rapport à l'accompagnement individualisé des apprenants, la Direction Innovation & Learner Experience de SKEMA a initié le projet « SKEMA AI Tutor ». Cet outil vise à **offrir une expérience d'apprentissage interactive et accessible, tout en renforçant l'autonomie des étudiants**.

Le « SKEMA AI Tutor » se concentre sur trois usages clés : approfondissement des notions pour clarifier les concepts vus en classe, résumé de documents (articles, lectures, interviews), génération de quiz personnalisés pour s'entraîner efficacement.

L'outil a été développé via une plateforme nocode (Glide) intégrant des composants utilisant des modèles d'IA différents (OpenAI, AzureML et Google Cloud) selon les usages. L'approfondissement de

notion utilise des fonctionnalités classiques de LLM (comme l'on pourrait le faire via ChatGPT) pour lequel des prompts ont été construits en amont pour orienter les réponses de l'IA en fonction des cours dans lesquels les étudiants sont inscrits. Pour le résumé de document et la génération de quizz, l'outil est basé sur un RAG et est connecté aux ressources documentaires du cours, préalablement choisies et formatées avec le professeur.

Développer une solution applicative propre permet, entre autres, de mieux maîtriser l'expérience utilisateur, d'assurer une cohérence globale dans la stratégie de prompting, d'être plus flexible et agile dans les développements et d'assurer l'égalité d'accès entre tous les étudiants.

Un pilote ciblé et riche en enseignements :

Le projet pilote a été lancé en janvier 2024 dans le cours « Grands enjeux géopolitiques » de Frédéric Munier, avec 170 étudiants répartis sur trois campus français. Pendant un semestre, l'outil a été intégré aux contenus pédagogiques et mis à disposition des étudiants. Des enquêtes et un focus group ont permis de collecter des retours précieux :

- Plus de 80 % des étudiants ont adopté le tuteur (2 750 requêtes)

- Approfondir des notions (10%) ;
- Résumer des documents (70%) ;
- Générer des quizz (20%),

4 000 questions ont été générées sur les 580 quizz déclenchés, la génération de quiz étant plébiscitée pour sa praticité et son efficacité.

- Les périodes de cours et d'examens ont connu des pics d'utilisation.

Des améliorations continues pour une adoption élargie et un projet stratégique pour l'avenir :

Les retours ont également mis en lumière des axes d'amélioration, tels que la qualité des réponses de l'IA, l'ergonomie, et l'ajout de fonctionnalités. Fort de ce succès, le « SKEMA AI Tutor » a été étendu dès septembre 2024 à 700 étudiants et sera progressivement intégré à d'autres matières comme le marketing et le droit. Des idées de nouveaux cas d'usage liées à la gestion de carrière, à l'apprentissage de langues, à la vie associative, ont pu être évoquées et sont intégrées à la roadmap de développement de l'établissement. Ce projet illustre l'ambition de SKEMA de placer l'innovation technologique au service de son ADN pédagogique, tout en répondant aux défis d'un enseignement supérieur en constante évolution.

6.1.3. Activités de recherche

L'IA générative peut également **fournir une aide précieuse aux chercheurs aux différentes étapes de leur travail**, qu'il s'agisse du design de la recherche, de la découverte ou formalisation de la revue de littérature, de la collecte et de l'analyse de données, de la finalisation ou édition de leur document final. L'IA peut par exemple servir pour brainstormer ou tester des idées, aidant à la génération de nouvelles réflexions théoriques ou méthodologiques. Elle peut aussi aider à la vulgarisation des travaux de recherche, en aidant à générer des explications simplifiées ou des infographies pour rendre des concepts complexes accessibles à un plus large public.

Là encore, **des outils spécifiques existent et peuvent permettre de gagner un temps précieux**, comme, par exemple, Semantic Scholar, scite.ai ou Elicit pour les revues de littérature (leurs fonctionnalités, variables selon les outils, vont par exemple de l'identification d'articles pertinents à l'extraction d'informations clés comme la méthodologie, la taille des échantillons ou les résultats principaux, en passant par la mise en relation entre les différents travaux de recherche), Ask Data pour l'analyse de données (fonctionnalité d'IA intégrée à Tableau qui permet aux utilisateurs de poser des questions en langage naturel sur leurs données et d'obtenir des visualisations en retour), etc.

Néanmoins, **leur praticité ne doit pas faire oublier qu'il ne s'agit là que d'outils complémentaires qui viennent renforcer l'arsenal du chercheur pour gagner en qualité, pertinence et efficacité.**

Comme tout autre outil, ils doivent donc être utilisés en conscience de leurs limites et risques associés, et notamment des considérations éthiques qui les accompagnent. De plus, il convient de bien faire attention aux bases de données sur lesquelles ils s'appuient pour générer leurs résultats. Certains de ces outils ne mobilisent par exemple que des articles publiés chez un éditeur en particulier, quand d'autres entretiennent un certain flou quant à la manière dont l'algorithme va sélectionner des articles mettant plus en avant un éditeur qu'un autre.

L'IA doit donc être vue comme un outil complémentaire qui améliore, plutôt que remplace, l'expertise et le jugement humains dans le processus de recherche comme dans toutes les autres activités évoquées ici.

Tableau 3 : Les différentes étapes de la recherche auxquelles l'IA peut contribuer



6.1.3. Activités administratives

Les outils d'IA générative peuvent également fournir **une aide non négligeable aux personnels administratifs des écoles, en améliorant à la fois leur productivité et la qualité du service rendu**. La diversité des profils et des missions de ces personnels fait que nous nous limiterons ici à n'en donner que quelques exemples, en essayant d'identifier des usages qui parleront à chacun.

1. Assistance à la réponse aux sollicitations des interlocuteurs : en générant des réponses automatisées et personnalisées aux questions fréquentes des étudiants, parents ou enseignants via des chatbots ou des modèles de traitement du langage naturel, les outils d'IA contribuent à réduire la charge de travail des équipes.

2. Aide à la planification d'événements : l'IA peut faciliter l'organisation d'événements en automatisant la gestion des disponibilités, la réservation de salles et l'envoi d'invitations, tout en proposant des rappels automatiques aux participants.

3. Aide à l'analyse d'emails : Des outils comme Copilot permettent de résumer les mails de la journée ou de la semaine, de proposer de les rassembler par thématique et d'enrichir les contenus sur la base des documents accessibles à l'utilisateur. Ils permettent également de lister les actions à mener et les différents degrés d'urgence des emails reçus.

4. Aide à la génération de documents (emails, attestations ...): Les LLMs permettent de générer rapidement des emails types, des attestations de présence, ou des courriers administratifs personnalisés en fonction des besoins de chaque service.

5. Aide à la rédaction de rapports et de comptes-rendus de réunions : L'IA peut produire automatiquement des comptes-rendus de réunions, aider à la rédaction de rapports de gestion ou d'audit, de synthèses financières ou de bilans d'activité en structurant et en résumant efficacement les données collectées (à vérifier, bien sûr – nous n'insisterons jamais assez sur ce point).

6. Aide à la traduction de documents : Les services des relations internationales peuvent utiliser l'IA pour traduire automatiquement des documents administratifs, voire des conversations orales en temps réel, facilitant ainsi les échanges avec les étudiants et partenaires étrangers.

7. Aide à la génération et à l'analyse d'indicateurs : Les outils d'IA peuvent extraire des indicateurs de performance clé à partir de bases de données, automatiser la création de tableaux de bord et fournir des insights sur la gestion des ressources.

Tout ceci est sans compter des usages spécifiques à certaines typologies de personnels. Ainsi, les équipes IT peuvent s'en servir comme **outils complémentaires de détection et de prédiction de comportements anormaux sur le réseau informatique**.

6.2. RAGs et Fine-Tuning : deux moyens complémentaires pour enrichir l'IA générative

Il est possible d'améliorer la qualité des résultats obtenus dans le cadre des quatre types d'usages précédents en affinant les outils d'IA générative utilisés. Parmi les multiples manières existantes, nous citerons ici le RAG (Retrieval-Augmented Generation) et le fine-tuning, complémentaires et combinables.

Le RAG est une technique qui permet aux LLMs d'**accéder à une base de connaissances externe pendant la génération de texte**. Lors de la génération, le modèle formule des requêtes pour aller chercher des informations pertinentes dans cette base de connaissances. Les passages retrouvés sont ensuite intégrés au contexte de génération du modèle pour l'aider à produire une réponse plus informée et factuellement correcte. **Cela permet d'augmenter les capacités des LLMs en leur donnant accès à des connaissances qui ne sont pas présentes dans leurs données d'entraînement de base.** C'est donc particulièrement utile pour les tâches nécessitant des connaissances spécifiques à un domaine, des informations à jour et pour la confidentialité des données lorsque le LLM tourne en local.

Quant au fine-tuning, il consiste à **continuer l'entraînement d'un LLM pré-entraîné sur une tâche ou un domaine spécifique, avec un jeu de données plus ciblé**. Cela permet d'adapter le modèle à une application particulière sans avoir à l'entraîner à partir de zéro, ce qui nécessiterait beaucoup plus de données et de puissance de calcul. Pendant le fine-tuning, seuls certains paramètres du modèle sont mis à jour, généralement les dernières couches. Les représentations de base apprises pendant le préentraînement sont ainsi préservées. Ainsi, le fine-tuning **permet d'obtenir de meilleures performances qu'un modèle générique pour la tâche visée**. On peut citer comme [exemple](#) de fine-tuning l'adaptation d'un LLM basé sur BERT par la NASA pour le spécialiser au domaine de l'aviation.

Ces deux techniques ont ainsi pour but d'améliorer les capacités des LLMs, de manière différente et complémentaire, le RAG pouvant être appliqué à un modèle déjà fine-tuné. Dans ce cas, le modèle fine-tuné avec accès à une base de connaissances via le RAG aura à la fois de bonnes performances sur sa tâche (grâce au fine-tuning) et des capacités étendues grâce à des informations externes (grâce au RAG). **Ces informations peuvent être mises à jour régulièrement pour assurer le maintien dans le temps de la qualité et de la fiabilité des réponses**. Cependant, qu'elles soient combinées ou considérées indépendamment l'une de l'autre, RAG et fine-tuning ne constituent pas une martingale absolue. Par exemple, elles ne garantissent pas l'absence d'hallucinations ou de biais dans les réponses apportées (même si celles-ci peuvent être réduites), peuvent avoir un coût computationnel important (stockage des données, capacité de traitement informatique des requêtes, etc.), et nécessitent d'apporter la plus grande attention à la qualité des données (d'entraînement pour le fine-tuning, des documents pour le RAG).

7. COMMENT SE LANCER ?

Contributeurs : *Alain GOUDEY & Bérénice DORIDO (NEOMA), Loïc PLÉ (IÉSEG), Emmanuelle CHAVANNE & Cécile WELTMAN (Arts et Métiers), Amandine DUFFOUX (EIGSI et Arts et Métiers)*

À ce stade de leur développement, sinon de leur compréhension, **il convient d'admettre que nous sommes encore collectivement en phase d'exploration plus que d'exploitation réelle des outils d'IA.** Cette exploration se fait à la fois au niveau institutionnel (définition de politiques liées à l'utilisation de l'IA générative) et individuel (exploration des usages potentiels par les professeurs, mais aussi les équipes administratives et, bien sûr, les étudiants).

7.1. « Unlearn & Relearn » : le nouveau monde de l'IA générative

ChatGPT est sans conteste l'outil d'IA générative qui a le plus attiré l'attention. La diffusion de dialogues entre l'outil et des utilisateurs, repris par de multiples médias (généralistes ou spécialisés, réseaux sociaux, etc.) a **généré toute sorte de fantasmes autour de l'impact qu'il pourrait avoir sur une infinité de secteurs d'activité, dont bien sûr celui de l'éducation.** Si les premières réactions concernèrent immédiatement la manière dont les étudiants pourraient s'emparer d'un tel outil pour échapper au travail qui leur est demandé, elles furent très vite suivies par des réflexions plus fines sur la manière dont ChatGPT pourrait contribuer à la fois à l'apprentissage (côté étudiants, donc) et à l'enseignement (côté professeurs).

Les modalités d'usage de ces outils impliquent une révision assez vaste des pratiques de ces deux publics, un vrai processus de type « unlearn & relearn ». **Savoir désapprendre et réapprendre sont en effet clés dans un monde où les modalités de création de contenus changent du tout au tout,** et s'ouvrent à des personnes ne disposant pas des compétences traditionnellement requises pour cela. En effet, avec l'IA générative, la création artistique (images, vidéos, etc.) ne dépend plus des pures compétences techniques (qui peuvent être proches de zéro), mais de la capacité d'une personne

à exprimer en langage naturel une demande à une IA générative, et à affiner progressivement cette demande pour arriver au résultat désiré (voire à mieux qu'attendu encore). **Désapprendre d'anciens réflexes et apprendre de nouvelles compétences et manières de penser, voilà donc le monde dans lequel nous plonge l'IA générative.**

Dans un tel contexte, **NEOMA** a mis en place sa stratégie d'IA générative entre novembre 2022 et juillet 2023, selon les cinq étapes suivantes :

- **La veille et la vision initiale :** Portée par le Learning Lab, l'équipe de NEOMA dédiée à la R&D qui comprend notamment un poste de curateur, et qui a alimenté l'ensemble des équipes liées à la transformation numérique de l'école.
- **La démarche de test & learn :** Réalisation de tests par des membres du Learning Lab et des professeurs de l'école pour évaluer la performance, la fiabilité et l'intérêt de l'IA générative, avec une mise en commun de leurs résultats.
- **La sensibilisation et l'organisation des équipes :** mise en place d'un groupe de travail composé de la Direction des Programmes, la Direction de la Faculté, la Direction des opérations, la Direction de la Pédagogie Innovante, le Learning Lab, la Direction des Systèmes d'Information mais aussi de la Library pour valider la vision stratégique et commencer à porter les changements nécessaires dans les modes opératoires de l'École.
- **La formation de la faculté et des étudiants :** Présentation des outils, de différents cas d'usage, élaboration de règles concernant l'utilisation par les étudiants. Sur ce point, il a été décidé qu'à NEOMA, les outils d'IA génératives sont utilisables (intelligemment) par les étudiants tout le temps : ils peuvent les mobiliser pour répondre aux devoirs et travaux demandés par les professeurs, tant que ces derniers ne les interdisent pas explicitement (dans

des cas de figure qui doivent faire sens). Les règles d'utilisation ont été rédigées pour éviter les risques d'utilisation abusive par les étudiants (utilisation de ChatGPT en faisant croire qu'il s'agit de sa propre production). Les étudiants ont aussi reçu une keynote sur le sujet à la rentrée 2023/2024. Côté professeurs, des conseils ont été donnés pour rédiger des consignes de devoirs, et deux ateliers de sensibilisation ont été organisés. Un groupe de travail dédié comprenant 14 professeurs volontaires pour des expérimentations approfondies fut aussi mis en place.

- **La mise à jour en continu des savoirs et des pratiques :** Devant l'évolution permanente et ultrarapide de ces outils, la veille, et la diffusion de celle-ci, continue et joue un rôle primordial.

7.2. Une vision collaborative et à 360° de l'IA générative selon un principe de test & learn.

L'arrivée des outils d'IA générative a suscité de colossales interrogations parmi la communauté académique, partagée entre les défis et les opportunités qu'ils soulèvent. Outre le sommet émergé de l'iceberg que représente ChatGPT, des centaines d'outils sont maintenant disponibles pour faciliter les processus d'écriture, générer du code, créer des images ou de la vidéo, etc. Un simple passage sur le site futurepedia.io, par exemple, permet de se rendre compte tant de l'ampleur du phénomène que de la diversité des secteurs d'activité – et, par ricochet, de la diversité des formations potentiellement concernées par ces outils.

Face à cette révolution, l'**IESEG** a opté pour une approche en plusieurs étapes. **Tout d'abord, une phase d'information et de sensibilisation,** débutée fin 2022. Le Directeur de l'Organisation et des Systèmes d'Information (DOSI) et le Directeur de la Pédagogie ont ainsi conjointement présenté ces outils au Comité de Direction de l'École, à l'ensemble des professeurs, permanents et intervenants extérieurs, et à tout le personnel administratif pour favoriser une prise de conscience et contribuer à l'acculturation collective vis-à-vis de la nature et des enjeux induits par cette mutation d'ampleur.

Cette présentation fut également diffusée auprès de certaines entreprises partenaires de l'École, qui souhaitaient comprendre les enjeux de l'IA générative.

Dans un deuxième temps, plusieurs groupes de travail ont été mis en place, composés soit d'administratifs, soit d'académiques. Pour les premiers, il s'agissait de **réfléchir à la manière dont ces outils peuvent contribuer à diminuer la lourdeur associée à des tâches répétitives et à enrichir la qualité du service fourni à leurs différents interlocuteurs internes et externes.** Les seconds ont été répartis selon trois thématiques, étroitement interreliées : la conception d'un cours (course design), la manière de le délivrer (course delivery) et les modalités d'évaluation des étudiants (course assignments and assessments). Les professeurs composant ces trois groupes se sont appuyés sur des recherches, des témoignages, leurs propres expérimentations, etc. pour produire des contenus qui ont alimenté la troisième phase du dispositif, à savoir le partage de ces connaissances acquises sur les usages possibles de ces outils avec les autres professeurs de l'IESEG. **De la sorte, les professeurs de l'IESEG seront à même de former eux aussi les étudiants à une utilisation raisonnée de ces outils,** que ces derniers doivent apprendre à maîtriser tout en les utilisant avec un regard critique, pour leur vie professionnelle future.

En parallèle, nous avons également développé, avec l'aide de plusieurs professeurs, **une charte encadrant l'utilisation de l'IA générative par les étudiants, afin de limiter, sinon éviter, tout comportement abusif assimilable à de la triche ou du plagiat.** Cette charte a été intégrée à la politique antiplagiat de l'école, et précise les conditions dans lesquelles l'usage des outils d'IA générative peut être considéré comme du plagiat, et celles dans lesquelles cet usage est autorisé, voire encouragé.

Dans un quatrième temps, nous avons **conçu et déployé des formations à destination des professeurs, des personnels administratifs et des étudiants.** Si certains éléments étaient communs aux trois publics, comme la présentation générale des LLMs et de leurs avantages et limites (avec un reformatage complet adapté aux attentes de la Génération Z pour les formations aux

étudiants) ; d'autres étaient spécifiques aux besoins des uns et des autres. Par exemple, la formation des enseignants incluait trois sections dédiées portant respectivement sur le design, l'animation et l'évaluation des cours.

En parallèle de ces quatre étapes, **différentes solutions répondant à des besoins explicitement exprimés ou identifiés par la Direction de la Pédagogie ou par la DOSI de l'IESEG ont été testées**. Ces tests, conduits avec des professeurs ou des personnels de différents services, nous ont amené à sélectionner ou à exclure des outils selon leur niveau de valeur ajoutée pour ces publics, mais aussi en fonction de leur contribution estimée à la qualité de l'expérience étudiante.

Au regard des développements continus et accélérés des outils d'IA générative, **ce travail se poursuit selon une logique de test & learn**, afin de continuer à accompagner ces développements de manière réfléchie, construite et collaborative avec les différents types d'acteurs concernés.

7.3. Accompagner et former le professionnel et le citoyen

Comme mentionné précédemment, l'IA générative peut être considérée comme un outil puissant permettant d'automatiser et de faciliter certaines tâches professionnelles. **Cependant, limiter son utilisation au seul cadre professionnel serait réducteur, car les étudiants et certains collègues utilisent également cet outil pour des tâches quotidiennes**. Cette absence de frontière claire entre l'utilisation privée et professionnelle de l'outil rend essentielle la formation de tous aux principes de fonctionnement et aux risques associés, notamment en matière de protection des données, et ce, au-delà des missions spécifiques de chacun.

Sous l'impulsion du [projet CAIRE](#) (Citizen-oriented Artificial Intelligence training for a Responsible Education), Arts et Métiers et ses partenaires – le CNAM, le CESI, PSB et UBE – visent à former un large public, qu'il s'agisse d'étudiants, d'enseignants, de personnels administratifs et techniques des établissements partenaires, de membres d'autres institutions ou de personnes en formation tout

au long de la vie, aux bases scientifiques des IA, à leurs risques éthiques et juridiques, ainsi qu'à leur impact environnemental. **L'objectif est de fournir à chacun les clés pour une utilisation raisonnée et responsable de ces outils, tant dans le domaine professionnel que privé.**

Un élément clé de la formation CAIRE est la proposition d'ateliers en présentiel, permettant la prise en main, des tests et des échanges autour des outils d'IA générative. **Ces ateliers réunissent différents publics, tant au niveau des apprenants que des formateurs, afin d'enrichir les utilisateurs par la diversité des points de vue**. Ainsi, les formateurs des premiers niveaux de la formation peuvent ne pas être des experts de la discipline, démontrant que tout le monde peut s'approprier ces outils et facilitant les échanges entre pairs. Les mixités sociales et professionnelles dans ces ateliers permettent des échanges riches et des accompagnements personnalisés, où les utilisateurs chevronnés peuvent aider les plus réticents ou craintifs à se familiariser avec l'outil, ce qu'ils ne feraient pas nécessairement seuls devant leur écran.

Ces mixités dans la formation, lorsqu'elle est proposée dans le cadre de la formation tout au long de la vie au sein d'un établissement d'enseignement supérieur et de recherche, permettent, entre autres, de **recréer du lien et de la discussion entre les différentes catégories de personnel au-delà des fonctions existantes**.

7.4. L'IA générative : des conseils pour se projeter dans un futur mouvant

Comme indiqué précédemment, les outils d'IA génératives constituent un vrai défi pour les établissements d'enseignement supérieur. Leur taux d'adoption, le rythme des avancées dans le domaine, les possibilités liées en font un des sujets majeurs de transformation de l'enseignement supérieur. Face à la profusion d'information, aux débats d'opinion, aux usages existants, il n'est pas toujours évident de savoir par quel bout se saisir du sujet et quelle méthodologie mettre en place pour accompagner les différents utilisateurs. Comme tout outil numérique et comme nous l'avons déjà indiqué dans ce chapitre, les outils d'IA générative ne sont que des outils au service de la stratégie d'établissement. Pour comprendre et appréhender en quoi ces outils peuvent accompagner le développement de nos établissements, il est donc important d'adopter une démarche structurée (Figure 6).

Figure 10 : Exemple de roadmap d'accompagnement dans la prise en compte des enjeux liés aux outils d'IA générative (source : IA et ESR, entre mythes et réalités, quelles limites aux outils existants et quels apports pour l'ESR, 2023, A. Duffoux, EIGSI)

01	Constituer un comité de pilotage	Avec un pilote dédié. Ce comité de pilotage doit être multidisciplinaire et intégrer les différentes parties prenantes de l'établissement (Ingénieurs pédagogiques, enseignants, apprenants, ...)
02	Organiser une veille thématique	Face à l'abondance de points de vue et articles d'opinions, sélectionner les bonnes sources (Commission numérique CGE, Ressources Canopé, Équipes d'appui pédagogique plus avancées, ...)
03	Faire un état des lieux	Pour mesurer le taux d'adoption des outils d'IA générative au sein des différentes populations et surtout leur niveau de connaissance.
04	Sensibiliser	Tous les acteurs de l'établissement : Équipes d'appui pédagogique, direction, enseignants, chercheurs, apprenants, personnel administratif . <i>Thèmes à aborder : Présentation générale de l'IA, IA génératives, biais et enjeux éthiques, hallucinations, coûts environnementaux, apports possibles, le prompt, impact sur les évaluations, ...</i>
05	Constituer des groupes thématiques	Exemples: Impact et adaptation des évaluations, usages pour les enseignants, usages pour les apprenants, usages pour les chercheurs, adaptation des process administratifs, adaptation des formations... Penser à intégrer les apprenants dans chaque groupe
06	Rédiger la politique de l'école	Mettre à jour le règlement des études et/ou la charte anti-plagiat Rédiger une lettre de cadrage à destination des personnels (enseignants et administratifs) Intégrer les impacts dans le schéma directeur de la transformation numérique de l'école
07	Proposer un plan de formation	Plan dédié et adapté à chaque acteur (Équipe de soutien pédagogique, enseignants, chercheurs, apprenants, administratifs)
08	Accompagner les projets	Accompagner les équipes métier (pédagogique, recherche, SI, fonctionnels) dans le suivi des projets intégrant des outils d'IA

CONCLUSION

Pour conclure, l'intégration des outils d'IA générative dans l'enseignement supérieur et la recherche offre **des opportunités inédites pour enrichir l'apprentissage et la recherche**, en permettant notamment une personnalisation des apprentissages et un gain de productivité et d'efficacité. Il est néanmoins **essentiel de bien mesurer les biais et les limites inhérents à ces technologies ainsi que leur impact environnemental afin de maximiser les fonctionnalités offertes.**

Comme pour tout nouvel outil, il convient en premier lieu de **réfléchir à l'intérêt de l'utiliser ou non** selon les besoins et à ce titre, les éléments fournis dans la partie « Quand utiliser le numérique (et quand ne pas l'utiliser!) » du [chapitre 1](#) « Inspirer et renouveler les pédagogies » sont tout à fait pertinents pour un déploiement d'outils d'IA au sein des grandes écoles.

Il convient de rester vigilant par rapport aux nombreux biais et aux nouveaux risques posés par l'usage des LLMs, qu'ils concernent la propriété intellectuelle des données et des productions, le respect de la confidentialité des [données utilisées](#) et saisies ou encore les risques de [tromperie](#) et de [désinformation](#) – sans parler des nouveaux usages des outils d'IA générative à des fins purement illicites (spams, phishing, etc.) ou les problématiques de [cybersécurité](#).

Dans ce contexte, **Il est essentiel d'adopter une organisation permettant d'effectuer une veille continue, tant le contexte technique et applicatif évolue rapidement.** Les initiatives lancées par les Grandes écoles montrent une volonté d'adopter ces technologies de manière réfléchie et responsable. L'approche à 360 degrés, qui englobe la formation, la sensibilisation des différents acteurs et la mise en place de cadres éthiques, est essentielle pour maximiser les bénéfices tout en minimisant les risques.

