PISTES POUR INTÉGRER L'IA GÉNÉRATIVE DANS SON ENSEIGNEMENT

12 décembre 2024

pascal.vangrunderbeeck@uclouvain.be



linkedin.com/in/vgpascal











AU PROGRAME

1. La sidération face à l'IA générative :

Comprendre cette nouvelle réalité

2. Réassurance et pédagogie :

Comment accompagner le changement, une question de posture?

3. Le temps des balises :

Repères, éthiques et critiques dans l'enseignement

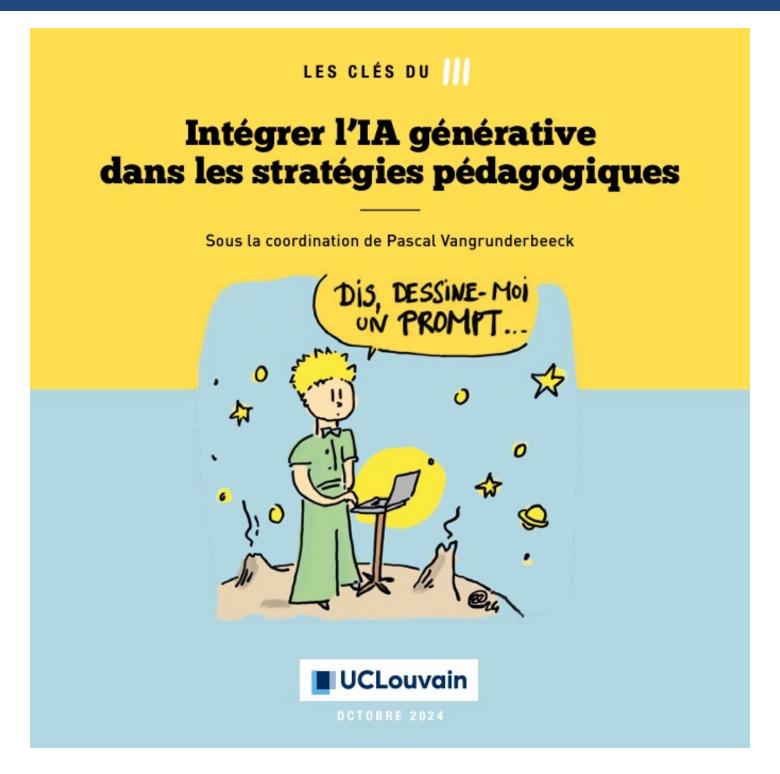
4. Assurer équité et transparence :

Former aux usages critiques de l'IA

5. Du deuil de la détection à la refonte des évaluations :

Se concentrer sur les compétences humaines

Pistes pour favoriser un travail réflexif avec l'IA

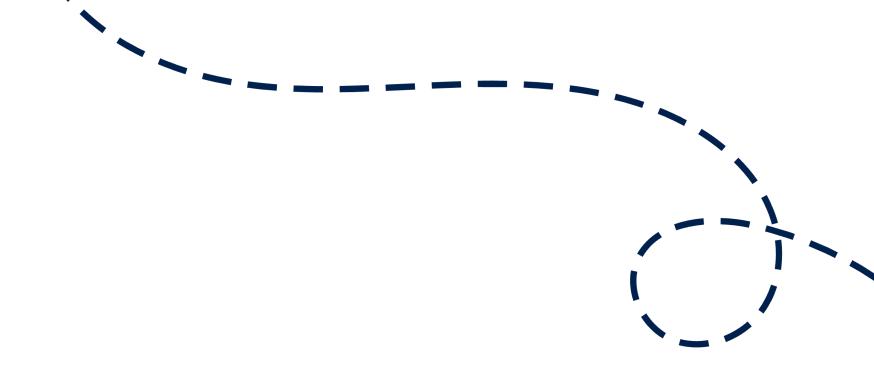


bit.ly/ia-III

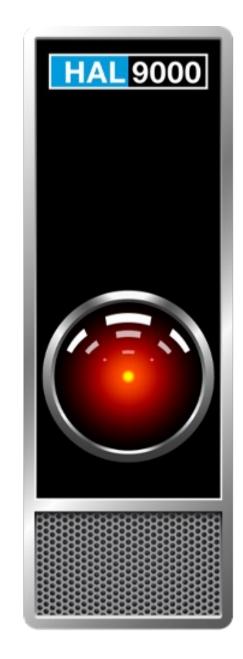
La sidération face à l'ia générative

Définir et comprendre cette nouvelle réalité



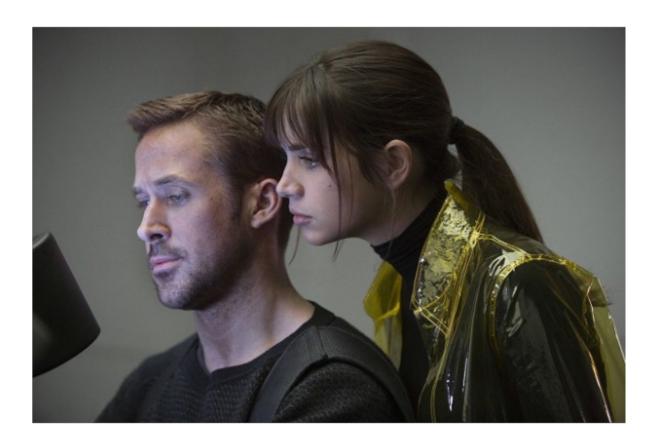


Parler à une machine... science-fiction?





« Her » de Spike Jonze (2013)



« Blade Runner 2049 » Denis Vileneuve (2017)

« 2001, A SpaceOdyssey » de Stanley Kubrick (1968)

Diversité et qualité des contenus générés

Generative

Cette technologie peut **générer de nouveaux contenus** (texte, image, audio, vidéo) de façon aléatoire à partir d'une information d'entrée (« prompt ») saisie par l'utilisateur. Un **prompt précis** et complet conduit à une **réponse plus affinée** et spécifique.



Transformers

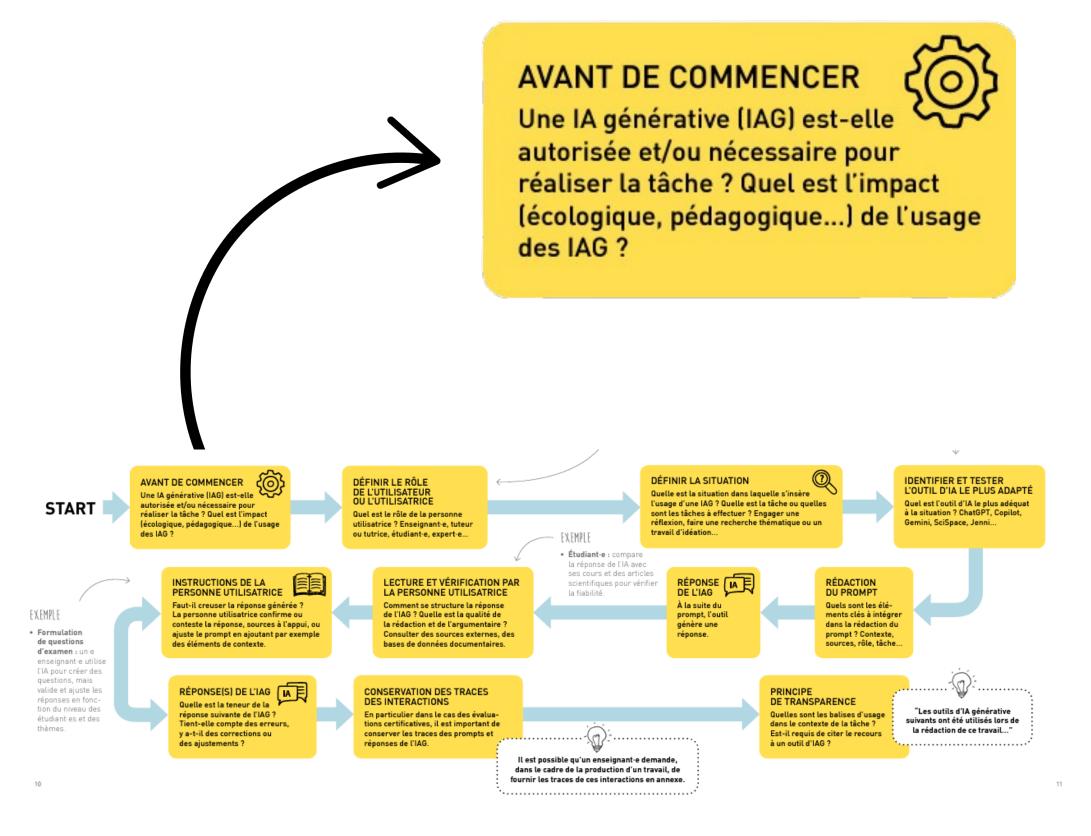
Un modèle de langage comme GPT-4 utilise une architecture de réseaux de neurones appelée « transformer » qui permet d'établir des relations mathématiques entre les mots et de déterminer le sens d'un mot en fonction du contexte, en se focalisant sur les mots les plus pertinents. Le contenu généré s'approche ainsi du langage naturel.

Pretrained

Cette IA est **pré-entraînée sur des données massives** (ensemble de Wi-kipédia, GitHub, réseaux sociaux, etc.). Ces données sont capitales dans l'apprentissage des modèles de langage et des modèles génératifs d'IA.

Source: Intégrer l'IA générative dans les stratégies pédagogiques - http://hdl.handle.net/20.500.12279/1089.3

Questions à considerer avant de commencer l'usage de l'IA



Tenir compte des enjeux de Développement Durable et da la Transition (DD&T)



"Image Latente" ©Quentin Sombsthay

Derrière l'IA générative, la violente réalité des travailleurs du clic

Empreintes écologiques

En équivalent CO₂ (kg)

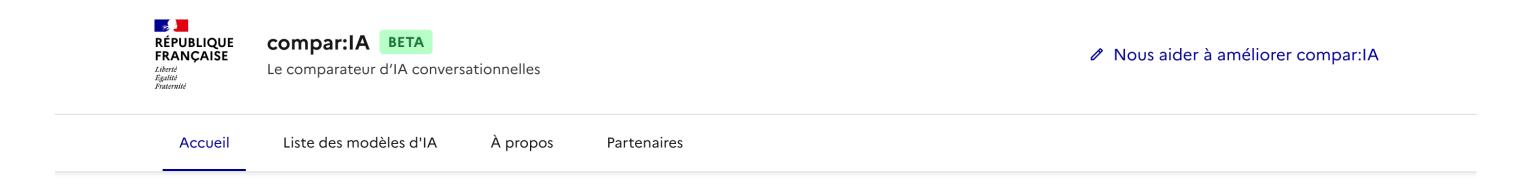


284 616

Source: Energy and Policy Considerations for Deep Learning in NLP, University of Massachusetts Amherst

https://arxiv.org/pdf/1906.02243 cité par https://www.lapresse.ca/affaires/economie/20a23-06-03/intelligenceartificielle/un-impact-environnemental-monstre.php

ComparlA pour comparer les réponses de plusieurs lA et leur impact envirronemental

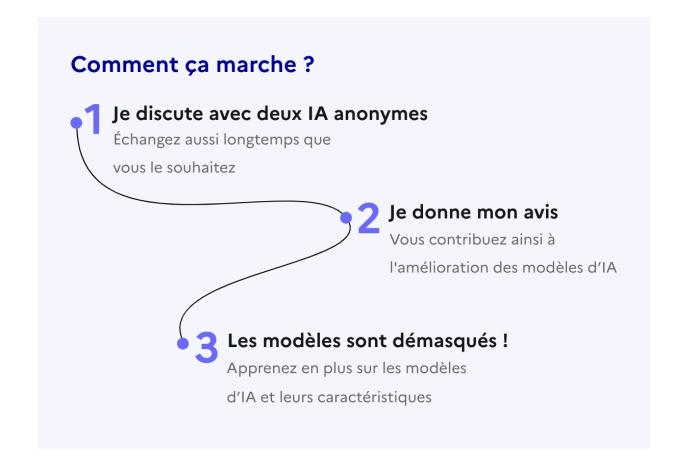




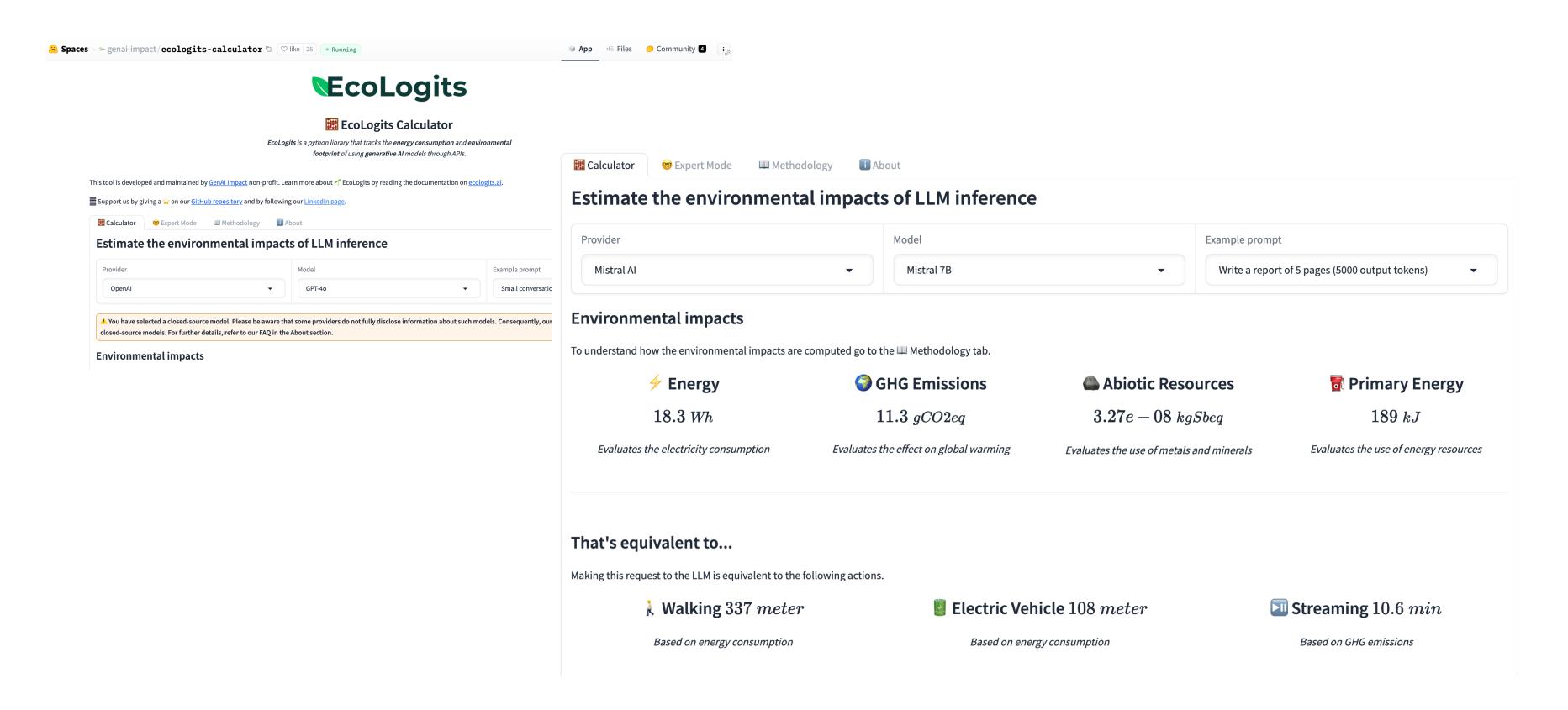
Discutez avec deux IA à l'aveugle pour croiser leurs réponses

J'accepte les modalités d'utilisation 2
Les données sont partagées à des fins de recherche

Commencer à discuter



Comprendre l'impact envirronemental huggingface.co/spaces/genai-impact/ecologits-calculator



La différence entre Google et ChatGPT?

« Quand vous posez une question à Google, Google donne des sources sans la réponse. »

« Quand vous posez une question à ChatGPT, ChatGPT donne une réponse sans rien sourcer. »

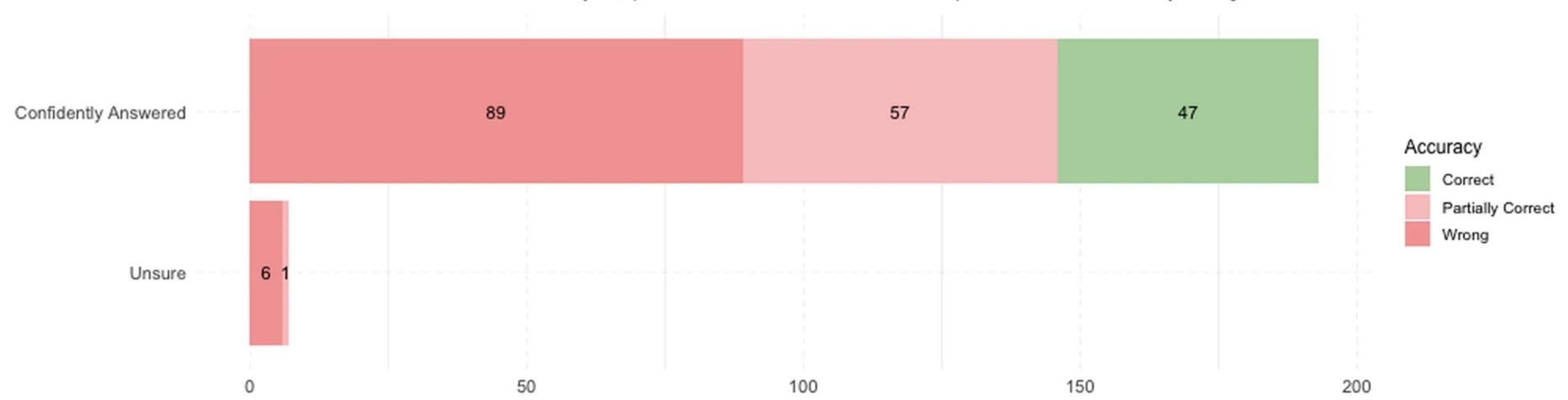


Vidéo au titre évocateur, <u>Thibaut Giraud</u> et <u>Lê Nguyên Hoang</u> (chercheur et médiateur scientifique à l'EPFL)

"Connect people with original, high-quality content from the web" (OpenAI) ... really?

ChatGPT's search feature was frequently wrong, but rarely unsure.





Source: https://www.cjr.org/tow_center/how-chatgpt-misrepresents-publisher-content.php

Diversité et qualité des contenus générés



Quelques points de vigilance

Compréhension limitée

Risque d'erreurs

Usage critique nécessaire

Prompting essentiel

Sécurité des données

Les LLM génèrent du contenu sans comprendre le sens du prompt ni du contenu généré.

Les LLM créent des contenus les plus probables sur base de son apprentissage.

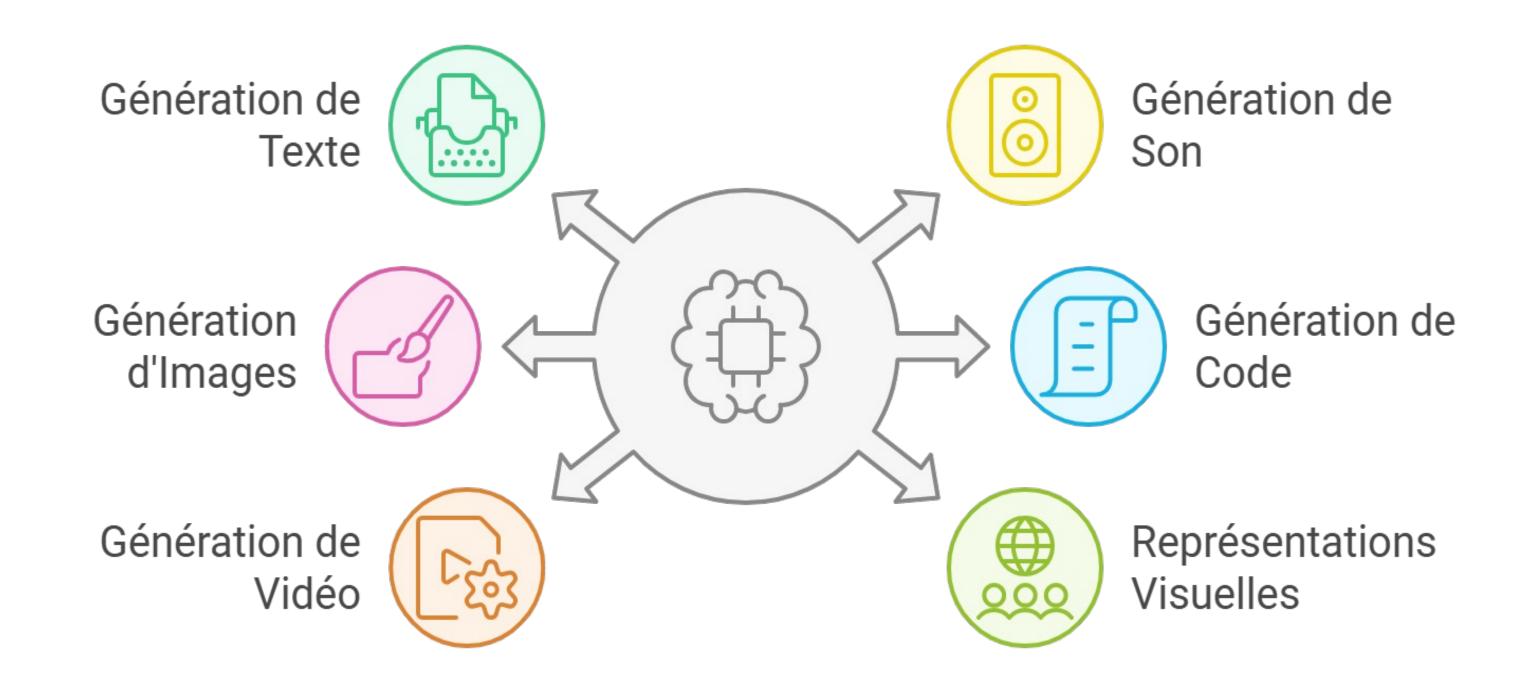
Certains contenus peuvent dès lors être incorrects ou biaisés.

L'utilisation des LLM demande une vérification des contenus générés et une réflexion critique.

La qualité des réponses du LLM dépend de la précision avec laquelle le prompt aura été rédigé.

L'usage du LLM ne garantit pas la sécurité des données, l'utilisateur ou l'utilisatrice reste responsable des données fournies dans ses prompts qui doivent respecter le RGPD et les droits d'auteur.

Diversité et qualité des contenus générés

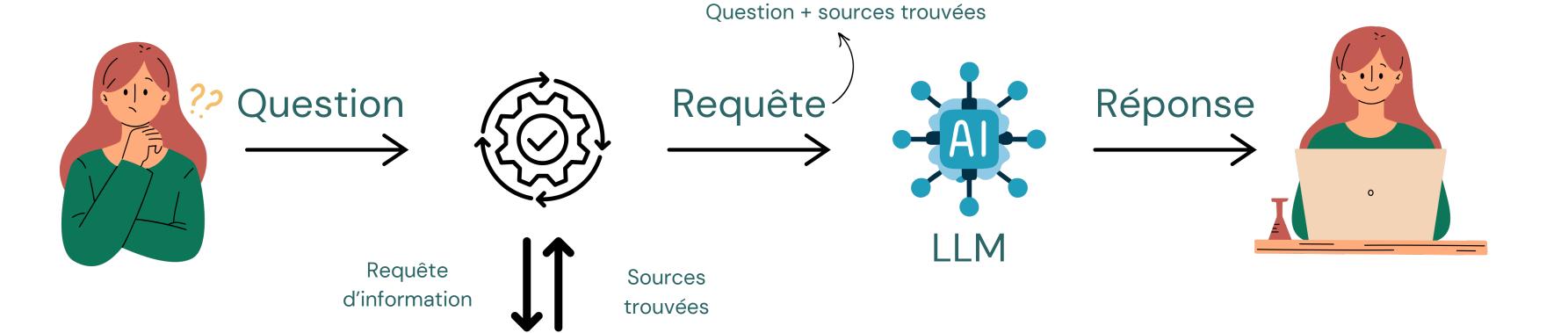


Mais pas uniquement via des chatbots généralistes ...

Les apports du RAG

RAG?

Retrieval augmented generation Génération augmentée par récupération





Base de connaissances



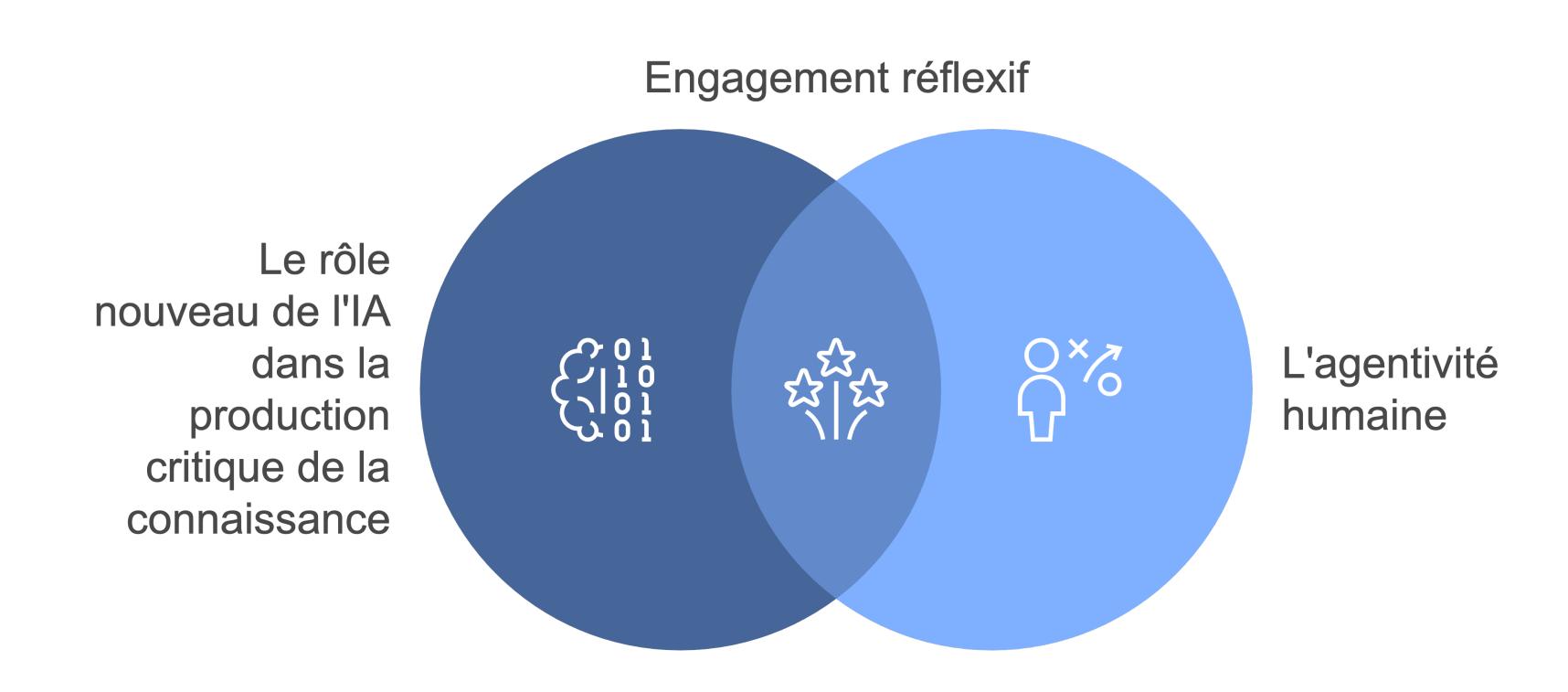
Les apports du RAG

Table 5: Literature Review Tools based on LLMs.

$ \mathbf{ID} $	Tool	Mode	Type	Website
1	Scite [162–164]	Web	Search Engine	https://scite.ai/
2	Elicit [165]	Web	Search Engine	https://elicit.com/
3	Consensus	Web	Search Engine	https://consensus.app/
4	EvidenceHunt	Web	Search Engine	https://evidencehunt.com/
5	MirrorThink	Web	Search Engine	https://mirrorthink.ai/
6	Perplexity	Web/App	Search Engine	https://www.perplexity.ai/
7	Scispace	Web	Search Engine	https://typeset.io/
8	Jenni.ai	Web/App	Writing Assistant	https://jenni.ai/
9	ResearchBuddies	Web	Writing Assistant	https://researchbuddy.app/
10	Silatus	Web	Writing Assistant	https://silatus.com/
11	Textero.ai	Web	Both	https://textero.ai/

Adapté de Warnier, L. et Vangrunderbeeck, P. (2024), d'après Waltzer, T., Cox, R., & Heyman, G. (2023) et Miller, M. (2023)

L'IA modifie notre interaction avec le savoir et son processus de production.

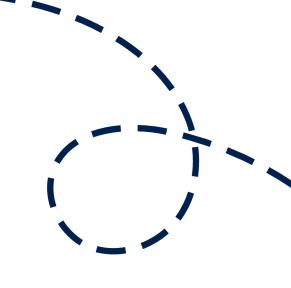


Réassurance et pédagogie

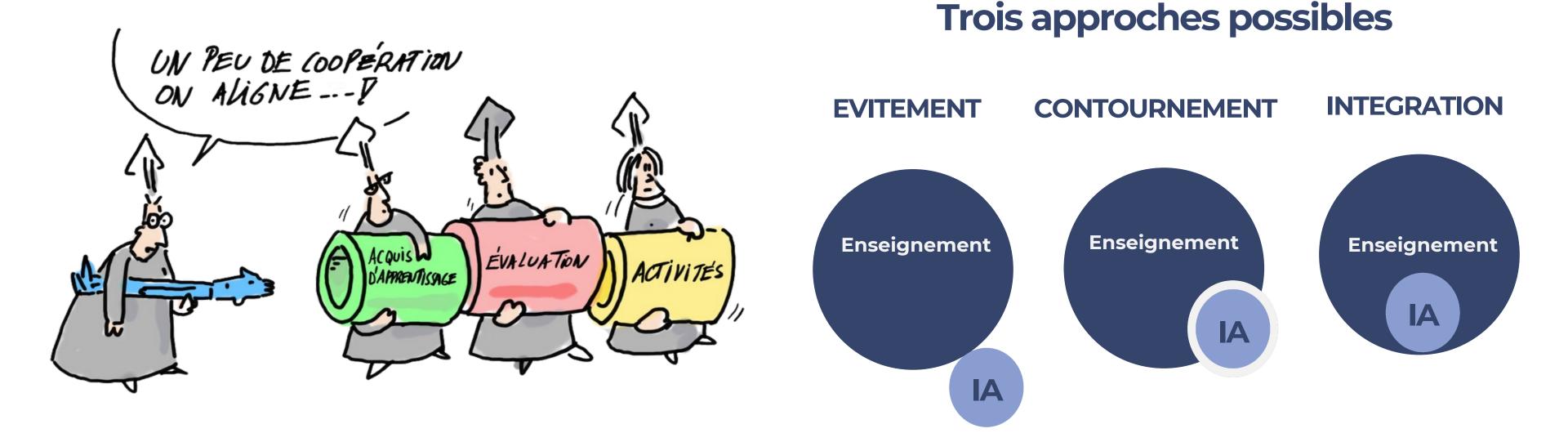
Comment accompagner le changement, une question de posture ?







Une question d'alignement



Options pour évaluer sans l'IA

- Examens en présence avec des limitations de temps.
- Exigences d'écriture manuscrite.
- Questions à développement court ou long, questions vrai/faux, questions à choix multiples, questions d'association ou textes à trous.
- Réalisations pratiques.

Pistes pour adapter les évaluations pour minimiser l'influence de l'IA

- **Contextualiser** les exercices et les questions en utilisant des éléments spécifiques mentionnés en cours.
- Demander une réflexion personnelle, ainsi que des exemples d'expériences vécues.
- Proposer des scénarios ou des études de cas qui exigent une réflexion critique et des solutions nuancées.
- Concevoir des questions qui nécessitent une compréhension approfondie du contexte.
- Proposer des situations complexes où la véracité et la qualité de l'information doivent être examinées, avec des arguments soutenus par une analyse rigoureuse.

Exemples d'activités pour évaluer en tenant compte de l'IA

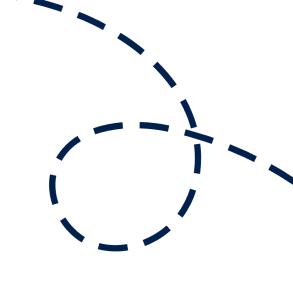
- Dans des contextes concrets d'apprentissage (travaux de terrain, simulations).
- Centrées sur un processus ou une démarche à intégrer (jeux de rôle, résolution de problèmes, forums).
- Évaluation formative et rétroactions pour tenir compte du processus d'apprentissage (projets d'équipe).
- Sur base d'interactions réelles avec des personnes concernées (présentations, entretiens).
- Sur le développement des habiletés métacognitives, incluant une autocritique du travail de l'étudiant (portfolios, appréciation entre pairs, autoévaluation).

Le temps des balises

Repères éthiques et critiques dans l'enseignement et la recherche







Balises et points d'attention à l'UCLouvain

Note institutionnelle

- **Principe de responsabilité** : Les étudiant · es sont entièrement responsables des travaux soumis. Ceux- ci doivent être le fruit d'une démarche personnelle.
- **Principe de transparence** : Là où c'est nécessaire, les étudiant es doivent clairement indiquer les aides et outils utilisés. Ces usages doivent respecter les principes d'intégrité académique.
- **Principe d'authenticité** : Les étudiant · es doivent veiller à ce que leur travail permette une évaluation des connaissances et compétences qu'ils ou elles ont authentiquement acquises. Ce principe vaut tant pour le résultat de leur travail que pour le processus et la méthode suivis pour le réaliser.

En fonction des facultés

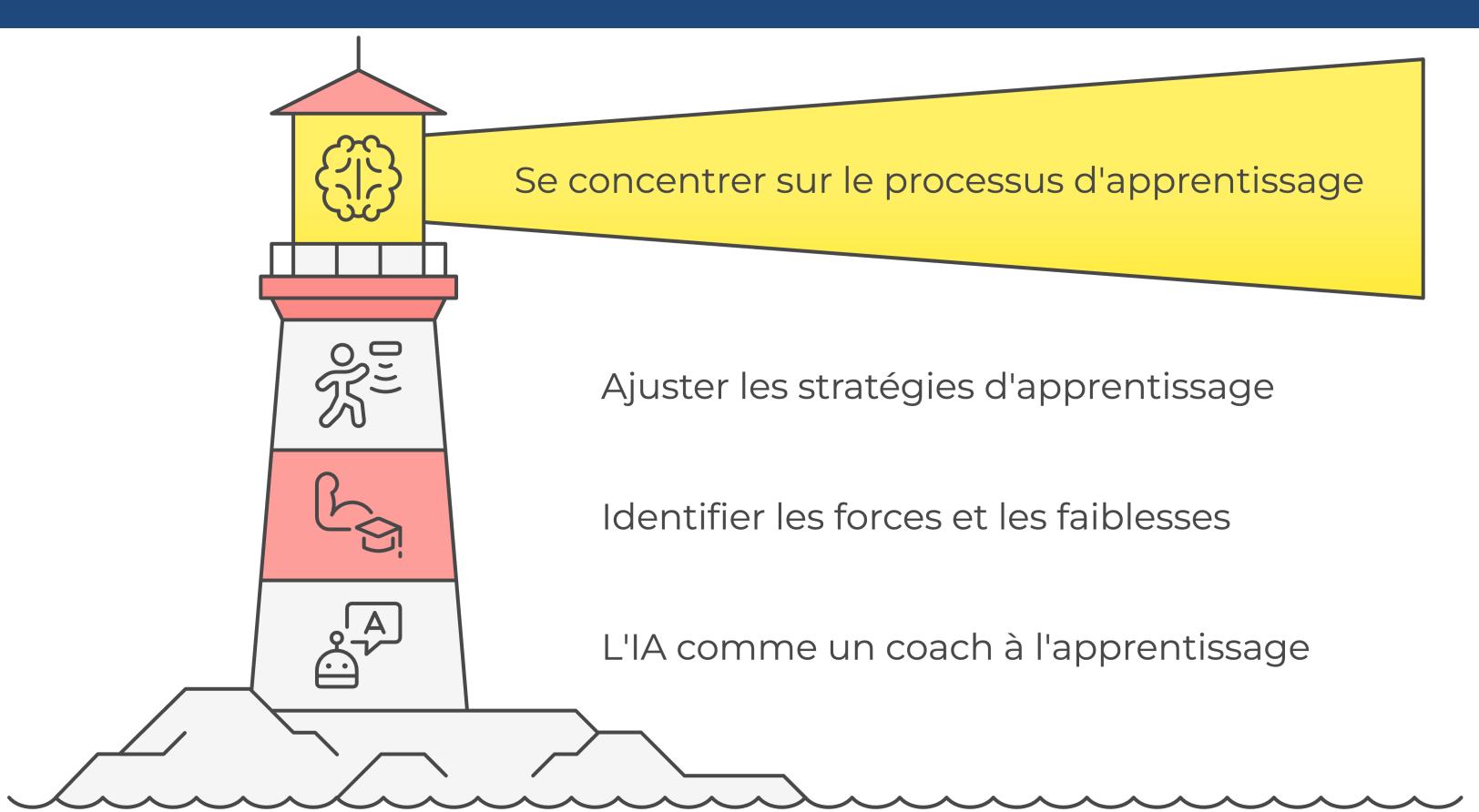
Principe de conservation

Point d'attention

• Respect du matériel protégé par le droit d'auteur, des données personnelles et des informations confidentielles. Ces données et informations ne peuvent être soumises à des outils n'offrant pas de garantie suffisante de confidentialité.

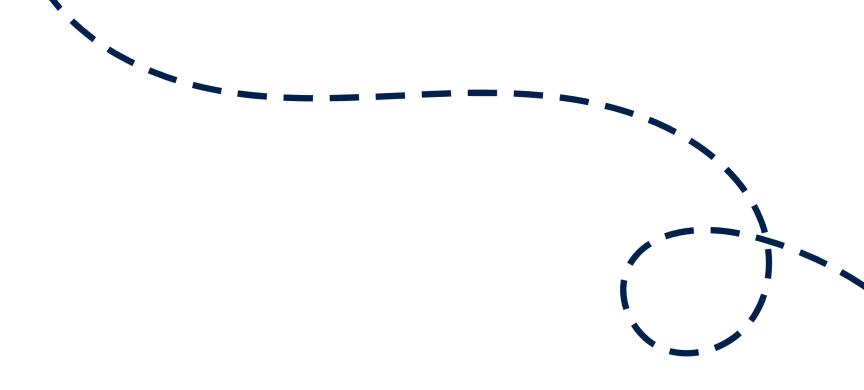
Source: Utilisation Responsable de l'Intelligence Artificielle Générative - http://hdl.handle.net/20.500.12279/1079

Pistes pour exploiter les conservations des interactions avec l'IA et favoriser un travail réflexif



Assurer équité et transparence Former aux usages critiques de l'IA

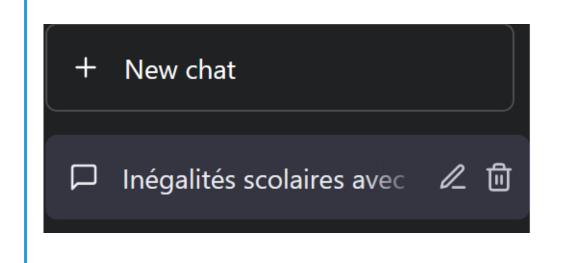






Source : Jansen, Lola. *Les inégalités en matière d'orientation : la réflexivité comme forme de capital culturel spécifique.* Faculté des sciences économiques, sociales, politiques et de communication, Université catholique de Louvain, 2021. Prom. : Draelants, Hugues.

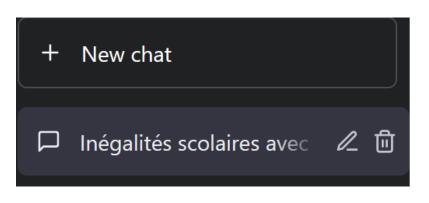
https://dial.uclouvain.be/memoire/ucl/object/thesis:28236





J'aimerais réaliser un mémoire de fin d'études sur la thématique des inégalités observables dans l'orientation scolaire aujourd'hui, en utilisant la théorie du capital social de Bourdieu. Peux-tu me proposer une structure ?





J'aimerais réaliser un mémoire de fin d'études sur la thématique des inégalités observables dans l'orientation scolaire aujourd'hui, en utilisant la théorie du capital social de Bourdieu.

	Chat GPT	Gemini	Perplexity	Scite.Ai
Plan / Structuration / Liens entre concepts				
Rédaction				
Références				

Observations globales:

- Observations
- Eléments distinctifs
- → Un bon compromis pourrait ...

Section	Al tool & URL	Description
Abstract	<u>Graphical Abstract Maker</u>	Simplifies complex research into understandable visuals.
Introduction	<u>Elicit.org</u>	Assists in formulating well-thought research questions during the early phases of a systematic review which can help to brainstorm a research question.
	OpenAl ChatGPT	Refines research questions by providing diverse perspectives and structured approaches, enhancing the brainstorming process.
		Summarizes text data into concise, meaningful abstracts, thus saving time
		Helps in structuring the introduction by generating suggestions based on provided information.
	ForeFront Al Chat	Allows the creation of unique writing personas and writing voices, enhancing the manuscript's style and tone.
	<u>Scite.al</u>	Finds citations related to GPT-generated responses, allowing for the user to find relevant sources to cite.
	<u>Jenni.al</u>	Facilitates the writing process by offering prompts and suggesting text, thereby making it easier to write the rationale.
	Research Rabbit App	Suggests related papers based on Zotero saved ones, promoting thorough research for the rationale.
	<u>Consensus.app</u>	Simplifies the research process by summarizing top papers and consensus from the scientific community, which can be used in the rationale.
	<u>EvidenceHunt</u>	Generates answers with cited articles, facilitating a swift and efficient research process for the rationale.
	<u>LitSuggest</u>	Recommends similar studies based on your research interest using machine learning.
	<u>PaperDigest</u>	Assists in building a solid literature review on any topic, providing a strong foundation for the rationale of the study and existing state of the literature.

Section	Al tool	Description
Methods	<u>Distiller SR</u>	A validated Al tool performs literature screening after training on a proportion of hits
	<u>Xtrct</u>	Employs semantic search to filter for eligibility criteria, thereby enhancing study selection efficiency.
		Finds at least five articles for MeSH terms, facilitating a targeted and specific search strategy.
	<u>WiseOne.io</u>	Ensures information reliability by presenting various sources discussing the same subject.
	<u>SearchSmart</u>	Assists in selecting the best research database, maximizing the relevance and quality of search results.
	<u>CitationChaser</u>	Identifies relevant research topics by tracing citation networks, thus expanding the scope of the search strategy.
	<u>Consensus.app</u>	Searches for relevant literature based on any question, ensuring a thorough search strategy.
	<u>PaperDigest</u>	Aids in constructing a comprehensive literature review, ensuring a solid understanding of the research field for the search strategy.
	<u>Thalia</u>	Enhances the efficiency of search results by allowing specific concept searches in PubMed.
	Covidence	Ranks articles by relevance using an active learning variant, enhancing the efficiency of the selection process.
	<u>RCT Tagger</u>	Estimates the likelihood of PubMed articles being RCTs, thereby distinguishing between RCTs and non-RCTs accurately.
	Rayyan.ai	Uses exclusion criteria to automatically identify articles for exclusion, streamlining the selection process.
	<u>ChatPDF</u>	Extracts data from research papers, though its accuracy should be manually verified due to its unvalidated status.
	RobotReviewer	Expedites the data collection process by automatically extracting data regarding the trial conduct (i.e, the 'PICO', study design and risk of bias) from research papers.
		Provides an automatic bias assessment in RCTs, recommended for semi-automatic use to verify accuracy.
	<u>PDFGear</u>	Extracts text and facilitates interactions with research papers by answering user's queries.
	<u>Humata.ai</u>	Allows users to interactively query about research papers, enhancing comprehension of the content.
	<u>HeyGPT.Chat</u>	Enables interactive conversations to query uploaded research papers.
	OpenAl ChatGPT	Assists in data analysis tasks, for example, generating code for statistical analysis in R or Python.

Source: Fabiano, N., Gupta, A., Bhambra, N., Luu, B., Wong, S., Maaz, M., Fiedorowicz, J. G., Smith, A. L., & Solmi, M. (2024). How to optimize the systematic review process using Al tools. JCPP advances, 4(2), e12234. https://doi.org/10.1002/jcv2.12234

Section	Al tool	Description
Results	<u>HeyGPT.Chat</u>	Interprets uploaded research papers and answers user's specific queries, for instance, about potential biases in studies.
	<u>OpenAl ChatGPT</u>	Assists in data analysis tasks, for example, generating code for statistical analysis in R or Python which can be used in result synthesis.
Discussion	<u>Scite.ai</u>	Finds citations related to GPT-generated responses, allowing for the user to find relevant sources to cite.
	<u>Jenni.ai</u>	Stimulates the writing process by providing prompts and suggesting text, enriching the discussion section.
	OpenAl ChatGPT	Used to generate counterpoints and arguments for the discussion, ensuring a balanced analysis.
	<u>CitationChaser</u>	Deepens the understanding of related topics by tracking citations, providing substance for the discussion.
	<u>Consensus.app</u>	Explores a topic thoroughly by searching for relevant literature to be used as citation sources.
	<u>Scholarcy</u>	Quickly summarizes any manuscript which can assist in the writing of a discussion section.

"Il est généralement accepté au sein de la communauté universitaire que les **détecteurs** d'utilisation d'outils d'IA générative **ne sont pas fiables.**

Il y aura toujours un risque d'accuser à tort un·e étudiant·e, car les erreurs de détection des passages suspects sont inévitables.

- Les **productions** peuvent souvent être **modifiées**, automatiquement ou manuellement, afin de **tromper le détecteur** ;
- L'évolution rapide des outils d'IA générative rend les outils de détection rapidement obsolètes ;
- L'usage d'un détecteur **peut nécessiter de partager** avec une société externe la production de l'étudiant·e, ce qui peut être éthiquement et **juridiquement risqué**.

Les détecteurs d'IA ne peuvent dès lors pas être utilisés pour déterminer avec précision ou donner un résultat définitif quant à la question de savoir si un texte est généré par une IA ou par un humain".

Les outils de détection ne sont pas fiables

■ Bloomberg Subscribe

The Al Race: Startups to Watch | Al Glossary | Model Collapse Risk | Al's Real Carbon F



Moira Olmsted was accused of using AI to write an assignment. Photographer: Nick Oxford/Bloomberg

Businessweek | The Big Take

Al Detectors Falsely Accuse Students of Cheating—With Big Consequences

https://www.bloomberg.com/news/features/2024-10-18/do-ai-detectors-work-students-face-false-cheating-accusations

January 31, 2023

New Al classifier for indicating Al-written text

We're launching a classifier trained to distinguish between Al-written and human-written text.

Our classifier is not fully reliable. In our evaluations on a "challenge set" of English texts, our classifier correctly identifies 26% of Al-written text (true positives) as "likely Al-written," while incorrectly labeling human-written text as Al-written 9% of the time (false positives). Our classifier's reliability typically improves as the length of the input text increases. Compared to our previously released classifier, this new classifier is significantly more reliable on text from more recent Al systems.

https://openai.com/index/new-ai-classifier-for-indicating-ai-written-text/

Traquer la "fake science" et développer l'intégrité académique de tous

Guillaume Cabanac, traqueur de fake science

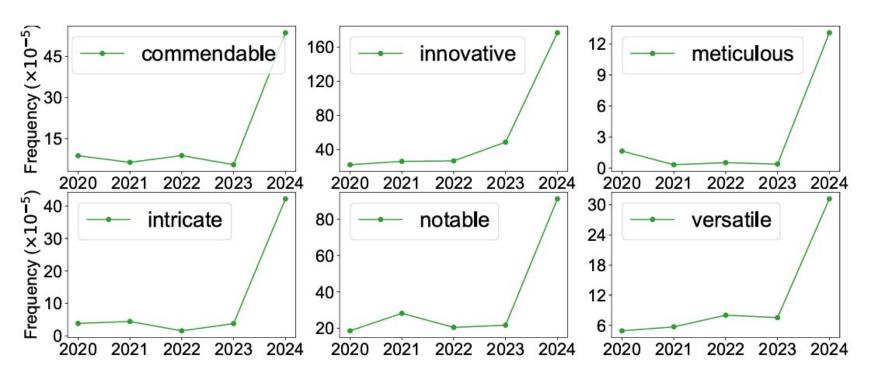
17.02.2022, par Gaël Hautemulle



https://lejournal.cnrs.fr/articles/guillaume-cabanac-traqueur-de-fake-science

Surveillance de contenu modifié par l'IA à grande échelle : une étude de cas sur l'impact de ChatGPT sur les évaluations par les pairs de conférences en IA

Environ 6,5 % à 16,9 % des évaluations par les pairs pourraient avoir été substantiellement modifiées par des LLM



Les auteurs constatent un changement significatif avec des adjectifs tels que « louable », « méticuleux » et « complexe » montrant des augmentations de probabilité dans la survenue dans une phrase.

https://arxiv.org/abs/2403.07183

Traquer la "fake science" et développer l'intégrité académique

LES PRINCIPALES CAUSES DU PLAGIAT (Audet, 2011)



https://pedagogie.uquebec.ca/le-tableau/lintegrite-academique-ca-sapprend

Une approche de prevention du plagiat nuancée

Pour les étudiants peu informés sur le plagiat

Il est crucial de renforcer la **communication** sur les risques et les conséquences du plagiat

Pour les étudiants confiants

Il est important de les amener à **s'autotester** et à objectiver leurs pratiques par rapport aux normes de l'intégrité académique

Pour les étudiants prudents

L'accent doit être mis sur le développement des compétences en matière de référencement et sur l'utilisation des outils d'aide à la rédaction Pour les étudiants moins équipés numériquement

Un accompagnement spécifique peut être nécessaire pour leur permettre d'accéder aux outils et aux ressources nécessaires à la production de travaux originaux

Cougnon, L.-A. Dresselaers, C., Jacques, J., Fastrez, P. (2024). PUPP - Partenariat universitaire sur la prévention du plagiat - 2024 - Bilan en Belgique francophone. Rapport de recherche, UCLouvain. https://doi.org/10.17605/OSF.IO/2UH98

Le cas Wikipedia



WIKIPROJECT AI CLEANUP

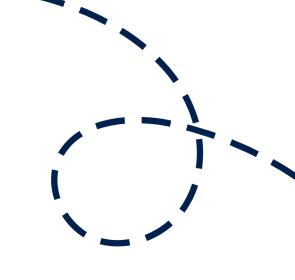
Objectif: suivre l'évolution des usages d'intelligence artificielle, essentiellement IA génératives à partir <u>d'essais de règlementation</u> pour un <u>contrôle des articles concernés</u> au niveau texte et images. Il est lié au projet anglophone <u>WikiProject Al Cleanup</u>.

Du deuil de la détection à la refonte des évaluations

Se concentrer sur les compétences humaines







Dans l'éducation, L'IA sera-t-elle l'acronyme de ...







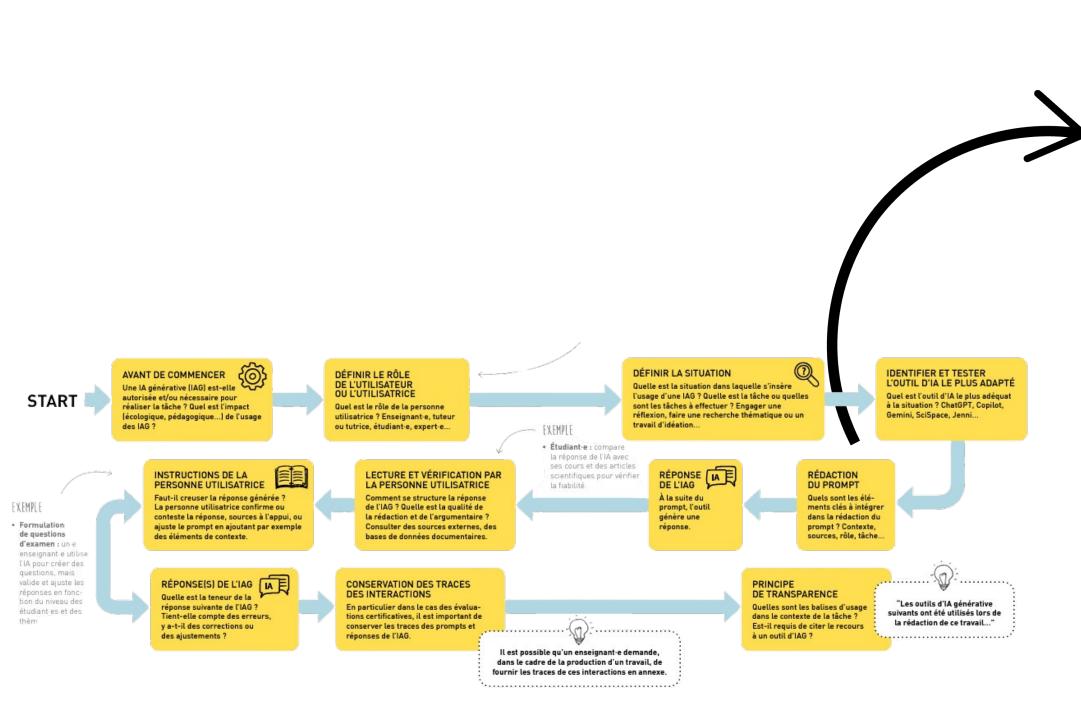
Intelligence Augmentée

Intelligence Appauvrie

Favorise l'engagement dans l'apprentissage

Conduit à la dépendance et à un effort réduit

Des interactions de qualité et critiques de l'IA

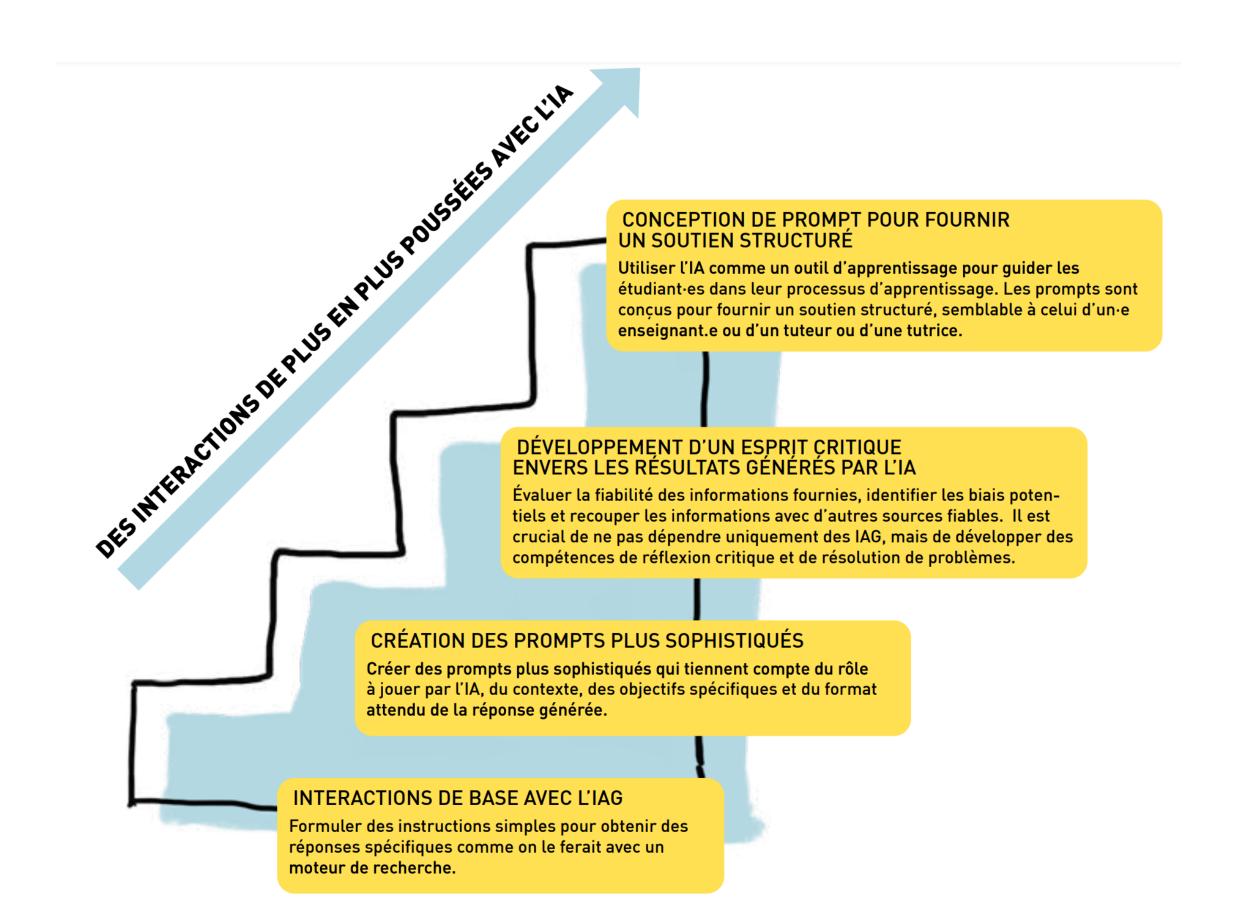


RÉDACTION DU PROMPT

Quels sont les éléments clés à intégrer dans la rédaction du prompt ? Contexte, sources, rôle, tâche...

Les clés du LLL "Intégrer l'IA générative dans les stratégies pédagogiques"

Pistes pour favoriser un travail réflexif avec l'IA



Scénarii d'usages critiques de l'IA comme une aide au tutorat (Mollick, 2023)

Usage de l'IA	Fonction	Bénéfices pédagogiques	Risques
MENTOR	Guider l'apprenant en donnant des rétroactions.	Les rétroactions fréquentes améliorent l'apprentissage, même si tous les conseils ne sont pas suivis	Ne pas faire preuve d'esprit critique en examinant le retour d'information, qui peut contenir des erreurs
TUTEUR	Accompagner la formation de manière spécifique	L'enseignement personnalisé est très efficace	Connaissances insuffisantes de l'IA. Risques d'hallucinations sur les formes de tutorat
COACH	Encourager l'apprenant à réfléchir et à s'améliorer par le biais de la réflexivité	Les occasions de réflexions et de régulations permettent d'améliorer les résultats de l'apprentissage	Le type d'accompagnement peut ne pas correspondre. Risques de conseils erronés
MEMBRE DE L'ÉQUIPE	Augmenter la performance de l'équipe	Proposer d'autres points de vue, aider les équipes apprenantes dans leur tâche d'apprentissage	Risque d'hallucination et d'erreurs. Risque de conflits de « personnalité » avec les autres membres de l'équipe
ÉTUDIANT	Se placer dans le rôle d'apprenant pour renforcer l'apprentissage par l'enseignement aux autres.	Enseigner aux autres est une puissante technique d'apprentissage	Les hallucinations et le raisonnement erronés peuvent réduire à néant les avantages de l'enseignement
SYSTÈME DE SIMULATION	Fournir des scénarios ou des exercices pratiques pour l'entraînement	Pratiquer et appliquer les connaissances favorise le transfert d'apprentissage	Risque d'hallucination et génération inappropriée des données de simulation
OUTIL	Faciliter l'accomplissement de tâches spécifiques	Aide les étudiants à accomplir plus de tâches dans le même temps imparti	Déléguer la réflexion à l'IA plutôt que de s'engager activement et de manière réflexive dans l'activité

Traduit par Pascal Vangrunderbeeck d'après Mollick, E., & Mollick, L. (2023). Assigning AI: Seven Approaches for Students, with Prompts. https://arxiv.org/abs/2306.10052

Prendre conscience des limites de la correction avec l'IA



Propriété intellectuelle

Le travail de l'étudiant constitue sa propriété intellectuelle. Sauf si une autorisation explicite a été obtenue, il est préférable de ne pas l'utiliser avec l'IA.



Manque de contexte concernant l'étudiant

L'outil d'IA ne connaît pas vos étudiants aussi bien que vous. Il risque de donner des résultats génériques et ne pas pouvoir évaluer aussi bien leur apprentissage.



Confidentialité

Les données personnelles de l'étudiant, telles que son nom, son numéro d'identification ou toute autre information sensible, ne doivent jamais être divulguées.



Impersonnel

Un outil d'IA peut produire des résultats convenus, mais vous avez la capacité de personnaliser vos commentaires d'évaluation et de progression.



Sécurité des données

Le contenu partagé avec un outil d'IA, en fonction de ses paramètres, risque d'être intégré à sa base de données et utilisé pour améliorer son entraînement.



Intégrité académique

Les personnes enseignantes doivent être des modèles de comportement éthique, notamment en assurant une évaluation transparente, juste et équitable.



Biais

Les outils d'IA sont connus pour être sujets à des biais. Les résultats générés par une application d'IA peuvent donc être biaisés, injustes, voire discriminatoires ou racistes.



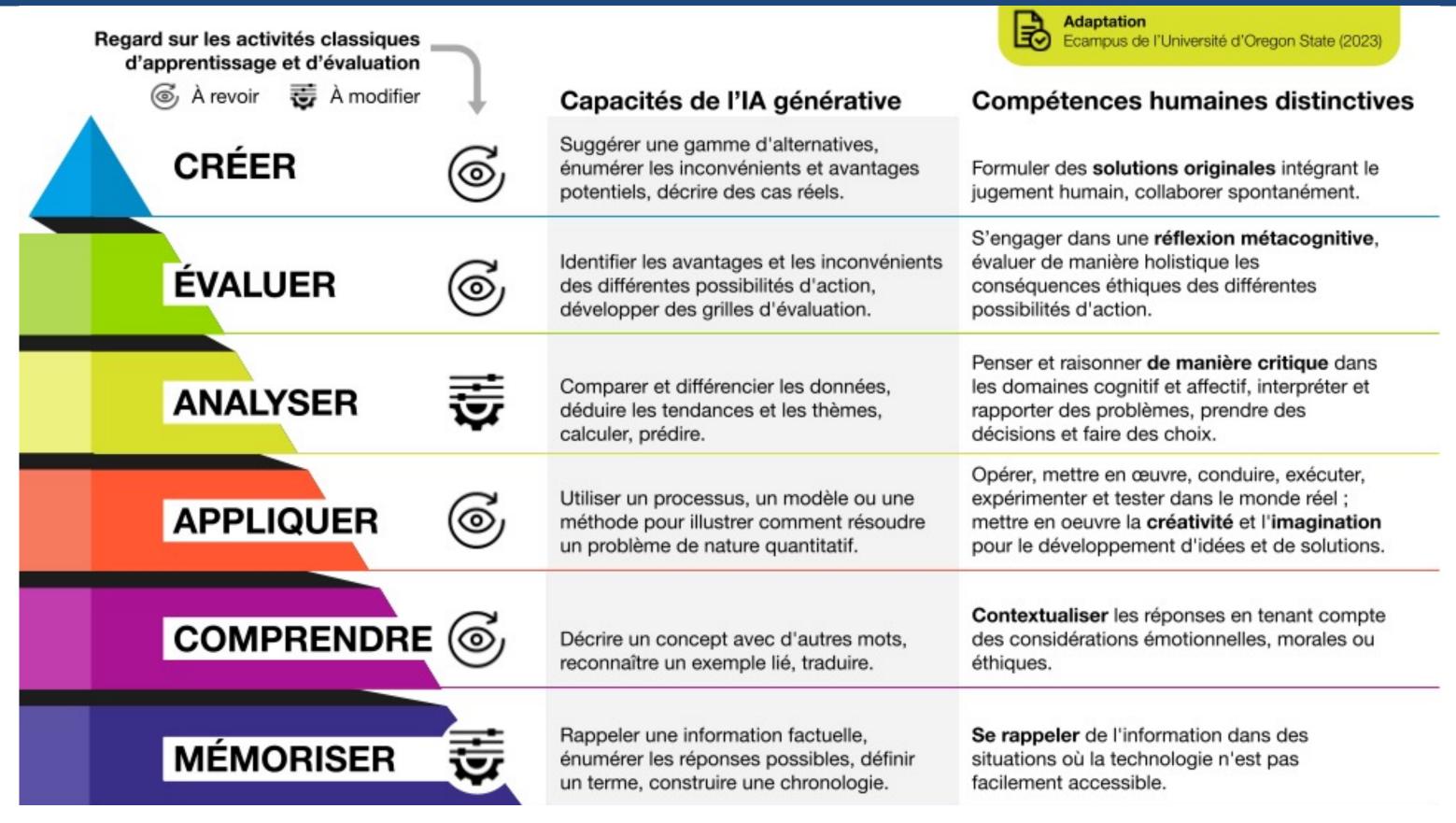
Responsabilités professionnelles

L'évaluation fait partie de vos responsabilités. Déléguer complètement cette tâche à une application d'IA pourrait compromettre l'exercice de votre jugement professionnel.

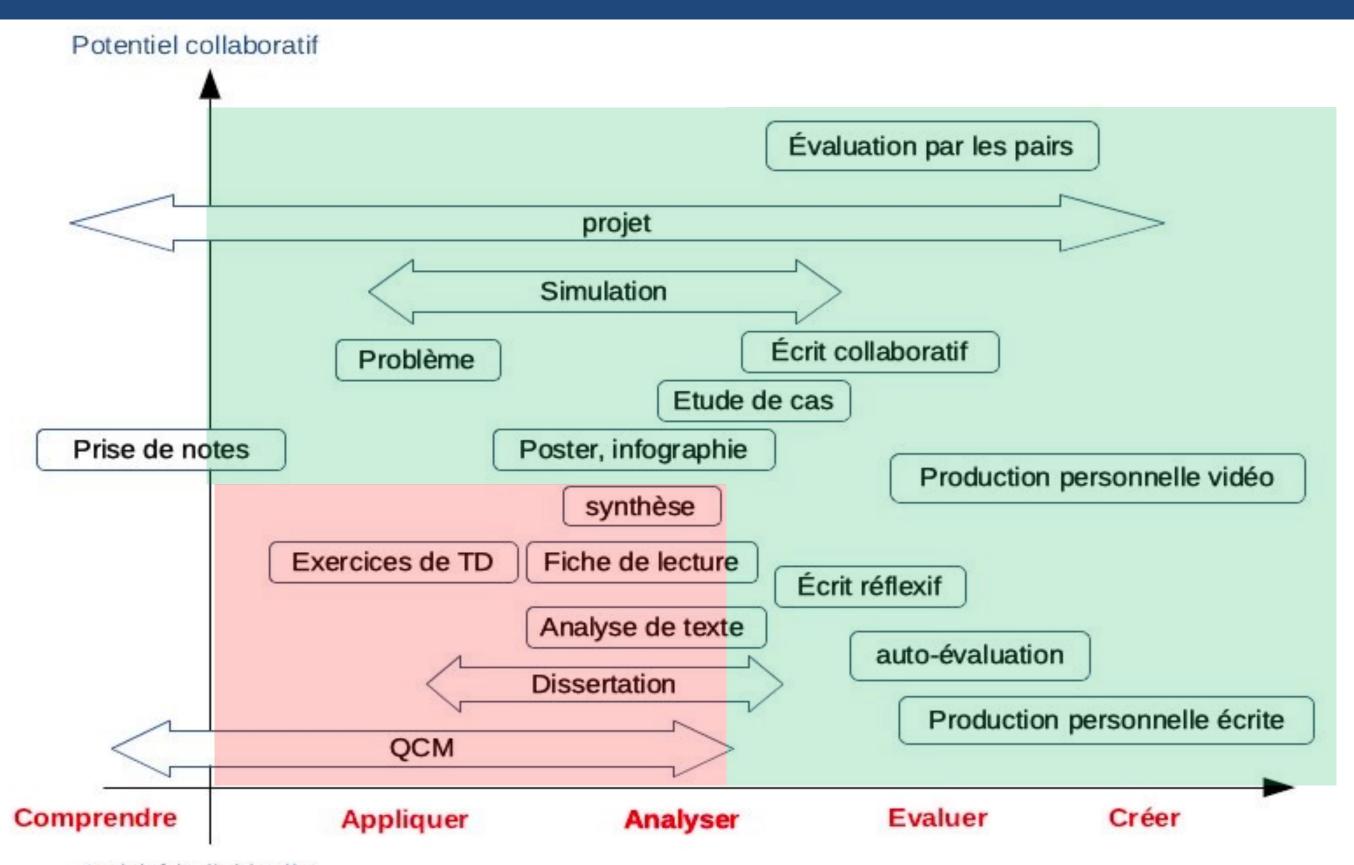


Munn, Y. (2024). Pourquoi le corps enseignant devrait éviter d'utiliser des outils d'IA pour l'évaluation de l'apprentissage, Le Carrefour UQAM (enseigner.uqam.ca), adapté et traduit de "Why Educators Should Avoid Using Al Apps to Help with Assessment" by Sarah Elaine Eaton, PhD, Werklund School of Education at the University of Calgary, CC BY-NC-SA 4.0

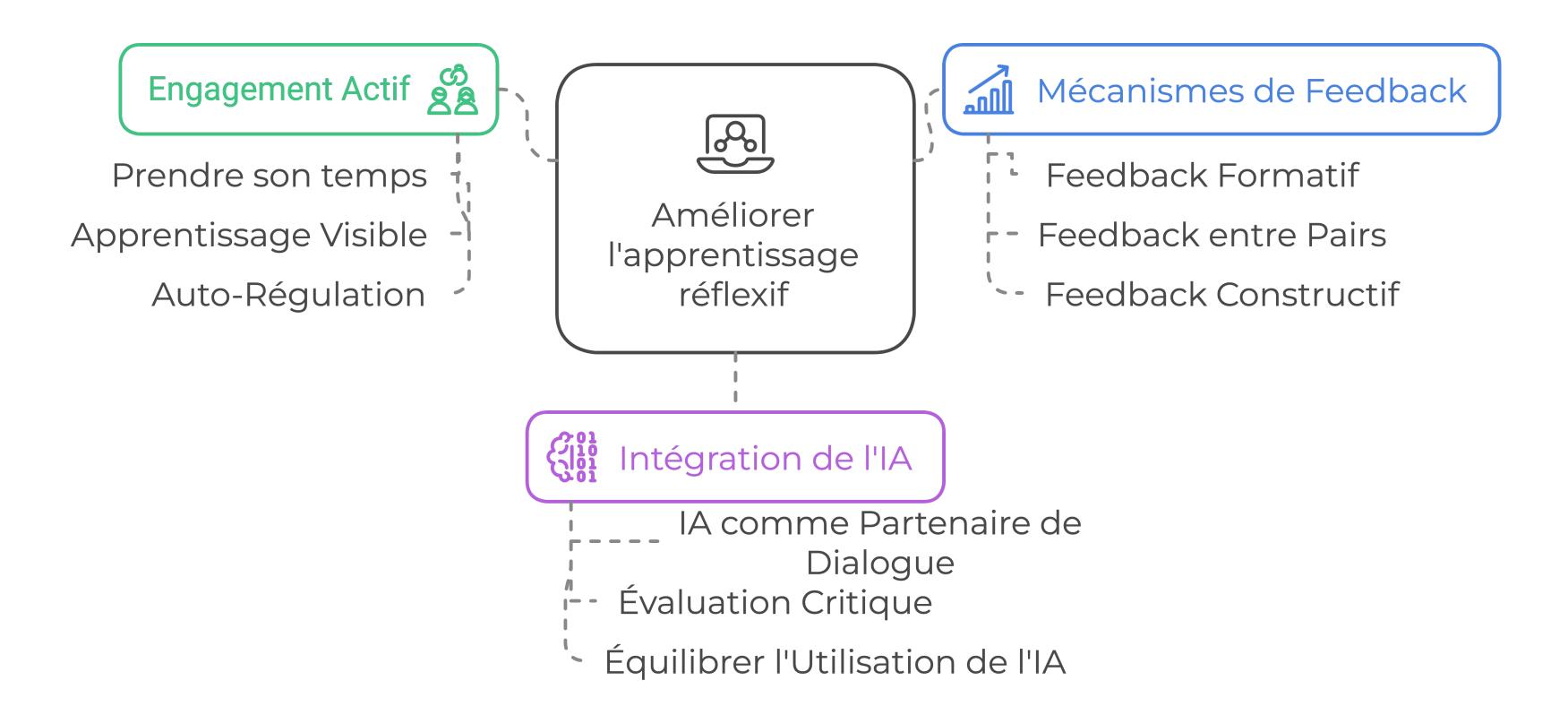
Evaluer les compétences humaines distinctives



Adapter les évaluations à l'ère de l'IA



Pistes pour mettre l'accent sur le processus d'apprentissage





MERCI POUR VOTRE ATTENTION

pascal.vangrunderbeeck@uclouvain.be

