



Si ChatGPT est la réponse, quelle est la question ?

**Repenser l'enseignement universitaire à l'ère
de l'intelligence artificielle générative et des modèles de langage.**

Pr. Germain Forestier

Université de Haute-Alsace / IRIMAS / ENSISA / Projet DemoUHA

Décembre 2024

Journée Pro3 - Orléans

<https://germain-forestier.info/orleans.pdf>

Plan

1. Introduction : Qu'est-ce qu'un modèle de langage ?
2. Comment fonctionnent ces modèles ?
3. Utilisation dans un contexte universitaire
4. Prompt engineering et usage diversifié
5. Challenges et préoccupations
6. Personnalisation des Modèles de Langages
7. Conclusion

Plan

1. Introduction : Qu'est-ce qu'un modèle de langage ?
2. Comment fonctionnent ces modèles ?
3. Utilisation dans un contexte universitaire
4. Prompt engineering et usage diversifié
5. Challenges et préoccupations
6. Personnalisation des Modèles de Langages
7. Conclusion

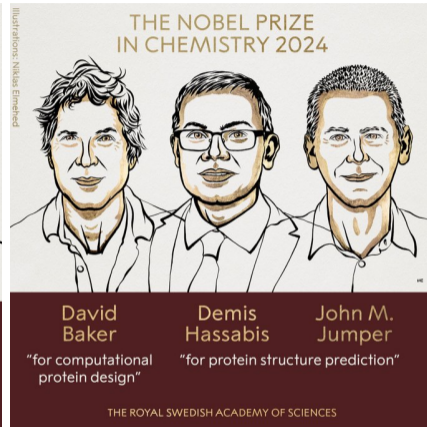
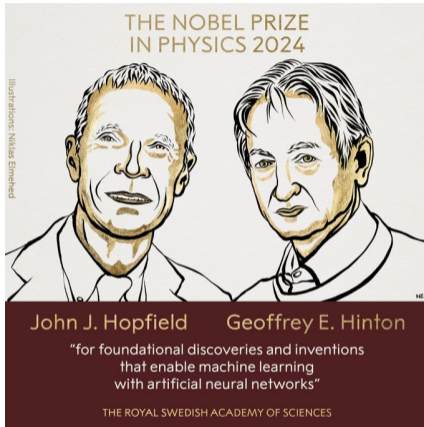
Modèle de langage :

Les modèles de langage sont des systèmes d'intelligence artificielle permettant à une machine de comprendre, générer, ou transformer du texte en langage naturel.

LLM : Large Language Models (terme de plus en plus discuté avec les capacités multimodales des modèles (image, son, vidéo))

Les LLM font partie des IA génératives qui font partie des IA.

Définition



L'intelligence artificielle représente des avancées majeures pour la science dépassant largement le cadre de simples outils informatiques (et les modèles de langage).

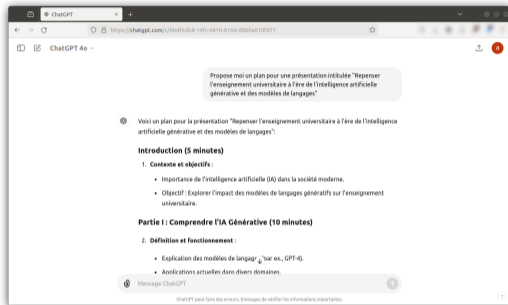
source : <https://www.nobelprize.org/>

Brève histoire et évolution (OpenAI)

- ▶ 2015 : Création d'OpenAI. Démonstrations avec des jeux Atari.
- ▶ 2017 : Google Brain "Attention Is All You Need"
- ▶ 2018 : OpenAI sort GPT-1
- ▶ 2019 : OpenAI sort GPT-2
- ▶ 2020 : OpenAI sort GPT-3
- ▶ 2022 : "Large Language Models are Zero-Shot Reasoners"
- ▶ 2022 : **30 Novembre : OpenAI sort ChatGPT**
- ▶ 2023/2024 : LLaMA, Mistral, Claude, Gemini, LLaMA 2/3, Phi, Qwen, Nova, etc.
- ▶ 2024 : GPT-4o (GPT-4 Omni) / Claude 3.5 Sonnet

<https://simonwillison.net/2023/Aug/3/weird-world-of-llms/>

Le plus connu : ChatGPT



<https://chat.openai.com/>

Juillet 2024 :

- Les IA génératives gagnent en notoriété auprès du grand public (78% en ont déjà entendu parler)
- 48% de ceux qui les utilisent dans ce contexte considèrent que leur entreprise les encourage dans cette lancée
- 79% des Français de 18 ans et plus déclarent être inquiets vis-à-vis de l'émergence des IA génératives

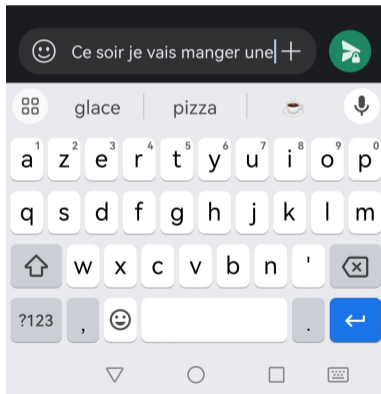
<https://www.ifop.com/publication/2eme-vague-du-barometre-sur-la-perception-et-lusage-des-ia-generatives-par-les-francais/> (juillet 2024)

Plan

1. Introduction : Qu'est-ce qu'un modèle de langage ?
- 2. Comment fonctionnent ces modèles ?**
3. Utilisation dans un contexte universitaire
4. Prompt engineering et usage diversifié
5. Challenges et préoccupations
6. Personnalisation des Modèles de Langages
7. Conclusion

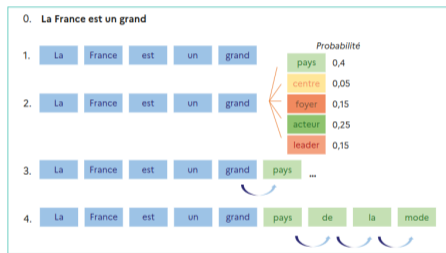
Comment fonctionnent ces modèles ?

Ces modèles apprennent à partir de vastes quantités de texte pour prédire le mot suivant dans une séquence donnée, capturant ainsi la structure et les nuances de la langue :



Comment fonctionnent ces modèles ?

- ▶ **Réseaux de neurones:** Modèles composés de "neurones" et de couches.
- ▶ **Mécanisme d'attention:** Le modèle évalue l'importance des différentes parties du texte.

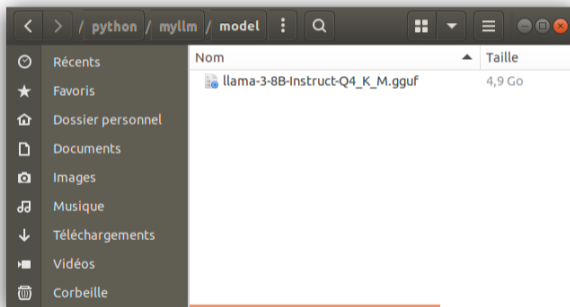


Au-delà de ce mécanisme, de nombreux travaux de recherche se poursuivent pour améliorer les performances en utilisant d'autres techniques avancées.

<https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/> / Rapport "25 recommandations pour l'IA en France".
<https://www.technologyreview.com/2024/03/04/1089403/large-language-models-amazing-but-nobody-knows-why/>

A quoi ressemble un LLM "en vrai" ?

C'est un "gros" fichier binaire contenant le modèle de langage, c'est à dire les paramètres "appris" (les poids) sur un ensemble de données fourni en entrée :



<https://huggingface.co/bartowski/Meta-Llama-3-8B-Instruct-GGUF>

<https://ai.meta.com/llama/>

Code Python pour exécuter le modèle

```
# load the large language model file
from llama_cpp import Llama
LLM = Llama(model_path="./llama-3-8B-Instruct-Q4_K_M.gguf")

# create a text prompt
prompt = "Write a short history of France"

# generate a response (takes several seconds)
output = LLM(prompt)

# display the response
print(output["choices"][0]["text"])
```

France, known as the land of love, art, and cuisine, has a rich and diverse history that spans over two millennia. The country's capital, Paris, is one of the most famous cities in the world and is often referred to as the City of Light. The earliest recorded human presence in what is now France dates back to the Paleolithic era, around 400,000 years ago. The Gauls, a Celtic people, inhabited the region from around 500 BCE until their defeat by the Romans in 52 BCE. The Roman occupation brought significant cultural, economic, and architectural changes to the region. After the fall of the Roman Empire, the Franks, a Germanic tribe, established their kingdom in what is now northern France. In 486 CE, Clovis I united the Frankish tribes under a single ruler, creating the Frankish Empire.

Code Python pour exécuter le modèle

```
# load the large language model file
from llama_cpp import Llama
LLM = Llama(model_path="./mistral-7b-v0.1.Q4_K.gguf")

# create a text prompt
prompt = "Comment expliquer l'intelligence artificielle a un enfant ?"

# generate a response (takes several seconds)
output = LLM(prompt)

# display the response
print(output["choices"][0]["text"])
```

L'intelligence artificielle (IA) est la branche de l'informatique qui vise à créer des systèmes ou des programmes qui fonctionnent en imitant les fonctions mentales d'une personne, tel que le raisonnement ou l'apprentissage. *Il est difficile pour les humains d'être sûrs d'obtenir des résultats plus efficaces que les machines*, et même d'obtenir une solution (que ce soit en mathématique, en physique, en chimie, en biologie, etc.) ou encore d'un problème, plus efficace que celle proposée par les systèmes d'IA.

Code Python pour exécuter le modèle

```
# load the large language model file
from llama_cpp import Llama
LLM = Llama(model_path="./mistral-7b-v0.1.Q4_K.gguf")

# create a text prompt
prompt = "Genere moi un poeme en Francais sur l'Universite d'Orleans."

# generate a response (takes several seconds)
output = LLM(prompt)

# display the response
print(output["choices"][0]["text"])
```

*Les rues de Loire, majestueuses,
S'élèvent vers le ciel éternel.
Au cœur de la ville, un bastion,
L'université d'Orléans, berceau de
savoir.*

*Dans ses bâtiments, vieilles pierres,
Des siècles d'histoire se lissent.
Les couloirs, aux murs peints de
rouge,
Accueillent les étudiants, étoiles
neuves.*

*Les salles de cours, émerveillées,
Sont des lieux de découverte, où
s'égraine
La curiosité, la créativité.
Les enseignants, guides inspirés,
Favorisent l'apprentissage, le progrès.
Les bibliothèques, temples du savoir,*

*Où les livres, trésors inestimables,
Attendent de révéler leurs secrets.
Les chercheurs, aventuriers de
l'esprit,*

*Orléans, ville aux mille visages,
Aujourd'hui, berceau d'études
florissantes.
L'université, coeur battant,
Fait vibrer la ville, avec son
dynamisme.*

*Je vous laisse, à l'Université
d'Orléans,
Où le savoir et la passion s'allient,
Pour créer un avenir, riche en
promesses,
Et une vie, étoilée de lumière.*

Illustration : Comment les LLMs sont entraînés

Texte complet :

La science (du latin scientia, « connaissance », « savoir ») est dans son sens premier « la somme des connaissances » et plus spécifiquement une entreprise systématique de construction et d'organisation des connaissances sous la forme d'explications et de prédictions testables. Faisant suite à la technique au niveau de son histoire, elle se développe en Occident au travers de travaux à caractère universel basés sur des faits, une argumentation et des méthodes qui varient selon qu'elles tiennent de l'observation, l'expérience, l'hypothèse, d'une logique de déduction ou d'induction, etc.

source : <https://fr.wikipedia.org/wiki/Science>

Exercice de complétion :

Mots à trouver : La science (du latin scientia, « connaissance », « savoir ») est dans son sens premier « la _____ des connaissances » et plus _____ une entreprise systématique de _____ et d'organisation des connaissances sous la forme d'_____ et de prédictions testables.

Phrase à compléter : Faisant suite à la technique au niveau de son histoire, elle se développe _____

Mieux comprendre comment les LLM fonctionnent

The screenshot displays the 'LLM Visualization' website interface. The browser address bar shows the URL <https://bbycroft.net/llm>. The page title is 'LLM Visualization' and the current chapter is 'Chapter: Overview'. The interface includes a navigation menu on the left with a 'Table of Contents' listing sections like 'Intro', 'Introduction', 'Preliminaries', 'Components', 'Embedding', 'Layer Norm', 'Self Attention', 'Projection', 'MLP', 'Transformer', 'Softmax', and 'Output'. The main content area features a flowchart of the GPT architecture, starting with 'How to predict text' (tokens: 3437, 284, 4327) and 'test tokens' (tokens: 46245, words: 3638). The flowchart shows the process from 'tok embed' and 'pos embed' through 'transformer i' (containing 'layer norm', 'multi-head, causal self-attention', 'layer norm', 'feed forward', 'layer norm') to 'linear' and 'softmax' layers. A 'Table of Contents' sidebar is visible on the right. At the bottom, a welcome message reads: 'Welcome to the walkthrough of the GPT large language model! Here we'll explore the model nano-gpt, with a mere 85,000'. Below this message are 'Continue' and 'skip' buttons. The right side of the page shows model selection tabs for 'GPT-2 (small)', 'nano-gpt', 'GPT-2 (XL)', and 'GPT-3', with 'tokens = 85,584' displayed below. A 3D visualization of the model stack is shown on the right side of the page.

LLM Visualization

<https://bbycroft.net/llm>

Mieux comprendre comment les LLM fonctionnent

The screenshot shows a web browser window titled "LLM Visualization" with the URL <https://bbycroft.net/llm>. The page is titled "Chapter: Self Attention" and contains the following text:

The self-attention layer is perhaps the heart of the Transformer and of GPT. It's the phase where the columns in our input embedding matrix "talk" to each other. Up until now, and in all other phases, the columns can be regarded independently.

The self-attention layer is made up of several heads, and we'll focus on one of them for now.

The first step is to calculate the Q, K, and V vectors for each of the columns from the input embedding matrix. These vectors are the Q, K, and V vectors:

- Q: Query vector
- K: Key vector
- V: Value vector

To produce one of these vectors, we perform a matrix-vector multiplication with a bias added. Each output cell is some linear combination of the input vector. E.g. for the Q vectors, this is done with a dot product between a row of the Q-weight matrix and a column of the input matrix.

The dot product operation, which we'll see a lot of, is quite simple: We pair each element from the first vector with the corresponding element from the second vector, multiply the pairs together and then add the results up.

Navigation buttons: Continue, skip

Model selection: GPT-2 (small), nano-gpt, GPT-2 (XL), GPT-3

Diagram illustrating the self-attention process:

- Input: Q Bias, K Bias, V Bias
- Q Bias and K Bias feed into Q Weights and K Weights respectively.
- V Bias feeds into V Weights.
- Q Weights, K Weights, and V Weights feed into Q vectors, K vectors, and V vectors respectively.
- Q vectors and K vectors feed into Attention Matrix.
- Vectors feed into Attention Matrix.
- Attention Matrix feeds into V Output.
- Vectors feed into Layer Norm.

LLM Visualization

<https://bbycroft.net/llm>

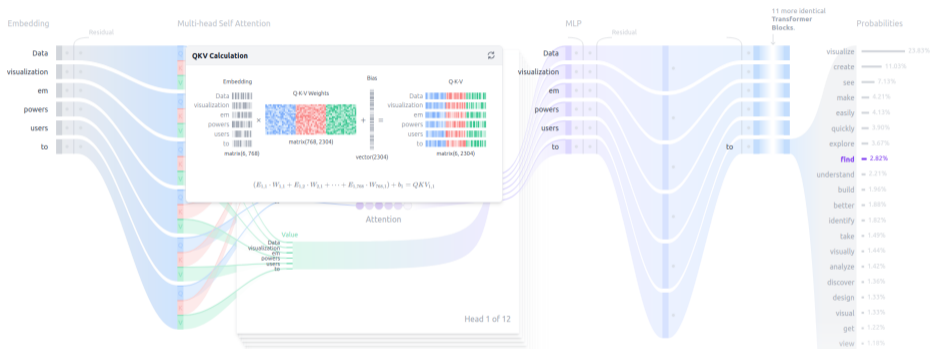
Mieux comprendre comment les LLM fonctionnent

TRANSFORMER EXPLAINER

Examples ▾ Data visualization empowers users to find

Generate

Temperature 1



Transformer Explainer

<https://poloclub.github.io/transformer-explainer/>

Quelles sources de données sont utilisées ?

Dataset	Sampling prop.	Epochs	Disk size
CommonCrawl	67.0%	1.10	3.3 TB
C4	15.0%	1.06	783 GB
Github	4.5%	0.64	328 GB
Wikipedia	4.5%	2.45	83 GB
Books	4.5%	2.23	85 GB
ArXiv	2.5%	1.06	92 GB
StackExchange	2.0%	1.03	78 GB

Table 1: **Pre-training data.** Data mixtures used for pre-training, for each subset we list the sampling proportion, number of epochs performed on the subset when training on 1.4T tokens, and disk size. The pre-training runs on 1T tokens have the same sampling proportion.

→ question des droits d'auteur

Plan

1. Introduction : Qu'est-ce qu'un modèle de langage ?
2. Comment fonctionnent ces modèles ?
- 3. Utilisation dans un contexte universitaire**
4. Prompt engineering et usage diversifié
5. Challenges et préoccupations
6. Personnalisation des Modèles de Langages
7. Conclusion

Article d'OpenAI le 31 Aout 2023 :

"Nous publions un guide à l'intention des enseignants qui utilisent ChatGPT dans leurs salles de classe, incluant des suggestions de questions, une explication du fonctionnement de ChatGPT et de ses limites, l'efficacité des détecteurs d'IA, et les biais."



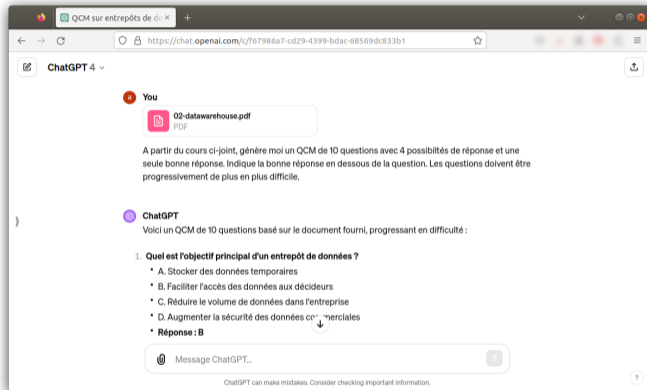
<https://openai.com/blog/teaching-with-ai>

Construire des quiz, des tests et des plans de cours à partir de matériel pédagogique

Fran Bellas, professeur à l'Universidade da Coruña en Espagne, recommande aux enseignants d'utiliser ChatGPT comme assistant pour créer des quiz, des examens et des plans de cours. Il conseille de partager d'abord le programme scolaire avec ChatGPT, puis de lui demander de nouvelles idées de quiz et de plans de cours qui utilisent des exemples modernes ou culturellement pertinents. Bellas se tourne également vers ChatGPT pour aider les enseignants à s'assurer que les questions qu'ils rédigent eux-mêmes sont inclusives et accessibles au niveau d'apprentissage des élèves. "Si vous allez sur ChatGPT et lui demandez de créer des examens de 5 questions sur les circuits électriques, les résultats sont très originaux. Vous pouvez prendre ces idées et les faire vôtres".

D'autres exemples sur <https://openai.com/blog/teaching-with-ai>

Enseignement : Préparation et support de cours avec IA



<https://chat.openai.com/>

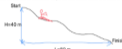
Exemple avec la génération de QCM.

Gemini, le modèle de Google, 2ème page du Tech. Report (sur 63):

Prompt

Here's a solution to a Physics problem by a student.

23. A skier slides down the frictionless slope as shown. What is the skier's speed at the bottom?



Solution:

The conservation of energy means that the energy of the skier is the same at the beginning and at the end of the slope.

Start: $E = mgh$

End: $E = \frac{1}{2}mv^2$

$$\Rightarrow mgh = \frac{1}{2}mv^2$$
$$\Rightarrow v = \sqrt{2gh} = \sqrt{2 \times 9.81 \times 40} = 39.6 \text{ m/s}$$

Try to reason about the question step by step.

Did the student get the correct answer? If the solution is wrong, please explain what is wrong and solve the problem.

Make sure to use LaTeX for math and round off the final answer to two decimal places.

Response

The student did not get the correct answer. The student made a mistake in the calculation of the potential energy at the start of the slope. The potential energy at the start of the slope is given by mgh , where m is the mass of the skier, g is the acceleration due to gravity, and h is the height of the slope. The student used the length of the slope, L , instead of the height, h , in the calculation of the potential energy.

The correct solution is:

The conservation of energy means that the total energy of the skier is the same at the beginning and at the end of the slope.

At the start of the slope, the skier has only potential energy, which is given by mgh .

At the end of the slope, the skier has only kinetic energy, which is given by $\frac{1}{2}mv^2$.

Therefore, we can write:

$$mgh = \frac{1}{2}mv^2$$

Canceling out the mass m on both sides, we get:

$$gh = \frac{1}{2}v^2$$

Solving for v , we get:

$$v = \sqrt{2gh}$$

Substituting the given values, we get:

$$v = \sqrt{2 \times 9.81 \times 40} = 28.01 \text{ m/s}$$

Therefore, the skier's speed at the bottom of the slope is 28.01 m/s.

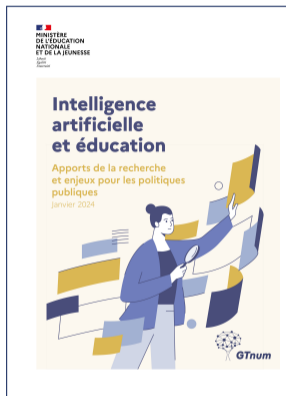
Figure 1 | Verifying a student's solution to a physics problem. The model is able to correctly recognize all of the handwritten content and verify the reasoning. On top of understanding the text in the image, it needs to understand the problem setup and correctly follow instructions to generate LaTeX.

<https://goo.gle/GeminiPaper> (19/12/2023)

Questions clés

- ▶ Comment utiliser l'IA pour la préparation des activités pédagogiques (présentation, quiz, textes, questions, exercices, podcasts, vidéos, etc.) ?
- ▶ Comment les étudiants utilisent-ils l'IA pour répondre à mes exercices, réviser, apprendre, gérer leurs projets, etc. ?
- ▶ Comment évaluer à l'heure de l'IA ? Est-ce que j'évalue l'IA, l'étudiant, ou sa capacité à utiliser l'IA ?
- ▶ Comment l'IA va-t-elle impacter les métiers visés par mes étudiants ?
- ▶ Comment mettre en place une réflexion critique face à ces outils ?

Promesses : automatisation, adaptation, personnalisation, disponibilité



" Intelligence artificielle et éducation : apports de la recherche et enjeux pour les politiques publiques (01/2024)" / Direction du numérique pour l'éducation - Ministère de l'Éducation nationale et de la Jeunesse

<https://edunumrech.hypotheses.org/10764>

Catégorisation proposée par la Commission européenne

Enseignement	Soutien à l'apprenant	Soutien à l'enseignant	Soutien au système éducatif
<ul style="list-style-type: none"> Environnement informatique pour l'apprentissage humain Systèmes tutoriels fondés sur le dialogue Applications d'aide à l'apprentissage des langues 	<ul style="list-style-type: none"> Environnements d'apprentissage exploratoire Évaluation formative des écrits Apprentissage collaboratif assisté par l'IA 	<ul style="list-style-type: none"> Évaluation sommative des écrits, notation Repirage de mots-clés dans les forums d'apprenants Assistants pédagogiques basés sur l'IA Recommandation de ressources pédagogique 	<ul style="list-style-type: none"> Exploration des données éducatives pour l'allocation des ressources Diagnostic des difficultés d'apprentissage Services d'orientation

Drapets (European Commission, Directorate General for Education, Youth, Sport and Culture, 2022)

Autre catégorisation issue des groupes thématiques numériques

- Systèmes adaptatifs et personnalisés (recommandation et sélection de contenus, de ressources), évaluation adaptative et correction automatique.
- Robots conversationnels et systèmes de tutorat intelligent, par ex. pour l'enseignement de la démarche d'investigation et de la démarche scientifique (Cied & Barois, 2019).
- Anticipation précoce, prédiction et prévention de l'abandon scolaire.
- Détection et remédiation aux difficultés d'apprentissage.
- Utilisation de savoirs produits par la recherche avec l'IA (analyse d'image, annotation, classification et résumé) (synthèse automatisées, traitement de données massives sur des corpus textuels ou iconographiques).
- Usages administratifs.

Drapets (DNE-TN2, 2021) (DNE-TN2, 2022) (Lepage, Alexandre, 2022) (Romero et al., 2021)

Taxonomie d'après la revue de littérature de (Holmes et al., 2022)

IA au service des élèves

Systèmes de tutorat intelligents
Applications assistées par l'IA (par exemple, mathématiques, synthèse vocale, apprentissage des langues)
Simulations assistées par l'IA (par exemple, apprentissage par le jeu, réalité virtuelle, réalité augmentée)
IA pour aider les apprenants à besoins éducatifs particuliers
Rédaction automatique d'essais
Agents conversationnels
Évaluation formative automatique
Orchestrateurs de réseaux d'apprentissage
Systèmes de tutorat basés sur le dialogue
Environnements d'apprentissage exploratoire
Assistant d'apprentissage tout au long de la vie assisté par l'IA

IA au service de l'enseignant

Détection de plagiat
Curatation intelligente du matériel d'apprentissage
Surveillance de la classe
Évaluation sommative automatique
IA d'assistance de l'enseignant (y compris assistant d'évaluation)
Orchestration de la salle de classe

IA au service des institutions

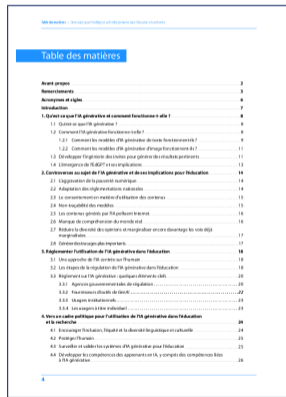
Admissions (par exemple, sélection des élèves)
Planification des cours – Planification des horaires – Programmation horaire
Sécurité des écoles
Identification précoce des décrocheurs et des élèves à risque
e-Proctoring (surveillance des examens à distance)

D'après (Holmes et al., 2022) traduction par (Bocquet, 2023)

” Intelligence artificielle et éducation : apports de la recherche et enjeux pour les politiques publiques (01/2024)” / Direction du numérique pour l'éducation - Ministère de l'Éducation nationale et de la Jeunesse

<https://edunumrech.hypotheses.org/10764>

Orientations pour l'intelligence artificielle générative dans l'éducation et la recherche



<https://www.unesco.org/fr/articles/orientations-pour-lintelligence-artificielle-generative-dans-leducation-et-la-recherche> (30/05/24)

Le rapport thématique n°3 : IA et éducation



**L'IA ET L'AVENIR
DU SERVICE PUBLIC**

RAPPORT THÉMATIQUE #3
IA ET ÉDUCATION

Christian Bruyen et Bernard Fialaire, rapporteurs

- Si le numérique a depuis longtemps fait son entrée à l'école et dans les apprentissages, le rôle croissant joué par l'intelligence artificielle (IA) vient bouleverser les approches de l'éducation en ouvrant considérablement le champ des possibles, notamment depuis le lancement de ChatGPT par Open AI en 2022 et le développement fulgurant des systèmes d'IA générative.

<https://www.senat.fr/rap/r24-101/r24-101-syn.pdf>

<https://www.senat.fr/rap/r23-491/r23-491-syn.pdf>


<https://www.senat.fr/travaux-parlementaires/office-et-delegations/delegation-a-la-prospective/lia-et-lavenir-du-service-public.html>

Le rapport thématique n°3 : IA et éducation

→ Le rapport thématique n° 1 : Impôts, prestations sociales et lutte contre la fraude

→ Le rapport thématique n° 2 : IA et santé

Les rapports thématiques et leurs rapporteurs

IA, impôts, prestations sociales et lutte contre la fraude (rapport publié) 

IA et système de santé (rapport publié) 

IA et éducation 

IA et environnement 

IA, territoires et proximité 

<https://www.senat.fr/rap/r24-101/r24-101-syn.pdf>

<https://www.senat.fr/rap/r23-491/r23-491-syn.pdf>

<https://www.senat.fr/travaux-parlementaires/office-et-delegations/delegation-a-la-prospective/lia-et-lavenir-du-service-public.html>

Plan

1. Introduction : Qu'est-ce qu'un modèle de langage ?
2. Comment fonctionnent ces modèles ?
3. Utilisation dans un contexte universitaire
- 4. Prompt engineering et usage diversifié**
5. Challenges et préoccupations
6. Personnalisation des Modèles de Langages
7. Conclusion

- ▶ **Qu'est-ce que le Prompt Engineering ?**
 - ▶ Technique pour formuler des instructions précises (prompts) afin de guider un modèle de langage vers la réponse souhaitée.
- ▶ **Importance :**
 - ▶ Les modèles comme ChatGPT répondent différemment en fonction du prompt donné.
 - ▶ Permet d'obtenir des réponses plus précises, pertinentes ou dans un format spécifique.
- ▶ **Techniques courantes :**
 - ▶ Spécifier le format de la réponse souhaitée.
 - ▶ Poser la question sous différents angles.
 - ▶ Utiliser un contexte ou un scénario pour encadrer la question.
 - ▶ Donner un rôle au modèle.

Exemples de Prompts pour les Enseignants

Prompt de base :

"Tu es un enseignant clair et pédagogue qui porte une attention particulière à adapter tes exercices en fonction du niveau de tes étudiants."

Applications du Prompt :

- ▶ **Création d'exercices** : "Propose-moi un exercice adapté pour des étudiants débutants en algèbre."
- ▶ **Feedback sur un travail** : "Comment répondrais-tu à un étudiant ayant des difficultés avec un concept précis en physique ?"
- ▶ **Conseils pédagogiques** : "Donne-moi des stratégies pour rendre mon cours sur la biologie plus engageant."
- ▶ **Planification de cours** : "Comment organiserais-tu une leçon sur la littérature romantique pour des étudiants intermédiaires ?"

Exemple de prompt utilisé par un de mes étudiants :

"Tu es un professeur des universités spécialisé en (insérer une matière ou thématique). Je te soumettrais des bouts de mon (spécifier le type de document), tu me diras si tu comprends les passages que je te sou mets. De plus tu feras une correction grammaticale, syntaxique et orthographique si elles sont nécessaires.

Tu répondras dans l'ordre suivant :

- ▶ *Dis d'abord si tu as compris ou non, si tu as des suggestions à apporter comme mentionné précédemment.*
- ▶ *Dis moi si il y a des erreurs orthographiques (majeures ou mineures) , donne la correction*
- ▶ *Dis moi si il y a des erreurs syntaxiques , donne la correction*
- ▶ *Dis moi s'il y a des fautes de conjugaisons*
- ▶ *Ensuite fait des corrections comme spécifiés au dessus, au besoin.*

Liste les avec les passages mentionnés et il faut que tu listes TOUTES les corrections"

Prompt utilisé par Bryan Dallest pour son rapport de stage (étudiant Master 2 Informatique 2023)

Quelles sont les tâches de votre quotidien qui pourraient être assistées par l'IA ?

▶ Génération de données synthétiques :

- ▶ Créer des données d'exemple pour des cas d'usage spécifiques.
- ▶ Générer des données réalistes mais anonymisées pour des contextes sensibles

▶ Conversion de formats :

- ▶ Convertir des formats, comme CSV vers JSON.
- ▶ Réorganiser et nettoyer des ensembles de données.

▶ Formatage de données :

- ▶ Mettre en forme un tableau copié-collé d'un PDF.
- ▶ Créer des visualisations à partir de données brutes.

▶ Génération de code :

- ▶ Produire des scripts Python pour l'analyse de données.
- ▶ Aider à déboguer ou à améliorer un code existant.

Plan

1. Introduction : Qu'est-ce qu'un modèle de langage ?
2. Comment fonctionnent ces modèles ?
3. Utilisation dans un contexte universitaire
4. Prompt engineering et usage diversifié
- 5. Challenges et préoccupations**
6. Personnalisation des Modèles de Langages
7. Conclusion

Éthique, biais et dépendance technologique

- ▶ **Biais dans les données** : Les modèles apprennent des données existantes et peuvent donc perpétuer des biais sociaux ou culturels.
- ▶ **Transparence** : Difficulté de comprendre comment un modèle prend une décision.
- ▶ **Responsabilité** : Qui est responsable lorsqu'un modèle fait une recommandation incorrecte ou nuisible ?¹
- ▶ **Sur-dépendance** : Risque de dépendre trop fortement des modèles, réduisant la pensée critique et l'acquisition de certaines compétences.
- ▶ **Conservation des compétences** : Risque de perdre certaines compétences.
- ▶ **Accès** : Barrières technologiques ou financières limitant l'accès pour certains étudiants, enseignants ou chercheurs.

¹<https://www.washingtonpost.com/travel/2024/02/18/air-canada-airline-chatbot-ruling/>

Les détecteurs d'IA fonctionnent-ils ?

"En bref, non. Bien que certains (y compris OpenAI) aient publié des outils prétendant détecter le contenu généré par l'IA, aucun d'entre eux n'a prouvé qu'il pouvait distinguer de manière fiable entre le contenu généré par l'IA et celui généré par l'homme.

De plus, ChatGPT n'a aucune "connaissance" de ce qui pourrait être généré par l'IA. Il invente parfois des réponses à des questions comme "as-tu écrit cet [essai] ?" ou "cela aurait-il pu être écrit par une IA ?". Ces réponses sont aléatoires et n'ont aucune base factuelle."

Même si ces outils pouvaient identifier avec précision le contenu généré par l'IA (ce qu'ils ne peuvent pas encore faire), les étudiants peuvent effectuer de petites modifications pour échapper à la détection.

<https://help.openai.com/en/articles/8313351-how-can-educators-respond-to-students-presenting-ai-generated-content-as-their-own>

"Do teachers spot AI? Evaluating the detectability of AI-generated texts among student essays" (Computers and Education: Artificial Intelligence)

*"The potential application of generative artificial intelligence (AI) in schools and universities poses great challenges, especially for the assessment of students' texts. Previous research has shown that people generally have difficulty distinguishing AI-generated from human-written texts; however, the ability of teachers to identify an AI-generated text among student essays has not yet been investigated. **Here we show in two experimental studies that novice (N = 89) and experienced teachers (N = 200) could not identify texts generated by ChatGPT among student-written texts.** However, there are some indications that more experienced teachers made more differentiated and more accurate judgments. Furthermore, both groups were overconfident in their judgments."*

<https://www.sciencedirect.com/science/article/pii/S2666920X24000109> (June, 2024)

EdTech Threats to Student Equity in the Age of AI

Rapport du *Center for Democracy and Technology* :

- ▶ *L'introduction rapide de l'IA générative dans les salles de classe a également érodé la confiance des enseignants dans les efforts de leurs élèves, 62% d'entre eux étant d'accord avec l'affirmation selon laquelle "[l]'IA générative m'a rendu plus méfiant quant au fait que le travail de mes élèves soit réellement le leur". Et la moitié des enseignants sont d'accord avec l'affirmation selon laquelle "[l]'IA générative m'a rendu moins enthousiaste à l'égard du travail de mes élèves, car je ne suis pas sûr qu'il s'agisse réellement du leur".*

<https://cdt.org/wp-content/uploads/2023/09/091923-CDT-Off-Task-web.pdf> (September 2023)

Hallucinations dans les Modèles de Langage

- ▶ **Définition** : Les hallucinations font référence à la capacité du modèle de langage à générer des informations qui ne sont ni exactes ni basées sur des données réelles.
- ▶ **Origine** :
 - ▶ Grand volume de données d'entraînement, fonctionnement des modèles.
 - ▶ Diversité des sources, incluant des informations erronées ou biaisées.
- ▶ **Conséquences** :
 - ▶ Production d'informations inexactes ou trompeuses.
 - ▶ Renforcement de biais ou de stéréotypes.
- ▶ **Contre-mesures** :
 - ▶ Fine-tuning avec des données plus spécifiques et fiables.
 - ▶ Évaluation régulière et mise à jour des modèles.
 - ▶ Sensibilisation des utilisateurs à utiliser le modèle de manière critique.

<https://www.rts.ch/info/sciences-tech/2024/article/1-intelligence-artificielle-nouvelle-menace-pour-l-integrite-de-wikipedia-28681292.html>

Hallucinations dans les Modèles de Langage

```
andrewvassili@Mac purpose_auto % cat 5e40c603c2b50c09d89928b
.asset/AssetData/metadata.json | grep hallucinate
"com.apple.textComposition.MailReplyLongFormBasic": "{ { specialToken.chat.role.system } } You are an assistant which helps
the user respond to their mails. Please draft a concise and natural reply based on the provided reply snippet. Please limit the a
nswer within 50 words. Do not hallucinate. Do not make up factual information. Preserve the input mail tone. { { specialToken.chat.
component.turnEnd } } { { specialToken.chat.role.user } } { { userContent } } { { specialToken.chat.component.turnEnd } } { { specialToken.ch
at.role.assistant } }"
andrewvassili@Mac purpose_auto %
andrewvassili@Mac purpose_auto %
andrewvassili@Mac purpose_auto % cat 8c3424876b79571b73d3814
.asset/AssetData/metadata.json | grep json
"com.apple.textComposition.MailReplyQA": "{ { specialToken.chat.role.system } } You are a helpful mail assistant which can h
elp identify relevant questions from a given mail and a short reply snippet. Given a mail and the reply snippet, ask relevant que
stions which are explicitly asked in the mail. The answer to those questions will be selected by the recipient which will help re
duce hallucination in drafting the response. Please output top questions along with set of possible answers/options for each of t
hose questions. Do not ask questions which are answered by the reply snippet. The questions should be short, no more than 8 words
. The answers should be short as well, around 2 words. Present your output in a json format with a list of dictionaries containin
g question and answers as the keys. If no question is asked in the mail, then output an empty list []. Only output valid json and
nothing else. { { specialToken.chat.component.turnEnd } } { { specialToken.chat.role.user } } { { userContent } } { { specialToken.chat.com
ponent.turnEnd } } { { specialToken.chat.role.assistant } }"
andrewvassili@Mac purpose_auto %
```



System prompt for Apple Intelligence

source : <https://x.com/burkov/status/1852169539124965490>

source : <https://arstechnica.com/gadgets/2024/08/>

do-not-hallucinate-testers-find-prompts-meant-to-keep-apple-intelligence-on-the-rails/

Coût énergétique et Numérique responsable

- ▶ **Coût énergétique élevé** : L'entraînement de grands modèles de langage peut consommer des quantités significatives d'énergie
- ▶ Le coût d'utilisation est limitée, mais peut augmenter rapidement avec le nombre d'utilisateurs (e.g. moteurs de recherche) et le type de modèle
- ▶ **Solutions envisagées** :
 - ▶ Optimiser les modèles pour réduire la consommation d'énergie.
 - ▶ Utiliser des modèles adaptés à la tâche demandée (texte vs. image)
 - ▶ Ne pas les utiliser 😊

"À titre de comparaison, le chargement d'un smartphone moyen nécessite 0,022 kWh d'énergie, ce qui signifie que le modèle de génération de texte le plus efficace consomme autant d'énergie que 9% d'une charge complète de smartphone pour 1 000 inférences, tandis que le modèle de génération d'images le moins efficace consomme autant d'énergie que 522 charges de smartphone (11,49 kWh), soit environ la moitié d'une charge par génération d'images. "

Luccioni, S. et al. (2024). Power Hungry Processing: Watts Driving the Cost of AI Deployment? The 2024 ACM Conference on Fairness, Accountability, and Transparency. <https://doi:10.1145/3630106.3658542>

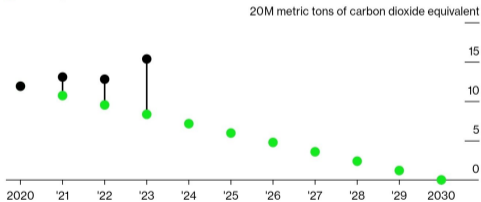
Microsoft's AI Push Imperils Climate Goal as Carbon Emissions Jump 30%

"The company's goal to be carbon negative by 2030 is harder to reach, but President Brad Smith says the good AI can do for the world will outweigh its environmental impact."

Microsoft's Emissions

Artificial intelligence is putting the tech giant's climate goals in peril

● Climate plan (simulated) ● Actual



Source: Microsoft (Scope 1, 2 and 3 "management criteria" data)

Note: Green dots represent linear decline to carbon negative goal.

source : <https://www.bloomberg.com/news/articles/2024-05-15/microsoft-s-ai-investment-imperils-climate-goal-as-emissions-jump-30>

microsoft-s-ai-investment-imperils-climate-goal-as-emissions-jump-30

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

Compar:IA BETA
Le comparateur d'IA conversationnelles

[Nous aider à améliorer Compar:IA](#)

[Accueil](#) [Liste des modèles d'IA](#) [À propos](#) [Partenaires](#)

Ne vous fiez pas aux réponses d'une seule IA

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

J'accepte les modalités d'utilisation [🔗](#)
Les données sont partagées à des fins de recherche

[Commencer à discuter](#)

Comment ça marche ?

- 1 Je discute avec deux IA anonymes**
Échangez aussi longtemps que vous le souhaitez
- 2 Je donne mon avis**
Vous contribuez ainsi à l'amélioration des modèles d'IA
- 3 Les modèles sont démasqués !**
Apprenez en plus sur les modèles d'IA et leurs caractéristiques

<https://languia.stg.cloud.culture.fr/>

Coût énergétique et Numérique responsable



The screenshot shows the 'Compar:IA BETA' interface, a tool for comparing conversational AI. At the top left is the logo of the French Republic with the text 'RÉPUBLIQUE FRANÇAISE' and 'États Généraux de l'Intelligence Artificielle'. To the right is the title 'Compar:IA BETA' and the subtitle 'Le comparateur d'IA conversationnelles'. A 'Accueil' link is in the top right. The main content area shows a user prompt: 'Donne moi un plan d'un cours de Fouille de données en Master'. Below this is a progress indicator 'Étape 1/3' and the heading 'Discussion avec deux IA anonymes'. A note says 'Prêtez attention au fond et à la forme, vous donnerez votre avis à la prochaine étape'. Two model results are shown: 'Modèle A' (purple dot) and 'Modèle B' (orange dot). Model A provides a detailed paragraph about a Data Mining course plan. Model B provides a structured list of course topics.

RÉPUBLIQUE FRANÇAISE
États Généraux de l'Intelligence Artificielle

Compar:IA BETA
Le comparateur d'IA conversationnelles

Accueil

Donne moi un plan d'un cours de Fouille de données en Master

Étape 1/3

Discussion avec deux IA anonymes

Prêtez attention au fond et à la forme, vous donnerez votre avis à la prochaine étape

Modèle A

Voici un plan de cours pour un module de Fouille de Données (Data Mining) adapté au niveau Master. Ce plan peut être ajusté en fonction des besoins spécifiques de votre établissement d'enseignement, de la durée du cours, et du niveau des étudiants.

Modèle B

Plan d'un cours de Fouille de Données en Master

- I. Introduction à la Fouille de Données
- A. Définition et importance de la fouille de données
- B. Histoire et évolution de la fouille de données

<https://languia.stg.cloud.culture.fr/>

Étape 2/3

Quel modèle d'IA préférez-vous ?

Votre préférence enrichit le jeu de données Compar:IA dont l'objectif est d'affiner les futurs modèles d'IA sur le français

Modèle A

Les deux se valent

Modèle B

<https://languia.stg.cloud.culture.fr/>

Coût énergétique et Numérique responsable



<https://languia.stg.cloud.culture.fr/>

Plan

1. Introduction : Qu'est-ce qu'un modèle de langage ?
2. Comment fonctionnent ces modèles ?
3. Utilisation dans un contexte universitaire
4. Prompt engineering et usage diversifié
5. Challenges et préoccupations
- 6. Personnalisation des Modèles de Langages**
7. Conclusion

Personnalisation des Modèles de Langages

- ▶ **Pour une université** : Un modèle adapté aux spécificités, à l'histoire et aux cours offerts par l'université, facilitant les inscriptions, les orientations et les communications internes.
- ▶ **Pour un laboratoire** : Modèle entraîné sur des publications et recherches spécifiques du labo, aidant à la génération d'hypothèses ou à la rédaction d'articles.
- ▶ **Pour un cours** : Modèle spécialisé dans le contenu du cours, offrant du tutorat 24/7, générant des quizz et des résumés.
- ▶ **Pour un enseignant** : Assistant personnel pour la préparation de cours, la correction, et la communication avec les étudiants.
- ▶ **Perspectives** : Avec la démocratisation de l'entraînement de modèles, chaque entité pourra avoir "son ChatGPT" personnalisé, adapté à ses besoins et son domaine d'expertise.

1. Personnalisation via des Prompts

- ▶ Utilisation de questions ou directives spécifiques pour orienter le modèle.
- ▶ Adaptation rapide sans modification du modèle sous-jacent.
- ▶ Exemple : Demander au modèle de répondre dans un format spécifique

2. Personnalisation via du Retrieval-Augmented Generation (RAG)

- ▶ Combinaison de génération de texte et de récupération d'informations.
- ▶ Le modèle génère du texte en s'appuyant sur une base de connaissances externe.
- ▶ Exemple : Utilisation de documents spécifiques pour enrichir les réponses.

3. Personnalisation via Fine-Tuning

- ▶ Réentraînement du modèle sur un jeu de données spécifique.
- ▶ Nécessite des ressources de calcul et un jeu de données pertinent.
- ▶ Permet une adaptation plus profonde aux besoins spécifiques.

Terms of use (OpenAI)

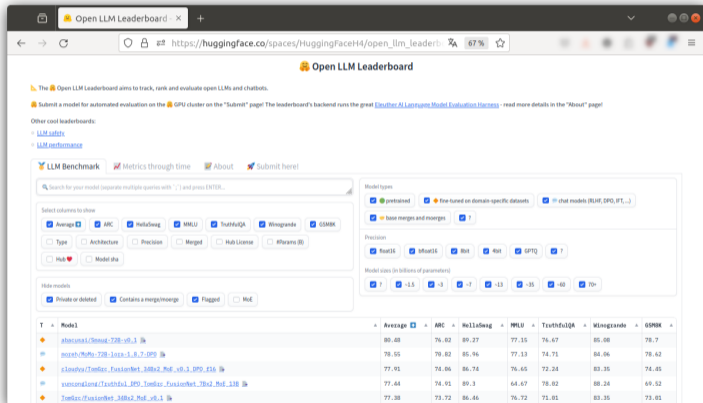
- ▶ Par défaut, tout ce qui est soumis à ChatGPT peut être utilisé par OpenAI pour entraîner ses modèles

When you use our services for individuals such as ChatGPT or DALL-E, we may use your content to train our models. You can opt out of training through our privacy portal by clicking on “do not train on my content,” or to turn off training for your ChatGPT conversations, follow the instructions in our Data Controls FAQ. Once you opt out, new conversations will not be used to train our models.

<https://help.openai.com/en/articles/5722486-how-your-data-is-used-to-improve-model-performance>

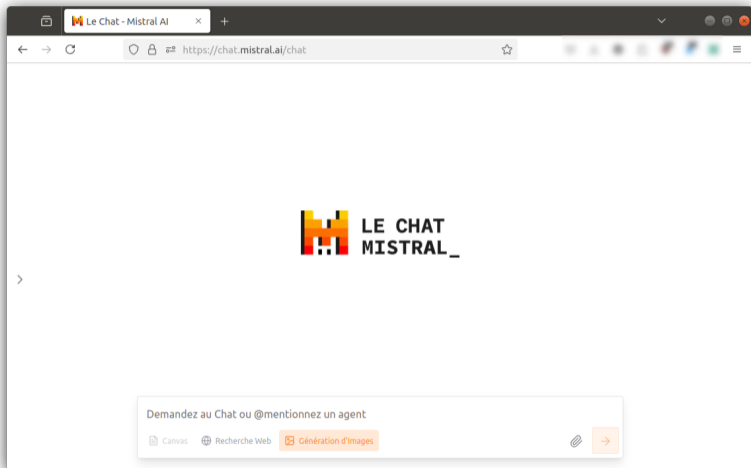
Personnalisation des Modèles de Langages

Le nombre de modèles "open weights" a fortement augmenté :



https://huggingface.co/spaces/HuggingFaceH4/open_llm_leaderboard

Une des alternatives à ChatGPT : Mistral



<https://chat.mistral.ai/>

Plan

1. Introduction : Qu'est-ce qu'un modèle de langage ?
2. Comment fonctionnent ces modèles ?
3. Utilisation dans un contexte universitaire
4. Prompt engineering et usage diversifié
5. Challenges et préoccupations
6. Personnalisation des Modèles de Langages
- 7. Conclusion**

Intelligence artificielle générative : Guide à l'intention de la communauté universitaire

1.	2.
5 Introduction à l'IA générative	17 Guides d'utilisation
6 Connaissances de base	18 Comment communiquer?
10 Une histoire de données	19 Structure d'une requête
11 Risques	22 Styles de travail
15 Bonnes pratiques	24 Choisir un usage
3.	4.
29 Compte rendu	37 Conclusion
30 Adoption à l'UNIGE	40 Remerciements
31 Écosystème	41 Références & ressources
34 Attentes	
35 Craintes et freins	



Connaissances de base

Parler d'intelligence artificielle (IA) induit le biais de se représenter l'intelligence selon notre perception humaine de ce qu'elle peut être. Ses contours sont mal définis, mais nous nous accordons généralement pour en avoir une idée commune qui se rapporte à notre capacité à effectuer certaines tâches avec un certain degré de compétence. L'analogie informatique par l'ajout du qualificatif «artificielle» ne permet pas de comprendre ce qu'est une IA, ni ce qu'elle peut faire, car elle ne partage pas la même cartographie de compétences.

Qu'est-ce que l'IA?

Cette locution désigne un domaine informatique. Les technologies qu'il englobe, souvent utilisées en combinaison, permettent aux ordinateurs d'exécuter un certain nombre de tâches spécialisées. La particularité principale, comparativement à un programme informatique classique, est l'autonomie de l'IA. Elle est conçue pour apprendre et s'améliorer avec le temps, elle peut faire un raisonnement, mais elle peut aussi percevoir des stimuli externes et interagir avec son environnement. Pour ces raisons, une IA est capable de réaliser un certain nombre de tâches complexes, habituellement propres à l'être humain, et généralement associées à notre conception de l'intelligence.

Exemple de recommandations

Intégrer l'IA générative dans les stratégies pédagogiques



https://oer.uclouvain.be/jspui/bitstream/20.500.12279/1089.3/6/CahierLLL_IAG_OKOER.pdf (10/2024)

Exemple de recommandations

L'utilisation pédagogique, éthique et légale de l'intelligence artificielle générative

**L'utilisation pédagogique, éthique
et légale de l'intelligence artificielle
générative**

GUIDE DESTINÉ AU PERSONNEL ENSEIGNANT

2024-2025

<https://cdn-contenu.quebec.ca/cdn-contenu/education/Numerique/Guide-utilisation-pedagogique-ethique-legale-IA-personnel-enseignant.pdf> (10/2024)

Exemple de recommandations



Rapport du Commission de l'Intelligence Artificielle (13/03/2024)

<https://www.elysee.fr/emmanuel-macron/2024/03/13/25-recommandations-pour-lia-en-france>

Exemple de recommandations

R6 : Généraliser le déploiement de l'IA dans **toutes les formations d'enseignement supérieur** et acculturer les élèves dans l'enseignement secondaire pour rendre accessibles et attractives les formations spécialisées.

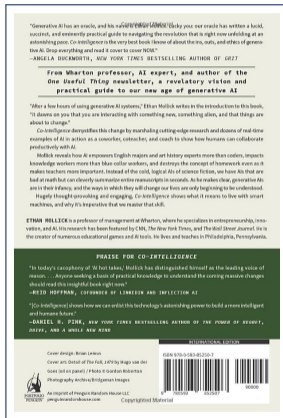
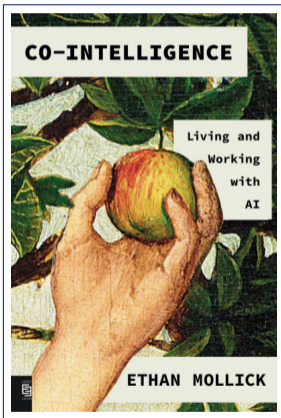
R7 : Investir dans **la formation professionnelle continue** des travailleurs et dans les dispositifs de formation autour de l'IA.

R8 : Former les professions créatives à l'IA, dès les premières années de **l'enseignement supérieur** et en continu.

R11 : Encourager l'utilisation individuelle, l'expérimentation à grande échelle et l'évaluation des outils d'IA pour **renforcer le service public de l'éducation** et améliorer le quotidien des équipes pédagogiques

Rapport de la Commission de l'Intelligence Artificielle (13/03/2024)

Co-Intelligence: Living and Working with AI (Ethan Mollick)



<https://www.oneusefulthing.org/>

<https://mgmt.wharton.upenn.edu/profile/emollick/>

Conclusion

Take-home message :

- ▶ L'impact de l'IA à l'Université dépasse la génération de texte avec ChatGPT
- ▶ Les modèles de langages sont là et les étudiants les utilisent
- ▶ Il n'existe pas de technique infaillible pour détecter si un texte a été généré par une IA
- ▶ Comment évaluer des productions réalisées en dehors d'un cours en présentiel ?
- ▶ **L'interdiction n'a pas de sens**, il faut former les étudiants, les enseignants, les chercheurs et les citoyens à cette nouvelle technologie
- ▶ Les avancées pour ces modèles vont (très / trop) vite

Merci !

contact : germain.forestier@uha.fr

