1

Une question, plusieurs réponses ☺

Je dois aller à la station de lavage pour laver ma voiture. Elle se trouve à 150 mètres de chez moi. Devrais-je y aller à pied ou en voiture ?

2

Risques d'hallucination des chatbots d'IA selon le type de tâche

Domaine de tâches	Problématique typique	Outputs problématiques*
Idéation & structuration	Hypothèses apparemment plausibles mais non fondées	< 5 %
Argumentation & textes explicatifs	Lissage logique, causalité raccourcie	10–25 %
Questions factuelles (closed-book)	Connaissances factuelles inventées ou erronées	15–40 %
Résumés	Omissions, glissements de sens	5–20 %
Sources & citations (sans retrieval)	Sources inexistantes ou incorrectes	20–50 %
Recherche & vue d'ensemble	Présentation sélective donnant une impression de complétude	10–30 %
Codage & programmation	Code fonctionnellement incorrect	5–20 %
Logique mathématique / formelle	Erreurs de calcul ou de dérivation	10–25 %
Événements actuels	Informations obsolètes ou incorrectement datées	15–40 %
Domaines décisionnels sensibles (p. ex. médecine, droit, etc.)	Recommandations inadmissibles	non quantifiable

Etat: janvier 2026

*Les plages de pourcentages indiquées sont des valeurs heuristiques à titre indicatif. Elles reposent sur des caractéristiques typiques des tâches et sur des schémas d'erreurs connus des IA génératives, et non sur des mesures systématiques ou des comparaisons de modèles. Elles indiquent la fréquence approximative de sorties potentiellement problématiques dans des conditions d'utilisation courantes.

3

4

Agenda

1. Brève explication des termes
2. Les quatre principales limites des chatbots
3. Métaprompting & Chain of Thought
4. Démonstration
5. Conclusion

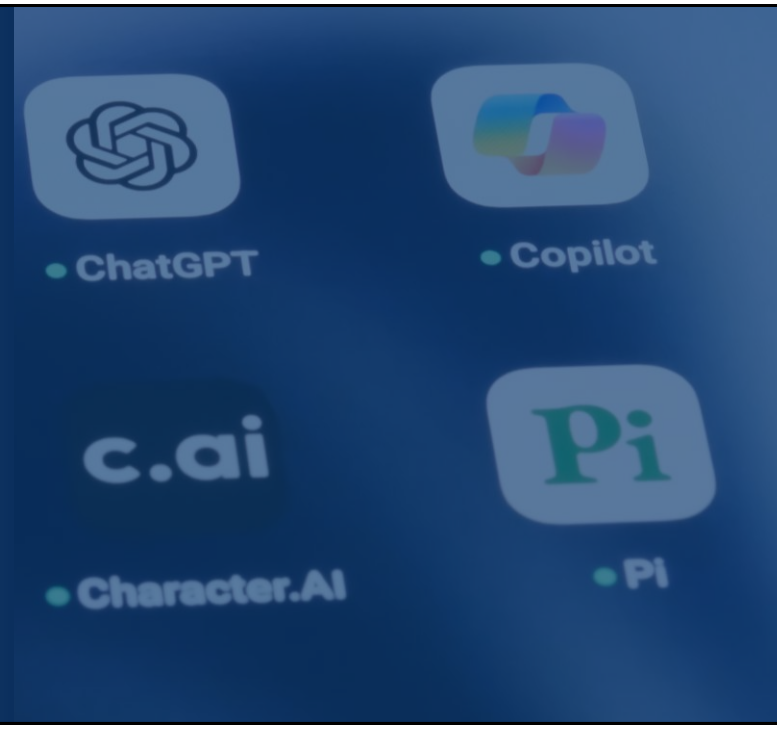
5

ChatGPT

G = generative

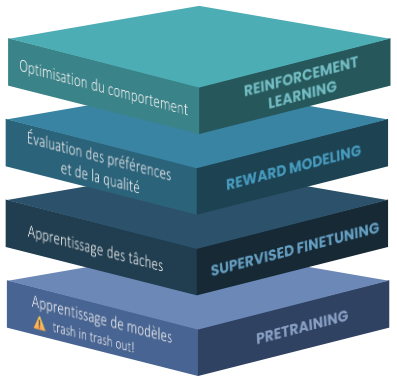
P = pre-trained

T = transformer



6

Le processus d'apprentissage des LLM passe par au moins quatre phases



4. Apprentissage renforcé-> Reinforcement Learning with Human Feedback

À la dernière étape, le modèle est affiné à l'aide de récompenses et de punitions. Il doit être capable d'évaluer dans quelle mesure une réponse du modèle correspond aux **attentes humaines**, par exemple en termes de politesse, d'utilité et de précision. Cette phase **est très exigeante en calcul**, car des milliards de paramètres sont à nouveau ajustés.

3. Modèle de récompense-> Reward Model

Au lieu d'apprendre de nouveaux comportements linguistiques, on entraîne à cette phase un modèle de récompense. Il doit être capable de classer les réponses comme **meilleures ou moins bonnes**. Pour cela, des humains comparent plusieurs réponses générées par un LLM et indiquent celle qu'ils préfèrent. Cette évaluation manuelle **nécessite un effort humain considérable**.

2. Ajustement fin supervisé-> Supervised Fine-Tuned Model

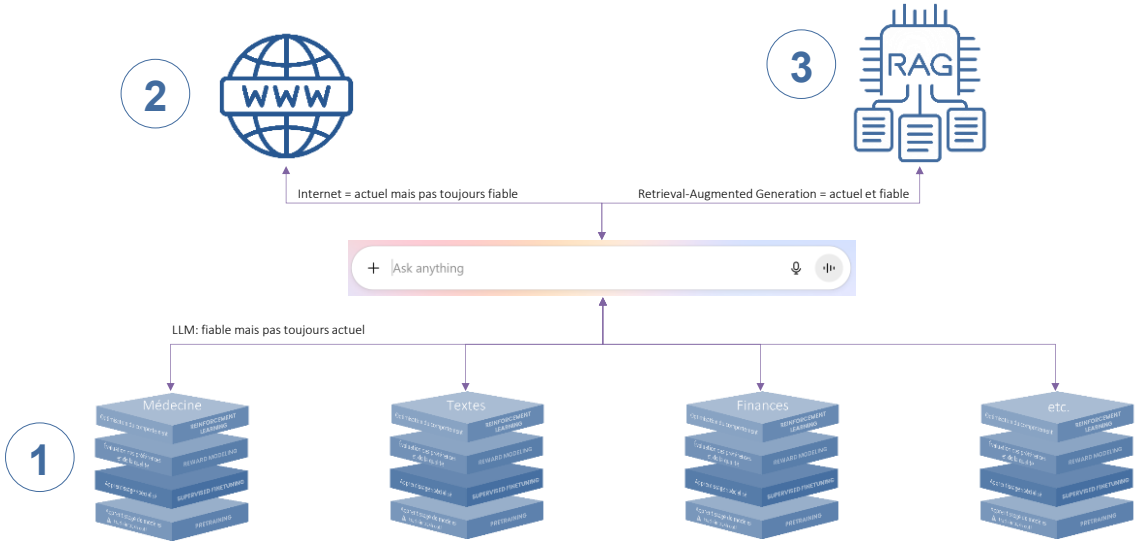
Dans cette phase, le modèle est adapté à l'aide de jeux de données supervisés et soigneusement sélectionnés par des humains pour **des tâches ou domaines spécifiques** – par exemple répondre à des questions médicales, analyser des textes juridiques ou apprendre un style d'écriture particulier (transfer learning). **La charge de calcul est nettement inférieure** à celle du pré-entraînement, car le volume de données est plus réduit.

1. Pré-entraînement -> Modèle de base (Foundation Model)

Lors du pré-entraînement, le modèle apprend de manière autonome à **partir d'énormes volumes de texte (corpus)**, provenant par exemple de contenus web, de Wikipédia, de livres ou de conversations en ligne, afin de capter les schémas linguistiques et les connaissances générales. Cette phase constitue la base de tous les apprentissages ultérieurs. GPT-3, par exemple, a été entraîné sur des milliards de phrases issues d'Internet, ce qui lui a permis d'apprendre à combler des lacunes textuelles. Cette étape **est particulièrement exigeante en calcul et consommatrice de ressources**.

7

Trois sources d'informations pour un chatbot



8

Agenda

1. Brève explication des termes
2. Les quatre principales limites des chatbots
3. Métaprompting & Chain of Thought
4. Démonstration
5. Conclusion

9

10

La règle de base en informatique = TITO

Trash in → Trash out

11

12

La langue des LLM, ce sont les « tokens ».

Pour l'IA, les mots ne sont pas du texte – ce sont des nombres. Et ces nombres forment des motifs géométriques dans l'espace.

Les tokens sont les éléments constitutifs des mots en nombres :

- *Pomme* = 1 token = [11756, 27849]
- *Je mange une pomme = 5 tokens*
- Un mot long peut comporter plusieurs tokens.

https://platform.openai.com/tokenizer

13

Les LLM ont un accès contextuel limité : ils « voient » seulement une partie

Modèle	Contexte max. (tokens)	Correspond approximativement à
GPT-3.5	env. 4'096 token	~ 3–4 pages texte
GPT-4 (standard)	env. 8'192–32'768 token	~ 6–25 pages texte
GPT-5 (128k)	Jusqu'à 128'000 token	env. 100 pages (mais techniquement complexe)

⚠ Le modèle « oublie » les parties plus anciennes de la conversation (elles sont coupées). Même pour les documents longs (p. ex. PDF), seule une partie est analysée.

14

15

Les LLM n’ont (encore) aucune mémoire.

Humain

Je me souviens

VS

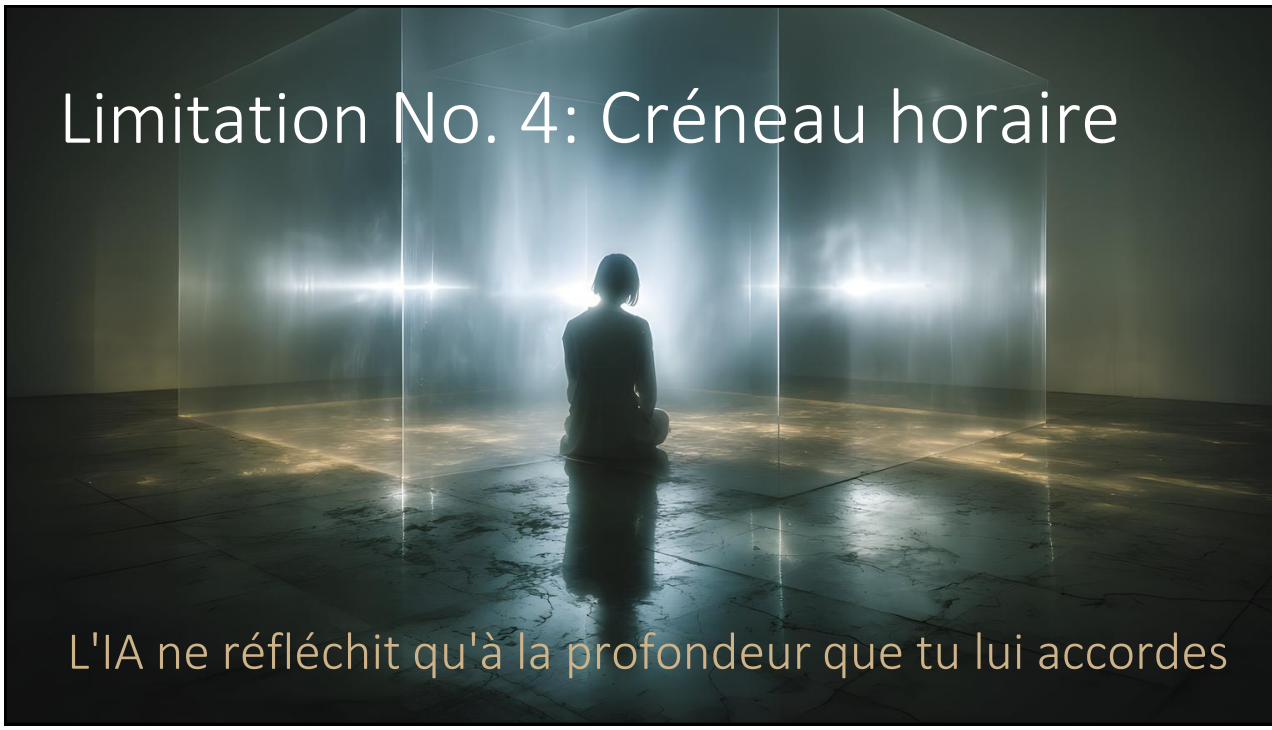
Intelligence artificielle

Je consulte

À chaque prompt, ChatGPT doit relire ce qui a été dit auparavant

ChatGPT 5 peut actuellement relire au maximum env. 128 000 tokens (env. 100 pages). Le reste est coupé !

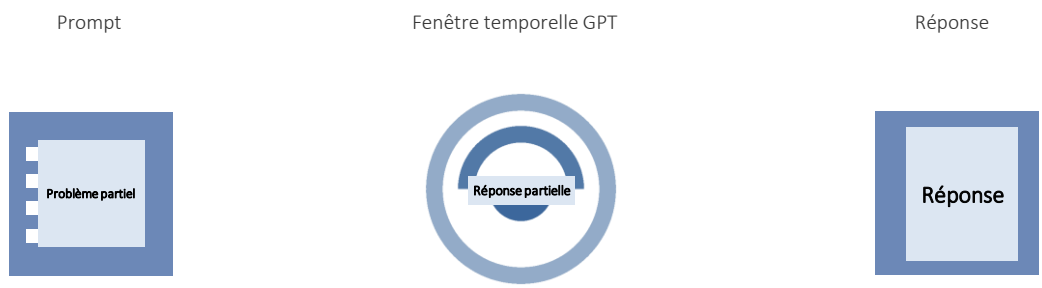
16



17

De meilleures réponses grâce à la « technique du salami » (segmenter)

« Résous le problème étape par étape et explique ton raisonnement ! »



Chain-of-Thought = CoT-prompting

→ Les fenêtres temporelles étant limitées, la structuration du problème en sous-problèmes améliore la qualité de la réponse.

18

Agenda

1. Brève explication des termes
2. Les quatre principales limites des chatbots
3. **Métaprompting & Chain of Thought**
4. Démonstration
5. Conclusion

19

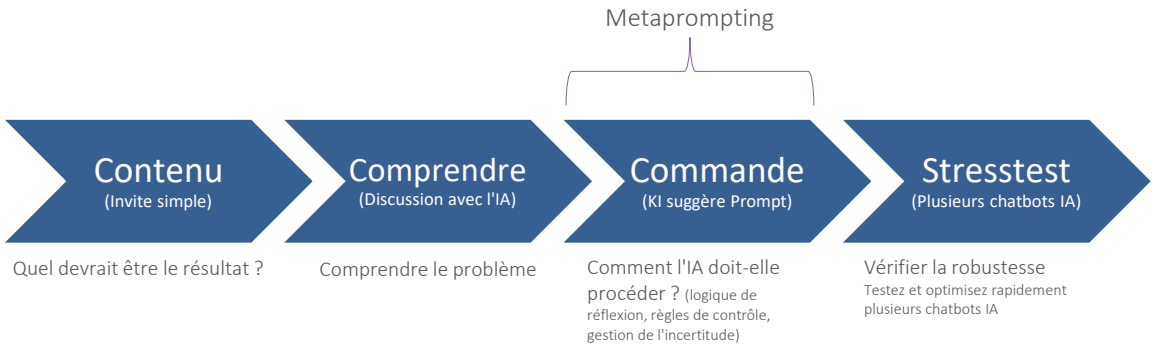
Les bonnes invites font gagner du temps, les mauvaises sont source de frustration

1	**Donne un rôle (set « persona ») et l'objectif…** (p. ex. agis en tant qu'écrivain, agis en tant qu'expert, agis en tant qu'apiculteur, agis en tant que Youtubeur, l'objectif est d'optimiser ma présence sur le web.)	Rôle & Objectif
2	**Explique la tâche …** (p. ex. rédige un essai universitaire, donne-moi une recette pour corriger ce texte, analyse le texte suivant)	Tâche
3	**Fixe les restrictions / le contexte…** (p. ex. donne la priorité à l'exactitude, écris avec des phrases courtes, utilises un style d'écriture poétique, soit concis, ajoutes beaucoup de détails, mets en évidence ce qui fait de chaque méthode un choix exceptionnel, voici un exemple)	Règles
4	**Définis le format de l'output…** (p .ex. présente le résultat dans un tableau, dans Word, dans Excel, dans PowerPoint, crée un graphique, rédige un résumé de deux pages)	Format
5	**Demande-lui comment procéder et s'il a tout compris…** (p .ex. si tu n'as pas bien compris ma demande, pose-moi d'abord des questions jusqu'à ce que tu sois sûr de toi..)	Questions? & Tools

- → **Sois spécifique :** Plus ta demande est précise, plus la réponse sera pertinente
- → **Ajoute des exemples :** Fournis un exemple du style ou du format attendu peut aider l'IA à mieux comprendre
- → **Testes et itère :** Si la première réponse n'est pas parfaite, précises ce qui doit être ajusté dans ton prochain prompt
- → **Pour les questions complexes**, demandez d'abord à l'IA comment elle souhaite procéder. Demandez-lui d'utiliser un outil spécifique.

20

Avec le métaprompting*, nous ne contrôlons pas la réponse, mais la manière dont l'IA arrive à ses réponses



*Le terme « métaprompting » n'a pas de définition uniforme. Dans cette présentation, il désigne le contrôle conscient du mode de pensée et de fonctionnement de l'IA, et pas seulement la description du résultat souhaité.

21

Forces et limites des chatbots d'IA

Domaine de tâches	ChatGPT (OpenAI)	Claude (Anthropic)	Gemini (Google)	Perplexity	Copilot (Microsoft)	Grok (xAI)
Structuration des problèmes & raisonnement heuristique *(ordonner un thème, esquisser des pistes de réflexion)*	🟢	🟢	🟡	🟡	🟡	🟡
Profondeur argumentative & cohérence *(argumenter de manière structurée, pour et contre)*	🟢	🟢	🟡	🟡	🟡	🔴
Textes spécialisés longs et factuels *(rapports, études, concepts)*	🟡	🟢	🟡	🔴	🟡	🔴
Synthèses & condensation de l'information *(extraire les messages clés de textes)*	🟢	🟢	🟡	🟢	🟡	🟡
Recherche systématique *(vue d'ensemble d'un thème, axes de recherche)*	🟡	🟡	🟢	🟢	🟡	🟡
Faits actuels & sources *(qu'est-ce qui est valable actuellement ? où vérifier ?)*	🟡	🟡	🟢	🟢	🟢	🟡
Mise en perspective discursive & pluralité des points de vue *(différentes interprétations et cadres d'analyse)*	🟢	🟢	🟡	🟡	🟡	🟡
Codage & programmation *(expliquer, déboguer et structurer du code)*	🟢	🟢	🟡	🔴	🟢	🟡
Tableurs & logique Excel *(formules, logique des données, analyses)*	🟡	🟡	🟡	🔴	🟢	🔴
Usages décisionnels à haut risque *(médecine, droit, décisions sensibles)*	🔴	🔴	🔴	🔴	🔴	🔴

Etat: janvier 2026

🟢 adapté 🟡 adapté sous conditions 🔴 inadapté

ChatGPT
- Très performant pour la structuration, l'argumentation et le raisonnement explicatif
- Large champ d'utilisation sur de nombreux types de tâches (outil polyvalent)
- Risque élevé de plausibilité trompeuse lorsque le contenu est factuellement incertain

Claude
- Forte cohérence argumentative et explicitation rigoureuse des hypothèses
- Très adapté aux textes spécialisés longs et neutres
- Parfois trop prudent et peu exploratoire

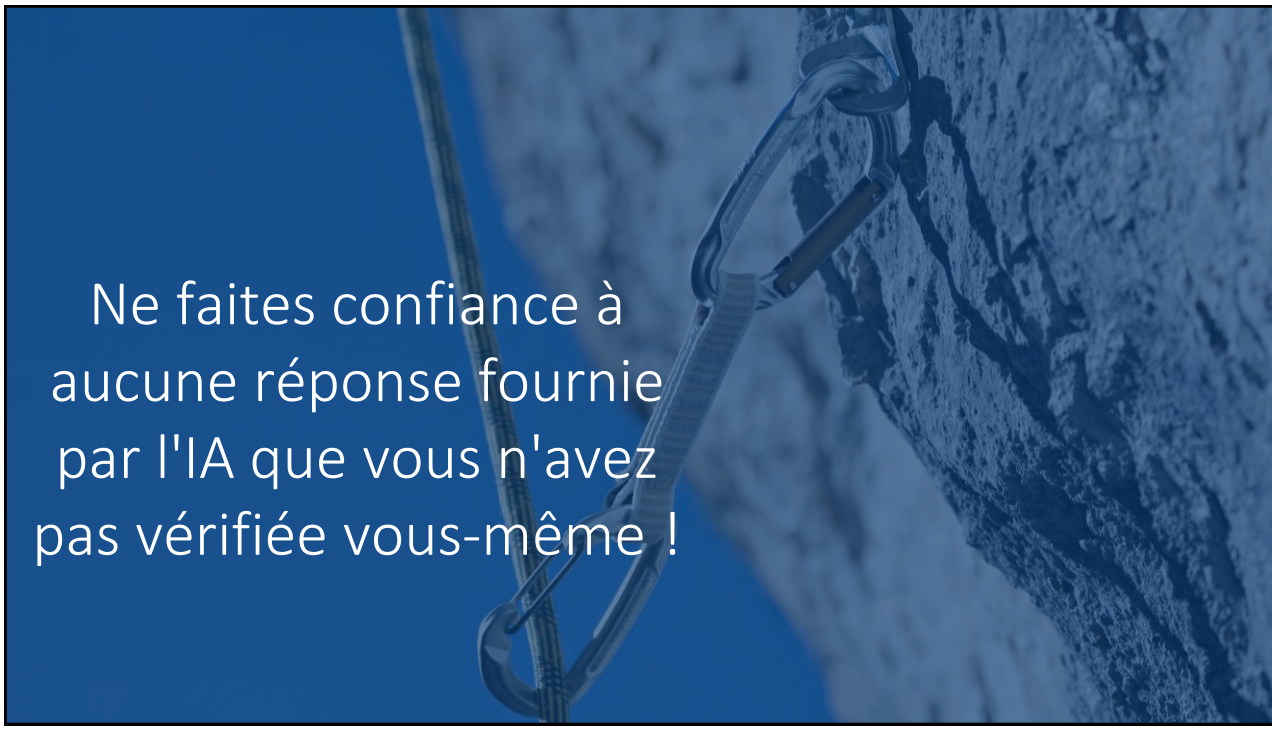
Gemini
- Bonne connexion à l'information récente et à la recherche
- Performance polyvalente solide pour les tâches standards
- Profondeur analytique variable selon les sujets

Perplexity
- Très performant en recherche, exploration et orientation par les sources
- Bonne aide pour s'orienter dans des domaines nouveaux ou peu connus
- Plus faible en synthèse et en analyse cohérente

Copilot
- Très efficace pour les workflows Office, Excel et le codage
- Faible barrière d'entrée pour les tâches productives du quotidien
- Profondeur conceptuelle et argumentative limitée

Grok
- Perspectives marquées, parfois originales ou décalées
- Rapide pour les sujets discursifs ou proches de l'opinion
- Instabilité méthodologique et fiabilité analytique limitée

22



23

Agenda

1. Brève explication des termes
2. Les quatre principales limites des chatbots
3. Métaprompting & Chain of Thought
4. Démonstration
5. Conclusion

24

Démo

1. Notebook LM
2. Metaprompting

25

Agenda

1. Brève explication des termes
2. Les quatre principales limites des chatbots
3. Métaprompting & Chain of Thought
4. Démonstration
5. Conclusion

26

Conclusion : L'avenir n'appartient pas à l'IA, mais aux personnes qui apprennent à penser avec elle

1. **Limites à connaître :** l'IA générative peut halluciner (inventer du contenu plausible mais faux) – ne jamais lui faire confiance aveuglément : toujours vérifier.
2. **Principe TITO & compétence de prompt :** « Trash in – Trash out » : des entrées de qualité et des prompts bien structurés sont indispensables pour obtenir des résultats utiles.

27

https://youtube.com/shorts/9VnYy-I5xsg?si=YPeOEGJI2Q8Ezrds

28

Merci de votre attention

Haute école de gestion
Dr. Serge Imboden
Techno-Pôle 3
3960 Sierre
+41 27 606 90 72
+41 79 217 06 08
serge.imboden@hevs.ch
www.2iManagement.ch

29