

La synthèse de NotebookLM

Points forts de Deepseek V3 :

- **Open source et gratuit** : contrairement à ChatGPT (OpenAI) et Claude (Anthropic), Deepseek V3 est open source, ce qui signifie que son code est accessible à tous. Il est également gratuit pour une utilisation en ligne, ce qui le rend accessible à un large public.
- **Performances élevées** : Deepseek V3 se classe parmi les meilleurs modèles de langage disponibles. Les benchmarks montrent qu'il surpasse ChatGPT 3.5 et GPT 4o dans certaines tâches, notamment le codage et les mathématiques. Il excelle également en raisonnement logique, comme le démontre sa capacité à résoudre le test « Combien de R il y a dans Strawberry ».
- **Capacités de recherche Web** : Deepseek V3 peut être connecté à Internet pour effectuer des recherches en temps réel et fournir des informations actualisées. Il présente les sources utilisées, permettant aux utilisateurs de vérifier les informations.
- **Prise en charge du multilingue** : Deepseek V3 est performant en chinois et prend en charge plusieurs langues, ce qui le distingue des IA souvent limitées à l'anglais.
- **Architecture « Mixture of Experts» (MoE)** : Deepseek V3 utilise une architecture MoE qui lui permet de mobiliser uniquement les "experts" nécessaires pour répondre à une requête. Cela optimise l'utilisation des ressources et réduit la consommation d'énergie.
- **Fonctionnalité « Deep Sync »** : Similaire à la fonctionnalité « Chain of Thought » (Chaîne de pensées) de ChatGPT, Deep Sync permet à Deepseek V3 de décomposer un problème en étapes et de présenter son raisonnement aux utilisateurs.
- **Coût de l'API très bas** : L'API de Deepseek V3 est accessible à un coût nettement inférieur à celui de ses concurrents, la rendant attractive pour les développeurs et les entreprises.
- Deepseek est disponible en mode bureau : <https://chat.deepseek.com/> et (déjà) en application Android à télécharger sur le Play Store et iOS, à télécharger sur l'Apple store.

Limites de Deepseek V3 :

- **Biais et censure** : en tant qu'IA chinoise, Deepseek V3 est soumis à la censure et à la propagande du gouvernement chinois. Il est prudent de ne pas partager d'informations personnelles sensibles.
- **Performances variables** : bien que Deepseek V3 soit performant dans de nombreux domaines, ses performances peuvent varier en fonction de la tâche. Il ne surpasse pas encore ChatGPT One Pro et peut donner des réponses erronées.
- **Faiblesses en rédaction** : Deepseek V3 est moins performant en rédaction que Claude 3.5 Sonnet. Son style est parfois jugé trop « IA » et manque de finesse littéraire.
- **Problèmes de fiabilité de la recherche Web** : Deepseek V3 peut fournir des informations incorrectes lors de l'utilisation de la fonction de recherche web.
- **Sécurité des données** : Deepseek V3 est soumis aux lois chinoises en matière de protection des données, ce qui peut poser un problème aux utilisateurs européens ou américains. La note de sécurité attribuée par ChatGPT est de 5/10. **N.D.L.R** : ChatGPT qui critique la protection des données chez un concurrent, c'est amusant.

En conclusion :

Deepseek V3 est une IA open source prometteuse qui défie les géants du secteur. Ses performances élevées, sa gratuité et son architecture innovante en font une alternative intéressante à ChatGPT et Claude. Cependant, il est important de garder à l'esprit ses limitations, notamment en matière de censure, de fiabilité et de sécurité des données.

N.D.L.R : Les I.A américaines sont plutôt mal placées pour donner des conseils s'agissant de la sécurité des données. Surtout maintenant, avec Donald Trump à la tête de l'Amérique qui, en plein accord avec tous les grands chefs de la Silicon Valley, va lâcher les chiens. De plus, Deepseek est sans doute contrôlée par la censure chinoise, ce qui est certainement un problème pour les Chinois, mais (hormis TikTok, que je fuis comme la peste) pas encore pour nous, tout de même.

Le podcast généré par NotebookLM, en français, à partir des URL des trois vidéos.

<https://notebooklm.google.com/notebook/603a77d7-5be5-4113-8021-d9ecec3e945/audio>

Mon avis sur Deepseek

Je l'utilise depuis quelques jours et j'ai été bluffé à chaque fois. Je précise que je vis littéralement dans l'intelligence artificielle depuis plus d'un an. C'est-à-dire que tous mes articles sont désormais rédigés en collaboration avec une I.A. Lesquelles ont souvent varié depuis quelques mois.

Comment je procède :

Je demande à l'IA un résumé, ou, comme c'est le cas cette fois avec NotebookLM, une synthèse et un podcast de trois vidéos sur Deepseek.

Je lis attentivement le résumé ou la synthèse et je lui pose des questions ou j'attire son attention sur un point omis, ou pas assez souligné à mon sens.

A l'issue d'une discussion dont la longueur peut être très variable, je demande à l'IA de synthétiser son résumé initial avec nos remarques qui l'ont suivi.

Cela n'a vraiment rien de compliqué et ne nécessite certainement pas d'être expert en informatique.

Je me suis abonné à Perplexity, ChatGPT, et Gemini extended.

A ce jour (tout peut changer très vite dans ce domaine) Deepseek présente le meilleur rapport qualité prix de toutes IA que j'ai pratiquées.

Pour ce qui est des plus, cette synthèse est éloquente.

Pour ce qui est des moins, invoquées par mes collègues blogueurs :

La censure chinoise : à ce jour, je n'ai rien vu de tel dans les prestations de Deepseek. Pour autant, il ne me viendrait pas à l'idée de demander son avis à Deepseek sur la politique chinoise. Pas plus que je ne demanderai jamais son avis à Micron sur la politique française de ces dernières années. Ou à Trump sur la politique américaine des prochaines années.

Dans une des vidéos, un collègue l'a fait ; dans un premier temps, Deepseek n'a pas voulu répondre. Normal ! La question portait sur le massacre de Tien An Men. Dans un second temps, l'IA a exposé successivement les points de vue occidentaux et chinois sur cette question !

Comment mieux répondre ? Essayez de poser la même question aux IA américaines sur les massacres perpétrés par leur armée, je doute fort qu'elles vous présentent le point de vue des populations exterminées.

Les erreurs dans ses réponses sur l'actualité : ma première discussion avec Deepseek a été de lui demander une revue de presse de la semaine écoulée, sur certains thèmes précis, et avec uniquement les médias indépendants dont je lui ai donné la liste.

Sa revue de presse, quasi instantanée, m'a fait tomber de ma chaise.

J'avais demandé la même chose à ChatGPT, Perplexity, et Gemini. Ces trois IA ont été incapables de me proposer des résultats satisfaisants. Perplexity, particulièrement réputée pour ses capacités Web, a fait nombre d'erreurs de dates, notamment, et a fini par m'avouer que même si elle a accès au Web, ce n'est pas son cœur de métier ! Seule Gemini m'avait fourni une revue de presse sans erreurs, mais ses choix d'articles n'étaient vraiment pas terribles. Je rappelle quand même que je demandais à une IA d'origine chinoise de choisir dans l'actualité française de la semaine dernière les articles les plus significatifs. Et, des articles de médias indépendants uniquement, donc tout sauf des perroquets.

Quant à la qualité des résultats fournis par Deepseek, je lui ai demandé de résumer **le livre de Laurent Gounelle : L'éveil** dont j'ai parlé sur ce site, un PDF de 19 MO. Non seulement son résumé, par chapitres, a été parfait, mais je lui ai demandé par la suite de retrouver dans ce livre toutes les allusions (cachées) au traitement désastreux de la crise de la Covid en France. Il en a retrouvé 12 ! Là, on était vraiment dans l'intelligence artificielle ! Je ferai un article prochainement sur cette prouesse et je poserai la même question à ChatGPT 01, pour voir.

Conclusion : je vais suspendre mes abonnements à Perplexity Pro, ChatGPT Plus, et Gemini Extended : 60 euros par mois ! Et, me cantonner à :

Deepseek, version ordinateur et version Android. Quand elle me le demandera, ce qui arrivera fatalement si je m'adresse qu'à elle, j'achèterai une API, ce qui évite un abonnement qui d'ailleurs n'existe pas pour Deepseek. Voir la note à la fin de l'article sur les prix de Deepseek par rapport à la concurrence. C'est très... chinois.

NotebookLM, parfait et pour l'instant gratuit. Pour cet article, je lui ai simplement donné 3 URL de vidéos Youtube sur Deepseek et il m'a fourni, dans la seconde, une synthèse parfaite et après quelques minutes, un podcast de 13 minutes **en français !**

Je continue également mon abonnement à **Recall.ai** qui fait rapidement des résumés très complets de vidéos, d'articles, de PDF, etc. pour seulement 10 euros par mois.

Je rappelle qu'à l'exception de Gemini Extended, toutes les autres I.A sont incapables de faire cela avec uniquement l'adresse Internet de la vidéo. Il faut impérativement leur fournir la transcription en mode texte ou PDF.

Avec Recall, je suis sur Youtube, je lui partage l'URL de la vidéo et j'obtiens dans la minute suivante le résumé, concis ou complet, dans la minute. Ou lorsque je lis un article qui m'intéresse, j'appelle l'extension Chrome et j'obtiens immédiatement le résumé.

Deepseek n'en est qu'à la version 3 ! Nul doute qu'elle va encore progresser.

NOTE* : prix de Deepseek par rapport à la concurrence.

DeepSeek-V3 **est 10 fois moins cher** que ses concurrents pour les tokens de sortie.

Par exemple, si ChatGPT coûte environ 3 à 5 dollars pour 1 million de tokens de sortie, DeepSeek-V3 coûte seulement 1 dollar pour la même quantité.

Pourquoi cette différence ?

DeepSeek-V3 utilise des techniques comme la Mixture of Experts (MoE), qui active seulement une fraction des paramètres du modèle à la fois, réduisant ainsi les coûts de calcul.

De plus, le modèle a été entraîné pour 6 millions de dollars, contre des centaines de millions pour des modèles comme GPT-4, ce qui se reflète dans les tarifs proposés.

N.D.L.R : les particuliers, hormis les compulsifs, ne sont pas concernés par cette tarification. En revanche, elle devrait intéresser les entreprises et les développeurs.