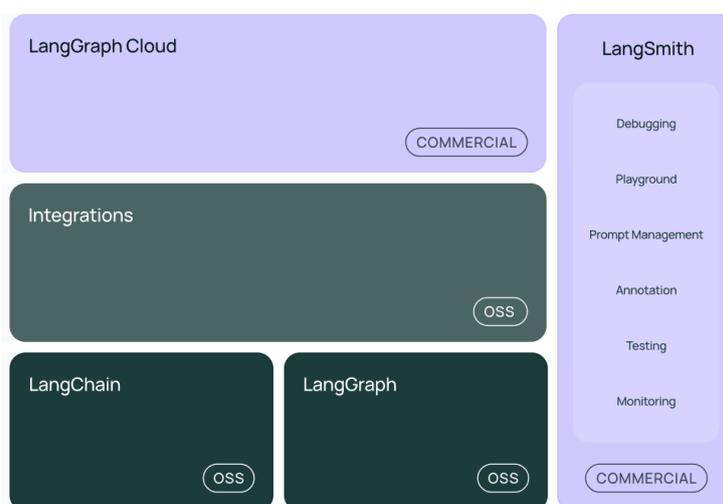


LangChain.js



Qu'est-ce que LangChain.js ?

LangChain.js est un outil de développement qui permet de créer des applications intelligentes à base d'intelligence artificielle. Il est écrit en JavaScript et TypeScript. Grâce à lui, on peut connecter des modèles de langage (comme ChatGPT) à d'autres services, comme des bases de données, des fichiers, des sites web ou des API. Le but est de créer des programmes capables de réfléchir, répondre et même agir automatiquement, comme un assistant virtuel.



Les nouveautés et changements avec LangChain.js.

Lire et comprendre différents types de fichiers et de données.

LangChain.js peut maintenant lire facilement des fichiers comme des PDF, Word, Excel, ou même des sites web. Il peut aussi se connecter à des bases de données ou à des services en ligne. Cela permet à l'intelligence artificielle d'utiliser ces données pour répondre à nos questions ou faire des résumés. Par exemple, on peut lui donner un dossier PDF et lui demander de faire un résumé ou d'expliquer certains passages.

Créer des assistants intelligents avec mémoire.

Avec LangChain.js, on peut créer des assistants qui se souviennent de ce qu'on dit. Ils peuvent raisonner, c'est-à-dire faire des étapes de réflexion pour trouver la bonne réponse ou résoudre un problème. Ils peuvent aussi choisir automatiquement les bons outils à utiliser, comme faire une recherche internet ou envoyer un message à notre place.

Fonctionne avec les IA les plus connues.

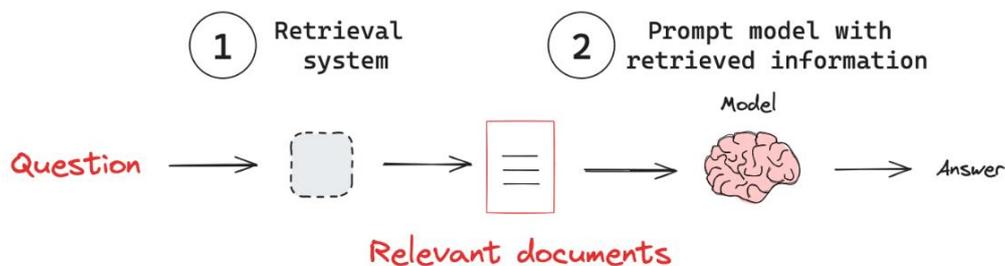
LangChain.js fonctionne avec les grands modèles d'intelligence artificielle du moment :

- GPT-3.5 et GPT-4 de OpenAI
- Claude 2 et 3 de Anthropic
- Mistral et Mixtral, des modèles open-source
- D'autres modèles hébergés sur Hugging Face

On peut donc choisir le modèle qu'on préfère ou celui qui est le plus adapté à notre besoin, sans changer tout le projet.

Recherche d'informations avec la technique RAG.

LangChain.js permet aussi d'utiliser une méthode appelée RAG (Retrieval-Augmented Generation). Cette méthode aide l'IA à chercher des documents ou des fichiers qui contiennent la réponse avant de répondre. Par exemple, si on pose une question sur un gros document, l'IA va chercher la bonne partie du texte avant de répondre. Cela rend ses réponses plus précises et à jour, surtout pour les données récentes.



(1) Système de récupération : Récupérer des informations pertinentes à partir d'une base de connaissances.

(2) Ajout de connaissances externes : Transmettre les informations récupérées à un modèle.

Déploiement facile avec LangServe.

LangChain.js a aussi un outil appelé LangServe. Il sert à transformer un assistant IA en service accessible sur internet. Grâce à ça, on peut connecter notre IA à une application mobile, un site web ou une interface sans trop de difficulté. On peut aussi l'utiliser sur des serveurs classiques, sur le cloud, ou même en edge computing (pour que ce soit plus rapide et sécurisé).

Conclusion.

LangChain.js est un framework très complet pour créer des applications avec de l'intelligence artificielle. Il ne se limite pas à faire des chatbots : il peut aussi agir, réfléchir, lire des documents, se connecter à des services. C'est un outil de plus en plus utilisé en 2024-2025, notamment dans les entreprises qui veulent intégrer l'IA dans leurs logiciels.

Mes sources.

<https://js.langchain.com>

<https://www.langchain.com>

<https://blog.langchain.dev>

<https://huggingface.co>

Rédigé le : 29/12/2024