

Projet LLM

Projet BUT 3 – Science des Données



Présentation



- 2019-2021 : **DUT STID** à Niort
- 2021-2024 : **CNAM Big Data et IA** à Niort
- Depuis 2024 : **Ingénieur IA** au sein du **Groupe IMA**

le cnam

- **Mon métier ?** Créer des applications connectées aux données de l'entreprise et embarquant de l'IA afin d'aider les métiers (équipes) de mon entreprise dans certaines de leurs tâches.



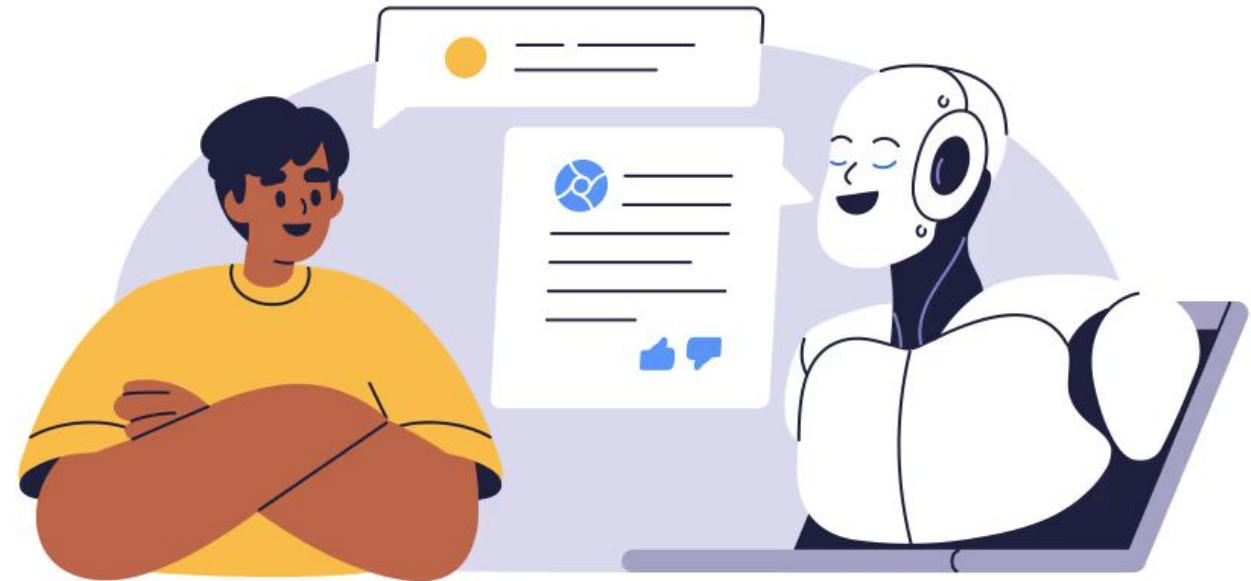
Le Projet

- **Création d'un portfolio "traditionnel"**
 - **Site web** avec sections : **compétences, expériences, projets, contact....**
- Malheureusement : Créer un portfolio est devenu "facile"

Se démarquer des autres ?

Le Projet

- ~~Portfolio statique~~ → Portfolio **INTERACTIF**
- **Créer un chatbot** permettant de discuter avec votre jumeau virtuel, qui **connait** tout de votre **profil professionnel**.
- On pourra poser des questions comme
 - Qui es-tu ?
 - Quelles sont tes compétences ?
 - Sur quels projets as-tu travaillé ?
 - ...



Comment faire ?

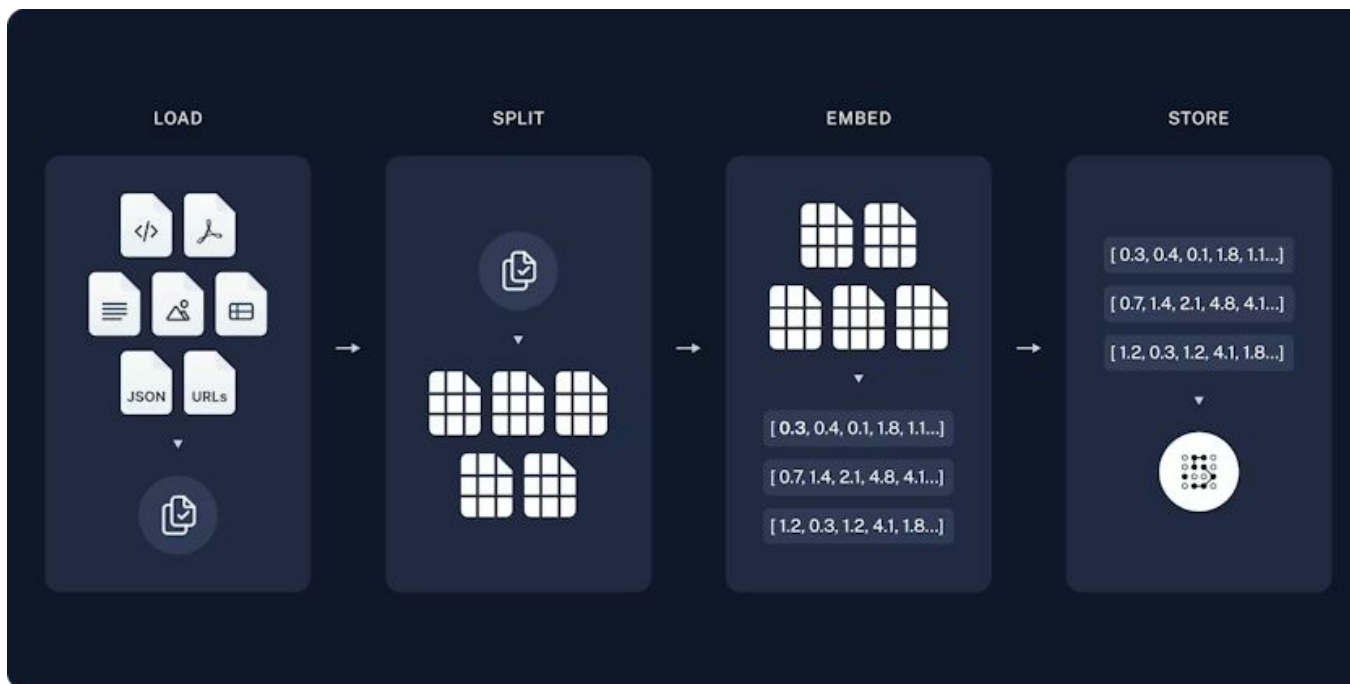
- Comment un **chatbot** peut-il **répondre correctement sur vous** ?
- Les **LLMs connaissent beaucoup** de choses, mais **pas les informations privées** (vos données, les données d'une entreprise...)

Retrieval **A**ugmented **G**eneration

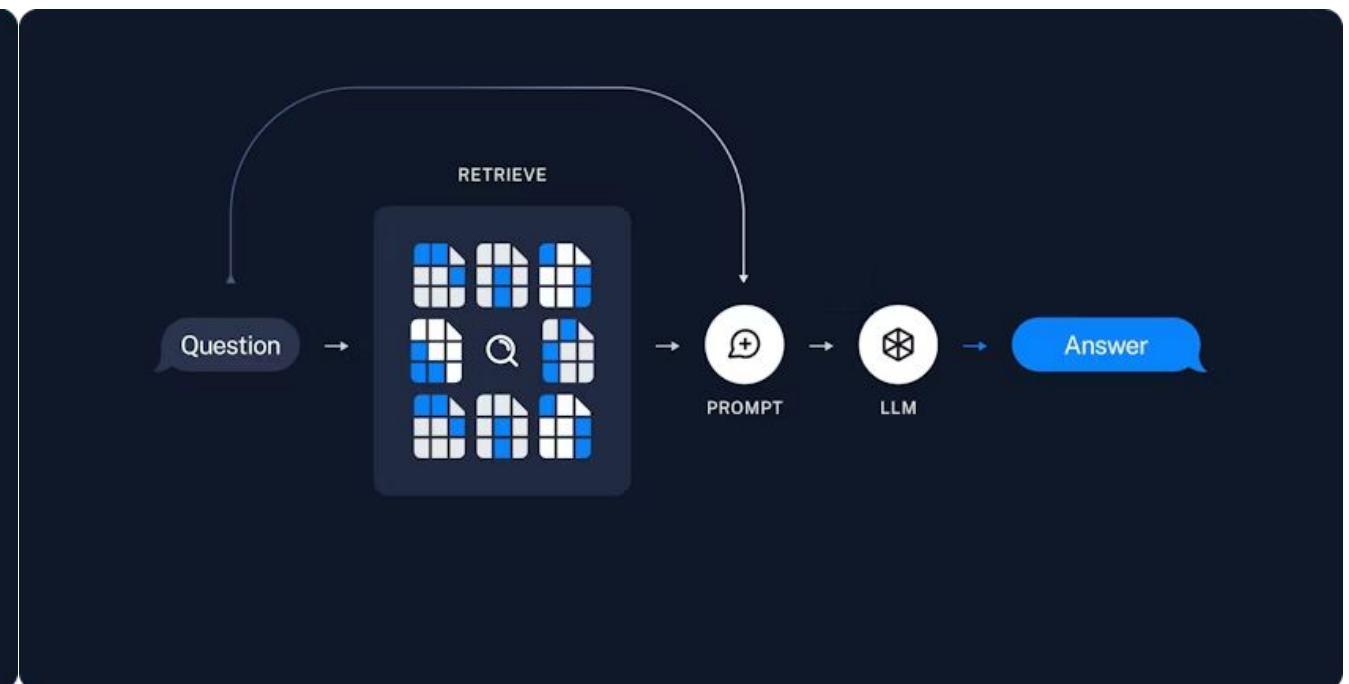
RAG, c'est quoi ?

- Technique qui consiste à améliorer les réponses des modèles d'IA générative en les alimentant avec des connaissances issues des bases de données internes/privées.

1 - Indexation



2 - Récupération & Génération

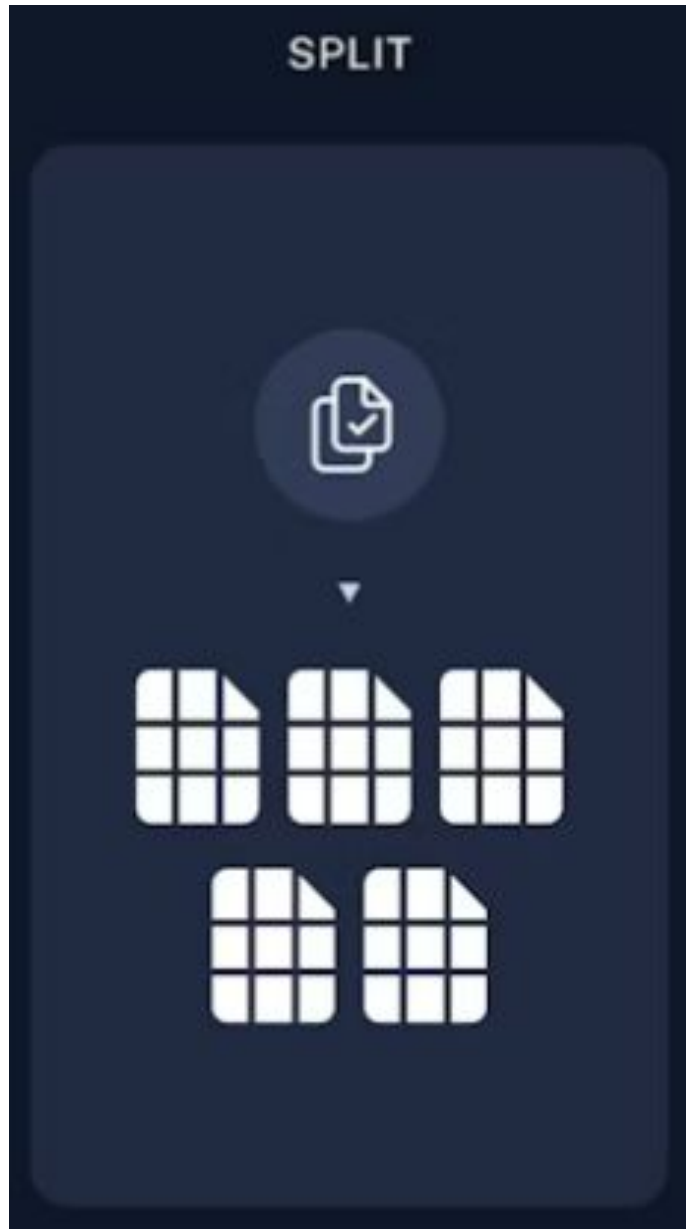


Indexation



1. Charger le contenu d'un document (PDF, JSON, DOCX, HTML, Images...)

Indexation



- Parfois le contenu des documents est trop long, il y a trop d'informations
- Découper / Séparer les informations

1^{er} chunk
(Introduction)



Le Groupe IMA (Inter Mutuelles Assistance) est un [groupement d'intérêt économique](#) (GIE) dont l'activité est l'[assistance](#) en Europe. C'est une filiale des mutuelles françaises [MACIF](#), [MAIF](#), [Matmut](#), etc.

2^{ème} chunk
(Histoire)



Histoire [[modifier](#) | [modifier le code](#)]

IMA a été créé en 1981 à la suite de la volonté de trois [mutuelles d'assurance](#) (la [MAIF](#), la [MAAF](#) et la [MACIF](#)) de mutualiser leurs ressources pour créer une [société d'assistance](#) et d'apporter à leurs sociétaires des prestations d'assistance.

Puis progressivement, d'autres mutuelles d'assurance les ont rejointes² : la [MAPA](#) et la [SMACL](#), suivies de la [Matmut](#), la [MAE](#), l'[AGPM](#), l'[AMF](#) et la [Mutuelle des Motards](#). En 2009, la MAAF a rejoint Société de groupe d'assurance mutuelle [Covéa](#) / groupe [Azur-GMF](#) qui possède sa propre société d'assistance [Fidelia](#).

Le Groupe IMA s'est développé à l'étranger et rassemble aujourd'hui différentes filiales.

3^{ème} chunk
(Activités)



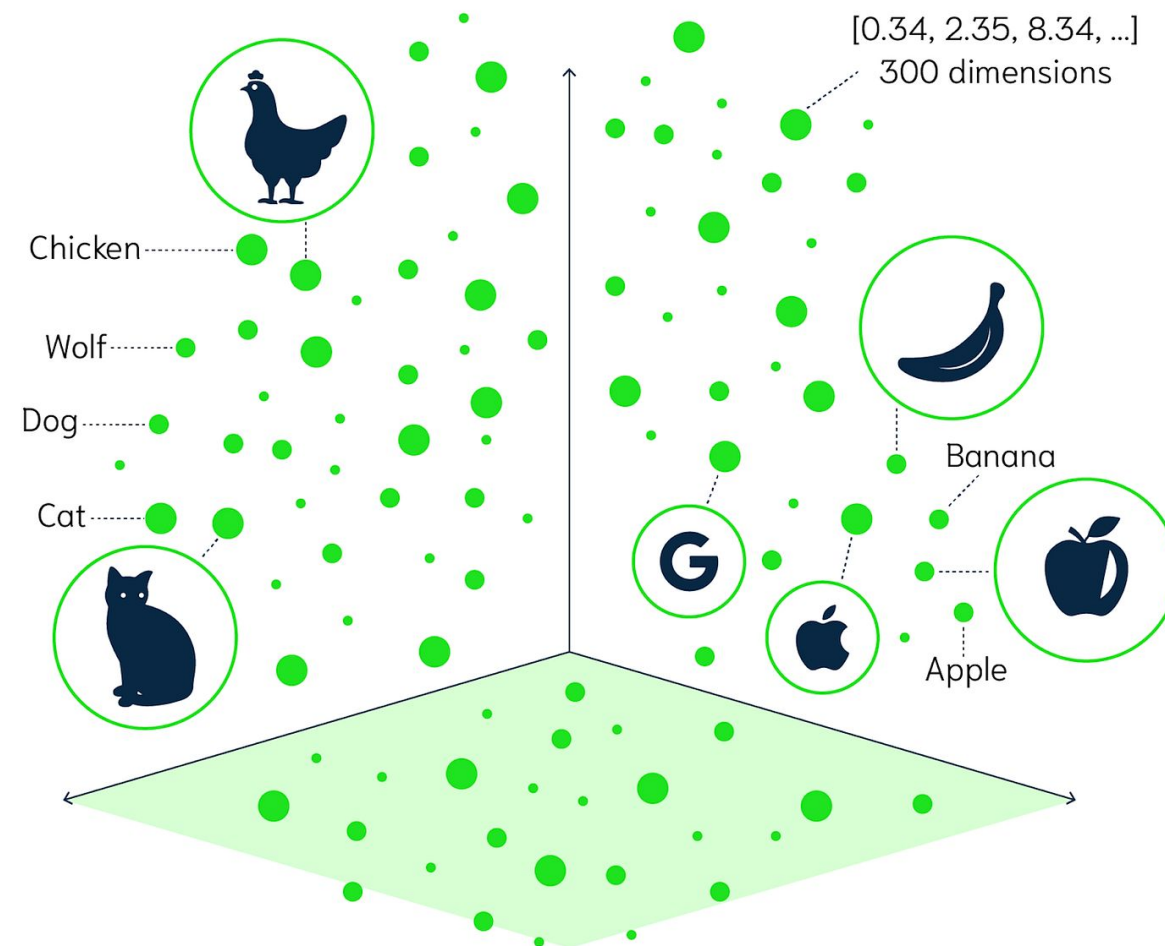
Activités [[modifier](#) | [modifier le code](#)]

Activité commerciale [[modifier](#) | [modifier le code](#)]

Basée à [Niort](#), dans les [Deux-Sèvres](#), l'entreprise est le numéro 2 de l'assistance en France³ (plus d'1 Français sur 2 est assisté par IMA), et le numéro 4 à l'international, selon le classement de [L'Argus de l'assurance](#) (2017).

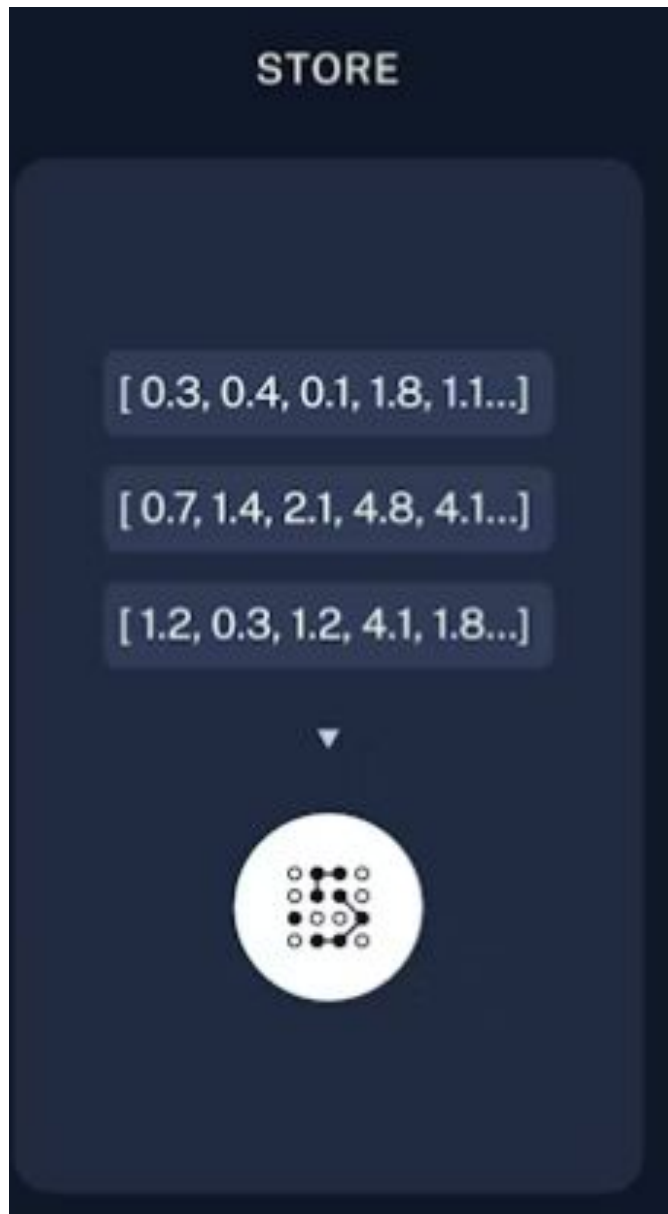
Indexation

- Transformer les chunks en vecteurs numériques compréhensibles par la machine



Indexation

- Sauvegarder les vecteurs dans une base de données vectorielle



Faiss



Pinecone



Chroma



elastic



drant

Récupération et génération

- **Récupérer les documents pertinents** via **recherche sémantique** à partir de la question de l'utilisateur afin de créer un « **contexte** »
- A partir du contexte récupéré, génération d'un **prompt** (Instruction) afin de formater/guider le comportement du **LLM** lors de la génération de la réponse

HUMAN

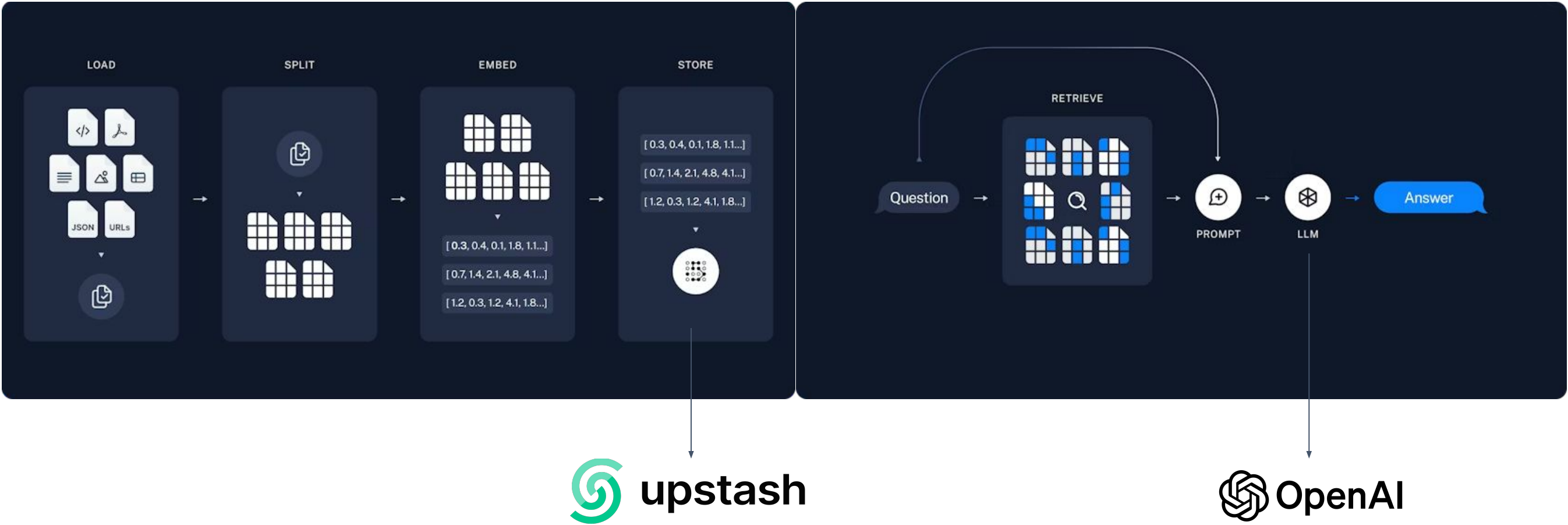
You are an assistant for question-answering tasks. Use the following pieces of retrieved context to answer the question. If you don't know the answer, just say that you don't know. Use three sentences maximum and keep the answer concise.

Question: {question}

Context: {context}

Answer:

Outils



Calendrier

- **8 janvier** : Présentation du projet + Vérifier que tout fonctionne sur les PC
- **15 janvier** : Projet
- **16 janvier** : Projet
- **25 janvier à 23h59** : Deadline



Grille d'évaluation

Thème	Points
Code	10
Qualité du code (simple, optimisé...)	3
Modularité du code	2
Déploiement en ligne (Streamlit Cloud)	2
Respect des bonnes pratiques (DocString, typage des variables des fonctions...)	3
Rapport (3 pages MAXIMUM)	10

Livrables

- Lien Repo Github
- URL Application Streamlit

git clone

<https://github.com/TheoKoess/projet-iut-potfolio.git>