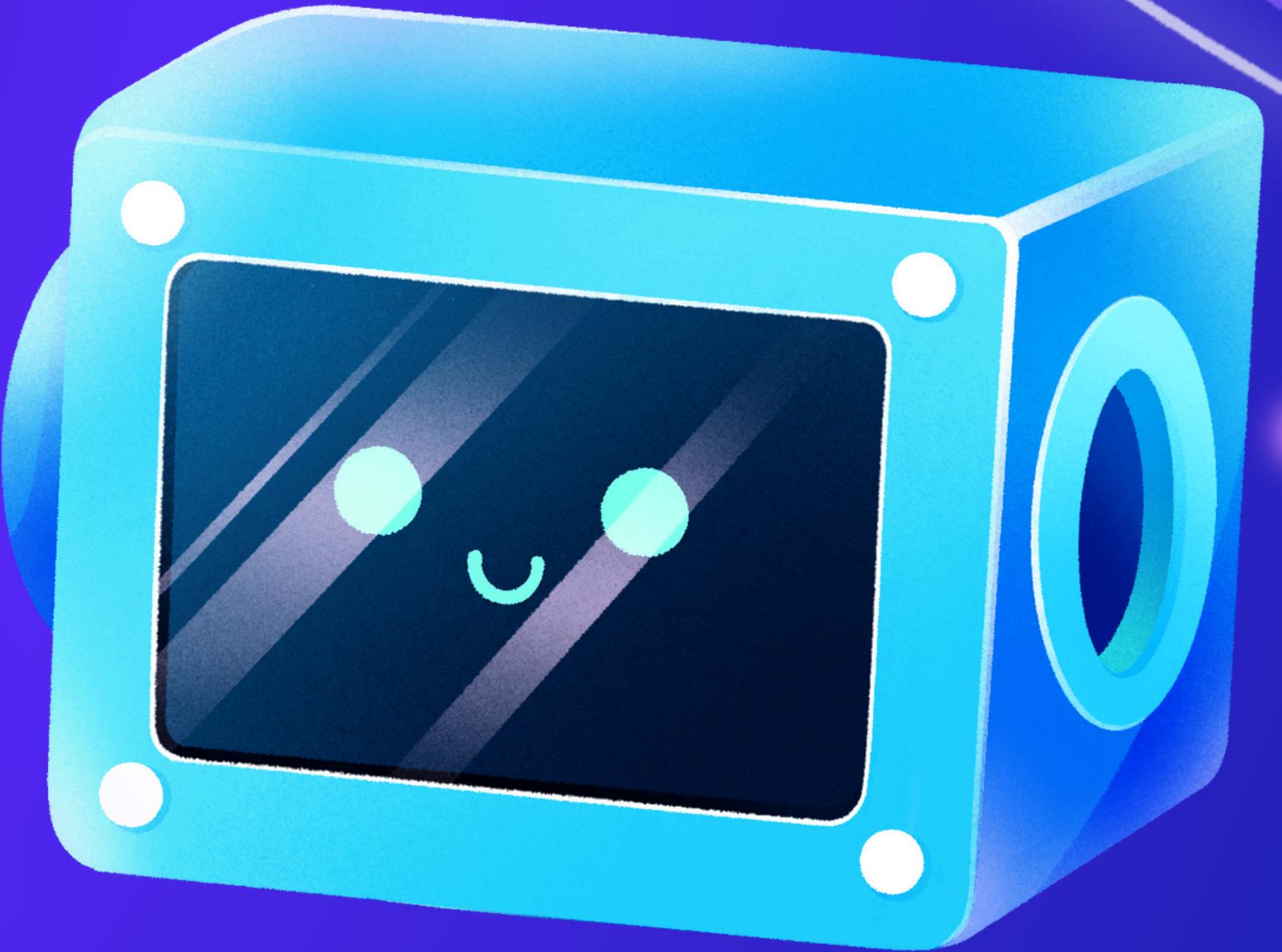


POC ARTE



BESOIN  
&  
MÉTHODO



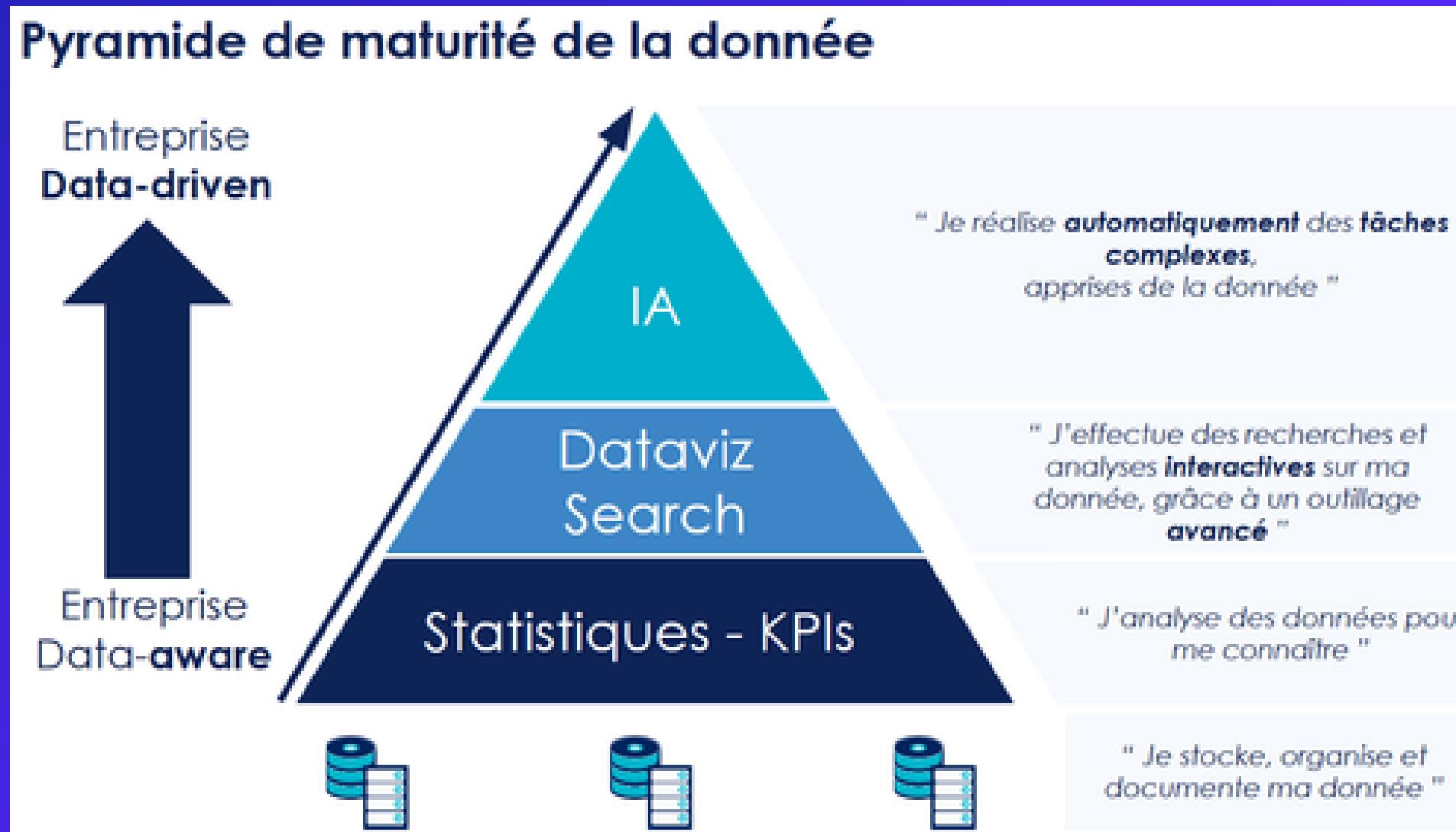
# MÉTHODOLOGIE DE PROJET IA

## Schema directeur Data IA

<b>1. Objectifs stratégiques</b>	<b>2. Diagnostique de l'existant</b>	<b>3. Cartographie des besoins</b>	<b>4. Analyse des solutions et impacts</b>	<b>5. Orientations stratégiques et priorisation</b>	<b>6. Trajectoire et plan d'action</b>
Identification des objectifs clés des directions	Référencement des initiatives data et IA	Ateliers métiers d'identification de cas d'usage	Evaluation stratégique des cas d'usage	Stratégie make or buy par verticale technologique et cas d'usage	trajectoire budgetisée de cas d'usage Data IA
Analyse des OKR	Cartographie des compétences	Diagnostic des processus back-office sous l'angle DAta IA	Identification et analyse des solutions Data IA	Fiches d'opportunité & business cases par cas d'usage Data IA	Lancement des 1ers quick wins
Priorisation des gisements de valeur	Diagnostic de maturité Data IA du SI	Analyse des prérequis et impacts d'architecture et de données :global & par cas d'usage	Evaluation des impacts métiers et process		

**7. Communication et conduite du changement de l'étape 2 à 6**

# MÉTHODOLOGIE DE PROJET IA



We've found that success lies in three critical components: data preparation, model optimization, and retrieval design. When building our system, we discovered that **high-quality training data wasn't just important - it was everything.**

# DÉFINITION DU BESOIN

**La demande doit venir des métiers**

**Importance de placer l'utilisateur au centre de la transformation**

**Data IA**

(Guide de l'IA dans l'entreprise de Stéphane Roder)

# PARAMÈTRES À PRENDRE EN COMPTE

- scalabilité
- multilingue (Français, Anglais, Allemand, Espagnol, Polonais, Italien, Roumain)
- frugalité
- respect de l'AI Act, du RGPD, de l'EAA
- utiliser principalement des technologies open-source
- réutilisabilité du projet par d'autres médias ?

# LES 5P

On souhaite discuter avec un chatbot en dessous des vidéos du dessous des cartes

Pourquoi ?

- Pour mieux comprendre des aspects de la vidéo qu'on a mal ou pas bien compris

*\*Pourquoi veut-on mieux comprendre ces sujets ?\**

- Parce que certains termes sont difficiles à comprendre (géopolitique) et parce que le format des vidéos est court

*\*Pourquoi a-t-on besoin d'explications complémentaires ?\**

- Parce que chaque viewer n'a pas le même niveau de culture géopolitique et a donc besoin d'un accompagnement personnalisé

*\*Pourquoi est-ce important d'avoir un accompagnement personnalisé ?\**

- Parce qu'un contenu plus accessible et interactif renforce l'engagement, l'apprentissage et l'envie de revenir sur Arte.tv

*\*Pourquoi veut-on renforcer l'engagement et l'apprentissage autour de Le Dessous des cartes ?\**

- Parce que cela permettrait à Arte de remplir pleinement sa mission éducative tout en fidélisant un public curieux qui cherche à comprendre le monde de manière critique et nuancée.

# LE BESOIN RÉSUMÉ

**On veut discuter avec un chatbot sous les vidéos du Dessous des cartes pour rendre le contenu plus compréhensible, accessible et engageant, afin de renforcer l'impact éducatif d'ARTE et fidéliser un public curieux.**



Bonjour,

On m'a demandé récemment de travailler sur un POC de chatbot qui serait lié à des vidéos de géopolitique dans le cadre d'un entretien d'embauche (<https://www.arte.tv/fr/videos/RC-014036/le-dessous-des-cartes/>) est-ce que vous connaissez l'émission le dessous des cartes d'Arte ?).

Arte a pour projet de créer des solutions d'IA ou de data dans les années à venir qui utiliseraient principalement des technologies open-source et qui à terme pourraient être utilisées par d'autres acteur.ices de l'audiovisuel (volonté de partage des technologies et des savoirs).

J'ai pris comme angle d'approche le fait d'avoir accès à des informations qui soient correctes factuellement et déontologiquement dans un but éducatif (de vulgarisation scientifique) pour pouvoir mieux informer les personnes qui regardent les vidéos sur les sujets qu'elles souhaiteraient apprendre.

J'aimerais pouvoir me baser sur un LLM d'un fournisseur français comme Mistral AI et faire du GraphRAG ou utiliser le principe du MCP pour communiquer avec des APIs comme celles de Wikipédia pour créer un chatbot scalable (en utilisant principalement les données provenant des sous-titres des vidéos) seulement voilà :

- **Comment peut-on prouver qu'une information générée par un chatbot est vraie ? (et pas une fake news) Est-ce que citer les sources suffit ?**
- **Comment peut-on prouver qu'une information générée par un chatbot n'a pas elle-même été générée par une IA générative ?**

- Est-ce qu'on peut créer des réseaux de données interopérable concernant des actualités politiques ? des ontologies ? est-ce qu'il en existe déjà ?  
- Dans un contexte de réponse dans différentes langues, est-ce qu'il vaudrait mieux que j'utilise une seule fois les sous-titre d'une vidéo pour les mettre dans une BDD vectorielle par exemple puis traduire le résultat ou à l'inverse utiliser les sous-titres dans tous les langages pour les mettre dans une BDD vectorielle pour avoir une réponse plus rapidement, pour avoir une réponse "mieux traduite" ?

Wikipédia est une bonne source de données typiquement mais n'importe qui peut en changer le contenu à n'importe quel moment ce qui en fait un danger pour la désinformation.

**Est-ce que vous auriez des pistes de réflexion ou d'autres personnes à contacter qui seraient à même de répondre à ces questions ?**

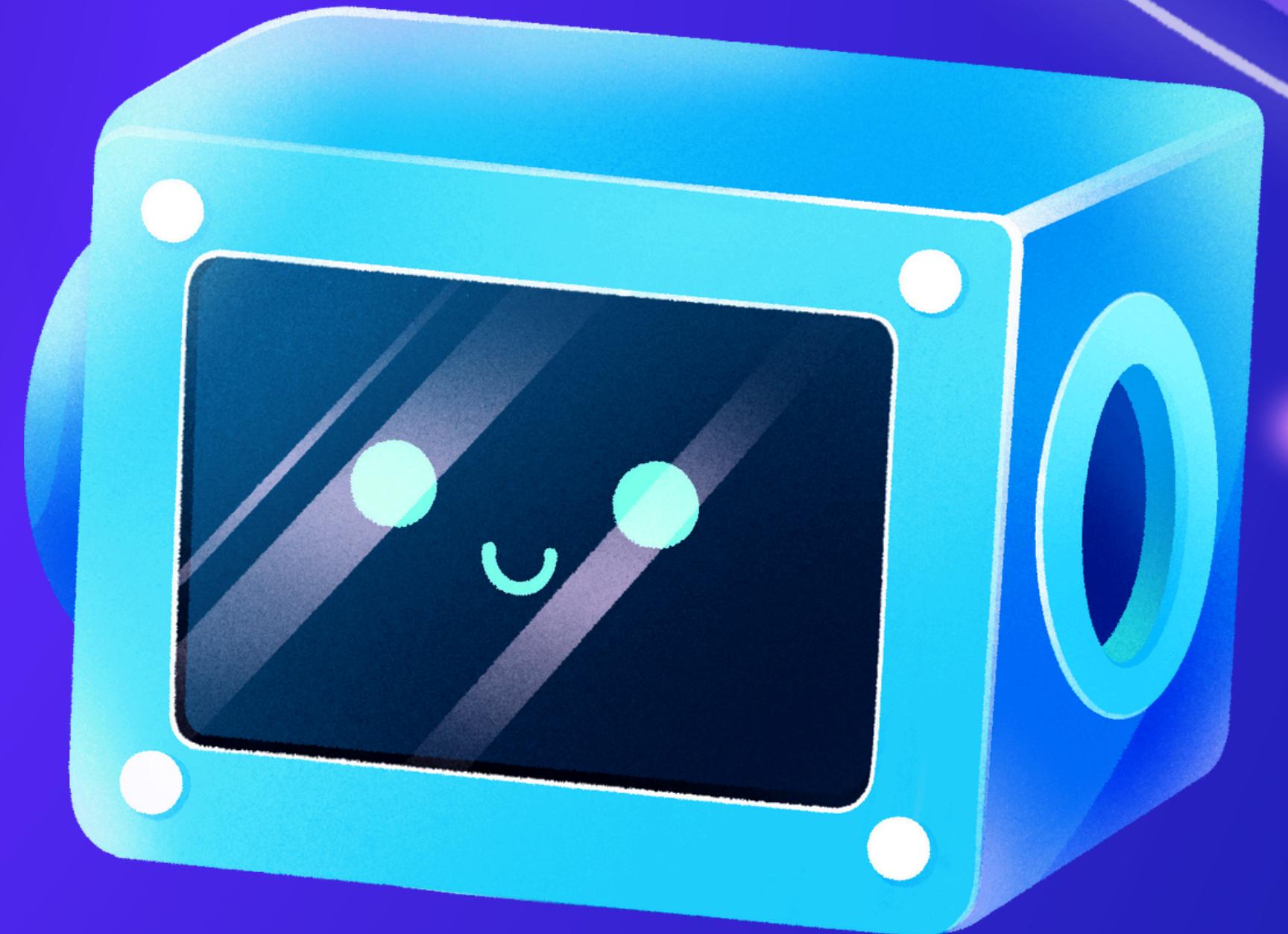
Merci beaucoup d'avoir pris le temps de lire ce message et merci par avance si vous y répondez !

Emilie Pacheco,  
AI and data engineer junior

# QUESTIONS POSÉES

Mails envoyés à l'équipe d'Artishau (Inria) qui travaille sur la fiabilité des IA et à certains membres de l'IRISA (équipe Kerdata – Systèmes sécurisés à large échelle)

ARCHI &  
DEV



# PREMIERE IDÉE D'ARCHITECTURE

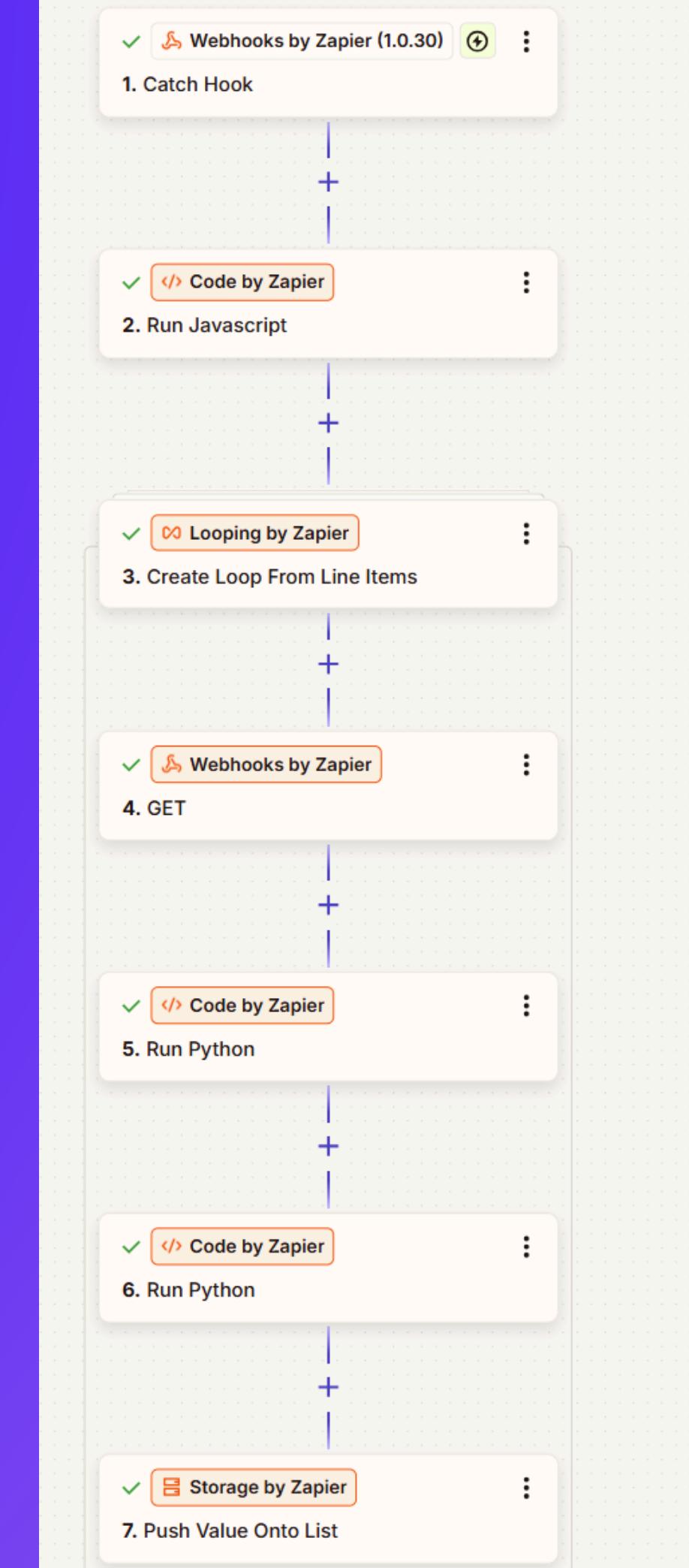


# WORKFLOW ZAPIER DE TRAITEMENT DES .VTT

1. Traitement qui prenait le fichier .json
2. Fetchait chacune des urls des sous-titres
3. Nettoyait les sous-titres (suppression de tous les caractères inutiles)
4. Créeait un webhook ou une API pour avoir accès aux données traitées

Problèmes avec le composant de looping ou lié au stockage des données durant le looping (les données n'étaient pas sauvegardées temporairement)

P



# WORKFLOW PYTHON DE TRAITEMENT DES .VTT

Système similaire mais en enregistrant les données dans un dataset Pandas puis dans un fichier .json en y ajoutant les métadonnées souhaitées avant d'être ingérées dans une vector db.

Par la suite création de chunks (parce que difficile à faire avec des outils no code) toujours en .json

Création d'un conteneur docker pour lancer le notebook Jupyter

# TESTS LANGFLOW

Tests avec Langflow en version hébergée par Datastax non concluantes (pb API keys en ligne, impossibilité pour moi d'héberger mes services sur un serveur)

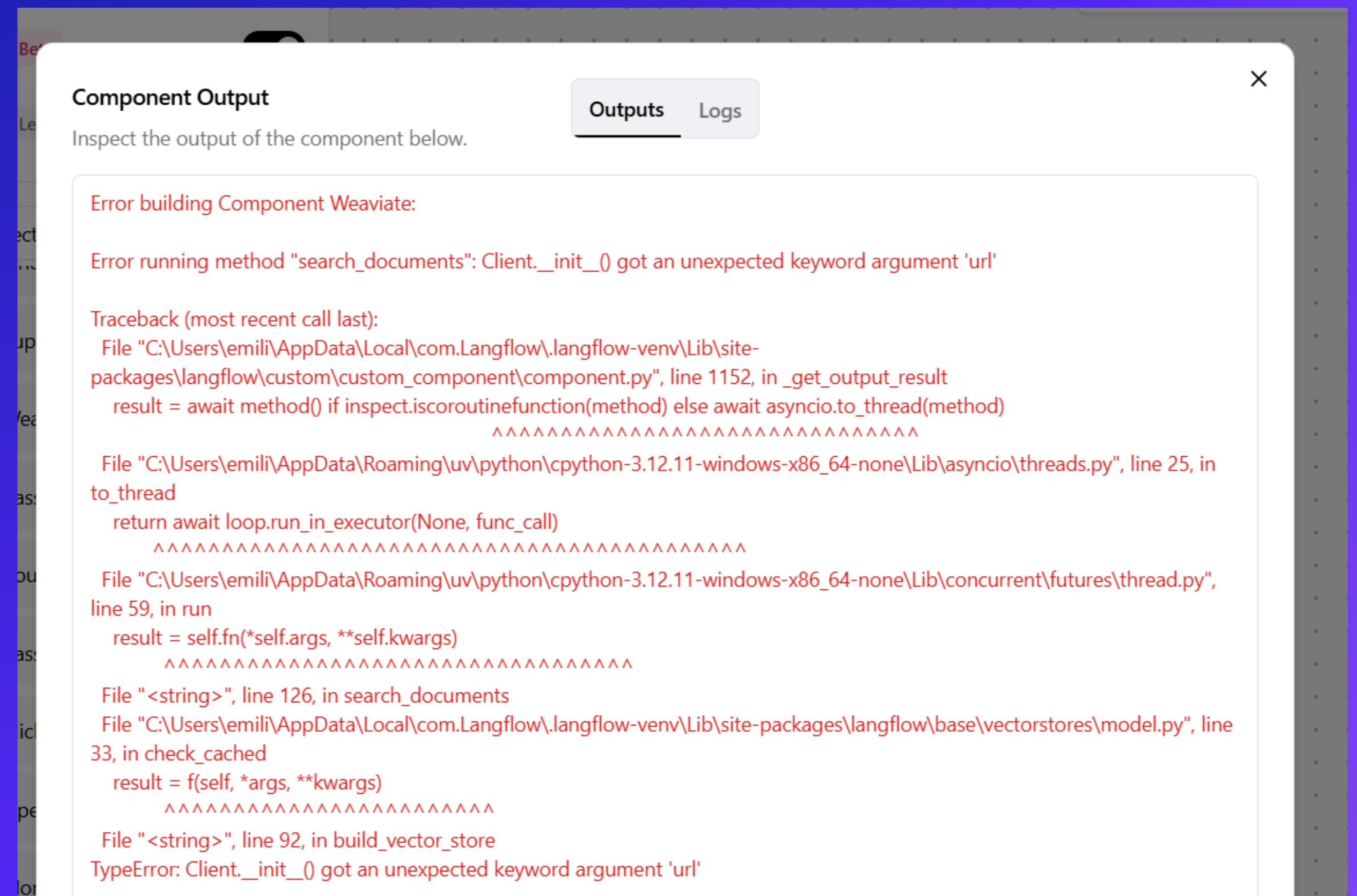
En local gros problèmes entre les composants anciens et nouveaux, nouvelles erreurs à tous les essais que ça soit pour utiliser un JSON au format classique ou sous forme de liste ou pour essayer d'accéder à une bdd vectorielle.

L'application sur PC Langflow rame beaucoup et bugue sur Windows

Langflow sous forme de code ?

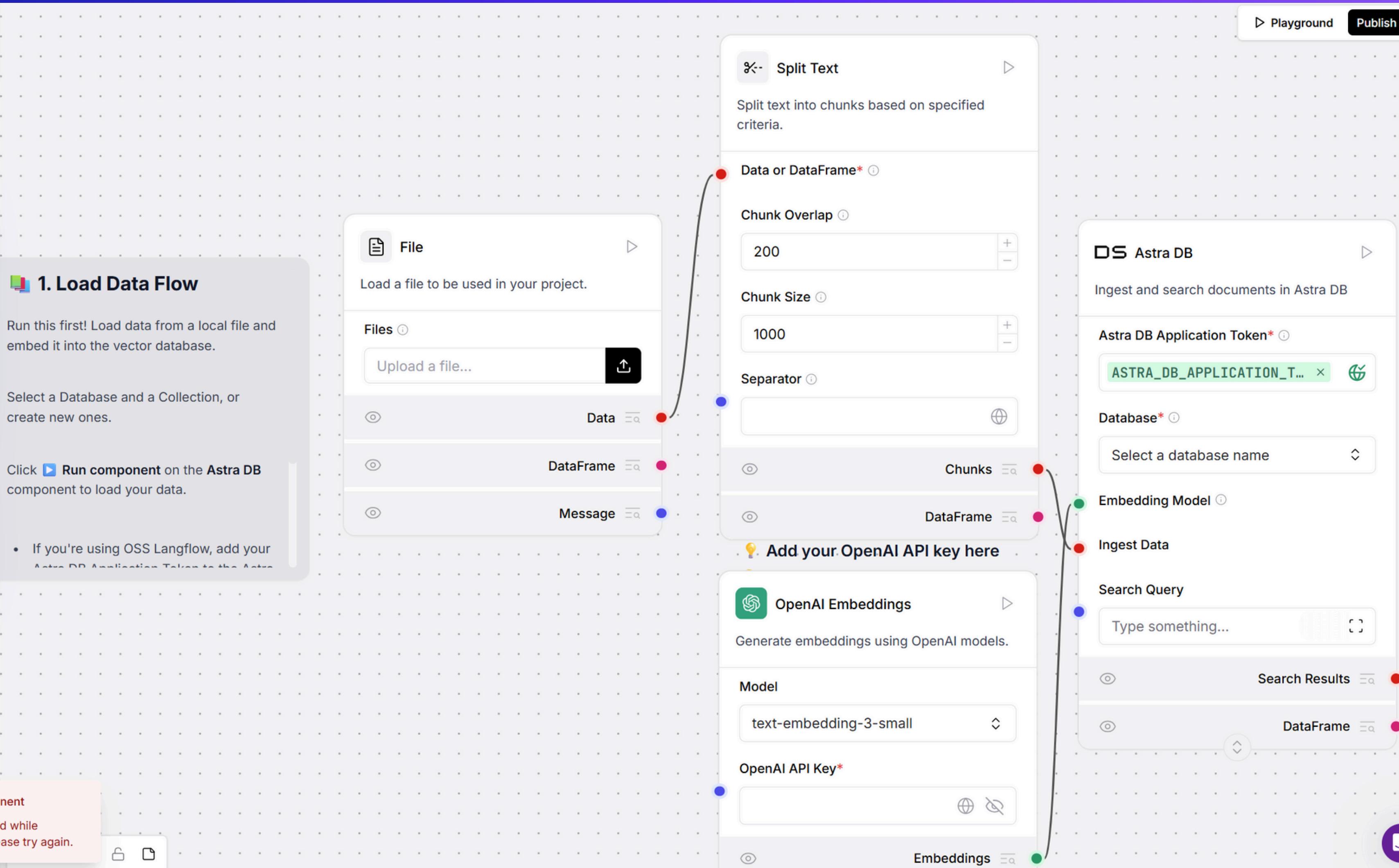
Installation de plusieurs vector db dans des conteneurs pour faire des tests (Weaviate, Qdrant)

# TESTS LANGFLOW

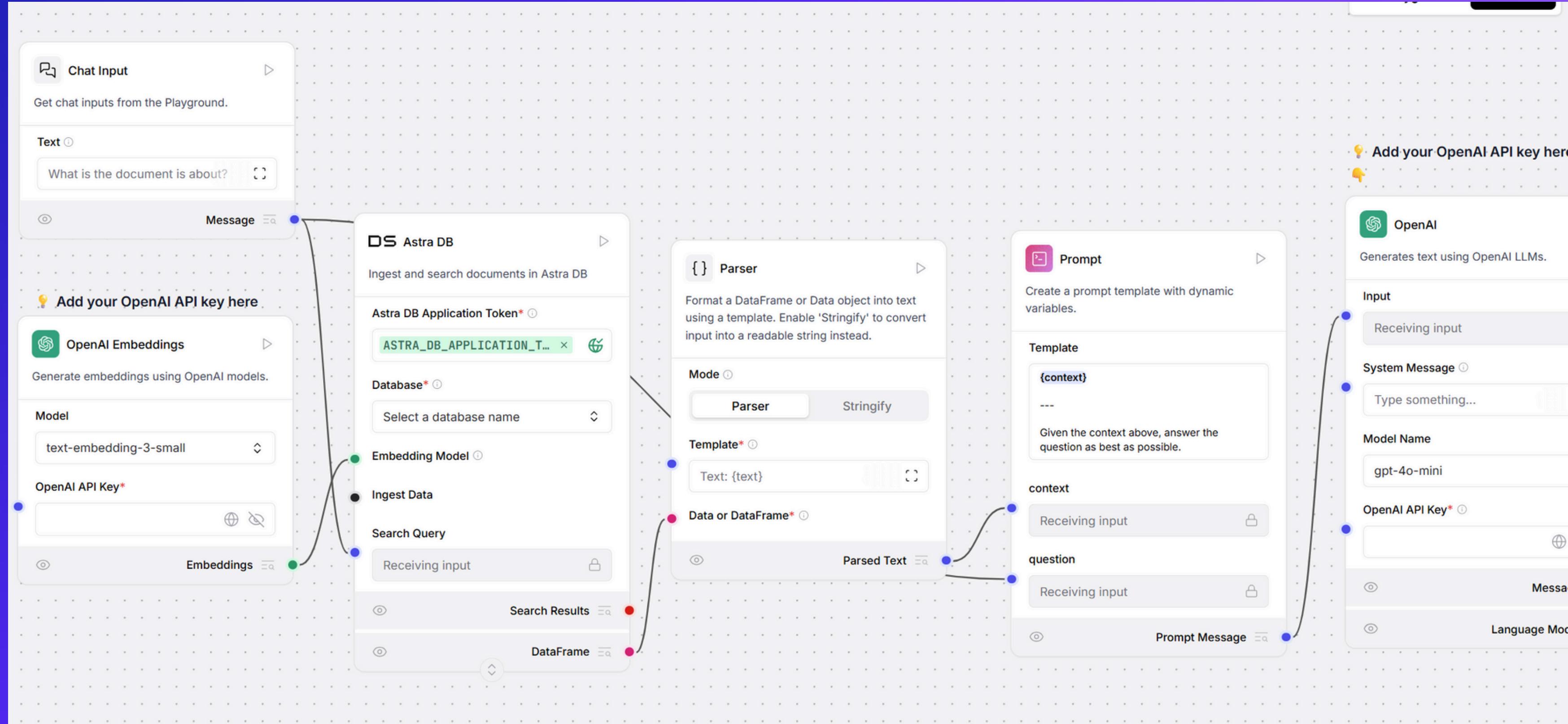


Problème de version de cli pour weaviate (présence d'un paramètre url dans les versions anciennes)

# TESTS LANGFLOW



# TESTS LANGFLOW



# ASTRA DB

Création d'une DB Astra et ingestion des données du fichier .json (au début tests en no-code avec des composants pour chunker) puis utilisation d'un fichier .json avec données déjà chunkées en python

ASTRA

DO

DO

DO

Keyspace *i*

arte\_test 1024 dimensions Cosine similarity 59 records

Load data

arte

Collections

- arte\_test
- poc\_arte
- subtitles

Create collection +

Search

Insert vector array [6.6463, 4.4971, 5.8337...]

+ Add filter

Cosine similarity scores *i*

Limit of 100

@1

@3

@10

@100

## Records

Table JSON

_id <i>i</i>	document_id	title	filename	chunk
b58435d4-958b-4084-8435-d4958be0844d	fc4e3ae0-2c58-488c-ae97-7949a0c7ba0b	Le dessous des cartes - Allemagne : changement d'époque	119961-017-A_st_VF-MAL.vtt	0
b3416058-0637-4e3e-8160-5806373e3e8d	baa21316-b31b-4aa8-ab0c-ea699fd2269a	Le dessous des cartes - La Pologne : nouveau leader européen	119961-018-A_st_VF-MAL.vtt	2
d7965659-623d-4cdf-9656-59623dbcd9d6	b0f2f027-d9c0-498a-b286-77586fae3fa3	Le dessous des cartes - Afrique : la France en berne	119961-021-A_st_VF-MAL.vtt	4
ed4d2f4a-9859-4a04-8d2f-4a98596a049b	1088e201-41ab-4d13-bd86-7d796988e52b	Le dessous des cartes - Économie russe : résiliente ou déclinante	119961-019-A_st_VF-MAL.vtt	1
ea8c3185-0d19-45d1-8c31-850d1995d104	70095962-804b-4104-b5ef-3d7630c68be6	Le dessous des cartes - Émirats arabes unis : comprendre le noi	123943-001-A_st_VF-MAL.vtt	10
b1429d35-072c-423a-829d-35072cf23a52	70095962-804b-4104-b5ef-3d7630c68be6	Le dessous des cartes - Émirats arabes unis : comprendre le noi	123943-001-A_st_VF-MAL.vtt	11
9e9733b0-2ce3-497e-9733-b02ce3097ea1	e86950d5-e72f-45f6-ba0d-65f2e69c28c7	Le dessous des cartes - La France et ses Outre-mer : des territo	119961-016-A_st_VF-MAL.vtt	3
59f5b7ad-c60a-4282-b5b7-adc60a528232	b0f2f027-d9c0-498a-b286-77586fae3fa3	Le dessous des cartes - Afrique : la France en berne	119961-021-A_st_VF-MAL.vtt	0
eaee0b4f-abe4-4772-ae0b-4fab43772d6	70095962-804b-4104-b5ef-3d7630c68be6	Le dessous des cartes - Émirats arabes unis : comprendre le noi	123943-001-A_st_VF-MAL.vtt	7
2becddff-b1fb-4068-acdd-ffb1fb90688c	9ebd68a9-2094-4e0e-8f39-2dff68adf8a	Le dessous des cartes - Smartphones : un enjeu de pouvoir	119961-015-A_st_VF-MAL.vtt	1
93fd0b75-fef1-4f30-bd0b-75fef13f30b8	70095962-804b-4104-b5ef-3d7630c68be6	Le dessous des cartes - Émirats arabes unis : comprendre le noi	123943-001-A_st_VF-MAL.vtt	8
5a370e05-f8fc-4f05-b70e-05f8fc7f053f	0c647994-2a6c-4436-99c5-e3005976725d	Le dessous des cartes - Le golfe de Guinée : un carrefour straté	119961-020-A_st_VF-MAL.vtt	3

Public Preview

# TESTS FLOWISE

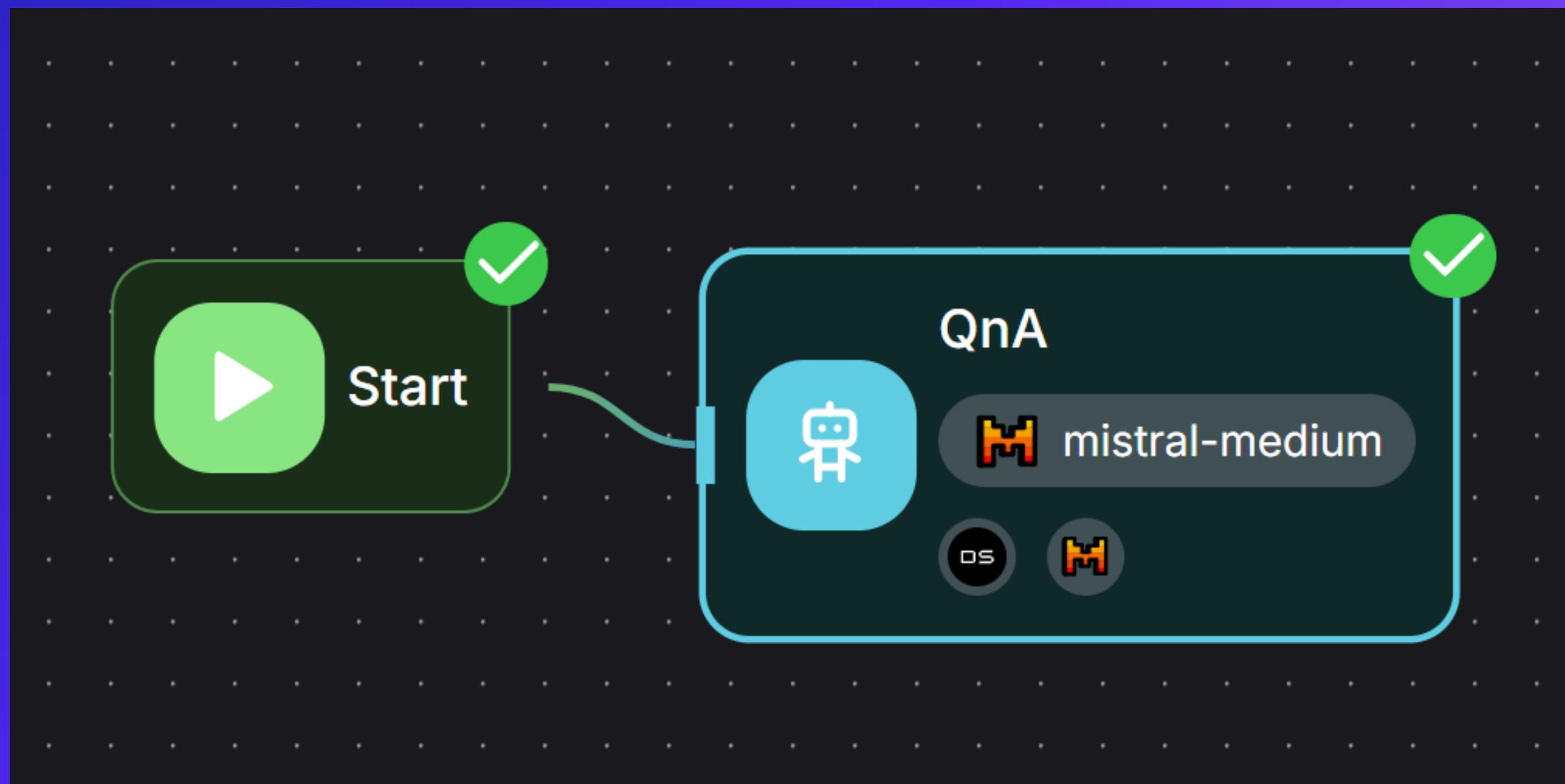
Problèmes de versions sur des templates fournis par l'app dans le cloud (v1, v2)

Clé Mistral qui ne fonctionnait pas

Installation en local pour tester le mcp de wikipédia mais aucun composant pour les connections SSE (streaming)

Problème de binding de port par l'app dans le conteneur docker (installation avec pip)

# TESTS FLOWISE



# TESTS FLOWTSE

Knowledge (Vector Embeddings) ⓘ

Vector Store \*

Astra

Astra Parameters

Embedding Model \*

MistralAI Embeddings

MistralAI Embeddings Parameters

Knowledge Name \*

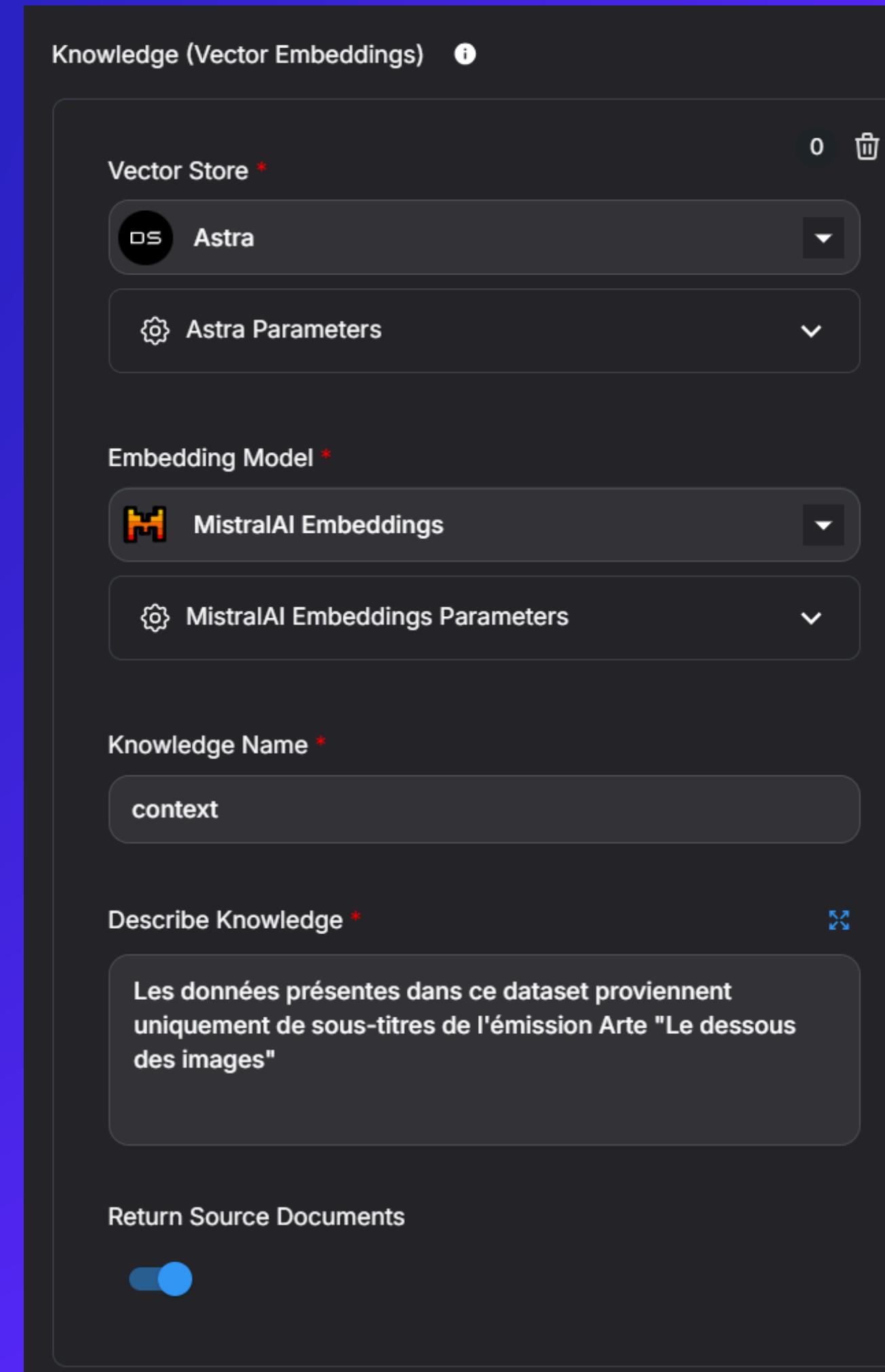
context

Describe Knowledge \*

Les données présentes dans ce dataset proviennent uniquement de sous-titres de l'émission Arte "Le dessous des images"

Return Source Documents

0 ⚡



Messages

Role \*

System

Content \*

(x) ⚡

Utilise les informations que tu trouveras dans ce contexte pour répondre.

Si aucune information pertinente n'est trouvée, dis

Role \*

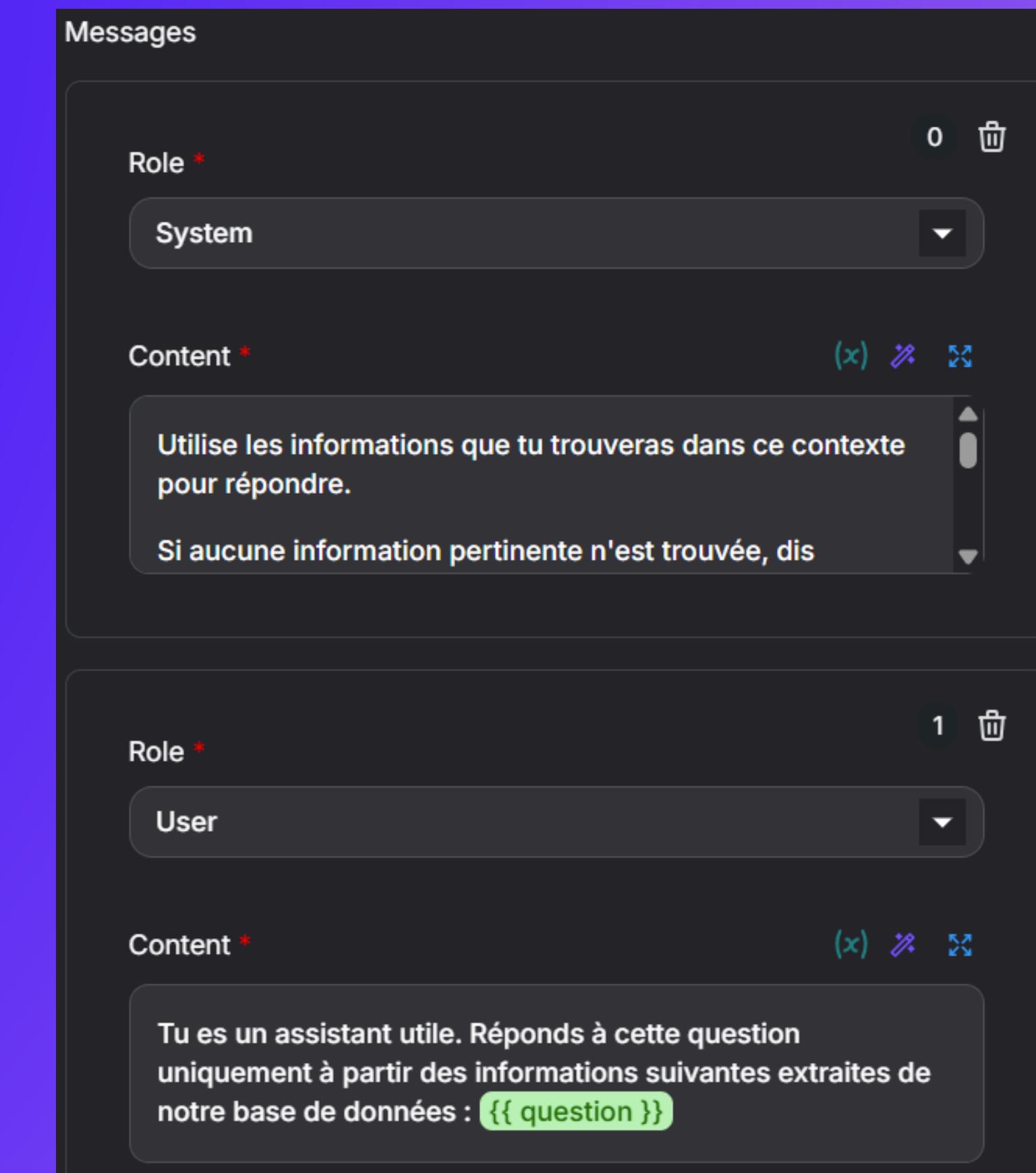
User

Content \*

(x) ⚡

Tu es un assistant utile. Réponds à cette question uniquement à partir des informations suivantes extraites de notre base de données : {{ question }}

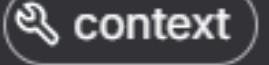
1 ⚡



# TESTS FLOWWISE

 Est-ce que l'Allemagne est le pays le plus puissant de l'Union européenne ?

  Process Flow 

 context

L'Allemagne est souvent considérée comme le pays le plus puissant de l'Union européenne en raison de sa forte économie, de sa population nombreuse et de son influence politique. Cependant, la notion de "puissance" peut varier selon différents critères tels que l'économie, la politique, la démographie, etc. D'autres pays comme la France jouent également un rôle significatif au sein de l'UE.

- Le dessous des cartes - L'Allemagne, une puissance européenne ?

 Est-ce qu'en 2013 le président français François Hollande a été à Bamako ?

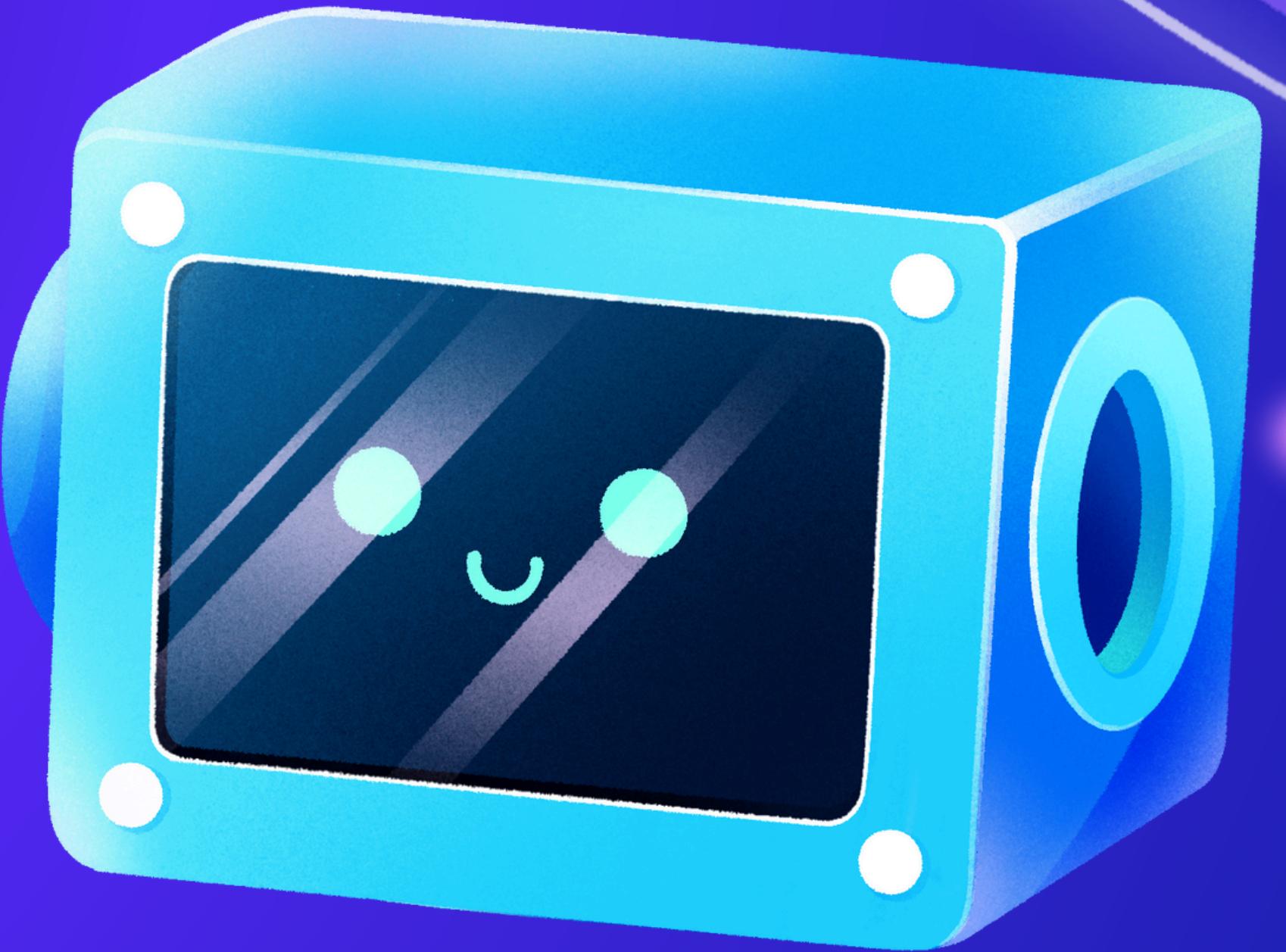
  Process Flow 

 context

Oui, François Hollande s'est rendu à Bamako en 2013.

- "Le 2 février 2013, François Hollande, président de la République française, est accueilli à Bamako par le président malien par intérim, Dioncounda Traoré."

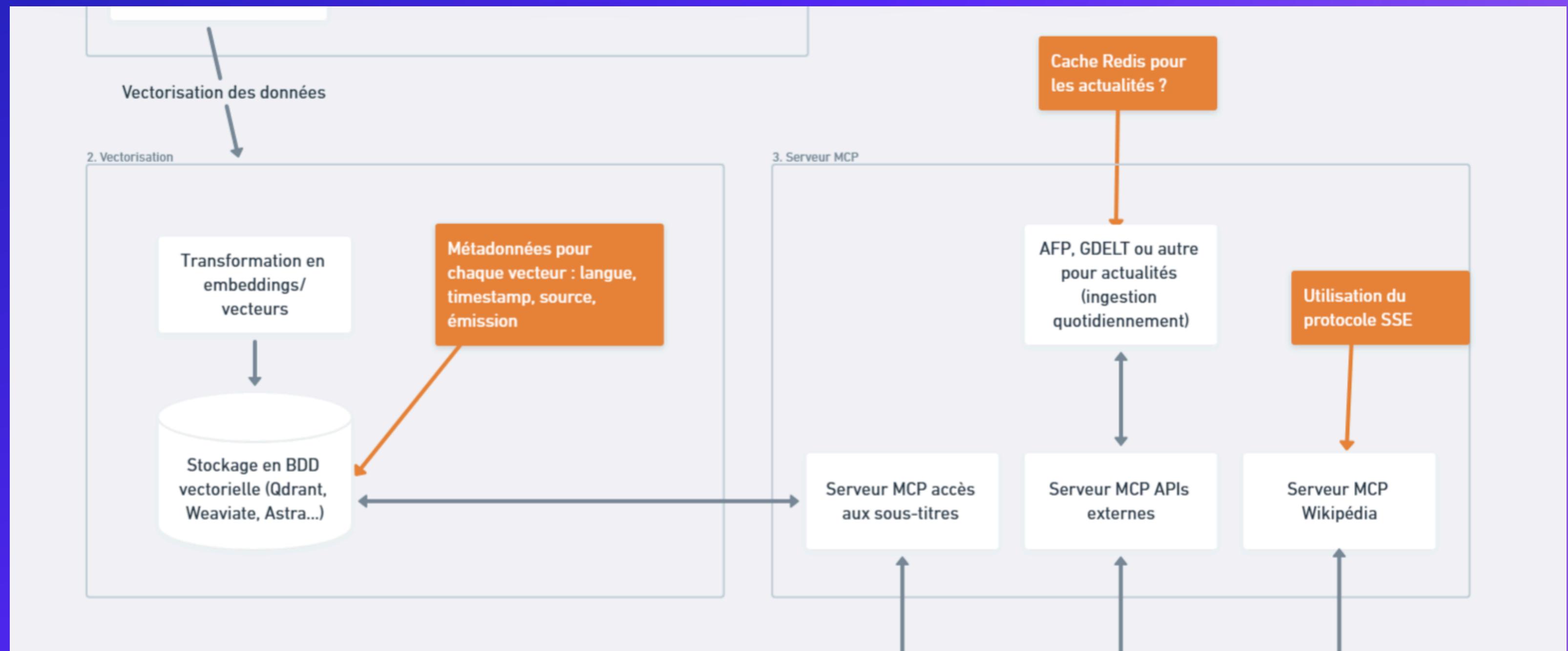
ARCHI  
FINALE



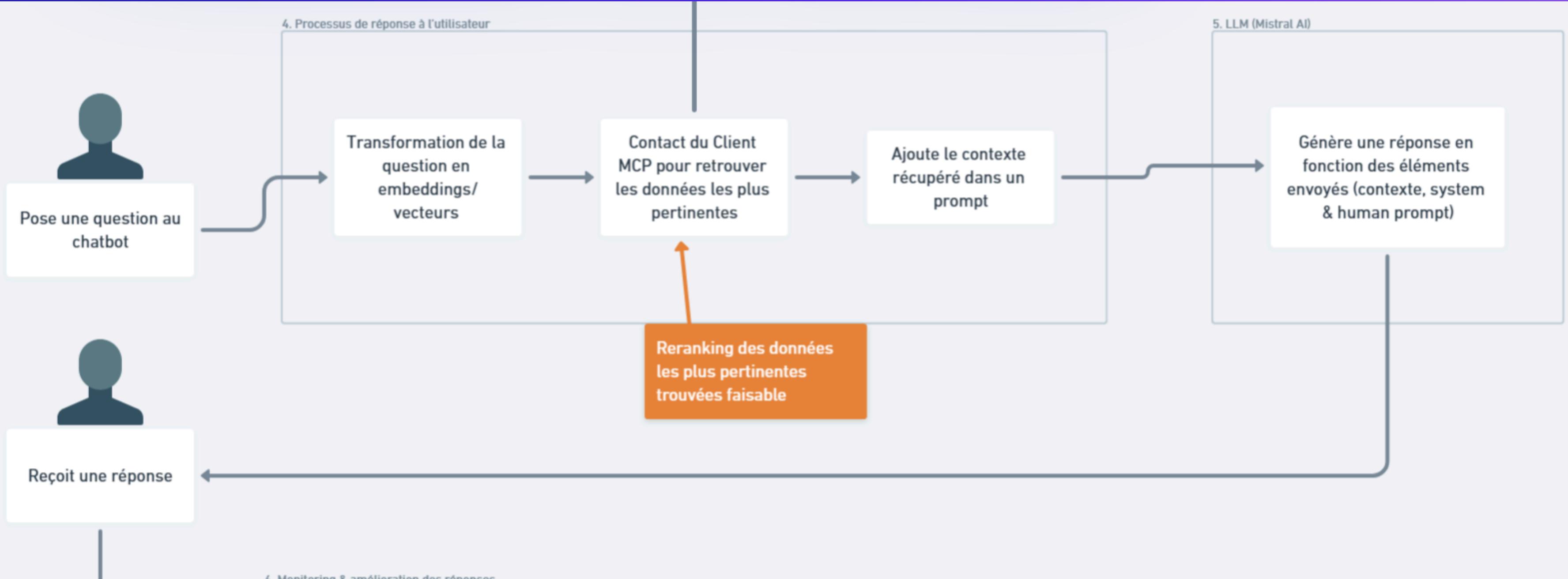
# TRAITEMENT DES DONNÉES



# VECTORISATION ET MCP



# PROCESSUS DE RÉPONSE



# MONITORING, MCO, OPS



MERCI

