

LLM-TALK

PHONE FEEDBACK AGENT

Réalisés Par :

- KACHMAR Mohamed
- EL OUALI HAMZA
- FAKRAOUI Ayoub
- Ennya Imane

Encadré Par :

- GAMOUH Hamza

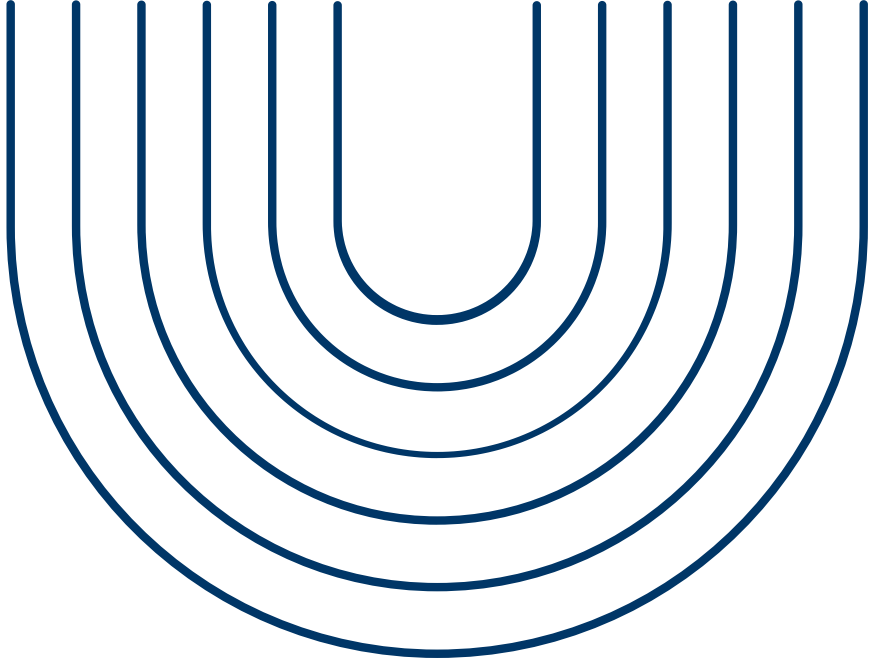
- 
- 01.** INTRODUCTION
 - 02.** TECHNOLOGIES UTILISÉES
 - 04.** ARCHITECTURE DU SYSTÈME
 - 07.** DÉMONSTRATION
 - 08.** CONCLUSION

TABLE DE MATIERE

01.

INTRODUCTION

Objectifs et Vision du Projet

OBJECTIFS ET VISION DU PROJET

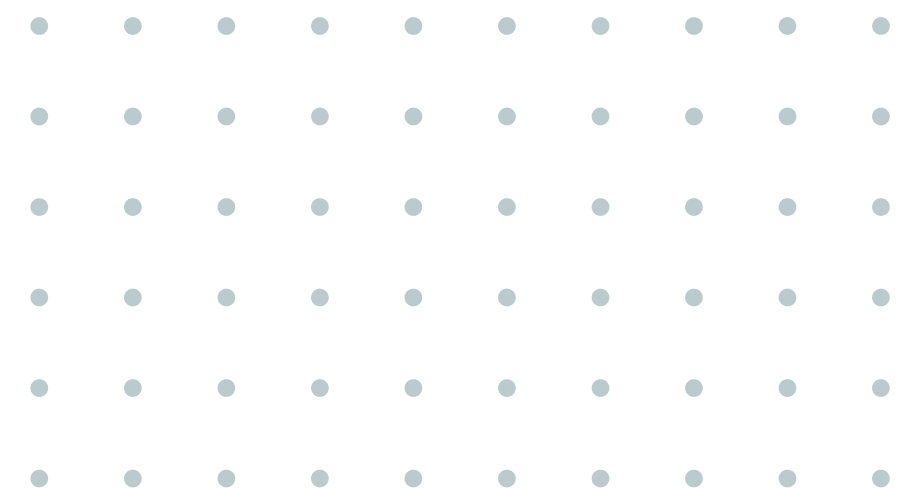
Notre projet vise à créer un système d'appels automatisés intelligents, combinant **LLM**, **TTS/STT** et **Hybrid RAG** pour mener des enquêtes vocales dynamiques.

L'objectif est d'offrir une solution adaptative capable de dialoguer, comprendre, et analyser les réponses en temps réel.

Nous aspirons à transformer les enquêtes classiques en interactions vocales intelligentes, modulaires et facilement déployables via Docker.

Le pipeline

- L'utilisateur crée une campagne (survey + KB)
- Données stockées dans MongoDB et Qdrant
- Twilio lance l'appel, le LLM orchestre la conversation.
- Réponses et analyses sont enregistrées



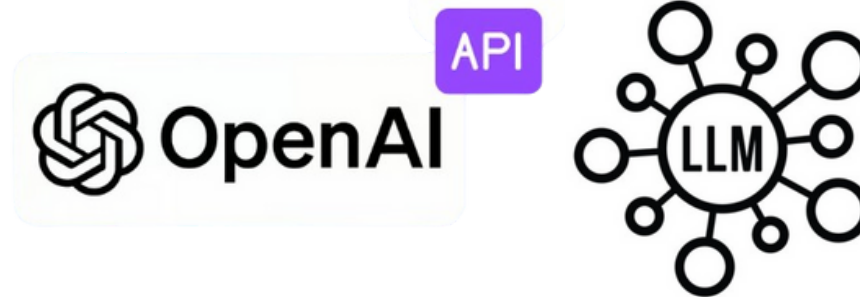
02.

ARCHITECTURE DU SYSTÈME

+ Stack Technologique

2/1. TECHNOLOGIES UTILISÉES

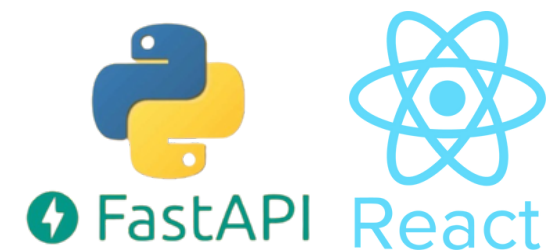
Couche Intelligence
Artificielle



Communication



Couche applicative



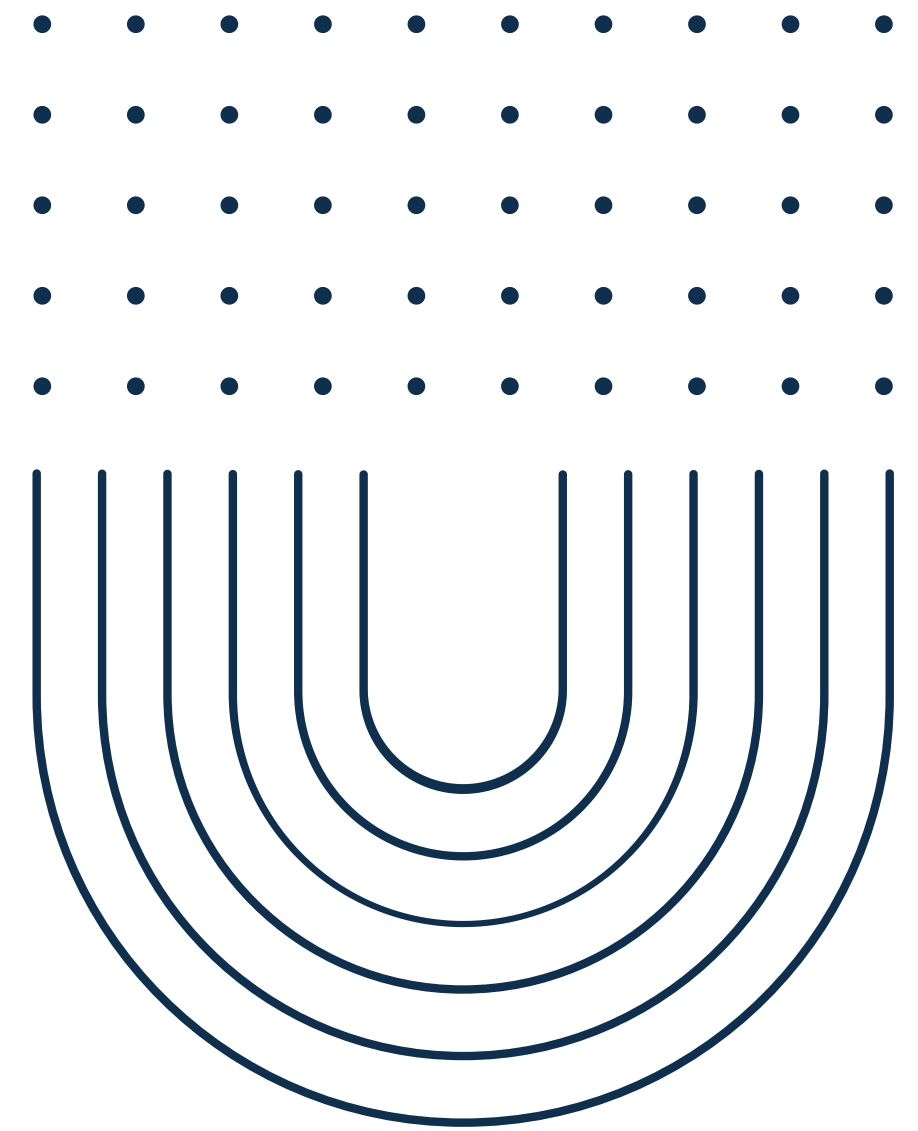
Stockage



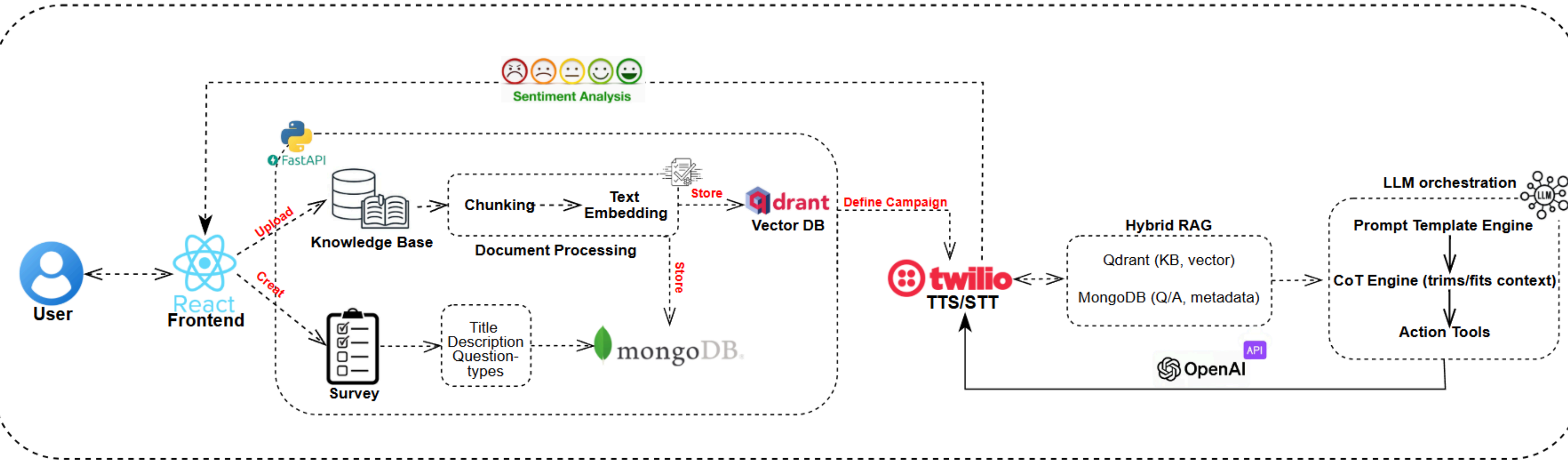
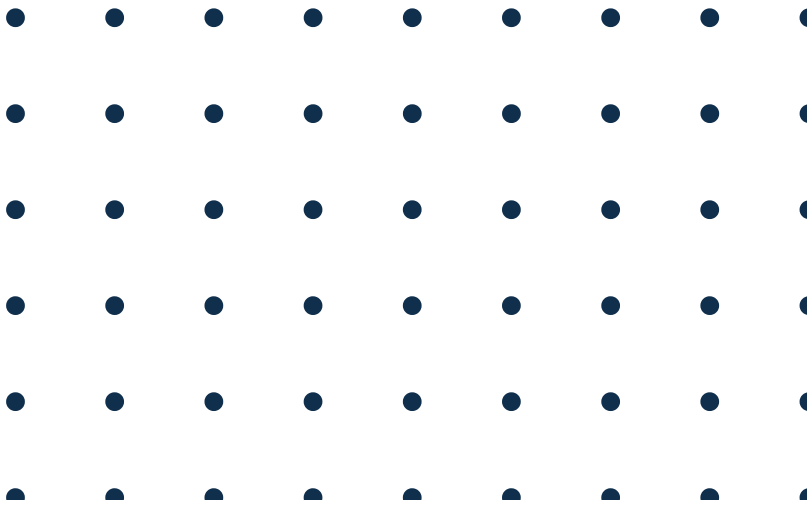
Deploiement



STACK
TECHNOLOGIQUE



2/2. ARCHITECTURE GLOBALE



ARCHITECTURE CONVERSATIONNELLE HYBRIDE AVEC RAG ET LLM

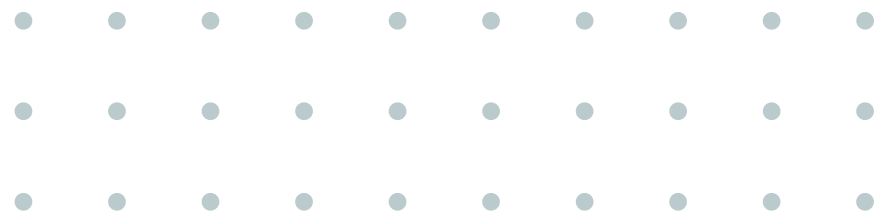
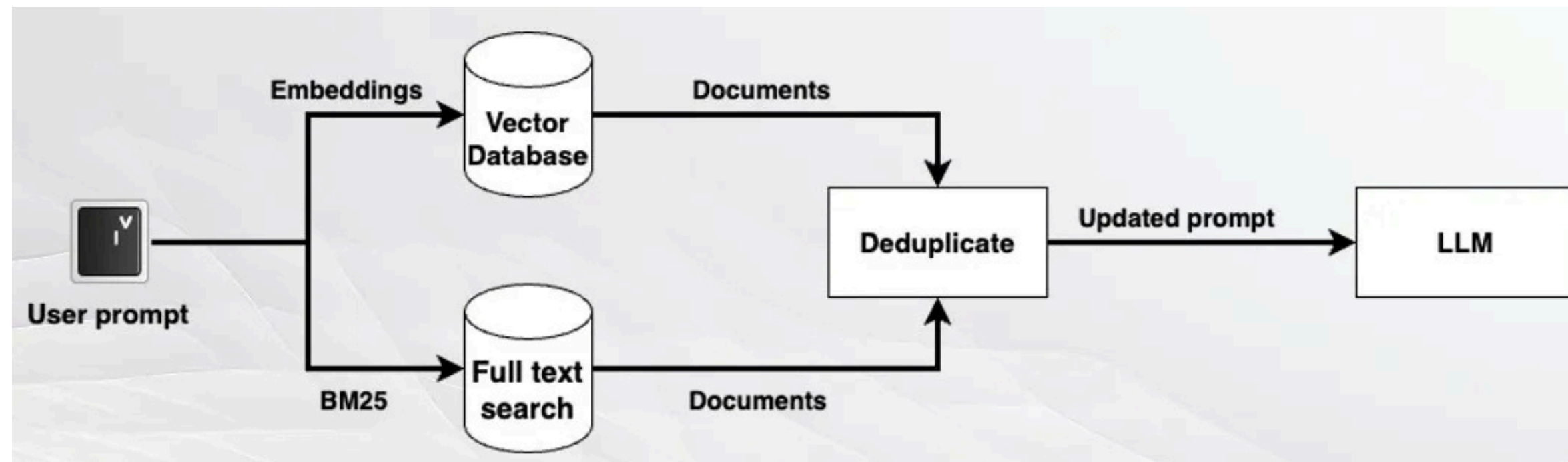
2/3. POURQUOI LE RAG ?

→ Pour fournir des réponses contextuelles, précises et enrichies à l'utilisateur pendant l'appel vocal, en combinant la puissance du LLM avec des données métier utiles.

Stratégies de récupération disponibles

L'algorithme choisit automatiquement une stratégie parmi :

- **SIMPLE** : recherche sémantique directe (vectorielle).
- **HYBRID** : combine recherche sémantique et par mots-clés.
- **CONTEXTUAL** : intègre l'historique de conversation pour adapter la requête.
- **ADAPTIVE** : choisit dynamiquement la meilleure stratégie selon le type de requête.



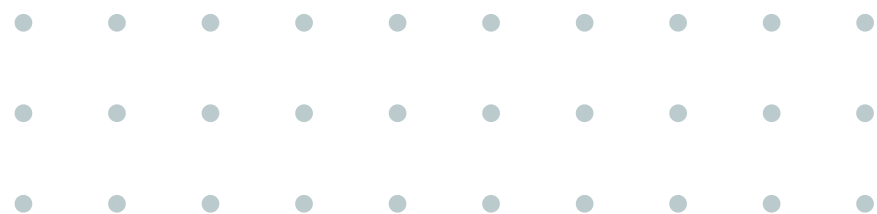
2/4. ANALYSE DE SENTIMENT – DÉFINITION & INTÉGRATION

L'analyse de sentiment permet d'évaluer automatiquement l'émotion exprimée par l'utilisateur dans sa réponse (positive, négative ou neutre), grâce à un modèle de langage (LLM). En cas d'échec, une méthode de secours basée sur des mots-clés est utilisée.

→ **Le score obtenu contribue à mesurer la satisfaction globale dans les enquêtes vocales.**

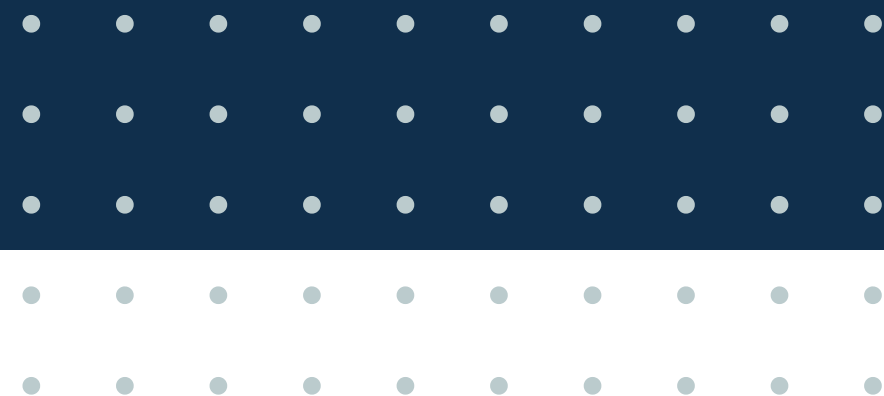
- **STT via Twilio : conversion voix → texte après la réponse utilisateur**
- **LLM (GPT-4) : analyse sémantique et scoring du ressenti**
→ Le score est basé sur : le ton émotionnel, le contexte, et les nuances du langage.

- **Fallback mots-clés : méthode de secours si le LLM échoue**
 - **Stockage MongoDB : score par question + moyenne globale**
 - **Usage métier : score exploité pour adapter le dialogue et mesurer la satisfaction**
- $$\text{Score} = \begin{cases} \min(0.8, 0.3 + 0.1 \times P) & \text{si } P > N \\ \max(-0.8, -0.3 - 0.1 \times N) & \text{si } N > P \\ 0.0 & \text{si } P = N \end{cases}$$



07.

DÉMONSTRATION



08.

CONCLUSION

OBJECTIF GLOBAL



Fournir une expérience conversationnelle automatisée, intelligente et évolutive, capable de s'adapter dynamiquement aux réponses utilisateur tout en capitalisant sur les données internes (KB, sondages).

PERSPECTIVES

Évolutions futures et extensions possibles de l'architecture. Intégration de technologies émergentes et amélioration continue des performances.

MACHINE LEARNING

Amélioration continue du scoring de sentiment via apprentissage supervisé sur les retours utilisateur.

INTÉGRATION OMNISCANALE

chatbot, WhatsApp, email, etc., via la même logique intelligente.



THANK YOU

*Special thanks to our Professor **Gamouh Hamza**
for all the efforts during this project.*

