

23 Mai 2025

# Mooc Project

Melody Duplaix, Dylan Chevallier, Arnaud  
Rambourg





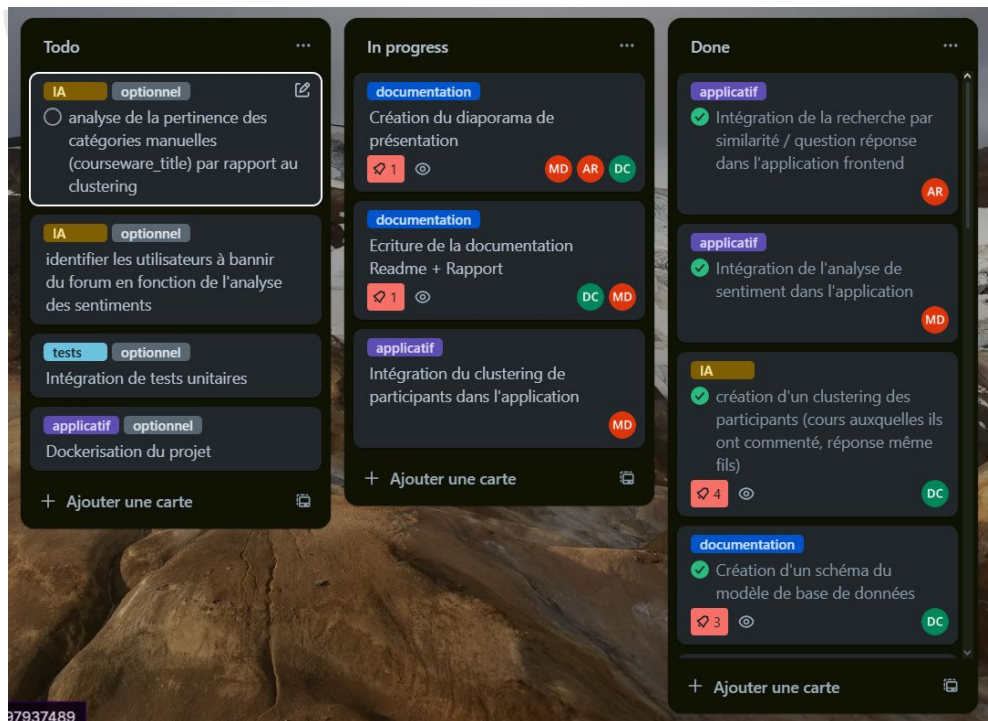
# Contexte, problématique et objectifs

Contexte : Avec la formation à distance et la gestion des MOOC, une très grande quantité de données est générée sous forme de fils de discussion sur les forums. Cette masse d'informations rend l'exploration et la recherche de contenus pertinents particulièrement complexe pour les utilisateurs

Problématique : Comment aider les nouvelles personnes et les anciennes à trouver les fils de discussion pertinents pour leur questions ?

Objectifs : Monitorer la formation, recommander des fils de discussion en fonction d'une question utilisateur

# Gestion du temps



→ création des schémas pour une bonne compréhension du sujet

→ création des tables dans une BDD locale avec ajout d'embedding + analyse de sentiments

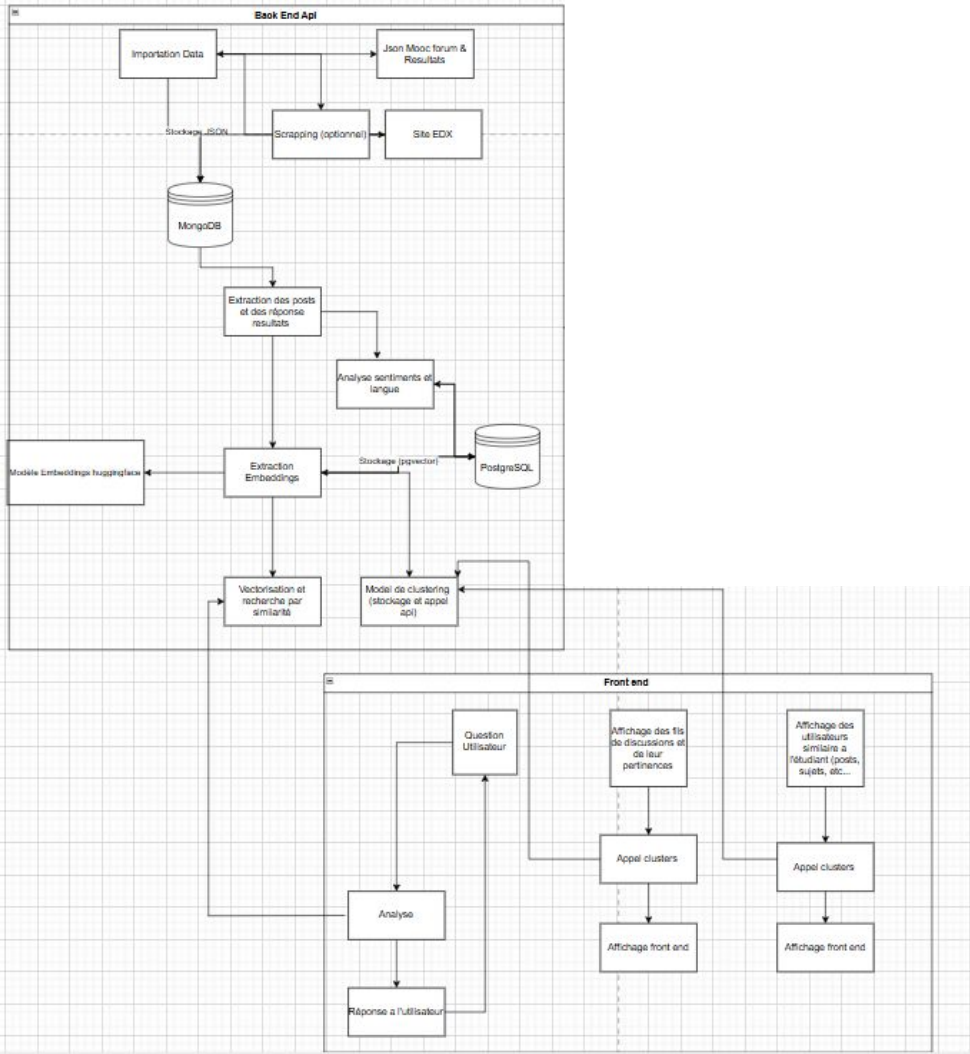
→ création clustering participants + messages

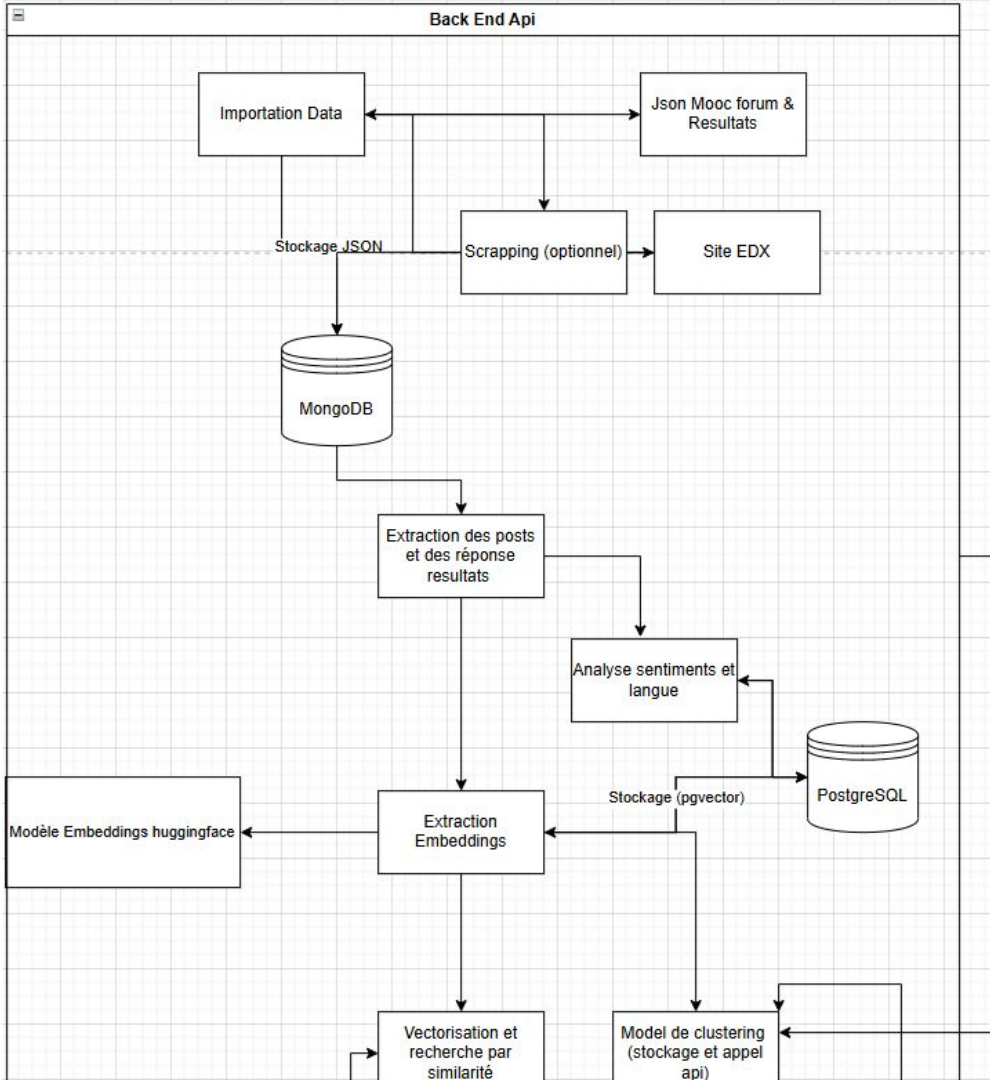
→ intégration dans API et application

# Schéma fonctionnel

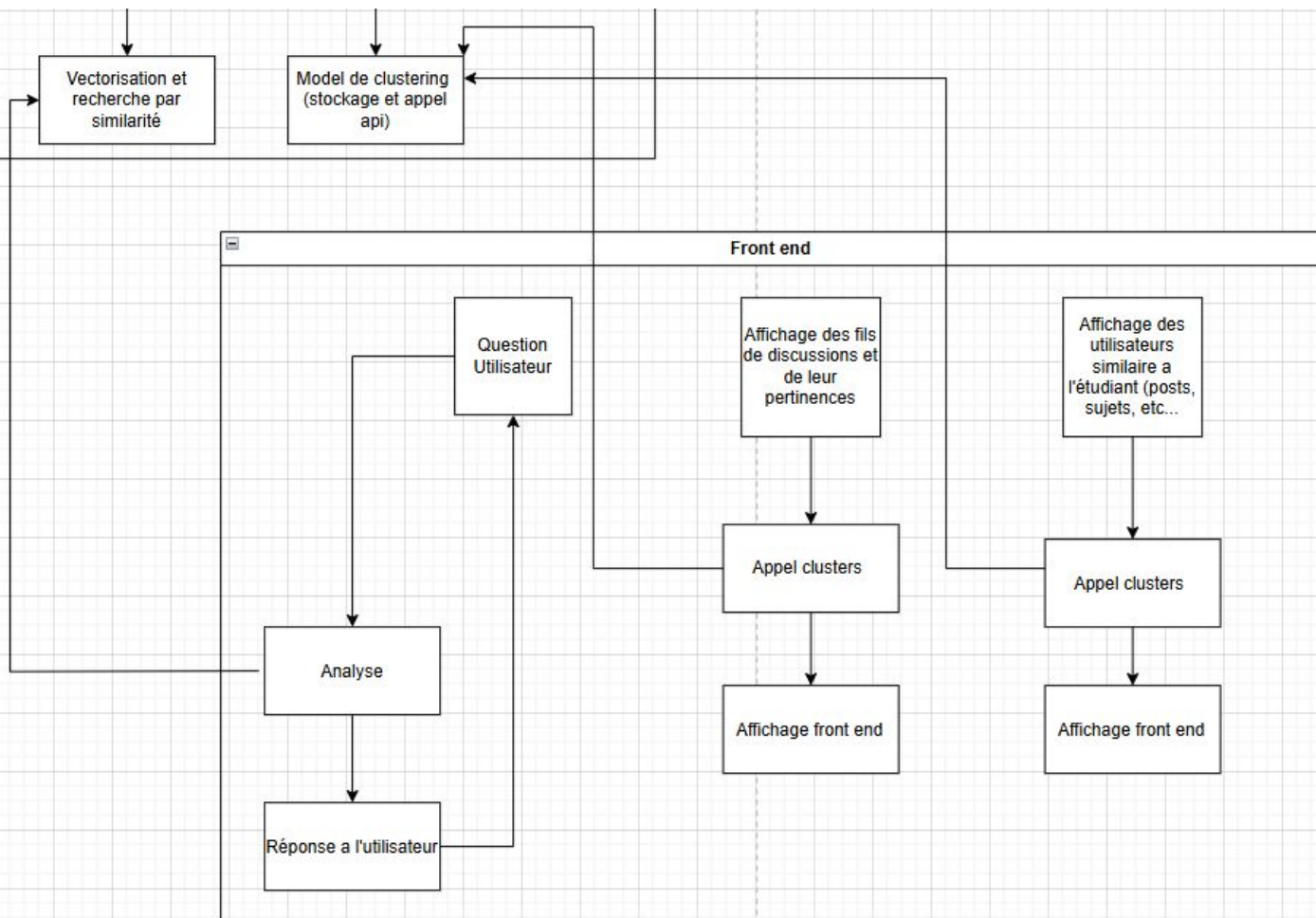
Partie 1 : back-END

Partie 2 : Front END





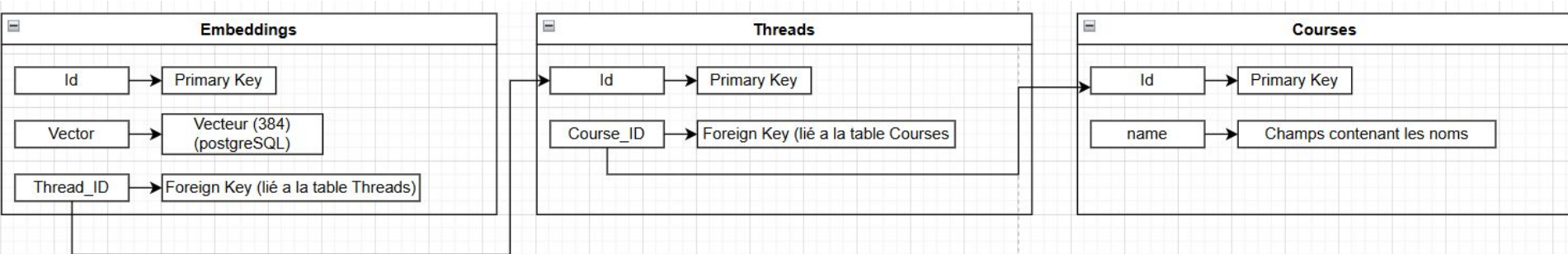
## Partie 1 : Back-END



## Partie 2 : Front-END



# MCD



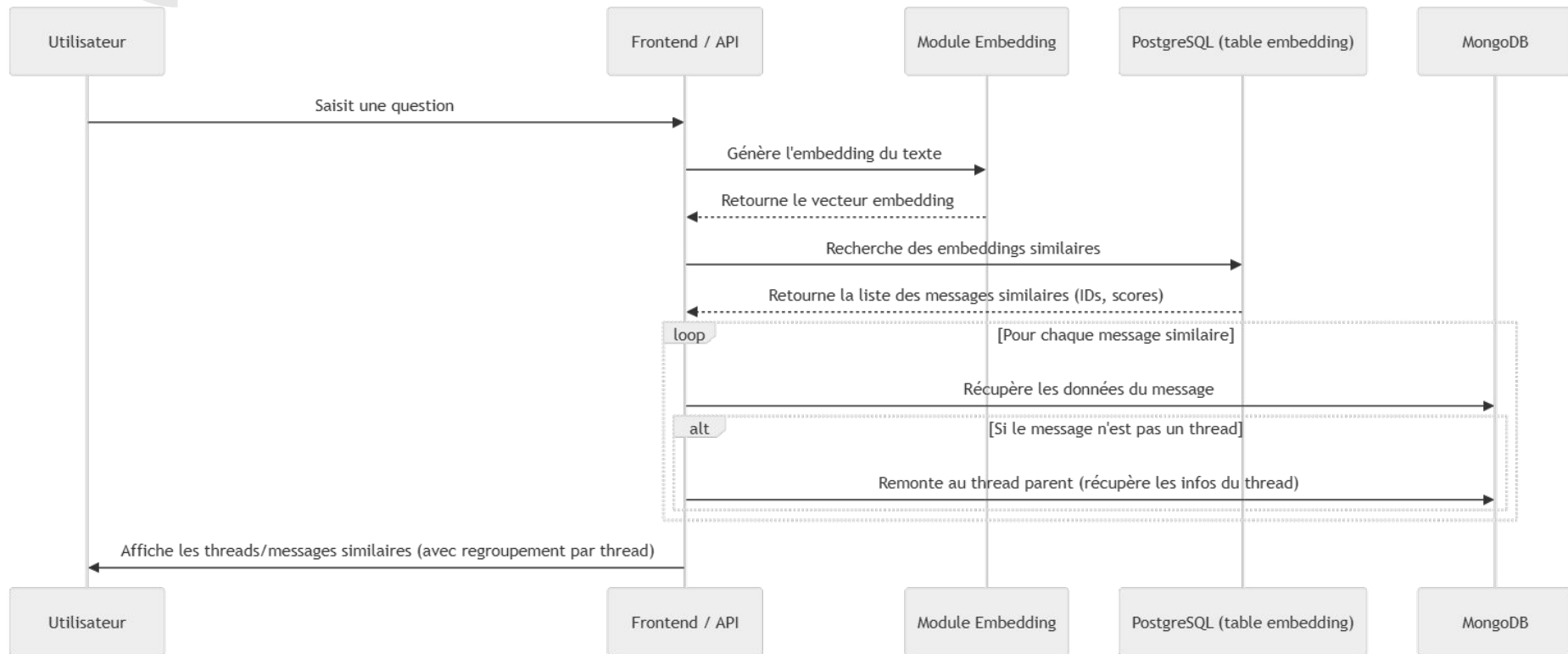


# Choix des modèles

- **Analyse de sentiment** : tabularisai/multilingual-sentiment-analysis (Hugging Face)
- **Embeddings** : paraphrase-multilingual-MiniLM-L12-v2
- **Clustering Messages** : BerTopic
- **Clustering Etudiants** : Kmeans



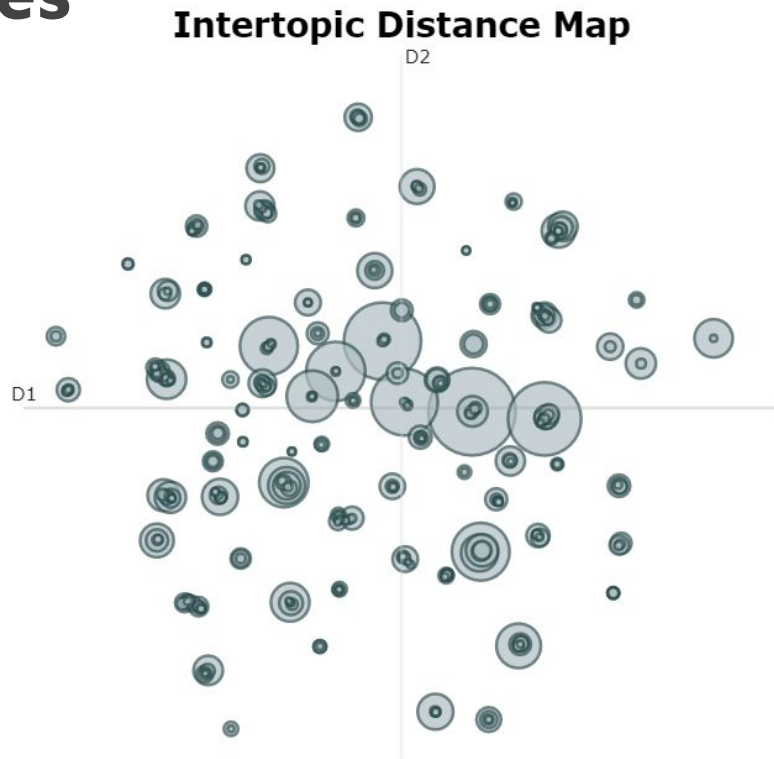
# Recherche de messages similaires





# Clustering des messages

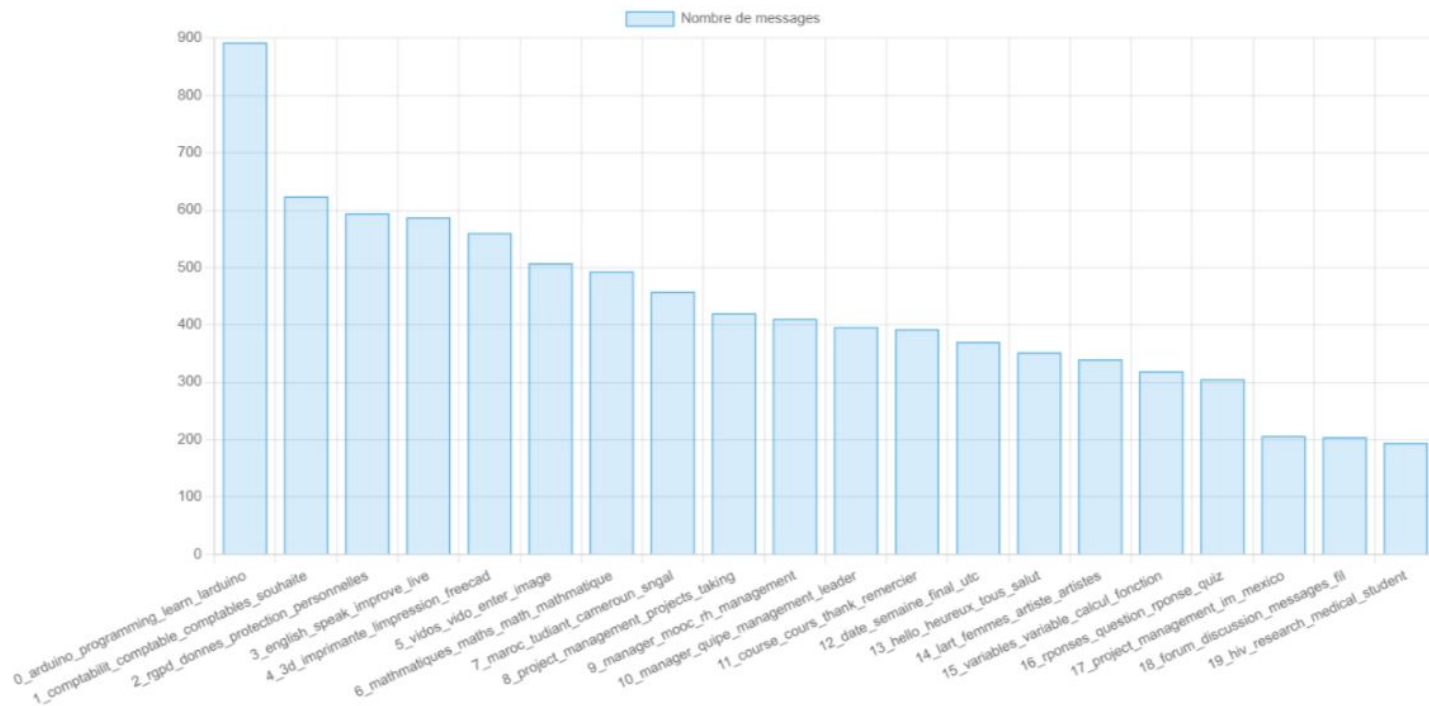
- Seulement sur les vecteurs de messages
- K-means → séparation des threads type présentation
- BERTopic → topic modeling



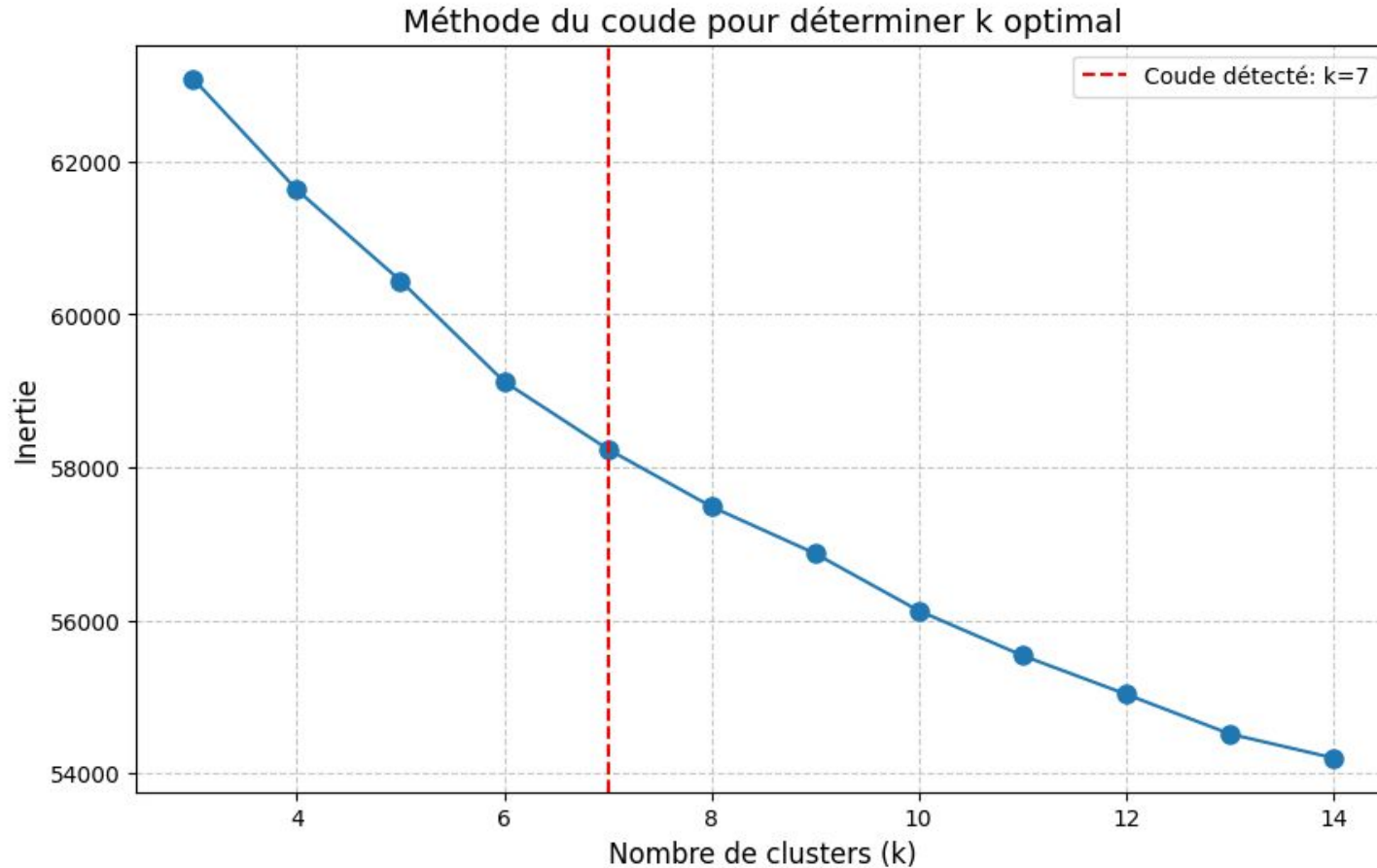


# Clustering des messages

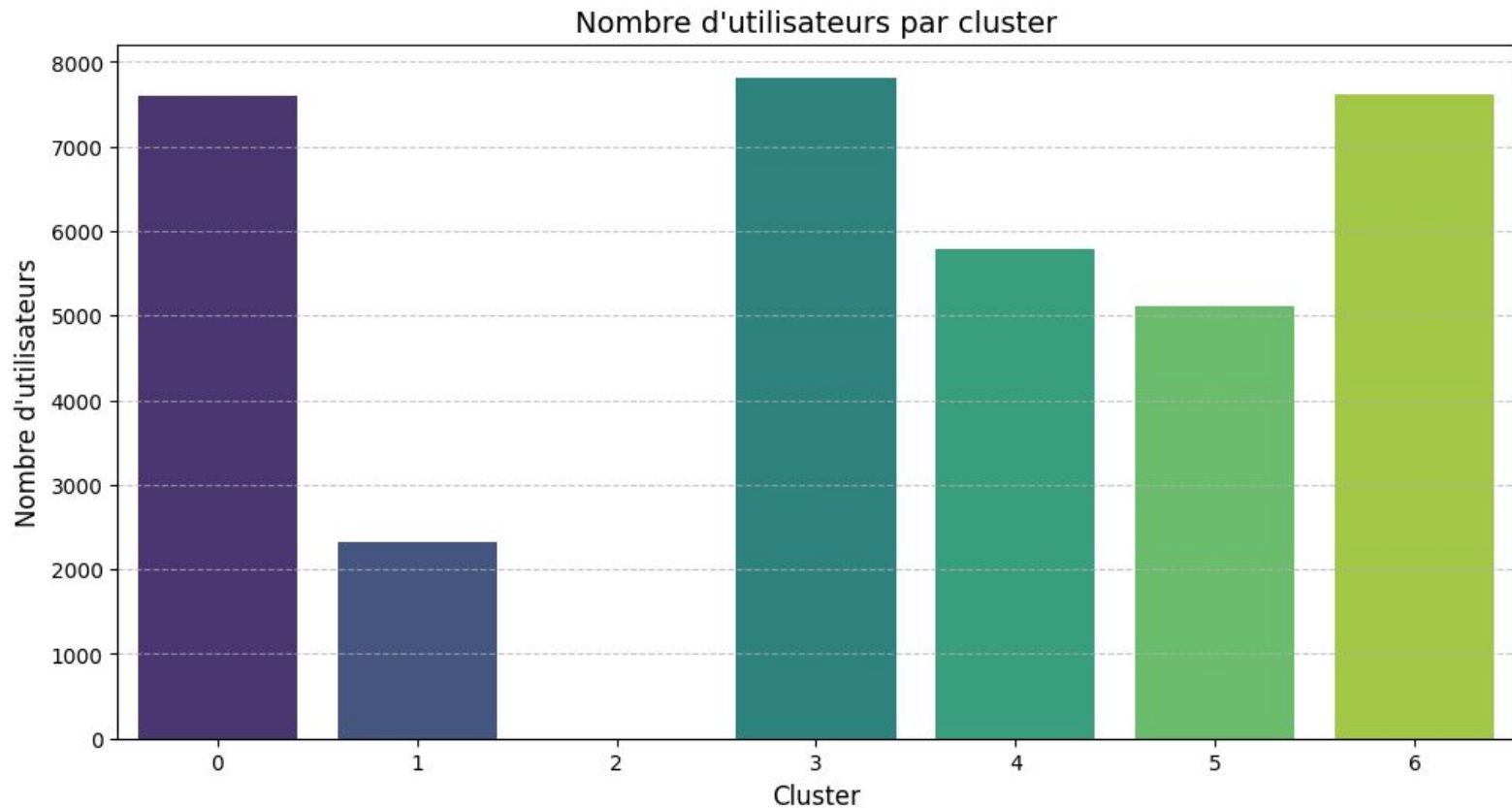
Histogramme des 20 Topics les plus fréquents



# clustering participants : Méthode du coude



# clustering participants : Répartition des utilisateurs par clusters





L'API du projet MOOC permet d'interagir avec les fils de discussion et les messages du forum. Elle est développée avec **FastAPI** et propose plusieurs points d'accès :

- **Récupération des threads** : obtenir la liste des fils de discussion
- **Recherche par ID** : afficher le contenu d'un fil à partir de son identifiant
- **Analyse de sentiments** : endpoint dédié pour analyser automatiquement le ton des messages (positif, neutre, négatif).
- **Ajout et recherche de messages** : possibilité d'ajouter ou de rechercher des messages similaires.



# application

L'application permet de **rechercher et d'analyser les fils de discussion** issus des forums de MOOC. Elle propose une interface web pour :

- **Poser** une question et visualiser les messages s'en rapprochant
- **Rechercher** un fil par son titre et visualiser ses messages et leurs sentiments
- **Visualiser** les statistiques sur les clusters de messages et participants
- **Visualiser** les topics et en choisir un pour parcourir ses messages
- **Chercher** un utilisateur et visualiser des utilisateurs similaires



**maquette figma**

FIGMA





# demo

## FUN MOOC

Application de suivi et d'analyse des forums

Poser une question

Fil de discussion

Regroupement des fils de discussion

Regroupement des participants au MOOCs

### Poser une question

Posez votre question ou recherche :

Rechercher

#### Résultats de la recherche :

User ID	Threads	Message
	<a href="#">apprendre</a>	je m'appel agathe je souhaite développer mes compétences c
21796	<a href="#">Francois de Lyon</a>	Bonjour, Je travaille en agence et j'ai la possibilité de prend
119410	<a href="#">De manager au leader</a>	Bonjour, Pour devenir un leader, le manager doit évoluer so
449973	<a href="#">Objectif</a>	Après 20 années d'expérience en management, je souhaite p
661827	<a href="#">Présentation</a>	Bonjour, je m'appelle Fares. Je suis ce mooc car je souhaite
1769797	<a href="#">Etre au fait des meilleures méthodes de management</a>	Bonjour, Je me suis inscrit dans ce Mooc car je souhaite être
3155213	<a href="#">Présentation</a>	Bonjour, je m'appelle Christelle. Je suis ce mooc car je soute
3832368	<a href="#">devenir leader : un chemin de vie</a>	devenir leader pour le manager est une évolution complète c
4726053	<a href="#">Présentation</a>	Grandement intéressé par le management, je me suis inscris
7392026	<a href="#">Présentation</a>	Bonjour , je suis développeur informatique et je voudrais se

Affichage de 1 à 10 sur 10 entrées

Précédente

1

Suivante

[Retour à l'accueil](#)

### Regroupement des fils de discussion

Nombre de topics:  
214

Nombre de messages:  
26222

Moyenne nb messages par topic:  
122.53

Médiane nb messages par topic:  
31.5

#### Tableau des Topics

Afficher 10 entrées		Rechercher :	
ID	Nom	Mots clés	Nombre de messages
0	<a href="#">0_arduino_programming_learn_larduino</a>	arduino, programming, learn, larduino, programmation, im, hi, want, use, darduino	892
1	<a href="#">1_comptabilite_comptables_comptables_souhaite</a>	comptabilite, comptable, comptables, souhaite, notions, connaissances, gestion, financiere, bases, mappelle	624
2	<a href="#">2_rapd_douzes_protection_personnelles</a>	rgpd, douzes, protection, personnelles, conformit, rglementation, dpo, mise, sein, clients	594
3	<a href="#">3_english_speak_improve_live</a>	english, speak, improve, live, im, like, france, want, uderstand, years	587
4	<a href="#">4_3d_imprimante_l'impression_freecad</a>	3d, imprimante, l'impression, freecad, modélisation, pices, fablab, imprimantes, fabrication, imprimer	560
5	<a href="#">5_videos_video_ester_image</a>	videos, video, ester, image, video, description, violence, tchager, videos, violences	507
6	<a href="#">6_mathematiques_maths_math_mathematique</a>	mathématiques, maths, math, mathématique, niveau, remettre, revoir, notions, am, bac	493
7	<a href="#">7_maroc_tadiant_cameroon_sagat</a>	maroc, tudiant, cameroon, sagat, congo, ivoire, sahat, nationalit, mappelle, connaissances	458
8	<a href="#">8_project_management_projects_taking</a>	project, management, projects, taking, course, career, looking, learn, forward, work	420
9	<a href="#">9_manager_mooc_rh_management</a>	manager, mooc, rh, management, competences, humaines, gestion, leader, formation, quipe	410

Affichage de 1 à 10 sur 213 entrées

Précédente

1

2

3

4

5

...

22

Suivante

Dernière



# Pistes d'améliorations

- Enrichissement des données pour avoir des données plus détaillées (notes, informations utilisateurs)
- Mise en place d'un système de RAG (Intégration d'un LLM)
- Intégration de tests unitaires et d'intégration
- Dockerisation du projet
- Amélioration des clusterings
- Intégration complète des fonctionnalités dans l'API



**Merci pour votre écoute**  
**Avez-vous des**  
**questions?**