

# Data Battle



PAU · PARIS

HERAS ESPINOSA Alejandro

PAPINI Julien

DE PONTON D'AMECOURT Marc-Henry

# Le sujet

Création d'une plateforme de révision pour des étudiants en propriété intellectuelle.

Les points importants :

Explicabilité

Facile d'utilisation

Écologique



# Notre solution



Plateforme de quiz  
interactive

QCM avec justification

Facile à prendre en  
main

# Questions existantes

Questions et réponses fiables

QCMs avec explication des réponses

Différents formats de fichiers

Identification du thème de la question

Minimiser la création de question



# Création des questions : extraction



Passage des données PDF à Markdown

Nettoyage des données inutiles

Meta données sur les chunks :  
Titre, section, références ...

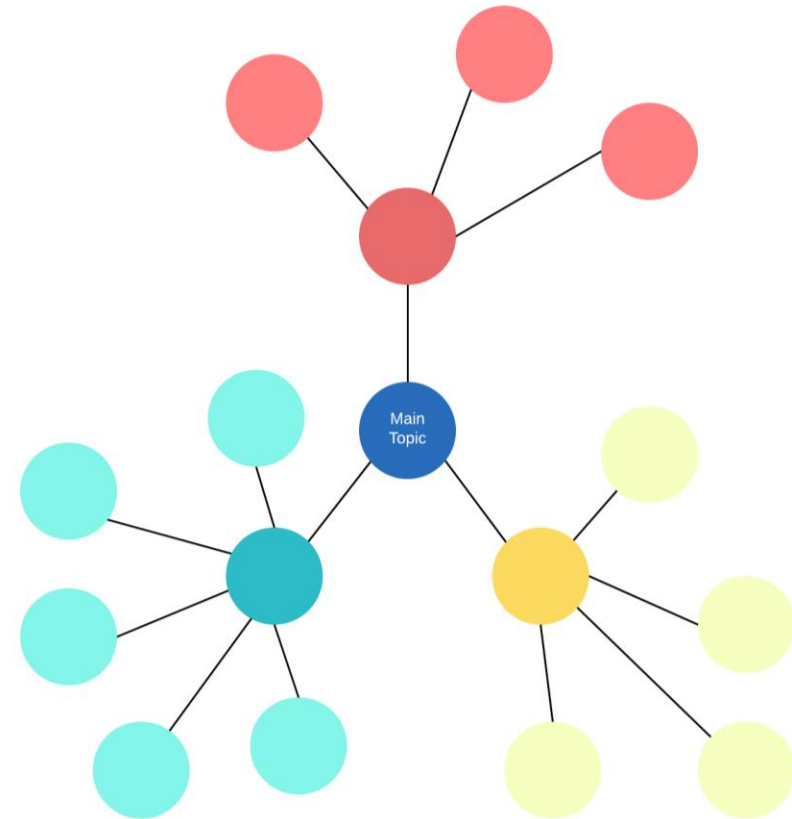
# Création des questions : stockage des chunks

Stockage des vecteurs et chunks dans Weaviate

Embedder utilisé : bge-small-en-v1.5

Attribution de thème(s) à chaque chunk

Création de clusters de plusieurs chunks pour chaque thème



# Création des questions : génération

Génération de questions par lots

Choix du thème lors de la génération

Basé sur 3 chunks aléatoires:  
explicabilité et aléatoire

Création des questions en avance  
pour vérification par un professeur

Modele : mistral-7b-instruct



# Création des questions : stockage des questions



Stockage des questions générées pour économiser des ressources

Stockage des sources

Base de données : Firebase



# Quiz

10 questions aléatoires selon un thème choisi

QCM avec justification

Évaluation de la réponse via NLP

Explication via un LLM si la réponse donnée n'est pas correcte

Modèle : all-MiniLM-L6-v2  
mistral-7b-instruct



# Impact environnemental

Minimisation des requêtes aux modèles :

Stockage des questions

LLM seulement si la réponse donnée n'est pas correcte

Petits modèles

**mistral-7b-instruct**

 **Energy**


**0.916 Wh**

*Evaluates the electricity consumption*

 **93 Paris ↔ NYC**

*Based on GHG emissions*

**chatgpt 4o**

 **Energy**

**21.9 Wh**

*Evaluates the electricity consumption*

 **2,209 Paris ↔ NYC**

*Based on GHG emissions*

# Démonstration

# Axes d'amélioration

Plateforme d'évaluation des questions

Suivi de la progression selon le thème

Modèle plus petit et plus performant

Choix de plusieurs catégories

Possibilité de signaler une question





Merci

Des questions ?