

# LE PROJET

# DATA FOR GOOD

3 thèmes possibles: L'écologie, la santé ou l'éducation



# LE PROJET

## Groupe de 5 personnes (voir 4)

- Le projet doit inclure **une ou plusieurs briques impliquant un LLM**.
  - **Au moins une** des briques doit être de la **RAG** (Chatbot ou autre)
- Un **dashboard** représentant des métriques pertinentes au sujet, ainsi que des variables systèmes: la latence de réponse, suivi de cout, et l'impact écologique.
- Une brique de **sécurité** pour prévenir les injections par prompt.
- Une **brique incluant du machine learning** (peut être celle de sécurité).
- Une **interface** d'interaction (Streamlit, Gradio, etc.)

**Inclure des briques non vues lors des TD sera récompensé lors de l'évaluation du projet.**

# CONSIDÉRATIONS

- **Il faudra pouvoir justifier le choix des modèles (LLM et Embeddings) utilisées:** le but est d'optimiser son architecture en utilisant exactement ce qui est nécessaire (attention à la facilité).
- **Plus vous serez sobre dans votre utilisation des ressources (taille des modèles, coûts, appels API, taille des modèles de ML), plus cela sera bénéfique sur la note finale du projet.** Le but du projet est d'avoir une application intelligemment construite, pas celle qui utilise les plus gros modèles.
- Même si nous utilisons des clés d'API gratuites, **pour votre suivi des coûts, considérez que nous avons une clé payante pour estimer le coût à l'utilisation.**

# UN BON PROJET EST UN PROJET BIEN DÉVELOPPÉ

Le respect de ces critères permettra de gagner des points sur le projet.

- **Documentez** et **typez** les fonctions (<https://medium.com/@moraneus/exploring-the-power-of-pythons-typing-library-ff32cec44981>);
- Passez **black** (<https://pypi.org/project/black/>) et **pylint** (<https://pypi.org/project/pylint/>) sur vos scripts;
- Essayez d'intégrer de la **programmation orientée objet** (<https://realpython.com/python3-object-oriented-programming/>) et les principes **SOLID** (<https://realpython.com/solid-principles-python/>);

# CALENDRIER DU PROJET

- Rendu du projet: 26/01/2025
- Présentation Orale: A définir

## CE QU'IL FAUT RENDRE

- Un repository Github avec tout le code du projet
- L'interface doit être hébergée sur HuggingFace et disponible à travers un space HuggingFace (vu en classe)
- Une présentation orale