

Le document traite du projet de recherche intitulé "Artist-run spaces : exploration et enrichissement des données textuelles",  
encadré par François-David Collin.

Sujet :

L'étude porte sur les artist-run spaces, c'est-à-dire des espaces d'art gérés et initiés par des artistes, souvent en marge des institutions traditionnelles comme les galeries ou les musées. Le projet vise à explorer et enrichir les données textuelles relatives à ces espaces à l'aide de modèles de langage avancés.

Contexte :

- Les artist-run spaces se sont développés ces 40 dernières années en réponse aux évolutions des politiques culturelles et de l'économie du secteur artistique.
- La plateforme artist-run-spaces.org recense ces espaces et les représente à travers des visualisations spatiales et temporelles, avec environ 330 entrées actuellement.
- Une collecte internationale de données supplémentaires est en cours via des contributions de type questions/réponses.

Objectifs et Méthodologie :

- Utilisation de modèles de langage pré-entraînés comme BERT ou GPT-3, et des techniques de few-shot learning pour identifier des relations

sémantiques et créer des visualisations interactives.

- Un pipeline sera développé pour gérer l'ajout de nouvelles données et annotations, avec une attention particulière portée à l'interactivité des visualisations (ex. nuages de mots, graphes de relations).

Missions :

1. Développer un outil exploratoire avec visualisation des clusters et des relations sémantiques en utilisant BERTopic.
2. Proposer de nouvelles annotations à valider par un expert.
3. Mettre en place un pipeline de few-shot learning pour généraliser les annotations et intégrer ces nouvelles classes dans les visualisations.

Outils/Technologies Utilisés :

- Python, Hugging Face, D3.js, entre autres.
- Dépôts notables : BERT, BERTopic, PET.

En résumé, ce projet se concentre sur l'utilisation de technologies de traitement du langage naturel (NLP) pour enrichir et visualiser des données relatives aux artist-run spaces, avec un accent sur l'interactivité et la flexibilité de l'outil développé.