

Sujet de stage M2 recherche: GraphRAG, graphes de contexte et LLM

Évaluation de l'apport des graphes de contexte et de GraphRAG pour la recherche texte-image d'œuvres d'art. Étude comparative sur un corpus artistique structuré en graphe de connaissances.

Les ensembles de données patrimoniales numériques, tels que ceux utilisés par les musées ou les bibliothèques culturelles, sont aujourd'hui organisés autour de vastes collections d'images d'œuvres d'art, enrichies par des métadonnées décrivant notamment l'artiste, la période, le style ou les thèmes iconographiques. Ces informations sont de plus en plus structurées sous forme de graphes de connaissances, qui constituent un support riche pour le raisonnement et l'exploration sémantique des collections. Toutefois, les systèmes de recherche d'images basés sur l'intelligence artificielle exploitent encore peu ces graphes, en se limitant le plus souvent à une association directe entre texte et image.

Les larges modèles multimodaux tels que CLIP (Radford et al., 2021) ou BLIP-2 (Li et al., 2023) offrent des performances remarquables pour l'appariement texte-image, mais leur déploiement pose des contraintes importantes en termes de ressources computationnelles et de latence.

Par ailleurs, l'émergence du GraphRAG (Graph-based Retrieval-Augmented Generation) (Edge et al., 2024) propose une approche novatrice combinant la génération augmentée par un contexte (RAG) avec l'exploitation de structures en graphe, offrant potentiellement une meilleure contextualisation des requêtes grâce aux relations sémantiques encodées dans le graphe de connaissances. Il s'agira ici d'adapter l'approche GraphRAG, qui génère un graphe à partir d'un ensemble de documents, à la notion que nous avons développée de 'graphe de contexte', un graphe préétabli, qui définit un contexte (Mimouni et al., 2024).

Ce stage a pour objectif d'étudier l'apport des graphes de connaissances et des approches récentes de type GraphRAG pour améliorer la recherche d'œuvres picturales par requêtes conversationnelles en langage naturel. Il s'appuie sur un corpus d'environ 60 000 peintures, structuré par un graphe de connaissances. Un pipeline de recherche texte–image de base servira de référence et sera enrichi par l'exploitation des relations entre les œuvres dans le graphe de contexte. Le travail consistera à concevoir et implémenter plusieurs pipelines de recherche intégrant ou non le graphe, afin de comparer leur capacité à exploiter les relations sémantiques entre œuvres, artistes, styles et périodes.

Le ou la stagiaire analysera différentes stratégies d'interrogation du corpus d'images, notamment à l'aide de modèles de langage, et évaluera l'impact de ces stratégies sur la pertinence, la cohérence et la richesse des résultats obtenus. Une attention particulière sera portée à la mise en œuvre d'un corpus de test reliant des requêtes complexes à des résultats attendus. Le stage donnera lieu à la réalisation d'un prototype de recherche conversationnelle, permettant de poser des questions complexes et de retourner des œuvres pertinentes en s'appuyant explicitement sur le graphe de connaissances.

Objectifs

1. Développer une architecture d'interrogation intégrant le graphe de connaissances du corpus
2. Implémenter et adapter une approche GraphRAG au contexte de la recherche d'images de peintures
3. Comparer les performances de plusieurs méthodes selon des métriques adaptées (avec ou sans graphe, avec graphe généré par le modèle, avec graphe de contexte)

Livrables

- Prototype fonctionnel de recherche conversationnelle
- Jeu de requêtes d'évaluation pour la recherche texte–image
- Code documenté et reproductible
- Mémoire de niveau recherche (état de l'art, analyse des résultats), présentations orales

Informations pratiques

- Lieu : CNAM, Paris – Laboratoire CEDRIC (équipe ISID), en collaboration avec Télécom Paris (équipe IDS)
- Durée : 4 à 6 mois
- Période : démarrage entre février et avril 2026
- Encadrement : Nada Mimouni (Cnam, laboratoire CEDRIC, équipe ISID), Jean-Claude Moissinac (Telecom Paris, équipe Images, Données, Signal (IDS))

Profil recherché

- Étudiant(e) en Master 2 en intelligence artificielle, data science ou domaine proche
- Bonnes bases en Python
- Connaissances en apprentissage automatique et traitement du langage naturel
- Intérêt pour les graphes de connaissances et les applications de l'IA au patrimoine culturel

Candidature

Les candidatures (CV, relevés de notes et court message de motivation) sont à adresser à Jean-Claude Moissinac (jean-claude.moissinac@telecom-paris.fr) avec en copie Nada Mimouni (nada.mimouni@cnam.fr)

Références

Nada Mimouni, Jean-Claude Moissinac. Towards Efficient Exploitation of Large Knowledge Bases by Context Graphs. SEMANTICS 2024, Sep 2024, Amsterdam, Netherlands. [⟨hal-04628484⟩](https://hal-04628484)

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., ... & Sutskever, I. (2021). Learning transferable visual models from natural language supervision. *International Conference on Machine Learning* (pp. 8748-8763). PMLR.
<https://arxiv.org/pdf/2103.00020>

Li, J., Li, D., Savarese, S., & Hoi, S. (2023). BLIP-2: Bootstrapping language-image pre-training with frozen image encoders and large language models. *International Conference on Machine Learning* (pp. 19730-19742). PMLR.

<https://arxiv.org/pdf/2301.12597>

Edge, D., Trinh, H., Cheng, N., Bradley, J., Chao, A., Mody, A., ... & Larson, J. (2024). From local to global: A graph rag approach to query-focused summarization. *arXiv preprint arXiv:2404.16130*.

Garcia, N., Renoust, B., & Nakashima, Y. (2020). Context-aware embeddings for automatic art analysis. *Proceedings of the 2020 International Conference on Multimedia Retrieval* (pp. 25-33). <https://arxiv.org/abs/1904.04985>

Mots-clés : Vision-Langage, GraphRAG, Recherche multimodale, Graphe de connaissances, Patrimoine culturel