



Appel à candidatures - Mission d'études et de recherches

Titre : Ingénieur·e en développement informatique

Type de poste : Ingénieur de Recherche - 12 mois

Prise de poste : à partir du 1er mars 2026 (négociable)

Employeur : Fondation des Sciences du Patrimoine (FSP)

Localisation : LIRIS, Lyon

Contexte

Crée en 2013, la Fondation des sciences du patrimoine (FSP) — regroupe des établissements d'enseignement supérieur, des institutions patrimoniales ainsi que des laboratoires et des unités de service qui dépendent, entre autres, du ministère de la Culture et du Centre national de la recherche scientifique. Elle assure la gouvernance du projet EquipEx+ ESPADON (ANR-21-ESRE-0050), dont l'objectif est d'introduire de nouvelles modalités d'accès aux connaissances sur les objets patrimoniaux et de nouveaux services numériques innovants pour leur étude pluri- et interdisciplinaire.

Le projet

Le projet TAC (Thésaurus Automatisé par Curation) a pour objectif la création d'un **outil** pour faciliter la documentation et la structuration de vocabulaires spécialisés pour le patrimoine culturel, en s'appuyant à la fois sur les pratiques documentaires existantes au sein des institutions mobilisées dans le domaine du patrimoine et sur les grands modèles de langages (LLMs). Intégré à l'écosystème ESPADON, il visera au versement et à l'agrégation automatisée de vocabulaires spécialisés vers son gestionnaire de thésaurus Opentheso. Plus particulièrement, il s'agira de faciliter la création de thésaurus à partir de listes de termes non structurées, en automatisant les indispensables étapes de curation, de définition, de structuration et d'alignement de concepts.

Mission

L'ingénieur ou l'ingénierie proposera un outil (nommé TAC) composé d'un microservice et d'une interface web. Le microservice s'appuyant sur un LLM pour assurer automatiquement la curation, la structuration, et l'alignement de concepts provenant de listes d'autorité sur les référentiels terminologiques exposés en SKOS via Opentheso. Une interface web devra être développée afin de permettre aux utilisateurs de valider et corriger les résultats obtenus avant de les exporter

et/ou les soumettre à Opentheso par le biais de son API. L'outil sera conteneurisé avec Docker et déployé dans l'écosystème numérique ESPADON.

Par exemple, un utilisateur envoie à TAC une liste de termes, non-structurée. Cette liste est ensuite traitée automatiquement par l'API à travers des étapes de classification des termes, d'extraction de leurs définitions, de formalisation à l'aide de SKOS, pour finalement aboutir à leur alignement aux autres thésaurus pré-existants. Les données de vocabulaires qui sont traitées en entrée par TAC, comme les listes de termes, les langages utilisés, les techniques de requêtes IA ou les connections vers des ressources externes sont restituées en sortie sous plusieurs formes : une documentation sur les vocabulaires, des propositions de concepts candidats pour Opentheso avec des mises à jour éventuelles, une gestion du versionnement et des éléments de synthèse pour la gestion documentaire.

Activités

Le projet s'articule en trois phases :

PHASES DU PROJET	DURÉE	TÂCHES
1/ De l'expérimentation au cahier des charges d'un outil fonctionnel : La première phase correspond à la mise en place du cahier des charges de l'API TAC dans la continuité de l'expérimentation LLMtheso. Il s'agira dans un premier temps de définir les cas d'usage qui permettront de généraliser l'approche expérimentale qui a déjà donné lieu à une preuve de concept.	2 mois	Choix des cas d'usage Définition et évaluation des prompts, tests RAG, tests <i>fine tuning</i> Définition du cahier des charges pour le développement de l'API et de l'UI Documentation du projet
2/ Développement de l'outil TAC : La deuxième phase prévoit le développement de TAC et d'une interface utilisateur (UI). Elle comprend également des tâches liées à la documentation du projet et du code en vue de son ouverture.	6 mois	Développement et documentation de l'API Développement de l'UI Documentation du projet
3/ Évaluation, débogage, déploiement et diffusion : la troisième phase de finalisation comprend les cycles de tests, débogages, et le déploiement de l'outil sur les serveurs d'ESPADON. Cette phase comprend également la communication et la diffusion des résultats du projet. Les codes sources afférents seront ouverts et mis à disposition via une forge logicielle.	4 mois	Évaluation de l'outil Débogages Déploiement ESPADON Documentation du projet Communication et publication

Compétences attendues :

Full-stack dev, profil généraliste :

- Langages de programmation (Python, Javascript)
- Gestion de conteneurs (Docker)
- Virtualisation

- Culture des sciences du patrimoine et humanités numériques.
- Anglais niveau B2–C1.

Livrables attendus (indicateurs) :

- Documentation
- Documentation technique du développement de la plateforme d'exposition des données

Savoir-faire opérationnels :

- Conseil et aide à la décision (qualité, gouvernance des données).
- Accompagnement, pédagogie et gestion de projet en humanités numériques.

Savoir-être :

- Travail en équipe, collaboration interdisciplinaire (numérique, SHS).
- Qualité relationnelle, sens de l'organisation et autonomie
- Capacité d'adaptation

Environnement et collaboration :

Le poste s'inscrit dans un consortium national (FSP, CNRS, Ministère de la Culture, universités) avec plateformes instrumentales, laboratoires SHS et un WP Data fournissant stockage, calcul et gestion des données. L'ingénieur ou l'ingénierie interagira avec la direction scientifique du projet, les développeurs et les dévelopeuses, et les partenaires (laboratoires de recherche et institutions patrimoniales).

Conditions du poste :

- Contrat : CDD 12 mois • Affectation : Lyon, LIRIS
- Des déplacements occasionnels à Paris et Marseille (pris en charge par le projet) sont à prévoir.
- Cadre open science : diffusion des résultats/prototypes en code/logiciels ouverts, avec déclaration d'invention/logiciel si nécessaire.

Pour candidater :

Date limite de candidature : 31/01/2026

Envoyer un document PDF (un seul fichier) incluant :

- Curriculum Vitae détaillé
- Lettre de motivation, accompagnée d'un résumé synthétique de vos activités de recherche ou de votre formation en relation avec les objectifs du projet
- Diplômes

à l'adresse suivante : olivier.malavergne@culture.gouv.fr