

POSTE D'INGÉNIEUR·E ARCHIVES ET TECHNOLOGIES SÉMANTIQUES

(CDD IGE 12 mois – Cat. A)

Héritière de la Bibliothèque de l'Université de Paris ouverte en 1770, la Bibliothèque interuniversitaire de la Sorbonne (BIS) est une bibliothèque de recherche spécialisée en lettres et sciences humaines, rattachée pour sa gestion à l'université Paris 1 Panthéon Sorbonne. Elle accueille un public d'étudiant·e·s (à partir de la troisième année de licence) et de chercheur·se·s et enseignant·e·s-chercheur·se·s. Dotée d'un budget annuel de 3,9 millions d'euros et d'un effectif de 128 agents titulaires renforcés par des contractuel·les, elle conserve dans ses collections et propose à ses lecteurs environ 2 millions de documents, dont 17 750 titres de périodiques et 3 500 manuscrits. La bibliothèque de l'institut de géographie lui est rattachée.

Depuis 2017, elle est bibliothèque délégataire du réseau CollEx-Persée pour l'histoire, la géographie, la philosophie et les sciences de l'Antiquité. Elle propose également de riches collections de littérature.

La BIS développe une politique ambitieuse de valorisation de ses collections et d'accompagnement à la recherche. Elle a créé à cette fin en 2019 un service dédié à ces enjeux, le service de la valorisation numérique des collections et du soutien à la recherche (SERVAL).

Elle porte des projets scientifiques et documentaires en lien avec des chercheurs (constitution de bases et d'entrepôts de données, mise en place de référentiels), et est partenaire de projets d'humanités numériques. Ces projets sont coordonnés ou menés au sein du service transverse de la valorisation numérique des collections.

Contexte :

ORESM (Oeuvres et Référentiel des Etudiants, Suppôts et Maîtres de l'Université de Paris au Moyen Âge), est un projet porté conjointement depuis 2020 par la Bibliothèque interuniversitaire de la Sorbonne (BIS) et par le laboratoire de médiévistique occidentale de Paris (LaMOP – UMR 8589) avec de très nombreux partenaires notamment les Archives nationales avec deux départements de la direction des Fonds qui sont responsables du traitement et de la valorisation des documents provenant de l'Université de Paris ; et le Lab, qui s'occupe de la R&D numérique de l'institution, est notamment en charge de ses projets *Linked Data* et est très impliqué dans les aspects numériques du projet. Le projet est soutenu et co-financé par le GIS CollEx-Persée.

Le projet ORESM vise à collecter, créer et valoriser des données relatives à l'histoire de l'Université de Paris, des collèges et écoles parisiens entre le 12^e et le 16^e siècle, et à les rendre consultables sur un portail web dédié.

Ce portail ambitionne entre autres :

- de réunir virtuellement les documents qui sont partie intégrante du fonds d'archives de l'université, et qui sont dispersés depuis la Révolution entre différentes institutions partenaires du projet, notamment la BIS, les Archives nationales (AN) et la Bibliothèque interuniversitaire de Santé (BIU Santé) ;

- de proposer un référentiel des personnes liées à l'université de Paris au Moyen Âge en s'appuyant sur les riches données de la base *Studium Parisiense* (gérée par le LaMOP) ;
- de proposer un répertoire des manuscrits liés aux personnes ayant gravité autour de l'université de Paris.

Missions du service :

Le Service de la valorisation numérique des collections et du soutien à la recherche (SERVAL) a pour mission de définir, coordonner et opérer des actions de valorisation des collections sur des outils numériques, promouvoir et porter des partenariats scientifiques en particulier, en lien avec des projets de recherche.

Au 01/09/2022, l'équipe comporte 3 conservatrices titulaires, 3 chargé-e-s de projets (cat. A), un ingénieur informatique à mi-temps et une gestionnaire de bibliothèque numérique (cat.B).

Missions principales du poste :

- Contribuer à la construction d'un graphe de connaissances RDF à partir des données produites ou réunies dans le cadre du projet, en utilisant, comme principale ontologie de référence, l'ontologie [ICA Records in Contexts – Ontology \(RiC-O\)](#) ;
- Mettre en place sur le web une base RDF (triplestore) pour stocker, gérer et enrichir ce graphe ; assurer l'administration de cette base RDF ; mettre en place un SPARQL endpoint pour l'interrogation du graphe par les utilisateurs humains et par les machines à distance ;
- Concevoir et réaliser une interface de recherche et d'exploration du graphe qui soit adaptée aux besoins des futurs utilisateurs du portail (par exemple en utilisant Elastic Search et/ou un éditeur visuel de requêtes SPARQL comme Sparnatural), et qui permette également d'exporter les données dans des formats divers tels que CSV ;
- Concevoir et développer une interface graphique permettant de consulter les données du graphe ;
- Assurer l'intégration de données dans l'inventaire virtuel du projet.

Missions et activités :

- Contribuer à la construction d'un graphe de connaissances RDF à partir des données produites ou réunies dans le cadre du projet (données issues des dépouillements d'archives (tableau CSV), données issues de réservoirs de données extérieurs ainsi que des inventaires existants (XML/EAD). Cela inclura : la participation à des travaux de modélisation, en prenant comme principale ontologie de référence RiC-O ; la participation au développement de scripts de conversion des données existantes en données RDF conformes au modèle choisi ; la réalisation d'un ou plusieurs pipelines de production du graphe (certaines des données source devant être mises à jour et enrichies pendant la durée du projet) ;
- Installer et configurer sur un serveur web une base RDF ; importer dans cette base le graphe de connaissances obtenu ; assurer l'administration des données du graphe, en définissant et mettant en œuvre des méthodes pour sa mise à jour et son enrichissement (par import de données ; par intégration de données RDF externes via des requêtes SPARQL fédérées ; ou par inférence en exécutant des opérations SPARQL de type INSERT DATA) et pour la gestion de sa qualité ;

- Mettre en place un SPARQL endpoint (point d'accès SPARQL) afin de permettre aux utilisateurs humains d'interroger le graphe directement en SPARQL et aux machines distantes d'utiliser le protocole SPARQL pour interroger ce graphe ;
- Concevoir et réaliser une interface de recherche graphique adaptée aux besoins des futurs utilisateurs du portail (par exemple en utilisant Elastic Search et/ou un éditeur visuel de requêtes SPARQL comme Sparnatural), permettant également d'exporter les données dans des formats divers tels que CSV ;
- Concevoir et développer une interface graphique permettant de consulter les données du graphe ;
- Réaliser l'intégration de données issues des dépouillements d'archives dans l'inventaire virtuel en XML-EAD du projet ;
- Rédiger la documentation technique des développements mis en œuvre, rendre le code compréhensible (commentaires) et accessible (logique d'Open Access) sur la plateforme GitHub ;
- Participer aux autres aspects du projet en fonction de son expertise ;
- Présenter le travail en cours aux partenaires, au sein des instances du projet (CS et GT) et lors de journées d'études.

La personne recrutée mènera à bien ces missions en concertation avec l'ensemble des personnes participant au projet, en particulier pour ce qui concerne la conception des interfaces de recherche et de consultation. Elle travaillera en collaboration avec le Lab des Archives nationales. Ces missions seront co-pilotées par la BIS et le Lab des Archives nationales.

Rattachement hiérarchique : sous la responsabilité de la cheffe du SERVAL, au sein de la Bibliothèque interuniversitaire de la Sorbonne, Université Paris 1 Panthéon-Sorbonne.

Liens fonctionnels :

En interne :

- Coordinatrice scientifique du projet à la BIS
- Département des manuscrits et livres anciens (DMLA)

En externe :

- Coordinateur scientifique du projet à l'université de Franche-Comté
- Archives nationales : départements du Moyen Age et de l'Ancien Régime et du Minutier central des notaires de Paris ; Lab
- Laboratoire de médiévistique occidentale de Paris (LaMOP)
- Chercheur·se·s et professionnel·le·s des archives et des bibliothèques participant au comité scientifique et aux groupes de travail du projet

Compétences :

- Méthodologie de conduite de projet ; maîtrise de Git et GitHub ;
- Formation en histoire médiévale, archivistique et humanités numériques. Une expérience de participation à un projet d'humanités numériques serait un plus.
- Anglais lu et écrit ;
- Connaissance au moins théorique des normes de description archivistique (ISAD(G), ISAAR(CPF), modèle conceptuel Records in Contexts) ;
- Maîtrise de XML, connaissance des formats XML/EAD, XML/EAC/CPF, maîtrise des technologies XML (notamment DTD, W3C Schema, langage XPath, transformation de documents XML avec le langage XSLT) ;
- Connaissances en gestion de bases de données et bases de graphes ;

- Expérience du développement applicatif (la maîtrise d'un langage de programmation type Python est attendu) et en conception et développement web (HTML, CSS, JavaScript) ;
- Maîtrise des concepts du web sémantique et des Linked Open Data ; expérience solide dans l'utilisation de RDF, des langages RDFS, OWL et SPARQL. Expérience significative en modélisation OWL, en gestion de graphe, en conception et réalisation d'application web de recherche et de consultation de graphe RDF.

Savoir-être :

- Excellent relationnel.
- Excellentes capacités d'expression écrite et orale.
- Aptitude à travailler en équipe.
- Rigueur et organisation.
- Capacités d'analyse, d'abstraction et de synthèse ;
- Autonomie et capacité d'initiative, capacité à rendre compte à ses encadrants

Contraintes liées au poste :

Régime 35h. Lieu de travail situé à Paris 5^e. Déplacements fréquents dans Paris et en Île-de-France.

Statut : CDD de 12 mois, rémunération selon expérience sur la base de la grille des ingénieurs d'étude.

Date limite de candidature : 31/10/2022

Prise poste souhaitée au : 01/01/2023

Contacts pour information et candidature :

Par courriel à :

- Laurence BOBIS, directrice de la bibliothèque, co-porteuse du projet ORESM : direction-bis@univ-paris1.fr
- Thierry KOUAME, professeur d'histoire médiévale à l'université de Franche-Comté, co-porteur du projet ORESM : thierry.kouame@univ-fcomte.fr
- Cécile OBLIGI, cheffe du SERVAL : cecile.obligi@bis-sorbonne.fr
- Laurie Aoustet, adjointe à la cheffe du SERVAL : laurie.aoustet@bis-sorbonne.fr
- Florence CLAUDAUD, responsable du Lab des Archives nationales : florence.clavaud@culture.gouv.fr