

Ingénieur-e de recherche en production, traitement et analyse de données

*Poste à pourvoir par contrat à durée déterminée de 24 mois,
à temps complet (38h45 par semaine)*

Informations générales

- *Spécialisation disciplinaire* : Sciences Humaines et Sociales
- *Emploi-type* : D1A41- Ingénieur-e de recherche en production, traitement et analyse de données
- *Niveau de recrutement* : Catégorie A
- *Niveau de rémunération mensuelle* : 3040 euros bruts
- *Localisation du poste* : Centre de recherches historiques (UMR 8558), EHESS, 54, boulevard Raspail, 75006 Paris.
- *Date prévisible de début du contrat* : 1er mai 2024

Contexte

L'École des hautes études en sciences sociales (EHESS) est un établissement public à caractère scientifique, culturel et professionnel, "grand établissement" assurant une mission de recherche dans le domaine des sciences humaines et sociales et de formation à la recherche. L'établissement accueille 3 000 étudiants dont 1 500 doctorants (avec près de 250 thèses soutenues par an dans toutes les disciplines des sciences sociales) et autant de masterants. L'EHESS accueille également 40 unités de recherche organisées autour d'un domaine, d'une approche ou, pour un tiers d'entre elles, d'une "aire culturelle". En leur sein, plus de 500 enseignants et chercheurs développent les savoirs, forment et encadrent les étudiants de master et les doctorants. L'École est membre fondateur de l'Établissement public Campus Condorcet, dont la mission est la coopération scientifique entre ses membres fondateurs, et notamment le soutien à leurs programmes de recherche, et la mise à disposition des chercheurs et des étudiants fréquentant le Campus de conditions de vie et de travail à la hauteur des normes internationales.

Biblissima+ (Observatoire des cultures écrites, de l'argile à l'imprimé) fait partie des équipements structurants pour la recherche EquipEx+ sélectionnés en 2020 dans le cadre des Investissements d'avenir. Il prend le relais de l'EquipEx Biblissima (Bibliotheca bibliothecarum novissima : observatoire du patrimoine écrit du Moyen Âge et de la

Renaissance, 2012-2021). Biblissima+ est porté par le Campus Condorcet, piloté par Anne-Marie Turcan-Verkerk (AOROC, EPHE-PSL) avec François Bougard (IRHT, CNRS), Marie-Agnès Lucas-Avenel (CRAHAM, Université de Caen) et Emmanuelle Morlock (HiSoMa, CNRS). Le programme bénéficie d'une aide de l'État gérée par l'Agence nationale de la Recherche. Il s'achèvera le 30 juin 2029. Biblissima+ fédère 16 établissements et une entreprise privée, réunis en un consortium et agissant par l'intermédiaire de leurs services ou d'unités de recherche ou de service placées sous leur tutelle. Avec leurs équipes, Biblissima+ crée une infrastructure numérique multipolaire de recherche fondamentale et de service consacrée à l'histoire de la transmission des textes anciens, des premières tablettes d'argile mésopotamiennes aux premiers livres imprimés, sur tous les supports et dans toutes les langues. L'équipe en charge du portail Biblissima+, composée de quatre personnes, est hébergée par l'EPCC dans l'Hôtel à projets du Campus Condorcet.

Présentation du projet

“Anthropologie historique du long Moyen Âge” (AHLOMA) est un groupe de recherche du Centre de recherches historiques (UMR 8558/ CNRS-EHESS) réunissant 14 chercheuses et chercheurs de l'EHESS et du CNRS. Le Centre de recherches historiques compte également trois ingénieurs de recherche et deux ingénieurs d'études développant des outils numériques de recherche. L'ingénieur-e de recherche, placé-e sous la responsabilité de la direction du Centre de recherches historiques, et sous la direction scientifique du responsable d'AHLOMA, aura pour mission de travailler sur le développement des bases de données iconographiques et textuelles du groupe. Celles-ci sont au nombre de trois : la base des enluminures médiévales, ThEMA (Thesaurus Exemplorum Medii Aevi) et Quodlibase.

1. La base des enluminures médiévales regroupe plus de 80 000 images de manuscrits indexées avec un vocabulaire contrôlé. Mise en œuvre à partir de 1984 par J.-C. Schmitt, J.-C. Bonne, J. Baschet et M. Pastoureau, elle s'inscrit dès ses débuts dans un triple projet de recherche, d'ingénierie et de pédagogie. Ce dernier s'incarne aujourd'hui encore dans un séminaire hebdomadaire de l'EHESS. Chaque indexation de manuscrit est le fruit d'un dialogue nourri entre chercheurs, ingénieurs et étudiants, l'ensemble étant cadré par le Thésaurus des Images Médiévales en ligne (TIMEL), contenant 4 576 concepts au standard SKOS (<https://datu.ehess.fr/timel/>). Le corpus indexé est largement polarisé par les manuscrits de trois collections de la Bibliothèque apostolique vaticane (Reg, Urb. et Lat.), pour un équivalent de 60 000 images issues de 20 000 folios de manuscrits. Il contient également 20 000 images de manuscrits conservés dans différentes institutions de conservation européennes et nord-américaines, dont la Bibliothèque nationale de France. L'ensemble concerne exclusivement des manuscrits occidentaux du VI^e siècle au XVI^e siècle. La base sera mise en ligne dans son intégralité en 2024 dans l'entrepôt de données ouvert de l'EHESS, Didomena, où les images nouvellement indexées sont directement déposées depuis 2020 (<https://didomena.ehess.fr/collections/3x816n47s>).

L'objectif du travail sur cette base est **(a)** de mettre en place les conditions d'une interopérabilité des données de la base avec les référentiels Biblissima+, en alignant les données EHESS sur les référentiels d'autorité Biblissima+ **(b)** de mettre en place les conditions d'une interrogation des données dans le portail Biblissima+, selon les modalités décrites dans le vademecum de Biblissima+ (<https://doc.biblissima.fr/vademecum-biblissima>) **(c)** de participer avec les autres ingénieurs du CRH et la DSI de

École des hautes études en sciences sociales

54, boulevard Raspail, 75006 Paris

l'EHESS au développement d'un portail web de consultation de la base, en lien avec l'API IIIF de Didomena.

Ce travail permettra d'initier une collaboration avec d'autres unités de recherche françaises et européennes disposant de bases indexées avec le thesaurus TIMEL. Pour ce faire, l'ingénieur-e sera amené-e à explorer la possibilité de collaboration avec des infrastructures d'humanités numériques nationales (Huma-num) ou européennes (DARIAH).

2. et 3. ThEMA (<http://thema.huma-num.fr/>) est une base de données d'anecdotes exemplaires du Moyen Âge, issues de textes de l'Europe, du Moyen Orient et de l'Asie. Il repose sur des fichiers XML stockés dans une base existdb. Quodlibase (<http://quodlibase.ehess.fr>) est une base de Quodlibets théologiques du XIIIe siècle au XIVe siècle, issus des séances solennelles au cours desquelles les maîtres des facultés de théologie répondaient aux questions des membres de la faculté sur n'importe quel sujet. Une nouvelle version du site, reprenant la même technologie employée pour ThEMA, sera mise en ligne en 2024. L'ingénieur-e de recherche pourra être sollicité-e pour intervenir, avec les autres ingénieurs-e-s du CRH, sur ces deux bases, pour des tâches de développement web ou de traitement des données.

Missions du poste

- Curation des données de la base des enluminures médiévales en lien avec les chercheurs travaillant sur la base.
- Alignement des données de la base des enluminures et du thesaurus TIMEL avec les référentiels Biblissima+ (cotes de manuscrits, organisations, lieux, descripteurs iconographiques).
- Partage des données de la base des enluminures selon les modalités décrites dans le vademecum Biblissima+, afin de permettre leur interrogation dans le volet iconographique du portail Biblissima+.
- Développement d'un portail Web pour la consultation de la base des enluminures médiévales, en lien avec l'API IIIF de *Didomena*, jusqu'à sa mise en production.
- Développements ou traitements de données ponctuels sur les bases Quodlibase et ThEMA.
- Dépôts des données des trois bases dans des entrepôts de données de la recherche ouverts.
- Accompagnement des porteurs scientifiques dans la définition de leurs besoins.
- Mise en place et suivi des opérations (administratives et techniques), des opérations spécifiques et des interventions sur les bases de données.
- Valorisation des réalisations *via* la participation à des journées d'études et l'écriture d'articles scientifiques.

L'ingénieur-e de recherche travaillera en lien avec le collectif "Sources et données de la recherche" du CRH, qui réunit les ingénieurs et les chercheurs de l'unité, ainsi qu'avec les pôles Numérique recherche et Infrastructures de la DSI de l'EHESS. Il-Elle sera amené-e à échanger avec l'équipe du portail Biblissima+ et à présenter l'avancement des travaux devant son bureau exécutif.

Conditions/sujétions particulières d'exercice

- Déplacements réguliers à l'INHA, 2, rue Vivienne, 75002 Paris, au bureau du groupe AHLOMA.
- Déplacements possibles au Campus Condorcet, 8, cours des Humanités 93322 Aubervilliers.
- Possibilité de télé-travailler

Compétences requises

- *Compétences générales*
 - Connaissance du secteur de la recherche publique ;
 - Connaissance des institutions patrimoniales ;
 - Reporting et circulation de l'information ;
 - Hiérarchisation des enjeux et priorités d'actions;
 - Aisance dans l'expression écrite et orale ;
 - Aptitude à travailler de façon polyvalente et en équipe.
- *Compétences numériques*
 - Capacité à s'autoformer et à s'adapter à de nouveaux outils ;
 - Maîtrise de l'écosystème du développement web (dont logiciel de gestion de versions, HTML, CSS, Javascript, déploiement, etc.)
 - Maîtrise d'un langage de programmation ;
 - Expérience en développement web requise ;
 - Des notions dans l'usage du logiciel OpenRefine et du standard SKOS, dans la manipulation des API et du standard IIF seraient appréciées, mais nullement obligatoires pour candidater.
- *Compétences linguistiques*
 - Anglais scientifique (B2-C1).

Profil recherché

- Diplôme : doctorat ou diplôme d'ingénieur ;
- Expérience en développement web requise.

Contacts

Pour toute information concernant le projet et les attendus scientifiques et techniques du poste, merci de contacter : Étienne Anheim <etienne.anheim@ehess.fr>, Pierre-Olivier Dittmar <pierre-olivier.dittmar@ehess.fr>, Raphaëlle Lapôtre <raphaelle.lapotre@ehess.fr>, Jean-Damien Généro <jean-damien.genero@cnrs.fr>.

Candidatures

Les candidatures (CV, lettre de motivation) doivent être déposées sur l'interface emploi de l'EHESS : <https://recrutement.ehess.fr/> au plus tard le vendredi 31 mars 2024.

Les candidats dont les dossiers auront été présélectionnés, seront convoqués pour une audition qui aura lieu la semaine du 1er avril 2024.