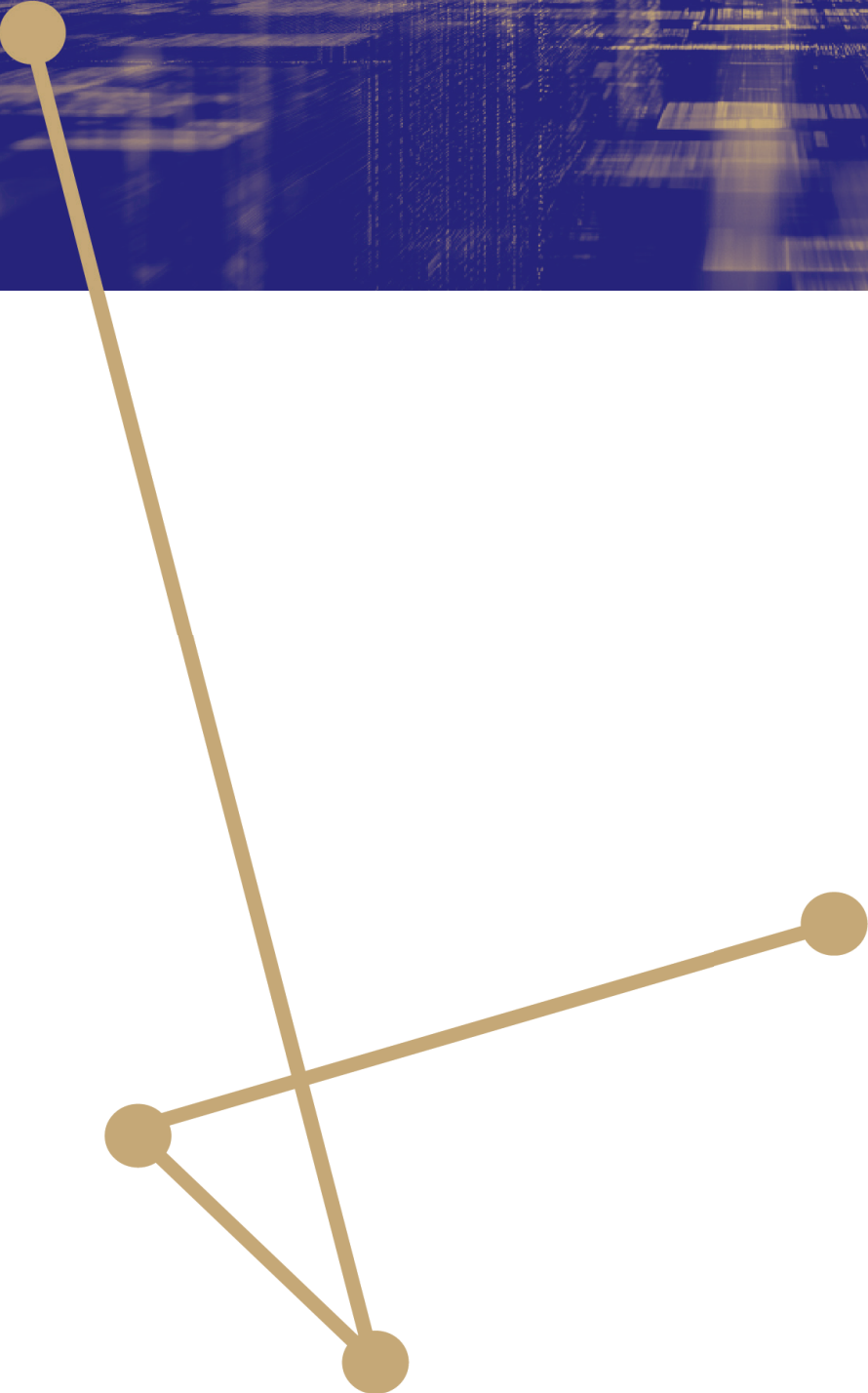


APPEL D'OFFRE N°2025-01



**Réponse jusqu'au
5 septembre 2025**



Objet : Accompagnement visant à faire émerger au sein d'un nombre limité d'institutions partenaires, des plateformes permettant d'expérimenter diverses technologies sur les corpus d'archives scientifiques dans le cadre de projets portés par des communautés de recherche.

Établissement : Muséum national d'Histoire naturelle
57 rue Cuvier
75005 Paris

Date limite de réponse : 05 septembre 2025

Type de marché : Prestation de service

Pour obtenir toute information complémentaire, les opérateurs économiques ont la possibilité de s'adresser à :

Judith SORIA, chargée de mission Archives scientifiques de CollEx-Persée
Bibliothèques du Muséum national d'Histoire naturelle
38 rue Geoffroy Saint-Hilaire 75005 Paris
judith.soria@mnhn.fr / 01 71 21 47 53

Contexte de la mission

Le groupement CollEx-Persée (<https://www.collexpersee.eu/>), s'est doté d'un programme national structurant dans le domaine de la documentation autour du signalement, de la diffusion et de l'exploitation des archives scientifiques et des matériaux pour la recherche, piloté par le Muséum national d'histoire naturelle (MNHN) et la Contemporaine.

Par l'intermédiaire administratif du MNHN, CollEx-persée ouvre une offre de mission d'accompagnement visant à faire émerger au sein d'un nombre limité d'institutions partenaires une plateforme de services permettant d'expérimenter diverses technologies d'accès à et d'exploitation de corpus d'archives scientifiques, dans le cadre de projets portés par des communautés de recherche.

Il a été constaté que les nouvelles possibilités d'analyse offertes par l'intelligence artificielle (HTR, OCR, NER, TDM, etc.) doivent pouvoir s'appliquer aux archives scientifiques pour offrir aux chercheurs de nouvelles possibilités d'analyse des fonds. Mais si les technologies sont désormais bien développées et pour certaines, presque mûres, il reste de nombreux obstacles en termes d'infrastructures, de ressources serveur, de compétences et d'organisation. Si des établissements de taille moyenne sont généralement en mesure de disposer de leur propre bibliothèque numérique, il leur est beaucoup plus difficile de déployer des technologies innovantes, qui doivent en outre être adaptées aux spécificités de chaque corpus et aux besoins de chaque communauté de recherche. D'un autre côté, il n'est pas forcément souhaitable de centraliser l'offre de service au sein d'une seule institution, solution qui fait peser le risque d'un manque de souplesse et d'une mauvaise prise en compte des spécificités disciplinaires.

Avant de lancer un appel à manifestation d'intérêt auprès de la communauté pour faire émerger de telles plateformes à un niveau régional (périmètre permettant de desservir le plus de communautés de recherches), il apparaît nécessaire de tester deux preuves de concept, en région Île-de-France et en région Grand-Est. Le programme veut se faire aider d'une mission d'accompagnement permettant de préciser les contours de ces expérimentations pratiques tout en pensant, d'emblée, leur transposition sur d'autres terrains géographiques.



Objet de la mission

La mission a donc pour objectif d'aider deux établissements pilotes, la direction des bibliothèques, de l'information et de la science ouverte (DiBISO) de l'Université Paris-Saclay et la Bibliothèque nationale et universitaire (Bnu) à Strasbourg, à préparer une preuve de concept.

La mission doit aider à cartographier les institutions partenaires susceptibles de participer et/ou d'héberger ce pôle technique toute la durée du programme, en vérifiant leur adhésion à un tel postulat. Les institutions partenaires devront donc être identifiées sur une logique d'expertises (en s'appuyant par exemple sur l'existence de datalabs, d'ateliers de la donnée, de MSH ou de mésocentres), mais aussi géographique et/ou disciplinaire, pour assurer un bon maillage territorial et une meilleure réponse aux besoins. La présence de machines et d'ingénieurs au service de la communauté de l'ESR sera aussi repérée. Dans la réflexion, les besoins des communautés de recherche seront pris en compte, afin de penser les outils, technologies et machines à mettre à disposition, ainsi que les compétences support permettant la production de nouveaux outils, tels que des modèles d'apprentissage d'écriture, d'extraction, d'analyse sémantique ou d'alignement par exemple. Ils devront en outre être pensés en conformité avec les principes de la science ouverte, c'est-à-dire être libres, partagés, réutilisables et à la disposition de toutes les communautés.

Enfin, la mission devra proposer des solutions de montage administratif ainsi qu'un modèle économique viable à long terme d'une telle plateforme de services. CollEx-Persée a en effet vocation à donner une impulsion de départ, mais l'infrastructure doit ensuite être reprise par l'établissement hôte, qui lui-même aura besoin d'un soutien financier de sources tierces (notamment de la part de porteurs ou financeurs des projets bénéficiaires).

Calendrier indicatif de la mission

La mission est établie pour la durée suivante : 15 octobre 2025 au 15 février 2026.

Montant de la mission

Le montant de la mission ne pourra pas dépasser 20 000€ HT.

Descriptif de la mission

Le prestataire aura donc les tâches suivantes à développer :

- Cartographier et indiquer les points forts des partenaires potentiels en région Grand-Est, d'une part, et en Île-de-France, d'autre part, susceptibles de former un consortium capable de porter ce type de plateformes techniques en mutualisant leurs compétences (sur la base de critères d'expertises à formaliser, d'infrastructures et d'ETP déjà présents et de rencontres avec ces institutions dont une première liste sera fournie au prestataire) et pointer les éventuels besoins à combler de ces partenaires afin de mener à bien le POC.
- Affiner les cas d'usages des chercheurs et donc les besoins en termes de ressources techniques et technologiques (quels modèles de Deep Learning et de TAL) à mettre à disposition des communautés de recherche concernées (sur la base des outils et services existants recensés, d'entretiens avec des ingénieurs en poste et d'équipes de recherche investies et dont une première liste sera fournie au prestataire). Il conviendra de prendre en compte tout le cycle de vie de la donnée jusqu'à sa préservation pérenne ou suppression définitive.
- En fonction de ces cas d'usage, préciser les technologies, services et infrastructures nécessaires à leur mise en œuvre, ainsi que les coûts induits (infrastructures, énergétiques, et RH).
- Offrir un repérage des ressources d'ingénierie privée spécialisée disponibles sur le territoire considéré, en vue de prestations ou de possibles partenariats public/privé.
- Conseiller un modèle administratif et économique pour le bon développement d'une telle plateforme de services (plusieurs scénarii possibles en distinguant la durée du groupement CollEx-Persée de la période post-financements CollEx-Persée).
- Formaliser dans un livrable la méthodologie développée dans cette mission pour qu'elle soit transposable de façon autonome à une autre région dès le premier semestre 2026 (par exemple Auvergne-Rhône-Alpes, Centre Val de Loire...)



Confidentialité

Le prestataire s'engage à ne pas divulguer les informations et documents de quelque nature que ce soit qu'il aurait pu recueillir, obtenir ou dont il aurait pu avoir connaissance au cours de l'exécution du présent marché.

De même, CollEx-Persée s'engage à garder strictement confidentiel l'ensemble des informations échangées, notamment les renseignements techniques et commerciaux.