Le Bnu_Lab

Présentation Comité CoRTecS | 16/06/2025

6 place de la République
BP 51029
67070 Strasbourg Cedex
03 88 25 28 00



La démarche Bnu_Lab

- Un lieu, des compétences, des corpus d'excellence en SHS.
- Des partenariats sur le terrain local : Estrades, ADELE helpdesk, Numistral.
- Une complémentarité avec l'écosystème actuel : PUD (Progedo), PHUN (HumaNum), Estrades (édition), ArkéoGis (disciplinaire), Num3D (MISHA).

> L'apport de Bnu_Lab : fourniture de jeux données à partir de corpus textuel patrimonial ou du nativement numérique (DL Web), enrichissement, exploitation dans le cadre de projets de recherche.

> Démarche CoRTecS = une opportunité exigeante et stratégique pour formaliser l'appui aux SHS et aux HN par la Bnu.

L’offre de services en 2025

- **Constitution** de corpus numériques, nativement numériques, jeux de données de qualité ;
- **Enrichissement** par segmentation, HTR/OCR, balisage de données en TEI, traitement automatique du langage (TAL) dont NER (entités nommées), mise en œuvre du protocole IIIF à partir de corpus issus des collections de la Bnu ;
- **Garantie** du respect de la science ouverte et des principes FAIR, en particulier l’interopérabilité et respects des standards avec des outils et bases existantes ;
- **Conseils** autour du TDM (text and data mining) sur des corpus sous-droits ;
- **Accompagnement** à l'utilisation des données du dépôt légal du web collecté par la BnF dans un projet de recherche : collecte, mise à disposition des données, fouille, à partir des extractions fournies par la BnF.

> **Guichet des projets** : travailler avec des partenaires, recherche des financements : AAP DataLab de la BnF, CollEx Persée, Equipex Biblissima+, ECHOES, etc.

Focus sur la formation et diffusion des savoirs

- Diffuser la culture de la donnée auprès du public SHS : programmation autour des outils, formats, codage, etc.
- Expérimentation en cours : appel à chercheurs en résidence à la Bnu : 2026 : «*Ovide en images à la Bnu : du livre ancien à la donnée*».
- Diffuser les savoirs : rencontres scientifiques ouvertes au grand public mêlant enjeux corpus/HN : ex. JE Éditer… et ensuite ? De l'édition numérique aux études littéraires computationnelles (18/06) ; Chronospedia (nov. 2025), CHI-KNOW-PO (février 2026), « IA et livre ancien : propositions méthodologiques », (2ème sem.).
- Journées **IA4LAM** pour le Grand Est : réunir une communauté de pratique en langue française autour de l'application des technologies d'intelligence artificielle dans les bibliothèques, archives et musées.

> Enjeux : structurer notre partenariat avec l'URFIST, se rapprocher du CCN, des Ecoles doctorales, de l'USIAS.

Les perspectives

- **Faire bénéficier nos partenaires de nos engagements nationaux** : membre du Consortium ReligiS, pilote axe archives scientifiques du GIS CollEx-Persée avec Université Paris Saclay , membre du groupe de travail Gallica Image de la BnF.
- **Nos attentes vis-à-vis de CoRTecS** : aide et conseil pour formaliser l'auditabilité des coûts des prestations de services, bénéficier de la démarche qualité et formation continue, approcher les acteurs du monde socio économique.
- **Questions ?** Monter ensemble des mécanismes transformants, par ex. résidences de chercheurs en humanités numériques ? Réaliser une étude de marché sur les relations entre SHS, HN et industrie ? De l'événementiel dédié ? Quelle place pour les plateformes en SHS dans ENACT ?



Contacts :

catherine.desos-warnier@bnu.fr
rosanne.wingert@bnu.fr
francois.quere@bnu.fr