

Journées annuelles du cluster 5b  
« Élaboration d'un laboratoire de textes »  
(jeudi 7 et vendredi 8 décembre 2023)

## PROGRAMME

### Jeudi 7 décembre 2023

Accueil des participants (9h30-10h00)

Atelier 1 (10h00-12h30) : « Segmentation des textes ».

Atelier 2 (14h00-15h30) : « Typologies textuelles ».

Atelier 3 (15h30-17h00) : « Besoins d'évolution du portail Biblissima ».

### Vendredi 8 décembre 2023

Atelier 4 (10h00-12h30) : « Apparat critique ».

Atelier 5 (14h00-16h00) : « Règles et recommandations des grandes collections d'édition : adaptation à l'édition numérique ».

**Lieu** : Campus Condorcet (Aubervilliers), Bâtiment Recherche Nord (salle à préciser).

### Description des ateliers :

#### Atelier 1 : « Segmentation des textes »

L'objectif de cet atelier est de recentrer la problématique de l'élaboration du laboratoire de textes en se concentrant avant tout sur ce qui est commun aux textes. Cela revient à essayer de caractériser la « couche de base », en écartant, dans un premier temps, ce qui appartient aux exceptions (traitées dans des modules spécifiques indépendants). Un moyen efficace d'y parvenir est de se mettre d'accord sur la segmentation des textes, c'est-à-dire sur la manière de découper nos textes et nos documents en unités logiques faisant sens.

Nous proposons donc de mener une réflexion sur les niveaux de segmentations à définir, et nous nous efforcerons pour cela de prendre de la hauteur, en faisant abstraction des questions techniques liées à l'encodage et au balisage des données. La première étape consiste à définir une liste de mots permettant de définir différents niveaux de segmentation (en fonction d'approches textuelles ou documentaires touchant soit au fond, soit à la forme). À partir de cette liste, nous discuterons ensuite de la pertinence d'adopter ou non tel ou tel type de segmentation. Il conviendra en particulier de définir collectivement un vocabulaire commun, adapté à tous les types de sources. Ce vocabulaire sera utile plus tard, lorsqu'il s'agira d'encoder nos textes en XML-TEI.

#### Atelier 2 : « Typologies textuelles »

Marie-Luce Demonet présentera d'abord brièvement la définition des typologies textuelles produites par les membres du consortium CAHIER (accessibles dans Opentheso, hébergé par l'infrastructure de recherche Huma-num : <https://opentheso.huma-num.fr/opentheso/?idt=43>). Puis nous proposerons aux participants de caractériser un ou plusieurs de leurs textes à l'aide de la définition de ces typologies textuelles. Chaque texte

peut en effet être caractérisé par un ou plusieurs mots-clefs correspondant à une typologie de texte. S'il s'avère que les typologies proposées par CAHIER ne suffisent pas (ce qui sera peut-être le cas pour certains textes historiques ou pour certaines sources anciennes [antiques ou médiévales]), ne pas hésiter à proposer d'étendre le modèle de CAHIER, qui continue à présent d'être enrichi par le consortium ARIANE.

Au terme de la session, les participants seront en mesure de définir une collection de mots-clefs permettant de caractériser leurs textes (et éventuellement les différentes sections de textes de leurs documents). Ces informations sont celles qui leur serviront plus tard à enrichir les métadonnées de leurs textes dans le laboratoire de textes (renseignement du `teiHeader` de leurs fichiers encodés en XML-TEI). Elles pourront aussi servir à caractériser les parties internes de leurs textes et elles permettront ainsi l'annotation de certaines spécificités.

### **Atelier 3 : « Besoins d'évolution du portail Biblissima+ »**

L'objectif de cet atelier est de réfléchir collectivement aux besoins d'évolution du portail Biblissima+ afin que les sources traitées dans le laboratoire de textes (cluster 5b) soient correctement prises en compte et qu'elles soient plus facilement recherchables à partir du moteur de recherche. Pour mener cette réflexion, nous nous servons du sondage suivant : <https://forms.gle/jVKCddBk4xxLobnCA>

Dans le portail Biblissima+ (<https://portail.biblissima.fr/>), il existe déjà la possibilité de filtrer les recherches par autorités (auteurs/nom de personnes physiques), par œuvres (titre), par manuscrits (cotes), etc. Il serait sans doute pertinent d'ajouter un filtre pour prendre en compte les typologies textuelles de nos sources (cf. atelier 2). Cela permettrait aux chercheurs de regrouper leurs sources par typologie et ainsi de facilement constituer des corpus typologiques. Il convient également de s'intéresser à la manière dont les éditions numériques seront référencées dans Biblissima+ pour les retrouver facilement dans le portail. D'autres besoins doivent sans doute être identifiés et définis collectivement : cet atelier est le lieu idéal pour en discuter.

Remarque : toutes les informations indispensables au portail Biblissima+ devront être soigneusement renseignées dans les métadonnées de nos sources encodées en XML-TEI afin de favoriser leur cherchabilité (alignement au moyen du format pivot de Biblissima décrit dans le *Vademecum*).

### **Atelier 4 : « Apparat critique »**

Cet atelier vise à tester et à prolonger les discussions et les choix pris lors de l'atelier « Apparat critique » organisé à Caen le 14 juin 2023 (voir le compte rendu de cet atelier, téléchargeable ici : <https://zenodo.org/records/8380666>).

Au terme de cette session, nous serons en mesure de définir collectivement les évolutions et les améliorations souhaitées dans le module « Apparat critique » du laboratoire de textes. Pour l'heure, celui-ci s'appuie très largement sur l'environnement de travail « Apparat critique » développé par le PDN dans le cadre du projet Malaterra (cf. le manuel utilisateur « Environnement pour réaliser des éditions avec appareil critique en XML-TEI P5 »), mais il doit aussi tenir compte des expériences acquises dans ce domaine par d'autres laboratoires (IRHT, CESR, CJM, HiSoMa, Ciham, etc.). Il serait donc intéressant que chacun de ces laboratoires soit représenté lors de cet atelier.

## **Atelier 5 : « Règles et recommandations des grandes collections d'édition : adaptation à l'édition numérique ».**

Afin d'accompagner les utilisateurs du laboratoire de textes jusqu'à la publication multi-support de leurs sources, il convient de veiller à définir une manière de transposer les règles et les recommandations préconisées par les grandes collections de textes pour des éditions papier (sources chrétiennes, Belles lettres, Ecole nationale des chartes, Oxford world's classics, etc.) dans le domaine du numérique. Il s'agit notamment de définir ce qui, lors du passage de l'édition critique imprimée à l'édition critique numérique, doit être conservé et reproduit, mais aussi ce qui doit être dépassé ou inventé.

Cet atelier vise à amorcer les discussions, mais aussi à définir, avec les personnes intéressées, un lieu et une date pour organiser une journée d'étude dédiée à ce sujet au cours du premier trimestre 2024.

Il serait par exemple intéressant de vérifier, à l'aide du projet Malaterra, s'il serait possible de porter l'environnement pour le rendre utilisable quelle que soit la collection de textes retenue par le chercheur en cas de publication multisupport (tenir compte du manuel « Environnement pour réaliser des éditions avec appareil critique en XML-TEI P5 »). Il faudrait parvenir à identifier ce qui est commun aux différentes collections (voir au passage si cela pourrait servir à la segmentation des textes ; cf. atelier 1), puis réfléchir ensuite à ce qui est particulier à telle ou telle collection pour déterminer comment ces différences peuvent être soit dépassées soit préservées, mais à moindres frais pour la construction du laboratoire de textes (identifier ce qui crée la différence : s'agit-il d'objectifs scientifiques différents ou de types de textes différents ?).

### **Inscription**

- Date limite : vendredi 24 novembre 2023.
- Conditions : avoir déjà participé à ou être actuellement impliqué dans un projet d'édition critique numérique.
- Comment : s'inscrire par mail auprès de Stéphane Lecouteux (stephane.lecouteux@unicaen.fr) et d'Emmanuelle Kuhry (emmanuelle.kuhry@irht.cnrs.fr), en précisant si vous souhaitez participer aux deux journées ou seulement à une seule (soit le jeudi, soit le vendredi).
- Logistique : voir avec Stéphane Lecouteux (stephane.lecouteux@unicaen.fr) pour l'organisation de votre mission (établissement des ordres de mission, réservation des billets de train, de la restauration et de l'hébergement, gestion des frais de mission, etc.).