

Le Livre médiéval au regard des méthodes quantitatives

École d'été
Paris, 15-19 juin 2020



Direction : Christine Bénévent, Emilie Cottreau-Gabillet,
Catherine Rideau-Kikuchi et Dominique Stutzmann
Coordination scientifique et pédagogique :
Emilie Cottreau-Gabillet et Octave Julien

*organisé par le Laboratoire de Médiévistique occidentale de Paris,
l’Institut de Recherche et d’Histoire des Textes, l’École nationale des
chartes et le laboratoire Dynamiques patrimoniales et culturelles*

*avec le soutien
du LabEx Histoire et anthropologie des savoirs, des techniques et des croyances
et la collaboration de la Bibliothèque de la Sorbonne*

L'état de l'art

**Lundi 15 juin – Campus Condorcet, Centre de colloque,
salle 50 (rez-de-chaussée)**

13h15 Ouverture

13h30 Le livre au regard des méthodes quantitatives : un état de l'art

par Émilie Cottreau-Gabillet et Octave Julien

14h30 Présentation et discussion des travaux des participants (première partie)

discussion animée par l'équipe de coordination

Ateliers d'application

Mardi 16 juin – École nationale des chartes, salle Delisle

9h Observer les livres pour quantifier : les manuscrits

atelier animé par Émilie COTTEREAU-GABILLET

Cet atelier a pour objectif de montrer, pour les différents éléments constitutifs d'un livre manuscrit (support, structure matérielle, réglure et mise en page, aspects périgraphiques, décoration) ainsi que pour les données relatives au colophon (forme matérielle, structure syntaxique) ou issues de ce dernier (lieu et date de production, commanditaire, copiste, etc.), la nature des paramètres susceptibles de faire l'objet d'une quantification au regard de leur exploitation. Dans cette perspective on reviendra, pour chacun de ces éléments, sur : la nature des observations qui peuvent être menées et les protocoles de relevés qui y sont associés ; la manière dont les données peuvent être formalisées ; les usages statistiques qui peuvent être envisagés ; et enfin les perspectives d'analyses historiques liées à cette utilisation des méthodes statistiques. L'enjeu est ici de faire le lien entre l'observation des volumes et leur interprétation, dans une démarche qui doit placer la questionnement (des données comme des résultats) au cœur du propos.

13h30 Observer les lives pour quantifier : les incunables

atelier animé par Christine BÉNÉVENT

Les incunables, livres imprimés avant la fin de l'année 1500, constituent un corpus quantitativement bien défini grâce aux nombreux catalogues et entreprises bibliographiques qui ont permis de les identifier et de les décrire. Il s'agit toutefois d'un corpus hétérogène, qui ne peut s'appréhender qu'en tenant compte des enjeux techniques propres à l'imprimerie mais aussi d'une dialectique d'imitation / opposition par rapport aux manuscrits. Dans la continuité de la séance consacrée au livre manuscrit, il s'agira de se pencher sur les éléments matériels constitutifs des incunables : support et structure matérielle (cahiers, formats certes,

mais aussi et surtout choix typographiques, mise en pages et inachèvement du produit, à l'origine conçu pour être complété par des éléments manuscrits. Il s'agira donc d'étudier ce faisant l'évolution de la structure et de la physionomie du nouveau produit au cours de la cinquantaine d'années que couvre la période incunable, mais aussi de constater des pratiques différencielles selon la nature des textes et le rapport avec l'existant manuscrit.

17h Présentation et discussion des travaux des participants (deuxième partie)

discussion animée par l'équipe de coordination

Mercredi 17 juin – Campus Condorcet, bâtiment de recherche Sud, salle 0.015 (rez-de-chaussée)

9h De la modélisation à l'exploitation : les manuscrits

atelier animé par Louis CHEVALIER

Les livres d'heures, copiés à partir du XIII^e siècle pour l'usage religieux des fidèles et conservés en grand nombre dans les institutions patrimoniales et les collections privées, constituent un précieux témoignage des traditions liturgiques médiévales et des pratiques de dévotion et des représentations sociales de leurs commanditaires. 7277 livres d'heures manuscrits et 246 livres d'heures imprimés sont aujourd'hui répertoriés ; ces livres comportent différentes sections : un calendrier, des péricopes évangéliques, des offices liturgiques (heures de la Vierge, office des morts...), des suffrages et des prières. Volumineux, luxueusement décorés, ils révèlent en outre l'art de nombreux ateliers d'Europe occidentale. Les livres d'heures sont à la fois identifiés par l'usage liturgique du diocèse ou de l'abbaye dont ils ont adopté le *cursus* et par leur lieu de production, qui peuvent être distincts.

Partant de données produites au sein du projet *HORAE (Hours - Recognition, Analysis, Editions. IRHT-CNRS, TEKLIA, LS2N. 2018-2021)*, dédié à l'analyse des livres d'heures et à la reconnaissance automatique des écritures médiévales, nous étudierons au cours de cet atelier les problématiques inhérentes à la formation d'un corpus de manuscrits, et appliquerons ces réflexions à un corpus de livres d'heures, type marginal de livre liturgique. Nous formaliserons ensuite les caractéristiques matérielles des manuscrits ainsi regroupés et utiliserons certaines fonctionnalités d'*Excel* et d'outils annexes tels qu'*OpenRefine* pour importer, classer, transformer, visualiser et analyser les données créées.

14h De la modélisation à l'exploitation : les incunables

atelier animé par Catherine RIDEAU-KIKUCHI

L'objectif de cet atelier est de proposer des éléments méthodologiques relatifs au traitement des données concernant les domaines de la production, de la diffusion et de la possession du livre incunable. On se concentrera sur la constitution et l'utilisation de deux bases de données complémentaires : l'*Incunabula Short Title Catalogue* (ISTC) et *Material Evidence in Incunabula* (MEI). Nous aborderons dans un premier temps la manière dont on peut utiliser les données de l'ISTC pour constituer une base de données et à la manière dont on peut retravailler les données du catalogue et formaliser de nouvelles informations. On verra comment il est possible de traiter ces données d'un point de vue quantitatif à travers quelques outils d'*Excel*. Nous verrons dans un deuxième temps comment la base de données MEI a compilé des données matérielles des livres pour retracer le fonctionnement du commerce des incunables. La manipulation de la base et de

son outil de visualisation permettra d'approfondir la question de la mise en série de données individuelles dans le cas de l'archéologie de l'objet-livre. Dans un dernier temps, on abordera un autre type de formalisation et de constitution de base de données à partir des priviléges octroyés par les autorités publiques ou religieuses : la base de données *Book Privileges in Venice*.

17h30 Conférence invitée. Eléonore CELLARD : Les manuscrits du Coran des premiers siècles de l'Hégire : perspectives de l'approche quantitative.

Jeudi 18 juin - LAMOP, Sorbonne

9h Introduction à l'analyse statistique : l'exploration d'un corpus

atelier animé par Octave JULIEN

Cet atelier consiste en une introduction aux concepts et outils statistiques utiles à la codicologie quantitative. On travaillera pour cela sur un corpus d'environ 200 recueils copiés en France et en Angleterre entre le XIII^e et le XV^e siècle, à la forme et au contenu variés. L'hétérogénéité de ce corpus suppose une première phase d'exploration des données pour en mesurer la diversité et les possibles biais. On verra quels indicateurs statistiques (moyenne, médiane, quartiles, écart type) et quelles représentations graphiques (diagrammes en barres, histogrammes, boxplots, courbes de densité) permettent d'appréhender un tel corpus. Dans le prolongement des ateliers précédents, on reviendra aussi sur les outils permettant d'analyser une corrélation (tris croisés, comparaison d'indicateurs par sous-populations). Les méthodes d'analyses factorielles seront enfin abordées afin de construire une typologie de manuscrits.

Cet atelier se fera en salle informatique sur *R*, un logiciel d'analyses statistiques puissant, libre et gratuit. On verra comment l'utiliser pour importer ses données, les explorer, les transformer grâce à quelques méthodes simples, et produire des tableaux et graphiques.

13h30 La construction d'un protocole d'analyse : étude et interprétation de corrélations statistiques

atelier animé par Octave JULIEN

En partant des analyses de Carla Bozzolo et Ezio Ornato concernant le lien entre format des manuscrits et disposition du texte, on cherchera à comprendre ce qui a déterminé les choix en matière de mise en page et de décoration dans le corpus de manuscrits utilisé le matin. La nature diversifiée de ce corpus conduit à étudier d'autres corrélations et à proposer d'autres interprétations pour rendre compte de ces choix fonctionnels et esthétiques. On réfléchira à la manière de quantifier les éléments décoratifs des livres manuscrits. Ce sera également l'occasion de mobiliser les méthodes et concepts vus le matin, et d'aller plus loin dans la démarche quantitative en utilisant des tests statistiques destinés à confirmer ou infirmer ces hypothèses (tests du χ^2 , coefficient R^2 de corrélation linéaire).

16h30 Présentation d'une sélection de manuscrits et d'incunables de la Bibliothèque de la Sorbonne

présentation assurée par Émilie COTTEREAU-GABILLET, Catherine RIDEAU-KIKUCHI et DOMINIQUE STUTZMANN.

Vendredi 19 juin – Campus Condorcet, Centre de colloque, salle 50 (rez-de-chaussée)

9h La paléographie à l'heure de l'intelligence artificielle

atelier animé par Dominique STUTZMANN

L'objectif de cet atelier est de mettre en œuvre des technologies d'analyse d'image par ordinateur fondées sur l'intelligence artificielle et de réfléchir à l'utilisation et à la signification historique des résultats générés suite à l'apprentissage machine. L'atelier s'ouvrira sur une présentation des enjeux de la paléographie comme discipline historique étudiant l'écriture dans ses formes et ses emplois socio-spatiaux. Plus particulièrement, le cas de l'écriture des livres d'heures permettra de faire le lien avec les ateliers des mardi et mercredi matin. Ensuite, dans la partie pratique, les participantes manipuleront des images disponibles selon le protocole IIIF (visualisation et import), puis exploiteront les résultats produits dans le cadre des compétitions internationales d'analyse d'image d'écritures médiévales (<https://clamm.irht.cnrs.fr/>) pour discuter les enjeux de la visualisation de données, de classification et de discréttisation dans les différentes tâches de la paléographie (paléographie d'expertise), en particulier la reconnaissance des mains, la classification des écritures, et la datation ou localisation des manuscrits.

Enfin les participantes testeront des logiciels, notamment Transkribus, qui ont été pré-entraînés pour classer les types de page, analyser leur mise en page ou aligner image et texte édité. Ils expérimenteront également les fonctions de reconnaissance d'écriture manuscrite (transcription automatique) et compareront les résultats sur différentes écritures.

14h Bilan du cours et remise des certificats de participation

Accueil limité à 25 inscrits. Formulaire d'inscription en ligne <https://forms.gle/n1cZyrsygrKuhng46>

Clôture des inscriptions : 28 mars 2020 - Sélection des dossiers : 3 avril 2020

Cette école d'été s'adresse, en priorité, aux chercheurs en formation (masterants, doctorants et post-doctorants). Des bourses, destinées à rembourser des frais de transport et de logement, pourront être distribuées, à hauteur d'un plafond qui sera déterminé lors de l'examen des dossiers, en fonction du nombre et de l'origine géographique des candidats sélectionnés.