

---

# *Le Livre médiéval au regard des méthodes quantitatives*

---

---

*École d'été*

---

*Paris, 12-16 juin 2017*

---



*Direction : François Foronda, Maria Gurrado,  
Christine Bénévent et Xavier Hermand  
Coordination scientifique et pédagogique :  
Émilie Cottereau-Gabillet,  
Octave Julien et Laura Albiero*

*organisé par le Laboratoire de Médiévisstique occidentale de Paris,  
l'Institut de Recherche et d'Histoire des Textes, l'École nationale  
des Chartes et l'Université de Namur,*

*avec le soutien*

*du LabEx Histoire et anthropologie des savoirs, des techniques et des croyances et  
du Conseil scientifique de l'Université Paris 1 Panthéon-Sorbonne*

---

# L'état de l'art

---

**Lundi 12 juin - IRHT**

**13h45 Ouverture**

*par François BOUGARD*

**14h Histoire culturelle et approche quantitative**

*par Jean-Philippe GENET*

**14h45 Mesure et imagination, les indispensables du quantitatif**

*par Ezio ORNATO*

**15h30 Formalisation, statistique, sémantique**

*par Alain GUERREAU*

**16h30 Principes de statistique**

*par Gaetan BONNOT*

**Mardi 13 juin - IRHT**

**9h Présentation des travaux des participants**

*discussion animée par François FORONDA et Xavier HERMAND*

---

## Ateliers d'application

---

**Mardi 13 juin - Bibliothèque de l'Arsenal**

**14h Présentation de quelques manuscrits et incunables**

**15h30 Analyse paléographique et méthodes quantitatives**

*atelier animé par Maria GURRADO et Denis MUZERELLE*

L'objectif de l'atelier de paléographie est de montrer comment regarder l'écriture et formaliser les données d'ordre paléographique en vue d'un traitement statistique. L'atelier se divise en deux temps. Le premier concerne la méthodologie d'analyse de l'écriture et le protocole d'observation. On se posera les questions suivantes : Comment regarde-t-on une écriture médiévale ? Existe-il plusieurs façons de procéder ? Est-ce que les variables à observer sont différentes pour chaque écriture et/ou pour chaque main ? Comment les choisir et comment les formaliser ? Existe-il une manière idéale de représenter graphiquement les données ? Comment choisir la représentation la plus adaptée ? Comment interpréter les données et leurs éventuelles corrélations ? Le deuxième temps consistera en des exercices pratiques sur l'écriture du XI<sup>e</sup> siècle en France et la question des critères de datation. Ces exercices porteront sur l'observation d'échantillons choisis, puis sur la mise en œuvre de tests d'application à partir de nouveaux échantillons. Les résultats seront à chaque fois vérifiés et discutés.

### 9h **Ouverture**

*par Michelle BUBENICEK*

### 9h15 **Observer les livres pour quantifier : les manuscrits**

*atelier animé par Emilie COTTEREAU-GABILLET*

Cet atelier a pour objectif de montrer, pour les différents éléments constitutifs d'un manuscrit (support, structure matérielle, réglure et mise en page, aspects périgraphiques, décoration), 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 ; sur la manière dont les données peuvent être formalisées ; sur les usages statistiques qui peuvent être envisagés ; et enfin sur les perspectives d'analyse liées à cette utilisation des méthodes statistiques.

### 14h **Observer les livres pour quantifier : les incunables**

*atelier animé par Ezio ORNATO et Christine BÉNÉVENT*

La séance consacrée aux incunables sera divisée en deux parties : dans la première, on insistera sur l'évolution de la structure et de la physiologie du nouveau produit, notamment en ce qui concerne les enjeux techniques et la dialectique d'imitation/opposition par rapport au livre manuscrit. Dans la seconde, on abordera l'analyse d'une production livresque très répandue – les recueils de sermons en langue latine – en introduction à l'étude quantitative qui sera présentée le jour suivant. L'exposé s'appuiera sur les très nombreuses reproductions d'incunables accessibles sur le Web, ainsi que sur les quelques exemplaires conservés à l'Ecole des chartes.

## Jeudi 15 Juin - Lamop, campus de la Sorbonne

### 9h **Ouverture**

*par Philippe BERNARDI*

### 9h15 **Le livre en données**

*atelier animé par Laura ALBIERO et Ezio ORNATO*

Cet atelier s'articule en deux parties complémentaires. La première a pour objectif d'exposer la manière de concevoir une base de données portant sur les manuscrits et de formaliser certaines informations codicologiques. Elle comprendra d'abord une présentation des possibilités d'exploitation des données à partir des recherches déjà menées et dont les résultats se sont avérés significatifs ; elle se focalisera ensuite sur les aspects dimensionnels des manuscrits et sur l'interrelation entre le texte et la structure matérielle. La deuxième partie est consacrée à l'analyse du corpus de 440 recueils de sermons imprimés avant 1501 présenté au cours de l'atelier du mercredi après-midi. A partir des données rassemblées (adresse typographique, format, nombre de feuillets, nombre de lignes par page, nombre et répartition géographique des exemplaires survivants), on montrera comment ce corpus pourrait rendre compte de certains aspects de l'évolution de la typographie européenne, et surtout témoigner de la réalité des pratiques éditoriales, s'échelonnant de la concurrence acharnée au partage implicite du marché.

## 14h **La construction d'une démarche statistique : la disposition du texte**

*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 le choix de la copie à longues lignes ou en colonnes dans un corpus d'environ 160 recueils datés du XIII<sup>e</sup> au XV<sup>e</sup> siècle. On verra ainsi les différentes étapes nécessaires à la construction d'une démarche statistique : exploration et description de l'échantillon, discrétisation et recodage des variables, mesure des corrélations entre variables grâce à des tests statistiques, et analyse comparée de ces corrélations pour évaluer l'importance des différents facteurs dans le choix de la disposition du texte. On s'initiera pour cela à R, un puissant logiciel d'analyses statistiques, en montrant comment importer ses données, les explorer, les transformer grâce à quelques méthodes simples, et produire des tableaux et graphiques.

**Vendredi 16 Juin - Lamop, Campus CNRS de Villejuif**

## 9h **Livres, textes et statistiques**

*atelier animé par Alain GUERREAU*

L'environnement technique totalement inédit qui s'est développé depuis le milieu des années 80 offre des possibilités elles aussi inédites pour l'étude des textes anciens. La condition préalable aux opérations analytiques est la construction de corpus de taille suffisante (plusieurs millions, voire plusieurs dizaines de millions de mots). Quelques corpus sont disponibles, d'autres en préparation, et les outils de construction ad hoc s'améliorent. La difficulté de l'élaboration de procédures statistiques efficaces tient au caractère strictement non-gaussien des distributions lexicales, et à l'insuffisance de notre compréhension des mécanismes sous-jacents qui régissent les fréquences et les consistences lexicales. Cependant, un champ de recherche très vaste et très actif s'est développé sous la bannière du "text-mining", les résultats sont impressionnants. Les historiens et les philologues doivent identifier dans ce champ les outils qui peuvent leur être utiles, et les adapter autant que nécessaire. L'examen des évolutions des fréquences relatives (correctement calculées) apporte déjà beaucoup. Mais la perspective la plus prometteuse est celle dite de la "sémantique distributionnelle", qui permet d'exploiter de diverses manières les fréquences des cooccurrences. C'est sur ce point qu'il convient d'insister, et de présenter les résultats déjà obtenus.

## 11h **Production, diffusion et possession du livre médiéval**

*atelier animé par Catherine 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 médiéval. En plus des remarques relatives aux bibliothèques, et plus particulièrement celles d'Angleterre, dans la conférence introductive sur la culture écrite, cet atelier se focalisera sur l'imprimerie vénitienne, premier pôle de production incunable. Sera envisagé plus spécialement ici la question de l'exploitation de bases de données bibliographiques (ISTC, GW, USTC) à partir de méthodes quantitatives. Seront abordées également des questions telles que la constitution de bases de données *ad hoc* à partir d'autres sources : privilèges, textes des colophons etc.

## 14h **Bilan du cours et remise des certificats de participation**

Accueil limité à 22 inscrits

Formulaire d'inscription à remplir en ligne <https://goo.gl/forms/dCd38JZ13g4E6xlq1>

Clôture des inscriptions : 13 avril 2017 - Sélection des dossiers : 20 avril 2017

Des remboursements de frais de transport et de logement pourront être réalisés, à hauteur d'un plafond qui sera déterminé en fonction du nombre et de l'origine géographique des inscrits