

**Retour sur le séminaire de l'École des Chartes
« Sciences du patrimoine/sciences du texte : confrontation des méthodes »**

Le séminaire « Sciences du patrimoine/sciences du texte : confrontation des méthodes », organisé par l'École des Chartes en partenariat avec IESF, s'est clos le 10 juin sur deux séances illustrant parfaitement l'intérêt du croisement entre ingénierie, numérique, sciences du texte et sciences du patrimoine.

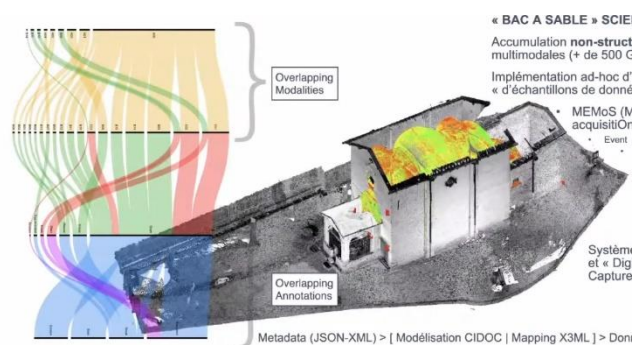


Ce séminaire est animé par Thibault CLÉRICE, responsable du Master « Technologies Numériques Appliquées à l'Histoire » (TNAH), membre de notre comité.

Il a duré 5 séances entre février et juin, nous vous en présentons les dernières, les premières ayant été évoquées dans un précédent Flash Info.

Si vous souhaitez revoir les séances, une playlist YouTube est mise à disposition à cette adresse : https://www.youtube.com/watch?v=Sx1IQBswFnQ&list=PLayqWLSo_nPU3_Ux7rBdDMMfzWQ3tUET6

Après des séances centrées autour de la donnée, de sa conservation et de son partage, incluant plusieurs membres de notre comité, dont Laurent ROMARY et son président Loïc BERTRAND, la question de l'analyse et de la valorisation fut mise au centre du débat dans deux séances.



Dans la séance du 8 avril (<https://www.youtube.com/watch?v=Sx1IQBswFnQ>), Nicolas Perreaux et Anthony Pamart nous ont présenté leur travail sur l'analyse de données. Ce dernier présenta ses travaux en cours sur l'accumulation de données 3D (suivant plusieurs technologies) et le problème que représente cette production de masse de données dans son traitement.

Dans cette perspective, il nous a proposé de nous plonger dans les questions de documentation de ces données en vue de leur fusion pour l'analyse y compris à travers des outils tels qu'AIOLI développé en interne à l'unité de recherche MAP (3495) dépendante du CNRS et du ministère de la Culture.

Puis, Nicolas Perreux, ingénieur de recherche, nous a montré comment, à partir de données textuelles il était possible de reconstruire ou questionner les modèles de pensées voire l'histoire économique du moyen-âge français.

À travers la fouille textuelle automatisée, et dans une tradition de l'histoire quantitative, il a proposé plusieurs exemples issus de ses travaux, comme celui de la remise en cause de la période principale pour l'apparition de moulins ou bien celui de l'évolution de l'habitat entre 600 et 1300 dans les traces textuelles.

Dans la séance du 20 mai (en cours de mise en ligne), les deux interventions se sont concentrées sur la question de la reconnaissance manuscrite automatisée de documents anciens et celle de la reconstitution semi-automatique de Versailles.

Alix Chagué, membre de l'INRIA, nous a présenté le projet de collaboration CREMMA (Consortium pour la Reconnaissance d'Écriture Manuscrite de Matériaux Anciens), <https://www.dim-map.fr/projets-soutenus/cremma/>) dont l'objectif est de mettre en commun infrastructure numérique, compétences traditionnelles des philologues pour la transcription de textes et compétences numériques des ingénieur-e-s.

En effet, depuis quelques années, la reconnaissance d'écritures manuscrites a largement profité du développement des réseaux de neurones et de l'apprentissage profond, au point où des études peuvent désormais se faire sans transcription manuelle de documents manuscrits du 13e siècle (cf. <https://www.youtube.com/watch?v=Pc0srBRx5zY>).

Ce développement et ce rassemblement de compétence sont centraux dans l'exploitation et la redécouverte pour le public des masses de données littéraires ou administratives présentes dans nos institutions de conservation du patrimoine, comme par exemple le minutier central de Paris sur lequel A. Chagué travaille en collaboration avec les Archives Nationales (<https://lectarep.hypotheses.org/>).

Puis, Michel Jordan et Nicolas Priniotakis (Université de Cergy, laboratoire ETIS) nous ont présenté leur travail sur le projet VERSPERA qui vise à reconstruire, à partir de plans d'archive numérisés du château de Versailles, les pièces de ce dernier, y compris temporaires comme le petit théâtre de Louis XV.

À travers un travail fin d'ingénierie visant à exploiter les plans 2D automatiquement en 3D puis à un travail de reconstitution par connaissance historique du mobilier et des décorations, les enseignant-e-s et étudiant-e-s de la Licence pro Patrimoine, visualisation et modélisation 3D ont permis de mettre en lumière des pièces n'ayant existé que sur plan (avec leurs défauts de conception !) ou bien ayant été détruit depuis leur construction. Pour voir des exemples de ces reconstitutions: <https://verspera.hypotheses.org/modelisation-3d/reconstitutions-virtuelles-partenariat-avec-la-licence-professionnelle-des-metiers-du-numerique-cy-cergy-paris-universite> .

Enfin, la séance du 10 juin (mise en ligne en cours) a mis à l'honneur la question du prédictif dans le contexte de la conservation des manuscrits. Tout d'abord, Jean-Baptiste Camps nous a présenté ses travaux alliant modèles de zoobiologie et transmission des textes et manuscrits anciens.

Si nous connaissons en effet Tite-Live, Sénèque, Virgile ou Homère, c'est en effet parce que les copistes, les bibliothèques et les imprimeurs ont permis leur survie à travers leur protection ou leur production.

Mais pour ces quatre auteurs que le grand public moderne connaît bien, combien ont disparu à travers les accidents, mais aussi l'évolution des goûts ou les priorités de tel ou tel lieu de copie ? C'est sur ce thème que le présentateur propose de retracer l'historique et les problèmes que cette question comprend tout en proposant une nouvelle approche liée aux théories de la sélection naturelle.

Puis, Valérie Lee et Michel Jordan de l'université de Cergy ont montré leurs premiers résultats sur la détection et classification automatique de défauts sur les reliures conservées aux Archives Nationales. Analyser une reliure et les problèmes qu'elle peut présenter représente aujourd'hui 10 min de travail par reliure, or, l'état de ces reliures doit être connu pour savoir si l'institution peut prêter ou laisser à disposition l'ouvrage.

Aux Archives Nationales seules, on compte environ 370 km (équivalent de Paris-Nantes) d'étagères d'archives (appelés kilomètres linéaires). En développant des outils basés sur la reconnaissance d'image, Valérie Lee et ses collègues ont réussi à proposer des premiers prototypes permettant de reconnaître automatiquement, sur la base d'une photographie, l'état de conservation des reliures d'un ensemble d'ouvrages, directement sur l'étagère.



Ces trois séances furent riches, y compris lors des phases de questions et de réponses. Pour le moment, le renouvellement du séminaire n'est pas décidé pour l'an prochain, mais si ces questions vous intéressent, nous vous conseillons les séminaires suivants :

- Les lundis numériques de l'Institut Nationale de l'Histoire de l'Art (<https://www.inha.fr/fr/recherche/programmation-scientifique/en-2020-2021/les-lundis-numeriques-2020-2021.html>)
- Les séminaires et conférences Humanités Numériques de l'École nationale des Chartes (<http://www.chartes.psl.eu/>)
- Le séminaire DH AI (<https://dhai-seminar.github.io/>)
- Les conférences et événements promus par l'association Humanistica des Humanités numériques francophones (www.humanisti.ca)
- Les événements du Domaine d'Intérêt Majeur Matériaux Anciens et Patrimoine (<https://www.dim-map.fr>)

Et en général, si ce domaine attise votre curiosité, le domaine des humanités numériques !

Thibault Clerice

Membre du Comité Patrimoine d'IESF