

Séminaire de l'Axe de recherche en histoire numérique 2023-2024

Extraire, visualiser et analyser des données

Les séances du séminaire se déroulent en hybride et sont ouvertes à toutes et tous sur inscription préalable auprès de vincent.alamercery@ens-lyon.fr. Précisez lors de votre inscription si vous souhaitez suivre la séance en visio pour recevoir le lien.

8 février 2024 10h-12h

Kraken, eScriptorium et l'ATR (automatic text recognition) : Naviguer dans les Océans de la Reconnaissance Automatique d'Écriture

L'émergence d'outils de reconnaissance automatique d'écriture avec une interface utilisateur de plus en plus intuitive, tels que Transkribus ou eScriptorium, ouvre aujourd'hui de nouvelles perspectives. De nombreuses équipes de recherche se sont lancées dans l'utilisation de l'ATR et l'entraînement de modèles de plus en plus performants sur les sources historiques, capables de reconnaître efficacement une grande variété de documents qu'ils présentent des écritures différentes, ou qu'ils proviennent de périodes historiques et d'aires linguistiques variées, modernes comme médiévales. Toutefois, malgré les progrès significatifs réalisés, des défis subsistent pour le traitement des documents historiques complexe ? Comment obtenir des transcriptions automatiques qui respectent des critères scientifiques ? Comment analyser la mise en page ? En conclusion, si la reconnaissance automatique d'écriture offre la possibilité d'explorer des tailles de corpus inédites, elle requiert toujours une certaine expertise pour relever les défis inhérents à la diversité et la complexité des documents médiévaux, limites et enjeux que nous aborderons dans cette présentation.

- Intervenante : [Ariane Pinche](#)
 - Chargée de recherche au CNRS, Ariane Pinche est spécialisée dans l'édition numérique de sources médiévales, la constitution de corpus et l'entraînement de modèles pour améliorer la reconnaissance automatique de texte (HTR). Elle est coresponsable de l'[axe transversal Corpor@Tech-Science ouverte](#) du CIHAM. Ariane Pinche a aussi enseigné l'édition numérique (XML TEI, XSLT) dans le master « Technologies numériques appliquées à l'histoire » de l'École des Chartes.

Extraire du texte des imprimés anciens : quels défis, quelles promesses ? Illustration par le projet AGODA

Intégralement retranscrits dans le *Journal Officiel* depuis 1881, les débats parlementaires anciens constituent des sources précieuses pour l'histoire. Ces documents numérisés et accessibles via Gallica, restent toutefois difficiles d'accès. Durant cette présentation, Marie Puren évoquera les défis rencontrés par le projet AGODA pour traiter et rendre accessibles ces documents.

- Intervenante : [Marie Puren](#)
 - Maîtresse de conférences à l'EPITA où elle dirige l'équipe [Méthodes numériques en Sciences Humaines et Sociales](#), Marie Puren est une historienne spécialiste d'histoire intellectuelle et politique et d'histoire de la littérature de jeunesse sous la Troisième République, et une historienne numérique. Elle est aussi porte-parole de l'association [Humanistica](#).
- Lieu : Lyon, [MSH Lyon-Saint-Étienne](#), salle Élise Rivet.

18 avril 2024 10h-12h

RiC-O et la question de l'articulation des ontologies pour les archives et la recherche [Titre provisoire] [résumé à venir]

- Intervenante : [Florence Clavaud](#)
 - Conservatrice générale du patrimoine, Florence Clavaud est responsable du [Lab des Archives nationales de France](#) à la Direction numérique et conservation et membre exécutif du Groupe d'expert sur la description archivistique du Conseil international des archives (ICA/EGAD) où elle dirige l'équipe de développement de [RiC-O](#). Florence Clavaud est aussi chercheuse associée au Centre Jean-Mabillon.
 - Lieu : Lyon, [MSH Lyon-Saint-Étienne](#), salle Élise Rivet.

30 mai 2024 10h-12h

Analyser les trajectoires socioprofessionnelles d'ouvriers gantiers et métallurgistes grenoblois à l'aide des outils numériques : enjeux et méthodes

Cette intervention est liée à une recherche doctorale qui se donne pour objectif d'analyser les déterminants familiaux et individuels de l'évolution des activités économiques à l'échelle du territoire grenoblois, en articulant les dimensions macroscopiques et les trajectoires socioprofessionnelles d'ouvriers gantiers et métallurgistes. Après avoir posé le contexte et les problématiques, il s'agira de revenir sur les sources, les méthodes et les outils employés pour analyser les cohortes de travailleurs constituées.

- Intervenant : Mathieu Rivero
 - Mathieu Rivero est doctorant à l'université Paris Nanterre, laboratoire [IDHES](#) où il réalise une thèse en histoire sous la direction de Xavier Vigna et de Laure Machu, "Les familles ouvrières en question(s). Expériences, trajectoires et pratiques (Grenoble, années 1850-Seconde Guerre mondiale)" incluant une forte dimension numérique, à la fois quantitative et qualitative.
 - Lieu : Grenoble [salle à venir].