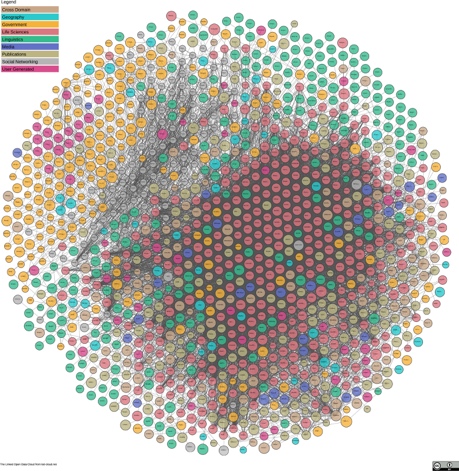
**Indexer les corpus numériques**

Atelier organisé par

Emmanuelle Perrin (UJM–IHRIM) et Fabienne Vial-Bonacci (CNRS-IHRIM)



**1er février et 29 mars 2019**

**IHRIM - Université Jean Monnet-Saint-Étienne**

L’édition critique est l’un des axes structurants de l’Institut d’histoire des représentations et des idées dans les modernités (IHRIM - UMR 5317). Le laboratoire mène en particulier de nombreux projets d’éditions numériques, pour lesquels il jouit d’une véritable reconnaissance. Dans ce domaine en pleine expansion des Humanités numériques, le partage des compétences et la diffusion des bonnes pratiques se révèlent essentiels.

Pour l’édition numérique de corpus, la constitution d’index est un véritable enjeu scientifique et technique, pour lequel il existe encore peu de documentation. Les index relèvent, avec l’annotation, de l’exploitation scientifique des textes édités : ils sont des outils de navigation dans le corpus et participent à sa compréhension en identifiant les personnes, les lieux et les œuvres mentionnés, ainsi que les matières traitées. L’exploitation de ces données aide par exemple à visualiser des réseaux de relations.

Les référentiels et les notices d’autorités permettent d’uniformiser et de consolider les index grâce à l’emploi d’un vocabulaire contrôlé et de formes normalisées. Ils apportent des solutions au traitement des variantes des patronymes, des toponymes et des titres d’ouvrages. Dans la perspective *du Linked Open Data*, l’alignement sur des référentiels aide à enrichir les données d’un projet en les complétant avec d’autres ressources et bases de données (data.bnf.fr, IdREF, VIAF, ISNI).

L’atelier, fondé sur le partage d’expériences d’ingénieurs et de chercheurs impliqués dans l’édition de corpus numériques, vise à définir les bonnes pratiques utiles à l’établissement d’index, pour identifier et documenter de manière univoque une notion ou une entité nommée : quels sont les référentiels à utiliser ? Comment valoriser au mieux son corpus ? Comment structurer ces données ? Quelles sont les solutions techniques ?

Le programme de cet atelier propose des présentations générales, des échanges autour de projets en cours et une table-ronde technique, animée par des ingénieurs. Les participants sont invités à suivre les deux journées pour faire part de leurs questions lors de la table-ronde.

*Contact*

[fabienne.vial@univ-st-etienne.fr](mailto:fabienne.vial@univ-st-etienne.fr) (Projet Marc Michel Rey)

[emmanuelle.perrin.touche@univ-st-etienne.fr](mailto:emmanuelle.perrin.touche@univ-st-etienne.fr) (Projet Charles Perrault)

**Programme provisoire**

**Séance 1 |1er février 2019 : Des référentiels à la publication numérique**

9H30 - 12H30 : Pratiques des référentiels

*Présidence de séance : Philippe Colantoni (UJM, VP numérique)*

1/ François Mistral (ABES, Responsable de d’IdRef) : De l’intérêt d’IdRef et des notices d’autorité.

2/ Francesco Beretta (LARHRA, Responsable de l'axe Pôle histoire numérique) : SyMoGIH et l’alignement avec les autorités IdRef.

3/ Catherine Fillon (UJM - Cercrid ) : Siprojuris, un exemple concret l’utilisation d’IdREf

14H – 17H : Indexation et corpus numériques (1)

*Présidence de séance : Thomas Lebarbé (Université de Grenoble, responsable du Consortium Cahier)*

1/ Emmanuelle Morlock (HISOMA, CNRS) : Naviguer dans un corpus numérique avec la TEI

2/ Thierry Joliveau et Noémie Boegli (UJM-EVS) : Pour une cartographie romanesque de Paris au XIXe siècle.

3/ Laurent Capelli (Huma-Num, CNRS) : Le traitement des données par la plateforme Isidore

**Séance 2 | 29 mars 2019 : Solutions techniques pour la constitution d’index**

9H30 - 12H30 : Schémas d’indexation

*Présidence de séance : Norbert Grataloup (UJM, directeur de la DSI)*

*1/ Conférence d’ouverture :* Lou Burnard*:* La TEI et la structuration des index (sous réserve)

Indexation et corpus numériques (2)

2/ Thierry Pécout : (UJM-LEM CERCOR) : Présentation du projet « Etudes angevines »

3/ Michel Bellet (UJM GATE) : Saint-Simon et l’économie

14H – 17H : table-ronde technique sur la question des index :

1/Léopold Boyer (IHRIM) : TEI et indexation dans le corpus de Charles Fontaine

2/Maud Ingarao et Pierre-Yves Jallud (ENS IHRIM) : solutions XML

3/ Gérald Chatagnon (DSI UJM) et Ahmad Fliti (CNRS-LEM-CERCOR) : Bestiole

4/ Dominique Rouger (BU UJM) : SolR

5/ Nathalie Arlin (AHN, ENS) : OAI-PMH

Synthèse du séminaire : Christelle Bahier-Porte (UJM-IHRIM)