

ATELIER DU CONSORTIUM CAHIER (version provisoire)

Rétro-numérisation de documents historiques et partage dans le Web sémantique : l'exemple de la lexicographie

26 - 29 juin 2018

L'atelier 2018 du Consortium CAHIER, qui se déroulera à Montpellier, s'adresse à des chercheurs, enseignants, ingénieurs ou étudiants impliqués dans des projets de valorisation de documents patrimoniaux. Il alternera des présentations de projets et des formations à des outils destinés à la numérisation, à la structuration et à l'annotation de corpus. Plusieurs environnements de mise en ligne de ressources lexicales, basés sur les récents standards du Web des données ou des technologies de la langue, seront abordés.

Lieu : Route de Mende, Université Paul-Valéry Montpellier 3, bâtiment Marc Bloch (E)

Accueil des participants : 25 juin, 19h30

Résidence Goélia Sun City (tramway : arrêt Place de l'Europe)

JOUR 1 - Mardi 26 juin

« Lexicographie et Web sémantique »

9h00 – 12h30 :

« TEI et dictionnaires, TEI-Lex0 » - Laurent Romary (Inria, DARIAH, Berlin)

« Projet Basnage » - Geoffrey Williams (LiCoRN, Lorient)

Table ronde

14h00 – 17h30 :

« One Document Does-it-all : comment maîtriser le tigre TEI ? » - Lou Burnard (TEI, Oxford)

en parallèle :

1. « VÉgA - Vocabulaire de l'Égyptien Ancien » - Frédéric Servajean (ASM, Labex ARCHIMEDE, UPV)
« Le projet LexArt, théorie de l'art » - Michèle-Caroline Heck (CRISES, UPV)
2. « ODD: session hands-on » - Lou Burnard (TEI, Oxford)

JOUR 2 - Mercredi 27 juin

« Projets d'humanités numériques développés à Montpellier »

9h00 – 12h30 :

« Données liées et annotation de corpus - ressources lexicales et ontologiques » - Francesca Frontini (Praxiling)

« Petit Thalamus, ACTO » - Gilda Caiti Russo (LLACS, UPV)

« Accès aux sources patrimoniales : le cas des projets MHARS et Philosophie cl@ndestine » - Maria Susana Seguin (IHRIM, IUF, UPV)

Table ronde (participation de la MSH Sud)

14h00 – 17h30 :

« Extraction automatique d'information lexicale à partir de ressources numérisées avec GROBID dictionaries (présentation) » - Mohamed Khemakhem (Inria, Centre Marc Bloch, Berlin)

« TXM : nouvelles fonctionnalités d'annotation de corpus et de création de ressources lexicales » - Alexei Lavrentev (IHRIM, ENS Lyon)

en parallèle :

1. « GROBID dictionaries (manipulation 1/2) » - Mohamed Khemakhem (Inria)
2. « TXM : travaux pratiques » - Alexei Lavrentev (IHRIM, ENS Lyon)

JOUR 3 - Jeudi 28 juin

Sciences du texte, analyse du discours et humanités computationnelles

9h00 – 12h30 :

« Projet Phraseorom, genres textuels » - Sascha Diwersy (Praxiling), Olivier Kraif (LIDILEM, Grenoble)

« NENUFAR : Nouvelles Éditions NUMériques de FAc-similés de Référence » - Hervé Bohbot (ASM, Praxiling), Agnès Steuckardt (Praxiling)

« Données diachroniques et classifications » - Sascha Diwersy, Giancarlo Luxardo (Praxiling)

14h00 – 17h30 :

en parallèle :

1. « GROBID dictionaries (manipulation 2/2) » - Mohamed Khemakhem (Inria)
2. Réunions des groupes CAHIER

JOUR 4 - Vendredi 29 juin à la MSH Sud

Site St Charles 2, rue Professeur Henri Serre (tramway : arrêt Albert 1^{er})

9h00 – 12h30 :

« De l'OCR à XML (FineReader, ScanTailor) » - Hervé Bohbot (ASM, Praxiling)

Bilans, perspectives et retours des réflexions des groupes de travail

Organisation : Praxiling UMR 5267

(Agnès Steuckardt, Francesca Frontini, Giancarlo Luxardo, Sascha Diwersy)

Contact : projet-nenufar@univ-montp3.fr

Inscription : <https://goo.gl/jGSxXp>