

Cosme² (Consortium Sources Médiévales 2)
Groupe de travail « Lemmes » - Atelier 3

Paris - IRHT - Salle Jeanne Vielliard - 40 avenue d'Iéna (métro Iéna)
10 décembre 2018 - 10h-18h

« Lemmes » est l'un des cinq groupes de travail du *Consortium des Sources Médiévales - Cosme²* - soutenu et financé par le TGIR Huma-Num du CNRS pour œuvrer à la constitution de corpus et d'outils numériques, consacré aux sources médiévales, piloté depuis l'IRHT. Le groupe de travail « Lemmes » a pour dessein le développement des outils existants et la diffusion de corpus lemmatisés.

L'objectif de ce troisième atelier est de continuer à développer les actions du groupe dans trois axes : 1) l'évaluation des outils de lemmatisation pour plusieurs langues médiévales (latin, français, anglais, allemand...); 2) la diffusion auprès de la communauté scientifique d'une information structurée sur les différents outils et paramètres de lemmatisation et sur des corpus lemmatisés librement disponibles ; 3) la formation des chercheurs et jeunes chercheurs aux techniques de formalisation et d'exploitation de corpus textuels médiévaux.

L'atelier est ouvert à tous les intéressés qui sont priés, pour des raisons d'organisation matérielle, de nous informer au plus vite de leur participation (écrire à eliana.magnani@univ-paris1.fr).

Programme indicatif

10h-12h30 - Diffusion et formation

- Tour de table introductif
- Tableaux et fiches sur les outils et paramètres de lemmatisation (et autres formes de diffusion)
- Organisation de la formation à la lemmatisation

12h30-13h30 - Buffet

13h30-18h - Outils de lemmatisation : développements, tests, comparaisons

- les développements de Collatinus pour l'analyse syntaxique, le latin médiéval et les fonctionnalités particulières du tagueur-LASLA (Y. Ouvrard, Ph. Verkerk)
- la constitution du corpus épigraphique bourguignon plurilinguistique et les premiers tests de lemmatisation - CBMA-CIFM (P. Brochard, E. Ingrand-Varenne, E. Magnani, A. Menudier, N. Perreaux)
- le lemmatiseur-tagueur Hydra pour l'allemand (L. Gagala)
- la constitution d'un corpus en allemand (S. Brisbois)
- les expériences de lemmatisation de l'anglais avec PALM (A. Mairey)

Organisation : Eliana Magnani (CNRS, LaMOP) et l'équipe administrative de l'IRHT.

Liens utiles :

<http://cosme.hypotheses.org/>

<https://www.huma-num.fr/consortiums>

<https://www.irht.cnrs.fr/>

<https://outils.bibliissima.fr/fr/collatinus-web/>

<https://github.com/Lukasz-G/Hydra>

<http://palm.huma-num.fr/PALM/>