

**Cosme<sup>2</sup> (Consortium Sources Médiévales 2)**  
**Groupe de travail « Lemmes » - Atelier 2**

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

« Lemmes » est l'un des sept groupes de travail du *Consortium des Sources Médiévales - Cosme<sup>2</sup>* - 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.

À l'occasion de ce deuxième atelier seront présentées deux recherches récentes sur les entités nommées, et on discutera de la mise en œuvre de corpus-test pour l'analyse comparée des différents paramètres et lemmatiseurs des langues médiévales (latin, français et anglais) ainsi que de la structure et diffusion des fiches descriptives des outils de lemmatisation et des corpus historiques lemmatisés librement disponibles.

**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](mailto:eliana.magnani@univ-paris1.fr)).

**Programme :**

- 10h : Tour de table introductif

**Autour des « entités nommées »**

- 10h30 : Approche multi-niveaux pour la reconnaissance des entités nommées en Moyen Français - Mourad Aouini (CNRS - CLT)
- 11h30 : Pause
- 11h45 : La récupération automatique des entités nommées dans les chartes médiolatines. Modélisation et perspectives d'utilisation - Sergio Torres (UVSQ - DYPAC)
- 12h45 : Buffet

**Autour de actions**

- 14h : Corpus-test et l'évaluation des outils de lemmatisation
- 15h : Opera latina, le corpus du Laboratoire d'Analyse Statistique des Langues Anciennes : contenu, formats de fichiers, méthodes de lemmatisation et d'annotation morphosyntaxique (LASLA) - Dominique Longrée, Margherita Fantoli (Université de Liège)
- 15h30 : Pause
- 15h45 : Actions de diffusion et d'information scientifique : fiches descriptives des outils de lemmatisation et des corpus historiques lemmatisés librement disponibles
- 17h30 : Bilan de l'atelier et perspectives
- 18h : Fin de l'atelier

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

**Liens utiles :**

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

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

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

<http://web.philo.ulg.ac.be/lasla/>

<https://lamop.univ-paris1.fr/menu-haut/presentation-de-lunite/annuaire/chercheurs-associes/aouini-mourad/>

<http://www.dypac.uvsq.fr/torres-aguilar-sergio-359917.kjsp>