

Atelier de création et utilisation de base de données textuelles lemmatisées en latin

9 et 10 avril 2013

10h-18h

Dijon, Maison des Sciences de l'Homme, salles R02 et R03

Le projet ANR Omnia (Outils et Méthodes Numériques pour l'Interrogation et l'Analyse des textes médiolatins) organise un atelier de deux jours d'initiation aux applications développées par le programme, sous la direction d'Alain Guerreau (CNRS).

Il s'agit de montrer, en pratique, comment mettre en forme des textes latins numérisés puis les exploiter. On aura donc deux volets, un jour sera consacré à chacun.

1. Premier jour : créer une base de données.

On part de fichiers numérisés, dans un format (presque) quelconque.

Il faut d'abord analyser de près le matériau disponible, puis mettre en oeuvre une suite d'opérations qui doivent donner d'un côté des textes lemmatisés et formatés, et de l'autre des informations sur les textes considérés, en particulier leur date.

Cela fait, on doit formater le résultat pour créer une base CWB, soit directement soit en passant par TXM.

2. Deuxième jour : exploiter une base de données textuelles sous CWB.

Trois interfaces sont disponibles : une interface en ligne de commande, puissante : cqp ; une interface totalement graphique, TXM ; une interface par l'intermédiaire du logiciel statistique R et de son espace de travail (rcqp).

On examinera en particulier la question de l'évolution des fréquences, et celle de la recherche et de la structuration des cooccurents.

Si toutes les conditions sont réunies, on consacra les matinées à une présentation schématique, et les après-midis à des exercices. Cela suppose que les participants soient munis d'ordinateurs (dont les utilisateurs aient les droits d'administration) et de fichiers de textes latins numérisés (les fichiers en format .txt des CBMA peuvent convenir - <http://www.artehis-cnrs.fr/page-documentaire-CBMA#EDITIONS>).

Le nombre de participants est limité à 12 personnes.

Les intéressés sont priés de s'inscrire auprès de Chantal Palluet (cnrs.auxerre@orange.fr) avant le 31mars 2013.

La prise en charge (transport et hébergement) d'un certain nombre de participants est prévue. Les candidats à cette prise en charge doivent envoyer une demande motivée, accompagnée d'un CV, avant le 15 mars 2013 à Chantal Palluet (cnrs.auxerre@orange.fr).

URL : <http://www.glossaria.eu/>

