



Lexicologie, terminologie, traduction

Journées de formation

13 et 14 septembre 2011

Du caractère au sens : méthodes et outils pour le TAL

Le déroulement d'une application de traitement automatique des langues (TAL) des textes écrits peut être envisagé comme un processus qui se découpe en plusieurs étapes successives. Il s'agit de transformer une suite indistincte de caractères – lettres, espaces, ponctuations – en un ensemble cohérent d'éléments de nature linguistique – mots, mots composés, syntagmes, phrases... – où chacun de ces éléments se voit attachés un ou plusieurs traits – nature, sens, domaine... Ce travail une fois réalisé, il est alors possible d'effectuer différents traitements : traduction, extraction d'information... À chaque étape, l'application d'une ressource linguistique, quelle qu'en soit la nature, est nécessaire.

L'objectif de ces journées de formation organisées par le laboratoire LDI (Lexiques, Dictionnaires, Informatique) est de proposer un ensemble de modules abordant différents aspects liés à ce continuum. Chacun des cours est construit de façon à être le plus autonome possible. Ainsi, chaque participant peut choisir de construire son parcours « à la carte ». Afin de faciliter les choix des participants, chaque cours fera l'objet d'une description précise : les objectifs, les pré-requis nécessaires et les outils mis en œuvre.

Seront abordés les thèmes suivants :

- construction et normalisation de corpus ;
- XML, TEI, bases de données XML ;
- corpus parallèles ;
- constitution de ressources morphologiques.

Organisation

Un questionnaire préalable aux journées sera proposé aux participants. Ce questionnaire permettra aux participants de choisir les différents cours.

Une plateforme web sera mise en place :

- lieu d'échange,
- ressources logicielles,
- ressources pédagogiques.

Il sera remis à chaque participant un polycopié contenant les supports (diapositives, tutoriels...) pendant les deux jours. Il pourra être distribué au début de la formation.

Les participants recevront en outre une clef USB contenant l'ensemble des outils utilisés afin de leur permettre de mettre en pratique, sur leurs propres sujets d'étude, les techniques et outils abordés.