

Séminaire MIDI : Pierre-Edouard Portier

15 Avril 2011, 14:00 – 15:30

Titre : "Documents multistructurés"

Intervenant : Pierre-Edouard Portier, LIRIS, Lyon.

Lieu : UCP, St-Martin 1, locaux ETIS.

Résumé :

J'étudie le processus de création de documents. Je pose que ces derniers sont les produits d'interactions qui se jouent au sein d'un système qui comprend plateformes informatiques et opérateurs humains. Un contexte applicatif particulier permet de valider mes hypothèses. Il s'agit de l'analyse critique et de la mise en valeur, par une équipe de chercheurs de l'ENS Lyon, du fonds d'archives des manuscrits du philosophe Jean-Toussaint Desanti. J'associe à tout document une structure générique de graphe : c'est un ensemble de composants documentaires interconnectés : des pages manuscrites numérisées, du texte transcrit et annoté, des collections de pages ordonnées, des zones d'images de manuscrits isolées, des relations nommées entre éléments quelconques, ... Une opération courante consiste à découper ce graphe en sous graphes qui sont d'ailleurs souvent des arbres. C'est au sens de cette opération que nous pouvons dire d'un document qu'il est multi-structuré. Dans ce contexte, je me donne pour problématique l'étude des processus de création de ces différentes structures. En même temps, je développe un système générique pour l'annotation sémantique de documents dans un environnement multi-utilisateurs.

Ainsi, trois contributions principales se distinguent. D'abord, une méthodologie pour la construction des documents multistructurés qui profite des événements structurels de chevauchement de termes d'annotation pour guider les utilisateurs dans la construction de vocabulaires de termes de plus en plus fins et qui formalisent de mieux en mieux leurs connaissances structurelles sur l'archive documentaire. Ensuite, un modèle de la trace des actions d'interaction qui permet une documentation dynamique des vocabulaires de termes d'annotation. Enfin, une plateforme philologique qui comprend, entre autres éléments, une implantation Web de ces deux dernières contributions.

Finalement, je présenterai un travail récent basé sur l'idée de tester la contrainte d'acyclisme sur le graphe orienté des relations qui relie des fragments documentaires quelconques afin d'assister l'utilisateur dans la construction d'ensembles cohérents de relations. J'appelle ces ensembles dimensions. En effet, la conception des dimensions comme ensembles de relations est à la base de travaux existants sur les hypertextes dimensionnels que mon approche vient compléter.

