

Modélisation des connaissances dans le cadre de bibliothèques numériques spécialisées

Pierre-Edouard Portier*, Sylvie Calabretto *

*Université De Lyon, INSA de Lyon, LIRIS
pierre-edouard.portier,sylvie.calabretto@insa-lyon.fr

Résumé. Nous présentons une application innovante de la modélisation des connaissances au domaine des bibliothèques numériques spécialisées. Nous utilisons la spécification experte de la TEI (Text Encoding Initiative) pour modéliser la connaissance apportée par les chercheurs qui travaillent sur des archives manuscrites. Nous montrons les limites de la TEI dans le cas d'une approche diachronique du document, cette dernière impliquant la construction simultanée de structures de données concurrentes. Nous décrivons un modèle qui présente le problème et permet d'envisager des solutions. Enfin, nous justifions les structures arborescentes sur lesquelles se base ce modèle.

1 Introduction

Ce travail est une synthèse de l'expérience acquise à côtoyer des chercheurs qui construisent une édition électronique des archives manuscrites du philosophe Jean-Toussaint Desanti. L'édition électronique recouvre l'ensemble du processus non seulement éditorial mais scientifique et critique qui aboutira à la publication d'une ressource sous forme électronique. Pour les manuscrits, la première ressource publiée est un facsimile numérique auquel s'ajoutera le travail de transcription et d'analyse critique des chercheurs. Nous avons échangé avec les responsables d'autres projets similaires pour isoler la problématique centrale et partagée de la construction de documents multistructurés. Nous montrons dans la seconde section que cette problématique apparaît naturellement avec l'utilisation de la TEI pour transcrire les manuscrits. Dans la troisième section nous proposons un modèle qui répond à cette problématique. Dans la dernière section nous formulons pour la première fois le problème de la construction de documents multistructurés.

2 La TEI et les documents multistructurés

2.1 Organisation de la TEI

La TEI, Text Encoding Initiative (Burnard et Bauman (2007)), est un consortium qui développe et maintient un standard pour la représentation des textes électroniques. Ses recommandations constituent une expertise dont peut profiter tout projet d'édition électronique. Elles sont exprimées sous la forme modulaire et extensible d'un schéma XML documenté.