

# Validation et enrichissement d'annotations : application à la veille médiatique

Olivier Carloni<sup>\*,\*\*</sup>, Michel Leclère<sup>\*</sup>, Marie-laure Mugnier<sup>\*</sup>

<sup>\*</sup>LIRMM 161 rue Ada, 34392 Montpellier  
{carloni, leclere, mugnier}@lirmm.fr  
<http://www.lirmm.fr>

<sup>\*\*</sup>Société Mondeca, 3 Cité Nollez 75018 Paris  
olivier.carloni@mondeca.com,  
<http://www.mondeca.com>

**Résumé.** Cet article présente un service de validation et d'enrichissement d'annotations conçu pour un outil industriel de gestion des connaissances basé sur le langage des Topic Maps (TM). Un tel service nécessitant la mise en œuvre de raisonnements sur les connaissances, nous avons été amené à doter le langage des TM d'une sémantique formelle. Ceci a été réalisé par l'intermédiaire d'une transformation des TM vers le formalisme logique des graphes conceptuels qui comme les TM dispose d'une représentation graphique des connaissances. La solution a été mise en œuvre dans une application conçue pour la veille médiatique. Des annotations sont extraites automatiquement de dépêches sur l'actualité économique puis ajoutées à la base de connaissances. Elles sont ensuite fournies au service de validation qui décide de leur validité, les enrichit et évalue leur pertinence afin de faciliter le travail du veilleur.

## 1 Introduction

Depuis une dizaine d'années, la connaissance a fait une entrée massive dans les organisations industrielles et administratives, répondant aux besoins (1) de capitalisation du patrimoine immatériel des organisations que constituent les savoirs et savoirs-faire de leurs membres et (2) d'exploitation à un niveau "sémantique" des systèmes d'information des organisations, c'est-à-dire s'appuyant sur la sémantique des informations échangées. Mondeca est un éditeur de logiciels qui développe ITM (Intelligent Topic Manager) un logiciel de gestion des connaissances structurant les connaissances représentées en 3 niveaux :

1. Le niveau méta contient l'ontologie de représentation. Il comporte toutes les primitives du langage d'ITM permettant de créer les ontologies de chaque "client" de l'outil ITM.
2. Le niveau modèle contient les ontologies créées pour chaque client, c'est-à-dire les types et des propriétés générales qui seront utilisées pour décrire les instances. Plusieurs sortes d'ontologies sont généralement définies : une ontologie de domaine spécifique au client, des ontologies "de services" (primitives de spécification de thesaurus, primitives de des-