

Tutoriel

Terminologie, ontologies ; vers le Web sémantique

Jean Charlet⁽¹⁾, Nathalie Aussenac-Gilles⁽²⁾ et Philippe Laublet⁽³⁾

(1) STIM – AP-HP – jc@biomath.jussieu.fr

(2) IRIT – UMR CNRS – aussenac@irit.fr

(3) LaLICC - UMR CNRS - Université de Paris-Sorbonne – Philippe.Laublet@paris4.sorbonne.fr

Le cours « terminologie, ontologies ; vers le Web sémantique » s’articulera en 3 parties autour de la notion d’ontologie. Nous aborderons d’abord l’apparition de cette notion dans le domaine de l’intelligence artificielle, ses fondements et les enjeux liés à la constitution et à la réutilisation de ces modèles. Nous préciserons le rapport des ontologies à la terminologie, en étudiant le positionnement des ontologies par rapport à différents produits terminologiques et leur construction à partir de corpus. Enfin, dans une dernière partie, nous aborderons la problématique du Web sémantique et l’apport des ontologies dans ce contexte.

Ontologies, enjeux et principes de constitution

La notion d’ontologie est devenue un élément clé dans toute une gamme d’applications faisant appel à des connaissances. Une ontologie est définie comme la conceptualisation des objets reconnus comme existant dans un domaine, de leurs propriétés et des relations les reliant. Leur structure permet de représenter les connaissances d’un domaine sous un format informatique en vue de les rendre utilisables pour différentes applications. Cette réutilisation, en pratique difficile, est limitée mais possible à condition que les principes de constitution de l’ontologie soient fondés et parfaitement clairs. Des exemples seront proposés, entre autres, dans le domaine médical, en particulier en relation avec les thésaurus médicaux.

Constitution d’ontologies à partir de textes

La constitution d’ontologies à partir de corpus textuels présente de réels intérêts, surtout lorsque ces ontologies sont utilisées au sein d’applications de gestion documentaire ou de recherche d’information :

- rendre compte au plus juste de la connaissance telle qu’elle est explicitée dans des documents consensuels et partagés,
- utiliser des termes réellement utilisés par une communauté d’utilisateurs dans un domaine d’activité,
- maintenir le lien entre l’ontologie et les éléments textuels qui justifient la modélisation choisie.

Cette partie du cours comportera un volet linguistique et traitement automatique des langues (comment repérer la connaissance dans les corpus) et un volet méthodologique (comment modéliser et formaliser cette connaissance, avec quels outils travailler), en prenant appui sur plusieurs expériences de constitution d’ontologies. Une réflexion soulignera en quoi ce type d’approche renouvelle le regard de l’IA sur la modélisation et la formalisation.

Méta-données, Ontologies et Web sémantique

Sous l'expression Web sémantique, se regroupe un ensemble de programmes de recherche et de travaux très diversifiés. Leur objectif commun est de permettre aux machines de mieux exploiter automatiquement les contenus des ressources Web afin de décharger leurs utilisateurs d'une partie de leurs tâches et de leur proposer une aide afin d'atteindre un niveau de services qualitativement nouveau. Ces ressources vont être ainsi munies de descriptions sémantiques, simples méta-données documentaires, schémas d'annotations plus riches jusqu'à des cartes complexes. Dans ce contexte du Web sémantique, les ontologies doivent trouver une place centrale puisqu'elles vont fournir le vocabulaire et les structures sémantiques, formelles ou non. Dans cette dernière partie du tutoriel, nous montrerons différents rôles que les ontologies et les méta-données / annotations peuvent prendre en les illustrant par des exemples. Nous situerons aussi les différents langages utilisables : langages de représentation comme RDF et les différents niveaux d'OWL, mécanismes inférentiels, langages de requêtes et différents outils associés. Nous soulignerons ainsi les opportunités, pour l'IA, offertes par le développement du Web sémantique.