



Conférence TOTh 2011

Terminologie & Ontologie : Théories et applications

www.porphyre.org/toth

Annecy, 26 - 27 mai 2011

PROGRAMME

Mardi 24 après-midi et mercredi 25 mai 2011 : Formation TOTh

Construction d'ontologies à des fins terminologiques

La formation vise à promouvoir la collaboration entre les professionnels de la langue et de la connaissance. Elle a également pour objectif de présenter les théories, méthodes et outils de construction d'ontologies à des fins terminologiques, dotant ainsi :

- les professionnels de la langue d'une approche méthodologique et systématique (raisonnement formel) pour la définition du système conceptuel, particulièrement utile pour l'opérationnalisation des terminologies ;
- les professionnels de la connaissance de compétences linguistiques indispensables pour l'acquisition de connaissances à partir de ressources linguistiques.

Jeudi 26 mai 2011 : Conférence TOTh

8h30 : Accueil

9h00 : Ouverture. *Christophe Roche*

9h15 : Conférence Invitée. *Michele Prandi*

Signes, signifiés, concepts : pour un tournant philosophique en linguistique

Pr. Michele Prandi, Università degli Studi di Genova, Italy

Au moment où la terminologie, après son tournant linguistique, prend conscience que l'analyse linguistique des textes de spécialité, loin de donner un accès direct à un univers de concepts, présuppose leur maîtrise préalable, Michele Prandi propose une réflexion parallèle qu'il poursuit depuis plusieurs années dans le domaine de la linguistique.

Tant la description lexicale d'une langue que l'étude du signifié des expressions complexes, et notamment des phrases, présupposent un accès indépendant à un système de concepts et de relations conceptuelles partagées. La grammaire des formes, pour ainsi dire, se double d'une grammaire des concepts, dont le noyau le plus stable est formé par ce que Michele Prandi a appelé une « ontologie naturelle ».

Biographie

Michele Prandi est professeur de Linguistique Générale à l'Université de Gênes, Faculté de Langues et Littératures Etrangères, et Docteur honoris causa de l'Université d'Uppsala.

Sa recherche porte sur la grammaire et la sémantique des expressions complexes – phrases, relations transphrastiques, textes – dont le contenu est vu comme le résultat d'une interaction complexe entre codage linguistique et inférence, grammaire des formes et grammaire des concepts, et sur l'analyse linguistique et conceptuelle du discours figuré.

10h15 : Pause

10h30 : Concepts as building blocks for knowledge organization – a more ontological and less linguistic perception of terminology

Klaus-Dirk Schmitz, Cologne University of Applied Sciences, Germany

In terminology science, concepts are defined as units of knowledge (ISO 1087-1:2000) or units of thinking (DIN 2342:2011). Concepts are mental representations which can be described by definitions. Ideally, definitions refer to other concepts and mention specific characteristics of the concept to be defined. On the basis of definitions and concept characteristics, it is possible to relate concepts to each other and to construct concepts systems representing the knowledge of a (sub-)domain in a systematic way. Therefore, concepts are the building blocks for any kind of (simple) knowledge organization system (SKOS), be it a thesaurus, an ontology, a classification or a terminological concept system.

For professional communication as well as for knowledge organization established terms are used to represent and refer to concepts. Of course, terms appear in technical texts not independently; they are embedded in a (socio-)linguistic context that has to be taken into account for any kind of lexicographical or – with a broader view – linguistic activity. And also terminology work has to care about the linguistic and socio-linguistic features of terms. But for any kind of terminology management, it is extremely important not to lose track of the concept part of terminology. Not only terminological concept systems but also all types of term bases and terminological data collections – seen as knowledge organization systems – have to follow a concept-oriented data modelling and working procedure approach. This approach is the necessary basis to guarantee that the (linguistic) knowledge of a subject field or a company (corporate language) is not accidentally scattered by linguistic features of the terms but organized and managed by the knowledge part of the concepts and the relations between concepts.

11h15 : Linking Specialized Knowledge and General Knowledge in EcoLexicon

Pamela Faber, Antonio San Martín, University of Granada, Spain

Ontologies have been criticized because they demand too much work or because they are not sufficiently flexible to capture the dynamism and complexity of reality (Kingson 2008). However, even though any representation of reality is imperfect, ontologies are the type of computational knowledge representation that best approximates the domain being conceptualized. In fact, they have increasingly come into focus because of the need for knowledge management and shared knowledge in both general and specialized knowledge domains. EcoLexicon is a frame-based visual thesaurus on the environment, whose knowledge is stored in a relational database, and which is gradually evolving towards the status of a formal ontology (Leon, Magaña, and Faber 2008; Leon and Magaña 2010). This paper describes the conceptual modeling techniques used in this knowledge resource, and the underlying theoretical premises that enable its contextualization and connection to general knowledge structures and resources.

12h00 : Repas

14h00 : Session 2. Président de session : Danielle Candel

14h00 : Terminus: a Workstation for terminology and corpus management

María Teresa Cabré, Rogelio Nazar, Institut Universitari de Lingüística Aplicada, Universitat Pompeu, Fabra, Spain

Terminus is a software developed at our Institute used by terminologists to go through the whole process of glossary creation, including tools for corpus compilation, corpus exploration and term management. In this paper, we present a description of its functions and examples of its application. Also, we present the algorithms of two new functions which are currently being added to a Beta version of the program. The first is the filtering of a set of documents according to level of specialization and relevance with respect to the analyzed thematic domain. The second algorithm is for automatic term extraction. Needless is to say these two represent fields of research that have received a lot of attention from the community of terminologists and corpus linguists in the last decades. We provide a description of our strategies and the results we have obtained with these experimental modules

14h45 : Le métier : son savoir, son parler

Caroline Djambian, Université Aix-Marseille 3, France

Le savoir de métier est ici au centre de nos réflexions. Connaissance expérimentielle collective formée et reformulée par le biais de concepts propres aux métiers, sa transmission est aujourd'hui en mutation dans les entreprises, posant des problématiques complexes comme dans le cas de la Division Ingénierie Nucléaire d'EDF. La réponse est dans la mise en forme de ce savoir dé-sormais médiaté : le langage. Nous en arrivons donc naturellement à la terminologie qui, en tant que représentation linguistique de la science d'un champ social, en fixe le langage. Les ontologies qui modélisent et représentent formellement le système notionnel des terminologies, en sont la suite logique. Nous illustrons notre réflexion en présentant une application concrète sur un domaine de la Division Ingénierie Nucléaire d'EDF, par la construction d'une base de connaissances constituée d'une terminologie et d'une ontologie du domaine centrée sur le sens métier.

15h30 : Acquisition automatique de termes et lexique scientifique transdisciplinaire

Patrick Drouin, Gabriel Bernier-Colborne, OLSST, Université de Montréal, Canada

La nature lexicale des textes scientifique fait l'objet de plus en plus de travaux dans le cadre de l'enseignement des langues secondes, mais elle est toujours peu explorée dans le cadre de la terminologie, plus particulièrement en terminologie computationnelle. Nous proposons, dans cet article, une première analyse de l'imbrication du lexique scientifique transdisciplinaire (LST) dans la terminologie d'un corpus de nature scientifique. Notre analyse a pour but de vérifier dans quelle mesure cette présence du LST peut être utilisée afin guider un logiciel d'acquisition automatique de termes dans son évaluation du potentiel terminologique.

16h15 : Pause

16h30 : Session 3. Président de session : Dardo de Vecchi

16h30 : Donner un nom propre au Faucon : portée taxinomique et philologique du terme Nom propre au XVIe siècle

Philippe Selosse, Université Lyon 2, France

Cet article explicite le sens du syntagme nom propre en français dans le discours scientifique de la Renaissance. Il décrit le fonctionnement sémantique de l'adjectif propre et son insertion dans l'ontologie aristotélicienne. Il établit l'emploi terminologique de nom propre (NP) : un NP est une appellation de rang spécifique (et non individuel comme en français moderne), porteuse de la propriété caractéristique (le « propre » aristotélicien) et opposée au nom commun, partie de discours porteuse des caractères accidentels. Le concept terminologique de NP est cependant conçu au sein des langues vulgaires. Dans celles de l'Antiquité, le NP est caractérisé par un principe de primarité et d'autorité. Dans celles du XVIe s., c'est le correspondant antique, établi selon un principe philologique de comparaison et d'« appropriation », qui valide le NP vulgaire. Le NP s'inscrit enfin dans une théorie de l'emprunt, partagée dans l'épistémè de la Renaissance. La traduction latine constante par « proprium nomen » conforte ces explications.

17h15 : Disputatio

17h15 : « Isagogè » de Porphyre

Christophe Roche, Equipe Condillac, Université de Savoie, France

Avec la Disputatio TOTh, les conférences TOTh renouent avec une forme d'enseignement et de recherche fondée sur la raison dialectique. L'objectif est ici de donner accès à un texte fondateur et à son enseignement par une étude commentée menée sous l'angle de la terminologie et de l'ontologie.

Le premier texte étudié sera l'Isagogè de Porphyre. Choisie en référence à la scolastique dont relève la disputatio, cette "introduction aux catégories d'Aristote" (plus justement aux prédicables) joue un rôle fondamental en terminologie et en ontologie. La définition en genre et différence s'y rattache tout comme l'organisation des connaissances sous la forme d'un arbre de Porphyre.

18h00 : Fin de la première journée

20h30 : Dîner TOTh

Bistrot de Bonlieu, Annecy

9h00 : Session 4. Président de session : Pierre Lerat

9h00 : Relier les niveaux terminologique et conceptuel dans le domaine juridique : hypothèse sur la méthodologie middle-out

Danièle Bourcier, Meritxell Fernandez, CERSA, CNRS Université Paris 2, France

Dans cet article une méthodologie de construction d'ontologies du droit qui tient compte des procédés du raisonnement juridique est proposée. La méthodologie middle-out est adaptée au domaine juridique par voie de la notion de qualification juridique comme niveau intermédiaire de conceptualisation. On propose une notion performative d'ontologie juridique vis-à-vis l'ontologie platonique descriptive et on souligne la relevance des catégories sémantiques fonctionnelles pour une ontologie du droit. On montre un exemple d'application de cette méthodologie dans le domaine de la régulation du bruit.

9h45 : Description de verbes juridiques au moyen de la sémantique des cadres

Janine Pimentel, OLST / CLUNL, Université de Montréal, Canada, Universidade Nova de Lisboa, Portugal

Dans un scénario tel que celui du jugement d'une cour suprême, le juge et l'appelant sont les acteurs principaux jouant des rôles distincts dans la poursuite d'un même objectif : la justice. Leur discours, pouvant être analysé au moyen des décisions écrites par les juges, invoque les motivations de chaque acteur ainsi que les actions entreprises tout au long du processus (et même du procès). Par conséquent, la terminologie des jugements des cours suprêmes est constituée non seulement par des noms mais également par des verbes dont il faut rendre compte. Dans cet article, nous proposons une méthodologie de description de verbes juridiques s'inspirant du cadre théorique de la sémantique des cadres (Fillmore 1977, 1982, 1985) et de son application FrameNet (Ruppenhofer et al. 2010) en vue de l'élaboration d'une ressource lexicale bilingue (portugais-anglais). Nous décrivons le travail d'implémentation et d'adaptation du cadre théorique aux objectifs du projet.

10h30 : Pause

10h45 : Session 5. Président de session : Loïc Depecker

10h45 : Cross language legal information retrieval: the semantic interoperability among thesauri as possible solution

Enrico Francesconi, Ginevra Peruginelli, ITTIG-CNR, Italian National Research Council, Italy

In the last few years crucial issues like cross-language legal information retrieval, document classification, legal knowledge discovery and extraction have been considered in theory and in practice. The availability of services allowing cross-language and cross-collection retrieval is a growing necessity. This paper focuses on the need to develop solutions for automatic, language-independent procedures to provide interoperability between mono/poly-lingual thesauri at national and European levels. This will guarantee sustainable and scalable services enabling to manage the multilingual complexity of the European Union legal context to be used for cross-language and cross-collection legal information retrieval. Wider use of the service can also be envisaged as support to legal translation services, as well as in general to promote integration and sharing of widespread and heterogeneous legal resources, providing new market opportunities for stakeholders to exploit the economic potential of public sector information in a multilanguage environment.

11h30 : Terminological Contributions in Ontology Building: The Informal Specification stage.

Claudia Santos, SCA - Aveiro University, Portugal

The need to manage, store, share and reuse knowledge has led to the creation of countless tools aiming to capture consensual domain conceptualizations which in turn would allow for a transformation of data into logical propositions understood by humans and systems alike. Terminology as a science and set of procedures intervenes in a decisive way in the informal specification stage of conceptualizations mainly through two methodologies: semasiology and onomasiology. This article presents results on the adoption of linguistic and extra-linguistic terminological approaches in the informal specification stage of ontology construction and, simultaneously, puts forward a mixed methodology proposition.

12h15 : Prix Jeune Chercheur

12h30 : Repas

14h30 : Textile terminologies - Projects of Centre for Textile Research, Denmark

Susanne Lervad, Pascaline Dury, Marie Louise Nosch, Termplus and Centre for Textile Research, Denmark

The Centre for Textile Research in Denmark launched its research programme with the aim to investigate Bronze Age textile production from an archaeological, experimental and linguistic point of view. It encompasses two parts: Tools and Textiles, and Text and Contexts. In the first part, the mission was to gather information on textile tools from various types of Bronze Age contexts and sites in the Eastern Mediterranean area. This provided a basic methodology for textile tool studies. Via tool studies, context studies, and experimental testing, the programme provided a clear picture of types and qualities of textiles, which derive from the tools. The second part, text and Contexts focuses on the written records of Eastern Mediterranean area in the 3rd and 2nd millennia BC, where we have references to a complex terminology of textiles, tools and techniques, decoration and specialised textile occupational titles. However, we often lack their precise meaning. The aim is thus to investigate textile terminology diachronically, and in a comparative approach. The Volume Textile Terminologies with different contribution is published 2010 at Oxbow Books

15h15 : Extraction de relations et de patrons de relations entre entités nommées en vue de l'enrichissement d'une ontologie

Yayoi Nakamura-Delloye, Rosa Stern, INRIA, France

Nous proposons dans cet article une méthode non-supervisée d'extraction des relations et des patrons de relations entre entités nommées, réalisée dans le cadre de la création et l'enrichissement d'une ontologie. La méthode proposée se caractérise par l'exploitation des résultats d'analyse syntaxique, notamment les chemins syntaxiques reliant deux entités nommées dans les arbres de dépendance. Les informations sur les relations syntaxiques présentes entre les composants sont mises à profit pour le calcul de la similarité employée pour la phase principale de classification. Nous présentons également le mécanisme conçu pour l'intégration des résultats obtenus dans une ontologie.

16h00 : Pot de départ