



Conférence TOTh 2010

Terminologie & Ontologie : Théories et applications

Annecy, 2-4 juin 2010

PROGRAMME DETAILLE

Mercredi 2 juin 2010

FORMATION TOTh

« Terminologie & Ontologie »

Jeudi 3 juin 2010

8h30 : Accueil

9h00 : Ouverture

9h30 : Conférence Invitée

Ontologie : Analyse des concepts ou métaphysique

Frédéric Nef

Ecole des hautes études en sciences sociales

Deux conceptions de l'ontologie s'affrontent : selon la première elle est réductible à l'analyse des concepts primitifs ; selon la seconde elle est le cœur de la métaphysique, une science transcendante de la réalité ultime. Je montrerai qu'une variété d'ontologie, fondée sur l'admission de structures ontologiques permet de comprendre cette opposition. Je montrerai que le réalisme structural est une voie qui permet de penser la nature de l'ontologie. Cependant il faut un réalisme fort des structures que je discuterai en critiquant le constructionnisme courant des structures ontologiques, réduites à des représentations logiques et mathématiques. L'ontologie formelle repose sur le développement d'une série d'outils formels de représentation, comme les calculs logiques, méréologiques etc., mais cela ne signifie pas qu'elle se confond avec la compilation de ces langages: elle a pour ambition de proposer un cadre unifié de connaissance des structures ontologiques de la réalité qui ne se réduisent pas à des structures sémantiques ou logiques. En ce sens, c'est une ontologie réaliste qui sera défendue.

10h30 : Pause

11h00 :

Le travail sur la représentation (visuelle) des connaissances en terminologie : un retour d'expérience
Dardo de Vecchi

Euromed-Management, Université Paris Diderot – EILA

La représentation graphique comme à la fois modalité et outil de communication et partage de connaissances est fréquente. Cependant les applications de l'outil liées à l'enseignement de la terminologie ou de la gestion de connaissance sont rares. L'article présente un retour d'expérience d'intégration de l'outil de travail sur le support visuel dans ces champs et en propose une adaptation au travail avec des malvoyants et aveugles. Dans le cas des entreprises, l'outil visuel invite à une réflexion sur l'organisation elle-même, son identité, et ses besoins d'information.

11h45 :

Une « ontoterminologie » pour les interprètes de conférence

Elisa Veronesi, Franco Bertaccini

Université de Bologne

La terminologie et la création d'ontologies s'avèrent de plus en plus utiles pour tous ceux qui travaillent dans les domaines de la traduction et de l'interprétation. Ce projet a donc pour but d'étudier l'usage d'arbres ontologiques en tant que support à l'interprétation simultanée. Nous avons essayé de mettre la terminologie au service des interprètes en concevant une ontologie dynamique et modulaire pouvant être utilisée par les interprètes directement en cabine. Pour ce faire, notre projet s'est basé sur la création d'une conceptualisation formelle d'un domaine spécialisé (une ontologie), développée à partir de ses termes et concepts fondamentaux. La méthodologie de sa structuration repose donc sur la façon dont les termes (et les concepts auxquels ils font référence) sont reliés entre eux d'un point de vue sémantique et logique. Pour cette raison, ce support pourrait être défini une « ontoterminologie ». Le domaine conceptualisé est celui du SIDA. Plus en particulier, notre analyse a porté sur l'étude des manifestations cutanéomuqueuses se présentant au cours de l'infection par le VIH. Nous avons créé deux corpora de langue spécialisée, un en français et un en italien, desquels nous avons extrait les concepts et les termes fondamentaux du domaine et du sous-domaine, ainsi que les relations sémantiques les reliant. Ensuite, nous avons développé une ontoterminologie bilingue italien-français très spécialisée et strictement liée au domaine. Le champ de la biomédecine étant très complexe et articulé, les instances, les relations, ainsi que la structure des diagrammes ont été définies manuellement, en utilisant le logiciel graphique Visio 2007. On a essayé, d'une part, de produire une ontologie flexible, synthétique mais exhaustive ; de l'autre, de faire en sorte qu'elle permette de prévoir (ou de suivre) l'évolution des discours de la conférence, en affichant non seulement les termes et les équivalents qui relèvent du sujet, mais aussi les relations sémantiques et logiques qui les relient. Enfin, nous avons mené un test empirique pour vérifier la « lisibilité » et l'utilité effective du support avec des interprètes professionnels. L'analyse de leurs performances et opinions nous a permis d'évaluer les avantages concrets apportés par l'outil réalisé.

12h30 : Repas

14h30 :

Une « ontoterminologie » pour les interprètes de conférence – Un outil développé au sein de l'environnement académique (Démonstration)

Elisa Veronesi, Franco Bertaccini

Université de Bologne

Dans un monde de plus en plus mondialisé et multiculturel, l'interprétation et la traduction ont largement recours à la terminologie et à la création d'ontologies de domaine pour pouvoir développer de bases de données dynamiques, modulaires et interactives. Ce projet a donc pour but d'étudier l'usage d'arbres ontologiques complexes en tant que support à l'interprétation simultanée. Nous avons essayé de mettre la terminologie au service des interprètes en concevant une ontologie dynamique et modulaire pouvant être utilisée par les interprètes en cabine, au cours du travail. Pour ce faire, nous avons créé une conceptualisation formelle d'un domaine spécialisé développée à partir de ses termes et concepts fondamentaux. La méthodologie de sa structuration repose par conséquent sur la façon dont les termes (et les concepts auxquels ils font référence) sont reliés entre eux d'un point de vue sémantique et logique. Pour cette raison, ce support pourrait être défini comme une « ontoterminologie ». Cet outil linguistique et onto-terminologique a été développé au sein du milieu académique avec des supports informatiques qui n'étaient pas complètement adéquats et peu de ressources financières. Il nécessiterait donc d'être davantage formalisé et mis au point pour l'adapter à différents domaines d'application et être utilisé par des interprètes professionnels. Il représente toutefois la modélisation d'un hypothétique système à base de connaissances pour l'interprétation simultanée, qui nous permet d'en postuler le développement et la réalisation par des moyens appropriés. La raison en est que nous avons pu démontrer, par le truchement d'un essai empirique, que ce type d'outil serait tout à fait utile et d'appoint aux interprètes non seulement au cours de la phase de préparation au travail, mais aussi lors de la traduction simultanée en cabine.

15h15 :

ITM, une infrastructure sémantique pour la maintenance du thésaurus multilingue Eurovoc (Démonstration)

Thomas Francart, Charles Teissèdre

Mondeca

Intelligent Topic Manager (ITM) est un outil propriétaire dédié à la gestion de connaissances. Dans le cas d'étude retenu pour la démonstration, l'outil sert de support logiciel pour maintenir et faire évoluer Eurovoc1, un thésaurus multilingue (disponible en 25 langues), couvrant les différents domaines d'activité de l'Union Européenne (politique, géographie, agriculture, etc.). Ce thésaurus est utilisé par les services documentaires des administrations européennes et nationales et permet d'indexer des documents en les arrimant à un vocabulaire destiné à servir de norme. Sa constitution et sa maintenance soulèvent des difficultés, qui tiennent autant à son évolution constante, qu'à l'édition d'un vocabulaire commun à différentes langues, ce qui renvoie au problème de la traduction, dont ici le paradigme retenu (qui ne va pas de soi) est que les concepts sont partagés par tous et que les termes qui les désignent sont spécifiques à chaque langue. Les fonctionnalités d'ITM utilisées dans ce contexte peuvent être réparties en quatre grandes catégories, chacune répondant à des besoins spécifiques pour la maintenance du thésaurus dans un contexte collaboratif et partagé.

16h00 : Pause

16h30 :

Semiotic Triangle Revisited for the Purposes of Ontology-based Terminology Management

Igor Kudashev, Irina Kudasheva

University of Helsinki

In this paper, we examine the limitations of the traditional semiotic triangle from the point of view of terminology management and suggest an alternative model based on the concept of terminological lexeme. The new model is being tested in the TermFactory project aimed at creating a platform and a workflow for distributed collaborative ontology-based terminology work.

17h15 :

INDEX d'Onomia : L'ontoterminologie pour la recherche d'information sémantique

Luc Damas, Christophe Tricot

Equipe Condillac, Université de Savoie

Onomia

Cet article propose une réflexion sur la recherche d'information sémantique dans les domaines scientifiques et techniques. Comment permettre aux collaborateurs d'obtenir rapidement de l'information pertinente ? Qu'est ce qu'une recherche pertinente et, par là même, un document pertinent ? Autant de questions qui restent ouvertes.

Nous commencerons par préciser les problématiques et ce que nous entendons par recherche pertinente. En effet, celle-ci ne peut plus se réduire à une simple recherche par mots-clés.

Nous abordons ensuite les différents aspects de la recherche d'information sémantique basée sur l'ontoterminologie, terminologie qui met en relation, tout en les distinguant, les systèmes conceptuel et linguistique qui composent toute terminologie. Nous montrerons comment exploiter ces deux systèmes sémiotiques et en particulier comment la qualité de la structure conceptuelle sur laquelle sont classées les informations (par exemple le modèle OK) influence l'analyse des documents et la manipulation des requêtes. Enfin nous abordons d'un point de vue formel les aspects de similarité et de pertinence et nous exemplifierons nos propos grâce à l'outil Index d'Onomia.

18h00 : Fin de la première journée

20h30 : Diner TOTh

« Bistrot de Bonlieu », Annecy

Vendredi 4 juin 2010

9h00 : Session 4 (Président : Franco Bertaccini)

9h00 :

Approche onomasiologique de la phraséologie transdisciplinaire des écrits scientifiques : la recherche sémantique dans les textes dans le cadre du projet Scientext (Démonstration)

Falaise Achille, Tutin Agnès

Equipe GETALP, LIG

LIDILEM, Université Grenoble 3

L'accès à la phraséologie, en particulier pour les applications en langue étrangère et seconde, se fait rarement à partir de corpus. Dans le cadre du projet ANR Scientext, nous avons élaboré un mode d'accès à la phraséologie transdisciplinaire des écrits scientifiques par un mode onomasiologique. Des requêtes prédéfinies portant sur la question linguistique du positionnement et du raisonnement ont été élaborées à partir de schémas syntaxiques et sémantiques, par exemple sur l'expression de l'opinion ou de l'évaluation. Ces grammaires sont ensuite appliquées à un large corpus d'écrits scientifiques balisé au plan structurel et au plan syntaxique (analyse de dépendance). L'utilisateur peut ainsi extraire, selon ses besoins, une phraséologie adaptée à une requête sémantique.

9h45 :

Modélisation des dénominations ontologiques

Benjamin Diemert, Marie-Hélène Abel, Claude Moulin

Heudiasyc, UMR 6599, Université de Technologie de Compiègne

L'échange de l'information numérique et sa réutilisation dans de nouveaux contextes pose un problème d'écarts des représentations entre le producteur et le récepteur. La production télévisuelle se tourne de plus en plus vers des modes de production collaboratifs incluant des amateurs ne maîtrisant pas le jargon audiovisuel. Dans ce papier, nous introduisons un modèle conceptuel capable d'articuler différentes dénominations autour d'un même élément ontologique (concept, instance, propriété). Nous définissons notre motivation à partir d'exemples d'échange d'information entre professionnel et amateur puis montrons comment le même problème s'applique également à la recherche multilingue. Nous rappelons l'apport des normes OAIS, FRBRoo et du modèle AXIS à la conceptualisation de l'échange d'information et les manques concernant la modélisation des écarts d'usages de la langue entre différentes communautés. Alors que ces modèles se concentrent sur l'échange de données, nous proposons de décrire et relier l'information, l'utilisateur et son contexte d'action. A partir d'une description d'un utilisateur et de son contexte, nous pouvons alors sélectionner la dénomination qui sera la plus adaptée. La définition de notre modèle permettra de mettre en évidence la généralisation de ce principe à des documents. Nous finissons par discuter nos choix de représentation en OWL.

10h30 : Pause

11h00 :

Ontoterminologie : principes et mises-en-œuvre (Démonstration)

Marie Calberg-Challot, Christophe Tricot

Onomia

L'association de l'ontologie et de la terminologie a abouti à l'introduction d'une nouvelle notion, celle d'ontoterminologie. En explicitant et en prenant en compte les deux dimensions conceptuelle et linguistique de la terminologie, ce nouveau paradigme offre de nouvelles perspectives pour l'opérationnalisation des terminologies, que ce soit à des fins de recherche d'information, d'aide à la traduction, de cartographie sémantique, de gestion documentaire ou de capitalisation des connaissances.

Dans le cadre de cette démonstration, nous verrons dans un premier temps l'outil iMap pour la construction de l'ontoterminologie, c'est-à-dire la définition du réseau conceptuel, des termes et des informations associées. La deuxième partie de la démonstration sera consacrée à la présentation des logiciels Codex pour la définition d'encyclopédies métier et Index pour la recherche sémantique d'information. Vision globale de l'ensemble des savoirs, exploration et navigation au travers de cartes et schémas, supports pédagogiques et plateforme ergonomique et collaborative sont autant de fonctionnalités offertes par Codex. Index, quant à lui, permet une interrogation des moteurs internes et externes (méta moteur), une optimisation de la pertinence, une expression de la requête en langage naturel ou via des cartes et une cartographie des résultats. Nous verrons ainsi que l'ontoterminologie joue un rôle central et nécessaire pour une capitalisation pérenne des connaissances.

11h45 :

Filtrage des Entités Nommées par des méthodes de Fouille de Textes

Mathieu Roche

Equipe TAL, LIRMM

Cet article présente une approche de Traitement Automatique du Langage (TAL) afin de filtrer les Entités Nommées à partir d'une liste de candidats à la collocation. La méthode proposée s'appuie uniquement sur des mesures statistiques associées aux ressources du Web. L'évaluation à partir de candidats à la collocation de type Nom-Nom issus d'un corpus en français (corpus de CVs) permet de valider l'approche et de discuter des limites de cette dernière.

12h30 : Repas

14h30 :

Libellex, environnement de gestion collaborative en ligne de terminologie au sein de communautés fermées (Démonstration)

François Brown de Colstoun, Estelle Delpech

Lingua et Machina

Libellex est une plateforme de travail multilingue pour tous les employés de l'entreprise, elle permet des rédactions et traductions exactes et sans fautes quelle que soit la langue. L'environnement est simple et intuitif, aucune formation n'est nécessaire. L'accès se fait par un navigateur via l'intranet de l'entreprise.

15h15 :

Ontologies des risques financiers – Continuité d'activité, gestion de crise, protection des « infrastructures critiques » financières

J. - Yves Gresser

GELM entreprises

Depuis la parution en 1995 du premier livre blanc sur « la sécurité des systèmes d'information dans la banque », le sujet des risques individuels ou systémiques dans la finance est toujours d'actualité.

Ces risques sont multiples. Dans cette communication, nous nous attachons à la représentation des risques liés aux infrastructures, à partir de deux exemples :

- La structuration du lexique de la continuité d'activité de la Banque de France,
- les ontologies de risques financiers développées dans le cadre du projet européen Parsifal.

Il n'existe pas à ce jour d'ontologie générale de la sécurité des systèmes d'information mais des développements spécifiques dans certains domaines. Il existe encore moins d'ontologies relatives au domaine bancaire ou financier. Notre point de départ fut un début de formalisation d'ouvrages de références :

- Critères communs pour la sécurité,
- Lexiques de la Banque de France et de la Banque centrale européenne.

Dans ce domaine comme dans d'autres la spécialisation des connaissances est un obstacle à la bonne appréciation des situations et au traitement approprié des menaces lorsqu'elles s'annoncent ou se matérialisent. Nous avons été les premiers à procéder au rapprochement formel des connaissances dans la sécurité et la finance. Un premier résultat tangible fut l'émergence de la notion de cycle de vie d'une crise et sa cartographie.

Nous avons développé les modèles dans plusieurs directions qui trouvent leur écho dans les crises ou problèmes actuels : vols d'identités, maîtrise des marchés financiers et tout simplement continuité d'activité, dans un monde de plus en plus dépendant des réseaux et de l'informatique.

Nous commençons à partager cette approche innovante et ses résultats avec les parties prenantes : financiers, spécialistes de la sécurité et stratégies politiques européens.

16h00 :

Vers Une ontologie pour le domaine de l'analyse de sécurité des systèmes de transport automatisés

Lassaâd Mejri, Habib Hadj-mabrouk, Patrice Caulier

LAMIH CNRS UMR 8530

INRETS, Institut National de Recherche sur les Transports et leur Sécurité

La résolution de problèmes dynamiques confronte l'opérateur humain à de nombreuses tâches à forte composante de prise d'information et de décision, sous contrainte temporelle. Si elle requiert des savoirs et des savoir-faire fondamentaux, la résolution de problèmes dynamiques nécessite aussi l'exploitation d'une expertise, à la fois statique et dynamique, trop peu souvent formalisée et capitalisée à des fins, notamment, d'assistance ou de supervision.

Une expérience de développement de systèmes d'aide à la résolution de problèmes dynamiques dans plusieurs domaines, notamment dans les domaines de l'analyse sécuritaire (Mejri ,95), (Mejri et Houriez, 98) , de la fiabilité du logiciel (Hadj mabrouk et al, 2000), du web sémantique (Mejri et al, 2008), de l'information multimodale (Mejri et Ben fraj, 2008) ou du génie logiciel (Kessentini et al ,2009) nous a amené à spécifier et modéliser un cadre générique de formalisation de l'expertise spatio-temporelle de résolution de problèmes dynamiques baptisé « scénario de résolution de problème » (Mejri et Caulier, 2005). En effet, le problème de l'analyse de sécurité des systèmes de transport automatisés est un problème typique où les experts du domaine proposent des situations d'insécurité potentielles conjointement par des descriptions statique et dynamique.

16h45 : Table Ronde, Pot de départ