

TALN 2015

Table des matières

Préface

Multilinguality at Your Fingertips : BabelNet, Babelfy and Beyond !.....	i-i
Pourquoi construire des ressources terminologiques et pourquoi le faire différemment ?.....	ii-ii

Session Extraction d'information

Apprentissage par imitation pour l'étiquetage de séquences : vers une formalisation des méthodes d'étiquetage « easy-first ».....	1-12
Stratégies de sélection des exemples pour l'apprentissage actif avec des champs aléatoires conditionnels.....	13-24
Identification de facteurs de risque pour des patients diabétiques à partir de comptes-rendus cliniques par des approches hybrides.....	25-36
Oublier ce qu'on sait, pour mieux apprendre ce qu'on ne sait pas : une étude sur les contraintes de type dans les modèles CRF.....	37-48

Session Compréhension et paraphrase

Analyse d'expressions temporelles dans les dossiers électroniques patients.....	49-58
Compréhension automatique de la parole sans données de référence.....	59-70

Session Désambiguïsation

Désambiguïsation d'entités pour l'induction non supervisée de schémas événementiels.....	71-82
Création rapide et efficace d'un système de désambiguïsation lexicale pour une langue peu dotée.....	83-94

Session Opinions et sentiments

Méthode faiblement supervisée pour l'extraction d'opinion ciblée dans un domaine spécifique.....	95-106
Une méthodologie de sémantique de corpus appliquée à des tâches de fouille d'opinion et d'analyse des sentiments : étude sur l'impact de marqueurs dialogiques et dialectiques dans l'expression de la subjectivité.....	107-118

Session Sémantique

Estimation de l'homogénéité sémantique pour les Questionnaires à Choix Multiples.....	119-130
Extraction automatique de relations sémantiques dans les définitions : approche hybride, construction d'un corpus de relations sémantiques pour le français.....	131-145
Déclasser les voisins non sémantiques pour améliorer les thésaurus distributionnels.....	146-157

Session Syntaxe et paraphrase

Grammaires phrastiques et discursives fondées sur les TAG : une approche de D-STAG avec les ACG.....	158-169
Noyaux de réécriture de phrases munis de types lexico-sémantiques.....	170-181
Extraction automatique de paraphrases grand public pour les termes médicaux.....	182-193
Analyse syntaxique de l'ancien français : quelles propriétés de la langue influent le plus sur la qualité de l'apprentissage ?.....	194-207

Session Classification et Alignement

Attribution d'Auteur : approche multilingue fondée sur les répétitions maximales.....	208-219
Mesurer la similarité entre phrases grâce à Wikipédia en utilisant une indexation aléatoire.....	220-231
Typologie automatique des langues à partir de treebanks.....	232-243

Session Traduction

Utilisation de mesures de confiance pour améliorer le décodage en traduction de parole.....	244-254
Multialignement vs bialignement : à plusieurs, c'est mieux !.....	255-266
Apprentissage discriminant des modèles continus de traduction.....	267-278

Session Plénière

Utiliser les interjections pour détecter les émotions.....	279-292
Comparaison d'architectures neuronales pour l'analyse syntaxique en constituants.....	293-304
...des conférences enfin disons des causeries... Détection automatique de segments en relation de paraphrase dans les reformulations de corpus oraux.....	305-316

Posters 1

Une méthode discriminant formation simple pour la traduction automatique avec Grands Caractéristiques.....	317-322
Natural Language Reasoning using Coq: Interaction and Automation.....	323-329
Vous aimez ?...ou pas ? LikeIt, un jeu pour construire une ressource lexicale de polarité.....	330-336
Étude des verbes introducteurs de noms de médicaments dans les forums de santé.....	337-343
Initialisation de Réseaux de Neurones à l'aide d'un Espace Thématique.....	344-349
FDTB1: Repérage des connecteurs de discours en corpus.....	350-356
ROBO, an edit distance for sentence comparison Application to automatic summarization.....	357-363
Classification d'entités nommées de type « film ».....	364-370
A critical survey on measuring success in rank-based keyword assignment to documents.....	371-376
Effects of Graph Generation for Unsupervised Non-Contextual Single Document Keyword Extraction.....	377-383
Adaptation par enrichissement terminologique en traduction automatique statistique fondée sur la génération et le filtrage de bi-segments virtuels.....	384-390
Une mesure d'intérêt à base de surreprésentation pour l'extraction des motifs syntaxiques stylistiques.....	391-396
Une Approche évolutionnaire pour le résumé automatique.....	397-403
Identification des unités de mesure dans les textes scientifiques.....	404-410
Évaluation intrinsèque et extrinsèque du nettoyage de pages Web.....	411-417
CANÉPHORE : un corpus français pour la fouille d'opinion ciblée.....	418-424
Extraction de Contextes Riches en Connaissances en corpus spécialisés.....	425-431
Traitement automatique des formes métriques des textes versifiés.....	432-438
Apprentissage automatique d'un modèle de résolution de la coréférence à partir de données orales transcrites du français : le système CROC.....	439-445
Vers un diagnostic d'ambiguïté des termes candidats d'un texte.....	446-452
Augmentation d'index par propagation sur un réseau lexical Application aux comptes rendus de radiologie.....	453-459
Détection automatique de l'ironie dans les tweets en français.....	460-465
Dictionnaires morphologiques du français contemporain : présentation de Morfetik, éléments d'un modèle pour le TAL.....	466-472

Posters 2

Une métagrammaire de l'interface morpho-sémantique dans les verbes en arabe.....	473-479
Création d'un nouveau treebank à partir de quatrièmes de couverture.....	480-486
Entre écrit et oral ? Analyse comparée de conversations de type tchat et de conversations téléphoniques dans un centre de contact client.....	487-493
Construction et maintenance d'une ressource lexicale basées sur l'usage.....	494-500
Utilisation d'annotations sémantiques pour la validation automatique d'hypothèses dans des conversations téléphoniques.....	501-507
Etiquetage morpho-syntaxique en domaine de spécialité: le domaine médical.....	508-514
Vers une typologie de liens entre contenus journalistiques.....	515-521
CDGFr, un corpus en dépendances non-projectives pour le français.....	522-528
Utilisation des réseaux de neurones récurrents pour la projection interlingue d'étiquettes morpho-syntaxiques à partir d'un corpus parallèle.....	529-536
Segmentation et Titrage Automatique de Journaux Télévisés.....	537-543
Un système hybride pour l'analyse de sentiments associés aux aspects.....	544-550
La ressource EXPLICADIS, un corpus annoté spécifiquement pour l'étude des relations de discours causales.....	551-557
La séparation des composantes lexicale et flexionnelle des vecteurs de mots.....	558-564
Traitements pour l'analyse du français préclassique.....	565-571
Classification de texte enrichie à l'aide de motifs séquentiels.....	572-578
Le traitement des collocations en génération de texte multilingue.....	579-585
Médicaments qui soignent, médicaments qui rendent malades : étude des relations causales pour identifier les effets secondaires.....	586-592
Exploration de modèles distributionnels au moyen de graphes 1-PPV.....	593-599
Apport de l'information temporelle des contextes pour la représentation vectorielle continue des mots.....	600-606
Etiquetage morpho-syntaxique de tweets avec des CRF.....	607-613
Caractériser les discours académiques et de vulgarisation : quelles propriétés ?.....	614-620
Extraction et analyse automatique des comparaisons et des pseudo-comparaisons pour la détection des comparaisons figuratives.....	621-627
Proposition méthodologique pour la détection automatique de Community Manager. Étude multilingue sur un corpus relatif à la Junk Food.....	628-634

Démonstrations

MEDITE : logiciel d'alignement de textes pour l'étude de la génétique textuelle.....	635-636
Phœbus : un Logiciel d'Extraction de Réutilisations dans des Textes Littéraires.....	637-639
YADTK : Une plateforme open-source à base de règles pour développer des systèmes de dialogue oral.....	640-641
TermLis : un contexte d'information logique pour des ressources terminologiques.....	642-643
Etude de l'image de marque d'entités dans le cadre d'une plateforme de veille sur le Web social.....	644-645
Building a Bilingual Vietnamese-French Named Entity Annotated Corpus through Cross-Linguistic Projection.....	646-647
Recherche de motifs de graphe en ligne.....	648-649
Un patient virtuel dialogant.....	650-651
Intégration du corpus des actes de TALN à la plateforme ScienQuest.....	652-653
Une aide à la communication par pictogrammes avec prédiction sémantique.....	654-656
Un système expert fondé sur une analyse sémantique pour l'identification de menaces d'ordre biologique.....	657-658
DisMo : un annotateur multi-niveaux pour les corpus oraux.....	659-661