



**IRIT**

Institut de Recherche  
en Informatique  
de Toulouse

Séminaire



**Mercredi 27 Janvier 2016**

**14h30 – 16h00**

**UT3 Paul Sabatier, IRIT, Auditorium J. Herbrand**

**Olivier FERRET**

**CEA LIST - Laboratoire Vision et Ingénierie des contenus  
(LVIC) (France)**

## **Combiner différents critères pour réordonnancer les thésaurus distributionnels**

*Résumé* : Bien que la notion de thésaurus distributionnel soit connue depuis longtemps, les améliorations dont ils font l'objet sont généralement la conséquence indirecte d'améliorations touchant les moyens - représentations ou mesures de similarité - de détecter la similarité sémantique entre les mots. Dans cette présentation, nous proposons de nous focaliser de façon plus directe sur l'amélioration des thésaurus distributionnels en tant que tels par l'intermédiaire de processus de réordonnancement des voisins associés à leurs entrées. Nous présenterons plus particulièrement deux types de méthodes pour un tel réordonnancement. L'un s'appuie sur la sélection non supervisée d'exemples pour construire un modèle permettant de réévaluer la similarité entre une entrée et ses voisins. L'autre repose la détection et le déclassement des voisins non similaires à leur entrée en exploitant un modèle discriminant de l'entrée en contexte s'inspirant de la désambiguïsation sémantique. Nous montrerons au travers d'évaluations à large échelle le caractère effectif de ces différentes méthodes, en particulier pour les entrées de fréquences moyennes et faibles, ainsi que l'intérêt de les combiner.

05 61 55 65 10

[info@irit.fr](mailto:info@irit.fr)

[www.irit.fr](http://www.irit.fr)



Organisé par  
**Département IA - Equipe MELODI**