

TAL et Linguistique outillée

Béatrice, DAILLE, LS2N – Université de Nantes, FRANCE

Mots-clés — TAL, Linguistique outillée, texte de spécialité

Le traitement automatique des langues est une discipline de l'informatique et des sciences du langage qui s'intéresse à la modélisation, au traitement et à l'analyse des productions et processus cognitifs langagiers. Comme discipline informatique, le TAL définit des algorithmes, construit des ressources langagières, élabore des outils et des architectures logicielles. Les outils développés pour ses recherches peuvent aussi être valorisés en linguistique appliquée. Après une présentation du domaine du traitement automatique des langues, nous présenterons quelques outils à l'origine développés pour le TAL qui sont devenus populaires en linguistique appliquée comme les concordanciers ou les plate-formes d'annotation pour l'enseignement, ainsi que quelques méthodes élaborées par les talistes pour la linguistique qui sont restées au niveau du prototype. Aujourd'hui, le TAL adopte des méthodes fondées sur les réseaux de neurones profonds qui ont permis à nombre de tâches du TALN, comme la traduction ou le résumé automatique, d'effectuer un bond qualitatif mais qui nécessitent des compétences techniques pointues. Les dernières architectures neuronales possèdent une capacité de prédiction importante qui permet non seulement une amélioration des outils précédents mais qui pourrait aussi produire une nouvelle génération d'outils plus performants pour l'aide à la lecture, la détection d'erreurs, ou encore la génération d'activités.

