Résumé

La construction des bases de données lexicales à partir de textes constitue un enjeu important pour le Traitement Automatique des Langues Naturelles (TALN).En effet les systèmes de TALN ne peuvent être efficaces que s'ils reposent sur des ressources lexicales bien construites pour le domaine et l'application concernés.

Dans ce mémoire nous avons développé un ensemble d'outils pour la construction d'une base de données lexicale pour la langue Arabe « AraWord ». Cette dernière peut être utilisée dans différentes applications pour le TALN.

Mots Clés : WordNet, Ontologie, Base de données lexicale, Construction d'ontologie, TALN Arabe.

ملخّص:

تهدف هذه المذكرة الى تطوير قاعدةبياناتمعجمية للغة العربية بالإعتماد على البنك العالمي للمصطلحات وورد نات "Wordnet ".لأجل هذا لقد قمنا بتطوير مجموعة من البرمجيات المساعدة على اكتشاف وضبط المصطلحات الموجودة في نصوص المدونة المستعملة لهذا الغرض. يمكن استعمال القاعدة المنجزة في عدة تطبيقات تخص المعالجة الآلية للغة العربية.

الكلمات المفتاحية :وورد نات، الأنطولوجية،بنك آلي لمصطلحات،تطوير الأنطولوجيات،المعالجة الآلية للغة العربية.

Abstract

The construction of lexical databases from texts is an important issue for Automatic Natural Language Processing (NLP). Indeed, NLP systems can only be effective if they are based on well-constructed lexical resources for the domain and application involved.

In this thesis we have developed a set of tools for building a lexical database for the Arabic "AraWord." It can be used in different applications for NLP.

Keywords: WordNet, Ontology, Lexical Databases,Ontology construction, TALN, Arabic NLP.