



5.00 crédits	22.5 h	Q1
--------------	--------	----

Enseignants	Fairon Cédric ;
Langue d'enseignement	Français > English-friendly
Lieu du cours	Louvain-la-Neuve
Préalables	/
Thèmes abordés	Le cours débute par l'étude de l'architecture d'un système complexe de traitement automatique du langage (reconnaissance, analyse, génération). Il se poursuit par l'étude de théories linguistiques et formalismes informatiques centraux dans le domaine du TALN. Une attention particulière est accordée à la présentation et à l'analyse d'applications réelles.
Acquis d'apprentissage	<p>A la fin de cette unité d'enseignement, l'étudiant est capable de :</p> <p>Le cours apportera à l'étudiant les bases théoriques nécessaires à la compréhension des objectifs et des défis actuels du traitement automatique du langage naturel (TALN). En parallèle, l'étudiant apprendra à analyser et à expliquer les limites pratiques et techniques que l'on rencontre dans la mise en place de systèmes informatiques destinés au traitement des langues (problèmes d'ambiguïtés, nécessité d'adaptabilité des ressources linguistiques, multilinguisme, etc.). Au terme du cours, l'étudiant aura eu un aperçu de " l'état de l'art " en matière de TALN, sera en mesure de porter un regard critique sur des applications de TALN et aura une connaissance générale des principales théories du domaine</p>
Modes d'évaluation des acquis des étudiants	<ul style="list-style-type: none"> - Deux travaux pratiques à réaliser au cours du quadrimestre [6 points] - Examen écrit (ou oral) portant majoritairement sur le cours et, dans une moindre mesure, sur les concepts importants issus des lectures obligatoires [14 points]
Contenu	<p>Le cours est donné sous forme d'exposé magistral interactif. Un dossier de lecture composé d'articles spécialisés permet aux étudiants de préparer les cours.</p> <p>Plan du cours :</p> <ul style="list-style-type: none"> - Le domaine du TAL (dénomination, aperçu historique, niveaux d'analyse) - Le codage et les pré-traitements - Les langages formels (expressions régulières, FSA) - Les modèles de langue probabilistes (notions de probabilité, modèles n-grams) - Les ressources lexicales (dictionnaires électroniques, etc.) - La lemmatisation - La détermination des natures (approche par règles, HMMs) - Les grammaires formelles (hiérarchie de Chomsky, grammaires non contextuelles) - Le parsing syntaxique (principes généraux, alternatives) - La sémantique lexicale (thésaurus, ontologies, WordNet) - La sémantique vectorielle (distributionnalisme, plongements lexicaux)
Faculté ou entité en charge:	FIAL

Programmes / formations proposant cette unité d'enseignement (UE)				
Intitulé du programme	Sigle	Crédits	Prérequis	Acquis d'apprentissage
Master [120] en science des données, orientation statistique	DATS2M	5		
Master [120] en langues et lettres françaises et romanes, orientation français langue étrangère	FLE2M	5		
Master [120] en linguistique	LING2M	5		