






5.00 crédits	30.0 h + 15.0 h	Q1
--------------	-----------------	----

Enseignants	Dupont Pierre ;
Langue d'enseignement	Anglais
Lieu du cours	Louvain-la-Neuve
Thèmes abordés	<ul style="list-style-type: none"> • Différents niveaux d'analyses linguistiques • Traitements de corpus • Etiquetage en parties du discours • Modélisation probabiliste du langage (Ngrams et Modèles de Markov Cachés) • Grammaires formelles et algorithmes d'analyse syntaxique • Traduction automatique, apprentissage profond • Applications en ingénierie linguistique telles que les logiciels de complétion automatique, d'étiquetage automatique, d'analyse ou de traduction automatique
Acquis d'apprentissage	<p>A la fin de cette unité d'enseignement, l'étudiant est capable de :</p> <p>Eu égard au référentiel AA du programme « Master ingénieur civil en informatique », ce cours contribue au développement, à l'acquisition et à l'évaluation des acquis d'apprentissage suivants :</p> <p>INFO1.1-3 INFO2.3-4 INFO5.3-5 INFO6.1, INFO6.4</p> <p>Eu égard au référentiel AA du programme « Master [120] en sciences informatiques », ce cours contribue au développement, à l'acquisition et à l'évaluation des acquis d'apprentissage suivants :</p> <p>SINF1.M4 SINF2.3-4 SINF5.3-5 SINF6.1, SINF6.4</p> <p>Les étudiants ayant suivi avec fruit ce cours seront capables de</p> <ol style="list-style-type: none"> 1 <ul style="list-style-type: none"> • décrire les concepts fondamentaux de la modélisation du langage naturel • maîtriser la méthodologie de l'utilisation de ressources linguistiques, en particulier sous la forme de corpus à grande échelle, éventuellement annotés ou structurés • appliquer de manière pertinente les techniques statistiques de modélisation du langage • mettre en oeuvre des méthodes récentes relevant de l'apprentissage automatique en lien avec le traitement du langage • développer des applications en ingénierie linguistique <p>Les étudiants auront développé des compétences méthodologiques et opérationnelles. En particulier, ils auront développé leur capacité à</p> <ul style="list-style-type: none"> • s'intégrer dans une approche pluridisciplinaire entre l'informatique et la linguistique, en utilisant à bon escient la terminologie, les outils et méthodes existants, • gérer au mieux le temps disponible pour mener à bien des projets de moyenne ampleur, • manipuler et exploiter de grandes quantités de données.
Modes d'évaluation des acquis des étudiants	<p>Les projets valent pour 30 % de la note finale, 70 % pour l'examen final (à livre fermé).</p> <p>Les projets ne peuvent pas être refaits en seconde session. Les notes des projets sont donc déjà fixées à la fin du quadrimestre et reprises tels quels dans la note finale en seconde session.</p> <p>L'examen final est, par défaut, un écrit (sur papier ou, le cas échéant, sur un ordinateur).</p>
Méthodes d'enseignement	<ul style="list-style-type: none"> • Cours magistraux • Projets pratiques implémentés en Python sur le serveur <i>Inginious</i>
Contenu	<ul style="list-style-type: none"> • Différents niveaux d'analyse linguistique • Traitement (automatisé) de corpus : formatage, tokenization, marquage des données • Modélisation probabilistes du langage : N-grams, HMMs • Etiquetage en parties du discours • Grammaires hors-contexte (probabilistes) : estimation des paramètres et algorithmes d'analyse

	<ul style="list-style-type: none">• Introduction à la traduction automatique• Introduction à l'apprentissage profond• Applications représentatives telles que la complétion automatique, le marquage automatique de parties de texte, l'analyse syntaxique ou la traduction automatique.
Ressources en ligne	moodle.uclouvain.be/course/view.php?id=1182
Bibliographie	One recommended textbook - un ouvrage conseillé : <ul style="list-style-type: none">• Speech and Language Processing, D. Jurafsky and J.H. Martin, Prentice Hall.
Faculté ou entité en charge:	INFO

Programmes / formations proposant cette unité d'enseignement (UE)				
Intitulé du programme	Sigle	Crédits	Prérequis	Acquis d'apprentissage
Master [120] en linguistique	LING2M	5		
Master [120] : ingénieur civil en science des données	DATE2M	5		
Master [120] : ingénieur civil en informatique	INFO2M	5		
Master [120] en science des données, orientation technologies de l'information	DATI2M	5		
Master [120] en sciences informatiques	SINF2M	5		
Master [120] en science des données, orientation statistique	DATS2M	5		