

En raison de la crise du COVID-19, les informations ci-dessous sont susceptibles d'être modifiées, notamment celles qui concernent le mode d'enseignement (en présentiel, en distanciel ou sous un format comodal ou hybride).






5 crédits	30.0 h + 15.0 h	Q1
-----------	-----------------	----

Enseignants	Dupont Pierre ;Dupont Pierre (supplée Fairon Cédric) ;Fairon Cédric ;
Langue d'enseignement	Anglais
Lieu du cours	Louvain-la-Neuve
Thèmes abordés	<ul style="list-style-type: none"> • Concepts de base en phonologie, morphologie, syntaxe et sémantique • Ressources linguistiques • Etiquetage en partie du discours • Modélisation statistique de la langue (N-grams et modèles de Markov cachés) • Algorithmes d'analyse robuste, grammaires hors-contexte probabilistes • Applications de l'ingénierie linguistique telles que les logiciels de correction orthographique ou grammaticale, l'étiquetage en partie du discours ,l'indexation automatique de documents ou la catégorisation de textes
Acquis d'apprentissage	<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> <ul style="list-style-type: none"> • INFO1.1-3 • INFO2.3-4 • INFO5.3-5 • INFO6.1, INFO6.4 <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> <ul style="list-style-type: none"> • SINF1.M4 • SINF2.3-4 • SINF5.3-5 • SINF6.1, SINF6.4 <p>1</p> <p>Les étudiants ayant suivi avec fruit ce cours seront capables de</p> <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 (corpus, dictionnaires, réseaux sémantiques, etc) et faire un choix argumenté entre les différentes ressources linguistiques • appliquer de manière pertinente les techniques statistiques de modélisation 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 à la frontière entre l'informatique et la linguistique, en utilisant à bon escient la terminologie et les outils de l'une ou l'autre discipline, • gérer au mieux le temps disponible pour mener à bien des mini-projets, • manipuler et exploiter de grandes quantités de données. <p>-----</p> <p><i>La contribution de cette UE au développement et à la maîtrise des compétences et acquis du (des) programme(s) est accessible à la fin de cette fiche, dans la partie « Programmes/formations proposant cette unité d'enseignement (UE) ».</i></p>
Modes d'évaluation des acquis des étudiants	<p>En raison de la crise du COVID-19, les informations de cette rubrique sont particulièrement susceptibles d'être modifiées.</p> <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> <p>Ces règles d'évaluation sont sujettes à d'éventuelles mises à jour en fonction de la situation sanitaire. En particulier, le poids relatif des projets et de l'examen final pourrait être adapté. De telles adaptations seraient alors notifiées aux étudiants via une annonce générale sur le site Moodle du cours.</p>

Méthodes d'enseignement	<p>En raison de la crise du COVID-19, les informations de cette rubrique sont particulièrement susceptibles d'être modifiées.</p> <ul style="list-style-type: none"> • Cours magistraux • Projets pratiques implémentés en Python <p>Les cours magistraux sont donnés, par défaut, en présentiel. Selon le nombre effectif d'étudiant.e.s inscrit.e.s au cours et l'évolution de la situation sanitaire, un dispositif pour pouvoir suivre le cours à distance sera également mis en place (enseignement co-modal).</p> <p>Les projets sont soumis en ligne et évalués sur la plateforme <i>Inginious</i>.</p>
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 • 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	http://moodleucl.uclouvain.be/course/view.php?id=7865
Bibliographie	<p>One recommended textbook - un ouvrage conseillé :</p> <ul style="list-style-type: none"> • Speech and Language Processing (2nd Edition), D. Jurafsky and J.H. Martin, Prentice Hall, 2009.
Faculté ou entité en charge:	INFO

Force majeure

Méthodes d'enseignement	<p>Les cours magistraux se donnent en distanciel. Les travaux pratiques continuent à être en ligne sur le serveur <i>Inginious</i>.</p>
Modes d'évaluation des acquis des étudiants	<p>L'examen final est un écrit à faire individuellement en ligne. Il est à livre ouvert.</p> <p>La matière sur laquelle porte cette évaluation est identique au cas normal (voir supports de cours).</p> <p>La pondération pour le calcul de la note globale du cours est revue comme suit : 50 % examen final, 50% projets effectués pendant le quadrimestre.</p> <p>La note des projets est définitive à l'issue du quadrimestre : il n'y a pas de possibilité de refaire ces projets en seconde session.</p>

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 linguistique	LING2M	5		
Master [120] : ingénieur civil en informatique	INFO2M	5		
Master [120] en sciences informatiques	SINF2M	5		
Master [120] : ingénieur civil en science des données	DATE2M	5		
Master [120] en science des données, orientation technologies de l'information	DATI2M	5		