

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 + 10.0 h	Q2
-----------	-----------------	----

Enseignants	Fouss François ;Françq Pascal ;
Langue d'enseignement	Français
Lieu du cours	Mons
Thèmes abordés	<p>L'objectif de cet enseignement est d'initier les étudiants à la collecte, au nettoyage et à l'analyse des données du web. Le cours abordera les éléments suivants :</p> <ul style="list-style-type: none"> • Une introduction, pour rappeler la nécessité de comprendre et donner du sens aux données issues du web • Les explications sur les différents types de données qu'on trouve sur le web, en incluant les métadonnées, et une sensibilisation aux différentes manières de les extraire. • Introduire ici les notions de code et sensibiliser aux différents langages existants. • Le cours introduira ensuite deux types d'analyses de données : • L'analyse lexicométrique en lien avec la linguistique • L'analyse des graphes en lien avec la théorie des réseaux sociaux
Acquis d'apprentissage	<p>Au terme de cet enseignement, l'étudiant sera capable :</p> <p>1</p> <ul style="list-style-type: none"> • De comprendre la structures des données numériques • De comprendre les grandes étapes de collecte, de nettoyage et d'exploitation des données • De réaliser une analyse lexicométrique et une analyse de graphes de base <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>L'évaluation (unique) a une composante pratique et une composante théorique:</p> <ul style="list-style-type: none"> • Evaluation continue (travaux pratiques cotés); • Réalisation d'un projet de groupe; • Défense du projet et examen oral. <p>!!! La note obtenue est unique et ne peut donc pas être améliorée durant l'année académique en 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"> • Exposés magistraux • Exercices intégrés au cours • Projet par groupe
Contenu	<p>Ce cours a pour objectif d'apprendre à appliquer une méthodologie adéquate d'analyse de données numériques issues du web, et d'apprendre à maîtriser les techniques et modèles ad-hoc, afin notamment de pouvoir les utiliser/ appliquer dans des situations réelles.</p> <p>Les thèmes principaux de ce cours sont:</p> <ul style="list-style-type: none"> • Recherche d'information sur le Web <ul style="list-style-type: none"> • Concepts de base (indexation, de compression, de recherche, de classification et de catégorisation) • Modèles de recherche d'information • Ingénierie du Web <ul style="list-style-type: none"> • Référencement • Collecter, nettoyer et analyser les données issues du Web • Outils d'analyse d'audience de sites Web

	<ul style="list-style-type: none"> • Link Analysis : recommander/analyser du contenu relevant sur base de la structure en hyperliens, formant un réseau/graphe <ul style="list-style-type: none"> • Éléments fondamentaux de la structure d'un réseau/graphe • Méthodologie d'analyse d'un réseau/graphe <ul style="list-style-type: none"> • Identification de sous-groupes cohésifs • Notions de similarité et de distance • Identification de nœuds prestigieux • Identification de nœuds centraux • Prédiction de nouveaux liens • etc.
Ressources en ligne	Voir Student Corner / Teams
Bibliographie	<ul style="list-style-type: none"> • MCILWRAITH D., MARMANIS H., BABENKO D. 2nd ed, Algorithms of the Intelligent Web, Manning Publications, 2016. • LANGVILLE A., MEYER C., Google's PageRank and Beyond : The Science of Search Engine Rankings, Princeton University Press, 2012. • FOUSS F., SAERENS M., SHIMBO M., Algorithms and Models for Network Data and Link Analysis, Cambridge University Press, 2016. • AMINI M.-R., GAUSSIER E., Recherche d'information : Applications, modèles et algorithmes, Eyrolles, 2013. • MANNING C. D., RAGHAVAN P., SCHÜTZE H., Introduction to Information Retrieval, Cambridge University Press, 2008. • MARTIN A., CHARTIER M., ANDRIEU O., Techniques de référencement web : Audit et suivi SEO, Eyrolles, 2016.
Faculté ou entité en charge:	COMU

Programmes / formations proposant cette unité d'enseignement (UE)				
Intitulé du programme	Sigle	Crédits	Prérequis	Acquis d'apprentissage
Bachelier en information et communication	COMM1BA	5		
Bachelier en sciences humaines et sociales	HUSM1BA	5		