

FICHE ECTS

Composante (dépt ou centre)	DGEI		
Code ECTS	I5SSIF11		
Crédits ECTS	4		
Enseignant responsable de l'UF	Nawal Guermouche		
Section CNU			
Libellé de l'UF (français)	Analyse et traitement des données, applications métier		
Libellé de l'UF (anglais)	Data analysis and processing, business applications		
Semestre :	Semestre 1 <input type="checkbox"/>	Semestre 2 <input type="checkbox"/>	Annuel <input type="checkbox"/>

DESCRIPTION GENERALE

VERSION FRANÇAISE

VERSION ANGLAISE

Objectifs, finalités ¹ (max 1000 caract.)	Objectives (max 1000 charact.)
<p>A la fin de ce module constitué de différentes thématiques principales, l'étudiant devra avoir compris et pourra expliquer (principaux concepts) :</p> <p><u>Algorithmique du traitement de données :</u></p> <p>Analyse Exploratoire/Confirmatoire des données. Complexité algorithmique, parallélisme et enjeux du développement incrémental d'une solution d'analyse. Représentation (parcimonieuse) de l'information</p> <p><u>Représentation sémantique :</u></p> <ul style="list-style-type: none"> - Qu'est-ce qu'une ontologie - Quels sont les éléments constitutifs d'une ontologie - Quels sont les avantages des données enrichies comparées aux données brutes <p><u>Ingénierie Logicielle</u></p> <ul style="list-style-type: none"> - Le cycle de vie d'un projet logiciel - Les enjeux du développement logiciel - Les différentes méthodes de gestion de projet, notamment la méthode agile <p>L'étudiant devra être capable de :</p> <ul style="list-style-type: none"> • Explorer un jeu de données, l'exploiter par rapport à une problématique et présenter les résultats de ses analyses dans un rapport. • Concevoir une ontologie pour formaliser un domaine de connaissances • Découvrir et s'approprier des sources de connaissance (ontologies, bases de connaissances) en ligne • Enrichir un jeu de données à l'aide de métadonnées sémantiques • Maîtriser la conduite d'un projet de développement logiciel mené en équipe, notamment en suivant la méthode agile • Mettre en pratique l'analyse des besoins à partir d'un cahier des charges : expression, analyse et transformation en exigences techniques • Maîtriser les activités d'assurance qualité du projet et la conduite des tests pour la validation et la vérification <p><i>Le verbe qui décrit l'objectif doit décrire de façon précise l'effet attendu, on conseille d'éviter des verbes vagues pour lesquels il sera difficile de vérifier si l'objectif est atteint (connaître, comprendre, apprendre, savoir...)</i> <i>On peut s'inspirer des verbes d'actions de la taxonomie de Bloom</i></p>	<p>At the end of this module, the student will have understood and be able to explain (main concepts):</p> <p><u>Data management:</u></p> <p>Exploratory/confirmatory data analysis. Algorithmic complexity vs. development costs, parallelism, software engineering notions (life cycle of a data analysis pipeline). Data visualisation techniques. Machine Learning notions.</p> <p><u>Semantic manipulation:</u></p> <ul style="list-style-type: none"> - What an ontology is - What are the main constituting elements of an ontology - What are the perks of enriched data compared to raw data <p><u>Software engineering:</u></p> <ul style="list-style-type: none"> - Software project lifecycle - The challenges of software development - Project management methods, including agile method <p>The student will be able to:</p> <ul style="list-style-type: none"> - Explore a dataset, leverage it to answer specific questions, and present the results of this analysis -incl. Its limits- in a synthetic written report. - Design an ontology to capture domain knowledge - Discover and reuse knowledge sources (ontologies, knowledge bases) online - Enrich a dataset with semantic metadata - Control the conduct of a software development project with a team by following the agile method • Perform requirement analysis: expression, analysis and transformation into technical requirements • Control of the quality assurance activities and the conduct of tests for validation and verification

Contenu (max 1000 caract.)	Description (max 1000 charact.)
<p><u>Algorithmique du traitement de données :</u></p> <ul style="list-style-type: none"> -R -EDA/CDA -Algorithmique, parallélisme -AI/ML <p>ingénierie logicielle</p> <ul style="list-style-type: none"> - Méthodologies de gestion de projet logiciel - Ingénierie des besoins: expression, transformation des exigences fonctionnelles en exigences techniques - Qualité logiciel: gestions de configuration, assurance et contrôle qualité, tests et validation <p><i>Vous pouvez aussi mentionner ci-dessus les documents remis aux étudiants</i></p>	<p><u>Data management:</u></p> <ul style="list-style-type: none"> -R -EDA/CDA -Algorithms vs implementation, parallelism -AI/ML <p>software engineering</p> <ul style="list-style-type: none"> - Software project management methodologies - Requirement engineering: expression, transformation of functional requirements into technical requirements - Software Quality: configuration management, insurance and quality control, testing and validation

Recommandation (max 1000 caract.)	Recommendation (max 1000 charact.)
<ul style="list-style-type: none"> - Combiner approche incrémentale (exploratoire) et développement logiciel fiable. - R 	<ul style="list-style-type: none"> - R - Statistics - Incremental exploration vs. reliable software

<ul style="list-style-type: none"> - Statistiques - Appropriation du vocabulaire (et des notions sous-jacentes) spécifique au web sémantique, complètement nouveau pour eux - Mise en relation avec des projets concrets qu'ils ont déjà rencontrés 	<p>development.</p> <ul style="list-style-type: none"> - Mastering the vocabulary (and the underlying knowledge), specific to the semantic web, that are completely new to them
<i>Principales difficultés habituellement rencontrées par les étudiants</i>	

<i>Pré-requis (Code UF + intitulé, sinon notions nécessaires) (max 200 caract.)</i>	<i>Necessary knowledge (UF Code + title, or required knowledge) (max 200 charact.)</i>
<ul style="list-style-type: none"> - Programmation et algorithmique - Notions de statistiques - Programmation en Java - Culture générale sur les technologies web 	<ul style="list-style-type: none"> - Algorithms and programming - Statistics (notions) - Java programming - Web technologies background knowledge

<i>Organisation, méthodes pédagogiques</i>			<i>Organisation, teaching methods</i>		
<p><u>Algorithmique du traitement de données :</u></p> <p>1/EDA/CDA 2/R 3/Algorithmique/parallélisme 4/AI/ML</p> <p><u>Représentation sémantique :</u></p> <ul style="list-style-type: none"> - Introduction aux besoins pour un web sémantique par rapport au web classique - Présentation des principes et des technologies du web sémantique, instanciés dans le cas de l'internet des objets - Ouverture sur des initiatives concrètes de web sémantique dans l'internet des objets - Mise en pratique des notions présentées en cours par la création d'une ontologie, puis la manipulation d'un jeu de données <p>ingénierie logicielle</p> <p>Des cours magistraux accompagnés de supports de cours seront donnés aux étudiants. Des travaux dirigés sont adossés aux cours magistraux durant lesquels les étudiants vont exercer les différents concepts liés à l'ingénierie logicielles. Les connaissances acquises seront exercées dans le cadre des projets intégrateurs.</p>			<p><u>Data management:</u></p> <p>1/EDA/CDA 2/R 3/Algorithms vs implementation, parallelism 4/AI/ML</p> <p><u>Semantic manipulation:</u></p> <ul style="list-style-type: none"> - Introduction to the need for a semantic web compared to the classic web - Presentation of the principles and technologies of the semantic web, instantiated with IoT use cases - Opening on concrete initiatives of semantic web of things - Creation of an ontology, and manipulation of a dataset according to the theoretical elements presented in lectures. <p>software engineering</p> <p>Lectures accompanied of course material will be given to students. During directed labs the students will practice the different concepts related to software engineering. The acquired knowledge will be exercised in the context of integration projects.</p>		
<i>Horaire présentiel (tel que l'enseignement est comptabilisé)</i>		<i>Travail personnel²</i>	<i>Contact hours</i>		<i>Personal work</i>
CM	13.75		Lectures	13.75	
TD	10		Tutorials	10	
TP	13.75		Lab work	13.75	
Projet			Project		
Examen formatif			Coursework		
Examen certificatif			Exam		

Format d'enseignement :	Présentiel <input type="checkbox"/>	Distanciel <input type="checkbox"/>	Hybride <input type="checkbox"/>
--------------------------------	-------------------------------------	-------------------------------------	----------------------------------

DESCRIPTION COMPLEMENTAIRE

Modalités d'évaluation (max 1000 caract.)		Assessment (max 1000 charact.)	
Comment évaluez-vous que ces objectifs sont atteints ?			
<p>Algorithmique du traitement de données :</p> <ul style="list-style-type: none"> - Projet final : exploration d'un jeu de données au choix et présentation d'un rapport. - Les étudiants doivent répondre à un certain nombre de questions dans un rapport remis à l'issu des TP <p>Ingénierie logicielle Suivie dans le cadre d'un vrai projet logiciel dans le cadre des projets intégrateurs</p> <p>Traitement et analyse de données : big data</p> <ul style="list-style-type: none"> - combine mise en œuvre des acquis et liberté «représentative» des cas pratiques d'analyse de données (ie. Pas de sujet « guidant l'étudiant » dans la vraie vie) - Les questions sont ouvertes, et permettent aux étudiants de manipuler précisément les concepts clés du cours, en attirant leur attention sur les objectifs pédagogiques. - La conduite d'un vrai projet permet aux étudiants d'appliquer et de confirmer leurs acquis et de confronter les difficultés liées au développement logiciel. <p>Examen écrit <input type="checkbox"/> oral <input type="checkbox"/> Rapport <input checked="" type="checkbox"/> Exposé <input type="checkbox"/> TP <input type="checkbox"/></p>		<p>Data management:</p> <ul style="list-style-type: none"> - Students will explore the dataset of their choice, and provide a short summary of interesting findings. - The students have to answer questions in a lab report <p>Software engineering Conduct a real software project</p> <p>Data management:</p> <ul style="list-style-type: none"> - It combines the necessity of leveraging acquired notions and the destabilising freedom typical of IRL data exploration scenarios. - The open questions allow the students to manipulate precisely the class key concepts, by making them focus on the objectives, - Driving a real project allows students to apply and confirm their achievements and confront the difficulties of software development. <p>Written ex. <input type="checkbox"/> Oral ex. <input type="checkbox"/> Report <input checked="" type="checkbox"/> Presentation <input type="checkbox"/> Labwork <input type="checkbox"/></p>	
Autre (préciser)		Other (please describe)	

Aides aux étudiants	Student aid

Public ciblé	Student aid
Type de formation	Formation initiale <input checked="" type="checkbox"/> Formation continue <input type="checkbox"/> Apprentissage <input type="checkbox"/> VAE <input type="checkbox"/>

Admission	Admission

Besoins particuliers	Particular needs

Langue(s) utilisée(s) pour l'enseignement :	Anglais / français
Langue(s) utilisée(s) pour le support de cours :	Anglais
Langue(s) utilisée(s) pour l'évaluation :	Français, si besoin en anglais

Mots clés :	Keywords :
Scientific data, big data, R, SOA , SOAP, REST, BPEL, ESB	

Bibliographie (auteur, titre, éditeur, année, ISBN)	Bibliography (author, title, publisher, year, ISBN)
<p>Bibliographie, webographie, polycopié...</p> <p>Il s'agit de documents accessibles aux étudiants pour mieux maîtriser votre enseignement</p> <p>Lecture slides handout</p>	

² à titre d'exemple, on peut multiplier le présentiel par un facteur fonction du type de pédagogie : 0,9 pour les CM, 0,7 pour les TD, 0,3 pour les TP, 1,5 pour les APP et autres pédagogies actives. Dans tous les cas, cette valeur doit être la plus authentique possible et s'appuyer sur des moyens appropriés pour guider le travail personnel de l'étudiant (exercices non corrigés lors des TD, préparation de TP, exercices « pour aller plus loin », grilles d'auto-évaluation, travail personnel à faire utilisant la bibliographie recommandée, ...)