

PÉRIODE D'ACCRÉDITATION : 2016 / 2021

UNIVERSITÉ PAUL SABATIER

---

# SYLLABUS MASTER

Mention Mathématiques et applications

M1 statistique et informatique décisionnelle

---

<http://www.fsi.univ-tlse3.fr/>  
[http://departement-math.univ-tlse3.fr/  
master-mention-mathematiques-et-applications-620690.kjsp](http://departement-math.univ-tlse3.fr/master-mention-mathematiques-et-applications-620690.kjsp)

2018 / 2019

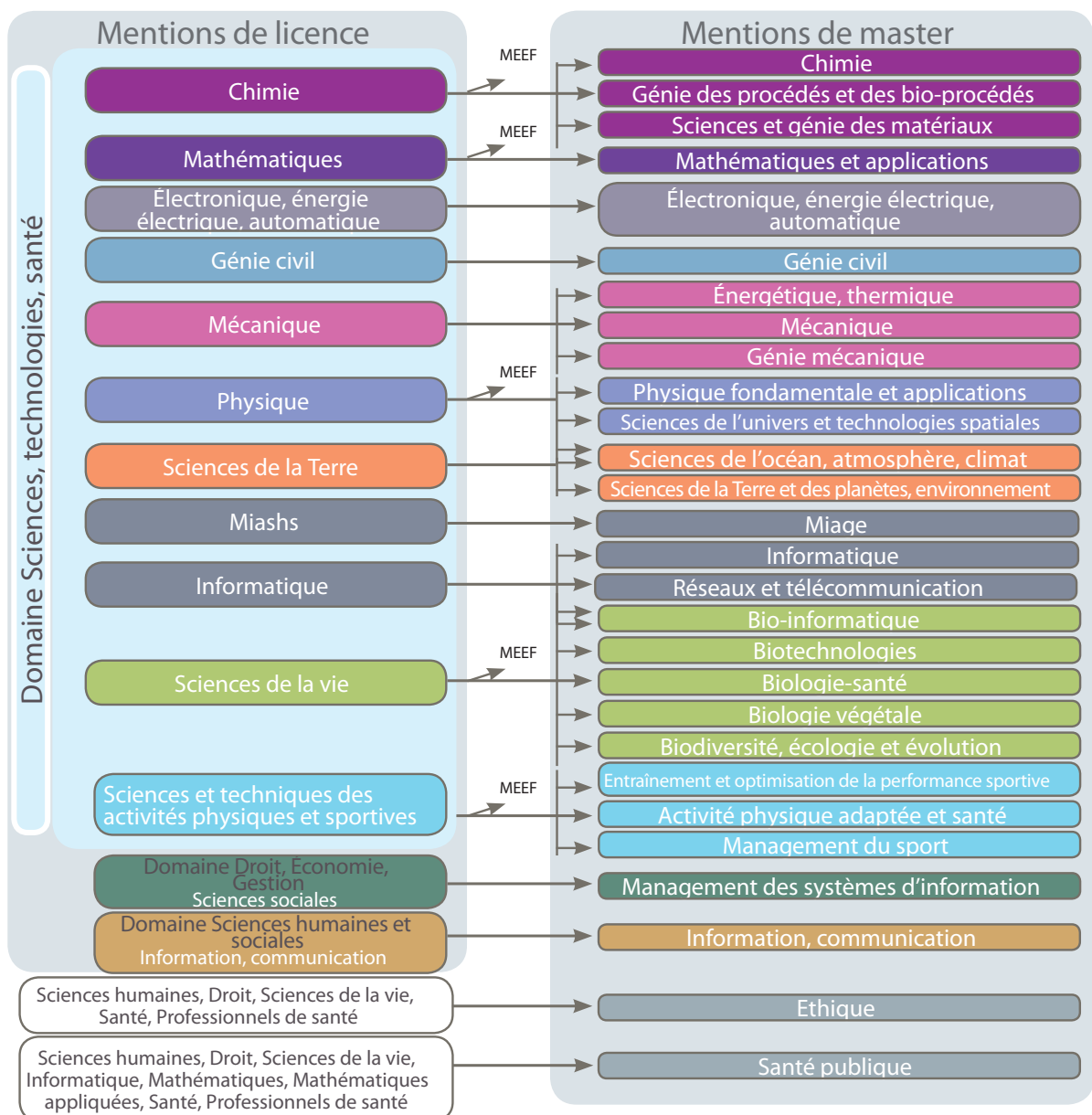
7 NOVEMBRE 2018

# SOMMAIRE

---

SCHÉMA ARTICULATION LICENCE MASTER . . . . .	3
PRÉSENTATION . . . . .	4
PRÉSENTATION DE LA MENTION ET DU PARCOURS . . . . .	4
Mention Mathématiques et applications . . . . .	4
Parcours . . . . .	4
PRÉSENTATION DE L'ANNÉE DE M1 statistique et informatique décisionnelle	4
RUBRIQUE CONTACTS . . . . .	6
CONTACTS PARCOURS . . . . .	6
CONTACTS MENTION . . . . .	6
CONTACTS DÉPARTEMENT : FSI.Math . . . . .	6
Tableau Synthétique des UE de la formation . . . . .	7
LISTE DES UE . . . . .	9
GLOSSAIRE . . . . .	32
TERMES GÉNÉRAUX . . . . .	32
TERMES ASSOCIÉS AUX DIPLOMES . . . . .	32
TERMES ASSOCIÉS AUX ENSEIGNEMENTS . . . . .	32

# SCHÉMA ARTICULATION LICENCE MASTER



# PRÉSENTATION

---

## PRÉSENTATION DE LA MENTION ET DU PARCOURS

### MENTION MATHÉMATIQUES ET APPLICATIONS

L'objectif du master mention Mathématiques et Applications est de former des mathématiciens pouvant travailler dans les métiers liés à l'**ingénierie**(parcours MApl3, SID, RO, SE, RI), à la **recherche**(parcours RI, RO, MApl3) et à l'**enseignement**(parcours ES),

Les métiers de l'ingénierie sont typiquement chefs de projets, chargés d'études, ingénieurs et chercheurs dans des secteurs d'activités tels que l'industrie, les services, le marketing.

Les métiers de l'enseignement concernent des postes de professeur de mathématiques en lycée, à l'université en passant par les classes préparatoires.

La recherche peut-être de nature académique, théorique et/ou appliquée, ou être tournée vers l'innovation et le développement dans le secteur privé.

Que ce soit pour les métiers de l'ingénierie, de l'enseignement ou de la recherche le nombre d'étudiants formés aux mathématiques en France est très inférieur au nombre de postes à pourvoir. De ce fait, l'insertion des étudiants titulaires d'un master en Mathématiques est excellente

### PARCOURS

Le parcours SID-Big Data (Statistique et Informatique Décisionnelle), labélisé CMI (Cursus de Master en Ingénierie) est une formation résolument tournée vers les métiers de l'ingénierie des données, qui en aborde tous les aspects : de la collecte à l'exploitation statistique et au machine learning en passant par le stockage et la gestion de bases de données massives et distribuées de type big data.

Il permet d'acquérir une double compétence en Statistique et en Informatique décisionnelle, complétée par une connaissance de l'entreprise et une pratique professionnelle acquise au travers des projets et de plusieurs stages. Cette double compétence, à l'ère du Big Data, est très appréciée sur le marché de l'emploi.

Les débouchés de la formation SID-Big Data (Licence et Master) se trouvent dans tous les secteurs où la mise en place, l'administration, l'exploitation et la valorisation de l'information sont primordiales pour la prise de décision.

## PRÉSENTATION DE L'ANNÉE DE M1 STATISTIQUE ET INFORMATIQUE DÉCISIONNELLE

Une fois acquis les concepts fondamentaux en Mathématiques/Statistique et Informatique/Base de données en L3 SID, les enseignements du Master SID-Big Data sont axés sur le perfectionnement et la spécialisation de méthodes plus avancées, avec un juste équilibre entre informatique/bases de données et mathématiques/statistique. En Master 1, les méthodes classiques dans le domaine des sciences des données sont approfondies et des méthodes plus novatrices, orientée vers la big data sont introduites. En Master 2, toutes ces compétences sont renforcées pour que les étudiants puissent maîtriser de bout en bout les différents aspects de leur futur métier dans le domaine des sciences des données.

Les principales thématiques abordées en Master SID-Big Data :

- Statistique exploratoire
- Modélisation aléatoire
- Séries chronologiques, sondages
- Apprentissage / Big data
- Exploitation et analyse de données multi-support
- Data mining / Text mining
- Optimisation numérique

- Base de données avancées
- Business intelligence

Au cours des deux années du Master SID-Big Data, en parallèle des enseignements classiques (cours, TD et TP), de nombreux projets sont mis en place, balayant l'ensemble des compétences enseignées.

En M1, un stage de 3-4 mois est obligatoire au 2nd semestre.

Les étudiants peuvent suivre le M2 SID-Big data selon deux formules :

- en formation initiale, sur un mode d'organisation classique : cours, TD, TP, projets, et enfin un stage de 5-6 mois couvrant tout le 2nd semestre.
- en alternance, avec des périodes (de 2 à 4 semaines) à l'université (avec les étudiants en formation initiale) ou en entreprise.

# RUBRIQUE CONTACTS

---

## CONTACTS PARCOURS

### RESPONSABLE M1 STATISTIQUE ET INFORMATIQUE DÉCISIONNELLE

CHHAIBI Reda

Email : [reda.chhaibi@math.univ-toulouse.fr](mailto:reda.chhaibi@math.univ-toulouse.fr)

MOKADEM Riad

Email : [riad.mokadem@irit.fr](mailto:riad.mokadem@irit.fr)

CHOUQUET Cécile

Email : [cecile.chouquet@math.univ-toulouse.fr](mailto:cecile.chouquet@math.univ-toulouse.fr)

Téléphone : 05.61.55.69.84

### SECRÉTAIRE PÉDAGOGIQUE

NICOLAS Clement

Email : [clement.nicolas2@univ-tlse3.fr](mailto:clement.nicolas2@univ-tlse3.fr)

## CONTACTS MENTION

### RESPONSABLE DE MENTION MATHÉMATIQUES ET APPLICATIONS

BARTHE Franck

Email : [barthe@math.univ-toulouse.fr](mailto:barthe@math.univ-toulouse.fr)

Téléphone : 05 61 55 82 06

VIGNAL Marie-Hélène

Email : [mhvignal@math.univ-toulouse.fr](mailto:mhvignal@math.univ-toulouse.fr)

Téléphone : 05.61.55.76.34

## CONTACTS DÉPARTEMENT: FSI.MATH

### DIRECTEUR DU DÉPARTEMENT

GARIVIER Aurélien

Email :

### SECRETARIAT DU DÉPARTEMENT

RODRIGUES Manuella

Email :

Téléphone : 05 61 55 73 54

Université Paul Sabatier

1TP1, bureau B13

118 route de Narbonne

31062 TOULOUSE cedex 9

# TABLEAU SYNTHÉTIQUE DES UE DE LA FORMATION

page	Code	Intitulé UE	ECTS	Obligatoire Facultatif	Cours	TD	TP	Projet	Stage
<b>Premier semestre</b>									
10	EMMAB1AM	ALGÈBRE LINÉAIRE	3	O	12	18			
11	EMMAB1BM	MATHÉMATIQUES POUR L'INGÉNIEURIE	3	O	12	12	6		
12	EMMAB1CM	STATISTIQUE EXPLORATOIRE	3	O	12	9	9		
13	EMMAB1DM	SÉRIES CHRONOLOGIQUES ET SONDAGES	3	O	12	9	9		
14	EMMAB1EM	APPRENTISSAGE ET COMPLEXITÉ	3	O	12	9	9		
15	EMMAB1FM	DATA WAREHOUSE ET OLAP	3	O	12	9	9		
16	EMMAB1GM	TRANSFORMATIONS DE MODÈLES	3	O	12	9	9		
17	EMMAB1HM	GÉNIE LOGICIEL, TRAVAIL EN GROUPE	3	O	12	18			
18	EMMAB1IM	MARKETING, TECHNIQUES DE COMMUNICATION, SÉCURITÉ, ASPECTS JURIDIQUES	3	O	12	18			
19	EMMAB1JM	EXTRACTION D'INFORMATION DANS LES DOCUMENTS TEXTES, AUDIO, VIDÉO	3	O	9	9	14		
20	EMMAB1TM	STAGE FACULTATIF	3	F					0,5
<b>Second semestre</b>									
21	EMMAB2AM	OPTIMISATION	3	O	12	9	9		
22	EMMAB2BM	MODÉLISATION ALÉATOIRE	3	O	12	6	12		
23	EMMAB2CM	ALGORITHMES DE CLASSIFICATION, DATA MINING ET TEXT MINING	3	O	12	6	12		
24	EMMAB2DM	PRINCIPES ET MÉTHODES DES SYSTÈMES DE BASES DE DONNÉES RÉPARTIES	3	O	12	12	6		
25	EMMAB2EM	PROJET	3	O				100	
<b>Choisir 1 UE parmi les 2 UE suivantes :</b>									
26	EMMAB2GM	STAGE EN ENTREPRISE	12	O					3
27	EMMAB2HM	STAGE EN ENTREPRISE	12	O					3
<b>Choisir 1 UE parmi les 4 UE suivantes :</b>									

page	Code	Intitulé UE	ECTS	Obligatoire Facultatif	Cours	TD	TP	Projet	Stage
28	EMMAB2VM	ANGLAIS	3	O		24			
29	EMMAB2WM	ALLEMAND	3	O		24			
30	EMMAB2XM	ESPAGNOL	3	O		24			
31	EMMAB2YM	FRANÇAIS GRANDS DÉBUTANTS	3	O		24			



---

## LISTE DES UE

---

<b>UE</b>	<b>ALGÈBRE LINÉAIRE</b>	<b>3 ECTS</b>	<b>1<sup>er</sup> semestre</b>
<b>EMMAB1AM</b>	Cours : 12h , TD : 18h		

### ENSEIGNANT(E) RESPONSABLE

SERRURIER Mathieu

Email : [Mathieu.Serrurier@irit.fr](mailto:Mathieu.Serrurier@irit.fr)

TAPIA Joseph

Email : [joseph.tapia@math.univ-toulouse.fr](mailto:joseph.tapia@math.univ-toulouse.fr)

Téléphone : (poste) 63.76

### OBJECTIFS D'APPRENTISSAGE

Apprendre à manipuler des formes bilinéaires, à construire des bases orthonormées, à travailler avec des normes. Effectuer une DVS

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Algèbre bilinéaire : matrices symétriques, formes bilinéaires et produits scalaires, bases orthonormées et procédé de Gram-Schmidt. Décomposition en valeurs singulières.

### PRÉ-REQUIS

algèbre linéaire élémentaire : espaces vectoriels, bases, applications linéaires, représentations matricielles, changements de base. Valeurs et vecteurs propres,

### RÉFÉRENCES BIBLIOGRAPHIQUES

S. Lipschutz, M. Lipson "Algèbre Linéaire", ed. Schaum's

### MOTS-CLÉS

produit scalaire, norme, Gram-Schmidt, DVS

<b>UE</b>	<b>MATHÉMATIQUES POUR L'INGÉNIERIE</b>	<b>3 ECTS</b>	<b>1<sup>er</sup> semestre</b>
<b>EMMAB1BM</b>	Cours : 12h , TD : 12h , TP : 6h		

### ENSEIGNANT(E) RESPONSABLE

PAUWELS Edouard

Email : [edouard.pauwels@irit.fr](mailto:edouard.pauwels@irit.fr)

### OBJECTIFS D'APPRENTISSAGE

Présenter des outils mathématiques pour la modélisation et la résolution de problèmes d'optimisation.

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

- 1) Calcul dans  $\mathbb{R}^n$  : généralités sur les fonctions de plusieurs variables, courbes de niveaux, gradient, hessienne, points singulier
- 2) Formules de Taylor à l'ordre  $n$ , développements limités, applications
- 3) Intégrales multiples : formules de changement de variables, théorème de Fubini. Applications aux probabilités.
- 4) Analyse de Fourier : introduction aux séries de Fourier, applications.

### PRÉ-REQUIS

base d'algèbre linéaire et d'analyse (calcul matriciel, calcul différentiel et calcul d'intégrales simples)

### RÉFÉRENCES BIBLIOGRAPHIQUES

Mathématiques pour les sciences de l'ingénieur, F. Bertrand, M. Maumy-Bertrand, S. Ferrigno, A. Muller-Gueudin, D. Marx, Edition Dunod

### MOTS-CLÉS

Intégrales multiples, développement limités dans  $\mathbb{R}^n$ , calcul différentiel, séries de Fourier

<b>UE</b>	<b>STATISTIQUE EXPLORATOIRE</b>	<b>3 ECTS</b>	<b>1<sup>er</sup> semestre</b>
<b>EMMAB1CM</b>	Cours : 12h , TD : 9h , TP : 9h		

### ENSEIGNANT(E) RESPONSABLE

CHOUQUET Cécile

Email : [cecile.chouquet@math.univ-toulouse.fr](mailto:cecile.chouquet@math.univ-toulouse.fr)

Téléphone : 05.61.55.69.84

### OBJECTIFS D'APPRENTISSAGE

L'objectif de ce module est de donner une introduction aux méthodes de statistique descriptive multidimensionnelle. On étudiera les méthodes factorielles classiques basées sur la projection des données sur des espaces propres associées à la diagonalisation de matrices estimées sur les données. Les méthodes de classification supervisée ou non supervisée feront l'objet de la dernière partie.

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Norms on vector spaces, orthogonal base changes and projections. Decomposition in singular values. Base notions of statistics (quantitative or qualitative variables, Gaussian distribution).

### PRÉ-REQUIS

Changement de bases orthogonales et projections. Decomposition en valeurs singulières. Notions de bases en statistiques (loi Gaussienne)

### RÉFÉRENCES BIBLIOGRAPHIQUES

Lebart, Morineau, Piron Statistique exploratoire multidimensionnelle

### MOTS-CLÉS

analyse multidimensionnelle, classification

<b>UE</b>	<b>SÉRIES CHRONOLOGIQUES ET SONDAGES</b>	<b>3 ECTS</b>	<b>1<sup>er</sup> semestre</b>
<b>EMMAB1DM</b>	Cours : 12h , TD : 9h , TP : 9h		

### ENSEIGNANT(E) RESPONSABLE

DUPUIS Jérôme

Email : [jerome.dupuis@math.univ-toulouse.fr](mailto:jerome.dupuis@math.univ-toulouse.fr)

Téléphone : (poste) 6984

### OBJECTIFS D'APPRENTISSAGE

Ce module a pour objectif d'introduire des techniques de modélisation statistique au travers de l'étude de séries chronologiques. Dans un premier temps, on présentera l'analyse déterministe des séries chronologiques et de leur prédiction (régression linéaire, lissage exponentiel. . .)

Dans un second temps, quelques modèles classiques de séries chronologiques stationnaires seront présentés (ARMA, GARCH. . .). L'étude de tels processus et de leurs indices caractéristiques permettra de présenter les techniques de prédiction.

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Analyse déterministe des séries chronologiques et de leur prédiction (régression linéaire, lissage exponentiel)

Modèles classiques de séries chronologiques stationnaires (ARMA, GARCH, ...) et présentation des techniques de prédiction

Bases de la théorie des sondages

### PRÉ-REQUIS

bases des probabilités et de la statistique, séries entières

### RÉFÉRENCES BIBLIOGRAPHIQUES

P. Brockwell and R. Davis. Time Series

### MOTS-CLÉS

séries chronologiques et sondages

<b>UE</b>	<b>APPRENTISSAGE ET COMPLEXITÉ</b>	<b>3 ECTS</b>	<b>1<sup>er</sup> semestre</b>
<b>EMMAB1EM</b>	Cours : 12h , TD : 9h , TP : 9h		

### ENSEIGNANT(E) RESPONSABLE

LOUBES Jean-Michel

Email : [loubes@math.univ-toulouse.fr](mailto:loubes@math.univ-toulouse.fr)

Téléphone : 05 61 55 85 73

### OBJECTIFS D'APPRENTISSAGE

Comprendre la notion d'apprentissage statistique supervisé et la notion de modèle complexe

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

\* Classification supervisée : définition de l'erreur d'ajustement et erreur de généralisation

\* Notion de sur-ajustement et sparsité

\* théorie de Vapnik-Chervonenkis

et entropie\*Pénalités liées à la complexité du modèle

\*Extension au cadre de la régression et sélection de modèle

### PRÉ-REQUIS

Probabilités

### RÉFÉRENCES BIBLIOGRAPHIQUES

\*Statistical Learning (Hastie & Tibshirani)

### MOTS-CLÉS

Apprentissage et complexité

<b>UE</b>	<b>DATA WAREHOUSE ET OLAP</b>	<b>3 ECTS</b>	<b>1<sup>er</sup> semestre</b>
<b>EMMAB1FM</b>	Cours : 12h , TD : 9h , TP : 9h		

### ENSEIGNANT(E) RESPONSABLE

HUBERT Gilles  
 Email : [hubert@irit.fr](mailto:hubert@irit.fr)

### OBJECTIFS D'APPRENTISSAGE

Concevoir un système décisionnel à base d'entrepôts et magasins de données

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Marché du décisionnel/Outils.  
 Problématique des Entrepôts de données.  
 Entrepôts de données et Magasins de données.  
 Modélisation multidimensionnelle.  
 Modélisation logique R-OLAP.  
 Modélisation physique.  
 Conception d'une solution décisionnelle

### PRÉ-REQUIS

Modèle conceptuel de données, Bases de données relationnelles, SQL, algorithmique et programmation

### RÉFÉRENCES BIBLIOGRAPHIQUES

The Data Warehouse Toolkit, R. Kimball and M. Ross.  
 OLAP Solutions : Building Multidimensional Information Systems, E. Thomsen.

### MOTS-CLÉS

Entrepôts de données, magasins, OLAP, Solution décisionnelle

<b>UE</b>	<b>TRANSFORMATIONS DE MODÈLES</b>	<b>3 ECTS</b>	<b>1<sup>er</sup> semestre</b>
<b>EMMAB1GM</b>	Cours : 12h , TD : 9h , TP : 9h		

### ENSEIGNANT(E) RESPONSABLE

BAHSOUN Wahiba

Email : [wbahsoun@irit.fr](mailto:wbahsoun@irit.fr)

Téléphone : 05 61 55 69 45

### OBJECTIFS D'APPRENTISSAGE

Maîtriser la conception objet UML.

Concevoir, mettre en œuvre et manipuler une base de données orientée objet.

Concevoir, mettre en œuvre et manipuler une base de données semi-structurées (json, XML).

Migrer une solution d'un modèle vers un autre.

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Concepts UML, Conception UML, de UML à SQL, Concepts des bases de données objet, Manipulation de base de données objet-relationnelle, Transformation d'une base relationnelle en base objet, Concepts des bases de données semi-structurées (XML, DTD...)

### PRÉ-REQUIS

Modèle Entité-association,.

Bases de données relationnelles

### MOTS-CLÉS

Transformation. UML. Bases de données relationnelles. Bases de données semi-structurées



<b>UE</b>	<b>GÉNIE LOGICIEL, TRAVAIL EN GROUPE</b>	<b>3 ECTS</b>	<b>1<sup>er</sup> semestre</b>
<b>EMMAB1HM</b>	Cours : 12h , TD : 18h		

### ENSEIGNANT(E) RESPONSABLE

BAHSOUN Wahiba

Email : [wbahsoun@irit.fr](mailto:wbahsoun@irit.fr)

Téléphone : 05 61 55 69 45

CHAUDET Christelle

Email : [Christelle.Chaudet@irit.fr](mailto:Christelle.Chaudet@irit.fr)

Téléphone : (poste) 84.56

### OBJECTIFS D'APPRENTISSAGE

Modèles de developpemen logiciel.

Langage de Modelisation UML.

Tests unitaires.

Métriques.

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

1. Ingénierie du Logiciel : Processus de développement ( Méthodes Agiles : RUP, Scrum, Open Up, collaboratif...). Gestion de Projet : différentes méthodes de modélisation (Gantt, Pert... ), Gestion de configuration, Intégration continue, Referentiels, Normes. Gestion des Risques . Définition des Exigences. Métrologie : Mesures et Diagnostic Qualité Logiciel ( diagramme de Kiviat...).
2. Qualité du logiciel :Langage de modélisation UML et programmation objet : Les , oncepts du langage UML et les règles applicables, Les principaux diagrammes utilisés.....Navigation dans un modèle.
3. Implémentation de la modélisation. Diagramme de séquence détaillée, classes participantes, implémentationn , métriques.

### RÉFÉRENCES BIBLIOGRAPHIQUES

Sotware Metrics Norman E Fenton, Software Engineering In Higher Education

### MOTS-CLÉS

Génie logiciel.

Gestion de projets.

<b>UE</b>	<b>MARKETING, TECHNIQUES DE COMMUNICATION, SÉCURITÉ, ASPECTS JURIDIQUES</b>	<b>3 ECTS</b>	<b>1<sup>er</sup> semestre</b>
<b>EMMAB1IM</b>	Cours : 12h , TD : 18h		

### ENSEIGNANT(E) RESPONSABLE

MOKADEM Riad

Email : [riad.mokadem@irit.fr](mailto:riad.mokadem@irit.fr)

### OBJECTIFS D'APPRENTISSAGE

L'objectif de cette UE est de sensibiliser, au delà des compétences techniques sur la gestion des données, aux questions afférentes : sécurité des données, problématiques juridiques, valorisation marketing.

En outre, les cours de communication poursuivent le développement chez les étudiants d'un savoir-être interpersonnel, en se focalisant progressivement sur la gestion de projet et le management d'équipe

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Communication : développement personnel : découverte de soi, gestion du temps, styles de management, management et animation d'une équipe, conduite de réunion.

Les étudiants devront aussi être capables de valoriser leurs talents et potentiels de travail lors d'un bilan de compétences, d'une réponse à des offres de stage ou d'emploi...

Sécurité : l'étude des problèmes liés à l'intégrité des bases de données (BD) actives puis décrire les mécanismes que fournit un système de gestion de BD (SGBD) pour assurer la cohérence des données ; étudier les protocoles d'accès à la BD permettant d'assurer la confidentialité des données selon un schéma défini dans les applications métiers, apprendre à écrire des programmes d'application qui garantissent la cohérence de la BD dans un environnement concurrentiel, étudier les mécanismes qui permettent la reprise de fonctionnement normal d'une BD après la survenue d'une panne. Les aspects juridiques relatifs aux données seront également évoqués à travers des conférences d'intervenants spécialistes de la thématique

### PRÉ-REQUIS

Communication : le cv vu en L3

### RÉFÉRENCES BIBLIOGRAPHIQUES

IT Governance : An International Guide to Data Security and ISO27001 / ISO27002 - Alan Calder et Steve Watkins

Traitement des données personnelles, guide juridique - de Fabrice Mattatia

### MOTS-CLÉS

Communication, marketing, sécurité, droit informatique

<b>UE</b>	<b>EXTRACTION D'INFORMATION DANS LES DOCUMENTS TEXTES, AUDIO, VIDÉO</b>	<b>3 ECTS</b>	<b>1<sup>er</sup> semestre</b>
<b>EMMAB1JM</b>	Cours : 9h , TD : 9h , TP : 14h		

### ENSEIGNANT(E) RESPONSABLE

MORENO José

Email : [jose.moreno@irit.fr](mailto:jose.moreno@irit.fr)

Téléphone : 05 61 55 63 22

### OBJECTIFS D'APPRENTISSAGE

Extraire à partir de tout support d'information (texte, image, son, vidéo), les caractéristiques de contenu et de forme

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

I. Objectifs, enjeux et applications.

II. Extraction d'informations à partir de textes : Préliminaires (Typologie des tâches d'extraction (entités nommées, relations, indices temporels, lexique...)) • Domaines d'applications • Systèmes d'extraction de l'information : architecture de base et technologies.

III. Principes et méthodes : Méthodes symboliques (expressions régulières, règles et patrons) • Méthodes supervisées (CRF, arbres de décisions),

IV. Extraction d'information de type audio : Présentation du signal sonore, Numérisation, Extraction de paramètres temporels et fréquentiels,

V. Extraction d'information de type image et vidéo : Présentation de l'image numérique, Quantification, Extraction de paramètres liés à la couleur,

VI. Évaluation : Méthodologie et métriques, Campagnes d'évaluation.

### PRÉ-REQUIS

Méthodes de classification

### RÉFÉRENCES BIBLIOGRAPHIQUES

1. Information extraction in a retrieval context : algorithms & prospects, Springer, M. F. Moens
2. Reconnaissance automatique de la parole : Du signal à son interprétation, Dunod, Haton
3. Image numérique couleur. Dunod, Trémeau

### MOTS-CLÉS

Extraction de paramètres, entité nommées, relations sur le texte, traitement du signal (audio et video)

<b>UE</b>	<b>STAGE FACULTATIF</b>	<b>3 ECTS</b>	<b>1<sup>er</sup> semestre</b>
<b>EMMAB1TM</b>	Stage : 0,5 mois minimum		

<b>UE</b>	<b>OPTIMISATION</b>	<b>3 ECTS</b>	<b>2<sup>nd</sup> semestre</b>
<b>EMMAB2AM</b>	Cours : 12h , TD : 9h , TP : 9h		

### ENSEIGNANT(E) RESPONSABLE

JAN Sophie

Email : [sophie.jan@math.univ-toulouse.fr](mailto:sophie.jan@math.univ-toulouse.fr)

Téléphone : (poste) 76.58, (dom)  
05.61.81.82.91

### OBJECTIFS D'APPRENTISSAGE

Je caractérise un minimum avec et sans contrainte dans un problème multivarié.

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Introduction à l'analyse convexe et optimisation, schémas numériques

### PRÉ-REQUIS

Algèbre linéaire, calcul différentiel

### RÉFÉRENCES BIBLIOGRAPHIQUES

Bierlaire M., Introduction à l'optimisation différentiable, Presse Polytechniques et Universitaires Romandes, 2006

### MOTS-CLÉS

Optimisation, analyse convexe, schémas numériques

<b>UE</b>	<b>MODÉLISATION ALÉATOIRE</b>	<b>3 ECTS</b>	<b>2<sup>nd</sup> semestre</b>
<b>EMMAB2BM</b>	Cours : 12h , TD : 6h , TP : 12h		

### ENSEIGNANT(E) RESPONSABLE

AZAÏS Jean-Marc

Email : [jean-marc.azais@math.univ-toulouse.fr](mailto:jean-marc.azais@math.univ-toulouse.fr)

Téléphone : 05 61 55 60 07

GAMBOA Fabrice

Email : [fabrice.gamboa@math.univ-toulouse.fr](mailto:fabrice.gamboa@math.univ-toulouse.fr)

Téléphone : 0561557276

### OBJECTIFS D'APPRENTISSAGE

- Découvrir plusieurs modèles statistiques utiles en présence de variables explicatives ou d'un processus temporel.
- Se confronter à plusieurs cas concrets et apprendre à les modéliser à l'aide de l'un de ces différents modèles.

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

1. Modèle linéaire généralisé :

- régression logistique
- modèle log-linéaire, régression de Poisson

2. Modélisation des processus :

- chaînes de Markov
- files d'attente

Des exemples concrets sont fournis pour chaque notion (suivis médicaux, etc.)

### PRÉ-REQUIS

modèle linéaire, maximum de vraisemblance, variables aléatoires discrètes et continues, indépendance, espérance conditionnelle, probabilité conditionnelle.

### RÉFÉRENCES BIBLIOGRAPHIQUES

J.J. Droesbeke, M. Lejeune, G. Sapota ; Modèles statistiques pour données qualitatives, 2005.

M. Tenenhaus ; Statistique : Méthodes pour décrire, expliquer et prévoir, 2007.

S. M. Ross ; Initiation aux probabilités, 2014.

### MOTS-CLÉS

modèle linéaire généralisé, processus stochastique

<b>UE</b>	<b>ALGORITHMES DE CLASSIFICATION, DATA MINING ET TEXT MINING</b>	<b>3 ECTS</b>	<b>2<sup>nd</sup> semestre</b>
<b>EMMAB2CM</b>	Cours : 12h , TD : 6h , TP : 12h		

### ENSEIGNANT(E) RESPONSABLE

MORENO José

Email : [jose.moreno@irit.fr](mailto:jose.moreno@irit.fr)

Téléphone : 05 61 55 63 22

PITARCH Yoann

Email : [yoann.pitarch@irit.fr](mailto:yoann.pitarch@irit.fr)

Téléphone : 05 61 55 6323

### OBJECTIFS D'APPRENTISSAGE

L'objectif de ce cours est double. Dans un premier temps, il s'agira de présenter les méthodes les plus populaires de classification supervisée. Ces méthodes introduites, nous étudierons comment ces classifieurs peuvent être combinés pour accroître la qualité de la prédiction. La seconde partie du cours sera consacrée à l'adaptation des techniques de machine learning étudiées jusqu'ici à un contexte de plus en plus populaire : le Web. Nous verrons que l'analyse de données Web regroupe 3 aspects bien différents (l'analyse du Web en tant que réseau, l'analyse de son contenu et l'analyse de ses usages) et que chacun oblige à repenser en partie les approches classiques d'apprentissage.

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Classification

1- Algorithmes de classification supervisée : regression logistique, SVM, methode des plus proches voisins, reseaux de neurones, arbres de decision.

2- Agrégation de classifieurs combinaison d'experts (super-learning), bagging, boosting, random forest.

3- Structure de graphes. Notions de centres de graphes, de clusters de graphes et de plus courts chemins), Web Mining (a- Fouille de structure : caractérisation d'un noeud dans un graphe (prestige, centralité, popularité) ; détection de communauté, extractions de sous-graphes sous contraintes, b- Fouille de contenu : classification de données textuelles, fouille d'opinion, analyse de sentiments, détection de contenu spam, c- Fouille d'usage : identification de sessions utilisateurs, clustering d'utilisateurs/de documents, détection de fraudes)

### PRÉ-REQUIS

Introduction au Machine Learning (L3)

Bases de la théorie des graphes

### RÉFÉRENCES BIBLIOGRAPHIQUES

The elements of statistical learning (Hastie Tibshirani) Web Data Mining : Exploring Hyperlinks, Contents, and Usage Data (Data-Centric Systems and Applications) par Bing Liu

### MOTS-CLÉS

Apprentissage statistique, machine learning, apprentissage supervisé, classification, régression, validation croisée, fouille du Web.

<b>UE</b>	<b>PRINCIPES ET MÉTHODES DES SYSTÈMES DE BASES DE DONNÉES RÉPARTIES</b>	<b>3 ECTS</b>	<b>2<sup>nd</sup> semestre</b>
<b>EMMAB2DM</b>	Cours : 12h , TD : 12h , TP : 6h		

### ENSEIGNANT(E) RESPONSABLE

HAMEURLAIN Abdelkader  
 Email : [hameurlain@irit.fr](mailto:hameurlain@irit.fr)

### OBJECTIFS D'APPRENTISSAGE

Présenter les principaux problèmes posés et d'introduire les méthodes proposées dans la conception et le développement des SGBD répartis.

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

1. Introduction aux BD réparties BDR.
2. Principales fonctions des SGBD répartis.
3. Architecture logiciel des SGBDR.
4. Approches et stratégies de conception d'une BDR.
5. Principe d'évaluation de requêtes réparties.

### PRÉ-REQUIS

Systèmes de bases de données relationnels

### RÉFÉRENCES BIBLIOGRAPHIQUES

Principles of Distributed Database Systems, February 2011. M Tamer Ozsu et Patick Valduriez, Springer-Verlag, 3rd, ed. 2011, ISBN-13 : 978-1441988331

### MOTS-CLÉS

BD distribuées, Fragmentation, Localisation, Traitement des requêtes réparties



<b>UE</b>	<b>PROJET</b>	<b>3 ECTS</b>	<b>2<sup>nd</sup> semestre</b>
<b>EMMAB2EM</b>	Projet : 100h		

### ENSEIGNANT(E) RESPONSABLE

MOKADEM Riad

Email : [riad.mokadem@irit.fr](mailto:riad.mokadem@irit.fr)

### OBJECTIFS D'APPRENTISSAGE

Gestion de projets et évaluation d'une compilation d'enseignements : de la modélisation d'un système d'information vers la visualisation de données

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Analyse des besoins.

Modélisation d'un système d'information.

Conception et mise en œuvre d'un système d'aide à la décision permettant de faire des analyses détaillées à partir des données issues d'une part, des bases de Données et d'autre part, de source d'informations hétérogènes.

Visualisation des données via le web.

Gestion de projet.

### PRÉ-REQUIS

Modèle conceptuel de données, interrogation de données, algorithmique et programmation

### RÉFÉRENCES BIBLIOGRAPHIQUES

Software Engineering In Higher Education

### MOTS-CLÉS

Gestion de projet, Modélisation, Interrogation et Visualisation de données

<b>UE</b>	<b>STAGE EN ENTREPRISE</b>	<b>12 ECTS</b>	<b>2<sup>nd</sup> semestre</b>
<b>EMMAB2GM</b>	Stage : 3 mois minimum		

### ENSEIGNANT(E) RESPONSABLE

MOKADEM Riad

Email : [riad.mokadem@irit.fr](mailto:riad.mokadem@irit.fr)

### OBJECTIFS D'APPRENTISSAGE

L'objectif est d'acquérir les compétences professionnelles aussi bien dans le domaine de la statistique que l'informatique, et d'appliquer les différentes méthodes dans tous les secteurs. Ce stage sera également l'occasion pour l'étudiant d'acquérir une certaine autonomie relative à son niveau d'études.

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Mise en place, administration et exploitation de l'information pour la prise de décision.

### PRÉ-REQUIS

Concepts fondamentaux en : programmation, bases de données et recherche d'informations.

### MOTS-CLÉS

Statistique, Informatique

<b>UE</b>	<b>STAGE EN ENTREPRISE</b>	<b>12 ECTS</b>	<b>2<sup>nd</sup> semestre</b>
<b>EMMAB2HM</b>	Stage : 3 mois minimum		

### ENSEIGNANT(E) RESPONSABLE

CHHAIBI Reda

Email : [reda.chhaibi@math.univ-toulouse.fr](mailto:reda.chhaibi@math.univ-toulouse.fr)

MOKADEM Riad

Email : [riad.mokadem@irit.fr](mailto:riad.mokadem@irit.fr)

### OBJECTIFS D'APPRENTISSAGE

Je sais mettre en oeuvre les connaissances et compétences acquises jusqu'en Master 1 pour répondre à différentes missions en science des données dans un contexte applicatif industriel ou recherche.

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Missions dans un contexte applicatif industriel ou recherche.

### PRÉ-REQUIS

Ensemble de compétences mathématiques et informatiques acquises jusqu'en M1 dépendant du contexte applicatif.

### MOTS-CLÉS

Mission, cadre applicatif, industrie, recherche

<b>UE</b>	<b>ANGLAIS</b>	<b>3 ECTS</b>	<b>2<sup>nd</sup> semestre</b>
<b>EMMAB2VM</b>	TD : 24h		

### ENSEIGNANT(E) RESPONSABLE

CHAPLIER Claire

Email : [claire.chaplier@univ-tlse3.fr](mailto:claire.chaplier@univ-tlse3.fr)

### OBJECTIFS D'APPRENTISSAGE

#### Niveau C1 du CECRL (Cadre Européen de Certification en Langues)

Développer les compétences indispensables aux étudiant/es en vue de leur intégration dans la vie professionnelle. Perfectionner les outils de communication permettant de s'exprimer dans le contexte international d'aujourd'hui et acquérir l'autonomie linguistique nécessaire à cette intégration.

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

- Outils de communication de l'écriture scientifique (dossier du projet, synthèse, compte-rendu (sur CO), abstract, rédaction, résumé...)
- Outils d'expression permettant de maîtriser une présentation orale ou une discussion critique dans le domaine scientifique (rhétorique, éléments linguistiques)
- Eléments de maîtrise d'un projet sur une thématique spécifique dans le domaine de spécialité
- Structures communicatives et linguistiques utilisées dans le cadre d'une simulation de tâche professionnelle

#### Compétences

Compréhension orale - Expression écrite - Expression orale - Compréhension écrite

- Savoir communiquer en anglais scientifique
- Savoir repérer les éléments constitutifs d'une communication écrite ou orale dans le domaine de spécialité
- Savoir prendre la parole en public (conférence ou réunion) dans le cadre d'un colloque, projet de recherche, projet professionnel

### PRÉ-REQUIS

N/A

### RÉFÉRENCES BIBLIOGRAPHIQUES

N/A

### MOTS-CLÉS

=12.0ptProjet - Repérer - Rédaction anglais scientifique - style - registre - critique - professionnel

<b>UE</b>	<b>ALLEMAND</b>	<b>3 ECTS</b>	<b>2<sup>nd</sup> semestre</b>
<b>EMMAB2WM</b>	TD : 24h		

### ENSEIGNANT(E) RESPONSABLE

SANTAMARINA Diego

Email : [diego.santamarina@univ-tlse3.fr](mailto:diego.santamarina@univ-tlse3.fr)

Téléphone : 05 61 55 64 27

### PRÉ-REQUIS

Niveau B2 en anglais

<b>UE</b>	<b>ESPAGNOL</b>	<b>3 ECTS</b>	<b>2<sup>nd</sup> semestre</b>
<b>EMMAB2XM</b>	TD : 24h		

### ENSEIGNANT(E) RESPONSABLE

SANTAMARINA Diego

Email : [diego.santamarina@univ-tlse3.fr](mailto:diego.santamarina@univ-tlse3.fr)

Téléphone : 05 61 55 64 27

### OBJECTIFS D'APPRENTISSAGE

Etre capable de travailler en milieu hispanophone ou avec des partenaires hispanophones

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Activités langagières permettant la maîtrise de l'espagnol général et de la langue de spécialité

### PRÉ-REQUIS

Niveau B2 en anglais - Pas de pré-requis particulier en espagnolEspagnol professionnel, le cours prend en compte les différents niveaux

### MOTS-CLÉS

Espagnol professionnel

<b>UE</b>	<b>FRANÇAIS GRANDS DÉBUTANTS</b>	<b>3 ECTS</b>	<b>2<sup>nd</sup> semestre</b>
<b>EMMAB2YM</b>	TD : 24h		

### ENSEIGNANT(E) RESPONSABLE

JASANI Isabelle

Email : [leena.jasani@wanadoo.fr](mailto:leena.jasani@wanadoo.fr)

Téléphone : 65.29

### OBJECTIFS D'APPRENTISSAGE

Cette UE est conseillée aux étudiants ayant un niveau très faible en français

### PRÉ-REQUIS

Niveau B2 en anglais

### MOTS-CLÉS

français scientifique

# GLOSSAIRE

---

## TERMES GÉNÉRAUX

### DÉPARTEMENT

Les départements d'enseignement sont des structures d'animation pédagogique internes aux composantes (ou facultés) qui regroupent les enseignants intervenant dans une ou plusieurs mentions

### UE : UNITÉ D'ENSEIGNEMENT

Unité d'Enseignement. Un semestre est découpé en unités d'enseignement qui peuvent être obligatoire, optionnelle (choix à faire) ou facultative (UE en plus). Une UE représente un ensemble cohérent d'enseignements auquel est associé des ECTS.

### ECTS : EUROPEAN CREDITS TRANSFER SYSTEM

Les ECTS sont destinés à constituer l'unité de mesure commune des formations universitaires de Licence et de Master dans l'espace européen depuis sa création en 1989. Chaque UE obtenue est ainsi affectée d'un certain nombre d'ECTS (en général 30 par semestre d'enseignement). Le nombre d'ECTS est fonction de la charge globale de travail (CM, TD, TP, etc.) y compris le travail personnel. Le système des ECTS vise à faciliter la mobilité et la reconnaissance des diplômes en Europe.

## TERMES ASSOCIÉS AUX DIPLOMES

Les diplômes sont déclinés en domaines, mentions et parcours.

### DOMAINE

Le domaine correspond à un ensemble de formations relevant d'un champ disciplinaire ou professionnel commun. La plupart de nos formations relèvent du domaine Sciences, Technologies, Santé.

### MENTION

La mention correspond à un champ disciplinaire. Elle comprend, en général, plusieurs parcours.

### PARCOURS

Le parcours constitue une spécialisation particulière d'un champ disciplinaire choisie par l'étudiant au cours de son cursus.

## TERMES ASSOCIÉS AUX ENSEIGNEMENTS

### CM : COURS MAGISTRAL(AUX)

Cours dispensé en général devant un grand nombre d'étudiants (par exemple, une promotion entière), dans de grandes salles ou des amphis. Au-delà de l'importance du nombre d'étudiants, ce qui caractérise le cours magistral, est qu'il est le fait d'un enseignant qui en définit lui-même les structures et les modalités. Même si ses contenus font l'objet de concertations entre l'enseignant, l'équipe pédagogique, chaque cours magistral porte la marque de l'enseignant qui le dispense.



## TD : TRAVAUX DIRIGÉS

Ce sont des séances de travail en groupes restreints (de 25 à 40 étudiants selon les composantes), animés par des enseignants. Ils illustrent les cours magistraux et permettent d'approfondir les éléments apportés par ces derniers.

## TP : TRAVAUX PRATIQUES

Méthode d'enseignement permettant de mettre en pratique les connaissances théoriques acquises durant les CM et les TD. Généralement, cette mise en pratique se réalise au travers d'expérimentations. En règle générale, les groupes de TP sont constitués des 16 à 20 étudiants. Certains travaux pratiques peuvent être partiellement encadrés voire pas du tout. A contrario, certains TP, du fait de leur dangerosité, sont très encadrés (jusqu'à 1 enseignant pour quatre étudiants).

## PROJET OU BUREAU D'ÉTUDE

Le projet est une mise en pratique en autonomie ou en semi-autonomie des connaissances acquises. Il permet de vérifier l'acquisition des compétences.

## TERRAIN

Le terrain est une mise en pratique encadrée des connaissances acquises en dehors de l'université.

## STAGE

Le stage est une mise en pratique encadrée des connaissances acquises dans une entreprise ou un laboratoire de recherche. Il fait l'objet d'une législation très précise impliquant, en particulier, la nécessité d'une convention pour chaque stagiaire entre la structure d'accueil et l'université.

