

PÉRIODE D'ACCREDITATION : 2022 / 2026

UNIVERSITÉ PAUL SABATIER

---

# SYLLABUS MASTER

Mention Mathématiques et applications

Master 2 Science et Ingénierie des Données (SID)

---

<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)

2022 / 2023

23 MARS 2023

# SOMMAIRE

---

PRÉSENTATION . . . . .	3
PRÉSENTATION DE LA MENTION ET DU PARCOURS . . . . .	3
Mention Mathématiques et applications . . . . .	3
Parcours . . . . .	3
PRÉSENTATION DE L'ANNÉE DE Master 2 Science et Ingénierie des Données (SID) . . . . .	3
Aménagements des études : . . . . .	4
Liste des formations d'UT3 conseillées : . . . . .	4
RUBRIQUE CONTACTS . . . . .	5
CONTACTS PARCOURS . . . . .	5
CONTACTS MENTION . . . . .	5
CONTACTS DÉPARTEMENT : FSI.Math . . . . .	5
Tableau Synthétique des UE de la formation . . . . .	6
LISTE DES UE . . . . .	7
GLOSSAIRE . . . . .	23
TERMES GÉNÉRAUX . . . . .	23
TERMES ASSOCIÉS AUX DIPLOMES . . . . .	23
TERMES ASSOCIÉS AUX ENSEIGNEMENTS . . . . .	24

# 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 M2 Statistique et Informatique Décisionnelle (SID) est la deuxième année du parcours SID et fait suite au M1 SID.

Ce parcours est ouvert à la fois dans la mention master en Mathématiques et Applications et dans la mention master en Informatique.

L'accès est donc possible dans les deux mentions.

Ce parcours est ouvert à l'alternance.

## PRÉSENTATION DE L'ANNÉE DE MASTER 2 SCIENCE ET INGÉNIERIE DES DONNÉES (SID)

Le parcours SID (Statistique et Informatique Décisionnelle) est une formation résolument tournée vers les métiers de l'ingénierie des données. Elle en aborde tous les aspects : de la collecte des données à l'exploitation statistique et au machine learning, en passant par le stockage et la gestion de bases de données massives de type big data. Le parcours SID se déroule en 3 ans **de la Licence 3 au Master 2**, avec un apprentissage progressif des concepts fondamentaux, des méthodologies et des outils utilisés en **science des données**. A l'ère de l'intelligence artificielle, le profil Ødata scientistØ des étudiants SID est valorisé par une double compétence statistique - informatique décisionnelle, avec notamment une spécialisation en machine learning et en bases de données. Ce type de profil est aujourd'hui très recherché dans divers domaines professionnels, comme le montre le taux d'embauche de 100% de nos étudiants dès leur sortie du Master 2.

La grande majorité des étudiants en M1 SID (respectivement en M2 SID) viennent du L3 SID (respectivement du M1 SID). De ce fait, les nouvelles entrées en M1 ou M2 SID sont assez rares (pas plus d'une petite dizaine par an). Les compétences clefs du Master SID-Big data

- Mettre en oeuvre une étude statistique de la planification jusqu'à l'analyse et la synthèse des résultats (Sondage, Marketing, Étude biomédicale, Maîtrise Statistique des Procédés).
- Concevoir et développer un système d'information (base de données relationnelle ou NoSQL) pour aider au fonctionnement d'une organisation.
- Extraire de l'information pertinente des sources de données textuelles ou structurées pour les valoriser (aide à la décision, recherche d'information, fouille de données) en entreprise, dans une administration, ou dans un environnement de recherche.

- Analyser des masses de données et construire des tableaux de bord pour le pilotage d'institutions (entreprises, administration, etc.)
- Maintenir et analyser un système décisionnel (entrepôt de données, OLAP, ERP) pour gérer les ressources d'une organisation
- Travailler en équipe en suivant des méthodologies de gestion de projets
- Échanger et s'exprimer aisément à l'oral et à l'écrit en français et en anglais, dans un contexte professionnel
- Construire et valider un modèle mathématique pour traiter des ensembles de données (provenant de sondages, de relevés, de capteurs, etc.), afin de développer des outils d'aide à la décision en entreprise, dans une administration, ou dans un environnement de recherche.

## AMÉNAGEMENTS DES ÉTUDES :

Le M2 SID est la dernière année de la formation SID-Big Dat. Les étudiants peuvent s'y inscrire en formation initiale ou en alternance :

- En alternance : Contrat de professionnalisation ou d'apprentissage d'une année (de septembre à septembre) répartie en 17 semaines à l'université et 34 semaines en entreprise (selon le planning établi à l'avance).
- En formation initiale : À l'université de septembre à mi-mars (avec projets), puis stage jusque fin août.

## LISTE DES FORMATIONS D'UT3 CONSEILLÉES :

M1 MAT/INF-SID 2021

# RUBRIQUE CONTACTS

---

## CONTACTS PARCOURS

### RESPONSABLE MASTER 2 SCIENCE ET INGÉNIERIE DES DONNÉES (SID)

BERTHET Philippe

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

HUBERT Gilles

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

### SECRÉTAIRE PÉDAGOGIQUE

NICOLAS Clement

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

## CONTACTS MENTION

### RESPONSABLE DE MENTION MATHÉMATIQUES ET APPLICATIONS

CHOUQUET Cécile

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

Téléphone : 05.61.55.69.84

COSTANTINO Francesco

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

MARECHAL Pierre

Email : [pr.marechal@gmail.com](mailto:pr.marechal@gmail.com)

Téléphone : (poste) 76.60

## CONTACTS DÉPARTEMENT: FSI.MATH

### DIRECTEUR DU DÉPARTEMENT

GAVRILOV Lubomir

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

Téléphone : 05.61.55.76.62

### SECRETARIAT DU DÉPARTEMENT

RODRIGUES Manuella

Email : [manuella.rodrigues@univ-tlse3.fr](mailto:manuella.rodrigues@univ-tlse3.fr)

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	semestre*	ECTS	Obligatoire Facultatif	Cours	e-Cours	e-TD	TD	e-TP	TP	Projet	Stage
<b>Premier semestre</b>													
8	KMAK9AAU	STATISTIQUE AVANCÉE : MODÉLISATION ET APPLICATIONS	I	6	O	28			20		12		
12	KMAK9ADU	BASES DE DONNÉES À GRANDE ÉCHELLE	I	6	O								
13	KMAK9AD1	BD massivement parallèles et cloud computing				14			10		6		
	KMAK9AD2	NoSQL				14			10		6		
9	KMAK9ABU	DONNÉES MULTIMEDIA	I	6	O	20			16		24		
10	KMAK9ACU	APPRENTISSAGE STATISTIQUE AVANCÉ ET DEEP LEARNING	I	6	O								
11	KMAK9AC1	Apprentissage statistique avancé				12			12		6		
	KMAK9AC2	Apprentissage profond				8			8		4		
14	KMAK9AEU	PILOTAGE EN ENTREPRISE : OUTILS ET TECHNIQUES	I	3	O	14	1	1	10	1	6		
<b>Choisir 1 UE parmi les 4 UE suivantes :</b>													
15	KMAK9AVU	ANGLAIS	I	3	O				24				
16	KMAK9AWU	ALLEMAND	I	3	O				24				
17	KMAK9AXU	ESPAGNOL	I	3	O				24				
18	KMAK9AZU	FRANCAIS LANGUE ETRANGERE	I	3	O				24				
<b>Second semestre</b>													
19	KMAKAAAU	COMMUNICATION, ÉTHIQUE ET ASPECTS JURIDIQUES LIÉS AUX DONNÉE	II	3	O	16			12		2		
20	KMAKAABU	EXPÉRIENCES EN SCIENCES DE DONNÉES (POUR NON-ALTERNANTS)	II	3	O	8			8		2	50	
21	KMAKAACU	STAGE	II	18	O								12
22	KMAKAADU	PROJET INTER-PROMO « CHEF D'OEUVRE »	II	6	O	13			10		4	100	

\* AN :enseignements annuels, I : premier semestre, II : second semestre

---

## LISTE DES UE

---

<b>UE</b>	<b>STATISTIQUE AVANCÉE : MODÉLISATION ET APPLICATIONS</b>	<b>6 ECTS</b>	<b>1<sup>er</sup> semestre</b>
<b>KMAK9AAU</b>	Cours : 28h , TD : 20h , TP : 12h	Enseignement en français	Travail personnel 90 h

[ Retour liste de UE ]

### OBJECTIFS D'APPRENTISSAGE

- 1- La première partie de ce cours vient en complément des enseignements sur le modèle linéaire concernant les situations plus complexes telles que les données répétées, les données censurées, les réponses multidimensionnelles, les modèles mixtes, les données manquantes et leur imputation.
- 2- La deuxième partie de ce cours est consacrée à la maîtrise statistique des procédés (MSP). Il s'agit de techniques descriptives et inférentielles pour le contrôle statistique des processus.

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Plan de cours :

Après des rappels sur les modèles linéaires gaussiens classiques (régression linéaire, analyse de variance, analyse de covariance) seront présentées des nouvelles méthodes :

- 1a- Analyse de variance multidimensionnelle (MANOVA) pour modéliser une variable multidimensionnelle
- 1b- Modèles mixtes pour prendre en compte des facteurs à effet aléatoire
- 1c- Modèles pour données répétées (variable réponse mesurée à plusieurs reprises au cours du temps pour un même individu)
- 1d- Traitement des données manquantes, méthodes d'imputation
- 1e- Traitement des données censurées, analyse de survie, modèle de Cox
- j Mise en oeuvre des méthodes sous le logiciel R : <https://rstudio.com/products/rstudio/download/>
- 2a- Rappels et compléments sur la distribution gaussienne. Méthodes d'estimation des paramètres.
- 2b- Capabilités, cartes de contrôles et applications.
- 2c- Plans d'échantillonnage.

### PRÉ-REQUIS

- Modèles linéaires gaussiens et généralisés
- Probabilités et statistique inférentielle

### RÉFÉRENCES BIBLIOGRAPHIQUES

- A.C. Rencher, "Methods of multivariate analysis", Wiley, 1995.
- T.W. Anderson, "An introduction to multivariate statistical analysis", Wiley, 2003.
- Mercier S. et Bergeret F. « MSP : Principes et cas industriels »

### MOTS-CLÉS

- Modélisation statistique, données répétées, données censurées, maîtrise statistique des procédés,
- Capabilité, cartes de contrôle, plan d'échantillonnage.

UE	DONNÉES MULTIMEDIA	6 ECTS	1 <sup>er</sup> semestre
KMAK9ABU	Cours : 20h , TD : 16h , TP : 24h	Enseignement en français	Travail personnel 90 h

[ [Retour liste de UE](#) ]

## OBJECTIFS D'APPRENTISSAGE

Des volumes massifs d'informations unimodales (texte, son, vidéo) ou multimodales (film, CR médical...) sont produits par un nombre sans cesse croissant d'applications (achat en ligne, télémédecine...). L'exploitation de ces informations brutes dans un processus décisionnel nécessite l'extraction d'éléments d'information de bas niveau puis leur analyse afin de dégager des descripteurs de haut niveau (sémantiques). L'objectif de ce module est de présenter les différentes méthodes d'extraction et d'analyse d'informations (texte, vidéo et son). Les compétences attendues sont : a) traiter des informations brutes en vue d'en extraire des caractéristiques de forme et de contenu, b) analyser ces caractéristiques pour en dégager des connaissances utilisables dans des applications.

## DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

- Objectifs
- Extraction d'informations textuelles
  - Typologie des méthodes (entités nommées, relations, indices temporels, lexique)
  - Principes d'extraction des entités nommées : méthodes symboliques, méthodes supervisées, méthodes basées sur l'apprentissage profond
  - Principes d'extraction de relations
  - Application : extraction d'opinions
- Analyse sémantique de textes : Représentation de mots, phrases (méthodes symboliques et distributionnelles), calcul de similarité sémantique entre textes (mesures orientées graphes de connaissances et représentations distributionnelles), application (résumé de textes)
- Extraction d'informations sonores : numérisation, extraction de paramètres temporels et fréquentiels
- Extraction et exploitation de contenus visuels (image, vidéo) : Image numérique, quantification, extraction de paramètres liés à la couleur et la texture, analyse de formes et de contour, classification et recherche d'images par le contenu (vers l'indexation vidéo...)
- Évaluation : méthodologie et métriques
- Exploitation de contenus sonores : Classification de sons (phonèmes, locuteurs), recherche de mots clefs, transcription automatique de la parole, indexation audio

## PRÉ-REQUIS

- Bases de l'apprentissage automatique
- Bases sur le pré-traitement de texte

## COMPÉTENCES VISÉES

Connaissances acquises :

- principes élémentaires du traitement automatique des médias : texte, audio, image et vidéo.

Compétences visées :

- Traiter des informations brutes de tout support d'information (texte, image, son, vidéo) en vue d'en extraire des caractéristiques de forme et de contenu,
- Analyser ces caractéristiques pour les exploiter dans des applications.

## RÉFÉRENCES BIBLIOGRAPHIQUES

Image numérique couleur. A. Tremeau, 2004

Information extraction in a retrieval context. M. F. Moens, 2011

Reconnaissance automatique de la parole. J.-P. Haton, 2006

## MOTS-CLÉS

Extraction de paramètres (texte/audio/vidéo), entités nommées, sémantique, traitement du signal, modélisation, indexation par le contenu

<b>UE</b>	<b>APPRENTISSAGE STATISTIQUE AVANCÉ ET DEEP LEARNING</b>	<b>6 ECTS</b>	<b>1<sup>er</sup> semestre</b>
<b>Sous UE</b>	Apprentissage statistique avancé		
<b>KMAK9AC1</b>	Cours : 12h , TD : 12h , TP : 6h	Enseignement en français	Travail personnel 100 h

[\[ Retour liste de UE \]](#)

## OBJECTIFS D'APPRENTISSAGE

Étendre les compétences acquises en apprentissage statistique. Cela passe par une étude approfondie de nouveaux paradigmes et de nouvelles approches pour l'apprentissage ainsi que par une compréhension plus fine des mécanismes qui sous-tendent le succès ou l'échec de la mise en œuvre d'une méthode d'apprentissage. Ces concepts sont illustrés sur des simulations ainsi que sur des données réelles afin de permettre une mise en pratique opérationnelle.

## DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Plan de cours :

- Erreur de Bayes, risque, compromis biais variance.
- Validation croisée, sélection de modèle, bootstrap.
- Régression polynomiale, régularisation.
- Méthodes d'ensemble pour l'apprentissage supervisé : bagging, boosting, forêts aléatoires, gradient boosting.
- Méthodes à noyaux : ACP, SVM, régression, one class SVM.
- Détection d'anomalies, principe général et évaluation.
- Méthodes de détection d'anomalie : plus proche voisin, Local Outlier Factor, K means, distance de Mahalanobis, mélange de gaussienne, estimation de densité à noyaux, isolation forest, one class svm.
- L'ensemble des concepts vus en cours est directement manipulé sur des simulations ou sur des données réelles avec la librairie scikit-learn en python.

## PRÉ-REQUIS

- Machine Learning : théorie et algorithmes classiques, bonne maîtrise de python.
- Modélisation aléatoire (statistique et probabilité).

## COMPÉTENCES VISÉES

Compétences :

- Mettre en œuvre un modèle d'apprentissage pour un problème donné.
- Proposer, évaluer et comparer différents modèles.

Connaissances :

- Spectre large de méthodes d'apprentissage, leur mise en place et leur évaluation.

## RÉFÉRENCES BIBLIOGRAPHIQUES

- Hastie-Tibshirani : Elements of statistical learning

## MOTS-CLÉS

Apprentissage statistique, apprentissage profond, détection d'anomalies, méthodes à noyaux.

<b>UE</b>	<b>APPRENTISSAGE STATISTIQUE AVANCÉ ET DEEP LEARNING</b>	<b>6 ECTS</b>	<b>1<sup>er</sup> semestre</b>
<b>Sous UE</b>	Apprentissage profond		
<b>KMAK9AC2</b>	Cours : 8h , TD : 8h , TP : 4h	Enseignement en français	Travail personnel 100 h

[\[ Retour liste de UE \]](#)

## OBJECTIFS D'APPRENTISSAGE

Connaître les méthodes d'apprentissage profond, des modèles théoriques à leur mise en place en pratique. Le cours couvre l'apprentissage des réseaux multicouche, les réseaux convolutifs, les réseaux récurrents et les modèles génératifs. Il porte aussi sur les limites et les difficultés de l'apprentissage profond. Les TP utilisent la bibliothèque tensorflow2/keras. Le but final de ce cours est d'utiliser les méthodes d'apprentissage profond dans un challenge (le défi IA).

## DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Pan de cours :

- 1- Descente de gradient : Perceptron de Rosenblatt (interprétation géométrique et descente de gradient); Rétropropagation du gradient (détail du calcul, approche matricielle); Algorithmes de descente de gradient; convergence
- 2- Réseaux Convolutifs : Convolutif vs dense sur les images; Pooling; Skip connections; Architectures modèles convolutifs
- 3- Réseaux récurrents et mécanisme d'attention
- 4- Approche pratique du deep learning
- 5- Modèles image to image : Auto-encoder, Variational auto-encoder et réduction de dimension ; Réseaux siamois (pour l'apprentissage de représentation) ; U-net, Mask RCNN et segmentation d'image
- 6- Modèles génératifs
- 7- Limites du deep learning

## PRÉ-REQUIS

Machine learning, optimisation, algèbre linéaire.

## COMPÉTENCES VISÉES

Mettre en œuvre un modèle d'apprentissage profond pour un problème donné. Proposer, évaluer et comparer différents modèles.

## RÉFÉRENCES BIBLIOGRAPHIQUES

- The elements of statistical Learning Trevor Hastie
- Deep learning Ian Goodfellow, Yoshua Bengio, Aaron Courville

## MOTS-CLÉS

Deep learning.

<b>UE</b>	<b>BASES DE DONNÉES À GRANDE ÉCHELLE</b>	<b>6 ECTS</b>	<b>1<sup>er</sup> semestre</b>
<b>Sous UE</b>	BD massivement parallèles et cloud computing		
<b>KMAK9AD1</b>	Cours : 14h , TD : 10h , TP : 6h	Enseignement en français	Travail personnel 90 h

[\[ Retour liste de UE \]](#)

## OBJECTIFS D'APPRENTISSAGE

Présenter les principaux problèmes posés et introduire les méthodes proposées dans la conception des systèmes de gestion de bases de données parallèles et de cloud computing.

## DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Plan de cours :

1. Objectifs des BD parallèles
2. Méthodes de répartition de données
3. Stratégies de parallélisation de requêtes décisionnelles
4. Introduction aux traitements de données dans les systèmes clouds
5. Evaluation et optimisation élastique de requêtes décisionnelles
6. Etude comparative entre les SGBD parallèles et les systèmes clouds

## PRÉ-REQUIS

Systèmes de BD relationnels, Evaluation et optimisation de requêtes, Systèmes de BD distribués.

## COMPÉTENCES VISÉES

Connaissances acquises :

- architecture des SGBDs parallèles et son évolution en environnement Cloud.
- approches et méthodes de répartition de données.
- allocation élastique de ressources, modèles de coûts.
- évaluation et optimisation de requêtes en environnements parallèle et Cloud.

Compétences visées :

- savoir choisir un système approprié pour une application.
- concevoir et mettre en oeuvre une BD en environnement parallèle ou cloud computing
- comprendre les principaux problèmes posés et les méthodes proposées dans la conception des SGBD parallèles et les systèmes Clouds.

## RÉFÉRENCES BIBLIOGRAPHIQUES

- Principles of Distributed Database Systems, T. Oszu P. Valduriez, 3rd edition, 2011
- Traitement parallèle dans les bases de données relationnelles : concepts, méthodes et applications, A. Hameurlain, P. Bazex, F. Morvan, 1996.

## MOTS-CLÉS

Répartition des données, Parallélisme, Parallélisation de requêtes, Systèmes Clouds, Scalabilité, Élasticité.

<b>UE</b>	<b>BASES DE DONNÉES À GRANDE ÉCHELLE</b>	<b>6 ECTS</b>	<b>1<sup>er</sup> semestre</b>
<b>Sous UE</b>	NoSQL		
<b>KMAK9AD2</b>	Cours : 14h , TD : 10h , TP : 6h	Enseignement en français	Travail personnel 90 h

[\[ Retour liste de UE \]](#)

## OBJECTIFS D'APPRENTISSAGE

- Connaître les principes des bases de données NoSQL et apprendre à les manipuler.
- Connaître et mettre en œuvre la programmation parallèle et distribuée.

## DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Plan de cours :

- Positionnement du NoSQL par rapport au SQL
- Fondements théoriques des bases NoSQL
- Panorama des différents types de bases NoSQL
- Concepts des bases NoSQL orientés graphe
- Concepts des bases NoSQL orientés document
- Principes de la programmation parallèle et distribuée
- En TD : différents cas d'étude de modélisation et de manipulation de données dans des bases NoSQL orientées graphe et orientées document
- En TP : mise en œuvre (création, manipulation, interrogation, programmation) de cas d'études dans des systèmes de bases de données NoSQL largement utilisés en entreprise ; mise en oeuvre de programmation parallèle et distribuée.

## PRÉ-REQUIS

Bases de données relationnelles, algorithmique et programmation, toutes les compétences du M1.

## COMPÉTENCES VISÉES

Compétences visées :

- modéliser un système d'information à partir d'une analyse des besoins dans un contexte de données massives,
- concevoir, mettre en œuvre et manipuler une base de données non structurées,
- comprendre les principaux problèmes posés et les méthodes proposées dans la conception des systèmes de gestion de bases de données massives.

Connaissances acquises :

Concepts des bases NoSQL, modèles NoSQL graphe et document, manipulation de bases NoSQL graphe et document, principes de la programmation parallèle et distribuée.

## RÉFÉRENCES BIBLIOGRAPHIQUES

- Les bases de données NoSQL : Comprendre et mettre en oeuvre - Rudi Bruchez, Eyrolles
- MongoDB - The Definitive Guide, Kristina Chodorow, O'Reilly

## MOTS-CLÉS

NoSQL, MongoDB, HBase, Neo4j, big data.

<b>UE</b>	<b>PILOTAGE EN ENTREPRISE : OUTILS ET TECHNIQUES</b>	<b>3 ECTS</b>	<b>1<sup>er</sup> semestre</b>
<b>KMAK9AEU</b>	Cours : 14h , TD : 10h , TP : 6h , e-Cours : 1h , e-TD : 1h	Enseignement en français : 1h	Travail personnel : 45 h

[\[ Retour liste de UE \]](#)

## OBJECTIFS D'APPRENTISSAGE

Ce cours s'intéresse au pilotage d'entreprise au travers de deux facettes : l'une sur l'évaluation Certification Logiciel (ECL) et l'autre sur la Business Intelligence (BI).

## DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

### 1. ECL

- Introduction au besoin de certification logicielle et aux exigences imposées par la nécessité de la certification
- Approche de la qualité du logiciel à travers le processus de développement (Modèle CMM, Standards et Normes de développements, ...) et le Produit (Norme ISO9126, Techniques de Tests, ...)
- Présentation des normes (DO2167, DO178), outils et techniques appliquées dans le domaine aéronautique, spatial et des systèmes "Safety of Life".
- Introduction de la notion de certification logiciel : Processus, Outils Case, Traçabilité des exigences qualité mises en œuvre dans un contexte de développement, Analyses statiques.
- Notions de base de test de logiciel : Test Structurel, Test Fonctionnel, Campagnes de Test
- Diagnostic Qualité, arbre de défaillance, Méthode d'évaluation de l'arbre qualité.

### 2. BI

- en cours : Business Intelligence au service du pilotage de l'entreprise ; Principes d'analyses OLAP ; Principes et modèles des systèmes de recommandation
- en TD : Différents cas d'étude d'analyses de données OLAP ; Différents cas d'étude de systèmes de recommandation
- en TP : Mise en place d'analyses BI dans un outil représentatif du marché (ex. Business Objects)

## PRÉ-REQUIS

- maîtriser les concepts et différents modèles de développement
- bases de données relationnelles, SQL, Entrepôts et magasins de données, principes et outils ETL

## COMPÉTENCES VISÉES

Connaissances acquises :

- approches de qualité du logiciel, normes, notion de certification logicielle, diagnostic qualité,
- principes de BI, principes OLAP, modèles de recommandation.

Compétences visées :

- donner un avis d'expert sur les critères d'évaluation de logiciel évoqués dans le cursus de la formation,
- mettre en œuvre un système d'aide à la décision permettant de faire des analyses détaillées à partir des données issues des bases de données et de sources d'information dans un contexte de données massives,
- connaître et maîtriser des outils de Business intelligence.

## RÉFÉRENCES BIBLIOGRAPHIQUES

Software Engineering in Higher Education : Successful Business Intelligence / Unlock the Value of BI & Big Data / SAP BusinessObjects BI 4.0 / SAP BusinessObjects Web Intelligence (WebI) BI 4.2 / Introduction to Software Testing

## MOTS-CLÉS

- certification, processus de développement, exigences, test
- Business Intelligence, OLAP, Business Object, Big data, systèmes de recommandation

<b>UE</b>	<b>ANGLAIS</b>	<b>3 ECTS</b>	<b>1<sup>er</sup> semestre</b>
<b>KMAK9AVU</b>	TD : 24h	Enseignement en français	Travail personnel 51 h

[ [Retour liste de UE](#) ]

## ENSEIGNANT(E) RESPONSABLE

CHAPLIER Claire

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

## OBJECTIFS D'APPRENTISSAGE

Niveau C1/C2 du CECRL (Cadre Européen Commun de Référence pour les Langues)/ Permettre aux étudiants de développer les compétences indispensables à la réussite dans leur future vie professionnelle en contextes culturels variés. Acquérir l'autonomie linguistique nécessaire et perfectionner les outils de langue spécialisée permettant l'intégration professionnelle et la communication d'une expertise scientifique dans le contexte international.

## DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Développer :

- les compétences liées à la compréhension de publications scientifiques ou professionnelles rédigées en anglais ainsi que les compétences nécessaires à la compréhension de communications scientifiques orales.
- les outils d'expression permettant de maîtriser une présentation orale et/ou écrite et d'aborder une discussion critique dans le domaine scientifique
- la maîtrise des éléments d'argumentation critique à l'oral et/ou à l'écrit d'une publication scientifique une réflexion plus large sur leur place, leur intégration et leur rayonnement en tant que scientifiques dans la société, abordant des questions d'actualité, d'éthique, d'intégrité...

## PRÉ-REQUIS

Niveau B2

## COMPÉTENCES VISÉES

S'exprimer avec aisance à l'oral, devant un public, en usant de registres adaptés aux différents contextes et aux différents interlocuteurs. Se servir aisément d'une langue vivante autre que le français : compréhension et expression écrites et orales :

- Comprendre un article scientifique ou professionnel rédigé en anglais sur un sujet relatif à leur domaine.
- Produire un écrit scientifique ou technique dans un anglais adapté, de qualité et respectant les normes et usages de la communauté scientifique anglophone.
- Interagir à l'oral en anglais : réussir ses échanges formels et informels lors des colloques, réunions ou entretiens professionnels.

## MOTS-CLÉS

Projet Anglais scientifique Rédaction Publication Communication esprit critique scientifique interculturel

UE	ALLEMAND	3 ECTS	1 <sup>er</sup> semestre
KMAK9AWU	TD : 24h	Enseignement en français	Travail personnel 51 h
Sillon(s) :	Sillon 1		

[\[ Retour liste de UE \]](#)

### ENSEIGNANT(E) RESPONSABLE

SANTAMARINA Diego

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

### OBJECTIFS D'APPRENTISSAGE

Donner les bases grammaticales et la maîtrise des éléments de la langue de spécialité pour pouvoir être opérationnel en milieu germanophone..

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

TD mutualisés permettant de travailler différentes activités langagières favorisant l'acquisition de compétences transversales.

### PRÉ-REQUIS

Pas de pré-requis particulier à part la motivation et la capacité à fournir un travail personnel conséquent. Autorisation préalable du responsable de filière .

### RÉFÉRENCES BIBLIOGRAPHIQUES

Les documents et les conseils bibliographiques seront directement donnés en cours par l'enseignant.

### MOTS-CLÉS

Allemand-compétences transversales

UE	ESPAGNOL	3 ECTS	1 <sup>er</sup> semestre
KMAK9AXU	TD : 24h	Enseignement en français	Travail personnel 51 h

[\[ Retour liste de UE \]](#)

### ENSEIGNANT(E) RESPONSABLE

SANTAMARINA Diego

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

### OBJECTIFS D'APPRENTISSAGE

Maîtrise des bases grammaticales essentielles et de la langue de spécialité. Travail de compétences transversales.

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

UE mutualisée.

### PRÉ-REQUIS

Autorisation préalable du responsable de filière.

### RÉFÉRENCES BIBLIOGRAPHIQUES

Les documents et orientations bibliographiques seront directement données par l'enseignant.

### MOTS-CLÉS

Espagnol-compétences transversales.

UE	FRANCAIS LANGUE ETRANGERE	3 ECTS	1 <sup>er</sup> semestre
KMAK9AZU	TD : 24h	Enseignement en français	Travail personnel 51 h

[\[ Retour liste de UE \]](#)

### ENSEIGNANT(E) RESPONSABLE

DULAC Céline

Email : [celine.dulac@univ-tlse3.fr](mailto:celine.dulac@univ-tlse3.fr)

GOFFINET Akissi

Email : [akissi.goffinet@gmail.com](mailto:akissi.goffinet@gmail.com)

### OBJECTIFS D'APPRENTISSAGE

Développer ses compétences langagières et interculturelles en français durant un séjour d'études en France.

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

- compréhension et expression orales du français général de niveau A1/A2, B1 ou B2+ selon le cours suivi
- acquisition de vocabulaire et de structures de niveau A1/A2, B1 ou B2+ selon le cours suivi
- éléments de prononciation et de prosodie du français
- réflexion sur les différences interculturelles

### PRÉ-REQUIS

Passation du test ELAO. L'étudiant-e suit le cours de son niveau (A1/A2, B1 ou B2).

### SPÉCIFICITÉS

**Ce cours est accessible uniquement aux étudiant-e-s étrangers-ères non francophones et en échange à l'UT3.**

### COMPÉTENCES VISÉES

Les compétences visées dépendent du niveau CECRL de l'étudiant-e; chaque cours est adapté en fonction des descriptifs du CECRL.

### MOTS-CLÉS

Français Langue Etrangère, Insertion, Interculturalité

<b>UE</b>	<b>COMMUNICATION, ÉTHIQUE ET ASPECTS JURIDIQUES LIÉS AUX DONNÉE</b>	<b>3 ECTS</b>	<b>2<sup>nd</sup> semestre</b>
<b>KMAKAAAU</b>	Cours : 16h , TD : 12h , TP : 2h	Enseignement en français	Travail personnel 45 h

[\[ Retour liste de UE \]](#)

### OBJECTIFS D'APPRENTISSAGE

Ce module a deux objectifs :- étudier les concepts de propriété intellectuelle et de droit des données- savoir préparer et mener un entretien d'embauche professionnel.

### PRÉ-REQUIS

Techniques de communication vues en M1 SID

### COMPÉTENCES VISÉES

Compétences attestées après validation du module :

- Rechercher de l'information sur les réglementations encadrant l'usage de tout type de donnée
- Connaître les grands principes du RGPD et son domaine d'application
- Avoir des alertes sur les risques et limites de l'utilisation de tout type de donnée
- Bien se préparer à un entretien d'embauche
- Connaître la marche à suivre et les outils utiles pour la recherche d'emploi.

Connaissances acquises après validation du module : réglementation sur les données (RGPD), types de droits de propriété intellectuelle, techniques professionnelles relatives à un entretien d'embauche

### RÉFÉRENCES BIBLIOGRAPHIQUES

- M. Bourgeois, Droit de la donnée, LexisNexis
- J. Foucault, L. Panhaleux, D. Renaud, P. Begasse, RGPD Le comprendre et le mettre en œuvre, ENI, collect. DataPro.
- P.-Y. Gautier, Propriété littéraire et artistique, PUF, collect. droit fondamental

### MOTS-CLÉS

RGPD, privacy, base de données, open data, données à caractère personnel, droit européen des données

<b>UE</b>	<b>EXPÉRIENCES EN SCIENCES DE DONNÉES (POUR NON-ALTERNANTS)</b>	<b>3 ECTS</b>	<b>2<sup>nd</sup> semestre</b>
<b>KMAKAABU</b>	Cours : 8h , TD : 8h , Projet : 50h , TP : 2h	Enseignement en français	Travail personnel 57 h

[ [Retour liste de UE](#) ]

### OBJECTIFS D'APPRENTISSAGE

Mettre en œuvre une étude statistique sur des données réelles issues du milieu industriel ou biomédical, en utilisant l'éventail des méthodes étudiées dans l'UE ØStatistique avancée : modélisation et applicationsØ.

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Ce projet d'analyse statistique sur des données réelles permet aux étudiants de revoir l'ensemble des méthodes statistiques vues pendant leur cursus, mais aussi d'appréhender de nouveaux outils mathématiques et statistiques spécifiquement adaptées aux données étudiées. Il s'agit de montrer comment faire face aux difficultés inhérentes aux données réelles, par exemple les données manquantes. Dans le cadre d'un projet sur données biomédicales, nous pourrions présenter en particulier les méthodes d'analyse de survie pour analyser les données censurées (Méthode de Kaplan-Meier et modèle de Cox).

### PRÉ-REQUIS

Statistique descriptive et inférentielle, modélisation, analyse des données.

### COMPÉTENCES VISÉES

Connaissances acquises :

- méthodes et outils d'une analyse statistique complète sur données réelles,
- méthodes de modélisation pour données répétées ou données censurées,
- méthodes d'imputation de données manquantes.

Compétences visées :

- Mettre en œuvre une étude statistique de la planification jusqu'à l'analyse et la synthèse des résultats,
- Identifier des méthodes de modélisation appropriées, les mettre en œuvre, en étudier la validité et interpréter les résultats.

### RÉFÉRENCES BIBLIOGRAPHIQUES

- Statistical Modeling for Biomedical Researchers : A Simple Introduction to the Analysis of Complex Data, William D. Dupont
- Statistical Analysis with Missing Data. R.J.A.Little et D.B.Rubin

### MOTS-CLÉS

Etude statistique, données réelles, analyse, modélisation

UE	STAGE	18 ECTS	2 <sup>nd</sup> semestre
KMAKAACU	Stage : 14 mois	Enseignement en français	Travail personnel 450 h

[\[ Retour liste de UE \]](#)

### OBJECTIFS D'APPRENTISSAGE

Mettre en application tous les concepts étudiés en formation, les mettre en œuvre sur les technologies utilisées par l'entreprise (parfois nouvelles) et de s'adapter au fonctionnement de l'entreprise.

### DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

Missions dans un contexte applicatif industriel ou recherche.

### PRÉ-REQUIS

Ensemble de compétences mathématiques et informatiques dépendant du contexte applicatif.  
Faire valider le contenu du stage par les responsables de formations.

### COMPÉTENCES VISÉES

Connaissances acquises :

Répondre à des objectifs tracés en s'appuyant sur des outils liés en l'ensemble des compétences de la formation (informatique décisionnelle, base de données, statistique, machine learning, gestion de projet).

Compétences visées :

- Mettre en œuvre les connaissances et compétences acquises au cours de la formation pour répondre à différentes missions dans un contexte applicatif industriel ou recherche,
- S'adapter à un nouvel environnement de travail.

### MOTS-CLÉS

Mission, cadre applicatif, industrie, recherche.

<b>UE</b>	<b>PROJET INTER-PROMO « CHEF D'OEUVRE »</b>	<b>6 ECTS</b>	<b>2<sup>nd</sup> semestre</b>
<b>KMAKAADU</b>	Cours : 13h , TD : 10h , TP : 4h , Projet : 100h	Enseignement en français	Travail personnel 123 h

[\[ Retour liste de UE \]](#)

## OBJECTIFS D'APPRENTISSAGE

L'objectif de ce projet regroupant les 3 promos de la formation SID (L3, M1 et M2) est de combiner les compétences acquises jusqu'ici dans la formation. Le projet montre la connexion entre la statistique et l'informatique décisionnelle, et permet de développer un système d'aide à la décision fournissant des analyses à partir des données textuelles issues des bases de données scientifiques et technologiques et de sources d'information hétérogènes. Ce projet PIP fait donc appel aux compétences en bases de données, en fouille de données, en apprentissage statistique, en présentation de l'information élaborée, ...

## DESCRIPTION SYNTHÉTIQUE DES ENSEIGNEMENTS

L'un des grands enjeux du projet PIP réside dans le fait que les étudiants travaillent en équipe (une dizaine d'équipes) et apprennent à gérer un projet important.

Il est réalisé dans un contexte client/fournisseur : un cahier des charges, souvent en partenariat avec une entreprise, est fourni aux étudiants, définissant les exigences et les contraintes à respecter.

Il incombe ensuite aux étudiants d'instaurer une démarche collaborative.

Ce projet de synthèse mobilise l'ensemble des compétences de la formation (informatique décisionnelle, statistique, Machine Learning, gestion de projet) pendant deux semaines à plein temps, en janvier, après une pleine semaine de préparation des groupes et la présentation du sujet et des données par l'entreprise, en décembre.

## PRÉ-REQUIS

Bases de données, programmation, statistiques, machine learning.

## COMPÉTENCES VISÉES

Compétences visées :

- mettre en œuvre les outils et concepts vus dans la formation sur un problème réel
- travailler en équipe et suivre des recommandations sur un projet
- utiliser des outils de gestion de projet et de communication.... (github... )

Connaissances acquises :

- en fonction des projets
- collecte et analyse de données / valorisation des données / visualisation de données / algorithmes de machine learning / algorithmes de traitements du texte (recherche, classification) / algorithmes de prédiction

## MOTS-CLÉS

Projet de synthèse - Chef d'oeuvre.

## TERMES GÉNÉRAUX

### SYLLABUS

Dans l'enseignement supérieur, un syllabus est la présentation générale d'un cours ou d'une formation. Il inclut : objectifs, programme de formation, description des UE, prérequis, modalités d'évaluation, informations pratiques, etc.

### DÉPARTEMENT

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

### UE : UNITÉ D'ENSEIGNEMENT

Un semestre est découpé en unités d'enseignement qui peuvent être obligatoires, à choix ou facultatives. Une UE représente un ensemble cohérent d'enseignements auquel sont associés des ECTS.

### UE OBLIGATOIRE / UE FACULTATIVE

L'UE obligatoire fait référence à un enseignement qui doit être validé dans le cadre du contrat pédagogique. L'UE facultative vient en supplément des 60 ECTS de l'année. Elle est valorisée dans le supplément au diplôme. L'accumulation de crédits affectés à des UE facultatives ne contribue pas à la validation de semestres ni à la délivrance d'un diplôme.

### ECTS : EUROPEAN CREDITS TRANSFER SYSTEM

Les ECTS constituent l'unité de mesure commune des formations universitaires de licence et de master dans l'espace européen. Chaque UE obtenue est ainsi affectée d'un certain nombre d'ECTS (en général 30 par semestre d'enseignement, 60 par an). Le nombre d'ECTS varie en 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 des formations de l'UT3 relèvent du domaine « Sciences, Technologies, Santé ».

### MENTION

La mention correspond à un champ disciplinaire. Il s'agit du niveau principal de référence pour la définition des diplômes nationaux. La mention comprend, en général, plusieurs parcours.

### PARCOURS

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

## LICENCE CLASSIQUE

La licence classique est structurée en six semestres et permet de valider 180 crédits ECTS. Les UE peuvent être obligatoires, à choix ou facultatives. Le nombre d'ECTS d'une UE est fixé sur la base de 30 ECTS pour l'ensemble des UE obligatoires et à choix d'un semestre.

## LICENCE FLEXIBLE

À la rentrée 2022, l'université Toulouse III - Paul Sabatier met en place une licence flexible. Le principe est d'offrir une progression "à la carte" grâce au choix d'unités d'enseignement (UE). Il s'agit donc d'un parcours de formation personnalisable et flexible dans la durée. La progression de l'étudiant.e dépend de son niveau de départ et de son rythme personnel. L'inscription à une UE ne peut être faite qu'à condition d'avoir validé les UE pré-requises. Le choix de l'itinéraire de la licence flexible se fait en concertation étroite avec une direction des études (DE) et dépend de la formation antérieure, des orientations scientifiques et du projet professionnel de l'étudiant.e. L'obtention du diplôme est soumise à la validation de 180 crédits ECTS.

## DIRECTION DES ÉTUDES ET ENSEIGNANT·E RÉFÉRENT·E

La direction des études (DE) est constituée d'enseignantes et d'enseignants référents, d'une directrice ou d'un directeur des études et d'un secrétariat pédagogique. Elle organise le projet de formation de l'étudiant.e en proposant une individualisation de son parcours pouvant conduire à des aménagements. Elle est le lien entre l'étudiant.e, l'équipe pédagogique et l'administration.

## TERMES ASSOCIÉS AUX ENSEIGNEMENTS

### CM : COURS MAGISTRAL(AUX)

Cours dispensé en général devant un grand nombre d'étudiantes et d'étudiants (par exemple, une promotion entière), dans de grandes salles ou des amphithéâtres. Ce qui caractérise également le cours magistral est qu'il est le fait d'une enseignante ou d'un enseignant qui en définit les structures et les modalités. Même si ses contenus font l'objet de concertations avec l'équipe pédagogique, chaque cours magistral porte donc la marque de la personne qui le crée et le dispense.

### TD : TRAVAUX DIRIGÉS

Ce sont des séances de travail en groupes restreints (de 25 à 40 étudiantes et étudiants selon les composantes), animées par des enseignantes et enseignants. Les TD 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 et les groupes de TP sont constitués de 16 à 20 étudiantes et étudiants. Certains travaux pratiques peuvent être partiellement encadrés ou peuvent ne pas être encadrés du tout. A contrario, certains TP, du fait de leur dangerosité, sont très encadrés (jusqu'à une enseignante ou un enseignant pour quatre étudiantes et 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 de 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.

## SESSIONS D'ÉVALUATION

Il existe deux sessions d'évaluation : la session initiale et la seconde session (anciennement appelée "session de rattrapage", constituant une seconde chance). La session initiale peut être constituée d'examens partiels et terminaux ou de l'ensemble des épreuves de contrôle continu et d'un examen terminal. Les modalités de la seconde session peuvent être légèrement différentes selon les formations.

## SILLON

Un sillon est un bloc de trois créneaux de deux heures d'enseignement. Chaque UE est généralement affectée à un sillon. Sauf cas particuliers, les UE positionnées dans un même sillon ont donc des emplois du temps incompatibles.

