

# DÉLIBÉRATION

## Conseil d'administration

Séance du 6 juillet 2021

Délibération  
n°135-2021  
Point 4.8.6.1.2

### Point 4.8.6.1.2 de l'ordre du jour

**Création du Master Parcours Sciences des données et intelligence artificielle, mention sciences du langage – Faculté des lettres**

#### EXPOSE DES MOTIFS :

Il est proposé au CA d'examiner la demande de création de Master Parcours Sciences des données et intelligence artificielle, mention sciences du langage.

La transformation numérique de la société et de ses besoins s'appuie de plus en plus sur le traitement pertinent, pointu de données disponibles de façon massive (Big data). Cette évolution est permise par la croissance exponentielle des capacités de stockages à la disposition des organismes privés et publics, comme des particuliers. La capacité à traiter de façon adéquate ces informations représente une richesse économique, scientifique et sociétale. La problématique constitue donc un enjeu majeur pour l'Université de Strasbourg, qui commence par la formation des étudiants. Le Programme Investissements d'Avenir soutient le déploiement de cet axe stratégique. Le cursus de formation nécessaire court de la licence au master, et rayonnera vers le doctorat. Il s'agit pour l'université de monter en puissance tant en recherche qu'en formation.

Chaque parcours AISD s'appuie sur un parcours de master existant dans lequel les étudiants suivent un ensemble d'UE pour un total de 90 ECTS, qui est complété par les 30 ECTS du parcours mutualisé en science des données.

Afin d'élargir et diversifier l'offre de formation, l'équipe pédagogique du master souhaite créer un deuxième parcours dans la mention, intitulé Sciences de la donnée et intelligence artificielle (SDIA), dans la mention Sciences du langage. Deux arguments majeurs étaient ce projet, celui du volume grandissant des données et celui de la diversité de leurs usages. Les institutions et les entreprises, on le sait, se tournent toujours davantage vers la production et l'exploitation des données en grande quantité, et font massivement appel aux algorithmes d'apprentissage automatique pour transformer ces données en informations utilisables. Il existe donc une demande importante, dans les secteurs de la recherche ou de ses applications, de profils de « data scientist », d'analyste de données ou d'analystes thématique capables de maîtriser les techniques sur des données variées (numérique, texte, images).

**Le 15 juin 2021, la Commission de la formation et de la vie universitaire a approuvé à l'unanimité, par 29 voix pour.**

**Délibération :**

Le Conseil d'administration de l'Université de Strasbourg approuve la création du Master Parcours Sciences des données et intelligence artificielle, mention sciences du langage – Faculté des lettres.

**Résultat du vote :**

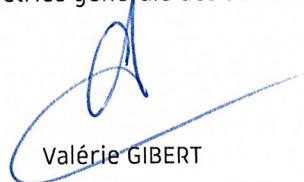
Nombre de membres en exercice	37
Nombre de votants	33
Nombre de voix pour	33
Nombre de voix contre	0
Nombre d'abstentions	0
Ne participe pas au vote	0

**Destinataires :**

- Madame la Rectrice déléguée pour l'enseignement supérieur, la recherche et l'innovation
- Direction générale des services
- Direction des finances
- Agence comptable

Fait à Strasbourg, le 8 juillet 2021

La Directrice générale des services



Valérie GIBERT

**Ouverture du parcours**

**Sciences des données et Intelligence Artificielle, Mention Sciences du Langage**

pour **2021/2022**

À faire parvenir par le directeur de composante, copie au responsable administratif de composante à la DES : [des-appui@unistra.fr](mailto:des-appui@unistra.fr), à Mme Bergmann ([pascale.bergmann@unistra.fr](mailto:pascale.bergmann@unistra.fr)), au Service de formation continue, le cas échéant : [dominique.schlaefli@unistra.fr](mailto:dominique.schlaefli@unistra.fr)

## **1. Nature de la demande**

Il s'agit d'une création ☒

Date d'approbation par le Conseil de composante : **9 mars 2021**

## **2. Exposé des motifs de la création**

La transformation numérique de la société et de ses besoins s'appuie de plus en plus sur le traitement pertinent, pointu de données disponibles de façon massive (Big data). Cette évolution est permise par la croissance exponentielle des capacités de stockages à la disposition des organismes privés et publics, comme des particuliers.

La capacité à traiter de façon adéquate ces informations représente une richesse économique, scientifique et sociétale. La problématique constitue donc un enjeu majeur pour l'Université de Strasbourg, qui commence par la formation des étudiants. Le Programme Investissements d'Avenir soutient le déploiement de cet axe stratégique.

Le cursus de formation nécessaire court de la licence au master, et rayonnera vers le doctorat. il s'agit pour l'université de monter en puissance tant en recherche, qu'en formation.

Une première étape a été la création en 2019-2020 d'un enseignement d'initiation à la science des données en 3<sup>e</sup> année de licence (approuvé en CFVU du 3 décembre 2019). Une seconde étape fut l'ouverture d'enseignements de science des données en Master, « transitoires » car ils ne concernent que la cohorte d'étudiants en M1 en 2020-2021 et en M2 en 2021-2022 (approuvé en CFVU du 30 juin 2020). La troisième étape fut la demande d'ouverture de 7 parcours de master AISD (Approche Interdisciplinaire en Science des Données) (approuvée en CFVU du 3 novembre 2020). L'étape actuelle consiste à organiser un parcours mutualisé d'enseignements en science des données accessible à partir de septembre 2021 dans le cadre des parcours AISD (Approche Interdisciplinaire en Science des Données) ouverts par les composantes qui le souhaitent.

Chaque parcours AISD s'appuie sur un parcours de master existant dans lequel les étudiants suivent un ensemble d'UE pour un total de 90 ECTS, qui est complété par les 30 ECTS du parcours mutualisé en science des données.

L'interdisciplinarité présente des difficultés telles que l'hétérogénéité des profils étudiants, les motivations et les attentes des étudiants. Nous encourageons les étudiants de L3 intéressés par les parcours AISD de master à suivre l'enseignement d'Initiation en Science Des Données, proposés en S5 et en S6. Cet enseignement est réalisé en petits groupes et en cours intégrés afin d'assurer un encadrement personnalisé et un ajustement de la pédagogie à la diversité des composantes d'origine des étudiants. Cet effort de pédagogie adaptée à des publics moins familiers du numérique est maintenu dans les enseignements de master. L'interdisciplinarité est aussi une richesse en soi car elle est l'occasion de s'ouvrir à d'autres

points de vue. Nous encouragerons les échanges de points de vue sur la science des données dans les différentes composantes de notre université en faisant intervenir des enseignants et des chercheurs représentatifs de ces nombreuses sensibilités : enjeux économiques, sociaux, juridiques, données massives, données médicales, données textuelles, etc. Nous profiterons de deux projets « transdisciplinaires » pour faire travailler ensemble 2 étudiants de composantes différentes sur un sujet encadré par un tuteur d'une troisième discipline. Au-delà de la confrontation de leurs cultures respectives, ces projets inciteront chaque étudiant à prendre un recul bénéfique pour aborder des problèmes nouveaux dans sa discipline d'origine plus tard.

Pour chaque étudiant, les débouchés sont d'abord ceux de son parcours disciplinaire existant. En effet, la production et l'utilisation des données deviennent une façon naturelle de travailler dans tous les métiers. Ainsi cette formation renforce l'employabilité de l'étudiant dans sa discipline d'origine. En plus, la formation suivie en science des données offre la possibilité de se spécialiser sur des métiers du numérique, à l'interface entre les métiers de l'informatique et de la discipline en lien avec le stockage, la modélisation et l'analyse de données. Ce sont donc de nouveaux débouchés qui s'ouvrent également à chaque étudiant.

Les effectifs visés sont de quelques étudiants par discipline afin d'une part de ne pas gêner l'organisation du parcours disciplinaire existant, et d'autre part de favoriser une grande diversité dans l'origine des étudiants regroupés lors des enseignements en science des données.

La **mention Sciences du langage**, créée en 2020 et actuellement composée d'un seul parcours, prépare des spécialistes dans le domaine général des Sciences du langage. La formation propose des enseignements fondamentaux dans les différentes branches de la linguistique, de la phonétique/phonologie et de la linguistique outillée. Elle inclut également des options de professionnalisation (écriture professionnelle et discours des entreprises/RSE) qui ouvrent des perspectives spécifiques aux diplômés de la mention.

Afin d'élargir et diversifier l'offre de formation, l'équipe pédagogique du master souhaite créer un deuxième parcours dans la mention, intitulé Sciences de la donnée et intelligence artificielle (SDIA), dans la mention Sciences du langage. Deux arguments majeurs étaient ce projet, celui du volume grandissant des données et celui de la diversité de leurs usages. Les institutions et les entreprises, on le sait, se tournent toujours davantage vers la production et l'exploitation des données en grande quantité, et font massivement appel aux algorithmes d'apprentissage automatique pour transformer ces données en informations utilisables. Il existe donc une demande importante, dans les secteurs de la recherche ou de ses applications, de profils de « data scientist », d'analyste de données ou d'analystes thématique capables de maîtriser les techniques sur des données variées (numérique, texte, images).

Par ailleurs, l'univers des données se diversifie et se complexifie à mesure que des domaines d'application nouveaux apparaissent. La donnée n'est plus seulement une catégorie mathématique abstraite, elle devient une valeur sociale, objet d'échanges et d'investigations. Une telle situation oblige à introduire davantage d'interdisciplinarité dans les méthodes et les compétences des experts, afin que les spécialistes de la donnée soient également linguistes, médecins, ingénieurs pédagogiques, biologistes, etc. C'est dans cette perspective qu'un parcours combinant compétences en linguistique et spécialisation en sciences des données, à même d'ouvrir de nouvelles opportunités sur le marché de travail pour nos étudiants de Sciences du langage, montre toute sa pertinence.

#### **Composante de rattachement : Lettres**

#### **Composantes ou services associées :**

Les enseignements en science des données sont assurés par l'UFR de mathématique et d'informatique, et pris en charge par un financement de l'IdEx. Les enseignements en linguistique et phonétique sont assurés par la Faculté de Lettres, ces cours étant mutualisés avec le parcours de Sciences du langage existant.

### **3. Responsable de la formation pour l'Université de Strasbourg**

Prénom, Nom : **Amalia TODIRASCU**

Grade : Professeure des Universités

CNU : 07

Adresse : 22, rue René Descartes, BP 80010, 67084 Strasbourg

Téléphone : 03 68 85 65 86

Email : todiras@unistra.fr

### **4. Conditions d'admission et publics concernés**

#### **4.1. Mode de recrutement / sélection<sup>1</sup>**

Le recrutement sera fait par la ou le responsable du parcours AISD disciplinaire dans la composante associée. Il tiendra compte du nombre de places ouvertes par chaque composante. Une concertation sera organisée par les responsables du parcours mutualisé en science des données, en particulier afin de gérer les candidatures simultanées à plusieurs parcours AISD. La sélection des candidat·es sera faite sur la base de leurs résultats académiques en licence, ainsi que sur l'exposé de leurs motivations à suivre une formation interdisciplinaire en science des données. À ce titre, avoir suivi l'UE d'Initiation à la Science Des Données en 3<sup>e</sup> année de licence sera un indicateur important, même si ce n'est pas un prérequis, à fortiori pour les candidats provenant d'autres universités.

Pour la faculté de lettres : la candidature se fait sur dossier (le même que celui demandé pour le master Sciences du langage : CV, lettre de motivation, relevés de notes, diplômes obtenus, certificat qui prouve un niveau B2 minimum en anglais ; pour les candidats étrangers, le niveau C1 minimum en français).

#### **4.2. Effectifs prévisionnels**

La capacité d'accueil est fixée par la somme des places ouvertes par chacun des parcours AISD. Pour la cohorte commençant en septembre 2021 et terminant l'été 2023, l'effectif maximal est de 34 étudiants. Pour la Faculté de Lettres, **3 places** ont été prévues et demandées.

#### **4.3. Modalités d'évaluation des étudiants**

Les enseignements mutualisés en science des données sont évalués en contrôle continu. Les autres enseignements du master Sciences du langage sont évalués encore en contrôle terminal, le master devra passer en ECI l'année prochaine.

#### **4.4. Équipe pédagogique**

En application de l'article L613-2, al. 2, la liste des enseignant·es intervenant dans les diplômes d'université est publiée sur le site internet de l'établissement.

<sup>1</sup>. Rappel : la sélection est possible à l'entrée en DUT, en M1, dans le cadre des diplômes d'établissement et/ou d'école.

## ► Enseignants universitaires

Nom et grade des enseignants-chercheurs, enseignants ou chercheurs		Section CNU	Composante ou établissement	Nombre d'heures assurées (HETD)	Enseignements dispensés	
Nom	Grade					
Lachiche Nicolas	MC	27	IUT Robert Schuman	28	Science des données	
Meillier Céline	MC	61	Télécom Physique Strasbourg	28	Science des données	
Dolques Xavier	MC	27	Mathématique et informatique	140	Science des données	
Bernhard Delphine	MC	27	Langues	14	Science des données	
Marc-Zwecker Stella	MC	27	Mathématique et informatique	14	Science des données	
Chapuis Yves-André	MC	63	Physique et Ingénierie	14	Science des données	
Schnedecker Catherine	PR	07	Lettres		Sciences du langage	
Lammert Marie	MC	07	Lettres		Sciences du langage	
Kuyumcuyan Annie	PR	07	Lettres		Sciences du langage	
Theissen Anne	MC	07	Lettres		Sciences du langage	
Gerhard-Krait Francine	MC	07	Lettres		Sciences du langage	
Sock Rudolph	PR	07	Lettres		Sciences du langage	
Todirascu Amalia	PR	07	Lettres		Sciences du langage	
Konstantzer Stephane	MC	11	Langues		Anglais	
Benninger Céline	MC	07	Lettres		Sciences du langage	
Revol Thierry	PR	07	Lettres		Sciences du langage	

## ► Enseignements

### Tronc commun

Ce tronc commun est intégré dans les UE4 et UE5, étiquetées AISD. Les projets interdisciplinaires sont intégrés en UE1

Intitulé de l'UE	ECTS	Coef.	Compétences attendues	CM	CI	TD	TP	TE	Total (HETD)
C1 Concepts, enjeux et transdisciplinarité	3	1	Connaitre le domaine de la science des données	14		10	4	5	33,7
C2 Éthique et droit	3	1	Connaitre les règlements et bonnes pratiques	14		10	4	5	33,7
M1 Méthodes d'apprentissage supervisé	3	1	Savoir utiliser les algorithmes d'apprentissage supervisé	6		8	14	10	26,3
M2 Méthodes d'apprentissage non	3	1	Savoir utiliser les algorithmes d'apprentissage non-	6		8	14	10	26,3

supervisé et d'apprentissage profond			supervisé et d'apprentissage profond						
M3 Apprentissage symbolique et numérique	3	1	Comprendre les techniques d'apprentissage symbolique et d'apprentissage numérique		14		14	10	26,8
D1 Collecte, fiabilité et visualisation des données	3	1	Savoir préparer et explorer des données	6		8	14	10	26,3
D2 Outils statistiques	3	1	Connaitre les outils statistiques	6		8	14	10	26,3
D3 Représentation et traitement des données	3	1	Savoir traiter les types de données courants	6		8	14	10	26,3
A1 Projet transdisciplinaire 1	3	1	Étudier un problème réel					75	
A2 Projet transdisciplinaire 2	3	1	Évaluer une solution à un problème réel					75	
<b>Total</b>									

CM = cours magistraux / CI = cours intégrés / TD = travaux dirigés / TP = travaux pratiques / TE = travail étudiant hors cours, TD et TP

#### Parcours mutualisé

L'architecture du parcours *Sciences de données et Intelligence Artificielle* de la mention *Sciences du langage* est construite sur le socle commun avec le parcours *Sciences du langage* (UE2 et UE3 sont mutualisés) et une partie d'UE1, en annexe.

Intitulé de l'UE	ECTS	Coef.	Compétences attendues	Matières	CM	CI	TD	TP	TE	Total (ETD)
S1 UE1 <b>Initiation à la recherche</b>	3	1	Connaitre les thématiques de recherche en linguistique	Séminaire SéMaSciLa			6			6
			Savoir faire une recherche bibliographique, gestion avec Zotero	Méthodologie de recherche			6			6
S1 UE2 <b>Outilage scientifique</b>	6	1	Connaitre les approches théoriques en TAL, les ressources, les applications	Introduction au Traitement automatique des langues (TAL)			12			12
			Maitriser les outils d'annotation automatique (morphosyntaxique et syntaxique)	Outils d'annotation automatique			12			12
			Acquisition de vocabulaire scientifique et approfondissement de l'étude des fonctionnements de la grammaire anglaise	Anglais scientifique			12			12
			Maitriser les outils de type concordanciers ; Maitriser le langage CQL et les expressions régulières	Méthodologie des corpus			12			12
S1 UE3	12	1	savoir procéder à des analyses	Principaux courants en	12					18

<b>Fondements scientifiques</b>			critiques de certains modèles (théoriques, conceptuels, physiques, statistiques, mathématiques...) de phonétique et de phonologie pour pouvoir confronter l'« ancien » au « nouveau », et illustrer le processus sous-jacent à l'émergence de ces théories phares dans le domaine de la production et de la perception de la parole.	phonétique et phonologie							
			12 1 Maitrise des déterminants et de leur rôle référentiel	Interface morphosyntaxe / sémantique	12					18	
			12 1 Maitrise des notions fondamentales en syntaxe	Syntaxe : Linéarisation de l'énoncé	12					18	
			12 1 Identifier les pronoms dits indéfinis ; Caractériser les fondements morphologiques, les traits syntaxiques et sémantico-référentiels des différentes formes étudiées et les repérer en discours	Sémantique de l'indéfinitude	12					18	
			12 1 retrouver les caractéristiques essentielles d'une langue romane dans un texte, - les mettre en rapport avec celles du français (ainsi que celles du latin et d'autres langues romanes).	Linguistique romane ou diachronique	12					18	
S2 UE1 <b>Projet professionnel</b>	6	1	Connaitre le fonctionnement d'une équipe de recherche et les thématiques de l'équipe	LiLPa (Séminaire de recherche)			6				
			Savoir citer une bibliographie, éviter le plagiat, apprendre les stratégies de reformulation	Méthodologie du mémoire (Présentation acad. / Plagiat)			8				
S2 UE2 <b>Outilage scientifique</b>	6	1	Savoir structurer des données en XML et écrire un schéma XML ou une DTD	Exploitation de corpus/XML			12			12	
			Mettre en application les techniques acquises	Projet tutoré d'exploitation de corpus					10		
			Acquisition de vocabulaire scientifique et des techniques de traduction ; Comprendre le sens global et détaillé d'un texte à caractère scientifique ; Savoir isoler les unités de traductions	Langue : Anglais scientifique			12			12	
S2 UE3 <b>Fondements scientifiques</b>	12	1	analyser du comportement de structures articulatori-acoustiques dans plusieurs contextes spatiotemporelles.	Phonétique : Organisation spatiotemporelle de la parole	12					18	
			Connaitre les procédés de nominalisation	Morphologie : La nominalisation	12					18	
			Analyser des phénomènes syntaxiques complexes	Syntaxe : Enchâssement /	12					18	

				emboitement						
				Analyser la classe d'adverbes au niveau morphologique et sémantique	Sémantique : L'adverbe	12				18
				construire un sujet de recherche diachronique ; exploiter des bases de données diachroniques	Linguistique romane ou diachronique	12				18
S3 UE1 <b>Recherche</b>	6	1	Connaitre le fonctionnement d'une équipe	LiLPa ou SéMaSciLa		8				8
			Savoir écrire un état de l'art	Méthodologie du mémoire		6				6
S3 UE2 <b>Outilage scientifique</b>	6	1	Appliquer des méthodes statistiques avec R pour l'analyse de données	Méthodes statistiques pour l'analyse de données textuelles		12				
			Analyser des noms d'humains dans les ressources lexicographiques et corpus	Lexicologie	24					30
			Connaitre les plateformes d'annotation manuelle de haut niveau (corréférence)	Annotations de haut niveau		12				
S3 UE3 <b>Fondements scientifiques</b>	9	3	savoir démontrer, d'après l'analyse de données sur des phénomènes anticipatoires et à partir du concept de la viabilité, que la parole est aussi bien un ensemble de mouvements rendus audibles (et visibles) qu'un ensemble de sons produits par des mouvements ; comprendre la nature de la spécification mutuelle des niveaux articulatoire et acoustique, un trait qui semble caractériser les systèmes cognitifs sensorimoteurs en général et, tout particulièrement, celui de la production-perception de la parole.	Phonétique et phonologie : Parole et cognition	12					18
			Connaitre les procédés de création lexicale	Morphologie : La néologie	12					18
			Connaitre les approches théoriques et pratiques en analyse de discours	Analyse du discours : Les dialogues	12					18
			Identifier les mentions (ou maillons) d'une chaîne de référence et les annoter ; Identifier les chaînes de référence d'un texte ; Utiliser les outils d'annotation adéquat pour repérer les chaînes d'un texte ; Exploiter les résultats de l'annotation (analyse quantitative et qualitative)	Sémantique : Chaînes de référence	12					18
			Lecture autonome de textes anciens. Utilisation de la	Linguistique romane ou diachronique	12					18

			bibliographie de référence.						
UE5 Professionnalisation	3	1	Savoir se présenter dans un CV ou une lettre de motivation	Bilan de compétences		8			8
			Savoir comment chercher un stage et mettre en place une convention de stage	Préparation au stage		3			3
			Savoir présenter une communication scientifique en anglais	Expression scientifique orale en anglais		12			12
UE1 Recherche	15	1	Savoir rédiger une synthèse argumentée et un projet de recherche	Mémoire					
UE1 Stage	15	1	Gérer un projet professionnel	Stage				Min 210 h max 6 mois	Min 210

► **Liste des UE disciplinaires enseignées en langue étrangère : *Une réflexion est en cours dans la mention à ce sujet***

► **Stage pratique d'application (inclus dans la formation) : *Volume minimal 210 heures ; durée maximale 6 mois***

## 5. Dispositifs de suivi de la formation

### 5.1. Évaluation des formations

Les formations seront évaluées conformément aux pratiques de l'université, cf. <https://evaluation-formation.unistra.fr>. Chaque parcours AISD sera évalué dans le cadre de sa composante. La partie des enseignements mutualisés en science des données sera évaluée dans le cadre des formations de l'UFR mathématique et informatique. Les enseignements disciplinaires seront évalués dans le cadre de la mention Sciences du langage.

### 5.2. Évaluation des enseignements

Un dispositif d'évaluation des enseignements par les étudiants est mis en place à la fin de chaque UE sur Moodle. Les réponses des étudiants seront analysées par l'équipe pédagogique dans un but d'amélioration continue de l'alignement pédagogique.

### 5.3. Conseil de perfectionnement

Un conseil de perfectionnement commun à tous les parcours AISD sera mis en place chaque année. Le conseil de perfectionnement de la mention Sciences du langage évaluera aussi le parcours AISD.

## 6. Autres dispositifs, le cas échéant :

► **Budget prévisionnel : *Non pertinent (inclus dans le BP et le Dialogue de gestion de la Faculté des Lettres***

► **Financement : *Cout constant***

► **Paramétrage des droits d'inscription : *Droits de base du diplôme***

LICENCE () ou MASTER (X)

Droits spécifiques : **Non** (sous réserves des décisions prises par l'Unistra)

**Pages 9 à 12 suivantes : Détail de la maquette du parcours (tronc commun et enseignements mutualisés confondus).**

## Master de Sciences du langage : Parcours « Sciences de la donnée et intelligence artificielle » (SDIA) / Semestre 1

S	UE	Type d'enseignement	ECTS	Coeff.	Intitulé de l'UE	Codes	Modules	Volume		Ventilation				
								Cours	TE*	CM	CI	TD	TP	Cout
S1	UE1	Recherche	3	1	Eléments transversaux	LT41GM11	SéMaSciLa (Séminaire de recherche)	6		6				9
						LT41GM12	Méthodologie documentaire (Initiation à Zotero)	6				6		6
	UE2	Outilage scientifique	6	2	Méthodologie disciplinaire	LT41GM22	Introduction au Traitement automatique des langues (TAL)	12				12		12
						LT41GM23	Outils d'annotation automatique	12				12		12
						LT41GM21	Langue : Anglais scientifique	12				12		12
						LT36GM10	Linguistique de corpus	12		12				18
						LT41GM30	Principaux courants en phonétique et phonologie	12		12				18
	UE3	Fondements scientifiques	12	4	Tronc commun disciplinaire	LT41GM31	Interface morphosyntaxe / sémantique	12		12				18
						LT41GM32	Syntaxe : Linéarisation de l'énonce	12		12				18
						LT41GM33	Sémantique de l'indéfinitude	12		12				18
						LT41GM34	Linguistique romane ou diachronique	12		12				18
						En attente	M-1 : Méthodes d'apprentissage supervisé	28	10	6		8	14	26,333
	UE4	Approches interdisciplinaires des sciences de la donnée (AISD)	6	2	Concepts et méthodes	En attente	C-1 : Concepts, enjeux et transdisciplinarité	28	5	14		10	4	33,667
	UE5	AISD	3	1	Données et outils	En attente	D-1 : Collecte, fiabilité et visualisation des données	28	10	6		8	14	26,333
<b>Totaux</b>			<b>30</b>	<b>10</b>				<b>204</b>	<b>25</b>	<b>104</b>	<b>0</b>	<b>68</b>	<b>32</b>	<b>245,33</b>

\*TE : Comptabilisation du temps de travail de l'étudiant·e consacré à un projet personnel, une réalisation collective, etc.

Master de Sciences du langage : Parcours « Sciences de la donnée et intelligence artificielle » (SDIA) / Semestre 2															
S	UE	Type d'enseignement	ECTS	Coeff.	Intitulé de l'UE	Codes	Modules	Volume		Ventilation					
								Cours	TE*	CM	CI	TD	TP	Cout	
S2	UE1	Projet professionnel	6	2	Eléments transversaux	LT41HM10	LilPa (Séminaire de recherche)	6		6				9	
						LT41HM12	Méthodologie du mémoire (Présentation acad. / Plagiat)	6		6				9	
						En attente	A-1 : Projet transdisciplinaire		75					0	
	UE2	Outilage scientifique	6	2	Méthodologie disciplinaire	LT41HM20	Exploitation de corpus/XML	12				12		12	
						Code à créer	Projet tutoré d'exploitation de corpus		10					0	
						LT41HM22	Langue : Anglais scientifique	12				12		12	
	UE3	Fondements scientifiques	9	3	Tronc commun disciplinaire	LT41HM30	Phonétique : Organisation spatiotemporelle de la parole	12		12				18	
						LT41HM31	Morphologie : La nominalisation	12		12				18	
						LT41HM32	Syntaxe : Enchâssement / emboîtement	12		12				18	
						LT41HM33	Sémantique : L'adverbe	12		12				18	
						LT41HM34	Linguistique romane ou diachronique	12		12				18	
	UE4	Approches interdisciplinaires des sciences de la donnée (AISD)	6	2	Concepts et méthodes	En attente	M-2 : Méthodes d'apprentissage non-supervisé et d'appr. profond	28	10	6		8	14	26,333	
						En attente	C-2 : Éthique et droit	28	10	6		8	14	26,333	
	UE5	AISD	3	1	Données et outils	En attente	D-2 : Outils statistiques	28	10	6		8	14	26,333	
Totaux			30	10					180	115	90	0	48	42	211

\*TE : Comptabilisation du temps de travail de l'étudiant·e consacré à un projet personnel, une réalisation collective, etc.

## Master de Sciences du langage : Parcours « Sciences de la donnée et intelligence artificielle » (SDIA) / Semestre 3

S	UE	Type d'enseignement	ECTS	Coeff.	Intitulé de l'UE	Codes	Modules	Volume		Ventilation				
								Cours	TE*	CM	CI	TD	TP	Cout
S3	UE1	Recherche	6	2	Eléments transversaux	LT41KM10 ou LT41KM11	LiLPa ou SéMaSciLa (Séminaire de recherche)	8		8				12
						LT41KM12	Méthodologie du mémoire (État de l'art et citation)	6		6				9
						En attente	A-2 : Projet transdisciplinaire		75					0
	UE2	Outilage scientifique	6	2	Méthodologie disciplinaire	LGC4KM24	Méthodes statistiques pour l'analyse de données textuelles	12				12		12
						LT41KM43	Lexicologie	24			24			30
						LT41KM20	Annotations de haut niveau	12				12		12
	UE3	Fondements scientifiques	9	3	Tronc commun disciplinaire	LT41KM30	Phonétique et phonologie : Parole et cognition	12		12				18
						LT41KM31	Morphologie : La néologie	12		12				18
						LT41KM32	Analyse du discours : Les dialogues	12		12				18
						LT41KM33	Sémantique : Chaines de référence	12		12				18
						LT41KM34	Linguistique romane ou diachronique	12		12				18
	UE4	Approches interdisciplinaires des sciences de la donnée	6	2	Méthodes et représentations	En attente	M-3 : Méthodes par apprentissage	28	10		14		14	26,833
						En attente	D-3 : Représentation et traitement des données	28	10	6		8	14	26,333
	UE5	Professionalisation	3	1	Connaissance du milieu professionnel	En attente	Bilan de compétences	8				8		8
						LT41KM50	Préparation au stage	3		3				4,5
						LGB2KM42	Expression scientifique orale en anglais	12				12		12
<b>Totaux</b>			<b>30</b>	<b>10</b>				<b>201</b>	<b>95</b>	<b>83</b>	<b>38</b>	<b>52</b>	<b>28</b>	<b>242,67</b>

\*TE : Comptabilisation du temps de travail de l'étudiant·e consacré à un projet personnel, une réalisation collective, etc.

Master de Sciences du langage : Parcours « Sciences de la donnée et intelligence artificielle » (SDIA) / Semestre 4													
S	UE	Type d'enseignement	ECTS	Coeff.	Intitulé de l'UE	Codes	Modules	Volume		Ventilation			
								Cours	TE*	CM	CI	TD	TP
S4	UE1	Recherche	15	5	Mémoire	LT41LM20	Mémoire		100				0
	UE2	AISD	15	5	Stage	En attente	Stage		210				0
Totaux			30	10				0	310	0	0	0	0

\*TE : Comptabilisation du temps de travail de l'étudiant·e consacré à un projet personnel, une réalisation collective, etc.

	Volume	Cours	TE
Semestre 1	204	25	
Semestre 2	180	115	
<b>Première année</b>	<b>384</b>	<b>140</b>	
<b>Total</b>	<b>524</b>		
Semestre 3	201	95	
Semestre 4	0	310	
<b>Deuxième année</b>	<b>201</b>	<b>405</b>	
<b>Total</b>	<b>606</b>		
<b>Première et Deuxième années</b>	<b>585</b>	<b>545</b>	
<b>Total</b>	<b>1130</b>		

<b>A-1</b>	<b>Formation de rattachement</b>	<b>Intitulé de l'enseignement</b>
	<b>Parcours Approche Interdisciplinaire en Science des Données</b>	<b>Projet transdisciplinaire 1</b>

**RESPONSABLE DE L'ENSEIGNEMENT**

<b>Nom/Prénom</b>	<b>Discipline</b>	<b>Composante / Courriel/Tél. (fixe)</b>
LACHICHE NICOLAS	Informatique	IUT Robert Schuman / <a href="mailto:nicolas.lachiche@unistra.fr">nicolas.lachiche@unistra.fr</a> / 54577
KIEFFER BRUNO	Biologie	ESBS, <a href="mailto:bruno.kieffer@unistra.fr">bruno.kieffer@unistra.fr</a> , 54722
DOLQUES XAVIER	Informatique	UFR Math-Info / dolques@unistra.fr

**MODALITES D'ORGANISATION DE L'ENSEIGNEMENT**

Semestre 2  
Projet de 75h par étudiant, en binôme, encadrement 4 HETD par binôme

**LANGUE ET NIVEAU DE LANGUE DE L'ENSEIGNEMENT : FRANÇAIS****DESCRIPTION DE L'ENSEIGNEMENT**

- Analyse du contexte et aspects sociaux et éthiques
- Etablissement d'un cahier des charges

**COMPETENCES VISEES****Objectifs en termes de connaissances**

- Aucun, c'est un projet de mise en application des connaissances acquises dans les autres UE du parcours

**Objectifs en termes de compétences**

- Appréhender un projet d'analyse de données dans sa globalité (éthique, technique)
- Interagir en pluridisciplinarité
- Rédiger un cahier des charges

**ENSEIGNEMENT – ORGANISATION HORAIRES**

<b>Titre de l'enseignement</b>	<b>CM</b>	<b>CI</b>	<b>TD</b>	<b>TP</b>
<b>Projet transdisciplinaire</b>				
<b>Eq. HETD=</b>				

**MODALITES D'EVALUATION DES ENSEIGNEMENTS**

(session initiale et de rattrapage)évaluation continue intégrale

Etudiants en situation spécifique : modalités d'évaluation continue adaptées à la situation

<b>A-2</b>	<b>Formation de rattachement</b>	<b>Intitulé de l'enseignement</b>
	<b>Parcours Approche Interdisciplinaire en Science des Données</b>	<b>Projet transdisciplinaire 2</b>

**RESPONSABLE DE L'ENSEIGNEMENT**

<b>Nom/Prénom</b>	<b>Discipline</b>	<b>Composante / Courriel/Tél. (fixe)</b>
LACHICHE NICOLAS	Informatique	IUT Robert Schuman / <a href="mailto:nicolas.lachiche@unistra.fr">nicolas.lachiche@unistra.fr</a> / 54577
KIEFFER BRUNO	Biologie	ESBS, <a href="mailto:bruno.kieffer@unistra.fr">bruno.kieffer@unistra.fr</a> , 54722
DOLQUES XAVIER	Informatique	UFR Math-Info / dolques@unistra.fr

**MODALITES D'ORGANISATION DE L'ENSEIGNEMENT**

Semestre 3  
Projet de 75h par étudiant, en binôme, encadrement 4 HETD par binôme

**LANGUE ET NIVEAU DE LANGUE DE L'ENSEIGNEMENT : FRANÇAIS****DESCRIPTION DE L'ENSEIGNEMENT**

- Exploration de solutions
- Analyse critique des résultats

**COMPETENCES VISEES****Objectifs en termes de connaissances**

- Aucun, c'est un projet de mise en application des connaissances acquises dans les autres UE du parcours

**Objectifs en termes de compétences**

- Mettre en oeuvre les méthodologies et techniques du parcours
- Développer une argumentation avec esprit critique

**ENSEIGNEMENT – ORGANISATION HORAIRES**

<b>Titre de l'enseignement</b>	<b>CM</b>	<b>CI</b>	<b>TD</b>	<b>TP</b>
<b>Projet transdisciplinaire</b>				
<b>Eq. HETD=</b>				

**MODALITES D'EVALUATION DES ENSEIGNEMENTS**

(session initiale et de rattrapage) évaluation continue intégrale

Etudiants en situation spécifique : modalités d'évaluation continue adaptées à la situation

<b>C-1</b>	<b>Formation de rattachement</b>	<b>Intitulé de l'enseignement</b>
	<b>Parcours Approche Interdisciplinaire en Science des Données</b>	<b>Concepts, enjeux et transdisciplinarité</b>

**RESPONSABLE DE L'ENSEIGNEMENT**

<b>Nom/Prénom</b>	<b>Discipline</b>	<b>Composante / Courriel/Tél. (fixe)</b>
LACHICHE NICOLAS	Informatique	IUT Robert Schuman, <a href="mailto:nicolas.lachiche@unistra.fr">nicolas.lachiche@unistra.fr</a> , 54577
KIEFFER BRUNO	Biologie	ESBS, <a href="mailto:bruno.kieffer@unistra.fr">bruno.kieffer@unistra.fr</a> , 54722

**MODALITES D'ORGANISATION DE L'ENSEIGNEMENT**

Semestre 1

2h étudiant par semaine, en présentiel (en visio en cas de distanciel)

Supports de cours et examens sur moodle

**LANGUE ET NIVEAU DE LANGUE DE L'ENSEIGNEMENT : FRANÇAIS****DESCRIPTION DE L'ENSEIGNEMENT**

- Principaux défis et enjeux
- Les principaux paradigmes des données et des traitements
- Vision transdisciplinaire (séminaires, présentation des projets)

**COMPETENCES VISEES****Objectifs en termes de connaissances**

- Connaître le paysage général -transdisciplinaire- et le vocabulaire de la science des données

**Objectifs en termes de compétences**

- Comprendre les enjeux sociétaux de la science des données
- Comprendre les enjeux et les limites de l'utilisation de l'IA

**ENSEIGNEMENT – ORGANISATION HORAIRE**

<b>Titre de l'enseignement</b>	<b>CM</b>	<b>CI</b>	<b>TD</b>	<b>TP</b>
<b>Concepts, enjeux, et transdisciplinarité</b>	4	20		4
<b>Eq. HETD=33,7</b>				

**MODALITES D'EVALUATION DES ENSEIGNEMENTS**

(session initiale et de rattrapage) évaluation continue intégrale

Etudiants en situation spécifique : modalités d'évaluation continue adaptées à chaque situation

<b>C-2</b>	<b>Formation de rattachement</b>	<b>Intitulé de l'enseignement</b>
	<b>Parcours Approche Interdisciplinaire en Science des Données</b>	<b>Éthique et droit</b>

**RESPONSABLE DE L'ENSEIGNEMENT**

<b>Nom/Prénom</b>	<b>Discipline</b>	<b>Composante / Courriel/Tél. (fixe)</b>
MARC-ZWECKER STELLA	Informatique	UFR math-info, <a href="mailto:stella@unistra.fr">stella@unistra.fr</a> , 50243
DELTORN JEAN-MARC	Droit	CEIPI, <a href="mailto:jm.deltorn@unistra.fr">jm.deltorn@unistra.fr</a> , 56938
MACREZ FRANCK	Droit	CEIPI, <a href="mailto:franck.macrez@ceipi.edu">franck.macrez@ceipi.edu</a> , 58800

**MODALITES D'ORGANISATION DE L'ENSEIGNEMENT**

Semestre 2  
 2h étudiant par semaine, en présentiel (en visio en cas de distanciel)  
 Supports de cours et examens sur moodle

**LANGUE ET NIVEAU DE LANGUE DE L'ENSEIGNEMENT : FRANÇAIS****DESCRIPTION DE L'ENSEIGNEMENT**

- Vie privée
- Anonymisation
- Droit

**COMPETENCES VISEES****Objectifs en termes de connaissances**

- Connaître les licences de réutilisation des données
- Comprendre les enjeux éthiques liés à la collecte et à l'exploitation des données

**ENSEIGNEMENT – ORGANISATION HORAIRE**

<b>Titre de l'enseignement</b>	<b>CM</b>	<b>CI</b>	<b>TD</b>	<b>TP</b>
<b>Concepts, enjeux, et transdisciplinarité</b>	4	20		4
<b>Eq. HETD=33,7</b>				

**MODALITES D'EVALUATION DES ENSEIGNEMENTS**

(session initiale et de rattrapage) évaluation continue intégrale

Etudiants en situation spécifique : modalités d'évaluation continue adaptées à chaque situation

<b>D-1</b>	<b>Formation de rattachement</b>	<b>Intitulé de l'enseignement</b>
	<b>Parcours AISD (Approches Interdisciplinaires en science des données)</b>	<b>Préparation et exploration des données</b>

**RESPONSABLE DE L'ENSEIGNEMENT**

<b>Nom/Prénom</b>	<b>Discipline</b>	<b>Composante / Courriel/Tél. (fixe)</b>
CHAPUIS YVES	Génie électrique, électronique, photonique et systèmes	Faculté de Physique et Ingénierie / <a href="mailto:ya.chapuis@unistra.fr">ya.chapuis@unistra.fr</a> / 03 88 10 63 29
DOLQUES XAVIER	Informatique	UFR Math-Info / <a href="mailto:dolques@unistra.fr">dolques@unistra.fr</a>

**MODALITES D'ORGANISATION DE L'ENSEIGNEMENT**

- Semestre d'automne du Master 1
- 2h étudiant par semaine en présentiel (visioconférence en cas de nécessité du distanciel).
- Supports de cours et corrections des exercices et évaluation sur Moodle
- Utilisation de l'application web open source Jupyter Notebook

**LANGUE ET NIVEAU DE LANGUE DE L'ENSEIGNEMENT : FRANÇAIS (B2 MINIMUM)****DESCRIPTION DE L'ENSEIGNEMENT****1. Préparation des données**

La « préparation des données » sera abordée pour maîtriser les opérations de collecte, sélection, nettoyage et transformation qui doivent être appliquées aux données brutes avant leur traitement de visualisation et analyse. Une phase d'initiation au langage Python et un apprentissage des packages pandas et numpy seront dispensée en amont de cet enseignement.

**2. Visualisation des données**

La « visualisation des données » sera enseignée pour aborder l'exploration visuelle et interactive de données de toutes volumétries, natures (structurées ou non structurées) et origines, et leur représentation graphique. Les étudiants aborderont ainsi les notions de séries temporelles et l'apprentissage des packages matplotlib, plotlib, cufflinks et seaborn.

**3. Cas pratiques**

Finalement, les étudiants mettront leurs compétences à contribution pour la visualisation et l'analyse de cas pratiques de jeu de données.

**COMPETENCES VISEES****Objectifs en termes de connaissances**

- Connaître les principaux formats standards de données.
- Comprendre les caractéristiques des représentations numériques et graphiques des données.

**Objectifs en termes de compétences**

- Mettre en forme des données en vue d'une analyse automatisée.
- Choisir une représentation graphique adaptée au type de données.

- |  |
|--|
| - Utiliser des critères statistiques simples, adaptés au volume et à la nature des données pour en évaluer la fiabilité. |
|--|

**ENSEIGNEMENT – ORGANISATION HORAIRE**

Titre de l'enseignement	CM	CI	TD	TP
<b>Préparation et exploration des données</b>	6		8	14
<b>Eq. HETD= 26,3</b>				

**MODALITES D'EVALUATION DES ENSEIGNEMENTS**

(session initiale et de rattrapage) évaluation continue intégrale

**Contrôle continu :**

2 épreuves sous forme écrite et 1 épreuve sous forme pratique

**Étudiants en situation spécifique :**

- Modalités d'évaluation continue adaptées à la situation
- Des exercices alternatifs pourront être proposés aux étudiants rencontrant des soucis d'assiduité.

D-2	Formation de rattachement	Intitulé de l'enseignement
	Parcours AISD (Approches Interdisciplinaires en science des données)	Outils statistiques

### Responsables de l'enseignement

Nom/Prénom	Discipline	Composante / Courriel/Tél. (fixe)
MAUMY-BERTRAND Myriam	Science des données	UTT-LIST3N/myriam.maumy@utt.fr/03.25.71.76.86
DOLQUES Xavier	Informatique	Unistra-ICUBE/dolques@unistra.fr/03.68.85.02.30

### Modalités d'organisation de l'enseignement

Semestre et UE de rattachement et organisation calendaire envisagée

Usage de l'enseignement à distance, ou autres outils numériques

Semestre 3 du parcours AISD – UE de rattachement : Données : de la collecte aux traitements

### Langue et niveau de langue de l'enseignement : Français (B2 minimum)

### Description de l'enseignement

Plan du cours

- Initiation au logistique libre R
- Statistique exploratoire
- Corrélations linéaire et non linéaire et tests statistiques adaptés à ces corrélations
- Sélection de variables par ACP ou ACM ou par sélection de modèles
- Classification
- Analyse discriminante linéaire

Les séances de TP avec le logiciel R seront l'occasion de mettre en œuvre les méthodes vues en cours et en TD sur des données en liens forts avec les domaines d'applications des étudiants.

Bibliographie éventuelle :

#### Initiation à la statistique avec R - 3e édition

Frédéric Bertrand Myriam Maumy-Bertrand

Collection : Sciences Sup, Dunod, 2018

#### Modélisation statistique par la pratique avec R

Frédéric Bertrand, Emmanuelle Claeys, Myriam Maumy-Bertrand

Collection : Sciences Sup, Dunod, 2019

### Compétences visées

#### objectifs en termes de connaissances

- Connaitre le vocabulaire de la statistique
- Connaitre les outils de la statistique exploratoire
- Connaitre les analyses statistiques classiques

#### **Objectifs en termes de compétences**

Être capable de :

- D'identifier des liaisons entre les variables des deux types (quantitatifs/qualitatifs)
- Réduire les variables d'un jeu de données
- Prédire un groupe d'appartenance à partir des caractéristiques données par le jeu de données

#### **Enseignement – organisation horaire**

Titre de l'enseignement	CM	CI	TD	TP
<b>Représentation et traitement des données</b>	6		8	14
<b>Eq. HETD= 26,33</b>				

#### **MODALITES D'EVALUATION DES ENSEIGNEMENTS**

(Session initiale et de rattrapage)

Exercices d'application durant les séances au cours du semestre. Pas de convocation.

Pas de session de rattrapage.

Étudiants en situation spécifique : les supports de cours et les corrections des exercices seront disponibles sur Moodle. Des exercices alternatifs pourront être proposés aux étudiants rencontrant des soucis d'assiduité.

## DESCRIPTIF DE L'UE

<b>D-3</b>	<b>Formation de rattachement</b>	<b>Intitulé de l'enseignement</b>
	<b>Parcours AISD (Approches Interdisciplinaires en science des données)</b>	<b>Représentation et traitement des données</b>

### RESPONSABLE DE L'ENSEIGNEMENT

<b>Nom/Prénom</b>	<b>Discipline</b>	<b>Composante / Courriel/Tél. (fixe)</b>
BERNHARD DELPHINE	Informatique	Faculté des Langues / dbernhard@unistra.fr/+33 3 68 85 66 29
MEILLIER CELINE	Traitement des images et des signaux - Statistiques	Télécom Physique Strasbourg / meillier@unistra.fr/+33 3 68 85 45 70

### MODALITES D'ORGANISATION DE L'ENSEIGNEMENT

Semestre et UE de rattachement et organisation calendaire envisagée

Usage de l'enseignement à distance, ou autres outils numériques

Semestre 3 du parcours AISD – UE de rattachement : Données : de la collecte aux traitements

### LANGUE ET NIVEAU DE LANGUE DE L'ENSEIGNEMENT : FRANÇAIS (B2 MINIMUM)

### DESCRIPTION DE L'ENSEIGNEMENT

Contenu, plan du cours, bibliographie éventuelle

- Accès aux données : bases de données versus données non structurées, *comprendre les différences et comment cela impacte les traitements à mettre en œuvre.*
- Notions de traitement d'images (*quelques exemples de traitement à partir d'exemples d'IA utilisant les images : vision par ordinateur, reconnaissance d'objets ou de caractères, etc.*)
- Notions d'analyse de langage naturel/texte : types de données textuelles, pré-traitements spécifiques, enrichissement des données à l'aide d'outils de TAL (Traitement Automatique des Langues)
- Données géolocalisées
- Analyse de séries temporelles

Les séances de TP seront l'occasion de mettre en œuvre les méthodes vues en cours et en TD sur des données en liens forts avec les domaines d'applications des étudiants.

### COMPETENCES VISEES

#### Objectifs en termes de connaissances

- Savoir exploiter des données non structurées et structurées
- Connaître les traitements de base pour des données de différents types (images, textes, série temporelle, données géolocalisées)
- Savoir utiliser les outils informatiques pour mettre en œuvre ces traitements dans divers contextes applicatifs

**Objectifs en termes de compétences**

A l'issue de cet enseignement l'étudiant devra être capable de :

- effectuer des traitements élémentaires d'images et de texte
- manipuler des données localisées dans le temps et l'espace
- organiser, étudier et synthétiser des données de divers types sous forme de résultats exploitables

**ENSEIGNEMENT – ORGANISATION HORAIRE**

Titre de l'enseignement	CM	CI	TD	TP
<b>Représentation et traitement des données</b>	6		8	14
<b>Eq. HETD= 26,33</b>				

**MODALITES D'EVALUATION DES ENSEIGNEMENTS**

(session initiale et de rattrapage)

Exercices d'application durant les séances au cours du semestre. Pas de convocation.

Pas de session de rattrapage.

Étudiants en situation spécifique :

Les supports de cours et les corrections des exercices seront disponibles sur Moodle. Des exercices alternatifs pourront être proposés aux étudiants rencontrant des soucis d'assiduité.

<b>M-1</b>	<b>Formation de rattachement</b>	<b>Intitulé de l'enseignement</b>
	<b>Parcours AISD (Approches Interdisciplinaires en science des données)</b>	<b>Méthodes d'apprentissage supervisé</b>

**RESPONSABLE DE L'ENSEIGNEMENT**

<b>Nom/Prénom</b>	<b>Discipline</b>	<b>Composante / Courriel/Tél. (fixe)</b>
LACHICHE NICOLAS	Informatique	IUT Robert Schuman / <a href="mailto:nicolas.lachiche@unistra.fr">nicolas.lachiche@unistra.fr</a> / 54577
DOLQUES XAVIER	Informatique	UFR Math-Info / <a href="mailto:dolques@unistra.fr">dolques@unistra.fr</a>
MEILLIER CÉLINE	Traitement du signal et des images	Telecom Physique Strasbourg / <a href="mailto:meillier@unistra.fr">meillier@unistra.fr</a>

**MODALITES D'ORGANISATION DE L'ENSEIGNEMENT**

Semestre d'automne Master 1  
 2h étudiant par semaine  
 Supports de cours et examens sur moodle

**LANGUE ET NIVEAU DE LANGUE DE L'ENSEIGNEMENT : FRANÇAIS (B2 MINIMUM)****DESCRIPTION DE L'ENSEIGNEMENT**

- Paradigme de l'apprentissage et de la fouille de données
- Apprentissage et classification supervisée : Arbres de décision, SVM

**COMPETENCES VISEES****Objectifs en termes de connaissances**

- Connaître les différents paradigmes de l'apprentissage
- Comprendre les principales méthodes d'apprentissage supervisé

**Objectifs en termes de compétences**

- Savoir utiliser les principales méthodes d'apprentissage supervisé et non-supervisé

**ENSEIGNEMENT – ORGANISATION HORAIRE**

<b>Titre de l'enseignement</b>	<b>CM</b>	<b>CI</b>	<b>TD</b>	<b>TP</b>
<b>Méthodes d'apprentissage</b>	6		8	14
<b>Eq. HETD= 26,3</b>				

**MODALITES D'EVALUATION DES ENSEIGNEMENTS**

(session initiale et de rattrapage) évaluation continue intégrale

Etudiants en situation spécifique : modalités d'évaluation continue adaptées à chaque situation

Les supports de cours et les corrections des exercices seront disponibles sur Moodle. Des exercices alternatifs pourront être proposés aux étudiants rencontrant des soucis d'assiduité.

<b>M-2</b>	<b>Formation de rattachement</b>	<b>Intitulé de l'enseignement</b>
	<b>Parcours AISD (Approches Interdisciplinaires en science des données)</b>	<b>Méthodes d'apprentissage non supervisé et d'apprentissage profond</b>

**RESPONSABLE DE L'ENSEIGNEMENT**

<b>Nom/Prénom</b>	<b>Discipline</b>	<b>Composante / Courriel/Tél. (fixe)</b>
LACHICHE NICOLAS	Informatique	IUT Robert Schuman / <a href="mailto:nicolas.lachiche@unistra.fr">nicolas.lachiche@unistra.fr</a> / 54577
DOLQUES XAVIER	Informatique	UFR Math-Info / <a href="mailto:dolques@unistra.fr">dolques@unistra.fr</a>
MEILLIER CÉLINE	Traitement du signal et des images	Telecom Physique Strasbourg / <a href="mailto:meillier@unistra.fr">meillier@unistra.fr</a>

**MODALITES D'ORGANISATION DE L'ENSEIGNEMENT**

Semestre de printemps Master 1.

2h étudiant par semaine

Supports de cours et examens sur moodle

**LANGUE ET NIVEAU DE LANGUE DE L'ENSEIGNEMENT : FRANÇAIS (B2 MINIMUM)****DESCRIPTION DE L'ENSEIGNEMENT**

Cet UE vise à faire connaître et savoir mettre en œuvre des méthodes d'apprentissage non-supervisé (clustering, association, motifs fréquents) et d'apprentissage profond (connaître la structure générale des réseaux de neurones et intérêt des réseaux de neurones convolutifs). Dans cette UE les étudiants seront amenés à adapter et appliquer ces différentes méthodes sur des données d'applications variées.

**COMPETENCES VISEES****Objectifs en termes de connaissances**

- Connaître les différents paradigmes de l'apprentissage non supervisé
- Comprendre les principales méthodes d'apprentissage non-supervisé et profond

**Objectifs en termes de compétences**

- Savoir utiliser les principales méthodes d'apprentissage non-supervisé et profond

**ENSEIGNEMENT – ORGANISATION HORAIRE**

<b>Titre de l'enseignement</b>	<b>CM</b>	<b>CI</b>	<b>TD</b>	<b>TP</b>
<b>Méthodes d'apprentissage</b>	6		8	14
<b>Eq. HETD= 26,3</b>				

**MODALITES D'EVALUATION DES ENSEIGNEMENTS**

(session initiale et de rattrapage) évaluation continue intégrale

Etudiants en situation spécifique : modalités d'évaluation continue adaptées à chaque situation

Les supports de cours et les corrections des exercices seront disponibles sur Moodle. Des exercices alternatifs pourront être proposés aux étudiants rencontrant des soucis d'assiduité.

<b>M-3</b>	<b>Formation de rattachement</b>	<b>Intitulé de l'enseignement</b>
	<b>Parcours AISD (Approches Interdisciplinaires en science des données)</b>	<b>Méthodes symboliques et numériques</b>

**RESPONSABLE DE L'ENSEIGNEMENT**

<b>Nom/Prénom</b>	<b>Discipline</b>	<b>Composante / Courriel/Tél. (fixe)</b>
LACHICHE NICOLAS	Informatique	IUT Robert Schuman / <a href="mailto:nicolas.lachiche@unistra.fr">nicolas.lachiche@unistra.fr</a> / 54577
DOLQUES XAVIER	Informatique	UFR Math-Info / <a href="mailto:dolques@unistra.fr">dolques@unistra.fr</a>
MEILLIER CÉLINE	Traitement du signal et des images	Telecom Physique Strasbourg / <a href="mailto:meillier@unistra.fr">meillier@unistra.fr</a>

**MODALITES D'ORGANISATION DE L'ENSEIGNEMENT**

Semestre d'automne Master 2.

2h étudiant par semaine

Supports de cours et examens sur moodle

**LANGUE ET NIVEAU DE LANGUE DE L'ENSEIGNEMENT : FRANÇAIS (B2 MINIMUM)****DESCRIPTION DE L'ENSEIGNEMENT**

Cet UE concerne la modélisation et l'apprentissage à partir de données hétérogènes multisources. Elle aborde en particulier la représentation et l'utilisation de connaissances, par le biais d'ontologies, ou par le partage de données, à l'aide du web sémantique ou du web des données. Dans cette UE les étudiants travailleront par exemple sur l'analyse de séries temporelles complexes, en lien avec l'UE D3.

**COMPETENCES VISEES****Objectifs en termes de connaissances**

- Comprendre et savoir utiliser des modèles de représentation de connaissances

**Objectifs en termes de compétences**

- Savoir représenter et manipuler des données symboliques et numériques

**ENSEIGNEMENT – ORGANISATION HORAIRE**

<b>Titre de l'enseignement</b>	<b>CM</b>	<b>CI</b>	<b>TD</b>	<b>TP</b>
<b>Méthodes d'apprentissage</b>	6		8	14
<b>Eq. HETD= 26,3</b>				

**MODALITES D'EVALUATION DES ENSEIGNEMENTS**

(session initiale et de rattrapage) évaluation continue intégrale

Etudiants en situation spécifique : modalités d'évaluation continue adaptées à chaque situation

Les supports de cours et les corrections des exercices seront disponibles sur Moodle. Des exercices alternatifs pourront être proposés aux étudiants rencontrant des soucis d'assiduité.