

## MASTER - Sciences du langage

### Sciences des données et intelligence artificielle

#### Pré-requis obligatoires

Une licence dans un de domaines suivants : Sciences du Langage, Lettres. Intérêt pour l'informatique et pour la programmation en Python. Niveau en anglais B2.

#### Objectifs du parcours

Ce parcours combine compétences en linguistique et en sciences des données, ouvre de nouvelles opportunités sur le marché de travail pour nos étudiant-es de Sciences du langage.

Les institutions et les entreprises 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 thématiques capables de maîtriser les techniques sur des données variées (numérique, texte, images).

Langue du parcours	Français		
ECTS	120 ECTS		
Volume horaire			
TP : 144h	TD : 216h	CI : 0h	CM : 769h
Formation initiale	Oui		
Formation continue	Non		
Apprentissage	Non		
Contrat de professionnalisation	Non		
Stage : (durée en semaines)	7		

#### Compétences à acquérir

Double compétence en linguistique et sciences de données:

- connaître les notions fondamentales dans les domaines des sciences du langage (morphologie, syntaxe, sémantique, phonétique et phonologie, linguistique romane et diachronique)
- représenter et manipuler des données variées (textuelles, images, son, vidéo)
- connaître des notions fondamentales d'apprentissage automatique (statistique, profond)

#### Codes ROME

- K2401 - Recherche en sciences de l'homme et de la société
- M1806 - Expertise et support technique en systèmes d'information

#### Stage et projet tutoré

Le stage peut avoir lieu en entreprise sur une durée de minimum 7 semaines (ou 210 heures à temps complet) à maximum 6 mois entre janvier et septembre.

#### Contact

Amalia Todirascu : [todiras@unistra.fr](mailto:todiras@unistra.fr)

# Master 1 - Sciences des données et intelligence artificielle

## Semestre 1

	ECTS	CM	CI	TD	TP	TE	Stage
<b>UE1 Éléments transversaux S1</b>	<b>3 ECTS</b>	6 h		6 h			
Séminaire Semascila		12 h					
Méthodologie documentaire		6 h					
<b>UE2 Méthodologie disciplinaire S1</b>	<b>6 ECTS</b>			36 h			
Introduction au TAL		12 h					
Outils d'annotation automatique				12 h			
Anglais scientifique S1				18 h			
Linguistique des corpus (méthodologie des corpus)		12 h					
<b>UE3 Tronc commun disciplinaire S1</b>	<b>12 ECTS</b>	12 h					
Principaux courants en phonétique et phonologie		12 h					
Morphologie S1		12 h					
Syntaxe S1		12 h					
Sémantique S1		12 h					
EC - linguistique romane ou diachronique		12 h					
<b>UE4 Concepts et méthodes S1</b>	<b>6 ECTS</b>						
MATI M1 - Méthodes d'apprentissage supervisé		6 h		8 h	14 h		
MATI C1 - Concepts, enjeux et transdisciplinarité		4 h	20 h		4 h		
<b>UE5 Données et outils S1</b>	<b>3 ECTS</b>						
MATI D1 - Collecte, fiabilité et visualisation des données		6 h		8 h	14 h		

## Semestre 2

	ECTS	CM	CI	TD	TP	TE	Stage
<b>UE1 Éléments transversaux S2</b>	<b>6 ECTS</b>	12 h					
LiLPa S2		6 h					
Méthodologie du mémoire							
MATI A1 - Projet transdisciplinaire 1							
<b>UE2 Méthodologie disciplinaire S2</b>	<b>6 ECTS</b>			12 h			
Exploitation du corpus/XML		12 h					
Anglais scientifique S2				18 h			
Projet tutoré d'exploitation de corpus							
<b>UE3 Tronc commun disciplinaire S2</b>	<b>9 ECTS</b>	12 h					
Organisation spatio-temporelle de la parole		12 h					
Morphologie S2		12 h					
Syntaxe S2		12 h					
Sémantique S2		12 h					
Linguistique romane ou diachronique S2		12 h					
<b>UE 4 Concepts et méthodes S2</b>	<b>6 ECTS</b>						
MATI M2 - Méthodes d'apprentissage non supervisé et d'apprentissage profond		6 h		8 h	14 h		
MATI C2 - Ethique et droit		4 h	20 h		4 h		
<b>UE5 Données et outils S2</b>	<b>3 ECTS</b>						
MATI D2 - Outils statistiques		6 h		8 h	14 h		

## Master 2 - Sciences des données et intelligence artificielle

### Semestre 3

	ECTS	CM	CI	TD	TP	TE	Stage
<b>UE1 Éléments transversaux S3</b>	<b>3 ECTS</b>	<b>14 h</b>					
Séminaires de recherche (1 séminaire au choix)							
Méthodologie du mémoire S3							
MATI A2 - Projet transdisciplinaire 2							
<b>UE2 Méthodologie disciplinaire S3</b>	<b>6 ECTS</b>		<b>24 h</b>	<b>24 h</b>			
Méthodes statistiques pour l'analyse de données textuelles				12 h			
Lexicologie S1 et S3		24 h					
Annotations de haut niveau				12 h			
<b>UE3 Tronc commun disciplinaire S3</b>	<b>12 ECTS</b>	<b>12 h</b>					
Parole et cognition		12 h					
Morphologie S3		12 h					
Syntaxe S3		12 h					
Sémantique S3		12 h					
Linguistique romane ou diachronique S3		12 h					
<b>UE4 Méthodes et représentations S3</b>	<b>6 ECTS</b>						
MATI M3 - Méthodes symboliques et numériques		8 h		8 h	14 h		
MATI D3 - Représentation et traitement des données		6 h		8 h	14 h		
<b>UE5 Connaissance du milieu professionnel S3</b>	<b>3 ECTS</b>	<b>3 h</b>		<b>20 h</b>			
Bilan de compétences				8 h			
Préparation au stage		3 h					

### Semestre 4

	ECTS	CM	CI	TD	TP	TE	Stage
<b>UE1 Mémoire S4</b>	<b>15 ECTS</b>						
Travail personnel de recherche						96 h	
<b>UE2 Stage S4</b>	<b>15 ECTS</b>						
Stage S4							7 sem