Université de Strasbourg

Arts, Lettres, Langues | 2026-2027

DIPLÔME NATIONAL DE MASTER CONTRÔLÉ PAR L'ÉTAT

Master Sciences du langage
Sciences du langage et sciences des données

Présentation

La mention Sciences du langage se situe dans la continuité de la licence Sciences du langage de l'Université de Strasbourg. Cette mention est la seule qui forme des spécialistes dans ce domaine à l'Université de Strasbourg, ouvrant certains de ses enseignements au parcours Technologies des langues (mention Traduction et Interprétation) qui, en échange, lui offre des enseignements spécifiques.

La formation s'appuie sur l'UR LiLPa (1339) et sur la Fédération de Recherche en Langage et communication (FRLC) (devenue Institut Thématique Interdisciplinaire LiRiC-Langage, Inclusion, Rémédiation, Interculturalité, Communication, à partir du 1er janvier 2025). En plus du parcours classique SDL orienté recherche, les parcours SDLSD et Communication des entreprises-RSE sont résolument tournés vers le monde de l'entreprise et répondent aux besoins du monde socio-économique. En ce qui concerne les SDL-SD, les offres relatives au métier d'analyste de données sont très nombreuses et peu de masters proposent une formation de ce type (les autres parcours Sciences de données ne prennent pas les textes comme support de données). En ce qui concerne le parcours M2 Communication des entreprises-RSE, il sera proposé en formation initiale et continue et en alternance. La plupart des secteurs économiques d'Alsace, et plus généralement de France (finances, automobile, prêt à porter, énergie, etc.) sont concernés par la communication RSE et manquent de spécialistes (les offres de stages sont nombreuses mais peu d'étudiants sont actuellement formés à la fois en communication et en RSE).

Les trois parcours préparent l'insertion professionnelle, d'une part par des cours de préparation à la recherche de stage, des bilans de compétences, des stages, optionnels en M1 et obligatoires en M2. Le nouveau parcours M2 Communication des entreprises-RSE facilitera cela d'autant plus que les étudiants alternants seront en situation professionnelle. D'autre part, des cours de méthodologie de la recherche, de méthodologie disciplinaire (notamment par la formation aux outils d'annotation et d'exploitation de données, écriture scientifique en anglais, méthodologie des écrits académiques), ainsi que le mémoire (via un suivi personnalisé et régulier) préparent aux métiers de la recherche, mais aussi développent les capacités de synthèse et d'analyse.

Les parcours SDL et SDL-SD sont construits sur un tronc commun (fondamentaux, méthodologie disciplinaire, anglais scientifique). Le parcours SDL propose des spécialisations en analyse de discours en synchronie et en diachronie dans les thèmes de l'UR LiLPa (lexicologie, analyse des discours normaux et pathologiques, TAL et linguistique outillée, linguistique romane et diachronique, variation). Une initiation à la RSE en vue du nouveau M2 et une option préprofessionalisante "Métiers de la réécriture" sont proposées. Le parcours SDL-SD propose un tronc commun avec les autres parcours orientés Sciences de données (apprentissage automatique supervisé et non supervisé, exploitation de données par des méthodes statistiques, visualisation). En revanche, le parcours M2 Communication des entreprises-RSE sera spécifique du point de vue de son organisation matérielle (une semaine de cours, trois semaines en entreprise) et pédagogique (cours spécifiques regroupés sur des plages dédiées, nombreuses interventions professionnelles).

Objectifs

• 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).
• Les diplômes vont acquérir une double compétence en linguistique et sciences de

Composante	• Faculté des lettres
Langues d'enseignement	• Français
Niveau d'entrée	BAC +3
Durée	2 ans
ECTS	120
Volume global d'heures	710
Formation à distance	Non, uniquement en présentiel
Régime d'études	• FI (Formation initiale)
Niveau RNCP	Niveau 7
RNCP	• RNCP38696 : Master Sciences du langage
Disciplines	• Sciences du langage • Informatique
Lieu	Université de Strasbourg (Le Portique, l'UFR Math-Informatique)
Campus	• Campus Esplanade
Secteurs d'activité	• Recherche-développement scientifique • Édition de logiciels • Activités de contrôle et analyses techniques • Traitement de données, hébergement et activités connexes ; portails Internet
Code ROME	• Chercheur / Chercheuse en sciences humaines et sociales • Product Owner
Stage	Oui
Alternance	Non

Aménagements pour les publics ayant un profil spécifique

Les Règles générales des licences et des masters de la Faculté des Lettres pour l'année 2023-2024, consultables dans leur totalité sur cette page, détaillent les aménagements possibles dans le cadre du Régime spécial d'études (RSE) et les démarches nécessaires pour en bénéficier :

"Dans [le contrat pédagogique], parmi d'autres clauses, le RSE adopté par l'Université (https//unistra.fr/rse) vise à permettre

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)
- Mettre en œuvre des techniques de la science des données pour classer, visualiser, structurer les données
- connaître des notions fondamentales d'apprentissage automatique (statistique, symbolique, profond)
- Recueillir, analyser et décrire des données en utilisant des méthodes, théories et modèles relevant de différents domaines des Sciences du Langage et de la sciences de la donnée
- Développer, se servir de façon autonome des outils et des ressources numériques avancés (outils d'annotation, concordanciers, bases lexicales, terminologiques etc) pour un projet professionnel ou de recherche en Sciences du langage
- Identifier et décrire la diversité des usages langagiers, la production du langage du point de vue du message et du contexte de sa production
- Construire un projet de recherche en sciences du langage ou un mémoire professionnel, en respectant les principes éthiques et déontologiques
- Communiquer à des fins de formation ou de transfert de connaissances, par oral et par écrit, en français et dans au moins une langue étrangère
- Construire et mener à bien un projet professionnel

Insertion professionnelle

le stage obligatoire en M2 aide à consolider le projet professionnel et souvent à trouver un emploi

Métiers visés

- Data scientist/Analyste de données

– Analyste thématique

– Consultant linguiste

– Linguiste informaticien

– Annotateur/Annotatrice

- Ingénieur d'études

Les diplômés du parcours SDL-SD peuvent travailler dans **les domaines suivants: informatique, systèmes d'information, TAL, Ingénierie/industrie de la langue : (développement des outils de dialogue homme machine; d'extraction d'informations; de veille technologique et documentaire, outils pur la fouille de textes, outils d'annotation)**

Pour connaitre en détail l'insertion professionnelle de nos diplômés, consultez cette page.

Les + de la formation

Il s'agit d'une formation interdisciplinaire, permettant de combiner connaissances théoriques et appliquées dans le domaine des Sciences du langage, du Traitement automatique des langues et du domaine des Sciences de la donnée. Les diplômés maîtrisent des méthodes par apprentissage automatique, des outils statistiques et de visualisation de données, et manipulent et structurent des données très variées (textes, signaux, vidéos, images etc.). Les étudiants peuvent acquérir une culture interdisciplinaire en travaillant sur des projets avec des étudiants d'autres cursus similaires, en s'appropriant données et approches qui ne sont spécifiques à leur domaine de compétences.

Critères de recrutement

Le niveau Bac+3, une licence en Sciences du langage, Lettres, Sciences humaines ou d'autres domaines, avec solides connaissances dans les domaines fondamentaux (morpho-syntaxe, syntaxe, sémantique, pragmatique, phonétique). Un niveau minimum C1 en français (pour les étudiants étrangers), un niveau B2 en anglais : certificats demandés. La maîtrise d'un langage de programmation (de préférence Python) est souhaitée.

Candidater

Pour consulter les modalités de candidature, consultez la page dédiée sur le site de l'Université de Strasbourg.

Prérequis obligatoires

Une licence dans un de domaines suivants : Sciences du Langage, Lettres. Niveau en anglais B2. Solides connaissances en linguistique et solide maitrise de la langue française, à l'oral et à l'écrit. Projet professionnel cohérent, en conformité avec le contenu de la formation. Pour les candidats étrangers, un certificat

de proposer des accompagnements en fonction des besoins particuliers de l'étudiant·e. Dans le cadre de la mise en œuvre de ce régime, la formation propose les aménagements suivants :
– Aménagement des horaires d'enseignement : l'étudiant·e bénéficie d'un choix facilité pour des groupes de TD compatibles avec ses contraintes
– Dispense d'assiduité et dispense de contrôle continu : l'étudiant·e bénéficie d'une dispense de présence aux enseignements et aux épreuves sans convocation, dans les conditions fixées aux rubriques Assiduité (pour la dispense d'assiduité) et Absence aux épreuves (pour la dispense de contrôle continu)
– Attribution d'un régime long d'études : l'étudiant·e bénéficie d'un étalement des études sur plusieurs années avec une dispense des limitations du nombre d'inscriptions
– Régime spécifique de conservation des notes : l'étudiant·e bénéficie de la conservation des notes au niveau de matières spécifiées, si le contrat pédagogique le prévoit
L'étudiant qui souhaite bénéficier d'un régime spécial d'études dans son contrat pédagogique dépose sa demande par formulaire, avec les pièces justificatives pertinentes, auprès de son service de scolarité. Cette demande est déposée au plus deux semaines après le début des cours du semestre concerné, ou dans les meilleurs délais après la survenance du fait pouvant justifier ce régime d'études. Au delà de quatre semaines après le début des cours du semestre concerné, les aménagements concernant les modalités d'examen que permettrait la formation ne sont plus possibles pour le semestre en cours.
En cas de refus d'octroi de régime spécial d'études, l'étudiant peut contester cette décision par une demande écrite et motivée déposée auprès de la Direction des études et de la scolarité.
Le contrat pédagogique peut être pluriannuel. Ses prévisions sont réexaminées, et éventuellement adaptées, en fonction de l'évolution de la situation de l'étudiant·e et de la formation."

Droits de scolarité

Pour consulter les droits de scolarité, consultez la page dédiée sur le site de l'Université de Strasbourg.

Contacts

Responsable(s) de parcours

- Amalia Todirascu

Autres contacts

Scolarité : formulaire de contact

attestant le niveau de français : C1 minimum requis. Anglais : niveau B2 minimum requis. Maîtrise des outils informatique de base (Word, Excel). La maîtrise d'un langage de programmation (Python) est nécessaire.
Il faut déposer un dossier de candidature sur la plateforme e-candidat (https://ecandidat.unistra.fr/) : CV, lettre de motivation qui montre votre intérêt pour cette formation, les relevés de notes, les dernières diplômes obtenus. Ouverture de la plateforme mars 2023.

Prérequis recommandés

Pré-réquis et admissions Le parcours est ouvert aux étudiants de licence en sciences humaines et sociales (sciences du langage, lettres, langues) qui ont des solides connaissances en linguistique. La connaissance des notions de base en programmation Python ou le suivi d'une UE d'initiation aux sciences de données (en L3, UE6) sont fortement recommandés, ainsi que le niveau B2 en anglais. Pour les candidats étrangers : le niveau minimum C1 en français est nécessaire.

Présentation et organisation de l'équipe pédagogique

Responsable pédagogique: Mme Amalia Todirascu

Programme des enseignements

Sciences du langage et sciences des données

Master 1 Sciences du langage - Sciences du langage et sciences des données

Semestre 1

		CM	TD	TP	CI
UE1 Méthodologie de la recherche	3 ECTS	-	-	-	-
Séminaire LiLPa		-	6h	-	-
Séminaire SéMaSciLa		-	6h	-	-
Méthodologie documentaire		6h	-	-	-
Introduction à la RSE		12h	-	-	-
UE2 Traitement de corpus et de données	9 ECTS	-	-	-	-
Introduction au TAL		12h	-	-	-
Outils d'annotation automatique		-	12h	-	-
Anglais scientifique S1		-	12h	-	-
Collecte de données écrites et outils d'exploitation de corpus écrits		12h	-	-	-
Collecte de données orales et outils d'exploitation de corpus oraux		12h	-	-	-
UE3 Fondements théoriques de l'analyse de discours S1	9 ECTS	-	-	-	-
Aspects phonétiques, phonologiques et prosodiques des genres oraux		12h	-	-	-
Aspects morphosyntaxiques et lexicaux des genres écrits		12h	-	-	-
Aspects sémantiques et pragmatiques des genres écrits		12h	-	-	-
Aspects syntaxiques des genres écrits		12h	-	-	-
Aspects diachroniques des genres écrits		12h	-	-	-
UE4 Concepts et methodes	6 ECTS	-	-	-	-
MATI M1 - Méthodes d'apprentissage supervisé		6h	8h	14h	-
MATI C1 - Concepts, enjeux et transdisciplinarité		24h	-	4h	-
UE5 Données et outils S1	3 ECTS	-	-	-	-
MATI D1 - Collecte, fiabilité et visualisation des données		6h	8h	14h	-

Semestre 2

		CM	TD	TP	CI
UE1 Méthodologie de la recherche	6 ECTS	-	-	-	-
LiLPa S2		-	6h	-	-
Méthodologie du mémoire		-	12h	-	-
MATI A1 - Projet transdisciplinaire 1		-	28h	-	-
Collecte de données : aspects éthiques et juridiques		6h	-	-	-
UE2 Traitement de corpus et de données	6 ECTS	-	-	-	-

		CM	TD	TP	CI
Structuration de données XML		-	12h	-	-
Anglais scientifique S2		-	12h	-	-
Analyse qualitative de données		12h	-	-	-
Édition de textes anciens		12h	-	-	-
UE3 Analyse de discours : méthodes et approches S2	9 ECTS	-	-	-	-
Approches par les genres oraux : Aspects phonétiques, phonologiques et prosodiques S2		12h	-	-	-
Approches par les genres écrits contemporains : Aspects morphosyntaxiques et lexicaux S2		12h	-	-	-
Approches par les genres écrits contemporains : Aspects sémantiques et pragmatiques S2		12h	-	-	-
Approches par les genres écrits contemporains : Aspects syntaxiques S2		12h	-	-	-
Approche des genres écrits anciens : aspects diachroniques S2		12h	-	-	-
UE4 Concepts et méthodes	6 ECTS	-	-	-	-
MATI M2 - Méthodes d'apprentissage non supervisé et d'apprentissage profond		6h	8h	14h	-
MATI C2 - Éthique et droit		4h	-	4h	20h
UE5 Données et outils S2	3 ECTS	-	-	-	-
MATI D2 - Outils statistiques		6h	8h	14h	-

Master 2 Sciences du langage - Sciences du langage et sciences des données

Semestre 3

		CM	TD	TP	CI
UE1 Éléments transversaux	6 ECTS	-	-	-	-
Séminaire SéMaSciLa		-	6h	-	-
MATI A2 - Projet transdisciplinaire 2		-	28h	-	-
Méthodologie du mémoire		-	12h	-	-
UE2 Traitement de corpus et de données	6 ECTS	-	-	-	-
Méthodes statistiques pour l'analyse de données textuelles		-	12h	-	-
Annotations de haut niveau corpus écrits		-	12h	-	-
Annotations des corpus oraux		-	12h	-	-
UE3 Analyse de discours : méthodes et approches S3	9 ECTS	-	-	-	-
Aspects phonétiques, phonologiques et prosodiques des genres oraux		12h	-	-	-
Aspects morphosyntaxiques et lexicaux des genres écrits		12h	-	-	-
Aspects sémantiques et pragmatiques des genres écrits		12h	-	-	-
Aspects syntaxiques des genres écrits		12h	-	-	-
Aspects diachroniques des genres écrits		12h	-	-	-
UE4 Méthodes et représentations	6 ECTS	-	-	-	-
MATI M3 - Méthodes symboliques et numériques		8h	8h	14h	-
MATI D3 - Représentation et traitement des données		6h	8h	14h	-
UE5 Connaissance du milieu professionnel	3 ECTS	-	-	-	-

	CM	TD	TP	CI
Préparation au stage	3h	-	-	-
Projet tutoré d'exploitation de corpus	-	-	-	-
Expression scientifique orale en anglais	-	12h	-	-

Semestre 4

	CM	TD	TP	CI
UE1 Mémoire S4 15 ECTS	-	-	-	-
Travail personnel de recherche	-	-	-	-
UE2 Stage S4 15 ECTS	-	-	-	-
Stage S4	-	-	-	-