



Master LITL



Domaine : Sciences humaines et sociales

Mention : Sciences du langage

Parcours : **Linguistique, Informatique, Technologies du Langage**

Principaux membres de l'équipe enseignante :
Cécile Fabre (resp.), Lydia-Mai Ho-Dac, Ludovic Tanguy

Plan de la présentation

- Objectifs du master
- Secteurs d'activité et types d'emplois
- Organisation et contenus
- Projets, mémoires
- Intégrer le master

Objectifs du master LITL

- Le master LITL associe **linguistique et informatique** pour former des spécialistes du **traitement automatique des langues (TAL)**
- Le TAL est un domaine qui vise à :
 - créer des outils capables de traiter de façon automatique des données langagières, et d'assister de nombreuses tâches nécessitant l'analyse du langage :
traduire, répondre à des questions, rechercher de l'information, détecter des connaissances nouvelles, des opinions, faciliter la gestion de données volumineuses (rapports, mails), etc.
 - répondre aux besoins aujourd'hui très importants d'accès à l'information et aux connaissances qui sont exprimées par des textes
- A la fois un domaine de recherche et un champ professionnel à l'importance croissante

Objectifs du master LITL

Le master LITL comporte en M2 une formation complémentaire à l'**ergonomie cognitive** : prise en compte des besoins utilisateurs et des usages, et intègre des **enseignements de professionnalisation**

Les diplômés du master LITL sont capables de développer des *ressources linguistiques* et de recourir à des *traitements informatiques* pour concevoir des systèmes de TAL

ET

ils sont sensibilisés à la prise en compte des situations de travail concrètes, des usages dans lesquels ces dispositifs vont s'intégrer

En M2 : des enseignements partagés avec les étudiants du M2 ECIT-FH de psychologie

Environnement



La formation s'appuie sur

- un laboratoire de recherche CNRS (**CLLE : Cognition, Langage, Langues, Ergonomie**) associant des chercheurs en linguistique et TAL (**CLLE-ERSS, axe CARTEL**) et en ergonomie cognitive (**CLLE-LTC**)
- le tissu industriel toulousain et un réseau professionnel (Airbus, Atchik, CNES, CFH, Inbenta, Orange, Prometil, Synapse, ...)
- un réseau d'anciens étudiants

Les secteurs d'activités concernés

- **La gestion et la valorisation des contenus textuels**
 - Rechercher, analyser, transformer l'information
 - Produire des documents de meilleure qualité, plus faciles à exploiter
- **Dans tous les domaines**
 - Gestion de la communication client
 - Gestion des risques
 - Veille technologique, commerciale, scientifique
 - Marketing, référencement web
 - Gestion de la documentation technique et des connaissances métier

Exemples d'emplois occupés par des diplômés du master

- *Dans des petites entreprises (startup, cabinet de conseil, société spécialisée en TAL) :*
 - *Ingénieur linguiste (Besedo, Paris)*
 - *Ingénieure linguiste (Syllabs, Paris)*
 - *Project Manager (Proxem, Paris)*
 - *Data scientist (Hubware, Toulouse)*
 - *Knowledge Manager (Synapse Développement, Toulouse)*
 - *NLP Engineer (Maluuba, Canada)*
 - *Linguiste (Inbenta, Toulouse)*
 - *Linguiste informaticienne (Safety Data - CFH, Toulouse)*
 - *Ingénieur linguiste informaticienne (ShareWizMe, Paris)*
 - *Linguist & Research engineer in Natural Language Processing (Prometil, Toulouse)*
 - *NLP specialist, Linguamatics, Cambridge UK*
- *Dans des grandes entreprises :*
 - *Terminologue (Airbus, Toulouse)*
 - *Ads Quality Associate in French (Google, Dublin)*
- *Dans des organismes publics :*
 - *Doctorante, contrat doctoral, CLLE, Toulouse*
 - *Doctorante, contrat CIFRE, Prometil & CLLE, Toulouse*
 - *Doctorant, financement ANR, ICAR, Lyon*
 - *Ingénieur d'études à l'INRA (Paris)*
 - *Ingénieur d'études au CNRS (Nancy)*
 - *Ingénieur pédagogie et TICE, COMUE de Toulouse*
 - *Responsable Numérique et Poursuite de carrière, Univ. de Toulouse*

Nature des activités professionnelles

- Développer des ressources linguistiques pour améliorer un système de TAL
 - Lexiques, grammaires, terminologies, ontologies
 - Traitement de corpus issus de domaines d'activité divers
- Evaluer des systèmes de TAL
 - Analyse de sorties de systèmes, évaluation, diagnostic
 - Recommandations d'amélioration en termes de traitements linguistiques
- Proposer des traitements innovants et adaptés à des besoins variés
 - Adaptation de techniques existantes à un nouveau domaine
 - Intégration de techniques plus riches en réponse à des demandes
- Etre expert linguiste sur toutes les questions liées au langage
 - Personne ressource pour les traitements linguistiques
 - Chef de projet, responsable recherche & développement
 - Polyvalence (suivi du développement des logiciels, tests, contacts clients/fournisseurs/équipes techniques, analyse des besoins...)

Exemples

- E. est **product manager junior chez Proxem (Paris)** : elle travaille sur des projets de réalisation d'agents conversationnels
- L. est **ingénieure linguiste à Besedo (Paris)** : elle travaille sur le développement de ressources et d'outils pour aider le travail de modération de contenu
- B. est **NLP engineer chez Maluuba (Canada)** : elle développe et améliore un système de dialogue oral en français sur les smartphones
- S. est **ingénieure linguiste informaticienne et responsable R&D chez Sharewizme (Lille)** : elle développe et utilise des systèmes informatiques qui analysent en temps réel les réponses des participants pour animer des séminaires
- L. est **terminologue chez Airbus (Toulouse)** : elle inventorie et organise les termes techniques utilisés dans l'entreprise et les utilise pour améliorer l'accès à la documentation
- C. est **ingénieure linguiste chez Inbenta (Toulouse)** : en lien avec les clients, elle développe des ressources langagières spécifiques pour adapter le moteur de recherche à leurs activités
- N. est **linguiste informaticien chez CFH/Safety Data (Toulouse)** : il y a réalisé son doctorat (convention CIFRE) dans lequel il a proposé un système innovant pour aider les experts en sécurité aérienne à analyser les rapports d'accidents
- F. a été **chercheur post-doctorant au CNRS (Paris)** : il travaillait sur un système qui analyse les forums de santé pour y identifier les effets indésirables des médicaments. Depuis il travaille dans la société **Linguamatics**, en tant que **NLP specialist**, à Cambridge.

Organisation de la formation

M1

300h

Enseignements spécifiques : informatique et TAL

Projet collectif tutoré

Enseignements communs avec LICOCO : linguistique et méthodologie des SDL

Mémoire de recherche

Langues

M2

350h

Programmation et TAL

Ergonomie

Enseignements de professionnalisation

Linguistique

Langues

Projet collectif tutoré

Stage long en entreprise

Période d'enseignement : sept-février

Stage de 4 à 6 mois ou mémoire de recherche

Détail des enseignements du M1

LITL

UE801 SLT0801V Mémoire M1 12 ECTS 25h	UE802 SLT0802V Ressources ling. pour le TAL 3 ECTS 25h	UE803 SLT0803V Informatique pour le TAL 2 3 ECTS 25h	UE804 SLT0804V Projet TAL 3 ECTS 25h	UE805 Option LICOCO 3 ECTS 25h	UE806 SL00806V Collecte de données et statistique 3 ECTS 25h	UE807 Langues 3 ECTS 25h	150h → 12h hebdo
UE701 SLT0701V Projet de recherche 4 ECTS	UE702 SLT0702V Méthodes et applications du TAL 7 ECTS 25h	UE703 SLT0703V Informatique pour le TAL 1 7 ECTS 25h	UE704 Option LICOCO 3 ECTS 25h	UE705 Option LICOCO 3 ECTS 25h	UE706 SL00706V Les données en SDL 3 ECTS 25h	UE707 Langues 3 ECTS 25h	150h → 12h hebdo

Modalités pédagogiques et évaluation

- Beaucoup de travail pratique
- Des évaluations en continu :
 - Exercices d'une séance à l'autre, compte-rendus, fiches de lecture...
- Travail personnel régulier en particulier pour l'apprentissage de la programmation
- Une dimension « promo » importante :
 - Projets collectifs
 - Créneaux d'accès en salle machine pour faciliter le travail en groupe
 - Tutorat des M1 par les M2

Le mémoire

- Une part très importante du travail à fournir dans l'année
 - Travail de longue haleine permettant de mettre en œuvre les compétences acquises
- Un travail visant à répondre à une question de recherche, en combinant :
 - Des traitements informatiques : manipulation de données linguistiques (corpus, lexiques), utilisation d'outils
 - Une analyse linguistique : annotations, descriptions, analyses

Le mémoire

- Sujets de mémoire proposés par l'équipe pédagogique élargie
 - Présentation pendant la réunion de pré-rentree
- Choix du sujet dans les 3 semaines qui suivent la rentrée
- Conseil : communiquer à C. Fabre avant l'été des vœux (mots-clés portant sur les thématiques, objets d'étude, que vous souhaitez privilégier)

Les projets tutorés

- **Projet TAL en M1**
 - Projet en équipe avec accompagnement pédagogique
 - Réalisation d'une chaîne de traitement automatique complète
 - Projets 2015-2016 et 2016-2017 : participation à une campagne d'évaluation internationale
 - Extraction des termes médicaux en français
 - Extraction des causes de décès à partir de rapports médicaux en français et en anglais
 - Projet 2017-2018 : analyse et classification de blogs sur le thème de la discrimination sexiste
 - Projet 2018-2019 : construction de ressources grammaticales à partir de corpus de sous-titres de films pour l'enseignement du français

Les projets tutorés

- **Projet conception en M2**
 - Projet collectif en réponse à une demande émanant d'un terrain professionnel
 - Analyse de la demande, proposition d'intervention et gestion des différentes étapes de conception d'une chaîne de traitement
 - Restitution écrite et orale au commanditaire
 - Ex de projets récents:
 - Evaluation de la lisibilité des articles de l'encyclopédie Wikidia, destinée à des jeunes lecteurs de 8 à 13 ans
 - Aide à la correction d'erreurs pour des apprenants dyslexiques
 - Implémentation d'un Wiki pour favoriser la transmission des connaissances au sein du CNES
 - Faciliter la navigation des utilisateurs du forum « Vivre sans thyroïde »

Intégrer le master LITL 1^{ère} année

- **Critères d'admission**

Le candidat doit :

- être titulaire d'une licence des mentions conseillées ou d'un diplôme de niveau Bac +3 en lien avec le master LITL
- faire état de compétences suffisantes en linguistique
- avoir une bonne maîtrise de l'expression écrite et orale en français (pour les étudiants non francophones : attestation du niveau C1 minimum)
- avoir une maîtrise suffisante de l'anglais pour lire des articles scientifiques dans cette langue
- posséder des compétences générales en informatique du niveau C2i (autonomie dans l'utilisation de l'ordinateur, maîtrise des outils de bureautique).

Le jury appréciera également la qualité du dossier universitaire ainsi que la cohérence du parcours et du projet présentés dans la lettre de motivation.

<http://sciences-du-langage.univ-tlse2.fr/>
Rubrique : MASTERS → LITL