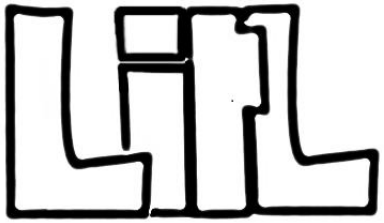
Master LITL

Domaine : Sciences humaines et sociales

Mention : Sciences du langage

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



Cécile Fabre, Mai Ho-Dac**, Ludovic Tanguy** (resp.)

Réunion d'information –2025

http://sciences-du-langage.univ-tlse2.fr/

Rubrique : MASTER → LITL

Plan de la présentation

- Objectifs du master et compétences développées
- 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) et de la linguistique outillée**
- 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
- Les méthodes du TAL sont également utilisées pour **outiller les travaux de recherche sur le langage** :
  - annotation automatique de corpus, fouille de données langagières massives, test de modèles théoriques
- A la fois un domaine de recherche et un champ professionnel, à l'importance toujours croissante, lié au développement de méthodes et d'outils issus de l'**intelligence artificielle**

Environnement

La formation s'appuie sur

- un laboratoire de recherche CNRS-UT2J (CLLE : Cognition, Langage, Langues, Ergonomie) associant des chercheurs en linguistique et en psychologie
  = laboratoire d'adossement du master
- le tissu industriel toulousain et un réseau professionnel à Toulouse (Airbus, Besedo, CNES, Coup de Puce, Inbenta, Prometil, Synapse) et en France (Eptica, Fortia financial solutions, Orange, Proxem, Syllabs…)
- un réseau d'anciens étudiants
- une formation en Intelligence Artificielle labellisée

Les secteurs d'activités

- Le traitement et l'analyse des contenus textuels
  - Rechercher, analyser, transformer …
- Dans tous les domaines
  - Gestion de la communication client
  - Gestion de la documentation technique
  - Ressources humaines
  - Gestion des risques
  - Veille technologique, commerciale, scientifique

Nature des activités professionnelles

- Développer des ressources linguistiques pour des systèmes de TAL
  - Corpus bruts et annotés (génériques ou issus d'activités spécifiques)
  - Lexiques, terminologies, ontologies, grammaires locales
- Evaluer et améliorer des systèmes de TAL
  - Développement et paramétrage de données de test
  - Analyse fine 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 d'emploi occupés

- *Ingénieure linguiste* chez Besedo (Toulouse) : R. travaille au développement de ressources et d'outils pour aider le travail de modération de contenus en ligne
- *Product manager junior* chez Proxem (Paris) : E. travaille sur des projets de réalisation d'agents conversationnels
- *Ingénieure en linguistique informatique* chez Alten (Toulouse) : C. développe un système d'analyse des rapports d'incidents du travail dans la construction aéronautique
- *Senior machine learning scientist* chez Ada (Montréal) : B. coordonne une équipe qui développe et adapte un agent conversationnel
- *Terminologue* chez Airbus (Toulouse) : M. inventorie et organise les termes utilisés dans l'aéronautique et les utilise pour améliorer l'accès à la documentation technique
- *Ingénieure d'étude NLP* à la SNCF (Paris) : S. développe des outils et méthodes pour analyser les rapports d'incidents
- *Chercheuse post-doctorante* (Orléans) : Après une thèse à Toulouse financée par un contrat doctoral (CDU), S. a obtenu un contrat de post-doctorat sur l'analyse de la coréférence dans les documents médicaux

Organisation de la formation

M1 300 h + mémoire		
Enseignements spécifiques	TAL et informatique	100h
Méthodologie	Corpus et linguistique outillée, statistiques	50h
Options de linguistique	Syntaxe, morphologie, sémantique, etc.	75h
Projet tutoré	Projet collectif encadré	25h
Langues	Anglais si C1 non acquis	50h
Mémoire	Travail autonome encadré	

Organisation de la formation

M2 250h + mémoire ou stage		
Enseignements spécifiques	Programmation et TAL	100h
Ergonomie	TAL et ergonomie, conception informatique et ergonomie	25h ou 50h
Options de linguistique	Syntaxe, morphologie, sémantique, etc.	25h ou 50h
Professionnalisation	Recherche de stage et projet professionnel	25h
Projet tutoré	Projet collectif encadré	25h
Langues	Anglais de spécialité	25h
Mémoire ou stage long	Travail autonome encadré	

Période d'enseignement : sept-février

N.B. démarrage 1 à 2 semaines avant la date de rentrée universitaire UT2J : pré-rentrée, soutenances de stages de la promotion précédente.

Stage de 4 à 6 mois ou mémoire de recherche

Modalités pédagogiques et évaluation

- Beaucoup de travail pratique
- Des évaluations en continu :
  - Exercices d'une séance à l'autre, comptes-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

Exemples de sujets de mémoire

- Expression de la neutralité dans les discussions Wikipédia
- Classification et analyse de l'expression des problèmes techniques dans les forums
- Comparaison des emplois du passif à l'écrit et à l'oral
- Evaluation de la qualité des questionnaires à choix multiples
- Evolution des structures discursives dans les écrits scolaires du CP à l'université

Exemples de sujets de stages

- Identification de profils de compétences similaires parmi les employés (ALTEN)
- Extraction d'information spécialisée dans les publications scientifiques (Airbus)
- Création d'un outil d'aide à l'analyse des entretiens sur la sécurité au travail (SNCF)
- Mesure de la richesse du discours produit par des apprenant FLE (Altean Technologies)
- Amélioration de la performance d'un chatbot pour le service client (Coddity)

Les projets tutorés

- Projet TAL en M1

  Projet en équipe avec accompagnement pédagogique

  – Projet 2020-21 : construction et enrichissement d'un corpus de textes procéduraux
  – Projet 2021-22 : participation au développement d'une plateforme d'annotation
  – Projet 2022-23 : prototype de système multilingue d'analyse d'opinion
  – Projet 2023-24 : analyse de discussions en ligne entre collégiens

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
  - Ex de projets :
    - Détection des risques dans les textes procéduraux (Safety Data)
    - Amélioration de l'accès aux ressources documentaires dans le domaine du droit (Revue Fiduciaire)
    - Développement de ressources multilingues pour un outil d'enseignement des langues (CheckYourSmile)

Intégrer le master LITL

Candidature via la plateforme nationale Mon Master (février 2025)

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 très 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 de base en informatique garantissant un premier niveau d'autonomie dans l'utilisation de l'ordinateur (stocker, organiser, traiter ses données, maîtriser les 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