

Master Traitement Automatique des Langues

Parcours Ingénierie Multilingue

Parcours Technologie de la Traduction et
Traitement des Données Multilingues

Parcours Recherche & Développement



2025-2026

Le MASTER TAL : l'informatique et l'IA au service des langues

Le Traitement Automatique des Langues (ou NLP pour Natural Language Processing) est un domaine à l'interface de la linguistique, de l'informatique et de l'Intelligence Artificielle axé sur les méthodes, algorithmes et technologies dédiées aux langues et aux textes. Les débouchés de ce domaine sont soit applicatifs (réalisation d'applications spécialisées en traduction automatique, agents conversationnels – chatbots, veille, filtrage et recherche d'information, extraction de connaissances, etc.) soit scientifiques (outils pour l'analyse linguistique automatique ou assistée, humanités numériques).

Le Master TAL, élaboré en partenariat avec les universités Sorbonne Nouvelle et Paris Nanterre, offre une formation complète en : constitution de corpus et jeux de données, modélisation, algorithmique, programmation, apprentissage neuronal, développement de chaînes de traitement automatisées. Sa spécificité est de se focaliser sur les problématiques du multilinguisme, de la variation linguistique et du traitement automatique des langues du domaine Inalco qui disposent souvent de peu de données disponibles et d'applications dédiées. La formation met donc l'accent sur les systèmes d'écriture, d'encodage, les caractéristiques linguistiques de ces langues dites « peu dotées ».

Le master TAL répond aux attentes économiques et sociales créées par la nécessité de penser le numérique en termes de diversité des langues et des cultures. Il forme aux métiers du numérique : ingénieur·e linguiste, ingénieur·e NLP, data scientist, ingénieur·e LLM, ingénieur·e IA, etc.

L'organisation du Master TAL

Après un tronc commun à l'Inalco, et aux universités Sorbonne Nouvelle et Paris Nanterre en 1^{re} année où sont enseignés les fondamentaux en linguistique pour le TAL, mathématiques, algorithmiques et programmation, la 2^e année permet de choisir entre 3 parcours :

- **Ingénierie multilingue :**

Formation aux métiers de l'ingénierie NLP et IA. Le parcours se déroule sur un semestre et demi et est suivi d'un stage en entreprise ou en laboratoire donnant lieu à un mémoire de recherche.

- **Technologies de la Traduction et Traitement des Données**

- Multilingues :**

Formation équivalente au parcours « Ingénierie Multilingue » mais spécialisée en traduction automatique pour les langues peu dotées. Attention, ce parcours ne sera pas ouvert à l'inscription en 2025.

- Recherche et développement :

Formation co-accréditée avec les universités Sorbonne Nouvelle et Paris Nanterre, « à la carte » destinée à préparer un projet de recherche doctorale. La formation s'étend sur un semestre suivi d'un stage donnant lieu à un mémoire de recherche. Le master TAL est multi-site. En master 1 et en Master 2 Recherche et développement les cours se répartissent entre le Pôle des langues et civilisations (Inalco, 13e), le Campus Nation (Sorbonne Nouvelle, 12e), le Campus Nanterre (Paris Nanterre, 92).

Les cours du Master 2 des parcours Ingénierie multilingue et Technologies de la Traduction et Traitement des Données Multilingues ont lieu à la Maison de la recherche de l'Inalco (7e).

Public visé

Le Master appuie sa démarche de formation sur les acquis en langues des étudiants. Il s'adresse :

- à des étudiants alliant des connaissances significatives en informatique et souhaitant valoriser des compétences en langue dans la perspective de travailler dans le domaine de l'ingénierie linguistique,
- à des étudiants en langues ou en linguistique n'ayant pas obligatoirement de compétences techniques mais désireux d'acquérir une formation technique approfondie favorisant leur intégration dans les métiers du numériques. Une semaine de formation intensive en pré-rentree est programmée pour la mise à niveau des candidats n'ayant pas suivi de formation informatique au préalable.

Conditions d'admission

Les étudiants sont admis sur dossier. Les étudiants doivent justifier d'un niveau suffisant dans une langue enseignée à l'Inalco et faire preuve de compétences en linguistique ou en informatique. Les étudiants étrangers doivent en outre justifier d'une connaissance suffisante du français.

Débouchés professionnels

Les métiers des industries de la langue sont multiples. Des possibilités s'offrent dans les entreprises du numérique ou ayant des secteurs spécialisés dans le développement d'outils TAL (NLP) et IA pour la conception et la maintenance de logiciels. Ces mêmes entreprises et bien d'autres (petites et moyennes, *startups*, organismes nationaux et internationaux), pour les services qu'elles proposent ou pour leurs besoins propres, peuvent également faire appel à des spécialistes

dans les domaines de la production, de l'organisation et de la gestion de l'information et de la connaissance. Les entreprises qui recrutent nos alumni sont par exemple Orange, Thales, Dassault System, Google, EDF, Ubisoft, Systran, Acolad, Viasema, Mondeca, Sématiweb, Aday, etc. Les métiers relèvent essentiellement de l'ingénierie informatique : ingénieur linguiste, ingénieur NLP, développeur NLP, ingénieur LLMs, ingénieur IA, data scientists et ingénieur text mining, etc.

Contacts

Mathieu VALETTE, Directeur de la filière, mathieu.valette@inalco.fr

Damien NOUVEL, directeur adjoint de la filière, damien.nouvel@inalco.fr

Bassir HAMID, responsable (administratif) de formation, bassir.hamid@inalco.fr
01 81 70 11 42

Alice RIMBAULT, gestionnaire pédagogique, alice.rimbault@inalco.fr
Bureau 3.25A – 01 81 70 11 36



Les enseignements

Les tableaux ci-après décrivent l'organisation des enseignements et la répartition des ECTS pour le Master 1 et le Master 2 parcours Ingénierie Multilingue. Les enseignements du parcours Recherche & développement sont à choisir parmi l'offre du 1er semestre du parcours Ingénierie Multilingue et dans celles des universités Sorbonne Nouvelle et Paris Nanterre.

Master 1 TAL, semestre 1	30 ECTS
UE 1 : LANGUES VIVANTES - 1 cours au choix	3
Langue de l'Inalco, 1	3
Paper reading, 1	3
NLP in English, 1	3
UE 2 : LINGUISTIQUE COMPUTATIONNELLE	9
Corpus multilingues et traduction	3
Linguistique pour le TAL	3
Corpus arborés et parsing	3
Un cours de linguistique au choix (Inalco, Sorbonne nouvelle ou Université Paris Nanterre)	
UE 3 : FONDAMENTAUX	9
Mathématiques pour le TAL, 1	3
Algorithmique et programmation, 1	3
Traitement symbolique des données	3
UE 4 : INGÉNIERIE	9
Introduction au TAL	3
Construction de données, 1	3
Projet de programmation encadré, 1	3

Master 1 TAL, semestre 2	30 ECTS
UE 1 : LANGUES VIVANTES - 1 cours au choix	3
Langue de l'Inalco, 2	3
Paper reading, 2	3
NLP in English, 2	3
UE 2 : LINGUISTIQUE COMPUTATIONNELLE	6
Analyse statistique des données textuelles	3
Sémantique distributionnelles	3
UE 3 : FONDAMENTAUX	9
Mathématiques pour le TAL, 2	3
Algorithmique et programmation, 2	3
Traitement statistique des données	3
UE 4 : INGENIERIE	12
Applications	3
Constructions de données, 2	3
Projet de programmation encadré, 1	3
Parole / oralités	3

Master 2ème année : 2 parcours

- Ingénierie multilingue
- Recherche et développement

M2 – Parcours Ingénierie multilingue – semestre 1	30 ECTS
UE 1 : LINGUISTIQUE	9
Linguistique pour le TAL multilingue, 1	3
Ecritures multilingues (encodage, langues peu dotées)	3
Corpus, langues, cultures	3
UE 2 : INGÉNIERIE	11
Techniques Web, programmation Web, réseaux et applications mobiles	3
Génération, 1 - traduction automatique	4
Extraction, 1 - fouille de textes, extraction d'information	4
UE 3 : MODELES ET FORMALISMES	10
Programmation et programmation objets	4
Apprentissage, réseaux de neurones profonds, modèles de langues, 1	3
Acquisition, modélisation et représentation des connaissances	3

M2 – Parcours Ingénierie multilingue – semestre 2	30 ECTS
UE 1 : LINGUISTIQUE	3
Linguistique pour le TAL multilingue, 2	3
UE 2 : INGÉNIERIE	9
Techniques Web, programmation Web, réseaux et applications mobiles	3
Génération, 1 - traduction automatique	3
Extraction, 1 - fouille de textes, extraction d'information	3
UE 3 : MODELES ET FORMALISMES	3
Apprentissage, réseaux de neurones profonds, modèles de langues, 2, RAG	3
UE 4 : STAGE ET MEMOIRE	15
Stage et mémoire	15

M2 - parcours Recherche et développement Semestre 1	30 ECTS
Les cours sont à sélectionner, en concertation avec l'équipe pédagogique, dans l'offre des 3 établissements du Master TAL. 3 cours de linguistique (9 ECTS) 7 cours d'ingénierie en TAL (21 ECTS)	
M2 - parcours Recherche et développement Semestre 2	30 ECTS
STAGE (en laboratoire ou en entreprise)	9
MÉMOIRE de Recherche	15
MÉTHODOLOGIE DE LA RECHERCHE	6