UNIVERSITY DIPLOMA IN LANGUAGE ENGINEERING APPLIED TO LANGUAGE PROFESSIONS

Main Language of Instruction:
French ○ English ☑ Arabic ○

Campus Where the Program Is Offered: Online

OBJECTIVES

As artificial intelligence redefines the ways language is produced, processed, and circulated, the role of language professionals is not diminishing but being fundamentally reconfigured. The University Diploma in Language Engineering Applied to Language Professions responds to this shift by positioning participants at the forefront of this transformation, equipping them with the expertise, agility, and critical insight required to operate at the intersection of linguistics, technology, and artificial intelligence.
Designed as a 20-credit specialization, the program moves decisively beyond a purely instrumental use of tools. It cultivates a deep understanding of the mechanisms underpinning AI-driven language systems while reaffirming the indispensable role of human expertise in safeguarding meaning, quality, and ethical responsibility.
Grounded in an interdisciplinary and technocritical approach, the program prepares participants not only to use language technologies but also to interrogate them, supervise their outputs with discernment, and actively contribute to their evolution.
The University Diploma in Language Engineering Applied to Language Professions therefore aims to:

- Develop advanced language and translation expertise combined with solid technological literacy in AI-driven environments
- Provide a comprehensive understanding of the principles of language engineering, including text processing, alignment, segmentation, and neural language models
- Train participants to integrate and critically evaluate language technologies and automated workflows with rigor and discernment
- Foster the ability to collaborate effectively within interdisciplinary environments involving developers, engineers, and data specialists
- Cultivate a technocritical and ethical mindset, enabling participants to assess the reliability, limitations, and societal implications of AI systems
- Strengthen competencies in supervising AI-mediated language production, including pre-editing, post-editing, quality assurance, and human-in-the-loop validation
- Develop strategic awareness of the evolving language industry, empowering participants to adapt to emerging roles and actively contribute to shaping the future of language professions

PROGRAM LEARNING OUTCOMES (COMPETENCIES)

- Design and implement automated workflows for effective multilingual project management
- Apply algorithmic foundations of language processing to optimize generative AI outputs in multilingual content production
- Develop interpersonal and transversal skills in AI-driven environments to foster multilingual creativity

ADMISSION REQUIREMENTS

- **Academic Eligibility**

Admission to the University Diploma in Language Engineering Applied to Language Professions is open to candidates holding a Master's degree in a relevant field. Eligible degrees include:

- Master in Translation
- Master in Interpretation
- Master in Languages

- Master in Applied Linguistics
- Master in Language Sciences
- Master in Language Engineering
- Master in Multilingual Communication
- Master in Localization
- Master in Language Technologies
- Master in Terminology
- Master in Natural Language Processing (NLP)

Candidates holding degrees in other disciplines may also be considered, subject to an individual review of their academic and professional background.

• **Admission Based on Qualifications or Professional Experience**

Applicants who do not hold one of the above-mentioned degrees may still be eligible for admission through recognition of prior professional experience. This applies to:

- Translators, writers, terminologists, interpreters, or instructors in translation and interpretation
- Professionals with a minimum of three years of relevant experience in translation or language services

Admission in such cases is granted following a thorough evaluation of the candidate's application file and professional achievements.

• **Application Process**

Admission is based on a comprehensive application file, which must include:

- A detailed curriculum vitae
- A statement of purpose (in English) outlining the candidate's interest in language engineering and translation and their motivation for pursuing the diploma
- Certified copies of academic degrees
- Employment certificates in cases where admission is based on professional experience
- Any supporting documents demonstrating familiarity with digital tools, language technologies, or collaborative work environments (considered an asset)

• **Required Language Proficiency**

Applicants must demonstrate:

- Excellent command of Arabic, along with strong proficiency in either French or English as a working language
- Solid skills in translation across a range of specialized domains

• **Basic Digital Competencies**

Candidates are expected to possess:

- Proficiency in standard office software (word processing, spreadsheets, presentations)
- Familiarity with digital work environments and online collaboration tools
- Prior experience or training in computer-assisted translation (CAT) tools is desirable, though not mandatory

PROGRAM REQUIREMENTS

20 Credits: Required Courses (20 credits)

Required Courses (20 Cr.)

Data Construction (4 Cr.). Ethics, Technocriticism, and Professional Responsibility (2 Cr.). Human Augmented LLM Prompting, Integration, and Evaluation (2 Cr.). Introduction to Multilingual NLP and MT (6 Cr.). Knowledge Engineering (2 Cr.). Mentored Project (4 Cr.).

COURSE DESCRIPTION

103DACOM1	Data Construction	4 Cr.

This module introduces students to the principles, methods, and best practices for data construction in natural language processing (NLP). Students will learn to collect, clean, annotate, and document linguistic corpora following industry standards as well as ethical and legal requirements (licensing, GDPR, data consent).
An introduction to the Python programming language will enable students to automate key stages of the data lifecycle, such as data collection, normalization, and corpus analysis, using beginner-friendly libraries.
The module covers:

- The data lifecycle (collection, normalization, segmentation, annotation, documentation)
- Standard formats (XML, JSON, TMX) and manual/semi-automatic annotation tools
- Quality issues (balancing, representativeness, bias) and strategies for low-resource languages
- Use of Python for corpus cleaning, transformation, and exploratory analysis
- Creation of datasets for training and evaluating NLP models
- Metadata planning to ensure corpus reusability and interoperability

103ETPRM1	Ethics, Technocriticism, and Professional Responsibility	2 Cr.

This module encourages students to adopt a critical and responsible stance toward language technologies and artificial intelligence (AI). Through case studies and guided readings, students will analyze the ethical challenges raised by algorithmic systems in translation, text mining, information extraction, named entity recognition, generative AI, and speech processing technologies.
Key focus areas include:

- Understanding algorithmic bias, misinformation risks, and privacy concerns
- Assessing the societal and professional impacts of NLP applications
- Drafting a code of best practices for the responsible use of AI in translation and localization projects
- Bridging translator ethics with regulatory frameworks (GDPR, ISO standards, ethical guidelines)

103HALMM1	Human Augmented LLM Prompting, Integration, and Evaluation	2 Cr.

This module equips language professionals with the knowledge and practical skills to work with large language models in a human-centered, critically informed manner, with a particular focus on the multilingual realities of Arabic in contact with English and French. It explores how linguistic variation across Modern Standard Arabic and Arabic dialects interacts with LLM behavior, especially in contexts involving code-switching, Arabizi, loanwords, and hybrid registers that characterize real-world communication.
Students learn to design effective and context-sensitive prompts that account for these complexities, guiding models to produce accurate, culturally appropriate, and audience-aware outputs. The module also introduces the integration of LLMs into professional workflows through APIs and automated pipelines, including applications in translation, localization, and multilingual content production. Particular attention is given to evaluating model outputs in linguistically diverse settings, identifying errors related to dialectal variation, script switching, and semantic drift, and refining outputs through iterative human intervention.
Throughout the module, emphasis is placed on human augmentation rather than automation: students develop the ability to supervise, correct, and strategically leverage LLMs while maintaining responsibility for meaning, quality, and ethical standards.

103IMNMM1	Introduction to Multilingual NLP and MT	6 Cr.

This module introduces students to the foundations of natural language processing (NLP) and machine translation (MT), with a focus on the evolution of MT technologies and their integration into professional translation and localization workflows. Through lectures and hands-on practice, students explore core NLP concepts, the transition from rule-based to neural MT, and gain experience with leading Translation Environment Tools (TEnTs), corpus management and text analysis software, AI voice generation tools, localization software, and computer-assisted interpreting tools.

103KNENM1	Knowledge Engineering	2 Cr.

This module provides a comprehensive introduction to knowledge engineering and its core methodologies. Students learn how to model, formalize, and structure knowledge extracted from textual corpora so that it can be operationalized in information retrieval and decision-support systems. The first part of the module introduces key knowledge representation models, including semantic networks, description logics, and ontologies, with a focus on current modeling standards such as RDF and OWL and their importance for data interoperability. The second part emphasizes practical application through case studies, including the visualization of event timelines from journalistic dispatches and the analysis of epistemic modality in specialized biology texts. These case studies illustrate how textual information can be extracted, validated, and transformed into structured knowledge usable by domain experts. The pedagogical approach combines theoretical instruction, corpus exploration, and hands-on use of natural language processing tools, offering students a holistic view that spans from information extraction to the creation of knowledge models.

103MEPRM1	Mentored Project	4 Cr.

This capstone project serves as the practical culmination of the University Diploma in Language Engineering Applied to Language Professions. Students are required to design, implement, and document a complete natural language processing (NLP) pipeline, drawing on the knowledge and skills acquired throughout the program. Under the guidance of a supervisor and with input from instructors of the different modules, students move from needs analysis to system design, implementation, and evaluation. The project involves data collection and preparation, the selection and deployment of NLP tools or models, the development of a user-friendly interface, and the production of comprehensive technical and methodological documentation.
The project concludes with an oral defense before a jury made up of the program's instructors. Assessment criteria include technical quality, methodological rigor, teamwork, ethical considerations, and the clarity of both oral and written communication. This course provides students with the opportunity to address real-world challenges and demonstrate their ability to deliver a fully operational project from design to deployment.