# MSC IN NLP COURSE DESCRIPTION SEMESTER 7 – S7

Updated for academic year 2024-2025

Study Unit – Unité d'Enseignement (UE) Course – Élément Constitutif (EC)

Mandatory 
 ILPC Track
 ILTrack
 Optional

<b>STUDY UNIT 701</b> – Probabilities, Statistics and Algorithms for Al	6 ECTS	COEF 6
• <b>PROBABILITIES AND STATISTICS</b> Marianne CLAUSEL — marianne.clausel@univ-lorraine.fr		٦
Mandatory for all M1 students.		
This course revisits the fundamental mathematical properties of proba covering classic data analysis methods and introducing more complex SVMs and discriminant analysis, with practical applications using real data	techniques s	such as
<ul> <li>Students will learn to:</li> <li>Identify statistical analysis methods appropriate for studying data</li> <li>Perform statistical calculations for the mathematical analysis of c</li> <li>Conduct qualitative data evaluations.</li> </ul>		
CM – 10 hours TD – 10 hours TOTAL – 20 hours	COEF -	1
• <b>PYTHON PROGRAMMING</b> Karën FORT — karen.fort@univ-lorraine.fr		
Mandatory for all M1 students.		
This course is designed to offer a comprehensive introduction to Python focus on deepening Python skills to solve complex problems and applications using various libraries for advanced learners.	-	
<ul> <li>Students will learn to:</li> <li>Model a problem in the form of algorithms.</li> <li>Develop a software solution in Python for data science problems.</li> <li>Implement and apply classical algorithms for processing numerication.</li> <li>Master the Python programming environment.</li> </ul>	al data.	

CM — 15 hours	TD — 15 hours	TOTAL — 30 hours	COEF — 1
STUDY UNIT 702 — Writter	n Corpora and Logic		6 ECTS COEF 6
WRITTEN CORPORA     Bruno GUILLAUME – bru	ıno.guillaume@univ-lc	orraine.fr	
Mandatory for all M1 stu	dents.		
written corpora, introduc	ring corpus linguistics various document f	chniques used in NLP for s, criteria for corpus constr ormats, and annotation p ation.	uction, data collection
<ul> <li>Standardise the</li> </ul>	ent of a written docum data of a corpus. iples and examples of		
CM — 15 hours	TD – 15 hours	TOTAL – 30 hours	COEF — 1
• LOGIC Philippe DE GROOTE — p	hilippe.de-groote@un	iv-lorraine.fr	
Mandatory for all M1 stu	dents.		
	g the syntax and se	th the fundamental conce emantics of propositional nd higher-order logic.	
	syntax and semantics oda calculus and high	-	
CM — 10 hours	TD — 10 hours	TOTAL – 20 hours	COEF — 1
STUDY UNIT 703 — Tools a	and Challenges of NLF	P Development	6 ECTS COEF 6
• ETHICS IN NLP Karën FORT — karen.fort	@univ-lorraine.fr		
Mandatory for all M1 stu	dents.		

This course introduces students to the ethical issues in NLP, including data privacy, algorithmic bias, information manipulation, and societal impact, through discussions and case studies that foster the ethical awareness necessary for designing responsible NLP systems and acting ethically in future AI-related work.

Students will learn to:

- > Master the development of solutions through an ethical approach.
- > Identify risk factors in the development of solutions.
- > Assess the consequences of developing an NLP solution.

CM = 7 hours ID = 8 hours IOIAL = 15 hours COEF = 1	CM – 7 hours	TD – 8 hours	TOTAL — 15 hours	COEF - 1
---	--------------	--------------	------------------	----------

# • PROJECT MANAGEMENT TOOLS

Maxime AMBLARD - maxime.amblard@univ-lorraine.fr

## Mandatory for all M1 students.

This course focuses on practical tools essential for effective project management, teaching students to use platforms like Git, work with the command line in Shell, collaborate on complex projects, and use LaTeX for documentation, enabling them to manage NLP projects professionally and collaboratively.

Students will learn to:

- Design and develop an automatic data management system based on a real-world case.
- Master techniques for gathering client requirements and modelling the corresponding functionalities.
- Master project management techniques as a whole, including financial analysis (costs and time), group facilitation, and conflict management.

```
CM – 10 hours TD – 10 hours TOTAL – 20 hours COEF – 1
```

## INTRODUCTION TO NLP

Maxime AMBLARD - maxime.amblard@univ-lorraine.fr

# Mandatory for all M1 students.

This course offers a comprehensive introduction to the fundamental concepts of NLP, while also exploring the relationship between NLP, computer science, and linguistics, and addressing the ecological impact of these advancements.

Students will learn to:

- > Understand the range of issues in NLP.
- > Use basic NLP tools on practical examples.
- > Implement a minimal solution to an NLP problem.

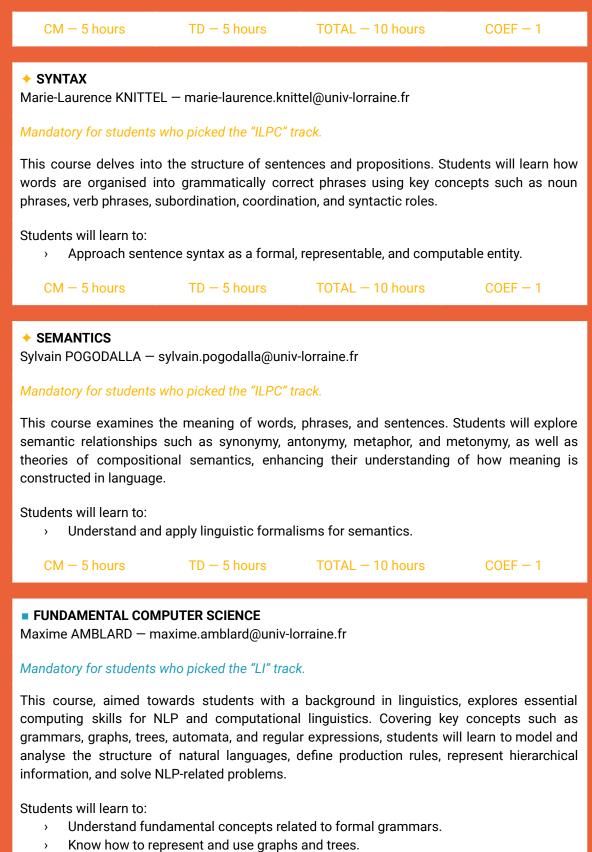
CM – 8 hours

TD – 7 hours

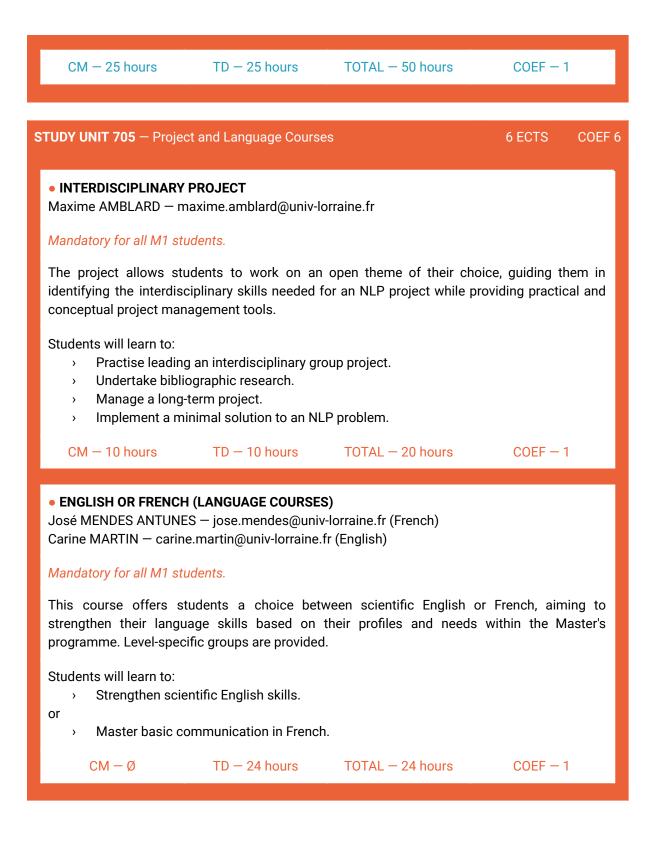
TOTAL – 15 hours

COEF - 1

STUDY UNIT 704 — Tracks (Specialisations)	6 ECTS	COEF 6
<ul> <li>GENERAL LINGUISTICS</li> <li>Mathilde DARGNAT — mathilde.dargnat@univ-lorraine.fr</li> </ul>		
Mandatory for students who picked the "ILPC" track.		
This course provides a comprehensive overview of general linguistics, intro the foundational concepts and issues in the field. It covers a wide range from field linguistics to psycholinguistics, equipping students with es knowledge and analytical skills for understanding language structure and us	of linguistic sential theo	topics,
<ul> <li>Students will learn to:</li> <li>Understand the broad spectrum of linguistic challenges.</li> <li>Use basic NLP tools on practical examples.</li> </ul>		
CM – 5 hours TD – 5 hours TOTAL – 10 hours	COEF — <sup>2</sup>	1
<ul> <li>Mélanie LANCIEN – melanie.lancien@univ-lorraine.fr</li> <li>Mandatory for students who picked the "ILPC" track.</li> <li>This course focuses on acoustic phonetics, targeting students with proceeding of speech acoustics, followed by practical errecognition, synthesis, and automatic processing of oral data, aiming to prova solid foundation in the technical aspects of phonetics.</li> <li>Students will learn to:         <ul> <li>Implement analytical techniques for linguistic data.</li> <li>Identify basic phonological properties.</li> </ul> </li> </ul>	xercises in a	speech its with
	OOEI	
<ul> <li>MORPHOLOGY</li> <li>Fiammetta NAMER — fiammetta.namer@univ-lorraine.fr</li> </ul>		
Mandatory for students who picked the "ILPC" track.		
This course explores the study of morphology, focusing on the regular co the form and meaning of complex words. Students will examine key morphemes, lexemes, derivation, and inflection, as well as the construction different languages.	concepts si	uch as
<ul> <li>Students will learn to:</li> <li>Implement analytical techniques for linguistic data.</li> <li>Identify the morphological properties of language.</li> </ul>		



- > Define automata and transducers.
- > Search for simple patterns in textual documents using regular expressions.



# MSC IN NLP Course description SEMESTER 8 – S8

Updated for academic year 2024-2025

Study Unit — Unité d'Enseignement (UE) Course — Élément Constitutif (EC)

Mandatory 
 ILPC Track
 ILTrack
 A Optional

# STUDY UNIT 801 — Machine Learning and Symbolic AI

6 ECTS COEF 6

# MACHINE LEARNING

Marianne CLAUSEL – marianne.clausel@univ-lorraine.fr

# Mandatory for all M1 students.

This course introduces Bayesian models to cover fundamental principles of machine learning, including supervision, classification, and optimisation, while addressing industry scenarios. It emphasises practical aspects and best practices for utilising machine learning methods, such as data analysis, preprocessing, visualisation, quality measurement, and managing large and streaming corpora.

Students will learn to:

- > Apply statistical mathematical theories to real-world data produced by humans.
- > Analyse the behaviour of automated models to assess their quality.
- > Apply logical mathematical theories to internet-derived data.
- > Extract information from this data and structure it into reusable knowledge.

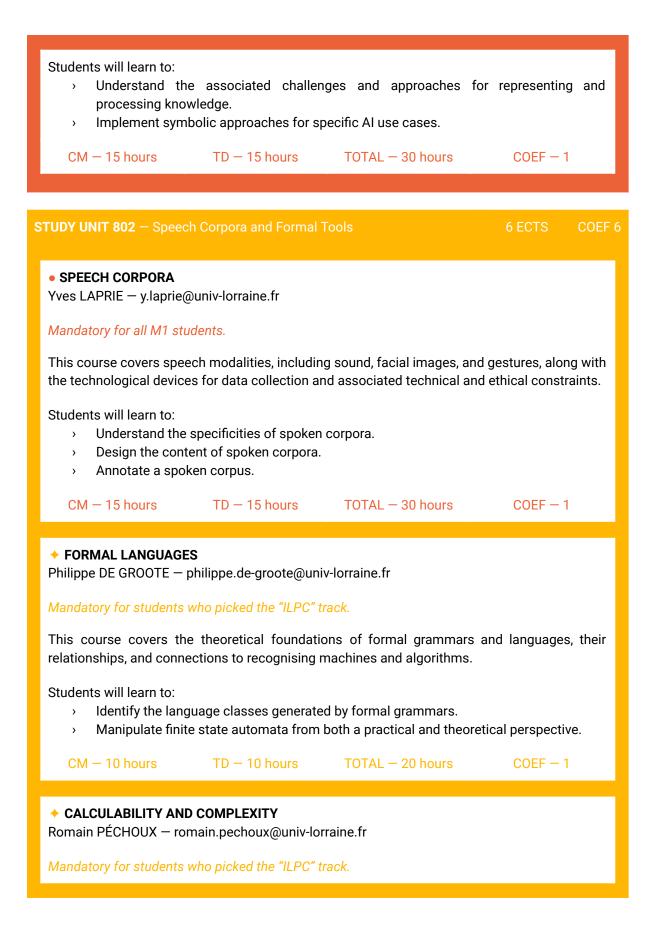
CM – 15 hours	TD – 15 hours	TOTAL – 30 hours	COEF - 1	
---------------	---------------	------------------	----------	--

## SYMBOLIC AI

Mathieu D'AQUIN - mathieu.d-aquin@univ-lorraine.fr

# Mandatory for all M1 students.

The Symbolic AI course focuses on 'knowledge-centric' approaches to AI, i.e. problem-solving, decision support and learning methods based on the explicit and formalised representation of knowledge. Starting from traditional knowledge-based systems and associated knowledge engineering methodologies, the course focuses on methods and technologies associated with the semantic web, including knowledge graphs and ontologies.



1

This course covers computability and algorithmic complexity, focusing on Turing machines, non-computable functions, undecidable problems, complexity measures (space, time), and distinct complexity classes (P, PSPACE, NP, EXPTIME).

- > Model a problem with a Turing machine.
- > Identify the complexity of an algorithmic problem.

CM – 10 hours	TD – 10 hours	TOTAL – 20 hours	COEF -

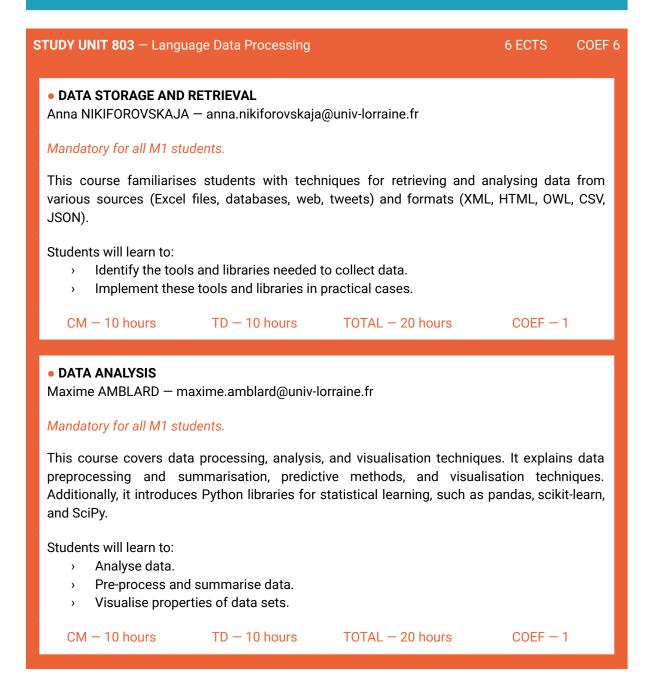
S	TUDY UNIT 802 — Speech Corpora and Applied Linguistics	6 ECTS	COEF 6
	• SPEECH CORPORA Yves LAPRIE — y.laprie@univ-lorraine.fr		
	Mandatory for all M1 students.		
	This course covers speech modalities, including sound, facial images, and ge the technological devices for data collection and associated technical and effective		-
	<ul> <li>Students will learn to:</li> <li>Understand the specificities of spoken corpora.</li> <li>Design the content of spoken corpora.</li> <li>Annotate a spoken corpus.</li> </ul>		
	CM – 15 hours TD – 15 hours TOTAL – 30 hours	COEF —	1
	<ul> <li>MORPHOPHONOLOGY</li> <li>Stéphanie LIGNON – stephanie.lignon@univ-lorraine.fr</li> <li>Mandatory for students who picked the "LI" track.</li> <li>This course briefly reviews major phonological theories and then explores le from a formal perspective and studies specific cases to explain a inconsistencies through morphophonological analysis.</li> <li>Students will learn to:</li> </ul>		
	<ul> <li>Identify the constraints for proper formation of lexemes in French.</li> <li>Understand the sense/shape interactions involved in morphological</li> </ul>	constructio	ns.
	CM – 10 hours TD – 10 hours TOTAL – 20 hours	COEF —	1
	LEXICON—SYNTAX RELATION AND POLYSEMY Alain POLGUÈRE — alain.polguere@univ-lorraine.fr		

# Mandatory for students who picked the "LI" track.

This course explores the semantic and syntactic properties of event-denoting nouns, both derived and non-derived. It focuses on how deverbal nouns inherit their syntactic and semantic features, including argument structure and aspectual structure, from their verb bases.

- Identify the syntactic and semantic properties of various noun classes to assess their type and mode of reference.
- > Justify analyses based on specific evidence.

```
CM - 10 hours 	TD - 10 hours 	TOTAL - 20 hours 	COEF - 1
```



NEURAL NETWORKS

Claire GARDENT - claire.gardent@univ-lorraine.fr

#### Mandatory for all M1 students.

This course introduces neural networks, focusing on their use in machine learning and artificial intelligence. It covers basic concepts, common architectures, and tools for quick and easy implementation.

Students will learn to:

- Design and implement neural networks, choosing the appropriate architectures for different tasks.
- Train neural networks using optimisation techniques, and understand the mechanisms of machine learning.
- > Evaluate the performance of neural network models using appropriate metrics and compare them to other approaches.

TOTAL – 20 hours

CM - 10 hours

#### **STUDY UNIT 804** – Linguistics

6 ECTS COEF 6

COEF - 1

# PROSODY

Mélanie LANCIEN – melanie.lancien@univ-lorraine.fr Samantha RUVOLETTO – samantha.ruvoletto@univ-lorraine.fr

TD – 10 hours

# Mandatory for all M1 students.

This course examines prosody, focusing on melody, rhythm, and tonal variations in speech. Students will study how prosodic features affect meaning and expressiveness, and learn to identify variations across languages and contexts, with a focus on language acquisition and bilingualism.

- > Identify variations in pitch, rhythm, and melody in speech.
- > Analyse intonational and rhythmic patterns in speech.
- > Understand how prosody influences the meaning and interpretation of discourse.

CM — 6	TD — 6	TOTAL – 12	COEF — 1
• PHONETICS Mélanie LANCIEN – me	lania lanaian@univ.lar	roine fr	
Mandatory for all M1 stu	C	Tame.n	

This course explores the acoustic properties of speech sounds. Students will analyze dialects and communication variations, focusing on the material aspects of speech and their role in effective communication.

Students will learn to:

- > Analyse the phonological rules governing the pronunciation of words.
- > Identify phonological and dialectal variations.
- > Compare the phonological systems of different languages.

CM – 6	TD — 6	TOTAL – 12	COEF - 1

# MORPHOLOGY

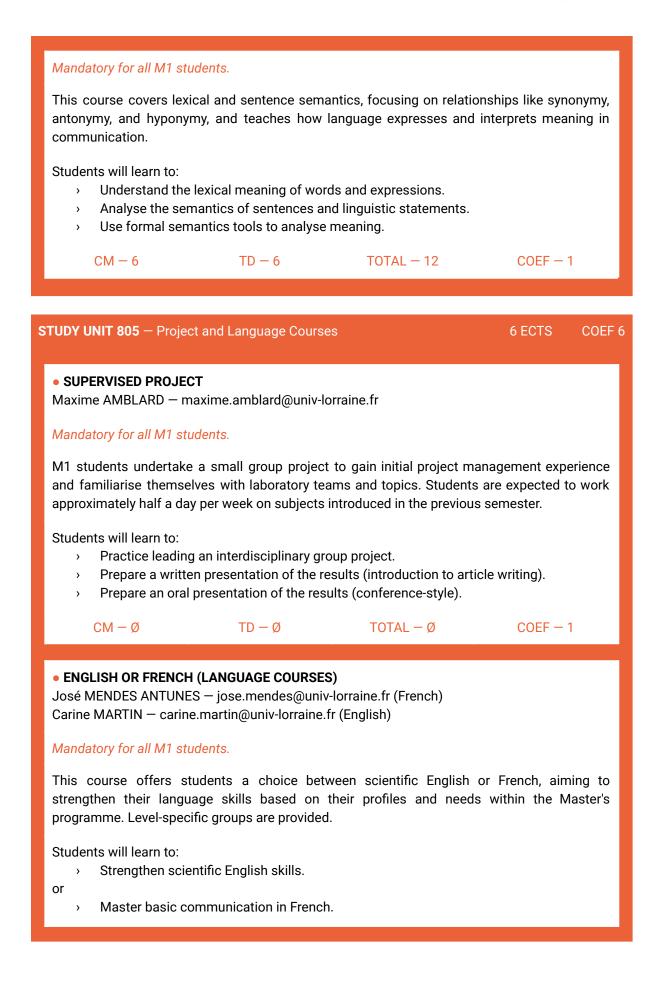
Fiammetta NAMER - fiammetta.namer@univ-lorraine.fr

Mandatory for all M1 students.

This course examines the relationship between word form and meaning in complex words. It focuses on modern morphological principles, particularly constructional morphology, and explores irregular word formations such as templates, cross-formations, subtractive morphology, and parasynthesis, using relational and paradigmatic approaches to analyse these constructions.

- > Analyse the morphological structure of words by identifying morphemes and lexemes.
- > Understand the processes of word formation, including derivation and compounding.
- > Apply morphological skills to the analysis of morphology in different languages.

CM — 6	TD — 6	TOTAL – 12	COEF — 1	
	dents. tence structure, with a al syntax within gener	ne.fr descriptive and typologic rative grammar in the sec		
<ul> <li>Students will learn to:</li> <li>Understand the fundamental principles of sentence structure.</li> <li>Analyse the syntactic rules that govern sentence structure.</li> <li>Identify the dependency relations between elements of a sentence.</li> </ul>				
CM — 6	TD — 6	TOTAL — 12	COEF — 1	
• SEMANTICS Philippe DE GROOTE — p	bhilippe.de-groote@un	iv-lorraine.fr		





# MSC IN NLP COURSE DESCRIPTION SEMESTER 9 – S9

Updated for academic year 2024-2025

Study Unit – Unité d'Enseignement (UE) Course – Élément Constitutif (EC)

Mandatory 
 ILPC Track
 ILTrack
 A Optional

STUDY UNIT 901 — Basics in NLP	9 ECTS	COEF 9
<ul> <li>NEURAL NETWORKS</li> <li>Emmanuel VINCENT — emmanuel.vincent@univ-lorraine.fr</li> </ul>		
Mandatory for all M2 students.		
This course provides an overview of neural network-based machine learn from multi-layer perceptrons to advanced architectures like determ networks, convolutional networks, and recurrent networks.		· · •
<ul> <li>Students will learn to:</li> <li>Master theories on neural networks.</li> <li>Apply these theories to develop a solution.</li> </ul>		
CM – 15 hours TD – 15 hours TOTAL – 30 hours	COEF —	1
<ul> <li>WRITTEN DATA PROCESSING</li> <li>Kamal SMAILI – kamal.smaili@univ-lorraine.fr</li> </ul>		
Mandatory for all M2 students.		
This course introduces learning methods and specialised libraries (e.g., N modelling various subtasks and applications in written text processing automatic translation and the Expectation-Maximisation algorithm.		
<ul> <li>Students will learn to:</li> <li>&gt; Use NLP tools for standard tasks.</li> <li>&gt; Implement techniques for machine translation.</li> </ul>		
CM – 10 hours TD – 10 hours TOTAL – 20 hours	COEF —	1

# LOW-RESOURCE LANGUAGES

Delphine BERNHARD - dbernhard@unistra.fr

# Mandatory for all M2 students.

This course focuses on the study and preservation of low-resource languages, exploring challenges in maintaining linguistic diversity and revitalising endangered languages through sociolinguistic, educational, and technological strategies.

- > Understand the challenges of preserving and revitalising low-resource languages.
- > Master strategies for linguistic documentation and revitalisation.
- > Contribute to the preservation and revitalisation of low-resource languages through concrete actions.

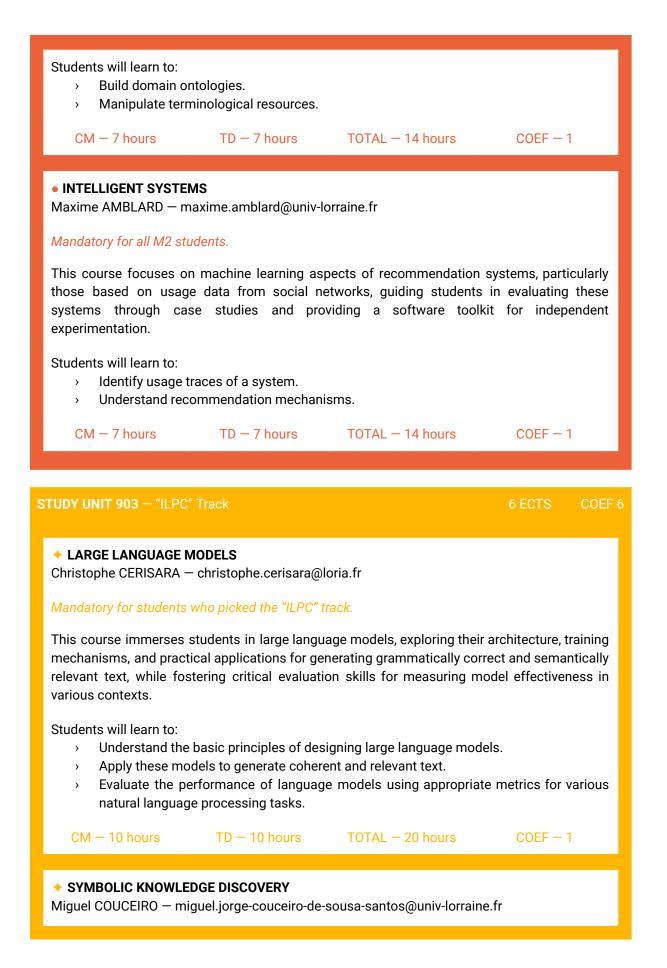
CM — 7 hours	TD — 7 hours	TOTAL — 14 hours	COEF — 1
• GENERATIVE AI Claire GARDENT — claire.	gardent@univ-lorrain	e.fr	
Mandatory for all M2 stud	lents.		
generating content from	data and text, includi	element of large language ng simplification and auton ic, and statistical approache	natic summarisation,
	les of text generation. tion of automatic ger		
CM — 10 hours	TD – 10 hours	TOTAL – 20 hours	COEF – 1
LEXICAL RESOURCES		unain a fu	
LEXICAL RESOURCES     Bruno GUILLAUME — brui	no.guillaume@univ-lo	rraine.fr	
		rraine.fr	
Bruno GUILLAUME – brun Mandatory for all M2 stud This course introduces k	lents. ey types of digital lex al ontologies, and	rraine.fr xical resources, including el transformer-based embed	
Bruno GUILLAUME – brun Mandatory for all M2 stud This course introduces k lexical databases, lexic foundational lexicology c Students will learn to:	lents. ey types of digital lex al ontologies, and oncepts. elf with existing lexica	tical resources, including el transformer-based embed	
Bruno GUILLAUME – brun Mandatory for all M2 stud This course introduces k lexical databases, lexic foundational lexicology c Students will learn to: > Familiarise onese	lents. ey types of digital lex al ontologies, and oncepts. elf with existing lexica	tical resources, including el transformer-based embed	



This course introduces various methods for representing the meaning of complex expressions at the sentence level, focusing on lambda calculus and higher-order logic, and exploring how the meaning of complex expressions is constructed from their components, as well as phenomena in discourse that require contextual and hierarchical considerations.

- Understand theories of semantic and discourse representation.
- > Apply lambda calculus and logic to semantics.





#### Mandatory for students who picked the "ILPC" track.

This course introduces students to advanced techniques for symbolic knowledge extraction from unstructured textual data.

Students will learn to:

- > Use techniques for extracting concepts and relationships from unstructured texts.
- > Create symbolic knowledge bases from textual data.
- Integrate these knowledge bases into AI systems to enhance understanding and decision-making.

CM - 10 hours $TD - 10$ hours	TOTAL – 20 hours	COEF — 1
-------------------------------	------------------	----------

# REASONING IN AI

Zied BOURAOUI – zied.bouraoui@cril.fr

Mandatory for students who picked the "ILPC" track.

This course explores the philosophy and concepts of symbolic artificial intelligence, focusing on symbolic representations, reasoning rules, and cognitive modelling, while analysing the strengths and limitations of symbolic AI and hybrid approaches, alongside ethical questions related to these systems.

Students will learn to:

- Analyse the philosophical foundations of symbolic AI and its implications for modelling artificial intelligence.
- Discuss the limitations and advantages of the symbolic approach compared to deep learning-based methods.
- Evaluate the ethical and societal implications of symbolic AI in the context of technological evolution.

CM – 10 hours TD – 10 hours TOTAL – 20 hours COEF – 1

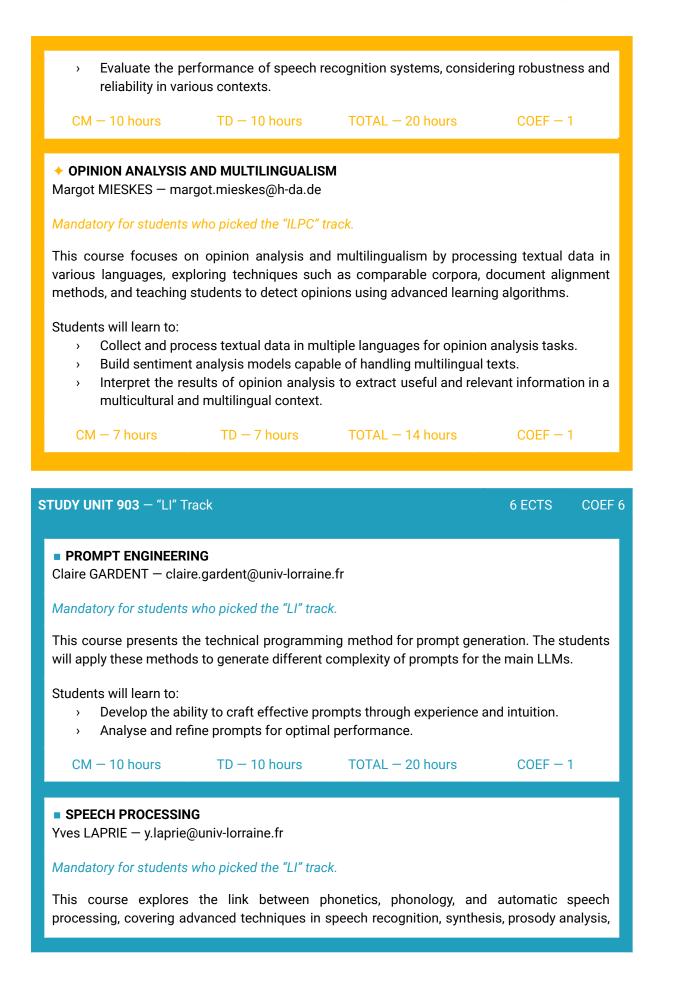
## ROBUST SPEECH PROCESSING

Romain SÉRIZEL – romain.serizel@univ-lorraine.fr

#### Mandatory for students who picked the "ILPC" track.

This course focuses on the robustness of speech recognition systems, addressing variability in speech, regional accents, background noise, and acoustic distortions, while teaching students to identify vulnerabilities and enhance system reliability in real-world situations through practical case studies.

- Identify the specific challenges related to speech variability and sources of acoustic disturbance.
- Develop techniques to improve the resilience of speech recognition systems under adverse conditions.



and addressing challenges like pronunciation variability, regional accents, and background noise management. Students will learn to: Understand the basic principles of speech recognition and speech analysis. > Apply advanced techniques for analysing the acoustic and prosodic features of > speech. Develop speech processing systems for applications such as automatic transcription > or speech synthesis. CM - 10 hours TD - 10 hours TOTAL - 20 hours COEF - 1PRAGMATICS Mathilde DARGNAT - mathilde.dargnat@univ-lorraine.fr Mandatory for students who picked the "LI" track. This course explores linguistic pragmatics, focusing on effective communication through concepts like reference, presupposition, implicature, and speech acts, while connecting these to formal representations in semantics and discourse theory. Students will learn to: Analyse speech acts and implicatures in real communicative contexts. > Understand how pragmatics influences the meaning of statements beyond syntax. > Apply pragmatic principles to analyse discourse and complex linguistic interactions. > CM - 7 hours TD - 7 hours TOTAL - 14 hours **COEF** - 1 LEXICOLOGY Yvon KEROMNES - yvon.keromnes@univ-lorraine.fr Mandatory for students who picked the "LI" track. This course introduces lexicology and lexicography, focusing on synchronic and diachronic analysis, differences between them, and translating lexicological research into dictionary descriptions.

- > Study the structure and formation of words in different languages.
- > Analyse the semantic properties of lexemes and their historical evolution.
- > Use lexical analysis methods to extract relevant lexical information from texts.

CM – 7 hours	TD – 7 hours	TOTAL – 14 hours	COEF — 1
TERMINOLOGY Alain POLGUÈRE – alain	n.polguere@univ-lorrair	ne.fr	

### Mandatory for students who picked the "LI" track.

This course focuses on terminology, teaching students to identify, extract, and standardise domain-specific terms, while exploring methods for building terminological databases and aligning multilingual terminologies.

Students will learn to:

- > Create glossaries and specialised terminology databases for specific domains.
- Manage multilingual terminologies and align technical terms across different languages.
- Apply terminological skills for specialised translation and information retrieval in professional contexts.

CM – 7 hours	TD – 7 hours	TOTAL – 14 hours	COEF - 1

#### DIALOGUE ENGINEERING

Maxime AMBLARD – maxime.amblard@univ-lorraine.fr

# Mandatory for students who picked the "LI" track.

This course presents the definitions of linguistic models of dialogue, construction of a corpus, definition of an annotation system and production of a resource usable for chatbots.

Students will learn to:

- > Design interactive dialogue systems tailored to user needs.
- > Develop chatbots capable of engaging in natural and useful conversations.
- Implement automated question-answering systems to effectively respond to a variety of questions in natural language.

TOTAL – 14 hours

```
CM - 7 hours
```

TD – 7 hours

STUDY UNIT 904 — Project and Language Courses

6 ECTS COEF 6

COEF - 1

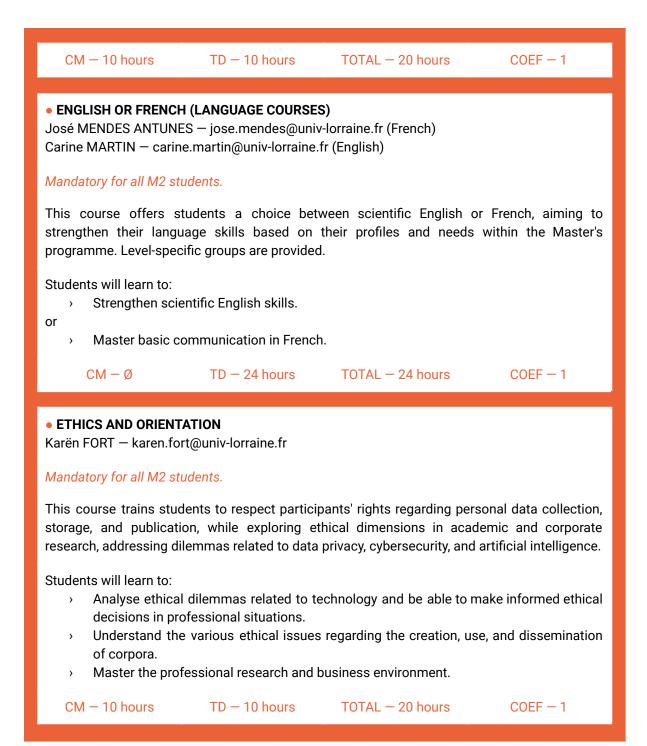
SOFTWARE PROJECT

Karën FORT – karen.fort@univ-lorraine.fr

## Mandatory for all M2 students.

This course focuses on developing essential programming and software engineering skills, teaching students to design, develop, and test software through practical projects while familiarising them with Agile methodologies and best practices in software quality.

- > Master programming techniques and be able to design functional software.
- > Apply software project management tools to plan and execute development projects.
- > Evaluate software quality by testing and debugging code effectively.



# MSC IN NLP COURSE DESCRIPTION SEMESTER 10 – S10

Updated for academic year 2024-2025

Study Unit – Unité d'Enseignement (UE) Course – Élément Constitutif (EC)

Mandatory 
 + ILPC Track
 ILPC Track
 A Optional

STUDY UNIT 1001 – Internship	6 ECTS	COEF 6
• INTERNSHIP Karën FORT — karen.fort@univ-lorraine.fr		
Mandatory for all M2 students.		
The second semester of the M2 is dedicated to an internship, either laboratory, culminating in a thesis presentation before a jury.	er in a compai	ny or a