

LISBON SUMMER SCHOOL IN LINGUISTICS

PhD Programme

Lisboa, July 4-8, 2022

Venue

Faculdade de Ciências Sociais e Humanas – NOVA
Avenida de Berna, 26-C
1069-061 Lisboa – Portugal



AREA 2: TERMINOLOGY AND LEXICOGRAPHY

Course 1: International standard for the representation of open lexical data –
Laurent Romary (INRIA, France) – LIMITED to 12 attendees

COURSE DESCRIPTION

The course will provide both a general background and practical insights about the representation of lexical information. It will address the various forms of lexical data and show how these have been tackled by international standardization initiatives. The course will focus on the use of the TEI Lex 0 recommendations and how they can be used to represent a variety of lexical phenomena at various levels of granularity.

Day 1

2h - Modélisation de données lexicales standard – cadre général (ce que c'est, termino-lexico, ISO TEI, W3C, Ontolex) – cf. Vienne/Parthenos

1h – Atelier : Introduction pratique à Oxygen et XML

Day 2

1h – Présentation du modèle TEI

2h – Atelier : Pratique de TEI Lex 0 sous Oxygen

Day 3

1h30 – Xpath et XSLT

1h30 – Atelier : XSLT pour la présentation de données lexicales

Day 4

1h – Science ouverte – publications, données (FAIR), logiciels

2h – Atelier: codage, transformation (suite)

Day 5

1h30 – De TEI Lex 0 à Ontolex

1h30 – Atelier : création de sorties Ontolex à partir de ses propres données

Course 2: Potencialidades no uso de corpos (linguísticos): estudos de tradução, extração de informação e leitura distante – Diana Santos (Linguatca & University of Oslo, Norway)

COURSE DESCRIPTION

Neste curso vamos ilustrar vários usos de corpos em português, usando os serviços da Linguatca.

Nos primeiros dois dias (segunda e terça) os alunos aprenderão a linguagem de procura e familiarizar-se-ão com a diversidade de recursos e de ferramentas acessíveis.

Nos dias seguintes terão a oportunidade de lidar com os três assuntos constantes do título: estudos de tradução, extração de informação e leitura distante.

Na quarta usaremos os corpos paralelos CorTrad (português-inglês) e PANTERA (português-norueguês), para fazer 3 mini-estudos: a) diferenças na tradução de emoções; b) a tradução de cores na culinária; e c) diferenças no uso dos tempos entre português original e português traduzido do norueguês.

Na quinta usaremos o corpo DHBB para estudar alguns assuntos de política brasileira moderna. Também extrairemos alguma informação sobre lugares na literatura portuguesa, usando o corpo Literateca.

Na sexta exploraremos o corpo Literateca, sob várias vertentes: trataremos pelo menos da relação entre a escola literária e algumas características linguísticas, e da identificação de personagens e outras pessoas.

Course 3: FAIR Terminology: design and implementation of terminological databases for open science – Federica Vezzani (University of Padova, Italy), Giorgio Maria Di Nunzio (University of Padova, Italy), Rute Costa (NOVA CLUNL)

COURSE DESCRIPTION

This course will focus on FAIR terminology as a new paradigm for the organization of Findable, Accessible, Interoperable, and Reusable terminological (meta-)data. We will present the theoretical and methodological background for the optimal representation of terminological data both from an abstract and implementation viewpoint. In this course, the attendees will have the possibility to put immediately into practice the methodological aspects and understand, step-by-step, how to tackle the main issues and make the right choices for the design of a terminological database fulfilling the Open Science requirement. Starting from scratch, each participant will be able to draw her/his own solution and compile the corresponding terminological records following the international ISO standards for terminology management by means of the FAIRterm Web Application. The main topics addressed are: Terminological Markup Framework (TMF) metamodel, TermBase eXchange (TBX) format, FAIR principles, Entity-Relationship model and Linked Open Data.

SCHEDULE

July 4-8	Formal and Experimental Linguistics	Terminology and Lexicography	Grammar & Text
9:30am – 12:30pm	Jacopo Torregrossa	Laurent Romary	Dominique Maingueneau
2:00pm – 5:00pm	Anabela Rato	Diana Santos	Álvaro Octavio de Toledo
5:30pm – 8:30pm	João Veríssimo	Federica Vezzani, Giorgio Maria Di Nunzio, Rute Costa	Ines Garcia Azkoaga