UniDive

2nd General Meeting

University of Naples L'Orientale, Naples, Italy, 8-9 February 2024

https://unidive.lisn.upsaclay.fr/

WG1, WG3

cost
EUROPEAN COOPERATION
IN SCIENCE & TECHNOLOGY

Funded by
the European Union

Universal Dependencies Treebank for Standard Albanian

Nelda Kote[1], Anila Çepani Sema[2], Alba Haveriku[1]

[1]Polytechnic University of Tirana, Tirana, Albania

[2]University of Tirana, Tirana, Albania

CONTRIBUTION

➢ A UD treebank for the Standard Albanian language, created in collaboration between linguistics and information technology experts.

➢ 25,000 tokens, 1,300 sentences.

ANNOTATED CORPORA

➢ Sentence segmentation;
➢ Words segmentation within a sentence;
➢ Lemmatization;
➢ Part-of-speech tags;
➢ Morphological features;
➢ Syntactic annotation.