

UDFest-BR 2022

Universal Dependencies Brazilian Festival

Proceedings of the Conference, Vol. 1

March 21, 2022

About the workshop

The Universal Dependencies Brazilian Festival (UDFest-BR) took place with the 15th edition of the International Conference on the Computational Processing of Portuguese (PROPOR 2022). It is a forum in which researchers involved with the study and application of the Universal Dependencies (UD) model, along with its adaptation to Portuguese, can meet and discuss best practices and strategies, as well as problems and solutions related to this topic.

UD is a cross-language international model for grammatical annotation (morphosyntactic tags, lexical/morphosyntactic features and syntactic dependencies), which results from the efforts of an open community of hundreds of researchers. Building upon a common framework, the model captures idiosyncrasies and similarities between different languages, aiming at fostering contrastive linguistic studies and the development of multilingual technologies for Natural Language Processing (NLP).

In recent years, this model has drawn considerable attention from the NLP community. This growing interest can be owed mainly to the fact that UD allows for the development of fairly precise taggers and parsers to a number of languages, based on the exploitation of syntactically annotated corpora (treebanks) and some Machine Learning algorithms. Currently, we find around 200 UD annotated corpora available in more than 100 languages.

Although there can be found some resources available in Portuguese, computational linguistic researches involving UD and this language are still incipient. As such, the existence of a workshop dedicated to this topic certainly helps to fill in this gap, giving more visibility to current efforts in this area as well as fostering new initiatives and collaborations.

This edition of the workshop includes six papers of renewed researchers in the area for the Portuguese language. There are theoretical discussions, reports on corpus annotation and the related challenges, and software for UD-related processing. The papers and their authors are:

- *Anotação de textos não canônicos: um estudo exploratório de Grande sertão: veredas pelas dependências universais*
André V. L. Coneglian, Ana Luísa A. R. Guimarães, Thiago C. Ferreira, Adriana S. Pagano
- *Polishing the gold – how much revision do we need in treebanks?*
Elvis de Souza, Cláudia Freitas
- *Que simples que nada: a anotação da palavra que em córpus de UD*
Magali Duran, Heloísa Oliveira, Clarissa Scandarolli
- *Shallow parsing of Portuguese texts annotated under Universal Dependencies*
Guilherme M. Oliveira, Paulo Berlanga Neto, Evandro Ruiz

- *Still on arguments and adjuncts: the status of the indirect object and the adverbial adjunct relations in Universal Dependencies for Portuguese*
Elvis de Souza, Cláudia Freitas
- *UDConcord: A Concordancer for Universal Dependencies Treebanks*
Lucas G. M. Miranda, Thiago A. S. Pardo

Organizing committee

This workshop is promoted by the [POeTiSA project](#), which is a long-term initiative that aims at growing syntax-based resources and developing related tools and applications for Brazilian Portuguese language, looking to achieve world state-of-the-art results in this area. The project is part of the Natural Language Processing initiative (NLP2) of the [Center for Artificial Intelligence](#) (C4AI) of the University of São Paulo, sponsored by IBM and FAPESP.

The following researchers organized this edition of the workshop:

Thiago Alexandre Salgueiro Pardo
University of São Paulo (USP)
São Carlos/SP, Brazil

Ariani Di Felippo
Federal University of São Carlos (UFSCar)
São Carlos/SP, Brazil

Norton Trevisan Roman
University of São Paulo (USP)
São Paulo/SP, Brazil

Program committee

Adriana Silvina Pagano (Brazil)
Cláudia Freitas (Brazil)
Daniel Zeman (Czech Republic)
Joakim Nivre (Sweden)
Jorge Baptista (Portugal)
Leonardo Zilio (UK)
Lucelene Lopes (USA)
Magali Sanches Duran (Brazil)
Maria das Graças Volpe Nunes (Brazil)
Valeria de Paiva (USA)