

CEUR Workshop Proceedings

Proceedings of the Doctoral Symposium on Natural Language Processing 2025 (NLP-DS-2025)

Papers

Integrating Subjectivity Analysis for Disinformation Detection in Spanish Alba Pérez-Montero	1-9
Determining and Evaluating the Quality of Corpora in the LLM Era Lucía Sevilla-Requena.....	10-18
Enhancing Small Open-Source Language Models for Ontology Generation through Metric-Guided Continual Pretraining Miquel Canal-Esteve.....	19-28
Bias Detection and Mitigation for the Development of Fair and High-Quality Automatic Text Simplification Corpora Victoria Muñoz-García.....	29-36
Diagnostic Assessment of Spanish as a Foreign Language using Technology based on Natural Language Processing María Victoria Cantero-Romero.....	37-42
Enhancing Trustworthiness in NLP Systems Through Comprehensive Explainability Approaches Santiago González-Silot.....	43-52
Separating Linguistic Competence from Factual Knowledge in Large Language Models Jaime Collado-Montañez.....	53-58

Automatic Counter-Narrative Generation	
María Estrella Vallecillo Rodríguez.....	59-67
Document Level Information Extraction in Low-Resource Languages	
Mikel Zubillaga.....	68-75
Early Explanatory Prediction of Cardiovascular Risk	
Ane G. Domingo-Aldama.....	76-84
Large Language Models as Educational Evaluators: Reasoning and Explainability in Low-Resource Language Assessment	
Ekhi Azurmendi.....	85-93
Optimization, Adaptation and Applications of Large Language Models	
David Ponce.....	94-103
Text Adaptation for Easy Read Content Generation	
Jesús Calleja.....	104-112
Language Variability in Basque: Data Creation, Evaluation and Systems for Dialect and Register Awareness in Natural Language Processing	
Jaione Bengoetxea.....	113-121
Multimodal Analysis of Emotion and Harmful Content in Online Communication	
Ronghao Pan.....	122-130
Textual Analysis of Political Trust in Spanish	
Tomás Bernal-Beltrán.....	131-138
Low-Resource Spanish Clinical Encoders: Architectures, Adaptation, and Evaluation under Computational Constraints	
Guillem G. Subies.....	139-145
Controlled Recipe Generation: Adapting Food Recipes to Meet Dietary Restriction	
Kemalcan Bora.....	146-154
Application of Natural Language Processing Techniques to Aid in Detecting and Monitoring People at Risk of Mental Illness	
Andreu Casamayor Segarra.....	155-162
Automatic Classification of the Economic Activity of a Company Using ML and DL Techniques	
Roi Santos-Ríos.....	163-171
Intelligent Assistants in the Era of LLMs: A New Methodology for Reusing Research Software from Documentation	
Carlos Utrilla Guerrero.....	172-180



IMPULSO DE LAS LENGUAS
EN LA INTELIGENCIA ARTIFICIAL



SEPLN

SOCIEDAD ESPAÑOLA
PARA EL PROCESAMIENTO
DEL LENGUAJE NATURAL



Adimen Artifiziala
eta Hizkuntzaren Teknologia
Katedra
Cátedra
de Inteligencia Artificial
y Tecnología del Lenguaje

HiTZ

Hizkuntza Teknologiako Zentroa
Basque Center for Language Technology

CEUR Workshop Proceedings

Proceedings of the Doctoral Symposium on Natural Language Processing 2025
(NLP-DS-2025)

Organization

Organizing committee

Alba Bonet Jover	Universidad de Alicante	alba.bonet@dlsi.ua.es
Beatriz Botella Gil	Universidad de Alicante	beatriz.botella@ua.es
Elena Lloret Pastor	Universidad de Alicante	elloret@dlsi.ua.es
M. Dolores Molina González	Universidad de Jaén	mdmolina@ujaen.es
Salud María Jiménez Zafra	Universidad de Jaén	sjzafra@ujaen.es

Edited in: Universidad de Jaén and Universidad de Alicante

Year: 2025

Editors

Alba Bonet Jover	Universidad de Alicante	alba.bonet@dlsi.ua.es
Beatriz Botella Gil	Universidad de Alicante	beatriz.botella@ua.es
Elena Lloret Pastor	Universidad de Alicante	elloret@dlsi.ua.es

M. Dolores Molina González

Universidad de Jaén

mdmolina@ujaen.es

Salud María Jiménez Zafra

Universidad de Jaén

sjzafra@ujaen.es

Published by: CEUR Workshop Proceedings

Program Committee

Alberto Bugarín

Universidad de Santiago de Compostela

Alberto Díaz Esteban

Universidad Complutense de Madrid

Paloma Martínez

Universidad de la Carlos III de Madrid

Andrés Montoyo Guijarro

Universidad de Alicante

Estela Saquete Boró

Universidad de Alicante

Isabel Espinosa

Universidad de Alicante

Robiert Sepúlveda

Universidad de Alicante

José Manuel Perea Ortega

Universidad de Extremadura

Manuel Carlos Díaz Galiano

Universidad de Jaén

Miguel Ángel García Cumbreiras

Universidad de Jaén

Eugenio Martínez Cámara

Universidad de Jaén

Arturo Montejo Ráez

Universidad de Jaén

Miguel Ángel Alonso Pardo

Universidad de A Coruña

Rafael Valencia García

Universidad de Murcia

Fermin Cruz

Universidad de Sevilla

José Antonio Troyano

Universidad de Sevilla

Gemma Bel

Universidad Nacional Autónoma de México

Luis Chiruzzo

Universidad de la República de Uruguay

Julio Gonzalo

Universidad Nacional de Educación a Distancia

Paolo Rosso

Universidad Politécnica de Valencia

Aitor Soroa

UPV/EHU

Rodrigo Agerri

UPV/EHU

Aitziber Atucha

UPV/EHU

Koldo Gojenola

UPV/EHU

German Rigau

UPV/EHU

Miguel Ángel Álvarez Carmona

CIMAT, México

Montserrat Cuadros

Vicomtech, España

Acknowledgements

These proceedings have been partially supported by projects with reference 2022/TL22/00215337, 2022/TL22/00215336, 2022/TL22/00215335 and 2022/TL22/00215334 funded by Ministerio de Asuntos Económicos y Transformación Digital and Plan de Recuperación, Transformación y Resiliencia - Funded by the European Union - NextGenerationEU; and Cátedra HiTZ de Inteligencia Artificial y Tecnología de la Lengua (TSI100923-2023-1) supported by Plan de Recuperación, Transformación y Resiliencia - Funded by the European Union, as well as the Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales.

ILENIA

IMPULSO DE LAS LENGUAS
EN LA INTELIGENCIA ARTIFICIAL



SEPLN

SOCIEDAD ESPAÑOLA
PARA EL PROCESAMIENTO
DEL LENGUAJE NATURAL



Adimen Artifiziala
eta Hizkuntzaren Teknologia
Katedra
Cátedra
de Inteligencia Artificial
y Tecnología del Lenguaje

HiTZ

Hizkuntza Teknologiako Zentroa
Basque Center for Language Technology

CEUR Workshop Proceedings

Proceedings of the Doctoral Symposium on Natural Language Processing 2025
(NLP-DS-2025)

Preamble

The Spanish Society for Natural Language Processing has organized the 8th edition of the Doctoral Symposium on Natural Language Processing held in Zaragoza on September 25, 2025 as part of the [XLI International Conference of the Spanish Society for Natural Language Processing](#) (SEPLN 2025).

The aim of the Doctoral Symposium is to promote scientific vocations in the field of Language Technologies and Natural Language Processing (NLP). To this end, it organizes a meeting point where PhD students in the area of NLP can present their work. Moreover, they receive constructive feedback and guidance from experienced researchers in their area of work.

In the 2025 edition, 21 PhD students participated. All participants submitted a paper describing their proposals, all of them being accepted after review by two members of the program committee and by the organizing committee.

November 2025

The editors

Copyright © 2025 for the individual papers by the papers' authors.
Copyright © 2025 for the volume as a collection by its editors. This volume and its papers are published under the Creative Commons License Attribution 4.0 International (CC BY4.0).