



**LANGUAGE, DATA and  
KNOWLEDGE 2025**

Proceedings  
of the 5th  
Conference on  
**Language, Data  
and Knowledge:  
Workshops**



UniorPress  
Naples  
2025

PROCEEDINGS OF THE 5TH CONFERENCE ON LANGUAGE, DATA  
AND KNOWLEDGE: WORKSHOPS

EDITORS:

Katerina Gkirtzou, ILSP “Athena” Research Center, Greece  
Slavko Žitnik, University of Ljubljana, Slovenia  
Jorge Gracia, University of Zaragoza, Spain  
Dagmar Gromann, University of Vienna, Austria  
Maria Pia di Buono, University of Naples “L’Orientale”, Italy  
Johanna Monti, University of Naples “L’Orientale”, Italy  
Maxim Ionov, University of Zaragoza, Spain



The proceedings are licensed under  
Creative Commons Attribution 4.0 International

ISBN 978-88-6719-334-9

UniorPress — University of Naples “L’Orientale”  
Via Nuova Marina 59 — 80133 Napoli (Italy)



UniorPress

# Table of Contents

## **TermTrends 2025: Bridging the Gap between Terminological Resources and Large Language Models**

<i>Introduction</i> .....	1
<i>The LegISTyr Test Set: Investigating Off-the-Shelf Instruction-Tuned LLMs for Terminology-Constrained Translation in a Low-Resource Language Variety</i> Paolo Di Natale, Egon W. Stemle, Elena Chiocchetti, Marlies Alber, Natascia Ralli, Isabella Stanizzi and Elena Benini .....	6
<i>Terminology Management Meets AI: The ISO/TC 37/SC 3/WG 6 Initiative</i> Mohamed Khemakhem, Cristina Valentini, Natascia Ralli, Sérgio Barros, Georg Löckinger, Federica Vezzani, Ana Salgado, Zhenling Zhang, Sabine Mahr, Sara Carvalho, Klaus Fleischmann and Rute Costa .....	21
<i>Inferring Semantic Relations Between Terms with Large Language Models</i> Giulia Speranza .....	30

# **Term Trends 2025**

## **TermTrends 2025: Bridging the Gap between Terminological Resources and Large Language Models**

The Trends in Terminology Generation and Modelling (TermTrends) workshop series started 3 years ago as a tutorial co-located with the 23rd International Conference on Knowledge Engineering and Knowledge Management (EKAW) 2022. This tutorial covered standardisation approaches for representing terminology, explored methods to accelerate terminology generation, demonstrated domain-specific use cases, and included a hands-on session with tools for terminology extraction, enrichment, and representation.

The concept of the tutorial quickly evolved into a workshop format, featuring a different topic each year based on the emerging trends and priorities within the field. Therefore, TermTrends 2023 embraced the Terminology in the Era of Linguistic Data Science, and it was co-located with the 4th Conference on Language, Data and Knowledge (LDK). In that edition, the papers presented addressed diverse challenges and innovations in multilingual and domain-specific terminology work, including the complexity of representing multilingual specialised knowledge, the development of terminology resources using OntoLex-Lemon, the theoretical and practical considerations of converting TBX to OntoLex-Lemon to ensure interoperability, and the formalisation of translation equivalence and lexically semantic relations in bilingual terminological resources, amongst other topics.

Last year, the workshop focused on Models and Best Practices for Terminology Representation in the Semantic Web, and it was co-located with the 3rd International Conference on Multilingual Digital Terminology Today. Design, representation formats and management systems (MDTT 2024). The contributions explored the intersection of lexical and domain knowledge, with a special focus on terminology representation and interoperability, including the linking of resources in the cybersecurity domain, the development of a three-layer multilingual terminology for oncological diseases, the representation of the TERMCAT glossaries in OntoLex-Lemon and the integration of the Lexical Markup Framework with the Terminological Markup Framework.

This year's edition of the workshop focused on a topic that has been rapidly adopted by nearly every sector of society and holds significant implications for language processing, knowledge representation, and artificial intelligence: Large Language Models. Therefore, the workshop is focused on Bridging the Gap between Terminological Resources and Large Language Models. In particular, this event intends to explore how terminological resources can be leveraged to enhance the performance of LLMs at different stages of their lifecycle, from pre-training and fine-tuning to evaluation and application. By identifying points of convergence between Linked Data-based terminological data and LLM processes, we seek to identify strategies to make these resources more interoperable, reusable, and impactful.

Specifically, TermTrends 2025 presents contributions on the performance of instruction-tuned Large Language Models in handling terminology-constrained translation for low-resource language varieties, the ISO/TC 37/WG 6 initiative that aims to integrate AI into standardised terminology management practices, and the capabilities of LLMs in inferring semantic relations between terms to enhance terminological resources.

Additionally, this year's edition of the workshop introduced a notable novelty: the inclusion of two keynote speakers, Giorgio Maria di Nunzio, Associate Professor at the University of Padova, and José Manuel Gómez Pérez, Director of Language Technology Research at expert.ai. Having such distinguished representatives from academia and industry, respectively, will provide comprehensive and contrasting insights into the application of terminological resources and LLMs, highlighting both theoretical advancements and practical implementations across different sectors.

With the papers presented and the insightful contributions from both academic and industry represen-

tatives, we have high hopes for this year's edition. TermTrends 2025 aims to bring together researchers and practitioners to explore practical ways of connecting terminological resources with Large Language Models, encouraging collaboration and adoption of these technologies. We look forward to productive discussions and new opportunities to advance the field together.

Patricia Martín Chozas, Elena Montiel Ponsoda, Sara Carvalho and Federica Vezzani  
TermTrends 2025 Organising Committee

## **Organizing Committee**

Patricia Martín Chozas, Universidad Politécnica de Madrid, Spain

Elena Montiel Ponsoda, Universidad Politécnica de Madrid, Spain

Sara Carvalho, University of Aveiro, Portugal

Federica Vezzani, University of Padova, Italy

## Program Committee

Ana Ostroški Anić, Institute for Croatian Language and Linguistics, Croatia  
Anas Fahad Khan, Istituto di Linguistica Computazionale, Italy  
Antonio San Martín, University of Quebec, Canada  
Beatriz Guerrero, Universidad de Salamanca, Spain  
Bruno Almeida, NOVA FCSH / NOVA CLUNL, Portugal  
David Lindemann, University of the Basque Country, Spain  
Elena Montiel-Ponsoda, Universidad Politécnica de Madrid, Spain  
Federica Vezzani, Università di Padova, Italy  
Francesca Frontini, Istituto di Linguistica Computazionale, Italy  
Giorgio Di Nunzio, Università di Padova, Italy  
Giulia Speranza, Università degli Studi di Napoli L'Orientale, Italy  
John McCrae, National University of Ireland, Ireland  
Jorge Gracia, Universidad de Zaragoza, Spain  
Laurent Romary, INRIA, France  
Manuel Fiorelli, Tor Vergata University of Rome, Italy  
Maria Pia di Buono, Università degli Studi di Napoli L'Orientale, Italy  
Miguel Sánchez Ibáñez, Universidad de Valladolid  
Natascia Ralli, Eurac Research, Italy  
Nava Maroto, Universidad Politécnica de Madrid, Spain  
Pamela Faber, University of Granada, Spain  
Patricia Martín Chozas, Universidad Politécnica de Madrid, Spain  
Patrick Drouin, Montreal University, Canada  
Paul Buitelaar, National University of Ireland, Ireland  
Penny Labropoulou, Institute of Language and Speech, Greece  
Pilar León Arauz, Universidad de Granada, Spain  
Sara Carvalho, Universidade de Aveiro / NOVA CLUNL, Portugal  
Sigita Rackevičienė, Mykolas Romeris University, Lithuania  
Silvia Piccini, Istituto di Linguistica Computazionale, Italy  
Špela Vintar, University of Ljubljana, Slovenia