LREC 2026

2nd Workshop on Evaluating Text Difficulty in a Multilingual Context (DeTermIt! 2026)

Workshop Proceedings

Editors
Giorgio Maria Di Nunzio, Federica Vezzani, Liana Ermakova, Hosein Azarbonyad, and Jaap Kamps

11 May 2026

Proceedings of the 2nd Workshop on Evaluating Text Difficulty in a Multilingual Context (DeTermIt! 2026)

ISBN 978-2-493814-63-0

Preface

Automatic Text Simplification (ATS) is the process of reducing the linguistic and conceptual complexity of a text in order to improve its comprehensibility and readability. ATS plays an important role in supporting clear and accessible communication, while also serving as a useful preprocessing step for tasks such as information extraction, retrieval, and knowledge access. More broadly, ATS has significant societal implications, especially for people with limited literacy, second language learners, and readers who encounter difficulties when dealing with specialized or terminology-rich texts.

One of the main barriers to text understanding is the presence of unfamiliar terminology and domain-specific concepts. In many real-world settings, accessibility depends not only on simplifying syntax or vocabulary, but also on identifying which terms and conceptual relations are likely to create difficulty for different audiences. Lexical simplification, terminology adaptation, and readability assessment therefore represent complementary perspectives on the same challenge: how to make texts more accessible while preserving their informational value. In recent years, these questions have gained further relevance with the increasing use of Large Language Models (LLMs) and Generative AI (GenAI), which are now widely used to generate, rewrite, and mediate access to information.

The *DeTermIt! Evaluating Text Difficulty in a Multilingual Context* workshop explores the theoretical and practical perspectives surrounding the evaluation and modeling of text difficulty in multilingual and terminology-rich contexts. In today's interconnected world, where access to knowledge increasingly depends on language technologies, it is essential to ensure that information remains understandable, faithful, and usable for audiences with different linguistic backgrounds and levels of domain expertise.

From a *theoretical* point of view, the workshop discusses models of text difficulty, lexical and conceptual complexity, and the interaction between terminology and readability across languages. It promotes reflection on how complexity can be identified, described, and evaluated, and on how multilingual and cross-linguistic perspectives can contribute to more robust and inclusive approaches. From a *practical* standpoint, the workshop considers methods and resources for text simplification, lexical complexity prediction, controllable text generation, and the evaluation of LLM- and GenAI-based systems. Particular attention is devoted to multilingual datasets, interpretable frameworks, domain-specific case studies, and evaluation protocols that support reproducible and accessible research.

The central questions addressed by this workshop concern, first, the linguistic and conceptual factors that make a text difficult and, second, the methods that can be used to assess, control, or reduce such difficulty while preserving meaning and domain adequacy.

This first edition of DeTermIt! 2024 was co-located with the LREC-COLING 2024 joint conference and held in Turin, on May 21, 2024.[1] This second edition of DeTermIt! 2026 is co-located with LREC 2026 and held in Palma de Mallorca, Spain, on May 11, 2026.[2]

The submitted papers went through a double-blind review process that required at least three reviews by members of the international scientific committee. We accepted 9 papers out of 12 submissions (75% acceptance rate): 8 long papers and 1 short paper.

The papers included in these proceedings reflect the variety of current research on text difficulty and simplification. Several contributions focus on multilingual readability and lexical

[1] https://determit2024.dei.unipd.it/

[2] https://determit2026.dei.unipd.it/

complexity, including cross-linguistic analyses of translated children's literature, multilingual lexical complexity prediction for language learning, and calibrated interpretable frameworks for multilingual text difficulty prediction. Other papers address simplification and accessibility in specialized settings, such as biomedical text simplification, patient-centered explanations of gut–brain axis concepts, and the use of small language models as evaluators of simplification quality. The program also includes work on translation as augmentation for difficulty assessment, the automatic generation of graded texts in Old Church Slavonic, and terminology-augmented generation for culturally grounded translation. Overall, these contributions highlight the growing convergence of text simplification, terminology studies, multilingual evaluation, and generative AI.

The keynote speaker of the workshop is Horacio Saggion (Universitat Pompeu Fabra, Barcelona, Spain), who will deliver a talk entitled "Text Simplification as a Tool for Facilitating Democratic Participation in the iDEM Project". The talk addresses the role of accessibility and easy language in supporting democratic inclusion, especially for people with language barriers or limited literacy. It will present work carried out in the Horizon Europe iDEM project, including the creation of language resources and the development of text simplification services integrated into a mobile application for deliberative democratic participation.

We would like to thank all authors for submitting their work, the members of the scientific committee for their careful and constructive reviews, and the LREC 2026 workshop organizers for hosting this second edition of DeTermIt!. We hope that these proceedings will contribute to the ongoing discussion on how terminology-aware, multilingual, and evaluation-driven approaches can support more accessible and effective communication in the age of generative AI.

Giorgio Maria Di Nunzio - University of Padua, Italy
Federica Vezzani - University of Padua, Italy
Liana Ermakova - University of Western Brittany
Hosein Azarbonyad - Elsevier, The Netherlands
Jaap Kamps - University of Amsterdam, The Netherlands

Organizing Committee

General Chairs
Giorgio Maria Di Nunzio - Università degli Studi di Padova, Italy
Federica Vezzani - Università degli Studi di Padova, Italy
Liana Ermakova - Université de Bretagne Occidentale, France
Hosein Azarbonyad - Elsevier, The Netherlands
Jaap Kamps - University of Amsterdam, The Netherlands

Scientific Commitee
Florian Boudin - Nantes University, France
Lynne Bowker - University of Ottawa, Canada
Sara Carvalho - Universidade de Aveiro, Portugal
Rute Costa - Universidade NOVA de Lisboa, Portugal
Eric Gaussier - University Grenoble Alpes, France
Natalia Grabar - CNRS, France
Ana Ostroški Anić - Institute of Croatian Language and Linguistics, Croatia
Tatiana Passali - Aristotle University of Thessaloniki
Grigorios Tsoumakas - Aristotle University of Thessaloniki
Sara Vecchiato - University of Udine, Italy
Cornelia Wermuth - KU Leuven, Belgium

Table of Contents