Preface

Federica Vezzani[1], Giorgio Maria Di Nunzio[2,*], Vanessa Bonato[1] and Gianmaria Silvello[2]

[1]Department of Linguistic and Literary Studies, University of Padua, Italy
[2]Department of Information Engineering, University of Padua, Italy

1. Foreword to the DETECH 2026 Proceedings

The International Workshop on Definition and Term Extraction Challenge (DETECH) 2026 was held on 24 June 2026 in Zadar, Croatia, as a co-located event of the 5th International Conference on Multilingual Digital Terminology Today (MDTT) 2026.[1] DETECH 2026 was organized as a HEREDITARY challenge devoted to the automatic extraction of domain-specific terms and the generation of natural language definitions for medical concepts. Its main objective was to advance research on explainable, data-driven medical terminology at the intersection of terminology, natural language processing, and biomedical text analysis.

The 2026 edition focused on the gut–brain interplay, a biomedical domain at the crossroads of gastroenterology, neuroscience, and genetics. This domain provides a challenging and realistic setting for terminology-oriented natural language processing, since it involves complex medical concepts, interdisciplinary terminology, multi-word expressions, abbreviations, and evolving scientific knowledge. DETECH 2026 was organized around two complementary tasks: Task A, devoted to the extraction of relevant single-word and multi-word terms from English biomedical texts, and Task B, devoted to the generation of natural language definitions for the concepts designated by the extracted terms.

The challenge material was distributed through the DETECH 2026 GitHub repository.[2] Participants were allowed to submit up to five runs per subtask. External resources, including pre-trained models, lexicons, ontologies, and terminology resources, were permitted, provided that their use was clearly documented. Task A was evaluated through Micro-F1 and Type-F1, in order to account for both occurrence-level extraction and type-level terminology identification. Task B was evaluated through BLEU, BERTScore, and additional qualitative checks addressing the informativeness and linguistic quality of generated definitions.

All papers included in these proceedings underwent a double-blind review process, with each submission evaluated by at least two members of the organizing committee. We received and accepted six papers in these proceedings which reflect a broad methodological spectrum. The accepted contributions address symbolic and rule-based term extraction, BERT-based sequence labelling, biomedical transformer models, customized generative AI, ontology-enhanced prompting, retrieval-augmented definition generation, and theoretically informed approaches to terminological meaning. Taken together, they show that term extraction and definition generation are best understood as connected knowledge-oriented tasks, requiring not only computational performance but also terminological adequacy, biomedical grounding, explainability, and expert-informed evaluation.

DETECH 2026 was chaired by Federica Vezzani, Giorgio Maria Di Nunzio, Vanessa Bonato, and Gianmaria Silvello from the University of Padua. The organizers gratefully acknowledge the authors, reviewers, participants, and all those who contributed to the preparation of the challenge, the evaluation

International Workshop on Definition and Term Extraction Challenge (DETECH) 2026, June 24, 2026, Zadar, Croatia
*Corresponding author.
federica.vezzani@unipd.it (F. Vezzani); giorgiomaria.dinunzio@unipd.it (G. Di Nunzio); vanessa.bonato@phd.unipd.it (V. Bonato); gianmaria.silvello@unipd.it (G. Silvello)
0000-0003-2240-6127 (F. Vezzani); 0000-0001-9709-6392 (G. Di Nunzio); 0009-0002-9918-282X (V. Bonato); 0000-0003-4970-4554 (G. Silvello)

[1]https://detech2026.dei.unipd.it/
[2]https://github.com/gmdn/DETECH2026/

process, and the workshop programme. We also acknowledge the support of the HEREDITARY project, the University of Zadar, the Institute for the Croatian Language, and NextGenerationEU. Their support made it possible to establish DETECH as a shared evaluation setting for research on definition generation, automatic term extraction, and terminology-aware biomedical NLP.