

EVALITA 2026

**The Ninth Evaluation Campaign of Natural Language
Processing and Speech Tools for Italian**

Proceedings of the Workshop

February 26-27, 2026

Copyright ©2026 for the individual papers by the papers' authors.
Copyright ©2026 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 BY 4.0).

These proceedings have been published in the CEUR Workshop Proceedings series.
The original papers are available at: <https://ceur-ws.org/Vol-4195>.
The papers are mirrored in the ACL Anthology.

ISBN 979-12-243-0586-6

Table of Contents

Preface

Francesco Cutugno, Alessio Miaschi, Alessio Palmero Aprosio, Giulia Rambelli, Lucia Siciliani and Marco Antonio Stranisci 1

EVALITA 2026: Overview of the 9th evaluation campaign of natural language processing and speech tools for italian

Francesco Cutugno, Alessio Miaschi, Alessio Palmero Aprosio, Giulia Rambelli, Lucia Siciliani and Marco Antonio Stranisci 5

Tu chiamale se vuoi emozioni: The impact of emotions on disinformation and sexism identification

Paolo Rosso 14

Kenji-Endo: a BabyLM @EVALITA

Calogero Jerik Scozzaro, Matteo Rinaldi, Gianluca Mittone and Marco Antonio Stranisci 16

DeSegMa-IT at EVALITA 2026: Overview of the "Detection and Segmentation of Machine Generated Texts in Italian"

Giovanni Puccetti, Andrea Pedrotti and Andrea Esuli 27

Gradient Descenders at DeSegMa-IT: Leveraging Monolingual Transformer for LLM-Generated Text Detection and Boundary Identification

Tran Phuoc Thanh Nhan, Bui Hong Son and Dang Van Thin 42

MINDS at DeSegMa-IT: Detecting Human-LLM Authorship Switches via Token-Level Classification

Flavio Giobergia 53

Nicla at DeSegMa-IT: DistilBERT for MGT Detection and LightGBM Regressor for HMT Segmentation

Nicla Auletta 59

Stochastic Gradient Descenders at DeSegMa-IT: Instruction-Tuned LLM and Token Classification for MGT Detection and Boundary Localization

Huynh Nguyen Phu, Bui Hong Son and Dang Van Thin 67

UniTor at DeSegMa-It: Analyzing Supervision and Encoder Representations for Italian Machine-Generated Text Detection

Federico Borazio, Giacomo De Luca, Daniele Pasquini, Danilo Croce and Roberto Basili 78

GSI:detect at EVALITA 2026: Overview of the Task on Detecting Gender Stereotypes in Italian

Gloria Comandini, Manuela Speranza, Sofia Brenna, Davide Testa, Stefania Cavagnoli and Bernardo Magnini 96

DIAG-Sapienza at GSI:detect: Joint Detection and Classification of Gender Stereotypes with Structured Prompting and Fine-Tuning

Olga E. Sorokoletova, Emanuele Musumeci and Daniele Nardi 109

Festa at GSI:detect: In-Context Learning vs. Fine-Tuning for Gender Stereotype Detection

Donato Festa 123

MINDS at GSI:detect: From Logits to Degrees of Agreement in Gender Stereotype Detection with LLMs

Flavio Giobergia 131

<i>StereoBusters at GSI:detect: LLM-Based Detection and Human Qualitative Analysis of Gender Stereotypes in Italian Short Text</i>	
Salvatore Greco, Moreno La Quatra, Marta Marchiori Manerba, Ricardo Muñoz Sánchez and Alessandra Teresa Cignarella	138
<i>Team Prisma at GSI:detect: Comparing PB&J Persona-Based and Few-Shot Approaches to Gender Stereotype Detection at EVALITA 2026</i>	
Claudia Zaghi	157
<i>Tiz at GSI:detect: Modeling Gender Stereotype Detection as Multi-Category Gender Stereotype Scoring</i>	
Tiziano Labruna	169
<i>MultiPRIDE at EVALITA 2026: Overview of the Multilingual Automatic Detection of Slur Reclamation in the LGBTQ+ Context Task</i>	
Chiara Ferrando, Lia Draetta, Marco Madeddu, Mae Sosto, Viviana Patti, Paolo Rosso, Cristina Bosco, Jacinto Mata and Estrella Gualda	178
<i>AIWizards at MULTIPRIDE: A Hierarchical Approach to Slur Reclamation Detection</i>	
Luca Tedeschini and Matteo Fasulo	196
<i>AVAHI at MultiPRIDE: Multilingual Reclaimed Language Detection via Knowledge Graphs and Retrieval-Augmented Generation</i>	
Tania Alcántara, Omar García-Vázquez, José A. Torres-León, Marco Cardoso-Moreno, Diana Jiménez and Luis Moreno-Mendieta	206
<i>Challenger at MultiPRIDE: Is It Hate Speech or Reclaimed?</i>	
Hadi Bayrami Asl Tekanlou, Mahdi Bakhtiyarzadeh and Jafar Razmara	229
<i>DataSummit at MultiPRIDE: Context-Aware Multilingual Detection of Slur Reclamation in LGBTQ+ Contexts</i>	
Federico Dingeo and Marco Viviani	238
<i>Ghavidel-Rajabi at MultiPRIDE: Identity, Toxicity, or Complexity? A Language-Specific Feature Selection Approach to Reclamatory Intent Detection</i>	
Houman Rajabi, Fatemeh Ghavidel and Kourosh Ghahremani	251
<i>GRUPPETTOZZO at MultiPRIDE: Detecting LGBTQ+ Reclamatory Intent via Context-Aware Transformers</i>	
Federico Traina, Alessandro Santoro, Gabriele Greco, Irene Siragusa and Roberto Pirrone	262
<i>HateItOff at MultiPRIDE: Linguistic and Sentiment Cues in Reclaimed LGBTQ+ Slur Detection</i>	
Greta Damo and Nicolás Benjamín Ocampo	273
<i>I2C At MultiPRIDE: Transformers And Generative LLMs For Multilingual Reclamation Classification</i>	
Laura Vázquez-Ramos, Jacinto Mata-Vázquez and Victoria Pachón-Álvarez	284
<i>INFOTEC-NLP at MultiPRIDE: Region-Aware Transformer Models for Spanish Reclaimed Language Classification</i>	
Jorge Gleaves and Guillermo Ruiz	294
<i>KIT-TIP-NLP at MultiPride: Continual Learning with Multilingual Foundation Model</i>	
Barathi Ganesh HB, Michal Ptaszynski, Rene Melendez and Juuso Eronen	302
<i>LlaNa at MultiPRIDE: A Fine-Tuning Approach with Cross-Lingual Augmentation</i>	
Alessio Mercurio, Giorgio Talluto, Irene Siragusa and Roberto Pirrone	315

<i>MilaNLP at MultiPRIDE: Evaluating Lexical, Transformer, and Retrieval-Augmented Models for Reclaimed Language</i>	
Ivana Crescenzi, Arianna Muti and Debora Nozza	322
<i>NamDang at MultiPride@EVALITA 2026: Multilingual Classification of Reclaimed Language in LGBTQ+ Discourse using Transformer-based Models</i>	
Nam Dang, Vo Tuan Kiet and Dang Van Thin	335
<i>NetGuardAI at MultiPRIDE: Multilingual Detection of Reclaimed Language</i>	
Noe Come Jacques Le Pollotec, Elena-Simona Apostol and Ciprian-Octavian Truică	344
<i>SaFe Tweets at MultiPRIDE: From Multi-Model Ensembles to Expert LLMs for Reclamation Detection</i>	
Sara Visconti and Federica Manzi	352
<i>The Hate Busters at MultiPRIDE: Automatic Identification of Reappropriated Slurs in Multilingual LGBTQ+ Discourse</i>	
Aurora Ciminelli, Giulia Corvino, Camilla Gentili and Marco Viviani	366
<i>UniBO-FICLIT at MultiPRIDE: Fine-Tuning an ELECTRA-Based Model for the Detection of Italian Reclaimed Slurs</i>	
Simone Casazza	383
<i>Cruciverb-IT at EVALITA 2026: Overview of the Crossword Solving in Italian Task</i>	
Cristiano Ciaccio, Gabriele Sarti, Alessio Miaschi, Felice Dell’Orletta and Malvina Nissim . .	392
<i>AC/DG at Cruciverb-IT: Retrieval-Based Approaches for Italian Crossword Clue Answering</i>	
Ali Yassine, Claudio Savelli, Davide Napolitano, Giuseppe Gallipoli, Luca Cagliero and Elena Baralis	410
<i>FFT-UniBa at Cruciverb-IT: Special Length Tokens and CSP for Italian Crossword Solving</i>	
Andrea Porcelli, Filippo Di Gravina, Emanuele Fontana, Mattia Curri and Francesco Damiano Di Gregorio	421
<i>MINDS at Cruciverb-IT: Solving Italian Crossword Clues with Masked Language Models and Candidate Pooling</i>	
Flavio Giobergia	433
<i>UNIBA at Cruciverb-IT: Solving Italian Crosswords with Encoder–Decoder Models and Beam Search</i>	
Pierpaolo Basile	442
<i>UniTor at Cruciverb-IT: Retrieval-Augmented Two-Step Reasoning for Italian Crossword Clue Answering</i>	
Andriy Shcherbakov, Danilo Croce and Roberto Basili	452
<i>SVELA at EVALITA 2026: Overview of the Selective Verification of Erasure from LLM Answers Task</i>	
Claudio Savelli, Moreno La Quatra, Alkis Koudounas and Flavio Giobergia	467
<i>Eraserhead at SVELA: Detecting LLM Forgetting via Logit-Space Statistics</i>	
Matteo Berta and Tania Cerquitelli	477
<i>Ita-Lib at SVELA: Detecting the Forgotten — Representation-Based Approach for Verifying Machine Unlearning</i>	
Ali Yassine, Hadi Ibrahim and Luca Cagliero	484
<i>MALTO at SVELA: A Specific-Attention-Head Approach for Membership Inference Attacks in LLMs Unlearning Evaluation</i>	
Evren A. Munis, Mattia Sabato, Erfan Bayat and Andrea Lolli	495

<i>priyam_saha17 at SVELA: A Feature-Centric Pipeline for Verifying Selective Forgetting in Large Language Models</i>	
Priyam Saha	509
<i>EVWSD-ITA at EVALITA 2026: Overview of the Enhanced Visual Word Sense Disambiguation for Italian Task</i>	
Elio Musacchio, Lucia Siciliani, Pierpaolo Basile and Giovanni Semeraro	522
<i>UniTor at EVWSD-ITA: Zero-Shot Visual Word Sense Disambiguation via Visual Question-Answering</i>	
Claudiu D. Hromei, Antonio Scaiella, Danilo Croce and Roberto Basili	532
<i>ATE-IT at EVALITA 2026: Overview of the Automatic Term Extraction Italian Testbed Task</i>	
Nicola Cirillo, Giorgio Maria Di Nunzio and Federica Vezzani	550
<i>DSBKTE at ATE-IT: From Token Classification to Zero-Shot Generation: Two Approaches to Italian ATE at EVALITA 2026</i>	
Danil Smirnov, Bitu Khashechian and Giorgio Maria Di Nunzio	563
<i>JTTE at ATE-IT: A CRF Model with Contextual Embeddings</i>	
Juliette Tonneau and Giorgio Maria Di Nunzio	572
<i>MKTE at ATE-IT: CRF-Based Term Extraction for Italian Waste Management Documents</i>	
Minseok Kim and Giorgio Maria Di Nunzio	581
<i>OATE at ATE-IT: A Hybrid Approach for Automatic Terminology Extraction in Italian</i>	
Omar Arab and Giorgio Maria Di Nunzio	587
<i>Peacemaker at ATE-IT: Automatic term extraction from Italian text for waste management data using encoder model</i>	
Mahdi Bakhtiyarzadeh, Hadi Bayrami Asl Tekanlou and Jafar Razmara	593
<i>SMTE at ATE-IT: Ensemble Term Extraction with Italian BERT, spaCy, and Vocabulary-Based Filtering</i>	
Sofia Maule and Giorgio Maria Di Nunzio	602
<i>TermNinjas at ATE-IT: Overview of the Term Extraction and Term Variants Clustering Task</i>	
Zunaira Hasnain and Fashad Ahmed Siddique	614
<i>TrietNLP at ATE-IT: A Hybrid Pipeline for Italian Waste Management Terminology Analysis</i>	
Nguyen Minh Triet and Dang Van Thin	623
<i>VVTE at ATE-IT: From Candidates to Terms: Hybrid Italian ATE with Dependency Heuristics, Gemini, and Random Forest Filtering</i>	
Valentine G. L. Vandervoort and Giorgio Maria Di Nunzio	633
<i>PFB at EVALITA 2026: Overview of the Prometeia Financial Benchmark</i>	
Alessandro P. Bardelli, Tolga Çekiç, Irem Demirtaş, Michele Filannino, Simona Scala, Andrea Galassi, Gianmarco Pappacoda and Paolo Torroni	639
<i>AMSN at PFB: Overview of the Automated Multilingual Economics Question Answering Task</i>	
Ali Saleh Mohammadabad and Mehregan Nazarmohsenifakori	646
<i>UniTor at PFB: Constrained Agentic Prompting and Self-Consistency for Financial Multi-Choice QA</i>	
Federico Borazio, Seyed Alireza Mousavian Anaraki, Shahid Iqbal Rai, Danilo Croce and Roberto Basili	654
<i>FadeIT at EVALITA 2026: Overview of the Fallacy Detection in Italian Social Media Texts Task</i>	
Alan Ramponi and Sara Tonelli	673

<i>Label at FadeIT: Fallacy-Aware LLM Reasoning for Score-Based Classification</i>	
Tiziano Labruna and Eleni Papadopulos	686
<i>MALTO at FadeIT: A BERT-Based System for Multi-Label Fallacy Detection in Italian Social Media</i>	
Matin Salami, Luca M. Rodia, Vladimir Schiau, Evren A. Munis, Claudio Savelli and Flavio Giobergia	696
<i>PuDy at FadeIT: Enhancing Fine-Grained Fallacy Detection with Hierarchy-Aware Training and Contextual Enrichment</i>	
Duy Dang Phu, Son Bui Hong and Dang Van Thin	709
<i>RBG-AI at FadeIT: Prompted LLMs with Label Abstraction for Logical Fallacy Detection</i>	
Meenakshi, Jairam R, Reshma U, Barathi Ganesh HB and Michal Ptaszynski	721
<i>TIGRO at FadeIT: E Pluribus Unum – A Multi-task Approach to Fallacy Detection and Span Identification</i>	
Stefano Atzeni, Gabriele Sarti, Tommaso Caselli and Malvina Nissim	734
<i>Unica at FadeIT: Adapting Large Language Models to Fallacy Identification in Social Networks</i>	
Matteo Fenu and Maurizio Atzori	748
<i>IMPOLS at EVALITA 2026: Overview of the IMPOLS Task</i>	
Lorenzo Gregori, Walter Paci and Valentina Saccone	759
<i>RES2 at IMPOLS 2026: Implicit Content Classification via Span-Awareness, Hierarchical Multi-Task Learning and Synthetic Data</i>	
Salvatore Spezia, Salvatore Ferrara, Roberto Dioguardi, Ernesto Davì, Irene Siragusa and Roberto Pirrone	770
<i>INDAQ2 - A Large Italian Narrative QA Benchmark: A CALAMITA 2026 Challenge</i>	
Luca Gioffré, Luca Moroni, Alberte Fernández-Castro, Elena Marafatto, Giacomo Garufi and Roberto Navigli	782
<i>ProverbIT - Easy to complete, hard to choose: A CALAMITA Challenge</i>	
Enrico Mensa, Lorenzo Zane, Calogero Jerik Scozzaro, Matteo Delsanto, Tommaso Milani and Daniele P. Radicioni	806