

Preface to the CLiC-it 2025 Proceedings

Cristina Bosco¹, Elisabetta Jezek², Marco Polignano³ and Manuela Sanguinetti⁴

¹University of Torino

²University of Pavia

³University of Bari "Aldo Moro"

⁴University of Cagliari

The Italian Conference on Computational Linguistics (CLiC-it) is the yearly conference organized by the *Associazione Italiana di Linguistica Computazionale* (Italian Association of Computational Linguistics, AILC). Its main goal is to promote and spread original and high-quality research on the diverse aspects concerning the automatic processing of natural language, both spoken and written. The eleventh edition of CLiC-it took place in Cagliari, from the 24th to the 26th of September 2025.

In line with previous editions, submissions to the conference could be of two types: regular papers, featuring original and unpublished contributions, and non-archival research communications, consisting of papers accepted in 2024 and 2025 by major publication venues, namely the major international Computational Linguistics (CL) conferences (workshops excluded) or international journals. Regular paper submissions were assigned to thirteen Senior Program Chairs based on the general area that covered the paper's topic. Paper assignments to reviewers were managed individually by the single Senior PCs, though resorting to a global pool of 151 available reviewers. We have received 138 submissions of regular papers, hitting once more the record number of submissions even compared to the previous edition in 2024, where regular papers submitted were 133. This confirms the vitality and growth of the Italian Computational Linguistics community. Along with regular papers, we also received 22 research communications. Among the regular submissions, 113 were accepted for presentation at the conference, resulting in a 81.8% acceptance rate, with respect to the 85.7% rate of CLiC-it 2024. Out of these, 55 were accepted as oral presentations and 58 as posters. After the author notification was sent, 4 papers were withdrawn by the authors themselves. As a result, the conference featured a total of 55 oral presentations and 54 posters. Finally, of the 22 research communications submitted – a clear sign of the vitality and quality of the research carried out within the community – 14 were included for poster presentation in a dedicated session.

CLiC-it 2025: Eleventh Italian Conference on Computational Linguistics, September 24 – 26, 2025, Cagliari, Italy

✉ cristina.bosco@unito.it (C. Bosco); elisabetta.jezek@unipv.it (E. Jezek); marco.polignano@uniba.it (M. Polignano); manuela.sanguinetti@unica.it (M. Sanguinetti)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

The selection was not based on an additional review process, but rather on the venue of publication. Even in this case, two research communications were withdrawn after the notification, hence 12 posters were presented at the conference.

The program also included two keynote talks, by **Karen Fort** (University of Lorraine) and **Edoardo Maria Ponti** (University of Edinburgh/NVIDIA):

- Karen Fort gave a talk titled "Large Language Models: the challenge of evaluation": *In the past five years or so, Natural Language Processing has witnessed a revolution. Not only have Large Language Models (LLM) submerged the domain, but they also invaded our societies: our systems now have real users and an impact on their lives. This dramatic change happened so fast that we -the research community- are still trying to catch up, especially concerning the evaluation of the real capabilities of these tools. In this presentation, I'll show what are the flaws of the present LLMs' evaluation and how ethics is a powerful leverage to improve it.*
- the talk by Edoardo Maria Ponti was titled "A Blueprint for Foundation Models with Adaptive Tokenization and Memory": *Foundation models (FMs) process information as a sequence of internal representations; however, the length of this sequence is fixed and entirely determined by tokenization. This essentially decouples representation granularity from information content, which exacerbates the deployment costs of FMs and narrows their "horizons" over long sequences. What if, instead, we could free FMs from tokenizers by modelling bytes directly, while making them faster than current tokenizer-bound FMs? To achieve this goal, I will show how to: 1) learn tokenization end-to-end, by dynamically pooling representations in internal layers and progressively learning abstractions from raw data; 2) compress the KV cache (memory) of Transformers adaptively during generation without loss of performance; 3) predict multiple bytes per time step in an efficient yet expressive way; 4) retrofit existing tokenizer-bound FMs into byte-level FMs through cross-tokenizer distillation.*

By blending these ingredients, we may soon witness the emergence of new, more efficient architectures for foundation models.

One session of the conference was devoted to a discussion with **Paola Merlo**, conceived as a space for critical discussion on key areas of research in the fields of Computational Linguistics and Natural Language Processing, with a focus on both theoretical developments and applied scenarios within these disciplines. The discussion took place in an interview format. With the aim of promoting active and inclusive participation, a call for interest was launched, open to all interested participants - with a special focus on young researchers - to collect questions in advance to be addressed to Paola Merlo during the interview.

The conference was also preceded by a tutorial held by **Sandro Pezzelle** (University of Amsterdam) titled "*Language-and-vision models: From image-language alignment to storytelling and narration*". The tutorial aimed at providing an accessible yet in-depth overview of language-and-vision models, ranging from traditional modular pipelines to the latest end-to-end pre-trained systems (VLMs). The tutorial introduced foundational concepts and architectures, then focusing on recent approaches and evaluation challenges.

AILC also renewed its support to the **Emanuele Pianta Award** for the Best Master's Thesis defended at any Italian university between August 1st 2024 and July 31st 2025, and addressing a topic in computational linguistics or its applications. This year we received 9 candidate theses for the award. Of these, 2 were not further considered for evaluation due to incomplete submission. The candidate theses were evaluated by a jury composed of a current chair of CLiC-it, specifically Elisabetta Jezek was designated for this role, a co-chair of the past edition, i.e. Rachele Sprugnoli (University "Cattolica del Sacro Cuore" of Milan), and a further member of the AILC Board, i.e. Danilo Croce (University "Tor Vergata" of Rome). The winner was awarded by the members of the jury during the closing session of the conference.

We would like to thank all the **institutions** involved in the organization of the conference and the people of these institutions that worked with us for creating a successful event. For the logistic support, our thanks go to the Faculty of Economics, Law and Political Sciences of the University of Cagliari, that hosted the conference in the Sant'Ignazio Campus, and Valentina Deidda in particular, who kindly assisted us in all the technical and logistic aspects concerning the organization of the event. For the organizational and financial support, our thanks also go to the Department of Mathematics and Computer Science of the University of Cagliari, whose researchers were involved in the development and management of the website and whose students worked hard to help run

things smoothly during the conference.

Our gratitude also goes to our **corporate supporters** for their generous provision of the financial resources and services that made this event possible: Ap-tus.AI, Talia, Aequa-tech, Almawave, Domynt, Elra, Logogramma, Prometeia, and Translated.

We express our deepest gratitude also to all Senior Program Chairs, all members of the Program Committee, and all participants, who contributed to the success of the event. Chairs and reviewers are named in the following pages.

Our final and special thanks go to **AILC's** Board, whose members constantly supported us, giving guidance and invaluable assistance throughout the whole organization of the event, and to the whole AILC's association members that make this scientific event more interesting and richer every year.

Cagliari, September 2025

Conference Chairs

- **Cristina Bosco**, University of Torino
- **Elisabetta Jezek**, University of Pavia
- **Marco Polignano**, University of Bari "Aldo Moro"
- **Manuela Sanguinetti**, University of Cagliari

Local Committee

- **Maurizio Atzori**, University of Cagliari
- **Andrea Loddo**, University of Cagliari
- **Davide Antonio Mura**, University of Cagliari
- **Alessandro Pani**, University of Cagliari
- **Alessandra Perniciano**, University of Cagliari
- **Luca Zedda**, University of Cagliari

Proceeding Chairs

- **Francesca Grasso**, University of Torino
- **Andrea Zaninello**, Fondazione Bruno Kessler

Publicity and Data Chairs

- **Alessandro Bondielli**, University of Pisa
- **Mirko Lai**, University of Eastern Piedmont

Booklet

- **Alessandra Perniciano**, University of Cagliari

Webmasters

- **Maurizio Atzori**, University of Cagliari
- **Andrea Loddo**, University of Cagliari

Registration Management

- **Sara Barcena**, freelance designer
- **Manuela Speranza**, Fondazione Bruno Kessler

Senior Program Chairs

- **Elisa Bassignana**, IT University of Copenhagen
- **Pierluigi Cassotti**, University of Gothenburg
- **Simone Conia**, University of Rome “La Sapienza”
- **Elisa Di Nuovo**, Joint Research Centre European Commission, Ispra
- **Claudiu Daniel Hromei**, University of Rome “Tor Vergata”
- **Antonio Origlia**, University of Naples “Federico II”
- **Ludovica Pannitto**, University of Bologna
- **Beatrice Savoldi**, Fondazione Bruno Kessler
- **Gabriele Sarti**, University of Groningen
- **Lucia Siciliani**, University of Bari
- **Irene Siragusa**, University of Palermo
- **Rossella Varvara**, University of Torino/University of Pavia
- **Alessandro Vietti**, University of Bolzano

Program Committee

Oscar Araque, Serena Auriemma, Pier Balestrucci, Matilde Barbini, Alberto Barrón-Cedeño, Pierpaolo Basile, Valerio Basile, Monica Berti, Siddharth Bhargava, Arianna Bienati, Andrea Bolioli, Alessandro Bondielli, Giovanni Bonetta, Tommaso Bonomo, Federico Borazio, Federico Boschetti, Sofia Brenna, Luca Brigada Villa, Dominique Brunato, Davide Buscaldi, Elena Cabrio, Sara Candussio, Luca Capone, Franco Alberto Cardillo, Alessio Cascione, Silvia Casola, Giuseppe Castellucci, Camilla Casula, Flavio Massimiliano Cecchini, Giuseppe G.A. Celano, Mauro Cettolo, Emmanuele Chersoni, Cristiano Chesi, Francesca Chiusaroli, Alessandra Teresa Cignarella, Lorenzo Cima, Davide Colla, Gloria Comandini, Claudia Roberta Combei, Lina Conti, Serena Coschignano, Danilo Croce, Francesco Cutugno, Lorenzo De Matteo, Angelo Maria Del Grosso, Pietro Dell'Oglio, Maria Pia Di Buono, Andrea Di Fabio, Maria Di Maro, Luca Dini, Luca Ducceschi, Andrea Esuli, Alfio Ferrara, Elisabetta Fersini, Komal Florio, Simona

Frenda, Francesca Frontini, Achille Fusco, Eleonora Ghizzota, Aivars Glaznieks, Claudiu Hromei, Alina Karakanta, Fahad Khan, Tiziano Labruna, Mirko Lai, Alberto Lavelli, Gianluca Leboni, Els Lefever, Alessandro Lenci, Elisa Leonardelli, Soda Marem Lo, Agnese Lombardi, Marco Madeddu, Bernardo Magnini, Francesco Mambrini, Chiara Manna, Raffaele Manna, Marta Marchiori Manerba, Andrea Marra, Claudia Marzi, Arianna Masciolini, Alessandro Mazzei, Enrico Mensa, Alessio Miaschi, Simonetta Montemagni, Johanna Monti, Luca Moroni, Elio Musacchio, Benedetta Muscato, Vivi Nastase, Roberto Navigli, Nicole Novielli, Debora Nozza, Antonio Origlia, Teresa Paccosi, Francesca Padovani, Alessio Palmero Aprosio, Sara Papi, Lucia Passaro, Marco Passarotti, Viviana Patti, Matteo Pellegrini, Nicolò Penzo, Federico Pianzola, Vito Pirrelli, Roberto Pirrone, Massimo Poesio, Mattia Proietti, Valeria Quochi, Giulia Rambelli, Federico Ranaldi, Leonardo Ranaldi, Irene Russo, Daniel Russo, Andrea Santilli, Loredana Schettino, Giulia Speranza, Rachele Sprugnoli, Marco Antonio Stranisci, Carlo Strapparava, Alice Suozzi, Luigi Talamo, Fabio Tamburini, Benedetta Tessa, Davide Testa, Sara Tonelli, Giulia Venturi, Guido Vetere, Serena Villata, Vincenzo Norman Vitale, Roberto Zamparelli, Andrea Zaninello, Roberto Zanolli.

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT and Grammarly for grammar and spelling check. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.