



Associazione Italiana di  
Linguistica Computazionale



## Preface

The eighth edition of the Italian Conference on Computational Linguistics (CLiC-it 2021) was held at Università degli Studi di Milano-Bicocca from 29th June to 1st July 2022.

After the edition of 2020, which was held in fully virtual mode due to the health emergency related to Covid-19, CLiC-it 2021 represented the first moment for the Italian research community of Computational Linguistics to meet in person after more than one year of full/partial lockdown. Although the conference was held in dual mode, we strongly suggested the participants to attend it coming to Milan. Indeed, we received a strong feedback on this aspect from the community, which was eager to meet in person and enjoy both the scientific and social events together with the colleagues. In total, 99 participants registered to the conference benefiting from the early registration fee, 91 out of which expressed their intention to attend the event in person, which we consider as a very positive indication of enthusiasm from the community, given the uncertain situation due to the evolution of the pandemic in Italy.

In total, we received 68 proposals, organized in the following specific tracks: Information Extraction, Information Retrieval and Question Answering, Computational Linguistics and Natural Language Processing for the Humanities, Computational Social Science and Social Media Dialogue, Discourse and Natural Language Generation, Ethics and NLP, Language Resources and Evaluation, Spoken Language Processing and Automatic Speech Understanding, Cognitive Modeling and Psycholinguistics, Linguistic Issues in CL and NLP, Machine Learning for NLP, Machine Translation and Multilingualism, Morphology and Syntax Processing, Pragmatics and Creativity, Research and Industrial NLP Applications, Semantics, Knowledge Representation, Vision, Robotics, Multimodal and Grounding.

During the reviewing process, each submission was reviewed by three independent members of the scientific committees of the tracks in single-blind fashion. At the end of the process, 59 proposals were accepted for presentation at the conference and publication in the proceedings, resulting in an acceptance rate of 86.76%. Out of the 59 accepted proposals, 26 were included in the program of CLiC-it 2021 as oral presentations and the remaining 33 were assigned to one of the three poster sessions of the conference. As usual, the criterion for assigning a proposal to an oral or a poster session was based on the contents and not on the quality of the proposal. Regardless of the format of presentation, all accepted papers are allocated six pages of content (one additional page w.r.t. the previous editions) plus unlimited pages of references in the proceedings, available as open access publication. In line with last editions, the conference is receiving considerable attention from the international community, with 21 (31%) submissions this year showing at least one author affiliated to a foreign institution. This amounts to a total of 43 authors over 221 (20%) affiliated to 13 foreign countries: Belgium, Czechia, France, Germany, Romania, Hong Kong, Greece, Iran, The Netherlands, Spain, Switzerland, Turkey, United States.

The program of CLiC-it 2021 is completed by 10 research communications selected after the reviewing process. Research communications are not published in the proceedings, but have been orally presented within dedicated sessions at the conference, in order to enforce dissemination of excellence in research.

In addition to the technical programme, this year we were honoured to have as invited speakers such internationally recognized researchers as Barbara Plank (IT University of Copenhagen), with a keynote entitled “Returning the L in NLP: Why Language (Variety) Matters and How to Embrace it in Our

Models” and Dan Roth (University of Pennsylvania & Amazon AWS AI) with the keynote “It’s Time for Reasoning”.

Two tutorials were organized. On 29th June, Christian Chiarcos (Goethe University Frankfurt, Germany) provided an introduction to Linguistic Linked Open Data, while on 1st July Daniel Zeman (Charles University of Prague, Czech Republic) introduced the project of Universal Dependencies.

Moreover, the program included a panel discussion about linguistic infrastructures with representatives of CLARIN (Monica Monachini, ILC-CNR), ELG (Bernardo Magnini, FBK), and DARIAH (Emiliano Degli’Innocenti, OVI-CNR).

This year we have received 8 candidate theses for the “Emanuele Pianta Award for the Best Master Thesis”. This special prize for the best Master Thesis (Laurea Magistrale) in Computational Linguistics, submitted at an Italian University, is endorsed by AILC. The candidate theses have been evaluated by a special jury, and awarded during the last session of the conference, by the members of the jury.

We thank all the people and institutions involved in the organization of the conference, all track chairs, reviewers, and all participants, who contributed to the success of the event. All track chairs and reviewers are named in the following pages. We are grateful to the following, who made CLiC-it 2021 possible and supported us greatly and generously in the processes of local organization, publication of the proceedings and publicity: Marco Cremaschi, Giulia Rizzi, Silvia Terragni (Local Organizing Committee); Alessandra Teresa Cignarella, Matteo Pellegrini, Danilo Croce (Publication Chairs); Rachele Sprugnoli (Publicity Chair).

We thank the sponsors of CLiC-it 2021, who generously provided funds and services that are crucial for the realization of this event, including: B13, Dipartimento di Informatica, Sistemistica e Comunicazione (DISCO), Università degli Studi di Milano-Bicocca, Elsevier, Red Software Systems, Yewno (Gold sponsors); AIO Proactive Systems, Almawave, CELI, expert.ai (Silver sponsors); datrrix, ELRA, oaxs (Bronze sponsors).

Finally, we want to thank very much the Associazione Italiana di Linguistica Computazionale (AILC), all the members of the Association Board and, in particular, the President Bernardo Magnini, who never let us alone in the sea of doubts and problems that organizing an hybrid event implies.

June 2022

Elisabetta Fersini, Marco Passarotti, Viviana Patti  
CLiC-it 2021 General Chairs

## **Conference and Programme Chairs**

Elisabetta Fersini, Università degli Studi di Milano-Bicocca, Italy  
Marco Passarotti, Università Cattolica del Sacro Cuore, Milano, Italy  
Viviana Patti, Università degli Studi di Torino, Italy

## **Track Chairs**

### **Information Extraction, Information Retrieval and Question Answering**

Pierpaolo Basile, Università degli Studi di Bari  
Elena Cabrio, Université Côte d'Azur

### **CL and NLP for the Humanities**

Francesca Frontini, Istituto di Linguistica Computazionale "A.Zampolli"  
Francesco Mambrini, Università Cattolica del Sacro Cuore, Milano

### **Computational Social Science and Social Media**

Valerio Basile, Università degli Studi di Torino  
Nicole Novielli, Università degli Studi di Bari

### **Dialogue, Discourse and Natural Language Generation**

Alessandro Mazzei, Università degli Studi di Torino  
Antonio Origlia, Università degli Studi di Napoli Federico II

### **Ethics and NLP**

Andrea Bolioli, CELI Language Technology  
Malvina Nissim, University of Groningen

### **Language Resources and Evaluation**

Simonetta Montemagni, Istituto di Linguistica Computazionale "A.Zampolli"  
Rachele Sprugnoli, Università Cattolica del Sacro Cuore, Milano

### **Spoken Language Processing and Automatic Speech Understanding**

Franco Cutugno, Università degli Studi di Napoli Federico II  
Daniele Falavigna, Fondazione Bruno Kessler

### **Cognitive Modeling and Psycholinguistics**

Cristiano Chesi, Scuola Superiore Universitaria Pavia – IUSS  
Tommaso Fornaciari, Università Bocconi

### **Linguistic Issues in CL and NLP**

Maria Teresa Guasti, Università degli Studi di Milano-Bicocca  
Elisabetta Jezek, Università di Pavia

### **Machine Learning for NLP**

Roberto Basili, Università di Roma, Tor Vergata  
Felice Dell'Orletta, Istituto di Linguistica Computazionale "A.Zampolli"

### **Machine Translation and Multilingualism**

Matteo Negri, Fondazione Bruno Kessler  
Manuela Sanguinetti, Università degli Studi di Cagliari

### **Morphology and Syntax Processing**

Cristina Bosco, Università degli Studi di Torino

Fabio Tamburini, Università di Bologna

### **Pragmatics and Creativity**

Valentina Bambini, Scuola Superiore Universitaria Pavia – IUSS

Carlo Strapparava, Fondazione Bruno Kessler

### **Research and Industrial NLP Applications**

Francesca Bonin, IBM Research Ireland

Marcello Federico, Amazon

### **Semantics, Knowledge Representation**

Alessandro Lenci, Università di Pisa

Fahad Khan, Istituto di Linguistica Computazionale “A.Zampolli”

### **Vision, Robotics, Multimodal and Grounding**

Raffaella Bernardi, Università di Trento

Danilo Croce, Università di Roma, Tor Vergata

### **Local Organizing Committee**

Marco Cremaschi, Università degli Studi di Milano-Bicocca

Giulia Rizzi, Università degli Studi di Milano-Bicocca

Silvia Terragni, Università degli Studi di Milano-Bicocca

### **Publication Chairs**

Alessandra Teresa Cignarella, Università degli Studi di Torino

Matteo Pellegrini, Università Cattolica del Sacro Cuore, Milano

Danilo Croce, Università di Roma, Tor Vergata

### **Publicity Chair**

Rachele Sprugnoli, Università Cattolica del Sacro Cuore, Milano

### **Reviewers**

Luca Anselma, Alessio Palmero Aprosio, Mattia Atzeni, Stefano Bannò, Pierpaolo Basile, Valerio Basile, Luisa Bentivogli, Monica Berti, Arianna Bisazza, Luca Bischetti, Yuri Bizzoni, Paolo Boldi, Andrea Bolioli, Marianna Bolognesi, Cristina Bosco, Davide Buscaldi, Elena Cabrio, Tommaso Caselli, Pierluigi Cassotti, Giuseppe Castellucci, Flavio Massimiliano Cecchini, Cristiano Chesi, Francesca Chiusaroli, Alessandra Teresa Cignarella, Andrea Cimino, Giovanni Colavizza, Davide Colla, Danilo Croce, Francesco Cutugno, Rossana Damiano, Lorenzo De Mattei, Thierry Declerck, Felice Dell’Orletta, Rodolfo Delmonte, Danilo Dessì, Luigi Di Caro, Mattia Antonino Di Gangi, Maria Di Maro, Giorgio Maria Di Nunzio, Claudio Fantinuoli, Simone Filice, Komal Florio, Greta Franzini, Diego Frassinelli, Simona Frenda, Francesca Frontini, Gloria Gagliardi, Marco Gaido, Emiliano Giovannetti, Martin Gleize, Claudio Greco, Shohreh Haddadan, Christian Hardmeier, Sadid A. Hasan, Aurelie Herbelot, Delia Irazu Hernandez Farias, Ignacio Iacobacci, Andrea Iovine, Elisabetta Jezek,

Charles Jochim, Roman Klinger, Alberto Lavelli, Matteo Magnani, Mirko Marras, Barbara McGillivray, Stefano Menini, Alessio Miaschi, Monica Monachini, Massimo Moneglia, Johanna Monti, Alessandro Moschitti, Mohamed Nabih, Debora Nozza, Renáta Németh, Antonio Origlia, Petya Osenova, Alessio Palmero Aprosio, Alessandro Panunzi, Sara Papi, Patrick Paroubek, Matteo Pellegrini, Michael Piotrowski, Marco Polignano, Valeria Quochi, Alessandro Raganato, Andrea Amelio Ravelli, Matteo Romanello, Marco Rospocher, Gaetano Rossiello, Irene Russo, Manuela Sanguinetti, Gabriele Sarti, Beatrice Savoldi, Flavia Sciolette, Lucia Siciliani, Gianmaria Silvello, Maria Simi, Manuela Speranza, Francesca Spezzano, Rachele Sprugnoli, Armando Stellato, Fabio Tamburini, Mariona Taulé, Sara Tonelli, Rocco Tripodi, Marieke van Erp, Rossella Varvara, Giulia Venturi, Serena Villata.

