

COLING

**International Conference on
Computational Linguistics**

Proceedings of the Conference and Workshops

COLING

Volume 29 (2022), No. 21

**Proceedings of the CRAC 2022 Shared Task on
Multilingual Coreference Resolution**

**The 29th International Conference on
Computational Linguistics**

October 17, 2022
Gyeongju, Republic of Korea

Copyright of each paper stays with the respective authors (or their employers).

ISSN 2951-2093

Message from the Program Chairs

This volume contains papers describing the CRAC 2022 Shared Task on Multilingual Coreference Resolution and the participating systems. The public edition of the multilingual collection CorefUD 1.0 was used as the source of training and evaluation data, spanning 13 datasets for 10 languages, namely Catalan, Czech, English, French, German, Hungarian, Lithuanian, Polish, Russian, and Spanish. Shared task participants were supposed to identify mentions in texts and to predict coreference relations between the identified mentions; only identity coreference is considered in this shared task.

A total of 5 teams participated in this task, with 8 submitted systems ranging from a very simple rule-based system to a deep-learning system trained jointly on all languages. In this volume, system description papers delivered by 3 teams are presented, preceded with an overview paper describing in more detail the task itself, the input data, the baseline system, the main evaluation metric, and global performance comparisons.

From our viewpoint, major goals of the shared task were reached: not only that new systems capable of coreference resolution in various languages were created, but the state-of-the-art performance for the given multilingual collection was improved considerably. We hope that this success will attract new researchers to the area of multilingual coreference resolution, and hopefully also new participants to the future editions of the present shared task.

We would like to thank all the participants for their efforts, and program committee members for reviewing the submitted manuscripts. Last but not least, we would like to thank organizers of the previous shared tasks for sharing their experience with us, and in general all authors of the involved coreference datasets for making the results of their work publicly accessible.

September 2022
Maciej Ogrodniczuk, Zdeněk Žabokrtský
on behalf of the shared task organizers

Shared task specification

<https://ufal.mff.cuni.cz/corefud/crac22>

Shared task organizers

- Charles University (Prague, Czechia):
 - Anna Nedoluzhko
 - Michal Novák
 - Martin Popel
 - Zdeněk Žabokrtský
 - Daniel Zeman
- Institute of Computer Science, Polish Academy of Sciences (Warsaw, Poland):
 - Maciej Ogrodniczuk
- Georgetown University (Washington D.C., USA):
 - Yilun Zhu
- University of West Bohemia (Pilsen, Czechia):
 - Miloslav Konopík
 - Ondřej Pražák
 - Jakub Sido

Program Committee

- Haixia Chai
- Christian Hardmeier
- Veronique Hoste
- Loic De Langhe
- Ekaterina Lapshinova-Koltunski
- Yaqin Yang
- Juntao Yu
- Yilun Zhu

Table of Contents

<i>Findings of the Shared Task on Multilingual Coreference Resolution</i> Zdeněk Žabokrtský, Miloslav Konopík, Anna Nedoluzhko, Michal Novák, Maciej Ogrodniczuk, Martin Popel, Ondřej Pražák, Jakub Sido, Daniel Zeman and Yilun Zhu	1
<i>Coreference Resolution for Polish: Improvements within the CRAC 2022 Shared Task</i> Karol Saputa	18
<i>End-to-end Multilingual Coreference Resolution with Mention Head Prediction</i> Ondřej Pražák and Miloslav Konopik	23
<i>ÚFAL CorPipe at CRAC 2022: Effectivity of Multilingual Models for Coreference Resolution</i> Milan Straka and Jana Straková	28

CRAC Shared Task session program

Monday, October 17, 2022

09:00–10:30 Shared Task papers

09:00–09:30 *Findings of the Shared Task on Multilingual Coreference Resolution*
Zdeněk Žabokrtský, Miloslav Konopík, Anna Nedoluzhko, Michal Novák, Maciej Ogrodniczuk, Martin Popel, Ondřej Pražák, Jakub Sido, Daniel Zeman and Yilun Zhu

09:30–09:45 *Coreference Resolution for Polish: Improvements within the CRAC 2022 Shared Task*
Karol Saputa

09:45–10:00 *End-to-end Multilingual Coreference Resolution with Mention Head Prediction*
Ondřej Pražák and Miloslav Konopík

10:00–10:15 *ÚFAL CorPipe at CRAC 2022: Effectivity of Multilingual Models for Coreference Resolution*
Milan Straka and Jana Straková

10:15–10:30 *Discussion*

10:30–11:00 Coffee Break

11:00–12:00 Invited Talk

11:00–12:00 *The recent developments in Universal Anaphora Scorer*
Juntao Yu

