



## **CLASSLA-Express 1.0 Plus: corpora meets AI**

### workshops on using CLARIN.SI corpora and AI tools in language research

This workshop introduces participants to the use of the CLASSLA web corpora in language research. Through hands-on exercises, participants will learn how to create queries in corpora for Bosnian, Croatian, Macedonian, Serbian, and Slovene, including basic searches and advanced CQL queries.

Building on the CLASSLA-Express 1.0 format, this **1.0 Plus edition** adds a special segment that introduces AI-based approaches and invites discussion on how large language models can be integrated into linguistic research. The workshop is thus designed not only to provide participants with practical corpus skills, but also to open space for critical reflection on the complementarity of established corpus methods and AI tools.

The workshop is aimed at university students, linguists, language teachers, lexicographers, corpus linguists, and digital humanities scholars. **Participants are kindly requested to bring their own laptops and ensure that they are fully charged.**

Participation is free of charge. The workshop is funded by CLARIN.SI – Slovene Common Language Resources and Technology Infrastructure and the Department of Slavic Studies at the University of Graz, and co-organized with the Regional Branch Zagreb of the Croatian Association of Applied Linguistics.

#### **Organizers:**

Ivana Filipović Petrović (Croatian Academy of Sciences and Arts, Croatia)

Jelena Parizoska (University of Zagreb, Croatia)

Taja Kuzman Pungeršek (Institute Jožef Stefan, Slovenia)

Nikola Ljubešić (Institute Jožef Stefan, University of Ljubljana, Slovenia)

Petya Osenova (AI and Language Technologies Division IICT-BAS Sofia, Bulgaria)

**Local host:** Marko Simonović, Department of Slavic Studies, University of Graz

**Date:** 10 October 2025

**Venue:** Meerzwecksaal (0033EG0010), Merangasse 70 (Ground floor), Graz, AT.

**Convenors:** Marko Simonović and Ivana Filipović Petrović

**Language of instruction:** English and Bosnian/Croatian/Montenegrin/Serbian

#### **Registration form:**

<https://docs.google.com/forms/d/e/1FAIpQLSe7kfjqmWq20ASjDkSwoZvdC3vUnkf7c6z-8hNiqxb-2PncQ/viewform?usp=preview>

**Contact:** Ivana Filipović Petrović (ifilipovic@hazu.hr)

## Schedule

11:00 – 11:30 Building and using CLASSLA corpora for linguistic research and a glimpse into AI tools (Ivana Filipović Petrović)

11:30 – 12:15 Exercises in CLASSLA corpora (Ivana Filipović Petrović)

12:15 – 12:45 **Coffee break**

12:45 – 13:30 The freshest corpus studies from the Slavic Studies in Graz (Jelena Stojković)

13:30 – 14:00 Prospects for new corpora: Burgenland Croatian (Marko Simonović)

14:00 – 15:00 **Free time / Lunch break**

15:00 – 15:30 Prospects for new corpora: Pannonian Rusyn (Marko Simonović)

15:30 – 16:30 The best of both worlds: corpora and AI in language research (Ivana Filipović Petrović)

16:30 – 17:00 **Coffee break**

17:00 – 18:00 AI-driven interfaces: Extracting and interpreting linguistic data (Ivana Filipović Petrović)

Discussion and concluding remarks