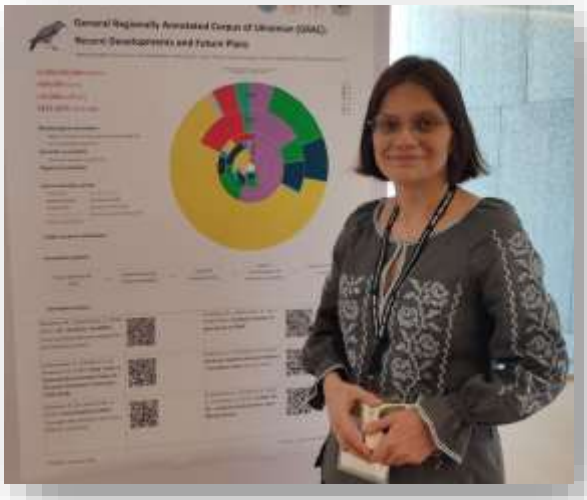
ЛЕКТОРІЙ: НАУКА З ПЕРШИХ УСТ

VORTRAGSREIHE: WISSENSCHAFT AUS ERSTER HAND

LECTURE SERIES: SCIENCE FIRST HAND



Dr. Maria Shvedova
photo: private

TOPIC: **An Ecosystem of Open Ukrainian Language Resources: Corpora, Annotation, and FAIR Compliance**

SPEAKER: **Dr. Maria Shvedova**, National Technical University "Kharkiv Polytechnic Institute" (Kharkiv, Ukraine)/ Potsdam University (Germany)/ Jena University (Germany)

MODERATOR: Dr. Olha Kanishcheva, Heidelberg University (Germany)/ SET University (Kyiv, Ukraine)

DATE: 24 November 2026, Tuesday

TIME: 18:00–19:30 (Berlin Time)

This lecture will be held in English.

Registration is open until November 23, 2026 at:
berlin@ukrainet.eu

The event will be held on the Zoom platform. During the discussion, you can also use the chat option, or formulate your questions in advance and send them to the following email address: berlin@ukrainet.eu. Please note that the presentation will be video-recorded, so we recommend keeping your microphones and cameras off during the lecture.

Dear ladies and gentlemen! Dear colleagues! Dear friends!

The German-Ukrainian Academic Society and the Embassy of Ukraine in the Federal Republic of Germany invite you to attend the lecture series "Science First Hand" on "**An Ecosystem of Open Ukrainian Language Resources: Corpora, Annotation, and FAIR Compliance**" by **Dr. Maria Shvedova**, National Technical University "Kharkiv Polytechnic Institute" (Kharkiv, Ukraine), Potsdam University (Germany), Jena University (Germany).

Summary: In recent years, Ukrainian language has gained a growing family of open language resources, many of them built on the General Regionally Annotated Corpus of Ukrainian (GRAC). This lecture offers an overview of this emerging ecosystem and the work behind it. It covers syntactically annotated data, including the Universal Dependencies treebanks; PluG, a copyright-free corpus of historical texts from 1816-1954; PAWUK, the Polish Automatic Web Corpus of Ukrainian Language (IPI PAN); ParlMix-UA-RU, a manually annotated Ukrainian-Russian code-switching corpus of parliamentary speech; and the first PARSEME corpus of Ukrainian multiword expressions. For each resource, the talk explains how it was

created, what annotation challenges it posed, and which research and NLP tasks it supports. The lecture also addresses practical questions of sharing and reuse that determine how easily these resources can be found, integrated into multilingual projects, and used by researchers and students. It draws on the experience of the CLARIN Knowledge Centre for Ukrainian NLP and Corpora (UkrNLP-Corpora), which supports resource creators in aligning their work with FAIR principles.

SPEAKER: Dr. Maria Shvedova is an Associate Professor at the Department of Intelligent Computer Systems at the National Technical University "Kharkiv Polytechnic Institute" in Ukraine. She is currently a DAAD research fellow at the University of Potsdam, Germany, where she is working on a project titled "Case Choice in Ukrainian Vocative Expressions from a Typological Perspective", funded by the German Academic Exchange Service (DAAD) within the programme Future Ukraine: Research Grants for Ukrainian Master's Students and Researchers. She is also a visiting researcher at the Friedrich Schiller University Jena.

Dr. Shvedova holds a Ph.D. in Philology and is the head of the GRAC corpus (General Regionally Annotated Corpus of Ukrainian), one of the largest open corpora of the Ukrainian language. Together with colleagues from Ukrainian and European institutions, she is involved in the development of a wide range of linguistic resources for Ukrainian, including PAWUK (Polish Automatic Web Corpus of Ukrainian Language, IPI PAN), UD_Ukrainian_ParlaMint, ParaRook (parallel corpora based on GRAC), PluG (a corpus of 19th- and early 20th-century Ukrainian texts), and the Ukrainian component of PressMint (Interoperable Corpora of Historical Newspapers). Her research interests include corpus linguistics, language resource development, linguistic annotation, and Ukrainian grammar.

https://www.uni-potsdam.de/de/slavische-linguistik/team/dr-maria-shvedova

The organizers of the event:

- German-Ukrainian Academic Society
- Embassy of Ukraine in the Federal Republic of Germany

Information about the event:
Topic: "An Ecosystem of Open Ukrainian Language Resources: Corpora, Annotation, and FAIR Compliance"

When: November 24, 2026, 18:00–19:30 (Berlin Time)
Registration is open until November 23, 2026: berlin@ukrainet.eu