

(see English version below)

Call for Presentations

Workshop Digitale Slawistik

im Rahmen der [48. Österreichischen Linguistik-Tagung](#)
vom 17. bis 19. Dezember 2024 an der Universität Innsbruck, Österreich

Mit dem Eintritt in die Big-Data-Ära des 21. Jahrhunderts haben sich die Forschungsmöglichkeiten und -methoden innerhalb der Geistes- und Sozialwissenschaften signifikant verändert. Das Training von Large Language Models (LLM) und die Entwicklung von Transformatoren wie BERT oder der GPT-Familie beeinflussen alle linguistischen Bereiche, insbesondere die Verarbeitung natürlicher Sprache (NLP), und die slawistische Linguistik ist hier keine Ausnahme. Das Ziel dieses Workshops ist es, die Auswirkungen von LLMs auf die Fragestellungen und Arbeitsweisen innerhalb der slawistischen Forschung auszuloten.

Dabei sind Beiträge zur Anwendung der neuen digitalen Methoden in der Slawistik (besonders der Transformer-Netzwerke und LLM) ebenso willkommen wie traditionelle Methoden (etwa die regelbasierte Lemmatisierung oder Wortart-Tagging), interdisziplinäre Arbeiten oder die kritische Auseinandersetzung mit den Möglichkeiten und Grenzen digitaler Methoden. Mögliche Vortrags- und Diskussionsthemen zur Anwendung digitaler Methoden sind:

- * Auf- und Ausbau von Korpora und Datenbanken für historische und moderne Sprachvarietäten
- * Diastatische, diatopische, diachrone und arealinguistische Dokumentations- und Darstellungsformen
- * Entwicklung von Toolkits für die Digitalisierung und Annotation slawischen Sprachmaterials: POS- und morphologisches Tagging, Lemmatisierung, Abhängigkeitsanalyse
- * Korpusbasierte und korpusgestützte Analyse
- * Maschinelle Übersetzungen für slawische Sprachen
- * Stylometrie, Klassifikation und Text-Networks-Analysen
- * Verarbeitung historischer Dokumente: HTR und bildgebende Verfahren (Pre- und Postprocessing, Modelltraining)
- * Textkritik
- * Benchmarks für slawische NLP
- * Regelbasierte Ansätze und KI
- * Grenzen und Desiderata

Präsentationssprache ist Deutsch oder Englisch.

Einreichungsfrist der Abstracts:

01.07.2024 (UTC+2, CET)

Wir bitten um Zusendung der Abstracts (nicht mehr als 1 A4-Seite, Times New Roman 12, einzeilig, exklusive Beispiele und Literaturangaben) an digitalislawistik@gmail.com.

Digital Slavic Studies Panel

At the [48. Österreichische Linguistik-Tagung](#)

17th to 19th December 2024

University of Innsbruck, Austria

Since the beginning of the 21st century, Big Data has reshaped research possibilities and methods in humanities and Social Sciences. The emergence of transformer neural networks, such as BERT, and the subsequent rise of Large Language Models (LLM) transformed all linguistic branches, especially Natural Language Processing (NLP). Slavic studies are not an exception. This workshop aims to explore the impact of LLMs on the problems and working methods within Slavic Studies.

This panel seeks contributions on novel digital methods (especially transformer networks and LLM) as well as on traditional approaches (such as rule-based lemmatization or part-of-speech tagging) within Slavic Studies, on the interdisciplinary exploration or on critical examination of the possibilities and limitations of digital methods. Among the possible topics are:

- * Corpus and database building for historical and modern language varieties
- * Digital documentation and representation of areal phenomena, as well as social, historical, and regional variation
- * Development of toolkits for digitisation and tagging of Slavic languages: POS- and morphological tagging, lemmatisation, dependency parsing
- * Corpus-based analysis
- * Machine translation for Slavic languages
- * Stylometry, classification and text network analysis
- * OCR and processing of historical documents: pre- and post-processing, model training
- * Textual criticism
- * Benchmarks for Slavic NLP
- * Rule-based approaches and AI
- * Limits and desiderata

Presentations should be in German or English.

Deadline for abstract submission:

01.07.2024 (UTC+2, CET)

Please send your abstracts (no more than 1 A4 page, Times New Roman 12, single-spaced, excluding examples and references) to digitalslawistik@gmail.com.