# C4DHI Anniversary workshop

TWO-DAY WORKSHOP ON THE OCCASION OF THE THIRD
ANNIVERSARY OF THE CENTRE FOR DIGITAL HUMANITIES INTEGRATION

# EVENT PROGRAM

**1 - 2 JUNE 2026**

JINONICE CAMPUS, ROOM C420

IMS FSV UK

C4DHI CENTRE FOR DIGITAL HUMANITIES INTEGRATION

4eu+ EUROPEAN UNIVERSITY ALLIANCE

ÚFAL

matfyz

LINDAT CLARIAH-CZ

# C4DHI Anniversary workshop

## 1 - 2 JUNE 2026

JINONICE CAMPUS

### *Centre for Digital Humanities Integration*

Center for Digital Humanities Integration (C4DHI) is a research centre working within the Institute of International Studies at the Faculty of Social Sciences, Charles University. C4DHI pursues the mission of ushering current realities of indispensability of the "digital" dimension of research and education in area studies and modern history at IMS. Through close collaboration with partner research institutions specialized in the technological aspects of DH applications, it seeds and nurtures transdisciplinary methodological integration in four key pillar activities:

1. internal DH skill and capacity building
2. external DH expertise influx and networking
3. direct DH application in research and publishing
4. integration of DH into the study curriculum.

IMS FSV UK

C4DHI CENTRE FOR DIGITAL HUMANITIES INTEGRATION

4eu+ EUROPEAN UNIVERSITY ALLIANCE

ÚFAL

matfyz

LINDAT CLARIAH-CZ

# C4DHI Anniversary workshop

## 1 - 2 JUNE 2026

JINONICE CAMPUS

**Barbora Hladká & Kateřina Motalík Hodková**
*(Charles University, Faculty of Mathematics and Physics)*
- From Legalese to Plain Czech -- Linguistic Features and LLMs in Legal AI

**Klára Kosová**
*(Charles University, Faculty of Social Sciences)*
- Metaphor Field Loop Theory and the Automation of Metaphoricity Detection

**Jan Švec**
*(University of West Bohemia)*
- LLM-Based Metadata Extraction from NATO Scanned Documents

**Jiří Kocián**
*(Charles University, Faculty of Social Sciences/Faculty of Mathematics and Physics)*
- Exploring parliamentary discourses through LLMs: cultural embeddedness of illiberalism and radicalisation

IMS FSV UK

C4DHI CENTRE FOR DIGITAL HUMANITIES INTEGRATION

4eu+ EUROPEAN UNIVERSITY ALLIANCE

ÚFAL

matfyz

LINDAT CLARIAH-CZ

C4DH

# Anniversary workshop

1 JUNE 2026
Room C420

## Words of welcome

9:30 - 10:00

Jiří Kocián - Head of the Centre for Digital Humanities Integration
Barbora Hladká - Head of the Institute of Formal and Applied Linguistics

10:00 - COFFEE BREAK

## Student presentations

10:30- 11:00

Semestral work presentations by the students of Faculty of Arts

## Session 1

11:00- 12:00

Barbora Hladká & Kateřina Motalík Hodková
*(Charles University, Faculty of Mathematics and Physics)*

- **From Legalese to Plain Czech -- Linguistic Features and LLMs in Legal AI**

12:00 - LUNCH BREAK

## Session 2

13:00 - 14:30

Klára Kosová
*(Charles University, Faculty of Social Sciences)*

- **Metaphor Field Loop Theory and the Automation of Metaphoricity Detection**

C4DH

# Anniversary workshop

2 JUNE 2026
Room C420

**9:00 - WELCOME COFFEE AND SNACK**

## Session 3

**9:30 - 11:00**

Jan Švec *(University of West Bohemia)*

- **LLM-Based Metadata Extraction from NATO Scanned Documents**

**11:00- LUNCH BREAK**

## Session 4

**12:00 - 13:30**

Jiří Kocián *(Charles University, Faculty of Social Sciences)*

- **Exploring parliamentary discourses through LLMs: cultural embeddedness of illiberalism and radicalisation**

**13:30 - CLOSING REMARKS**

# C4DHI Anniversary workshop

## SPEAKERS

**1 - 2 JUNE 2026**

Institute of International Studies
Faculty of Social Sciences - Jinonice Campus
U Kříže 8, Prague 5 - Jinonice
Room C420

IMS FSV UK
C4DHI CENTRE FOR DIGITAL HUMANITIES INTEGRATION
4eu+ EUROPEAN UNIVERSITY ALLIANCE
ÚFAL
matfyz
LINDAT CLARIAH-CZ

C4DH

# Anniversary workshop

SPEAKERS

## Barbora Hladká

Barbora Hladká is an associate professor in the Institute of Formal and Applied Linguistics, the head of which she has been since February 2021, at the Computer Science School of the Faculty of Mathematics and Physics, Charles University. Her research focuses on issues from natural language processing, mainly corpus annotation, information extraction, and application of machine learning in various domains especially in the digital humanities. In 1998, she was awarded the Post Workshop'98 Research Project Award by the Center for Language and Speech Processing, Johns Hopkins University, Baltimore, MD, USA where she was working as a postdoc in 2000-2001. She has been teaching both national and international undergraduate courses on machine learning and supervising undergraduate and graduate students.

## Kateřina Motalík Hodková

Kateřina Motalík Hodková completed her doctoral studies in Romance Languages at the Faculty of Arts of the University of South Bohemia in České Budějovice. He currently works as a researcher at the Institute of Formal and Applied Linguistics (Faculty of Mathematics and Physics, Charles University). She focuses on semantics, terminology, professional and legal language, and the specifics of their translation.

## Klára Kosová

Klára Kosová is interested in combining the social sciences and humanities with computational science to address complex societal challenges. Her primary research interests lie in cognitive linguistics and political discourse analysis. She also focuses on questions of legitimacy crises in contemporary democracies, with particular attention to post-socialist countries in Central-Eastern and Southeastern Europe.

C4DH

# Anniversary workshop

SPEAKERS

## Jan Švec

Jan Švec is a researcher and AI developer at the University of West Bohemia in Pilsen. His work focuses on innovative solutions in speech recognition, machine learning, and the processing and understanding of complex textual and archival materials. He collaborates on projects that connect academic research with practical applications, including digital humanities and the large-scale analysis of historical documents. His current work also explores how large language models can be used to extract structured metadata from a broad range of documents. His goal is to push the boundaries of current methods and find new ways to apply them in meaningful real-world contexts.

## Jiří Kocián

In my research, while focusing on the area of Central and Southeast Europe, I deal with transitions to democracy, issues of populism and democratic backsliding, transnational history, (non-)national identities, the history of Jews in and the Holocaust with special attention to oral history, and the implementation of DH methods in SSH. With colleagues from CUNI, we have recently designed and launched an interdisciplinary Certified study program in Digital Humanities.

# C4DHI Anniversary workshop

# ABSTRACTS

## 1 - 2 JUNE 2026

Institute of International Studies
Faculty of Social Sciences - Jinonice Campus
U Kříže 8, Prague 5 - Jinonice
Room C420

IMS FSV UK
C4DHI CENTRE FOR DIGITAL HUMANITIES INTEGRATION
4eu+ EUROPEAN UNIVERSITY ALLIANCE
ÚFAL
matfyz
LINDAT CLARIAH-CZ

C4DH

# Anniversary workshop

ABSTRACTS

## *Barbora Hladká & Kateřina Motalík Hodková:* *From Legalese to Plain Czech -- Linguistic Features and LLMs in Legal AI*

Legal documents are frequently challenging for non-experts to comprehend. This presentation introduces an automated approach to assessing the readability of Czech legal texts by combining rule-based linguistic analysis with large language models (LLMs). Our aim is to explore how artificial intelligence can contribute to clearer legal communication and make legal information more accessible to the public.

## *Klára Kosová: Metaphor Field Loop Theory and the Automation of Metaphoricity Detection*

The presentation examines the role of conceptual metaphors as performative tools in political discourse. Drawing on Conceptual Metaphor Theory (Lakoff & Johnson), Critical Metaphor Theory (Musolff), and the metaphor field loop framework (Valdivia & Valdivia), it argues that metaphors do not merely describe political reality but actively shape how political world is constructed, contested in public discourse. The presentation will then use this theoretical model to interrogate a key methodological challenge: can metaphoricity be detected automatically? While supervised machine learning approaches trained on annotated corpora show promise for identifying lexical metaphors, they tend to underperform on contextually and culturally embedded instances — precisely those most relevant to political discourse analysis. Unsupervised methods offer scalability but lack sensitivity to metaphor's discursive function within broader narratives. The presentation argues that the gap between computational detectability and theoretical salience remains a critical open problem for corpus-based metaphor research.

## *Jan Švec: LLM-Based Metadata Extraction from NATO Scanned Documents*

This presentation will focus on how scanned NATO archival documents can be transformed into structured, research-ready data. I will show how large language models can support the extraction of metadata such as document titles, dates, institutions, archive codes, correspondence information, and thematic tags. The talk will also address the practical challenges of working with noisy historical documents and multilingual archival material. In the final part, I will briefly discuss how this kind of software can be designed and developed efficiently with Speckit.

## *Jiří Kocián: Exploring parliamentary discourses through LLMs: cultural embeddedness of illiberalism and radicalisation*

Since LLM usage in SSH got normalised during the recent years, social scientists and humanities scholars became proficient in analysing data on a scale previously difficult to imagine to the majority of them. After the initial shock and period of experimentation, more advanced methods, such as RAG and multi-model processing became common knowledge, enabling production of vast amounts of empirical data for a mixed-method research. In my presentation, I would like to discuss approaches and efficient usage pathways of large language models for more complex interpretive tasks in a theory-oriented comparative cross-cultural research, discussing three examples from my own research practice.