


2nd Budapest Methods
Workshop on

Large Language Models and Generative AI: Social Science Applications and Legal Aspects



**HUN-REN Centre for
Social Sciences, Budapest**

November 21-22,
2024

 4 Tóth Kálmán Street, H-1097, Budapest, Hungary



November 21, 2024 – Thursday

Schedule

10:00-12:45

Registration (Lobby)

Start your day by registering for the workshop and collecting your materials.

10:00-12:45

PhD Tutorials (T.2.37-38.)

BSc, MSc and PhD students can register for a consultation with the poltextLAB research team.

12:45-13:00

Welcome by Miklós Sebők, K011-K012)

Our host Miklós Sebők will kick off the workshop with some opening remarks.

13:00-14:30

Panel A: LLM Applications to Central and Eastern European Politics (K011-K012)

The detailed programme is found on page 4.

14:30-15:00

Coffee break (1st floor)

Enjoy some sandwiches and coffee on the first floor. We make group photos in this break.

15:00-16:30

Panel B and C (parallel sessions)

Panel B: Emotions and bias in AI (K011-K012)

Panel C: Open Science and LLMs (T.2.37-38.)

16:45-17:45

Keynote speech by Michal Parížek

Tracing the Narratives of International Conflicts Using Large Language Models

19:00-21:00

Match-making session @Vígvarjú restaurant - INVITE ONLY!

2 Vigadó tér, H-1051, Budapest, Hungary

Do you need help?

Find Laura Seben!
laura.seben@tk.hu



4 Tóth Kálmán Street, H-1097, Budapest, Hungary

November 22, 2024 – Friday

Schedule

9:30–11:00

Panel D and E (parallel sessions)

Panel D: Regulatory and Ethical Implications of Generative AI (T.2.37–38.)

Panel E: GenAI and LLM applications (K011–K012)

11:00–11:15

Coffee Break (1st floor)

Enjoy some sandwiches and coffee on the first floor.

11:15–12:45

Roundtable: Open science in the age of LLMs (K011–K012)

Moderator:

Levente Littvay (HUN-REN Centre for Social Sciences)

Panelists:

Clément Bénése (OpSci.ai)

Hubert Plisiecki (Polish Academy of Sciences)

Miklós Sebők (HUN-REN Centre for Social Sciences)

12:45–13:00

Closing remarks and next steps (K011–K012)

Miklós Sebők (HUN-REN Centre for Social Sciences)

13:00–14:00

Lunch @ HTK Restaurant

4 Tóth Kálmán Street, H-1097, Budapest, Hungary

Ground Floor

Do you need help?

Find Laura Seben!
laura.seben@tk.hu



4 Tóth Kálmán Street, H-1097, Budapest, Hungary

PANELS ON THURSDAY

13:00-14:30 Panel A: LLM Applications to Central and Eastern European Politics (K011-K012)

Chair: **Krzysztof Rybinski**

- **Jakub Szabo**: Analyzing Sentiments towards the European Union in Slovak Parliamentary Speeches (1994–2023)
- **Jakub Stauber**: The Narratives of the War in Ukraine in Czech News Media
- **Margaryta Khvastova**: Democracy in Donbas: Using Text Analysis for Studying Local Policy Adjustments to Hybrid Threats
- **Csaba Molnár**: The ParLText database of the V-Shift Momentum project

15:00-16:30

Panel B: Emotions and bias in AI (K011-K012)

Chair: **Jakub Szabo**

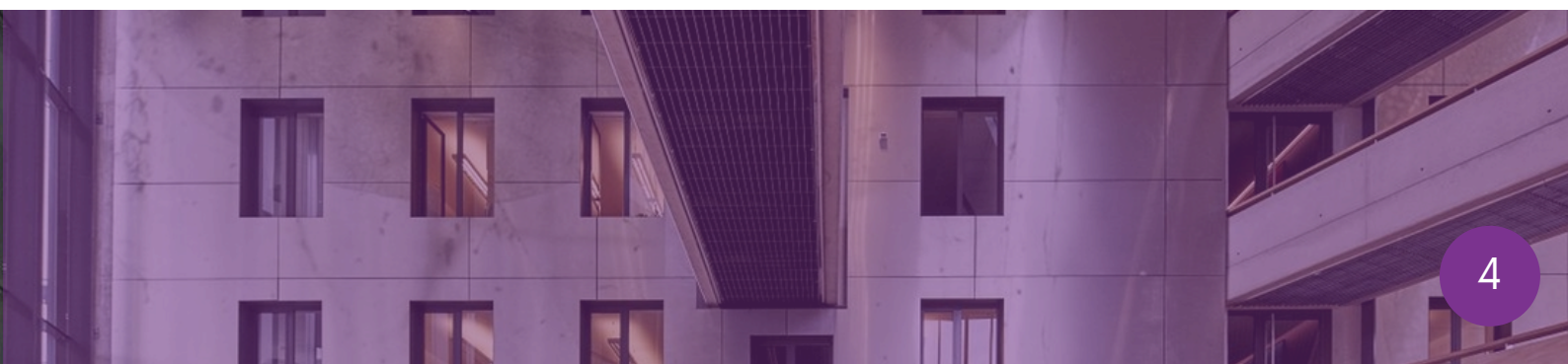
- **Hubert Plisiecki**: Beyond the Black Box: Addressing Participant Bias in Machine Learning and Lexicon Models for Emotion Prediction from Text
- **Krzysztof Rybinski**: Leveraging large language models for comprehensive psychological analysis: insights from four theoretical frameworks
- **Janice Butler**: Large Language Models in the Social Sciences: An Example Using Humour
- **Zsolt Boda**: What kinds of emotions are mobilised by different policy fields? A text mining analysis of parliamentary speeches

15:00-16:30

Panel C: Open Science and LLMs (T.2.37-38.)

Chair: **Miklós Sebők**

- **Maximilian Weber**: Large Language Models in the Wild: A Comparative Analysis of AI Discourse in Traditional Media Post-November 2022
- **Yen-Chieh Liao**: Electoral Systems and Legislative Preferences: Pork Barrel Statements in Parliamentary Questions
- **Viktor Kovács**: Leveraging Open Large Language Models for Multilingual Policy Topic Classification: The Babel Machine Approach.
- **Levente Littvay**: The Zeus Protocol: Automating Survey Experimental Social Science Research Discovery with Large Language Models
- **Marcello Carammia, Stefano Iacus**: Rethinking Scale: The Efficacy of Fine-Tuned Open-Source LLMs in Large-Scale Social Science Research



NOVEMBER 21 (THURSDAY)

16:45-17:45 Keynote speech (K011-K012)

Michal Parížek: Tracing the Narratives of International Conflicts Using Large Language Models

As both military and non-military conflicts proliferate in world politics of today, a need is emerging in scholarly research to develop the ability to detect how the conflicts are narrated in politically relevant textual content, such as in news and social media, in political statements, or in debates in relevant political forums. What narratives of the conflicts prevail in political debates around the world shapes the conflicts themselves, but it also conditions their long-term impact on the international order. This talk will discuss how Large Language Models and Generative AI can be used to detect the competing narratives of international conflicts empirically. It will present the results of the investigations of the prevalence of the competing narratives of the Russia-Ukraine war using novel data on the textual content of online news media worldwide in 2022–2024 from project GLOWIN.

19:00-21:00 **Match-making session**
@Vígvarjú restaurant

**2 Vigadó tér,
H-1051,
Budapest,
Hungary**



PANELS ON FRIDAY

9:30-11:00

Panel D: Regulatory and Ethical Implications of Generative AI (T.2.37-38.)

Chair: **Kitti Mezei**

- **Márton Varju:** Regulating General Purpose AI: A Comparative Analysis of Regulatory Models in the UK, the EU, the USA, and Australia
- **Gergely Csősz:** Copyright in the Age of Generative AI
- **Anna Zsófia Horváth:** Black box or glass box? Transparency challenges concerning generative AI systems
- **Kitti Mezei:** Securing the Future: Addressing Cybersecurity Risks, Misuse and Regulatory Challenges in Generative AI and Large Language Models

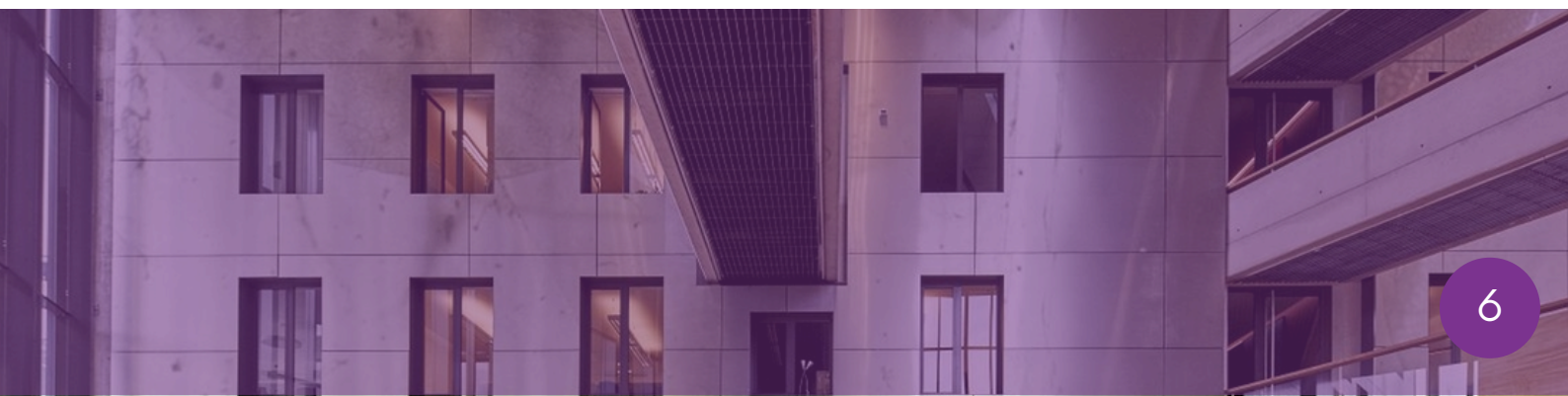
9:30-11:00

Panel E: GenAI and LLM applications (K011-K012)

Chair: **Jakub Stauber**

- **Hossein Kermani:** Computational Identification Of Incivility: A comparison of three methods on Farsi tweets
- **Massimo Terenzi:** Using Semantic Visual Clustering with LLMs to Detect Coordinated Manipulative Narratives in Online Gambling Networks
- **Anna Takács:** Crisis-Driven Frames or Narratively Manufactured Crises? A Research Agenda for the Comparative Study of Illiberal Policy Frames
- **Clément Bénése:** PROMPT: A Multilingual, LLM-powered pipeline for Narrative Trajectories
- **Eduardo Ryo Tamaki:** Chrono-Sampling: Generative AI Enabled Time Machine for Public Opinion Data Collection

11:15-12:45 **Roundtable: Open science in the age of LLMs**
(K011-K012)



ABSTRACTS - PANEL A

Jakub Szabo (Postdoctoral Researcher at Comenius University Bratislava)

Analyzing Sentiments towards the European Union in Slovak Parliamentary Speeches (1994–2023)

The era of digitalisation has expanded opportunities for political science research. In recent years, the growing availability of political texts and advancements in computational tools have facilitated large-scale analyses of political speeches. However, such approaches remain underutilized in the Slovak context. This paper introduces a novel dataset comprising over 420,000 speeches from the National Council of the Slovak Republic (1994–2023), pre-processed for quantitative text analysis. In the second step, the paper examines Slovak politicians' sentiments toward the European Union, revealing that these sentiments peaked around Slovakia's accession and have steadily declined since. Following the 2010 euro crisis, EU-related sentiments reached historic lows and have continued to deteriorate. Lastly, structural topic models are applied to identify the topics most associated with positive and negative EU-related sentiments.

Jakub Stauber (Assistant Professor at Charles University, Prague)

The Narratives of the War in Ukraine in Czech News Media

This study examines how pro-Russian and pro-Western narratives about the war in Ukraine are represented across Czech news media. Czechia's diverse media landscape, encompassing mainstream outlets, tabloids, public broadcasters, and disinformation sources, makes it an ideal case study for narrative analysis. Using a supervised machine learning (SML) approach combined with topic modeling based on document embeddings, the paper analyses how these narratives vary by media type and evolve over time.

The methodology involves two custom SML models based on the deBERTa-V3 architecture, fine-tuned on training datasets generated through a combination of human validation and GPT-4 labeling. These models classify sentences as reporting on the war from a Russian perspective, a Western perspective, or as neutral/informative. In the second phase, BERTopic is applied to extract topics that are characteristic of each narrative. To refine the generated topics, open-source large language models are used to enhance topic representations.

The analysis is based on a unique dataset covering major Czech news media outlets from early 2022 to March 2024. By identifying and analyzing the dynamics of war-related narratives, this research provides a nuanced view of how media outlets in Czechia shape public perceptions of the Ukraine conflict. The findings contribute to broader discussions on media influence and the role of disinformation in reproducing narratives during times of international crisis.

ABSTRACTS - PANEL A

Margaryta Khvostova (Postgraduate Research Student at Margaryta Khvostova)
Democracy in Donbas: Using Text Analysis for Studying Local Policy Adjustments to Hybrid Threats

The paper explores how democratic institutions can adjust their practices to build resilience to hybrid threats. It assesses the steps taken by the Ukrainian central and local legislative institutions to use the window of the hybrid war low-intensity phase to strengthen their legitimacy, especially on the local level. The paper provides an empirical examination of this process by presenting a case study focusing on two frontline towns in Ukraine—Bakhmut and Lysychansk—during the hybrid phase of the Russo-Ukrainian War in 2014-2022. It uses the text analysis methods of topic modelling and word embedding to trace the changes in the legislative practices of local governments in the two towns. The research demonstrates that the local legislation reacted to central policy changes and substantially altered decision-making focus to adjust the new identity-related policies to the local context and react to the hybrid threats.

Csaba Molnár (Research Fellow at HUN-REN Centre for Social Sciences)
The ParlText database of the V-Shift Momentum project

The availability of ready-made textual corpora for research is crucial for social scientists, especially in the current era of rapid advancements in Natural Language Processing (NLP) and Artificial Intelligence (AI) methods. Despite various useful contributions that address issues of accessibility and standardization when it comes to such corpora, in many cases, they have limitations related to scope, geographical coverage, and time frame. This concern is particularly significant in the context of political research on Central-Eastern Europe (CEE), for which such deployment-ready databases are few and far between. We bridge part of this gap by making available a new database: ParlText CEE. The database, prepared under the auspices of the V-Shift Momentum project at the HUN-REN Centre for Social Sciences, covers almost 1.3 million text vectors and metadata for parliamentary speeches, bills, and laws for Czechia, Hungary, Poland, and Slovakia for the period started in 1990-1991 up until 2022-2024. The datasets encompass relevant dates, texts, titles, and, in the case of the speech corpora, parliamentary agendas, speaker names, and parties. All data are also linked based on a unique identifier following the ParlLawSpeech standard. This paper introduces the specifics of the 1.0 release of ParlText CEE and contemplates its possible use cases.

ABSTRACTS - PANEL B

Hubert Plisiecki (Doctoral Candidate at the Polish Academy of Sciences)

Beyond the Black Box: Addressing Participant Bias in Machine Learning and Lexicon Models for Emotion Prediction from Text

Predicting / inferring emotions from text is a popular method of enriching text corpora in social sciences and beyond. It has typically been attempted via a variety of methods, ranging from lexicon-based approaches to machine learning models (ML) and mixed methods. When it comes to choosing one over the other, typical arguments focus on the issues with the black-box nature of the ML approaches, counterbalanced by the lack of syntactic nuance of the lexicon-based methods. These arguments, however, are rarely expanded upon, leading to a simplistic understanding of the consequences of this choice. In my work I have pinpointed specific problems with the propagation of participant bias to the predictions of the black-box ML models. This finding is an important cautionary note for using ML methods research applications. In my presentation you will learn about 1) the actual issues associated with both ML and lexicon models, and 2) the available alternatives that avoid reliance on simplistic lexicon approaches.

Krzysztof Rybinski (Professor at Vistula University)

Leveraging large language models for comprehensive psychological analysis: insights from four theoretical frameworks

The rapid advancement of generative Artificial Intelligence (AI) has significantly transformed various research domains. This paper introduces a novel, fully automated methodology for applying Large Language Models (LLMs) to psychological text analysis. The approach includes prompt design for zero-shot and few-shot learning, model internal consistency analysis, autonomous machine evaluation, and additional human validation. Applied to four psychological theories—Self-Determination Theory, the Big Five Personality Traits, Psychological Well-being, and Cognitive Behavioral Therapy—this methodology is tested on a dataset of 25,780 emails written by a senior executive over 16 years. The analysis involves extracting psychological characteristics from the emails and regressing these characteristics against personal, professional, and environmental factors. The results demonstrate that the methodology provides unique insights into the examined psychological theories, offering a detailed understanding of how various factors influence psychological states and traits over time. This research highlights the potential of LLMs in capturing and analyzing complex psychological patterns in large text corpora, contributing a robust framework for future studies and practical applications in psychological assessment and intervention. The findings underscore the transformative impact of generative AI in psychological research, opening new avenues for understanding human behavior through advanced language models.

ABSTRACTS - PANEL B

Janice Butler (Researcher at the European University Institute)

Large Language Models in the Social Sciences: An Example Using Humour

A growing challenge in social science research is conducting efficient, large-scale content analysis. Traditional computer-aided text analysis (CATA) techniques, while useful, often fail to capture nuanced patterns in complex datasets. Large Language Models (LLMs) offer a promising alternative, enabling rapid, high-fidelity analysis while overcoming limitations associated with manual coding, such as inter-coder reliability.

This paper demonstrates the use of fine-tuned LLMs to automate a particularly challenging task, namely the detection of humour-type and degree. A classifier, trained to detect various forms of humour, is applied to a dataset of over 420,000 political and 250,000 non-political social media posts, illustrating the potential for very large-scale text processing.

This work illustrates how LLMs can address the non-trivial task of humour detection, transcending word-level analysis by identifying deeper linguistic structures. Such advancements underscore the role of LLMs as essential tools in the social sciences, enhancing our ability to interpret complex language phenomena across large datasets.

The ability to quantify reliably “the packaging of the message” is in any case a crucial step forward in automated content analysis. The paper not only explains the potential contributions of LLMs to both social science scholars and policy researchers, but also investigates practical challenges, limitations and recommendations, including those on state-of-the-art tooling.

Findings suggest humour significantly enhances message propagation, with specific humour types (e.g., irony) showing a particularly strong impact for politicians and journalists. Results indicate that while humour is sparingly used in political contexts, it remains an effective tool for engagement, warranting further exploration within social and policy communication.

ABSTRACTS - PANEL B

Zsolt Boda (General Director and Research Professor at HUN-REN Centre for Social Sciences)

What kinds of emotions are mobilised by different policy fields? A text mining analysis of parliamentary speeches

Recently, the role of emotions in defining or influencing behaviour, including political behaviour, has been acknowledged, and research is increasingly addressing how affective processes shape our attitudes, actions, and decision-making. Policy studies have also started to analyse how emotions are reflected in policy discourses and how they influence policy change and support for policies. Most of these studies use qualitative methods. Our paper seeks to contribute to the field by conducting quantitative, text-as-data analysis to identify the emotional content of policy discourses. The aim is to give a descriptive analysis of which emotions are mobilised by different policy fields, which emotions are used by the government and the opposition when framing policies, and how the emotional patterns of policy discourses have changed over time. The parliamentary speech databases of the Hungarian Comparative Agendas Project are analysed using state-of-the-art large language models fine-tuned for emotion analysis. The time frame of the project covers the period 1998-2022. Preliminary findings of the computational analysis confirm the tendency of emotionalisation: the manifestations of emotions increase over time, which is especially true for joy and fear.

ABSTRACTS - PANEL C

Maximilian Weber (Researcher and lecturer at Mainz University)

Large Language Models in the Wild: A Comparative Analysis of AI Discourse in Traditional Media Post-November 2022

In November 2022, the release of ChatGPT and other large language models (LLMs) marked a new era in the widespread use of artificial intelligence and natural language processing applications. This study explores how the discourse surrounding LLMs evolved following the release of ChatGPT, which became a topic of discussion in both traditional and social media. While some individuals lauded the benefits of LLMs, others raised concerns about potential negative impacts, such as spreading false information, perpetuating biases, affecting the job market, and compromising privacy. This study investigates how traditional media discussions evolved after the launch of LLMs to a wider public and whether different news outlets addressed the associated risks of LLM adoption in distinctive ways. Drawing on a dataset of over 22,000 newspaper articles from the past five years, this analysis uses Llama and BERTopic to uncover trends and thematic shifts in media coverage.

Yen-Chieh Liao (Research Fellow at the University of Birmingham)

Electoral Systems and Legislative Preferences: Pork Barrel Statements in Parliamentary Questions

Electoral systems significantly shape legislators' preferences and their expressions of constituency service. This paper quantitatively investigates how electoral rules influence pork-barrel statements in parliamentary questions, using a comprehensive dataset of 148,764 written parliamentary questions from Taiwan's Legislative Yuan during its transition from multi-member districts (MMD) to single-member districts (SMD). To systematically analyze legislative preferences through questioning behavior, I fine-tune three major Chinese language transformer models (Chinese BERT, MacBERT, and Chinese ALBERT) specifically for detecting pork-barrel features in parliamentary questions over time. Using regression analysis, I test the impact of the electoral reform towards SMD while controlling for differences between district magnitude and legislators' individual attributes throughout the reform process. The evidence shows that the electoral reform led to significant changes in legislators' propensity for pork-barrel statements. Analysis reveals that legislators under multi-member districts were more likely to express preferences for particularistic benefits when submitting questions to ministerial offices. The impact of the reform varies significantly across parties, particularly between governing and opposition parties.

ABSTRACTS - PANEL C

Viktor Kovács (Junior Researcher at HUN-REN Centre for Social Sciences)

Leveraging Open Large Language Models for Multilingual Policy Topic Classification: The Babel Machine Approach.

The article presents an open-source and freely available natural language processing system for comparative policy studies. The CAP Babel Machine allows for the automated classification of input files based on the 21 major policy topics of the codebook of the Comparative Agendas Project (CAP). By using multilingual XLM-RoBERTa large language models, the pipeline can produce state-of-the-art level outputs for selected pairs of languages and domains (such as media or parliamentary speech). For 24 cases out of 41, the weighted macro F1 of our language-domain models surpassed 0.75 (and, for 6 language-domain pairs, 0.90). Besides macro F1, for most major topic categories, the distribution of micro F1 scores is also centered around 0.75. These results show that the CAP Babel machine is a viable alternative for human coding in terms of validity at less cost and higher reliability. The proposed research design also has significant possibilities for scaling in terms of leveraging new models, covering new languages, and adding new datasets for fine-tuning. Based on our tests on manifesto data, a different policy classification scheme, we argue that model-pipeline frameworks such as the Babel Machine can, over time, potentially replace double-blind human coding for a multitude of comparative classification problems.

Levente Littvay (Research Professor at HUN-REN Centre for Social Sciences)

The Zeus Protocol: Automating Survey Experimental Social Science Research Discovery with Large Language Models

With advances in AI technology, prominent accelerationists, futurists, and AI developers are predicting an “intelligence explosion” happening not within decades but years. At the core of these arguments lies the automation of scientific advancement. Early attempts to automate scientific advancement in domain-specific ways are already widespread. Alphafold recently won the Nobel Prize for advances in Chemistry, AlphaChip’s AI is advancing the layout of the microchips it runs on, while sakura.ai developed fully automated AI scientists that follow all steps of the scientific inquiry from literature review through project ideas to paper writing and peer review in the domain of algorithmic advancement. Automation of scientific inquiry could ensure fast and efficient scientific advancement with little to no human assistance. While automation attempts are ongoing, few such efforts have received adequate scientific scrutiny to assess their effectiveness, especially in the social sciences. This article aims to automate a narrow slice of social scientific inquiry: survey experiments, from research goals to pre-testing with silicon samples by taking survey experimental research goals, producing AI-proposed mechanisms to achieve those goals, and automatically designing promising experimental treatments. Finally, AI pre-tests the proposed experiments on silicon samples in anticipation of human subject experiments. We apply the procedure to the problem of political depolarization, an especially difficult contemporary societal problem where survey experiments yielded few good solutions. This analysis aims to take us one step closer to automating scientific inquiry in the social sciences.

ABSTRACTS - PANEL C

Marcello Carammia (Associate Professor at the University of Catania), **Stefano Iacus** (Director of Data Science and Product Research at the Institute for Quantitative Social Science, Harvard University)

Rethinking Scale: The Efficacy of Fine-Tuned Open-Source LLMs in Large-Scale Social Science Research

Large Language Models (LLMs) are distinguished by their architecture, which dictates their parameter size and performance capabilities. Social scientists have increasingly adopted LLMs for text classification tasks, which are difficult to scale with human coders. While very large, closed-source models often deliver superior performance, their use presents significant risks. These include lack of transparency, potential exposure of sensitive data, challenges to replicability, and dependence on proprietary systems. Additionally, their high costs make them impractical for large-scale research projects.

In contrast, open-source models, although available in various sizes, may underperform compared to commercial alternatives if used without further fine-tuning. However, open-source models offer distinct advantages: they can be run locally (ensuring data privacy), fine-tuned for specific tasks, shared within the research community, and integrated into reproducible workflows. This study demonstrates that small, fine-tuned open-source LLMs can achieve equal or superior performance to models such as ChatGPT-4. We further explore the relationship between training set size and fine-tuning efficacy in open-source models. Finally, we propose a hybrid workflow that leverages the strengths of both open and closed models, offering a balanced approach to performance, transparency, and reproducibility.

ABSTRACTS - PANEL D

Márton Varju (Research Professor, HUN-REN CSS Institute for Legal Studies)

Regulating General Purpose AI: A Comparative Analysis of Regulatory Models in the UK, the EU, the USA, and Australia

Driven by the core dilemmas of regulating rapidly changing technologies, policy-makers in different countries have developed different strategies for regulating AI. These strategies also have an impact on how regulation may be developed regarding emergent General Purpose AI technologies, which are considered by policy-makers as a particularly problematic new wave of artificial general intelligence technologies. In this talk, we will compare the general regulatory strategies of AI in four jurisdictions, and introduce and then analyse on that basis regulatory proposal and solutions - if they are available - for General Purpose AI.

Gergely Csósz (Head of the Copyright Administrative Unit at the Hungarian Intellectual Property Office, Member of the Council of Copyright Experts)

Copyright in the Age of Generative AI

In the age of generative AI, where algorithms learn from vast datasets of copyrighted works to create new content, the fundamental principles of copyright are being challenged. This presentation provides an overview of the key legal questions surrounding generative AI and copyright, specifically addressing the use of protected works for training AI models and the ownership of copyright of the AI-generated outputs in the current legal framework. The presentation aims to raise awareness about the complexities of copyright protection in the age of AI and to stimulate further discussion on how to balance innovation with the rights of creators.

ABSTRACTS – PANEL D

Anna Zsófia Horváth (Associate, CMS CEE)

Black box or glass box? Transparency challenges concerning generative AI systems

The widespread use of generative artificial intelligence (AI) systems, such as text and image generators, brings about significant data protection challenges in practice. Due to the complex technical solutions inherent in AI, the intricate processes involved, and the systems' capacity for extensive data processing, such systems often operate in a way that is opaque to affected individuals. Additionally, providers of such systems often only disclose certain aspects of their data processing practices, creating obstacles for individuals seeking to exercise their data protection rights. This lack of transparency is particularly problematic for individuals who do not interact directly with such systems (e.g. in cases where deepfake content is created using photos or videos of third parties).

In light of the above, this presentation aims to outline the data protection requirements related to transparency and the practical challenges in enforcing these requirements in the context of generative AI systems. Participants will also gain insights into solutions and strategies aimed at enhancing transparency in generative AI operations, thus supporting individuals in exercising their data protection rights.

Kitti Mezei (Research Fellow, HUN-REN CSS Institute for Legal Studies)

Securing the Future: Addressing Cybersecurity Risks, Misuse and Regulatory Challenges in Generative AI and Large Language Models

The rapid evolution of Generative AI and Large Language Models (LLMs) presents unprecedented opportunities and challenges across industries. However, alongside their potential, these technologies introduce significant cybersecurity risks, including misuse, exploitation, and the potential for large-scale threats. This talk explores the critical intersection of generative AI, LLMs, and cybersecurity, examining the vulnerabilities unique to these technologies and the potential misuse scenarios. Additionally, it delves into the regulatory landscape, analysing current and emerging frameworks aimed at mitigating risks while promoting responsible AI development and deployment.

ABSTRACTS - PANEL E

Hossein Kermani (MSCA post-doctoral researcher at the University of Vienna)

Computational Identification Of Incivility: A comparison of three methods on Farsi tweets

Incivility is a pertinent problem in the contemporary social media environment. While there is a substantial line of study in this field, using computational methods in incivility identification is still uncertain. More significantly, computational endeavors in detecting incivility on social media are mainly focused on mainstream languages like English and German. As a result, our knowledge of how automated techniques work in other languages, such as Farsi, is niche. In this paper, I will compare three different methods to better understand the strengths and weaknesses of computational methods in incivility identification: Human-driven qualitative coding, supervised machine learning (SML), and large language models. The empirical analysis of this study concentrates on WLF, e.g., the #MahsaAmini movement on Persian Twitter. WLF was a large-scale protest against the Islamic regime in Iran sparked by the murder of Mahsa Amini for 'violating hijab' by the so-called 'morality police'. I have collected all popular tweets, i.e., tweets with more than 1k likes per day, during the first months of the movement (N= 38,936). Five human coders coded this dataset qualitatively to identify uncivil messages, comprising 45% of the dataset. In the next step, I will use BERT as an SML model and ChatGPT as the baby poster of LLMs to understand the extent to which these machine algorithms could identify uncivil tweets. Findings will enhance our understanding of how computational methods work on understudied languages. The relevance of the research to the workshop theme is clear enough, I think, as it exactly focuses on the benefits, challenges, and limitations of computational models in social science research, particularly with its focus on marginal languages and contexts.

ABSTRACTS - PANEL E

Massimo Terenzi (Postgraduate researcher at Università degli Studi di Urbino)

Using Semantic Visual Clustering with LLMs to Detect Coordinated Manipulative Narratives in Online Gambling Networks

"This study investigates coordinated inauthentic behavior within Facebook groups promoting online gambling schemes, focusing on a methodological framework based on large language models (LLMs) and semantic visual clustering (Arminio et al. 2024). Our approach aims at identifying and interpreting manipulative patterns in image content shared across these groups, which appears to be centrally orchestrated to exploit user attention through recurrent, visually engaging media (Simon 1971; Wu 2016; Giglietto et al. 2023).

Leveraging OpenAI gpt-4o VLLM, we processed over 2,100 images by first generating descriptive labels and converting these to embeddings with OpenAI text-embedding-3-large model, allowing for semantic similarity analysis in a high-dimensional space. Dimensionality reduction techniques, specifically PCA and UMAP, were then applied to streamline these embeddings while preserving critical relational structures. Subsequently, we employed HDBSCAN clustering, yielding 11 distinct thematic clusters, representing different narrative strategies aimed at user manipulation. Each cluster unveils specific persuasive techniques, such as financial incentives or aspirational imagery, designed to maximize user engagement and drive traffic to betting platforms.

The insights from this methodological framework underscore the potential of LLMs to manage and analyze extensive multimedia datasets effectively, enabling the detection of coordinated manipulation in social platforms. By refining these techniques, this study contributes to the broader application of LLMs in social science research, particularly in the automated analysis of visual data within digital ecosystems."

Anna Takács (Junior Researcher at HUN-REN Centre for Social Sciences)

Crisis-Driven Frames or Narratively Manufactured Crises? A Research Agenda for the Comparative Study of Illiberal Policy Frames

The pervasive and growing illiberal movement is, perhaps, the greatest global challenge to democracy today. Scholars argue that domestic and international crises have played an important role in perpetuating illiberalism among leaders and growing its support among their populace. In this paper, we analyse how illiberal leaders have used policy crises to communicate their policy ideas through illiberal policy frames (IPFs). Our analysis includes two steps. First, we define and measure IPFs in four countries (Austria, Germany, Hungary and the United States) regarding two policy issues (migration and COVID-19). Second, we assess the extent to which these frames are genuine proposals to address extant policy crises or are used to manufacture policy crises. In taking these steps, we develop a new methodology to tackle these questions based on a novel codebook of IPFs and state-of-the-art quantitative social science methods, including large language models.

ABSTRACTS - PANEL E

Clément Bénése (Senior AI Researcher at Opsci.ai)

PROMPT: A Multilingual, LLM-powered pipeline for Narrative Trajectories

Progress in the field of Generative AI (GenAI), and especially of Large Language Models (LLMs), deeply changed the landscape of social networks and the manner we consume online content. In particular, disinformation and narratives created from malicious actors hold a dual relationship with GenAI. On the one hand, modern tools and techniques enable wide dissemination of synthetic narratives and empower malicious actors by lowering the cost of precision targeted campaigns. On the other hand, those same tools allow journalists and social practitioners to fact-check narratives at scale or have metrics of authenticity or coordination, including for cross-platform content. While computational toolkits cannot replace human journalists, they can empower them in a positive feedback loop akin to Reinforcement Learning from Human Feedback (RLHF).

The Predictive Research On Misinformation & Propagation Trajectories (PROMPT) project is aimed at such an empowerment of journalists. In this presentation, we explore the pipeline we adopted. While research in this field is often interested in multimodal and cross-platform content, our specificity is the concern about multilingual analysis, especially in Retrieval Augmented Generation (RAG) steps. Indeed, narratives trajectories, especially when it comes to disinformation, are often intertwined with Foreign Information Manipulations & Interferences (FIMIs) and being able to track such content over several languages allow for better understanding of its author's identity and intent. Given this, adequate performances across various platforms and languages is mandatory. Complementary to this, our pipeline remains as model-agnostic as possible in order to follow technological advances, especially when it comes to open-source and proprietary models. Considering the ever-increasing cost of computation for edge-models, we rather approach this issue with classical techniques from statistical fields. We use medium-sized finetuned models, specialized for topics and platforms of interest, along with Mixture-of-Experts in order to obtain results at a lesser and modular cost, while keeping the flexibility for changes in the narrative landscape.

ABSTRACTS - PANEL E

Eduardo Ryo Tamaki (Research Fellow at the German Institute for Global and Area Studies (GIGA))

Chrono-Sampling: Generative AI Enabled Time Machine for Public Opinion Data Collection

This paper introduces "Chrono-sampling," a novel method leveraging Large Language Models (LLMs) to simulate historical survey respondents, enabling social science researchers to explore past public opinions as if they had access to a "time machine." The study builds on recent advancements in generative AI, particularly LLMs like OpenAI's GPT, which have demonstrated the ability to mimic human attitudes and behaviors. By employing techniques such as "time-gating" and "Clio contexts," we restrict LLMs' knowledge to specific historical periods and provide them with context-rich backstories to enhance the realism of their simulated responses. Utilizing data from the American National Election Studies (ANES), we replicated sociopolitical attitudes from key historical periods, all the way back to the Reagan era. Our results indicate that LLM-generated "silicon" samples can effectively mirror the dynamic relationships observed in human responses, particularly in how retrospective and prospective economic evaluations shift with political and economic changes. This method opens new avenues for historical research, allowing scholars to generate and analyze synthetic data from periods and contexts where traditional data collection is unfeasible. This pilot study also highlights the potential and limitations of using AI in social science research, emphasizing the need for careful methodological considerations when interpreting AI-generated data.

CONTACTS

Name	Affiliation	Email
Clément Bénése	Opsci.ai	clement@opsci.ai
Zsolt Boda	HUN-REN Centre for Social Sciences	boda.zsolt@tk.hu
Janice Butler	European University Institute	janiceelizabeth.butler@eui.eu
Marcello Carammia	University of Catania	marcello.carammia@unict.it
Gergely Csősz	Hungarian Intellectual Property Office	gergely.csosz@hipo.gov.hu
Hossein Kermani	University of Vienna	hosseinkermani@gmail.com
Margaryta Khvostova	University of Surrey	m.khvostova@surrey.ac.uk
Yen-Chieh Liao	University of Birmingham	davidycliao@gmail.com
Levente Littvay	HUN-REN Centre for Social Sciences	littvay.levente@tk.hu
Kitti Mezei	HUN-REN Centre for Social Sciences	kitti.mezei@tk.hu
Csaba Molnár	HUN-REN Centre for Social Sciences	molnar.csaba@tk.hu

CONTACTS

Name	Affiliation	Email
Michal Parížek	Charles University	michal.parizek@fsv.cuni.cz
Hubert Plisiecki	Polish Academy of Sciences	hubert.plisiecki@gssr.edu.pl
Orsolya Ring	HUN-REN Centre for Social Sciences	orsolya.ring@tk.hu
Krzysztof Rybinski	Vistula University	rybinski@rybinski.eu
Miklós Sebők	HUN-REN Centre for Social Sciences	miklos.sebok@tk.hu
Jakub Stauber	Charles University	jakub.stauber@fsv.cuni.cz
Jakub Szabo	Comenius University	jakub.szabo@uniba.sk
Anna Takács	HUN-REN Centre for Social Sciences	anna.takacs@tk.hu
Eduardo Ryo Tamaki	CEU Democracy Institute	TamakiE@ceu.edu
Massimo Terenzi	University of Urbino	massimo.terenzi@uniurb.it
Márton Varju	HUN-REN Centre for Social Sciences	varju.marton@tk.hu
Maximilian Weber	University Mainz	weber.aca@gmail.com