PHD SEMINAR „COMPUTER TECHNOLOGIES AND APPLICATIONS“

Seminar 1 | 16 April 2026

ABSTRACTS

1

Ivan Bekyarov

A SURVEY OF INTELLIGENT DATABASE MONITORING SYSTEMS

ABSTRACT

This survey examines modern intelligent database monitoring systems used in practice, including solutions such as Prometheus, Netdata, Percona Monitoring and Management, and others. The aim is to present the current state of technologies and the key trends in the development of intelligent monitoring.

Miglena Dodleva

USING AZURE MCP SERVER FOR AI-ASSISTED CLOUD DEVELOPMENT

ABSTRACT

In this paper we explore the integration of AI agents and large language models with software systems and development environments. While modern AI assistants such as GitHub Copilot and ChatGPT significantly improve developer productivity, they typically lack direct access to external systems and cloud resources, which limits their practical use in production environments. To address this limitation, the Model Context Protocol (MCP) introduces a standardized approach for connecting AI agents with external tools and services. MCP enables large language models to interact with applications, data sources, and operational environments through a unified mechanism for integrating with software systems. This paper presents the architecture and capabilities of Azure MCP Server, an open-source implementation that exposes services from Microsoft Azure as MCP-compatible tools. The proposed approach illustrates how MCP can support the application of AI capabilities in real-world applications and development environments.

Kristiyan Kolev

CONTEXTUAL ARCHITECTURE FOR THE APPLICATION OF LANGUAGE MODELS IN CLOUD INFRASTRUCTURE MANAGEMENT

ABSTRACT

This paper examines the integration of language models into the management of internal infrastructure platforms. A key challenge is that general-purpose models do not possess knowledge of the specific technology stack, policies, and operational constraints of a given organization. The proposed architecture combines semantic search, vector databases, and staged workflow mechanisms to refine user intent against actually available resources. Kubernetes serves as the control plane, while organizational knowledge is structured as contextual memory. The main conclusion is that the architecture surrounding the model is of greater importance than the model itself.

ABSTRACTS

2

Vasil Kostadinov

FORMAL MODELS AND PRACTICAL APPLICATIONS OF RATIONAL AGENTS IN MODERN ARTIFICIAL INTELLIGENCE

ABSTRACT

This presentation examines the concept of rational agents as one of the fundamental models in the field of artificial intelligence. A rational agent is defined as a system that perceives its environment through a sequence of percepts and takes actions aimed at maximizing a predefined utility function or performance measure. In this context, the main components of agent architectures are analyzed, including perception mechanisms, decision-making models, and goal formalization. Special attention is given to the interaction between the environment, the agent, and its actions, as well as the role of internal representations and knowledge. The presentation also addresses the problem of uncertainty inherent in real-world environments, introducing probabilistic methods for modeling and reasoning, including the use of Bayesian networks and stochastic processes. Furthermore, approaches to learning and adaptation are analyzed, with emphasis on machine learning methods and reinforcement learning, which enable agents to optimize their behavior based on accumulated experience and interaction with the environment. In conclusion, the key challenges and future perspectives in the development of agent-based models are outlined, including the need for integrating different paradigms, improving interpretability, and ensuring the reliability of intelligent systems. Contemporary applications of rational agents in autonomous systems and cyber-physical systems are also discussed, along with potential directions for their application in dissertation research.

Pavel Kyurkchiev

IMPACT OF GRAPH TOPOLOGY ON COMPUTATIONAL COMPLEXITY IN SEMANTIC INDEXING OF SYSTEM DYNAMICS MODELS

ABSTRACT

The integration of semantic search methods utilizing vector embeddings in computer modeling platforms requires significant computational resources. This study investigates the computational challenges associated with the serialization and semantic indexing (the Write/Upsert phase) of System Dynamics models. By profiling a cloud microservices architecture, a comparative latency analysis is conducted on the processing of graph structures with varying topological complexity. Experimental data reveal an interesting phenomenon: the computational complexity during topological sorting and linearization does not depend on the absolute number of components, but rather on the density of overlapping algebraic loops (feedback loops). The results demonstrate that a densely interconnected ecological model with only 6 nodes generates an exponential spike in backend processing time (over 12 seconds), whereas larger, linear models (with 10 nodes) are processed in approximately 2.8 seconds. Concurrently, the AI vectorization via local ONNX inference maintains a constant bottleneck of about 5.4 seconds. These findings demonstrate that the direct application of monolithic architectures in intelligent simulation systems would lead to severe user interface blocking (nearly 18 seconds) during write operations. Consequently, the implementation of an asynchronous, event-driven microservices architecture is essential to isolate heavy topological and NLP operations, ensuring the scalability and responsiveness of modern cloud knowledge repositories.

ABSTRACTS

3

Viktor Matanski

MULTI-AGENT SYSTEMS FOR CREATIVE TASKS: ARCHITECTURES, COORDINATION MECHANISMS, AND EVALUATION CHALLENGES

ABSTRACT

Multi-agent systems are increasingly used in creative artificial intelligence because they decompose complex creative processes into coordinated roles such as perception, planning, generation, critique, memory management and revision. In contrast to monolithic models, these systems enable modular, iterative refinement together with better control over outputs across tasks such as text generation, visual art synthesis, music composition, design ideation and transformation. This report examines the main architectures of multi-agent systems for creative tasks, including hierarchical, collaborative, debate-based and human-in-the-loop configurations. It also analyzes the coordination mechanisms that govern their operation, such as task delegation, shared memory, feedback loops, and role specialization. The report argues that the main advantage of the multi-agent approach lies in its capacity to structure creativity as a process of constrained, verifiable, and revisable generation.

Zlatomila Milcheva

THE EVOLUTION OF TIME SERIES FORECASTING: FROM STATISTICAL BASELINES TO COMMERCIAL FOUNDATION MODELS

ABSTRACT

Time series forecasting is a critical component of enterprise decision-making, yet the underlying technologies driving these predictions are undergoing a rapid and fundamental transformation (Hyndman & Athanasopoulos, 2021). This study provides a comprehensive overview of the evolving landscape of applied time series analysis. We trace the evolutionary path of forecasting methodologies, beginning with traditional, localized statistical models such as ARIMA (Box & Jenkins, 1976) and progressing to the deep learning era dominated by recurrent neural networks, including Long Short-Term Memory models (Hochreiter & Schmidhuber, 1997), as well as Temporal Convolutional Networks designed for sequential data modeling (Bai, Kolter & Koltun, 2018). The focus then shifts to the emerging paradigm of Time Series Foundation Models (TSFMs). Mirroring the trajectory of Natural Language Processing, the field is increasingly moving toward large-scale pre-trained models capable of cross-domain generalization and zero-shot forecasting (Bommasani et al., 2021; Zhou et al., 2023). To ground this trend in practical application, the study examines how major cloud providers are commercializing these advancements. In particular, we highlight enterprise forecasting architectures currently deployed by Amazon Web Services (Chronos) (Ansari et al., 2024), Google Cloud (TimesFM) (Das et al., 2023), and emerging forecasting capabilities within Microsoft Azure's AI ecosystem. These platforms aim to significantly reduce engineering complexity while enabling scalable and accessible forecasting solutions for enterprise environments.

ABSTRACTS

4

Bozhidar Samokovski

COMPARATIVE ANALYSIS OF EEG FUNDAMENTAL MODELS

ABSTRACT

This paper presents a comparative analysis of the emerging category of fundamental models for electroencephalographic signal processing. With the advent of large-scale architectures trained on large datasets, the paradigm in neurotechnology is shifting from specific classifiers to universal neural encoders. The study compares leading models as of 2026, including ZUNA, LaBraM, and NeuroLM, and examines their specific approaches. The analysis focuses on critical metrics such as noise robustness, hardware independence, and language decoding capacity (Brain-to-Text). The results of the review synthesize the advantages and limitations of different architectures, offering a framework for selecting the optimal model depending on the application needs of brain-computer interfaces.

Detelinka Trifonova

METHODOLOGIES FOR DOMAIN-SPECIFIC ADAPTATION OF LARGE LANGUAGE MODELS

ABSTRACT

Although general-purpose large language models (LLMs) demonstrate remarkable reasoning and linguistic capabilities, their performance often falters in specialized sectors such as medicine, law, and high-tech engineering. These “out-of-the-box” domains require precision, specialized terminology, and adherence to unique logical structures that are absent from general pre-training data. Here, we provide a comprehensive overview of the current landscape of methodologies used to infuse domain expertise into LLMs, ranging from inference-time adjustments to architectural modifications. We categorize these methodologies into four main layers based on their computational cost and depth of knowledge integration:

- Contextual learning and hint engineering: The most accessible layer, using multi-step hinting and complex system instructions to guide the model’s behavior without changing the underlying weights.
- Retrievable Augmented Generation (RAG): A dynamic approach that connects LLMs to external, domain-specific databases. We discuss how RAG mitigates hallucinations by providing verifiable context at inference time.
- Parametrically efficient fine-tuning (PEFT): Methods such as low-rank adaptation (LoRA) and adapters allow the injection of domain knowledge by training a subset of the model parameters, offering a compromise between performance and resource efficiency.
- Domain-adaptive pre-training (DAPT): The most intensive approach, involving extensive pre-training on large corpora of domain-specific text to align the model’s internal representations with specialized nomenclature and semantics.

The presentation concludes with a comparative analysis of these strategies, assessing the trade-offs between data requirements, hardware constraints, and the resulting “domain-fluency”.