

Aleksandar Gavric

 aleksandargavric.com  Google Scholar  aleksandargavric  aleksandar.gavric@tuwien.ac.at  gavric97@gmail.com
 Vienna, Austria  0009-0005-1243-7722  @AcaGavric  LinkedIn  +43 667 777 7542  +41 77 277 5804

RESEARCH INTERESTS AND PERSONAL STATEMENT

I am a researcher working on advancing **world-model construction** by integrating **multimodal process mining** with AI-driven **mixed-reality (MR) simulation** and **generative video techniques** for the discovery of **tacit expertise**. My research focuses on extracting precise and tacit process knowledge from heterogeneous multimodal data streams—including video, audio, interaction logs, and environmental signals—and converting these insights into actionable world models. I develop methods and tools that enable **validation**, **explainability**, and **training** of downstream **AI systems** that learn **real-world processes** and their **latent behavioral patterns** end-to-end, unifying conceptual modeling and knowledge elicitation within a single continuous learning loop. From my personal life, I play piano, drums, guitar, cajón, ukulele, and am currently learning tenor saxophone; I have also co-organized one of Austria's historic balls in the main royal palace, dating back to 1846 multiple times.

EDUCATION

- | | |
|---|---|
| • PhD in Computer Science (Dr. techn.) | Technische Universität Wien, TU Wien |
| <i>Thesis: Enhancing Conceptual Modeling through Multimodal Data Analysis and MR</i> | 2022 – Present · Expected Feb 2026 |
| • M.Eng. in Electrical Engineering and Computing — Data Engineering | University of Nis |
| <i>GPA: 3.96; Research Focus: Medical MLLM for HoloLens-Based Radiology Imaging Transparent AI</i> | 2020 – 2024 |
| • M.Sc. in Information Systems | Union University, Belgrade |
| <i>GPA: 4.00; Research Focus: Data Marts and Data Warehousing AI in Wireless Sensor Networks</i> | 2020 – 2021 |
| • B.Eng. with Honors in Electrical Engineering and Computing | University of Nis |
| <i>GPA: 3.73; Research Focus: Apache Spark Implementation of the Constrained Clustering Algorithm</i> | 2016 – 2020 |

PROFESSIONAL EXPERIENCE

- | | |
|---|--------------------------------|
| • TU Wien, The Business Informatics Group | Vienna, Austria |
| <i>University Assistant</i> | January 2023 – present |
| ○ Research, Software Engineer, Teaching & Examination: Co-designed and co-organized the international workshop Multimodal Process Mining (MMPM) ; taught several conceptual modeling-related courses; served as departmental webmaster; and co-supervised multiple Master's theses in conceptual modeling and Mixed Reality. Researched world models , multimodal data analysis , video generation , and mixed reality simulations . | |
| ○ Technology Stack: Python, C++, C#, Three.js, PM4Py, PyTorch, OpenCV, AlgoVision, Genie, Open-Sora, CogVideoX, Pika, Luma Dream Machine, Unity, MRTK, Node.js, Next.js, WebXR, WebGL, HuggingFace Diffusers, Whisper, CLIP, SAM. Custom video–image–audio preprocessing pipelines; Retrieval-Augmented Generation with vector databases (Pinecone, Weaviate), Open AI Real-Time Voice API, and large-scale pipelines for training and evaluating open-source video generation , world-modeling , and multimodal architectures. | |
| • EPFL, Mackenzie Mathis Lab | Lausanne, Switzerland (remote) |
| <i>Teachning Assistant</i> | September 2022 – November 2022 |
| ○ DeepLabCut Cajal Course: Guided students in analyzing custom or publicly relevant video datasets using DeepLabCut for animal/human behavioral research; supervised hands-on work in tracking , pose estimation , action segmentation , kinematic analysis , and behavior modeling . | |
| ○ Research Discussion & Crossmodal Foundations: Explored Stable Diffusion and generative vision models to neural data imaging, including joint crossmodal embeddings , representation alignment, and multimodal decoding for computational neuroscience pipelines. | |
| • Union University, The School of Computing | Belgrade, Serbia |
| <i>University Assistant</i> | October 2021 – August 2022 |
| ○ Research, Teaching & Examination: Delivered practical sessions and created/grated exams for courses including Deep Learning , Machine Learning , Algorithms and Data Structures , Compiler Construction , and Integrated Information Systems . Researched various data lakes to compare storage formats, indexing strategies, metadata management, and integration capabilities. | |
| • Rice University, Kavraki Lab | Houston, TX, USA |
| <i>Researcher</i> | July 2021 – October 2021 |
| ○ Peptide–HLA Structural Clustering: Conducted large-scale clustering of peptide–HLA complexes using geometric embeddings , t-SNE, and UMAP to uncover recurrent structural motifs and presentation patterns . | |
| ○ Technology Stack: Python, BioPython, MDAnalysis, MDTraj, PyRosetta, scikit-learn, NumPy/SciPy, GROMACS/AMBER trajectory parsing, PCA/t-SNE/UMAP. | |
| • Rice University, Rice Networks Group | Houston, TX, USA (remote) |
| <i>Researcher</i> | July 2020 – October 2020 |
| ○ Adaptive Swarm Learning: Working on ASTRO (astro.rice.edu) with Dr. Edward Knightly and Dr. Ahmed Boubrima. Created regression and calibration models to analyze correlations between low-cost gas sensors and high-precision reference stations , including spatio-temporal interpolation , and uncertainty estimation . | |
| ○ Technology: Python, C, C++, PyTorch, scikit-learn, NumPy/SciPy, Kalman filtering, GeoPandas, rasterio. | |
| • Self-Employed, AI-for-Hotels.com | Toronto, Canada (remote) |
| <i>Founder & Lead Architect</i> | July 2017 – January 2023 |
| ○ Robotics & Holographic Systems: Deployed UniTree Go2 robots using SLAM, and API task control; created interactive holographic displays using transparent LCD cabinets, and spectroscope-based phantom imaging. | |
| ○ Multimodal Intelligence: Engineered multimodal RAG pipelines using the OpenAI Real-Time Voice API . | |

PEER-REVIEWED PUBLICATIONS (SELECTED)

- **Gavric, A., Bork, D., & Proper, H. A. (2025). Towards the Enrichment of Conceptual Models with Multimodal Data:** 33rd International Conference on Information Systems Development (ISD 2025).
Description: Proposed embedding-based **fusion of multimodal data** to automatically enrich conceptual models; introduced **crossmodal alignment** techniques for mapping raw evidence to conceptual constructs.
- **Gavric, A., Bork, D., & Proper, H. A. (2025). Petri Net of Thoughts: A Structure-Enhanced Prompting Approach for Process-Aware Artificial Intelligence:** 15th International Workshop on Enterprise Modelling and Information Systems Architectures (EMISA 2025), pp. 105–110.
Description: Introduced a Petri net structure-guided **prompting paradigm**, improving **reasoning over process structures**; demonstrated gains in **trace classification**, **next-event prediction**, and **process-aware reasoning**.
- **Gavric, A., Bork, D., & Proper, H. A. (2025). Petri Net Structure-Driven Video Generation:** NeurIPS 2025 Workshop “What Makes a Good Video”.
Description: Built a **structure-conditioned video generation** pipeline using Petri nets as **causal scaffolding** for diffusion models; enabled controllable, process-driven synthetic video creation.
- **Gavric, A., Bork, D., & Proper, H. A. (2025). Beyond Logs: AI's Internal Representations as the New Process Evidence:** 23rd International Conference on Business Process Management (BPM 2025), Process Technology Forum, pp. 232–246. Springer. Seville, Spain.
Description: Introduced a methodology for process mining on latent AI representations as process evidence; defined and implemented a **relaxed process discovery** and **conformance checking** paradigm for process flows **discovered from AI models' internal states**.
- **Gavric, A., Bork, D., & Proper, H. A. (2025). Surgery AI: Multimodal Process Mining and Mixed Reality for Real-time Surgical Conformance Checking and Guidance:** 7th Central European Workshop on Services and their Composition (ZEUS 2025). Heilbronn, Germany.
Description: Developed a multimodal process-mining pipeline integrating surgical video, instrument-tracking, and OR sensor streams; implemented real-time conformance checking and procedural guidance in immersive mixed-reality environments.
- **Gavric, A., Bork, D., & Proper, H. A. (2024). Stakeholder-specific jargon-based representation of multimodal data within business process:** Companion Proceedings of PoEM 2024. Stockholm, Sweden.
Description: Designed NLP pipelines that translate multimodal observations into stakeholder-specific jargon, creating **domain-adaptive representations** and **improving explainability**.
- **Gavric, A., Bork, D., & Proper, H. A. (2024). How Does UML Look and Sound? Using AI to Interpret UML Diagrams Through Multimodal Evidence:** Advances in Conceptual Modeling: ER 2024 Workshops, Springer Nature. Pittsburgh, PA, USA.
Description: Applied vision-language models to relate UML diagram elements with **synthetic visual or acoustic evidence**; introduced a user study for linking UML fragments to verbal descriptions and observed interactions.
- **Gavric, A., Bork, D., & Proper, H. A. (2024). Enriching Business Process Event Logs with Multimodal Evidence:** Proceedings of PoEM 2024, Stockholm, Sweden.
Description: Augmented traditional event logs with multimodal signals; Developed a tool that uses **crossmodal encoders** to infer missing or tacit activities.
- **Gavric, A., Bork, D., & Proper, H. A. (2024). Multimodal Process Mining:** 26th International Conference on Business Informatics (CBI 2024). Vienna, Austria.
Description: Formulated multimodal process mining as a representation-learning task; developed unified embeddings combining video, audio, and UI interactions for robust discovery under ambiguity.
- **Marjanović, M., Veljković, S., Mitrović, N., Živanović, E., Gavric, A., & Danković, D. (2024). Modified SPICE-Compatible Model Integrating NBTI and Self-Heating Effects for VDMOS Transistors:** IcETRAN 2024, pp. 1–6. IEEE. Nis, Serbia.
Description: Built physics-informed models predicting transistor degradation; combined simulation data with ML-supported curve-fitting for accuracy under thermal and stress conditions.
- **Ali, S. J., Gavric, A., Proper, H., & Bork, D. (2023). Encoding Conceptual Models for Machine Learning: A Systematic Review:** MODELS-C 2023, pp. 562–570. IEEE. Västerås, Sweden.
Description: Surveyed ML-ready encodings of BPMN/UML/Petri nets; categorized graph-neural, image-based, and text-based approaches; highlighted open challenges in multimodal interoperability.
- **Gavric, A. (2023). Enhancing process understanding through multimodal data analysis and extended reality:** Companion Proceedings of PoEM/EDOC 2023, Vienna, Austria.
Description: Demonstrated XR-based reenactments for generating high-quality multimodal logs; evaluated video-driven activity recognition for process discovery.
- **Gavric, A., Merlinšky, E. A., & Stanimirović, A. (2023). A System for Detection and Tracking of Oculo-Vestibular Complications Associated with Extended Reality Headset Usage:** MIEL 2023, pp. 1–4. IEEE.
Description: Examined the various health issues associated with XR headset usage and introduced a novel system for detecting and monitoring XR-related ocular and vestibular complications, followed by guidance and actions for preventing further complications.
- **Merlinsky, E. A., Gavric, A., Stojković, H., & Živanović, E. (2023). Physics-Driven Methods for Adaptive Optics Effect in Extended Reality:** MIEL 2023, pp. 1–4. IEEE.
Description: Surveying a landscape where the amalgamation of adaptive optics, eye-tracking, and light field displays provides an immersive experience without the constraints of traditional visual corrective aids.
- **Gavric, A., Vujošević, D., Radosavljević, N., & Prvulović, P. (2022). Real-Time Data Processing Techniques for a Scalable Spatial and Temporal Dimension Reduction:** INFOTEH 2022, pp. 1–6. IEEE.
Description: Shows experimentally that it is possible to build more successful predictive models by reducing dimensions and discuss the potential advantages of reducing the spatial and temporal dimensions of sensor measurements in different applications.
- **Gavric, A., Stanimirović, A., & Stoimenov, L. (2021). Identification of Air Pollution Sources using Predictive Models and Vehicular Sensor Networks:** ICIST 2021 Proceedings, pp. 124–127.
Description: Designed and implemented a predictive ML model for a distributed system that applies machine learning algorithms over data streams for efficient estimation of dominant pollution sources in real-time.