

ALEKSANDAR TOMAŠEVIĆ

Data Scientist / Applied AI Researcher

ML/NLP researcher working on online-community analysis, AI evaluation, and multi-agent simulation. Fulbright fellow; author of *transforEmotion*, an R package with Python backend and 9k+ CRAN installs.

Novi Sad, 21000, Serbia

✉ atomashevic@gmail.com | 🌐 www.atomasevic.com | 🐙 [atomashevic](https://github.com/atomashevic) | 🌐 [atomasevic](https://www.linkedin.com/company/atomasevic)

WORK EXPERIENCE

Institute of Physics Belgrade

Belgrade / Serbia

Senior Research Associate, Scientific Computing Laboratory

Jul 2025 – Present

- Built Python and R workflows for large-scale data analysis, simulation, and model evaluation.
- Developed and validated machine learning, NLP, and network-analysis pipelines for behavioral and online-community data.
- Evaluated trust, toxicity, conversation quality, and agent behavior in AI systems and multi-agent simulations.
- Built experiment infrastructure around PostgreSQL, Docker, Linux, and reproducible research workflows.

University of Novi Sad

Novi Sad / Serbia

Assistant Professor / Research Associate

Feb 2016 – Present

- Led data-intensive projects using Python and R to analyze large behavioral, textual, and network datasets.
- Designed machine learning and NLP workflows for online-community research, including text classification, sentiment analysis, toxicity-related analysis, and simulation support.
- Built automation scripts and reproducible pipelines for data cleaning, modeling, visualization, reporting, and experimental analysis.
- Partnered with physicists, psychologists, and sociologists on behavioral, text, and network analysis projects.

University of Virginia, School of Data Science

Charlottesville, VA / USA

Fulbright Visiting Researcher

Oct 2023 – Jan 2024

- Developed *transforEmotion*, an R package with Python backend for multimodal emotion recognition across text, image, and video data.
- Built reproducible machine learning workflows connecting transformer-based models, statistical analysis, and large-scale media data processing.

Eipix Entertainment

Novi Sad, Serbia

Medior Game Designer

May 2014 – Feb 2016

- Worked with cross-functional production teams on game systems, content design, and delivery across 10+ shipped titles.

EDUCATION

University of Novi Sad

Novi Sad, Serbia

PhD in Quantitative Methods

Oct 2012 – Mar 2019

- Research and training centered on statistics, computational methods, network analysis, and large-scale behavioral data in social sciences.

Master in Social Sciences, University of Novi Sad

Oct 2011 – Oct 2012

Bachelor in Social Sciences, University of Novi Sad

Oct 2007 – Jun 2011

PROJECTS

🔗 YSocial Multi-Agent Social Platform Simulation

Institute of Physics Belgrade / Research collaboration Sep 2025 – Present

- Built and evaluated components of an LLM-powered social media twin spanning microblogging and forum-style simulations for trust, toxicity, and conversation-quality research.
- Contributed to a Python/Flask/SQLAlchemy stack with SQLite and PostgreSQL backends; used Dockerized services and OpenAI-compatible backends including Ollama and vLLM.
- Investigated failure modes such as repetitive outputs and thread-structure breaks, using real community data to guide validation and calibration.

📦 MADOC Dataset and Tools

Institute of Physics Belgrade Jan 2025 – Present

- Built a research dataset and tooling stack for large-scale online-community analysis, with emphasis on reproducible ingestion, cleaning, and downstream modeling workflows.
- Implemented NLP pipelines for text classification, toxicity detection, and sentiment analysis over cross-platform community text.
- Developed pyMADOC, a Python package for dataset access and text processing using pandas, NumPy, and scikit-learn.

🔗 transforEmotion R package

Personal collaboration Oct 2023 – Present

- Built an R package with a Python backend for multimodal emotion recognition across text, image, and video, with more than 9k CRAN installs.
- Integrated transformer-based models into reproducible data-processing workflows for large-scale media analysis.
- Maintained the package as production-quality open-source software with GitHub Actions CI/CD and cross-language tooling.

🔗 COVID-19 Data Analysis and Visualization Platform

Data Versus Corona initiative Apr 2020 – Jun 2020

- Engineered R data pipelines to clean, analyze, and visualize epidemiological data from multiple sources.
- Built a real-time data fetching and processing API for automated updates and reproducible reporting.
- Created interactive visualizations for communicating complex health data clearly to broad audiences.

SELECTED SKILLS

Research Focus	online-community analysis, AI evaluation, trust and toxicity, multi-agent simulation
NLP & Modeling	text classification, sentiment analysis, embeddings, transformer fine-tuning, multimodal emotion recognition
Engineering	Python, R, package development, pandas/NumPy, scikit-learn, PyTorch, SQL, Flask
Systems & Infra	PostgreSQL, Docker, Linux, HPC, GitHub Actions, LLM serving (Ollama, vLLM), OpenAI-compatible APIs
Languages	English (C1), French (B1), Serbian (Native)