Anna Smirnova

Lausanne, Switzerland · anna.smirnova@unil.ch · +41 78 234 86 14 · github.com/mechanicpanic

Research interests

Deep learning for economics and finance; large language models as measurement tools in empirical economics; evaluation of NLP systems.

Education

University of Lausanne 2024 – present (expected 2028)

PhD candidate in Economics. Advisor: Prof. Simon Scheidegger. Research: deep-learning methods for economics and finance.

Swiss Program for Beginning Doctoral Students in Economics, Study Center Gerzensee, 2025.

Saint Petersburg State University BSc 2020 · MSc 2022

Faculty of Mathematics and Mechanics. BSc and MSc in computer science (Software and Administration of Information Systems).

Working paper

M. Igami, S. Kusaka, S. Scheidegger, **A. Smirnova**. *Patents as Indicators of Innovation: Large Language Models Meet Ground Truth in the LCD Industry.* August 2026.

Publications

1. **A. Smirnova**, A. Kopan, V. Makeev, G. Chernishev. *What Aggregate Scores Hide: Per-Rule Evaluation of Russian Grammatical Error Correction.* BEA Workshop at ACL 2026.
2. A. Kopan et al. *Citation Style Classification: a Comparison of Machine Learning Approaches.* AI4SciSci Workshop at ICDM 2023.
3. Y. Kuzin et al. *Query Processing and Optimization for a Custom Retrieval Language.* Pan-DL Workshop at COLING 2022.
4. **A. Smirnova** et al. *Situation-Based Multiparticipant Chat Summarization: a Concept, an Exploration-Annotation Tool and an Example Collection.* ACL-IJCNLP Student Research Workshop 2021.

Research projects

- **LLMs as measures of innovation.** LLM classification of all US LCD patents, evaluated against firm-product-fab ground truth and structural welfare estimates (2001–2011). Regression and random-forest prediction of innovation incidence and welfare gains, leave-one-regime-out out-of-sample validation, and reproduction with an open-weight model. LLM-based measures raise the adjusted R^2 for welfare gains from near zero to 12–14%. Data pipeline in Python.
- **Global solution methods for dynamic economic models.** Deep-learning solver for recursive equilibrium models that keeps nonlinearities intact, with Euler-equation accuracy diagnostics; released as an open-source JAX library, DEQN-JAX.
- **Rule-level evaluation of grammatical error correction.** Synterr, an open-source framework for grammar-informed synthetic data and per-rule evaluation of GEC systems, including LLM-based ones; basis of the BEA 2026 paper.
- **Pattern queries over conversational text.** PrismQL, a query language for temporal patterns across message corpora, with an ANTLR grammar and a Python engine.

Teaching

University of Lausanne 2023 – present

Teaching Assistant, Data Science and Advanced Programming and Advanced Data Analytics.

Deep Learning for Economics and Finance, Geneva 2026

Delivered lectures, substituting for one of the instructors.

Industry

Universe Data (formerly Unidata), Saint Petersburg 2021 – 2023

Data Scientist. Entity-resolution pipelines for enterprise customer data; analysis of failure modes of standard deduplication methods on noisy production data.

Languages

Russian (native) · English (C2) · French (A1)