

VALENTIN VELEV

+49 176 59993513 — valentin.velev@gmx.de — valentinsvelev.github.io



Konstanz, Germany

RESEARCH INTERESTS

My research focuses on developing accessible language models and language model-based systems for multilingual and context-aware real-world applications. I am particularly interested in exploring how such models can be used for social good, including supporting education, advancing climate science, improving medical information access and communication, and contributing to our understanding of cognition.

EDUCATION

MSc Social and Economic Data Science

University of Konstanz

October 2023 – Present

Konstanz, Germany

- Preliminary average grade: 1.1 (German grade)
- Courses: Deep Learning (1.0), Word Representations and Language Models (1.0), Statistical Learning (1.3), Data Mining (1.0)
- Master's thesis (in progress): "Benchmarking Retrievers for Robustness and Generalizability in Retrieval-Augmented Generation Systems" (Supervisors: Prof. Dr. Andreas Spitz and Prof. Dr. Daniel Keim)

BA Political Science

Goethe University Frankfurt

October 2019 – January 2023

Frankfurt am Main, Germany

- Final grade: 1.2 (with distinction; German grade)
- Courses: Statistics (1.3), Data Science for Human Geography (1.0)

EXPERIENCE

Teaching Assistant

Department of Computer and Information Science, University of Konstanz

August 2025 – Present

Konstanz, Germany

- Led tutorials and provided technical guidance for the master's-level course "Word Representations and Language Models," covering embedding models, sequence-to-sequence models, the Transformer architecture (including BERT), and generative language models
- Developed and maintained reproducible teaching materials using Python, Git, and \LaTeX

Data Science Working Student

Codify AG

March 2025 – Present

Kreuzlingen, Switzerland

- Designed and implemented an evaluation pipeline for agentic AI systems using Python, MCP servers (for agent tools), LLM-as-a-judge metrics, and Arize Phoenix (for agent traces)
- Built a conversation evaluation framework for clients to assess model performance on multi-turn tasks
- Integrated end-to-end evaluation workflows and visualizations into production systems using Python, TypeScript, and SQL
- Designed prompt injection challenges for workshop at the AIXCHANGE Conference 2025

Student Research Assistant

Center for Data and Methods, University of Konstanz

April 2024 – February 2025

Konstanz, Germany

- Conducted independent quantitative analyses and literature reviews for projects in computational social science, contributing to three peer-reviewed journal publications
- Designed data visualizations and reproducible analysis pipelines in R and Python for publication-quality outputs
- Collaborated on research design, data preprocessing, and statistical modeling using R, Python, Stata, and \LaTeX

PUBLICATIONS

Velev, V., & Spitz, A. (2025). gglyph: Network-Style Visualization of Directed Pairwise Relationships. *CRAN package*.
<https://doi.org/10.32614/CRAN.package.gglyph>

Velev, V., & Schmidt-Catran, A. W. (2024). Economic Inequality and Public Support for Redistribution in Europe: A Cross-Sectional and Longitudinal Multilevel Analysis. *International Journal of Public Opinion Research*, 36(4), edae055.
<https://doi.org/10.1093/ijpor/edae055>

Velev, V. (2023). Replication of "Multilevel Models for the Analysis of Comparative Survey Data Common Problems and Some Solutions" with R. *SocArXiv*. <https://doi.org/10.31235/osf.io/phq56>

SKILLS

Programming: Python, R, Stata, L^AT_EX, Git, MS Excel, Java, HTML, CSS, SQL (PostgreSQL, MySQL), Tableau, Docker, JavaScript, TypeScript, C++

Languages: English (C2), German (mother tongue), Bulgarian (mother tongue), Dutch (elementary proficiency)

FURTHER EDUCATION & CERTIFICATES

Advanced Data and Information Literacy (ADILT) Certificate

October 2023 – September 2025

University of Konstanz

Konstanz, Germany

- Interdisciplinary certification program on data analysis, AI ethics, and digital responsibility
- Courses: Digitalization and Law, AI in Democracy

Google Data Analytics Professional Certificate

June 2021 – July 2021

Coursera

Online

- Introduction to R, SQL, MS Excel, and Tableau
- Application of technical knowledge in capstone project