Dmitry Pronin

Digital humanities researcher · Computational linguist · Python developer

HSE University · School of Philological Studies · Moscow, Russia

i@ddpronin.ru | ddpronin.github.io | Google Scholar | ORCID

Profile

I develop interpretable methods for analysing literary style, authorship, and text corpora. I work as a programmer and visiting lecturer at HSE University and am a core developer of SOCIOLIT and PROZIMETRON. My publications include a peer-reviewed article in Digital Scholarship in the Humanities and a textbook on digital philology. At DH2026, I presented research and chaired a session.

Research interests: interpretable stylometry; authorship attribution; digital literary studies; corpus linguistics; explainable NLP; large language models.

Digital Humanities 2026

Presenter

From Genes to Tokens: a GWAS-inspired Approach for Interpretable Stylometric Analysis, with Evgeny Kazartsev. Session S072, 30 July 2026, Room 107. Open preprint

Session chair

Session S026 on linked open data, historical information retrieval, and rumour discovery. 29 July 2026, Room 103.

Academic and research experience

HSE University | 2025–present

Programmer and visiting lecturer · School of Philological Studies, Faculty of Humanities · Moscow

- Develop statistical and NLP methods for literary-text analysis, with an emphasis on interpretability, validation, and reproducibility.
- Core developer of SOCIOLIT and PROZIMETRON: analytical modules, corpus pipelines, APIs, and interfaces for researchers.
- Co-instructor of the BA course Literature and Digital Humanities in 2025/26.

Bauman Moscow State Technical University | 2024–present

Team lead and Python developer · Research and Education Center · Moscow

- Lead research-software development, testing, and integration of Python components.
- Develop statistical text-analysis methods and large-language-model applications.

Education

HSE University | 2025–present

MA programme in Digital Methods in the Humanities · Moscow.

Bauman Moscow State Technical University | 2021–2025

BSc in Control in Technical Systems, specialising in Intelligent Control and Data Processing Systems · Moscow.

Selected research projects

SOCIOLIT

A Russian literary analytics platform. Contributed a corpus of 755 works by 180 authors spanning the eighteenth to twenty-first centuries. SOCIOLIT platform

PROZIMETRON

Comparative analysis of verse and prose: accentual annotation, rhythm modelling, probabilistic comparison, and multilingual workflows.

Interpretable Delta measures

Rank-Turbulence and Jensen–Shannon Delta, token-level decompositions, and stability validation. Code and notebooks

GWAS-inspired stylometry

Logistic-regression association testing with multiple-comparison correction to study lexical associations with authorship. Code and notebooks

Selected publications

Pronin, D., & Kazartsev, E. (2026). Rank-Turbulence Delta and interpretable approaches to stylometric Delta measures. Digital Scholarship in the Humanities, 41(3), 1616–1630. DOI 10.1093/llc/fqag072

Pronin, D., & Kazartsev, E. (2026). From Genes to Tokens: a GWAS-inspired Approach for Interpretable Stylometric Analysis. Short paper presented at DH2026; preprint arXiv:2606.09543. Open preprint

Kazartsev, E., Pronin, D., & Kirina, M. (forthcoming). Digital Literary Studies at the New Russian Computational Platform SOCIOLIT. In Proceedings of the 37th Conference of Open Innovations Association FRUCT. IEEE. HSE publication record

Kazartsev, E., & Pronin, D. (2024). Fundamentals of Digital Philology: Methods and Principles of Computer-Assisted Text Analysis. St Petersburg: Politekhnika. 178 pp. In Russian.

Technical profile

Python, SQL, Git; R for research; pandas, NumPy, SciPy, scikit-learn, statsmodels; Django and REST APIs. Stylometry, corpus processing, statistical modelling, clustering, data visualisation, and LLM applications.

Science communication and awards

Science communicator for System Block. Winner of BEST Hack 2023 in the Data Science track. Writing on System Block