

Дмитрий Пронин

Исследователь в digital humanities · Компьютерный лингвист · Python-разработчик

НИУ ВШЭ · Школа филологических наук · Москва, Россия

i@ddpronin.ru | ddpronin.github.io | [Google Scholar](#) | [ORCID](#)

Профиль

Разрабатываю интерпретируемые методы анализа литературного стиля, авторства и текстовых корпусов. Работаю программистом и приглашённым преподавателем НИУ ВШЭ, участвую в разработке СОЦИОЛИТА и ПРОЗИМЕТРОНа. Соавтор рецензируемой статьи в Digital Scholarship in the Humanities и учебника по цифровой филологии. На DH2026 выступил с докладом и председательствовал на секции.

Научные интересы: интерпретируемая стилометрия; атрибуция авторства; цифровое литературоведение; корпусная лингвистика; объяснимая обработка естественного языка; большие языковые модели.

Digital Humanities 2026

Докладчик

From Genes to Tokens: a GWAS-inspired Approach for Interpretable Stylometric Analysis, совместно с Евгением Казарцевым. Секция S072, 30 июля 2026 года, аудитория 107. [Открытый препринт](#)

Председатель секции

Секция S026 о связанных открытых данных, поиске исторической информации и выявлении слухов. 29 июля 2026 года, аудитория 103.

Академический и исследовательский опыт

Национальный исследовательский университет «Высшая школа экономики» | 2025—настоящее время

Программист и приглашённый преподаватель · Школа филологических наук факультета гуманитарных наук · Москва

- Разрабатываю статистические методы и инструменты обработки естественного языка для анализа литературных текстов с акцентом на интерпретируемость, проверку результатов и воспроизводимость.
- Один из основных разработчиков СОЦИОЛИТА и ПРОЗИМЕТРОНа: аналитические модули, обработка корпусов, API и интерфейсы для исследователей.
- Участвую в преподавании бакалаврского курса «Литература и цифровая гуманитаристика» в 2025/26 учебном году.

МГТУ им. Н. Э. Баумана | 2024—настоящее время

Руководитель команды и Python-разработчик · Научно-образовательный центр · Москва

- Руководжу разработкой исследовательского ПО, тестированием и интеграцией компонентов на Python.
- Разрабатываю статистические методы анализа текста и приложения на основе больших языковых моделей.

Образование

НИУ ВШЭ | 2025—настоящее время

Магистратура «Цифровые методы в гуманитарных науках» · Москва.

МГТУ им. Н. Э. Баумана | 2021–2025

Бакалавриат по направлению «Управление в технических системах», профиль «Интеллектуальные системы управления и обработки данных» · Москва.

Избранные исследовательские проекты

СОЦИОЛИТ

Платформа анализа русской литературы. Участвовал в создании корпуса из 755 произведений 180 авторов XVIII–XXI веков. [Платформа СОЦИОЛИТ](#)

ПРОЗИМЕТРОН

Сравнительный анализ стиха и прозы: акцентная разметка, моделирование ритма, вероятностное сопоставление и работа с разными языками.

Интерпретируемые меры Delta

Rank-Turbulence и Jensen–Shannon Delta, разложение результатов на вклады токенов и проверка устойчивости интерпретаций. [Код и ноутбуки](#)

Стилометрия по аналогии с GWAS

Проверка связей между лексическими признаками и авторством с помощью логистической регрессии и поправки на множественные сравнения. [Код и ноутбуки](#)

Избранные публикации

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. Краткий доклад на DH2026; препринт arXiv:2606.09543. [Открытый препринт](#)

Kazartsev, E., Pronin, D., & Kirina, M. (в печати). Digital Literary Studies at the New Russian Computational Platform SOCIOLIT. In *Proceedings of the 37th Conference of Open Innovations Association FRUCT*. IEEE. [Запись в библиографии ВШЭ](#)

Казарцев Е. В., Пронин Д. Д. (2024). Основы цифровой филологии: методы и принципы компьютерного анализа текста. Санкт-Петербург: Политехника. 178 с.

Методы и технологии

Python, SQL, Git; R для исследовательских задач; pandas, NumPy, SciPy, scikit-learn, statsmodels; Django и REST API. Стилометрия, обработка корпусов, статистическое моделирование, кластеризация, визуализация данных и приложения на основе больших языковых моделей.

Научная коммуникация и награды

Автор «Системного Блока». Победитель BEST Hack 2023 в направлении Data Science. [Тексты в «Системном Блоке»](#)