Artem Matveev

matfu21@yandex.ru
LinkedIn
GitHub
Google Scholar
Moscow, Russia

About Me

Research scientist and engineer with 4+ years of experience in recommender systems, with accepted papers at KDD, SIGIR, and RecSys. Currently a Senior Research Scientist at AI VK. B.Sc. from Lomonosov Moscow State University; M.Sc. candidate at the Higher School of Economics.

Research Interests

Information retrieval, conversational recommendation, large language models, generative recommendation, in-text reasoning for recommendation, reinforcement learning, and the mathematical foundations of deep learning.

Work Experience

Mar 2026 – Present **Senior Research Scientist, AI VK**

Research team led by Prof. Dr. Alexander D'yakonov, formerly world's #1 rank on Kaggle.

- Worked on long-term RL optimization for generative retrieval in recommender systems; first author of *Long-Term Optimization for Large-Scale Generative Retrieval with Off-Policy REINFORCE* (KDD 2026).
- Researching PLUM-style adaptation of LLMs to generative recommendation via Semantic IDs.
- Contributed to online A/B tests in VK Video, delivering $+x\%$ Total Viewing Time (TVT).
- Made major contributions to the internal research framework, speeding up experiment iteration several-fold.

Jul 2025 – Mar 2026 **RecSys R&D Team Lead / Senior Research Engineer, Yandex**

- Led a three-person R&D team responsible for pretraining generative recommendation models, internal training framework development, and inference optimization.
- Worked on pretraining a OneRec-style generative recommender that replaces the entire cascade pipeline (candidate generation, ranking, and re-ranking) with a single end-to-end neural network.
- Introduced FSDP (Fully Sharded Data Parallel) training, scaling recommender transformers up to 7 billion parameters with improved model quality.
- Built the training framework and infrastructure for large-scale generative recommender models.
- Worked on adapting sub-quadratic attention mechanisms for generative retrieval in recommender systems (paper).

May 2023 – Jul 2025 **RecSys Research Engineer, Yandex**

- Investigated graph neural networks for item representation in recommender systems (talk).
- Developed a transformer-based candidate generation model for Yandex Market and conducted A/B tests.
- Co-developed Argus, a billion-parameter generative recommender transformer; achieved +2.26% Total Listening Time and +6.37% item like rate at Yandex Music in online A/B tests; later adapted across Yandex products (e.g., e-commerce +1.7% GMV, food-tech +2.1% GMV).
- Contributed to Argus adaptation for Yandex Ads ranking, delivering +2% revenue uplift in online A/B tests (talk).
- Mentored two interns, who have since become a Senior RecSys Researcher and a core developer at YandexGPT.

Feb 2022 – Jul 2022 **Backend Developer Intern, Yandex**

- C++ backend development: optimized runtime performance of search algorithms and data-preparation pipelines.
- Implemented new ML features and conducted A/B tests.

Education

Sep 2024 – Present **Higher School of Economics (HSE)**, Moscow, Russia
Master's Program in Modern Computer Science.
Thesis: *Long-Term Optimization for Large-Scale Generative Retrieval.*

Sep 2024 – Present **Yandex School of Data Analysis (YSDA)**, Moscow, Russia
Selective postgraduate program in machine learning, deep learning, and applied mathematics.

Sep 2020 – Jun 2024 **Lomonosov Moscow State University (MSU)**, Faculty of CMC, Moscow, Russia
B.Sc. in Applied Mathematics and Computer Science.
Thesis: *Graph Neural Networks for Candidate Retrieval in E-Commerce.*

Selected Publications

KDD 2026 Long-Term Optimization for Large-Scale Generative Retrieval with Off-Policy REINFORCE, *5th Workshop on End-to-End Customer Journey Optimization*
Artem Matveev, Sergei Makeev, Aleksei Krasilnikov, Vladimir Baikalov, Sergei Liamaev, Kirill Khrylchenko

Scaling Recommender Transformers to One Billion Parameters
Kirill Khrylchenko, **Artem Matveev**, Sergei Makeev, Vladimir Baikalov

SIGIR 2026 Gated Bidirectional Linear Attention for Generative Retrieval
Artem Matveev, Vladislav Tytskiy, Sergei Makeev, Sergei Liamaev

RecSys 2025 Correcting the LogQ Correction: Revisiting Sampled Softmax for Large-Scale Retrieval
Kirill Khrylchenko, Vladimir Baikalov, Sergei Makeev, **Artem Matveev**, Sergei Liamaev

arXiv preprint Embedding Items at Scale: Comparing GNN-Based and ID-Based Item Embeddings in the Yandex Ecosystem
Sergei Makeev, **Artem Matveev**, Vladimir Baikalov, Kirill Khrylchenko

Sona Technical Report
Alexandr Udeneev, Aleksei Krasilnikov, Alexey Nadtochiy, Andrey Semenov, Andrey Tsyrkunov, Anna Krivonos, Anna Lipkina, **Artem Matveev**, et al. (34 authors)

Teaching

2024 – 2026 Teaching Assistant, Yandex School of Data Analysis (YSDA)
DeepRecSys Course. GitHub (140+ stars).
- Seminars on evaluation metrics, neural RecSys framework design, neural ranking, and reinforcement learning.
- Authored homework assignments for the course.

2025 – 2026 Teaching Assistant, Higher School of Economics
Recommender Systems Course. GitHub (90+ stars).

2025 Teaching Assistant, Higher School of Economics
Deep Learning Course. GitHub (100+ stars).
- Seminar on neural ranking.

Talks

2025 *Do We Need Graph Neural Networks to Encode Items in Recommendations?*, Yandex For ML
YouTube video.

Competitions

2023 Kaggle "ICR — Identifying Age-Related Conditions"
Silver Medalist.

2022 ACM ICPC, 2022 ICPC Northern Eurasia Finals
Diploma of the 3rd degree.

2020 ACM ICPC, Final of Northern Eurasia 2020
Diploma of the 2nd degree.

Technical Skills

Languages Current: Python, Triton, CUDA. Past: C++, Assembly.
Frameworks PyTorch.
ML Engineering Full production cycle: data preparation, experimentation, model export & conversion, performance benchmarking, Triton Inference Server deployment, real-time history processing.

Professional Service

2026 UMAP 2026 — PC member, Industry Track

Patents

Dec 2025 Application for invention 2025136959, Federal Service for Intellectual Property
Yandex filed a patent application for Argus (see *Scaling Recommender Transformers to One Billion Parameters*).