

Dmitry Balobin

Senior NLP / LLM Engineer

Open to opportunities in the UK and Serbia
Available for relocation

hanaan@internet.ru
linkedin.com/in/d0rj
huggingface.co/d0rj

Summary

Senior NLP / LLM Engineer with experience building, fine-tuning and deploying large language model systems across search, banking, maps and conversational AI products. Strong background in LLM training pipelines, RLHF, SFT, RAG, summarisation, AI agents and production ML infrastructure. Experienced in research experimentation, applied NLP product development and technical interviewing.

Work Experience

Yandex Jan 2026 – Present

Senior ML Engineer, Alice on Search

- Developing, maintaining and training custom LLMs across multiple GPU clusters for search-integrated conversational AI use cases.
- Supporting large-scale model training workflows, experimentation and production operations for LLM-based systems.

Raiffeisen Bank Sep 2025 – Jan 2026

Senior AI Engineer, Corporate & Investment Banking

- Developed AI copilots and autonomous agents for online trading workflows in a CIB environment.

2GIS Jan 2024 – Sep 2025

Senior NLP Data Scientist

- Integrated LLM capabilities into the user-generated content section of the application, including review summarisation and context-aware summarisation for online search.
- Contributed to production NLP features for search and UGC scenarios, bridging model quality, product requirements and deployment constraints.
- Conducted technical interviews across NLP and adjacent ML teams, assessing candidates in NLP, ML engineering and applied data science.

Sber Jun 2022 – Dec 2023

ML Engineer, NLP B2C Team

- Developed the RLHF training stage for the GigaChat LLM and ran research experiments to improve model behaviour and alignment.
- Improved the SFT training pipeline for GigaChat and adapted internal models to product-specific requirements.
- Built pilot integrations of GigaChat into B2C and internal banking products.
- Developed retrieval components for RAG-based business use cases powered by GigaChat.

Sber Dec 2021 – May 2022

Data Scientist Intern, NLP B2C Team

- Worked on speech analytics for telephone calls, including clustering and visualisation of call data.

Education

MIPT, School of Applied Mathematics and Informatics <i>PhD Candidate, Computer Science and Informatics</i>	2025 – 2028
MIPT, School of Applied Mathematics and Informatics <i>MSc, Computer Science; Machine Learning and Data Analysis</i>	2023 – 2025
MIREA – Russian Technological University <i>BSc, Applied Mathematics; Data Analysis</i>	2019 – 2023

Publication

D. Y. Balobin, O. D. Somov. Improved Detection Methods for Generated Texts	2025
---	------

Technical Skills

Frameworks & Tools: PyTorch, Transformers, DeepSpeed, LangChain, PydanticAI, Langfuse, Kubernetes, FastAPI.