Semen Korobko

semenkorobko2@gmail.com · Telegram: @UncleSema · unclesema.top

Research engineer focused on predictive representations, multimodal reasoning, and planning. Leads tool-use post-training for GigaChat Omni; previously established VLM development at Yandex Autonomous Vehicles and built the team's full model release pipeline.

Interests: world models and latent dynamics; predictive and self-supervised representation learning; reasoning, planning, and decision-making; model-based reinforcement learning.

Experience

Sber · GigaChat Omni

Lead, Reasoning and Tool Use · Sep 2026–present

- Lead the new tool-use post-training effort for GigaChat Omni as its sole technical contributor, covering web search, thinking with images, calculator use, and MCP tools.
- Assembled cold-start training datasets and enabled tool execution within model evaluations.
- Developed an RL training recipe using **GRPO/CISPO** with judge-based and algorithmic rewards; initial experiments are ongoing.

Yandex Autonomous Vehicles · ML R&D

Software Engineer · Jul 2024–Sep 2026

- Established the team's VLM effort and built the full model release pipeline: raw-log mining, video/LiDAR preprocessing, training, evaluation, simulator testing, and vehicle deployment.
- Developed and deployed a VLM-based planner on a vehicle, producing **50-step trajectories over a 5 s horizon with 200 ms inference latency**.
- The planner passed test-track safety and maneuver tests, including cone avoidance, while meeting GPU-memory requirements.
- Built data-mining and annotation tools for archives containing hundreds of petabytes of raw logs; ran training and inference workloads on hundreds of GPUs.

T-Bank · Core Technologies

Software Engineer · Dec 2021–Jul 2024

- Developed internal ML services and software tools for bank employees.

Research Projects

JEPA-Based Trajectory Planning

Yandex Autonomous Vehicles · Research prototype

- Investigated planning through sequential prediction of future latent states; built a compact planner with a **ViT-B backbone and a ~20M-parameter latent predictor**.
- Outperformed a substantially larger autoregressive planning baseline on selected driving scenarios.

Inference-Time Compute Allocation in Multimodal Models

ITMO University / Sber GigaChat R&D · Master's thesis, ongoing

- Extended latent visual reasoning to variable-length latent and textual trajectories. Implemented up-front budget prediction and learned stop heads as alternative length-control strategies.
- Reproduced LVR, HyLaR, and LUT; evaluating the extension on Qwen3.5 and Qwen3.8 (27B).

VLM-Based Error Mining and Rare-Object Data Selection

ITMO University / Yandex Autonomous Vehicles · Bachelor's thesis, 2025

- Built a VLM-based pipeline to mine perception errors and select data for rare-object detector fine-tuning. Compared Yandex VLM, Qwen2.5-VL, and InternVL2.5 on a **50,000-image evaluation set**.
- Prompt tuning reduced object-counting MSE by **31% (15.55 → 10.73)** relative to the same Yandex VLM 35B baseline without adaptation.

Education

ITMO University · M.Sc. in Applied Mathematics and Computer Science

2025–present · Expected graduation: 2027

Deep Learning and Generative Artificial Intelligence. Integrated Master's-to-PhD track.

ITMO University · B.Sc. in Applied Mathematics and Computer Science

2021–2025

Computer Science and Programming.

Teaching

T-Bank · Algorithms and Data Structures

Course Creator · Spring 2024, 2025, and 2026

Created the course for university students across three spring offerings.

ITMO University · Teaching Assistant (Seminars)

Courses led by Georgiy Korneev

Introduction to Programming — Fall 2024; Programming Paradigms — Spring 2025.

Publications & Presentations

N. D. Prigodich and **S. S. Korobko**. Application of Software Methods for Automated Processing of Sources of Personal Origin. *Historical Informatics*, no. 1, pp. 1–9, 2023. DOI: 10.7256/2585-7797.2023.1.40376. (In Russian.)

S. S. Korobko. Preprocessing Digital Copies of Historical Sources. 12th Congress of Young Scientists, ITMO University, 2023. Talk and published abstract. (In Russian.)