

Lev Brelovskiy

Data Scientist | ML Engineer

+7 909 570 42 09 | Barbarng0@gmail.com | telegram: [@LevBrell](https://t.me/LevBrell) | github.com/Golden-dreamer | brellev.com |
Petrozavodsk, Russia | Open to relocation | Remote OK | Requires work permit

ML Engineer with experience developing production systems for document analysis. Specializing in OCR, NLP, and Document AI: building end-to-end pipelines from model training to **FastAPI** services and **Docker** deployment.

Experience

May 2026 – present: developing a GraphRAG system (see Projects)

LIGRES

Data Scientist

March 2024 – February 2026

Moscow, remote

- Legal documents with handwritten text were processed manually. Task: automate document processing. Built a production system for automated processing (OCR + NER + classification) on **FastAPI** + **Docker** + **RabbitMQ** + **S3** — 10,000+ documents/day.
- Implemented CV preprocessing for OCR quality: increased the pool of documents processed automatically by improving data extraction from low-quality scans. **OpenCV**, **torchvision**. Experiments — **ClearML**, **Hydra**.
- Result: 80% of documents processed without human intervention (previously — all manual). Reduced manual review time for lawyers by 2 hours/day, team processed 300 more documents.

July 2023 – February 2024: self-studying ML/CV, pet project on segmentation, hackathon (see Projects)

Military Innovative Technopolis ERA

Data Scientist

July 2022 – June 2023

Anapa

- Internal OCR prototype for paper documents: classification and field extraction on low-quality scans and photocopies. **Tesseract**, **EasyOCR**, **OpenCV**, **Python**.
- Deployed the prototype on **FastAPI** + **PostgreSQL** with monitoring: low-confidence documents routed to manual review, accuracy tracked via **PostgreSQL**.

Skills

ML/DL: **PyTorch**, **torchvision**, **Hugging Face Transformers**, **PEFT/LoRA**, **Scikit-learn**

NLP & CV: **Tesseract**, **EasyOCR**, **spaCy**, **NLTK**, **OpenCV**

MLOps & Infra: **FastAPI**, **Docker**, **RabbitMQ**, **S3**, **vLLM**, **GitLab CI**, **ClearML**

Other: **Python** (NumPy, Pandas, asyncio), **PostgreSQL**, **Neo4j**, **Linux** (Ubuntu/Debian), **Nginx** | English — B2

Projects & Achievements

Knowledge Graph QA System (GraphRAG)

2026

Built a GraphRAG system for answering questions over structured data: parsed data into graph database **Neo4j** (212k nodes, 899k edges), created a web UI for natural-language queries — LLM generates Cypher queries and returns answers grounded in the knowledge graph. Stack: **FastAPI**, **JavaScript**, **Docker**, **LLM**. github.com/Golden-dreamer/animeGraphRAG

Railway Segmentation

Oct. 2023

Railway track segmentation using U-Net. Dataset: 89 labeled images, expanded with augmentation to 984. **PyTorch**, **OpenCV**. github.com/Golden-dreamer/railway-segmentation

Hackathon

2023

3rd place, All-Russian Hackathon — semantic segmentation of urban scenes.

Education

Petrozavodsk State University

2018 – 2022

B.Sc., Applied Mathematics and Computer Science

Thesis: real estate price prediction using ML (regression models, feature engineering on listing data).