

DIAS KHALNIYASOV

+7 (708) 773-75-82 ♦ Almaty, Kazakhstan

[Telegram](#) ♦ diaskhalniyasov@gmail.com ♦ [GitHub](#) ♦ [DevPost](#) ♦ [LinkedIn](#)

OVERVIEW

AI Engineer with 4+ years of experience with a proven track record of delivering impact in AI startups, particularly in AI agents and large language models (LLMs). Multiple-time hackathon winner

SKILLS

Technologies Python, TypeScript, GCP, LangChain, LangGraph, FAISS, RAG, Git, FastAPI, Docker, PyTorch
Communication English (C1), Italian (B1), Russian (native), Kazakh (B1)

EXPERIENCE

Software Engineer Nov 2025 - Now
[String AI](#) - End To End Web Extraction Platform *Remote, US*

- Build evaluation for AI Parsing Agents

AI Research Engineer Nov 2024 - Oct 2025
[Zencoder](#) - The AI Coding Agent startup [Crunchbase profile](#) *Remote, US*

- Build and evaluate AI Agents for software development code generation
- Main contribution to E2E Agent (AI QA testing tool), Web Development Agent (full SDLC agent)

AI Engineer - R&D team Oct 2023 - Sep 2024
SOAX - British startup that raised \$6.3 million in 2019 [Crunchbase profile](#) *Remote, UK*

- Developed a multi-agent system, coordinating multiple agents for AI Web Scraper to auto-parse websites and extract structured information [conference talk \(in Russian\)](#); [PH release](#)

Data Scientist Dec 2021 - June 2023
Domusware Srl Roma - IT team of Yard Reaas SpA [Crunchbase profile](#) *Rome, Italy*

- Constructed an **OCR** model with a Word Error Rate of **0.2**, **80% accuracy** per word for CAPTCHA recognition.
- Build a pipeline to match properties (SentenceBERT + PySpark + FAISS) and APIs (Pydantic + FastAPI)

Junior Data Scientist Jun 2020 - Mar 2021
Galamat Technologies - speech recognition and synthesis team *Almaty, Kazakhstan*

- Developed and deployed a data collection service, collecting over **10,000** hours of annotated speech in 3 months.
- Trained text-to-speech models in Russian and Kazakh languages as research baselines.

HONORS AND AWARDS

- [Siemens Hackathon Tech for Sustainability - 1st place](#) - Interactive Energy Data Map with AI chat assistant
- [HackUPC HP Supply Chain Optimization - 1st Place](#) - BI solution with predictive model for inventory for next 13 weeks, including AI Assistant and scenario simulator. Won among 600 participants
- [Halyk RAG Challenge - Best Advanced Retrieval](#) - Multimodal RAG (PDF, Excel supported) with hybrid search

EDUCATION

MSc Data Science, Sapienza University of Rome 2020 - 2023
B.S. in Mathematical and Computer Modeling, Kazakh-British Technical University 2016 - 2020

EXTRA-CURRICULAR ACTIVITIES

- [Speaker on "Beetech 2025" - AI track](#) - Topic: "How to parse websites with LLM agents"
- [Speaker on the Forum "Gem of Data Science"](#) - Topic:- Intelligent chat bots.