Igor Buyanov

Email: buyanov.igor.o@yandex.ru
Github: github.com/Astromis/
LinkedIn: https://www.linkedin.com/in/igor-buyanov/
Personal site: astromis.ru

Executive Summary

Principal NLP/ML Engineer with 6+ years in production conversational AI for voice channels. Built 10+ NER/Intent models (NER F1-macro $\approx 90\%$, 400 intents F1-macro $> 75\%$) for a telecom chatbot serving millions. Led LLM-based dialogue quality evaluation at 120k+ sessions/day, deployed MLOps/labeling workflows (42% faster). Experience in scaling ML systems end-to-end (Triton, ONNX).

Experience

- **MWS AI (Former MTS AI)** Remote
  Principal developer, NLP team (Full-time) *August 2024 - Present*
  - **Chatbot NLU models development**: maintained NLU components for a voice-channel telecom chatbot used by millions; shipped 10+ NER/Intent models (NER F1-macro $\approx 90\%$; 400 intents with F1-macro $> 75\%$).
  - **Chatbot quality assessment pipeline**: built an LLM-based dialogue quality estimator enabling automatic evaluation of $> 120,000$ sessions/day, improving QA coverage and feedback loops.
  - **Mentoring**: led a junior developer's project that introduced phonological augmentation, helped with onboarding for newcomers
- **MTS AI** Remote
  Senior developer, NLP team (Full-time) *March 2023 - August 2024*
  - **Data programming**: applied a weak supervision to filter and annotate business defined geographic entities. The model achieved a $\approx 95\%$ F1-macro score.
  - **Model inference optimization**: deployed a model optimization pipeline (ONNX), reducing inference latency by 20% across key services.
  - **Utterance end detection (endpointing) for voice chatbot**: Researched and deployed utterance endpointing for voice bots (don't-cut detection), achieving $\approx 70\%$ F1-macro on target scenarios.
- **MTS AI** Remote
  Senior Developer, Labeling team (Full-time) *June 2021 - March 2023*
  - **End-to-end ML labeling workflow**: introduced an end-to-end labeling workflow (Label Studio + pre-annotation + QC library), reducing labeling errors and speeding throughput by 42% company-wide.
  - **MLOps platform**: deployed MLOps solution and introduced it into 4 ML teams' workflow. Integrated WandB with data storage and designed a Python kit to load and save multimodal data. Participated in the designing of the internal machine learning platform.
  - **Crowdsourcing**: developed a methodology to use Yandex.Toloka (analogous to Amazon Mechanical Turk), handled automation through the provided API. Three Voice datasets were collected using this methodology.
- **MTS AI** Remote
  Middle developer, NLP team (Full-time) *June 2020 - June 2021*
  - **Token-based regex engine with NLP ops (C++ core, Python bindings)**: developed a token-based regex engine with NLP operations (C++ with Python bindings) for production text processing.
  - **Adapting academic research to production**: searched and found a working method for noise reduction in datasets, which was actively used by the NLP team. The result of the work also was published in a corporate blog.
  - **Automatic dataset construction**: created a methodology for building the dataset with a hierarchical labeling structure from Wikipedia automatically. The dataset was used in a comparative test of NLP models. The methodology was shared in the corporate blog.

- **MTS AI** Remote
  Junior developer, NLP team (Full-time) *March 2019 - June 2020*
  - **Building text classifiers**: built a classifier for news that was used in Voice Assistant and for customer support department.
  - **Code optimization**: optimized the code of the stress module for speech synthesis, which allowed the module to work with a delay of 0.3 seconds, as well as code of number normalizer.

Projects and Volunteering

- **Kitoboy - a suicide risk prevention ecosystem (NLP)**: The system is designed to assess the suicidal risk of the user by his social networks' posts in order to make a decision about the intervention. Founder, lead of 7 developers team, pitcher, and product manager of this project. Released the datasets, models and web-application. Planned that the system will allow volunteers and psychologists to process about 50 social account per week.
- **Psytechlab - the RnD team with the focus on solving mental health problems with machine learning**: Founder and head of the team. Released the Kitoboy system, dataset translating pipeline. Support university students by proving the team related topics for their theses.

Publications

- **Paper: The dataset for presuicidal signals detection in text and its analysis**: The paper says about the dataset for presuicidal signal detection in Russian posts from social media. To the best of our knowledge, it is the first dataset of such type for this language. We develop a collection methodology and conduct a linguistic analysis of the completed dataset. We also build a classification baseline with machine learning models to solve the detection task. (June '22)
- **Paper: Transferring Natural Language Datasets Between Languages Using Large Language Models for Modern Decision Support and Sci-Tech Analytical Systems**: In this work, we investigated how one can use large language models (LLMs) to transfer the dataset and its annotation from one language to another. This is crucial since sharing knowledge between different languages could boost certain underresourced directions in the target language, saving lots of effort in data annotation or quick prototyping. We experiment with English and Russian pairs, translating the DEFT (Definition Extraction from Texts) corpus. (April '25)
- **Paper: The methodology of constructing the large-scale dataset for detecting presuicidal and anti-suicidal signals in social media texts in Russian**: In this work we have presented a detailed methodology of building the dataset for detecting texts that describe presuicidal and anti-suicidal signals. This methodology describes the process of instruction and class table creation, the process of annotation, verification and post-annotation correction. Guiding by this methodology, we collect and annotate a large-scale Russian dataset with more than 50 thousand texts from social media. (November '25)
- **Corporate and personal posts on Habr.ru**: I have 18 posts on Habr, which is the largest IT blog platform in Russia. The posts are about machine learning that I wrote for a corporate blog as results of my work or my personal results that I achieved in my research and projects.

Education

- **Moscow Technical University of Communication and Informatics** Moscow, Russia
  Specialist - Information Security of Telecommunication Systems *September 2013 - December 2018*
- **FRC CSC RAS** Mosc
  PhD Student - Thesis: Methods for Suicide Prevention In Social Media (in progress) *September 2021 - Sept*

Skills Summary

- **Languages**: Python, Bash
- **ML**: PyTorch, Scikit-learn, Transformers
- **LLM**: LangChain, Llama.cpp, Ollama, Unsloth
- **Data & MLOps**: Docker, Triton Inference Server, ONNX, ClearML,W&B, Label Studio
- **Tools**: Docker, Git, ClearML, Ollama, Llama.cpp, PostgreSQL, SQLite
- **Databases**: PostgreSQL, SQLite
- **Platforms**: Linux
- **Soft Skills**: Tutoring, Mentoring, Writing, Public Speaking