

Ширяев Антон Дмитриевич

Deep Learning Engineer (CV, LLM & VLM)

📍 Владивосток (GMT +10)
📅 22.03.1990
☎ 8 914 971 10 73
✉ med-phisiker@yandex.ru
📧 @med_phisiker

🔗 <https://github.com/medphisiker>
🔗 <https://gitlab.com/users/med.phisiker>
🔗 <https://www.kaggle.com/medphisiker>
🔗 <https://medium.com/@med.phisiker>
🔗 https://leetcode.com/u/med_phisiker/

О СЕБЕ

Разрабатываю системы компьютерного зрения на нейронных сетях для детекции, сегментации и трекинга объектов в режиме реального времени (опыт 3 года). Последние полгода занимаюсь разработкой сервисов с использованием VLM и LLM. Осуществлял полный цикл ML разработки от сбора данных до действующего ML-сервиса с микросервисной архитектурой.

Имею опыт научных исследований в Российской академии наук (8 лет). Интересны проекты с LLM и VLM и разработка evaluation frameworks.

SKILLS

- Алгоритмы и структуры данных, успешно прошел собеседование на белой доске в Яндекс.Дзен (VK).
- Цифровая обработка акустических сигналов (спектральный и корреляционный анализ).
- Опыт научной работы (исследований и подготовки научных публикаций в области акустики)
- Английский язык - B1

НАУЧНЫЕ ПУБЛИКАЦИИ

Мои профили в индексируемых базах рецензируемых научных статей (акустика, по ML публикаций нет):

- E-library ([ссылка](#))
- Scopus ([ссылка](#))
- Web of science ([ссылка](#))

TECH SKILLS

Programming: Python(venv, conda, poetry, uv), Fast API

Deep Learning: PyTorch, Lightning, Hugging Face Transformers

LLM & Multimodal: vLLM, Unsloth, LangChain, LangGraph, Qdrant VLMEvalKit, Arize Phoenix

Machine Learning: LightAutoML, CatBoost, XGBoost, Optuna, scikit-learn

MLOps & Monitoring: MLFlow, Weights & Biases, TensorBoard, Prefect, ONNX, TensorRT

DevOps & Infrastructure: Linux, tmux, Git / GitHub / GitLab, Git Submodules, Commitizen, IDE VS code, Docker, NVIDIA Container Toolkit, Docker Compose, uv workspaces

Data & Tools: NumPy, SciPy, Pandas, Matplotlib, Seaborn, Streamlit, OpenCV, PIL, kornia, ablumentations, ultralytics, CVAT, Label Studio, Datumaro, FiftyOne, MinIO S3, RabbitMQ

ОПЫТ РАБОТЫ

10.2021 – Present Разработчик систем компьютерного зрения АО Кашалот

- Разрабатывал микросервисный сервис с использованием мультимодальной модели Qwen2.5-VL-7B для классификации российских документов и извлечения структурированных данных. Использовал vLLM (serving), RabbitMQ (очереди), MinIO (хранилище) и Arize Phoenix (управление промптами). Реализовал human-in-the-loop: сотрудники корректируют предсказания, данные агрегируются для будущего дообучения. Архитектура предусматривает разделение на контейнеры dev (быстрые эксперименты через uv workspaces и git submodules) и prod (воспроизводимость, фиксированные версии).
- разработал систему компьютерного зрения для телеуправляемого необитаемого подводного аппарата (ТНПА) (платформа nVidia Jetson ARM64) для обнаружения представителей морской фауны в интересах мониторинга марикультуры. На ВЭФ 2022 года был представлен наш аппарат, привожу ссылки на материалы ([демо](#), [СМИ](#), [Первый канал](#)).
- разработал систему компьютерного зрения для автоматизации конвейерной сортировки рыбоперерабатывающего завода с помощью дельта-робота (платформа AMD64 с Nvidia GPU Linux Ubuntu Server): сегментация, SORT-трекинг, оценка длины рыбы ([демо](#), [демо с роботом](#)).

07.2012 – 02.2022 Научный сотрудник лаборатории акустической томографии ТОИ ДВО РАН

- за время работы в институте мной были получены компетенции в области применения Python для задач цифровой обработки акустических сигналов и статистического анализа. Есть опыт подготовки статей в рецензируемые научные журналы и выполнении научных грантов.
- подробнее вы можете прочитать обо мне на [сайте ТОИ ДВО РАН](#).

ОТКРЫТЫЕ ПРОЕКТЫ

Пет проект, Public Lifelong Learning Assistant — Платформа для непрерывного обучения и подготовки к собеседованиям по ML/DL и алгоритмам https://github.com/Lifelong-Learning-Assistant/lifelong_learning_assistant

Микросервисная платформа (FastAPI, LangGraph и Docker) для интерактивной подготовки к техническим собеседованиям с LLM-ассистентом. Система включает: LLM-агента с WebSocket-поддержкой (LangGraph), гибридный RAG-поиск, генератор квизов из Markdown, Docker-песочницу для запуска и тестирования кода по алгоритмам на python и React-интерфейс в стиле Cyberpunk. Архитектура использует внешние Docker-сети и сетевые алиасы для надёжной DEV/PROD-сборки. PROD образы публикуются в GitHub Container Registry. Написаны интеграционные тесты для ключевых пользовательских сценариев (Quiz, RAG, Chit-chat), запускаемые внутри контейнеров.

Компоненты: гибридный RAG-поиск и генератор квизов из Markdown были реализованы не мной.

Пет проект, Public VLMHyperBench — Open-source фреймворк для оценки возможностей Vision Language Models (VLM) в распознавании русскоязычных документов <https://github.com/VLMHyperBenchTeam/VLMHyperBench>

Инструмент позволяет сравнивать модели, которые нельзя запустить в одном окружении, в том числе на разных фреймворках инференса, оценивать промпты для разных типов документов и полей, а также легко добавлять свои данные, новые модели и метрики.

Проект выиграл грант «Yandex Open Source 2025» в треке «Машинное обучение». Руководил командой магистров 1 курса при разработке проекта.

Пет проект, Public Проект по аудио-визуальному распознаванию эмоций человека для сервисов онлайн видео звонков (Zoom, Skype и д.р.) https://gitlab.com/group_19200719

В рамках проекта я собрал решение, которое использует State of art модель (papers), показывающую лучший результат на датасете RAVDESS (dataset) на данный момент времени intermediate transformer fusion из статьи "Self-attention fusion for audiovisual emotion recognition with incomplete data" (arxiv).

Благодаря проекту занял 1-ое место в лидерборде (лидерборд) курса "MLOps и production подход к ML исследованиям 2.0" (курс) от ГазпромНефть на платформе ODS.

Пет проект, Public CareerRank, сервис для подбора подходящих друг другу вакансий и резюме https://github.com/medphisiker/maching_cv_and_vacancy

Прототип сервиса для подбора подходящих друг другу вакансий и резюме. Используются нейросетевые языковые модели (LM) для подбора подходящих друг другу вакансий и резюме на основе их текстовых описаний. Использовал FAISS для vector search.

Демо видео работы данного сервиса (демо).

ОБРАЗОВАНИЕ

9.2023 - Present Магистратура ИТМО «Искусственный интеллект» AI Talent hub ИТМО
В партнерстве с специалистами «Sber AI Lab», «СеверСталь», «NapoleonIT», «AIRI».

19.02.2016 Ученая степень к.ф.-м.н по специальности 01.04.06 – акустика. ТОИ ДВО РАН
Решением диссертационного совета ТОИ ДВО РАН от 19.02.2016 г. № 3-2016 присуждена ученая степень к.ф.-м.н по специальности 01.04.06 – акустика.

9.2012 - 6.2015 Аспирантура по специальности «Акустика» ТОИ ДВО РАН
Принимал участие в выполнении научно-исследовательских контрактов и грантов (ТОИ ДВО РАН)

9.2007 - 6.2012 Физик ДВФУ
Присуждена квалификация «Физик» по специальности «Медицинская физика», диплом с отличием

СОРЕВНОВАНИЯ

- Хакатон «Лидеры цифровой трансформации 2023», участвовал в составе команды Baseline Solution, выбрали задачу «Система видеодетекции вооруженных людей» (задача). Стали финалистами хакатона и вошли в топ10 участников (финал). Мы заняли 4-тое непризовое место (результат).
- Хакатон «AI Talent Hackathon 2023», выбрали задачу «Классификация характеристик товара на основе фотографии ценника» (задача) 2023 г. Видео с демо решения (демо). Ссылка на репозиторий (геро). Сертификат призера (сертификат).
В треке «Machine Learning in Retail» мы вошли в тройку лучших.
- Полный список соревнований/хакатонов/буткемпов доступен по ссылке (ссылка).

СЕРТИФИКАТЫ И ВЫСТУПЛЕНИЯ

- Грант «Yandex Open Source 2025» с проектом VLMHyperBench (Хабр).
- Выступление на Data Fest 2025, с проектом VLMHyperBench, секция Open Source (ODS, ВК видео).
- Выиграл именные стипендии в конкурсе Selectel Career Wave 2023 и 2024 годов в треке «Разработчики и IT-инженеры» (стипендия, 2024, 2023, сертификат 2024, сертификат 2023).
- Курс МФТИ «Глубокое обучение в NLP», получено удостоверение гос. образца (ссылка)
- RuCode 6.0 ФПМИ МФТИ — интенсивы по спортивному программированию, удостоверение гос. образца (сертификат).
- Полный список сертификатов (ссылка).
- Полный список выступлений (ссылка).