HEAD OF AI / CTO

Владимир Евграфов

Прикладные AI/ML-системы, production engineering и безопасность LLM-агентов



evgrafov.su | LinkedIn | GitHub | Google Scholar

ОПЫТ РАБОТЫ

Head of AI

Сентябрь 2026 - настоящее время | Москва | биомедицина

- Руковожу AI-стратегией, прикладными исследованиями и разработкой в сфере биомедицины.
- Проектирую LLM-системы и агентные архитектуры с RAG; развиваю практики оценки моделей и AI-инфраструктуру.
- Сочетаю практическую разработку с техническим руководством, выстраиваю процессы внедрения AI от экспериментов до эксплуатации.

Виасат Тех - Технический директор (CTO)

Январь 2022 - сентябрь 2026 | Москва | viju

- Отвечал за AI/ML, рекомендации, архитектуру, разработку и инфраструктуру; построил департамент из 35 человек и 7 команд.
- Перевел legacy-бэкенд на Go-микросервисы: платформа выдерживает 100-кратный рост нагрузки при четырехкратном снижении потребления ресурсов.
- Обеспечил 100+ релизов на 7 платформах, повысил выполнение квартальных планов с 40% до 83% и в 4 раза снизил затраты на планирование.
- Внедрил инженерные стандарты, observability, SSO, релизные календари и KPI; снизил инфраструктурные расходы.
- Снизил Android crash rate в 14 раз, вдвое сократил время старта видео и достиг iOS crash-free 99,88%.

Atlas Biomed - Backend Team Lead

Апрель 2020 - январь 2022 | healthtech

- Разрабатывал системы обработки медицинских, геномных и микробиомных данных, ML-механизмы и научные алгоритмы.
- Проектировал вычислительные пайплайны и переносил исследовательскую логику в надежные production-сервисы.

TRMPLN Media - Golang Developer

Октябрь 2018 - апрель 2020 | adtech

- Разрабатывал высоконагруженную систему показа рекламы, алгоритмы принятия решений и безопасную обработку персональных данных.

Multy - Golang Developer

Март 2018 - октябрь 2018 | fintech

- Разрабатывал микросервисы для интеграции с финансовыми системами.

Okmeter - Golang Developer

Сентябрь 2017 - декабрь 2017 | monitoring

- Участвовал в разработке сервиса мониторинга Linux-устройств и процессов.

SolidWall - Python Developer

Июль 2015 - июнь 2017 | application security

- Развивал Python-кодовую базу, систему сборки и внедрения ПО, CI и мониторинг.

ПРОФИЛЬ

Руковожу AI-направлением в сфере биомедицины: AI-стратегия, прикладные исследования и разработка. Занимаюсь LLM-системами, агентными архитектурами, RAG, оценкой моделей и AI-инфраструктурой.

CTO с техническими корнями: вырос из Go-разработчика до руководителя департамента. Построил команду из 35 человек, перевел legacy на Go-микросервисы и отвечал за архитектуру, инфраструктуру, бюджет, риски и delivery.

КЛЮЧЕВЫЕ РЕЗУЛЬТАТЫ

35
человек в департаменте

7
команд и платформ

100x
рост выдерживаемой нагрузки

14x
снижение crash rate Android

КОМПЕТЕНЦИИ

- AI/ML и научные алгоритмы
- Engineering leadership
- Стратегия, бюджет и риски
- Архитектура и production engineering
- Найм, процессы и delivery
- LLM-агенты и AI security

Владимир Евграфов

Прикладные AI/ML-системы, production engineering и безопасность LLM-агентов

ОБРАЗОВАНИЕ

Аспирантура - искусственный интеллект

МГУ имени М.В. Ломоносова, ВМК | защита ожидается в 2027

Диссертация о безопасности агентных систем на базе больших языковых моделей: атаки на LLM-агентов и методы защиты.

Магистратура - искусственный интеллект

МГУ имени М.В. Ломоносова, ВМК | 2021

Метод преобразования классических нейронных сетей в спайковые: сохранение точности при более энергоэффективном событийном выводе.

Бакалавриат - прикладная математика и информатика

МГУ имени М.В. Ломоносова, ВМК | 2019

Грамматико-ориентированный фаззинг шаблонизаторов веб-страниц для поиска XSS, серверного выполнения кода и отказов в обслуживании.

ИССЛЕДОВАНИЯ

Исследую безопасность автономных LLM-систем: prompt injection, атаки на память, инструменты и протоколы, мультимодальные угрозы и методы защиты во время инференса.

ИЗБРАННЫЕ ПУБЛИКАЦИИ

2026 **Auditing Inference-Time Defense Evaluation for Multimodal Large Language Models**
arXiv, cs.CR

2026 **Attack Methods and Defenses in LLM-Based Agentic Systems**
International Journal of Open Information Technologies 14 (5)

2023 **Обзор возможных рисков использования систем с мультимодальными языковыми моделями**
Тихоновские чтения

2021 **Спайковые нейронные сети**
International Journal of Open Information Technologies 9 (7)

ТЕХНИЧЕСКИЙ ФОКУС

- Machine learning
- Biomedical AI
- Python
- Go
- C++
- Data pipelines
- Distributed systems
- Microservices
- DevOps
- Observability
- English C1

НАУЧНЫЕ ИНТЕРЕСЫ

- LLM agent security
- Inference-time defenses
- Multimodal models
- AI evaluation
- Spiking neural networks

ЦЕЛЕВЫЕ РОЛИ

Открыт к full-time ролям CTO и Head of AI, а также к частичной и проектной работе как fractional CTO или технический советник.

ПРОФИЛИ

evgrafov.su

LinkedIn

GitHub

Google Scholar