

Roman Beltsov

QA Lead / Head of QA

15 years in testing, 4 years managing teams — up to 15 people at Kaspersky. Infosec, fintech, embedded systems; FSTEC certification. Since 2025 bringing LLMs into QA: RAG over documentation, AI agents, LLM evaluation. Shipped my own Android app with a local LLM.

Looking for	Head of QA / QA Lead role with a team of direct reports
English	technical — fluent reading; spoken — basic
Contacts	t.me/beltsov · dell.oxl@gmail.com · linkedin.com/in/rbeltsov · cv.beltsov.ru/en · AI experience: cv.beltsov.ru/en/ai

KEY ACHIEVEMENTS

- Led a team of up to 15 SDETs, hired 7 engineers
- Two FSTEC certifications — first attempt, on schedule
- Smoke cycle: 1 month → 2-5 days
- Onboarding: 2 months → 1 week
- Test stability: 10% → 97%
- AI test design with RAG: 4 h → 25 min per case
- 100% API autotests with LLMs: 1 month vs a year
- Built and shipped an LLM app for Android

EXPERIENCE

Personal project “I Hear You”Jul 2026 — present

Android app with a local LLM · solo, built with AI agents · published on Google Play

646 commits, 30+ releases, 207 requirements, ~1,000 autotests

- LLM inference optimization: reply 24-28 s → 1.5-2.6 s, found the cause by measuring
- LLM evaluation (eval): LLM-as-a-Judge, a golden dataset, a blind pairwise judge on DeepEval — quality regression on every prompt change
- Guardrails: crisis messages, prompt leaks — enforced in code, covered by tests
- Prompt engineering: one-variable experiments; a short context prompt cut replies from 2,164 to 281 chars and won with the blind judge
- AI agents (Claude Code) and spec-driven development: every test traced to a requirement, 30+ releases on Google Play
- Product metrics in GA4 — no user text

Kaspersky LabDec 2023 — Jul 2026

QA Team Lead · up to 15 people

KISG — cyber-immune IoT gateway on KasperskyOS · Dec 2023 — Jul 2024 · 15 SDETs as direct reports, ~30 developers

- Built test design and test plans from scratch: 100% traceability across 400 requirements
- Prepared the testing evidence base for two FSTEC certifications — the IoT gateway (firewall) and KasperskyOS, two separate releases — first attempt, on schedule
- Cut the smoke cycle from 1 month to 2-5 days, test bench deployment from 2 weeks to 1-3 days
- Built test infrastructure of 13 test benches with centralized management
- Simplified test cases so a Junior can run them instead of a Senior

NGFW — next-generation firewall · Aug 2024 — Jul 2026 · functional lead, ~100 developers, 20-30 in testing

- Replaced one-person expert testing with domains and sub-teams: grew leads who own feature areas
- Introduced requirement tagging — each team's release scope became unambiguous
- Planned releases from actual team capacity: measured where time goes, evened out coverage across all features
- Introduced test plans with suites of varying depth and product configurations instead of formal test runs
- Set up weekly quality reporting: coverage, unverified requirements, defects by priority — the basis for release decisions
- Rebuilt the 2022-2025 backlog, set up a test bench and automation for stability testing

AI in testing · within NGFW · 2025–2026

- RAG over internal documentation: local LightRAG (chunking, embedding, reranker) + MCP
- The LLM generates test design and API autotests grounded in the docs; an engineer reviews every result
- Test design: 4 h → 25 min per case; 100% of the API (40 endpoints, 500+ tests) under autotests in a month instead of a year
- Bench usage monitoring (React + FastAPI) against “forgot to release”: +30% utilization, no new purchases
- Measured automation payoff: e2e caught a fraction of a percent of defects → checks moved down the pyramid

Team · both products

- Hired 7 engineers: screening, technical interviews, offer decisions
- Rebuilt onboarding: first independent test run after 1 week instead of 2 months
- Ran one-on-ones, performance reviews, retrospectives

VK Pay

May 2020 — Aug 2023

Automation QA Lead · up to 5 people · payment system of Russia's #1 social network

- Defined the testing strategy and Quality Gates
- Raised test and environment stability from 10% to 97%, made e2e runs 6× faster
- Prepared acquiring for the AliExpress 11.11 sale: load testing (Locust), 6× throughput
- Set up performance monitoring: JTL Reporter, Grafana

Luxoft

2017 — 2020

QA Automation & CI/CD Expert · project for Russian Post

- Cut the release cycle to 30 minutes by building a full CI/CD pipeline
- Ran load testing (Yandex.Tank)
- Developed an internal communications service for 130,000 users

Earlier

2009 — 2017

Oracle — Senior QA: automation framework, −30% manual testing, mentored 3 people · vKalendare — own project · RADICO — QA for the nuclear industry · Paragon Software

SKILLS

Management	testing strategy and test pyramid, hiring, onboarding, quality metrics, FSTEC certification
AI	RAG, MCP, AI agents, LLM-as-a-Judge, DeepEval, Evaluation-driven development, Guardrails, Prompt engineering, LLM inference optimization, local LLMs: Qwen, Gemma, Llama, Mistral (oMLX, LM Studio, llama.cpp, LiteRT-LM), cloud LLMs: Claude (Opus, Sonnet, Haiku), MiniMax, GLM, OpenRouter, Claude Code, Codex, OpenCode, Hermes Agent, OpenClaw, MiniMax Code, LightRAG, OpenSpec
Automation	Python, PyTest, FastAPI, Selenium, Locust, Allure, Yandex.Tank, JUnit, Maestro
CI/CD	GitLab, Azure DevOps, Docker
IoT	MQTT, Zigbee, Home Assistant, industrial data gateways

OUTSIDE WORK

- Smart home: Home Assistant + Zigbee, 60+ devices — designed and maintained myself
- Travel and hiking: 18 countries, Moscow State University hiking club, hiking and road trips, forest music festivals
- Electric unicycle, snowboarding, drums

EDUCATION

National Research Nuclear University MEPhI, Cybernetics, 2011
New Manager School, Kaspersky, 2024 · Management Intensive, VK, 2021