

Ivan Blinov

Data & AI Platform Architect

contact@ivanblinov.com · linkedin.com/in/ivanblinov · github.com/littleBro
ivanblinov.com · Remote · Americas, UTC-3 to UTC-6

I help teams put AI agents to work on their real data and infrastructure, safely, and make money-critical data platforms fast and correct. Twelve years in software. Today I am the hands-on tech lead for the billing data of millions of sellers at one of the largest marketplaces in Eurasia (100+ TB); I built that team's AI-agent platform and co-founded a health-tech startup in Brazil.

WHAT I DO

AI agents for engineering teams

MCP servers over your databases, orchestration, observability, Git and tracker; capability-level guardrails; AI on-call and AI code review. 2-6 weeks

Migrations and lakehouse

Cluster moves, key-type migrations, Iceberg / Trino / Spark design with ADRs and a working proof of concept. 3-8 weeks

Fractional data & AI architect

One or two days a week for startups and scale-ups: architecture, reviews, hiring, roadmap. retainer

Data platform performance and reliability

ClickHouse, PostgreSQL, Greenplum, Kafka: the mechanism behind slowdowns and incidents, a fix plan, verification. 1-3 weeks

Billing correctness

Executable specifications and formal checks (Dafny, TLA+, Z3, Lean) for money logic in SQL and Go. 2-6 weeks

How I work

I work with coding agents and verify everything they produce: the speed of a small team with one accountable engineer. Safety by capability, not by prompt. Fixed scope, retainer or hourly; full-day overlap with US time zones, mornings with Europe.

SELECTED WORK

An AI-agent platform for a 12-engineer data team

Context. A billing data team on Greenplum, ClickHouse, PostgreSQL, Kafka and Airflow wanted AI help without putting production at risk. **Work.** An MCP server with 100+ tools, most of them read-only; personal read-only database roles; tool and egress allowlists; an autonomous on-call agent and an agentic merge-request reviewer, all under a 21-ADR security architecture. **Result.** The whole team works with agents. The reviewer caught 6 of 7 reference issues versus ~2.5 for diff-only LLM review and reported 15 production findings with zero false positives; the on-call agent's first sweep found a stream that had been stuck for three weeks.

Dashboard timeouts traced to a lock inside ClickHouse

Context. Reporting queries waited ~60 s at peak while most of the CPU sat idle. **Work.** Measured lock acquisitions per query, found a global context mutex saturating at ~58K acquisitions/s per node in ClickHouse 23.8, confirmed that the upstream fix was missing in that version and benchmarked three versions. **Result.** A redesigned query path delivered ~5x throughput on the same version, plus a measured case for the upgrade.

Racing an int32 overflow across a billing platform

Context. Product IDs reached 93.5% of the 32-bit range, growing by up to 20 million a day, across ClickHouse, Greenplum and PostgreSQL. **Work.** Found every consumer in query logs and grants, ordered deployments by how each client breaks, migrated three ClickHouse clusters in a week, and copied 4.7 TB of partitioned PostgreSQL tables into 64-bit twins under a WAL budget. **Result.** Five live functions that would have failed first were caught in time.

Correctness you can check: BendKV and formal methods on billing

Built BendKV with a team of AI coding agents in twelve days: a Redis-compatible server whose core is formally verified against 68 laws, passing 1,769 of Redis's 1,812 own tests; differential testing independently found nine bugs in Redis and Valkey. The same methods on production billing surfaced 42 findings.

RESEARCH

Open Kernel Challenge 2026 — LLM-inference kernels on an NVIDIA H200, built with a fleet of agents.
heilbronn-triangle-bounds — Lean proofs with exact certificates.

BACKGROUND

Wildberries — Tech Lead, 2024-
Groupsi — Co-founder & CTO, 2026-
MDF AG — Engineering Manager, team 6→20
Netfactory — Team Lead, 10 engineers
Sketchfab, Smartbank — backend, B2B SaaS

STACK

Python, Go, SQL, TypeScript ·
Greenplum, ClickHouse, PostgreSQL,
Kafka, Airflow, Spark, Trino, Iceberg ·
Kubernetes, Terraform, Azure, GCP ·
MCP, Claude Code, Codex, OpenCode ·
TLA+, Dafny, Z3, Lean

Languages: English, Spanish, French — full professional; Portuguese — professional working; Russian — native