

Leo Vaychik *Systems & Data Engineer*

For legal purpose: Oleksii Belochenko Calgary, Canada hi@zeelex.me linkedin.com/in/zeelexes
https://www.zeelex.me/

PROFILE

Systems & Data Engineer, Applied Mathematics background. Seven years on one recurring problem: recovering reliable state from unreliable measurement, and keeping that recovery running under production load. Wind resource assessment from raw SODAR and met-mast signal; clinical data platforms turning unstructured hospital records into verifiable schema; adaptive control for high-altitude flight under non-linear turbulence. I build the model and the infrastructure it runs on, because they're the same problem.

WORK EXPERIENCE

Software Developer, *GE Vernova*

Aug 2024 – Dec 2025 | Barcelona, Spain

- Built and maintained the Python ecosystem behind wind resource assessment - SODAR (Sonic Detection and Ranging) and met-mast processing, vectorized in NumPy.
- Implemented detrending and temporal decomposition to analyze air-mass density stability and hub/mast positioning, feeding turbine siting decisions.
- Rewrote critical Airflow DAGs around vectorized algorithms: 5–10X faster task execution and far fewer pipeline failures.
- Migrated 4+ core services to Python 3.x, writing LLM-based static analysis to automate type-hinting, signature refactoring, and dependency resolution (UV/Ruff), and built pymolt to automate behavioral verification.
- Built the observability stack - Grafana, Loki, Prometheus — over GPU-accelerated Kubernetes clusters, cutting incident resolution time through real-time tracking.

Data Engineer, *Trialing Health S.L.*

Dec 2023 – Aug 2024 | Barcelona, Spain

- Set up an AWS S3 data lake and rebuilt the Airflow ETL around it, cutting data-to-insight latency from 120 minutes to 40.
- Designed a multi-tier validation framework with a dedicated quarantine zone for manual verification, holding data integrity across engineering and data-management teams.
- Built an ingestion engine (Regex, Pydantic) that turned non-standardized scientific records into a high-fidelity schema, letting physicians track drug efficacy and treatment outcomes in real time.
- Deployed a unified medical data platform aggregating unstructured research archives from hundreds of hospitals and clinical centers across Spain and Portugal.

DevOps and MLOps, *Lemay.ai*

Dec 2022 – May 2023 | Toronto, Canada

- Administered bare-metal Proxmox hypervisors (LVM partition resizing, hardware allocation) alongside core AWS cloud infrastructure (EKS, EC2, S3).
- Implemented Apache Kafka & Kafka Streams for microservice event streaming, establishing baseline observability across event lifecycles.
- Designed a secure data ingestion layer enforcing strict Role-Based Access Control (RBAC) across dataset onboarding.
- Deployed self-hosted GitLab CI/CD pipelines and managed version-controlled Nginx reverse-proxy configurations.

Data & MLOps Engineer, *Contract (DataArt, Inmost, Virtuace)*

Jul 2021 – Aug 2023

- Delivered infrastructure and ML work for international clients on cloud migration and high-load systems: disaster recovery on AWS to 99% availability, and 15–25% lower storage and processing costs through S3/Lambda resource optimization.
- Led on-premise to cloud migrations, optimized PostgreSQL aggregation schemas, and standardized CI/CD across client engineering teams.

Backend Engineer, ML Infrastructure

Oct 2019 – Jul 2021 | Odesa, Ukraine

Sembly AI — full-time, concurrent with BSc studies

- Architected a high-throughput Kafka audio ingestion system for real-time stream processing, driving a 30% improvement in system responsiveness.
- Decoupled monolithic AI platform into fault-tolerant microservices and overhauled the NLP inference path — slashing deployment latency by 40% and boosting inference speed by 25%.
- Established end-to-end automated model training and validation workflows, cutting model time-to-production by 50%.

Elliscope (co-founded) - part-time, concurrent with Sembly AI

R&D venture on autonomous high-altitude airships and hybrid energy propulsion. Presented at WebSummit 2021.

<https://elliscope.zeelex.me>

- Developed adaptive control algorithms - Kalman filtering and Model Predictive Control - compensating for non-linear turbulence and buoyancy loss at 10–12 km, and assessed controllability across atmospheric profiles.
- Modeled hydrogen fuel-cell integration and energy balance for a theoretical 5+ week flight endurance, identifying the critical lift-to-weight constraints for zero-emission propulsion.
- Designed and tested scale-model demonstrators validating stability and sensor-integration hypotheses; published findings within the Ukrainian Space Association.

EDUCATION

Graduate coursework – MSc in Modeling for Science and Engineering

2025 – 2026 | Barcelona, Spain

Universitat Autònoma de Barcelona (UAB)

Mathematical modeling, dynamical systems and complexity, high-performance computing (HPC)

BSc, Applied Mathematics and Computing, Odesa I.I.Mechnikov National University

2019 – 2023 | Odesa, Ukraine

Numerical methods, optimization and control theory, machine learning, econometrics

Completed while working full-time

PROJECTS

pymolt [↗](#), *Python Migration Tool*

2026

- Runtime tracing tool built on sys.setprofile with streaming JSONL and a diff engine, detecting behavioral changes across dependency version upgrades.
- Architecture: pluggable TraceBackend, JsonISink/MemorySink, three-level early-exit cascade for tracing overhead, TraceScope isolation.

MathMegaMind (M3), how machines can understand mathematics

2026

- Testing whether systematic arithmetic depends on an agent's interface rather than model capacity.
- Built a graph-rewriting substrate (interaction combinators, Python, ~235k rewrites/s) with no numeric primitives; ruled out four stabilization mechanisms.
- Migrating to conservation-law dynamics in chemlambda, per convergent evidence from two independent substrates.
- Target: spontaneous emergence of persistent incrementable linear structures — an observer-independent criterion.

Brixel / companion & engine, local-first webassembly micro-app runtime

2026

- Local-first execution engine translating natural language prompts into sandboxed micro-apps composed from verified WebAssembly modules, eliminating subscription bloat and data leaks.
- Zero-trust capability model blocking network, filesystem, and system APIs by default with deterministic CPU fuel metering and memory-capped threads.
- Architecture: Wasmtime host runner, .brix package spec, WASI syscall virtualization, zero-copy linear memory boundaries, and offline SQLite CRDT vault.

SKILLS

Programming Languages — **Expert:** Python | **Competent:** C/C++, Rust, Java

Mathematical Modeling & HPC — Numerical Methods | Optimization & Control Theory | High-Performance Computing (HPC) | Time-Series Analysis | Vectorized Algorithms | Dynamical Systems

State Estimation & Signal Processing — Kalman filtering | Model Predictive Control (MPC) | time-series analysis & decomposition | detrending | signal processing | control theory | data validation architecture

Data Engineering — Apache Airflow | Apache Spark | Apache Kafka | Celery | dbt | ETL/ELT | data lakes (Iceberg/Delta) | data modeling | data quality | distributed systems

Cloud & Infrastructure — AWS (S3, Lambda, EC2, EKS, Glue) | Kubernetes | Docker | Terraform | Helm | CI/CD (GitHub Actions, GitLab CI) | Linux | microservices | infrastructure as code

ML & AI — PyTorch | TensorFlow | scikit-learn | MLflow | NumPy | Pandas | MLOps | model serving | NLP | computer vision | LangChain | DSPy

Observability — Prometheus | Grafana | Loki | Promtail | monitoring | incident response

Back-end & Databases — FastAPI | Flask | Django | Django Rest Framework | ClickHouse/Apache Doris | PostgreSQL | MySQL | MongoDB