

Daniel Klein 魏但以

email: daniel.l.klein@pm.me • web: valueof.info • github: othercriteria

Experience

Independent work: agentic systems & quantitative trading

2024 -- present

Sole architect and operator

- Built and operate **Decapod**, an agentic portfolio-management system trading live capital: LLM portfolio managers behind SRE-grade rails: SLOs, daily reconciliation, canary PMs, milestone gates, and a human-operator playbook.
- Simulated shop at ~\$1.1M scale runs at ~1.6 factor-adjusted Sharpe net of costs under minimal attendance: a daily operator sign-off against an automated blotter, with agent autonomy scored down whenever human review is skipped. Earlier, higher headline Sharpe was largely momentum and AI-capex beta on leverage; the factor-adjusted number is the one I stand behind.
- Run the substrate: 5-host NixOS fleet with Prometheus/Grafana/Loki observability, Nix binary cache, self-hosted CI runners, GitOps deployment.
- Reading-group replications (2026), each a reproducible Nix/uv repo with tests and hash-pinned artifacts: MaxEnt inverse RL (Ziebart et al. 2008) end to end, including a substitute route-choice pipeline on the 1.7M-trip Porto taxi corpus when the original data proved unavailable; de Blanc (2011) on ontological crises, where I found a dimensional error in the printed objective, showed equal-loss ontology maps transfer different rewards, and proved a sufficient condition for path-independent reward transfer that does not need state-space invertibility.
- Selected open work: *stiletto* (multilingual ML pipeline: content-addressed artifacts, graph clustering, SFT/adaptor training; live at stiletto.valueof.info); *vibectl* (LLM-driven kubectl with structured-output rails, exercised adversarially via CTF and Chaos Monkey demos).

Oracle Corporation (Advertising and Customer Experience, Moat)

2019 -- 2024

Principal Data Scientist

- Developed, instrumented, and maintained streaming pipelines for Invalid Traffic detection, consuming ~250,000 events / sec.
- Migrated production systems cross-cloud (AWS to OCI) and cross-architecture (x86 to arm64).
 - Shifted ETL on EC2 VMs to k8s / Argo Workflows / Spark, saving ~\$40K/month and reducing P99 delay ~90%.
 - Replaced legacy real-time reporting with Spark Streaming (Kafka to Trino), saving ~\$150K/month.

DMG Blockchain Solutions (Blockseer)

2018 -- 2019

Senior Software Engineer

- Led rewrite of Bitcoin address clustering (EMR, Scala, GraphX); 1.2 TB blockchain processed in ~6 hours.
- Validated novel Bitcoin transaction classification method using custom DeepWalk implementation.

RealScout

2016 -- 2017

Data Scientist (first data science hire)

- Productionized insights from agent-agent modeling into user-facing Recommended Listing feature.
- Built out Bayesian models for sales price and days-on-market prediction with near-Zillow/Redfin accuracy.

Earlier

2003 -- 2016

Data science apprenticeship rotations (Catenus / 8VC: Radius Intelligence, RealScout, Neighborly, 2016); scientific computation for lab automation at TransForm Pharmaceuticals (2006 –2008); ecology field research and modeling (Williams College, University of Minnesota).

Education

Brown University, Ph.D. program, Applied Mathematics

2008 -- 2016 (did not complete)

- Advised by Matt Harrison; research on statistical inference in settings of extreme data sparsity or imbalance.

Williams College, B.A. Mathematics and Biology, departmental honors in Biology

2006

Systems & methods

- Streaming / batch data systems: Kafka, Spark (incl. Structured Streaming), Trino, Argo Workflows, Kubernetes.
- Operations: SLO design, reconciliation-first ops, Prometheus/Grafana/Loki; NixOS, GitOps, infrastructure-as-code.
- ML & agents: PyTorch; SFT / LoRA adapters; embedding and clustering pipelines; LLM agent harnesses (claude-agent-sdk, MCP), structured-output rails, eval design; inverse RL and reward-transfer replications.
- Statistics: Bayesian modeling; inference under extreme sparsity / imbalance; factor models and performance attribution.

Comparative advantage

- I'm constructively skeptical of data, design guarantees, and monitoring; these add value iff they track reality.
- I'm unsentimental about using dumb methods over theoretically optimal ones or those I've built myself.
- I look to the non-technical contexts of product, growth, and markets to frame technical choices. I read The Diff.