

Christopher Gorski

Data Engineer | AI Experts | AWS & GCP Specialists

<https://www.linkedin.com/in/christopher-gorski-7a762038a> | chris.dev1472@outlook.com

Work Experience

Senior Software Engineer – Allegro

April.2024 – Present

- Built an event-driven pricing/checkout data plane: Kafka topics partitioned by merchant+SKU; idempotent consumers with Debezium CDC from OLTP into ClickHouse for sub-second analytics.
- Real-time materialized views: aggregated price/stock snapshots (Parquet on S3 + Iceberg tables) refreshed via Airflow + incremental upserts; cut query time from 2.8 s → 320 ms for dashboard drill-downs.
- RAG-ready catalog enrichment: generated product embeddings offline, stored in pgvector with HNSW; enabled semantic matching for substitutes/upsells increasing attach-rate by +3.1% on A/B.
- Data quality SLAs: Great Expectations suites on critical DAGs (freshness, null %, schema drift) + data contracts with producers; prevented 4 broken deploys and reduced downstream incident rate -40%.
- Cost optimization: column pruning + ZSTD compression, ClickHouse TTLs, Iceberg compaction; saved ~\$6.2k/mo on storage & scan costs.
- Anomaly detection: seasonality-aware z-score on checkout funnel + payment auth declines; detected 3 silent regressions pre-release (rolled back within 15 min).
- End-to-end traces: OTEL spans stitched from browser RUM → API → Kafka → Airflow task → ClickHouse; mean debug time 45 → 12 min across 12 services.
- Data privacy: implemented PII tokenization (format-preserving) and row-level security for BI; fulfilled DSAR exports in < 5 minutes.
- Frontend analytics portal: React + Ant Design with query templates and saved views; reduced ad-hoc SQL requests to data team by -55%.

Senior Software Engineer – Artisan AI

Mar.2024 – Mar.2025

- Search & RAG platform: hybrid retrieval (BM25 + pgvector dense) with Redis Bloom prefilters and re-rankers; F1 +13%, tail p99 520 → 260 ms on multilingual corpora.
- Feature store & offline metrics: bootstrapped Feast over ClickHouse/Postgres for offline/online consistency; label joins sped up 9x; leakage bugs dropped to zero.
- Model lifecycle: MLflow registry + staging/prod canaries via KServe; automated shadow traffic and A/B with p-value gates; rollback in < 5 min.
- Batch & stream inference: Airflow for nightly batch, Kafka for online scoring; autoscaled consumers to hit p95 ~180 ms @ 2.1k rps while keeping GPU util > 75%.
- Guardrails & governance: prompt sanitizers, profanity/PII filters, grounding checks; hallucinations -28%, policy violations -62%; red-team playbooks documented.
- Evaluation harness: BLEU/ROUGE + task-specific metrics and human-in-the-loop labeling UI (React); cut weekly eval cycle -50%.
- Vector hygiene: periodic HNSW rebuilds, dedupe, and drift monitors; false-positive rate -18% with stable latency.
- Data mesh alignment: published dataset contracts + ownership; on-call load to data platform team -30% by pushing transformations to domains.
- Cost: Triton dynamic batching + mixed precision; A100 throughput +41%, request unit cost -24%.

DevOps Engineer – AioCare

Oct.2021 – Feb.2024

- Rolled out a Kubernetes platform (3 envs, 8 node pools) with Terraform + Helm + Argo CD, cutting deploy time 45 min → 7 min (-84%) and enabling ~20 prod releases/week; verified in Argo CD history and Jira releases.

- Built progressive delivery (blue/green + canary 5%→25%→100%) on GitHub Actions → Argo Rollouts, reducing failed deploys by ~72% and MTTR 38 min → 11 min; confirmed via PagerDuty + Grafana incident timelines.
- Introduced SLOs (p95 API < 250 ms, error budget 1%/30d) with burn-rate alerts; lifted API uptime to 99.95% over 12 months and cut MTTD 12 min → 3 min; Prometheus/Grafana dashboards as evidence.
- Optimized cloud spend -28% YoY using HPA/VPA, spot nodes, storage lifecycle policies, and right-sizing; environment spin-up shrank 2 days → 2 hours with reusable Terraform modules; validated by Cost Explorer/Billing + CI timestamps.
- Shipped end-to-end observability (OpenTelemetry traces, Prometheus metrics, Loki/ELK logs) and golden dashboards per service, lowering “unknown root cause” postmortems by ~60% and after-hours pages by ~50%; PagerDuty analytics + RCA tags.
- Hardened SDLC for medical data: SBOM + image signing (Syft/Grype, Cosign), OPA/Gatekeeper policies, Vault/KMS secrets; reduced critical CVEs in images by >90% within 2 quarters; security scan trends used for proof.
- Delivered DR runbooks and quarterly game-days achieving RPO ≤ 5 min **and** RTO ≤ 30 min via cross-region backups and infra-as-code restores; results documented in DR drill reports.

Software Engineer – Spyrosoft

Jul.2016 – Sep.2020

- Designed, developed, and maintained enterprise and consumer-facing applications for multiple international clients, such as a fintech dashboard for BNP Paribas, an e-commerce platform for a retail client, and a healthcare management tool for a telemedicine startup; delivered scalable solutions using React, Angular, and Vue on the front end, with FastAPI, Django, Java, and Golang powering high-performance back ends, ensuring seamless integration, security, and excellent user experience.

Education

Master's Degree in Computer Science
Lublin University of Technology

10/2014 - 03/2016

Bachelor's Degree in Computer Science
Lublin University of Technology

05/2010 - 09/2014

Skills

- Data Platform: PostgreSQL, ClickHouse, BigQuery, Snowflake (basics), Kafka (Streams/KSQL), Debezium CDC, Airbyte/Fivetran, dbt Core, Airflow
- Storage & Lakehouse: S3/MinIO, Parquet, Iceberg/Delta Lake tables, Glue/BigQuery Metastore
- Vector/RAG: pgvector, Redis (Bloom/TF-IDF), Qdrant/Pinecone (familiar), re-ranking (ColBERT/monoT5 style)
- ML/MLOps: Python (FastAPI), Feast feature store, MLflow, Triton Inference Server, KServe, batch & stream inference, model registry & canaries
- Data Quality & Governance: Great Expectations, data contracts, lineage (OpenLineage), PII masking & tokenization, GDPR/DSAR workflows
- Observability: OpenTelemetry, data SLOs, anomaly detection, Prometheus/Grafana, Datadog
- Frontend: React, Angular, Vue for analytics portals & labeling tools
- Infra: Docker, Kubernetes, Argo CD/Rollouts, Terraform, IAM/VPC, Vault, SSO/OIDC