

Sean Kraemer

Chicago, IL | seankraemer.com | linkedin.com/in/sean-kraemer | github.com/SeanKraemer

EXPERIENCE

Zurich North America

Data Scientist

Schaumburg, IL

Oct. 2023 – Present

AI Document Understanding · Sep. 2025 – Present

- Leading development of the LLM agent that rates corporate compliance with Zurich's underwriting sustainability guidelines, classifying each provision red/yellow/green with every verdict cited to its source; shipped the underwriter-facing app and am driving productionization to an Azure Function App
- Generates a report in 1-2 minutes against a 4-hour manual review; designed for a sub-15-minute human-reviewed decision
- Engineered its retrieval (RAG) pipeline: chunking, byte-aware batching, and idempotent upsert/delete against an HNSW vector index on Azure AI Search, with async retries and Pydantic-validated JSON outputs
- Built a second agent for middle-market underwriting guidelines: a context-management agent infers which account-data sections a query needs; a routing orchestrator runs agents concurrently (asyncio)

Quantitative Pricing & Technical Underwriting · Oct. 2023 – Sep. 2025

- Rebuilt the U.S. Workers' Compensation pricing model; built its Medical-Only submodel (Tweedie GLM, 4-fold CV), now in Zurich's production pricing systems and used by underwriters, plus a parallel Databricks batch scoring pipeline for model reviews
- Owned the Manufacturing and Products submodels in the U.S. General Liability refresh and re-engineered Zurich Canada's GL model (MLflow-tracked) for the company's re-entry into that market; both in production
- Presented WC, General Liability, and Auto model-performance reviews, including models I did not build, converting lift and predictor-drift findings into model iterations and underwriter adjustments

Associate Data Scientist

Sep. 2022 – Oct. 2023

- Optimized a core Workers' Compensation modeling pipeline, cutting runtime 70% (5h to 1.5h) by porting Pandas code to PySpark on Databricks

Actuarial Analyst

Aug. 2021 – Sep. 2022

- Co-developed a data scorecard (SQL, R, SAS) that improved pricing-team reporting; enhanced a core data tool with Power Query and Power Pivot. Passed actuarial exams 1/P, 2/FM, and 3/IFM (2020 to 2021)

PROJECTS

Distributed Stream Processor | *Go, Docker*

github.com/SeanKraemer/distributed-stream-processor

- Built a distributed stream-processing system in Go: SWIM-style failure detection with gossip dissemination, a consistent-hashing replicated file system, and an exactly-once leader-worker streaming engine on a 10-node Docker cluster
- SWIM suspicion and confirm windows give a 6s detection bound; the file system re-replicates lost blocks in a measured 1.0–1.35s after a node failure

Devil's Advocate | *React, FastAPI, Socket.IO, Gemini*

github.com/SeanKraemer/devils-advocate

- Built a real-time AI debate platform that stress-tests startup ideas: Socket.IO streaming backend, Gemini Live voice and text debaters with retrieval (RAG) memory, on Firebase Hosting and Cloud Run with a public demo; every external client sits behind a deterministic mock mode so the demo runs keyless, with 81 backend tests and ruff/pytest/ESLint/Vitest in CI

EDUCATION

University of Illinois Urbana-Champaign

Master of Computer Science (MCS) · GPA 3.93/4.0

May 2026

B.S. in Actuarial Science, with Distinction

May 2021

SKILLS

Languages: Python, Go, SQL, Java, R

ML & AI: LLM agents & orchestration, RAG, LangChain, Pydantic structured outputs, vector search, MLflow, PySpark, scikit-learn, statsmodels

Systems & Cloud: Azure (OpenAI, AI Search, Functions, Databricks), Docker, FastAPI, asyncio, pytest, GitHub Actions CI, Linux