Sean W. Evans

+1 (516) 330-2517
Website

E-Mail
Github
Linked In

Summary

Systems-level engineer specializing in High-Performance Computing (HPC), Database Internals, and Machine Learning. Expert in bridging theoretical modeling with production engineering using CUDA, Rust, and C++. Proven track record of optimizing high-throughput pipelines and developing novel open-source architectures, including GPU-accelerated physics engines and secure CI/CD static analysis tools.

Skills

Languages: Rust, C++, C, CUDA, Python, SQL (Postgres/T-SQL), TypeScript

Systems & HPC: GPU Kernels (CUDA/AVX), Distributed Consensus (Raft), Linux, Compiler Design (AST), Multiprocessing

Data & Databases: PostgreSQL Internals, Extension Development, Vector Search, ETL Pipelines, XML (DOM/SAX)

Machine Learning: PyTorch, TensorFlow, Detectron2, Computer Vision (OpenCV), Structural Analysis

Open Source Engineering

fluid-sims (High-Performance Computing) *CUDA, C++, SPH, Navier-Stokes*

– Engineered a GPU-accelerated Computational Fluid Dynamics (CFD) engine. Implemented Smoothed-Particle Hydrodynamics (SPH) and Shallow Water Equations with log-depth stability, achieving real-time simulation of 100k+ particles.

pg_os & pg_gpt2 (Database Internals) *C, PL/pgSQL, Postgres Extensions*

– Developed a custom runtime environment within PostgreSQL exposing OS primitives (processes, scheduler). Built a C-based extension to embed GPT-2 inference directly into the database engine for zero-latency batch processing.

Ghast (Security Engineering) *Python, AST Analysis, GitHub Actions*

– Built a static analysis tool for CI/CD pipelines preventing 'Poisoned Pipeline Execution' and enforcing immutable action pinning. Integrates with SARIF for native GitHub Security alerts.

raft-vm (Distributed Systems) *Rust, Consensus Algorithms, Actor Model*

– Implemented a fault-tolerant virtual machine leveraging the Raft consensus algorithm and Actor model to ensure state replication and consistency across distributed nodes.

Experience

Data Conversion Laboratory **Remote**
Machine Learning Engineer *May 2021 – Feb 2025*

– Designed a Detectron2-based document segmentation pipeline reducing manual classification by 50 %.
– Built a CNN-RNN LaTeX OCR system with near state-of-the-art accuracy.
– Automated document styling using NLP + MS Office SDK, reducing a 40-hour workflow to 2 hours.
– Implemented an OpenCV-based checkbox detector achieving >96 % accuracy across thousands of formats.
– Developed high-throughput pipelines capable of processing millions of pages per week.

Data Conversion Laboratory **Remote**
Software Engineer *Sep 2020 – May 2021*

– Architected a PDF cleaning and OCR preprocessing system scaling to 100k+ pages per week per server.
– Automated LaTeX and JATS XML correction workflow, replacing a full-time manual process.
– Developed robust .docx to XML conversion tooling using MS Office InterOp and C #.

Data Conversion Laboratory **Queens, NY**
Lead Technology Analyst *Sep 2019 – Sep 2020*

– Led workflow optimization initiatives, mentored junior engineers, and coordinated technical requirements with stakeholders.

Data Conversion Laboratory **Queens, NY**
Technology Analyst *May 2018 – Sep 2019*

– Developed custom data-conversion tools for production support and maintained high-volume enterprise pipelines.

Education

- **Long Island University**
  Bachelor of Science in Mathematics *Fall 2025*