

SEAN W. EVANS

+1 (516) 330-2517
SeanWEvans.com

E-Mail 
Github 
Linked In 

SUMMARY

Systems-level engineer with 6+ years building end-to-end ML, vision, and document-analysis pipelines, plus experience in C/C++/CUDA, runtime systems, Postgres tooling, and mathematical modeling. I specialize in rapid prototyping across abstraction layers (theory → GPU → product), automating complex workflows, and shipping high-reliability systems at scale.

SKILLS

Languages: Python, C, C++, CUDA, C#, Perl, SQL (PostgreSQL/T-SQL)

Machine Learning: PyTorch, TensorFlow, Detectron2, CNN/RNN OCR, NLP (SpaCy)

Computer Vision: OpenCV, document segmentation, structural analysis

Systems & Performance: Linux, Git, Make/CMake, multiprocessing, GPU kernels

Data & SDKs: PostgreSQL, XML (DOM/SAX), PDF libraries, MS Office SDK, Adobe SDK

Scientific & Typesetting: L^AT_EX, MathJax, numerical & geometric modeling

EXPERIENCE

Data Conversion Laboratory

Machine Learning Engineer

Remote

May 2021 – Feb 2025

Designed and deployed a Detectron2-based document segmentation model (equations, tables, figures) achieving industry-comparable accuracy and reducing manual classification by **50%**.

Built a CNN–RNN equation OCR engine (Torch7/Lua) converting images to L^AT_EX with near state-of-the-art accuracy in production.

Automated large-scale document styling with NLP + MS Office SDK, reducing a recurring 40-hour manual workload to **under 2 hours**.

Implemented a robust OpenCV checkbox detection pipeline (>96% accuracy) handling rotation, noise, and thousands of formatting variants.

Developed high-throughput processing systems capable of millions of pages/week across OCR, normalization, and structural analysis.

Data Conversion Laboratory

Software Engineer

Remote

Sep 2020 – May 2021

Built a scalable PDF cleaning + OCR preprocessing system (deskew, despeckle, rotate) processing **100k+ pages/week/server**.

Authored automated L^AT_EX correction and JATS XML tooling replacing a full-time manual process with a reliable automated pipeline.

Developed a comprehensive .docx → XML converter via MS Office InterOp, normalizing heterogeneous client inputs into a unified representation.

Applied TDD, Agile workflows, and robust XML parsing (DOM/SAX) to ship reliable, production-grade tools.

Data Conversion Laboratory

Lead Technology Analyst

Queens, NY

Sep 2019 – Sep 2020

Led optimization initiatives across ingest, transformation, and QA pipelines; mentored junior developers; served as technical stakeholder liaison.

Data Conversion Laboratory

Technology Analyst

Queens, NY*May 2018 – Sep 2019*

Built custom data conversion and document-normalization utilities used in large-scale document migration projects.

Maintained and enhanced high-volume production pipelines supporting multiple enterprise clients.

EDUCATION

- **Long Island University**

Bachelor of Science in Mathematics

Long Island, NY*2003–2026*