Stephen Sequenzia

Orlando, FL | stephen@sequenzia.com | 407.718.2010 | linkedin.com/in/sequenzia | github.com/sequenzia

ACTIVE SECRET SECURITY CLEARANCE

PROFESSIONAL SUMMARY

Senior Staff AI/ML Engineer & Architect with 20+ years building production ML, data, and software systems end to end, now leading AI/ML architecture and technical strategy for production agentic AI systems in secure, large-scale defense environments. At Lockheed Martin, help lead multiple engineering teams across 8–10 concurrent AI/ML efforts, drive agentic AI enablement reaching 200+ engineers, and ship production agents, bespoke Model Context Protocol (MCP) servers, evals, and guardrails with Pydantic AI and LangChain/LangGraph on Kubernetes and AWS GovCloud — alongside an enterprise MLOps framework that cut production deployment time by roughly 85% and deep learning optimizations that cut training time by over 94%. Author of three open-source ML frameworks (Photon, Maxwell, Dyson) and a growing ecosystem of agentic tools, with an owner-operator background pairing deep hands-on engineering with technical leadership that turns cutting-edge AI into reliable, measurable business value.

TECHNICAL SKILLS

Agentic AI: Pydantic AI, LangChain, LangGraph, MCP, RAG Pipelines, Multi-Agent Systems

AI Engineering & Tooling: Pydantic Evals, Pydantic Logfire, Claude Code, Codex, FastAPI, MLflow

AI/ML Frameworks: PyTorch, TensorFlow, Keras, Scikit-learn, NumPy, Pandas

Data Engineering: Apache Airflow, Apache Arrow, Apache Parquet, Dask, Ray, Flyte, PostgreSQL

Cloud & Infrastructure: AWS, Kubernetes, Docker, GPU/CUDA Acceleration

Languages: Python, TypeScript, C++, SQL

PROFESSIONAL EXPERIENCE

Lockheed Martin Corporation
Oct 2022 — Present

Orlando, FL

Senior Staff AI/ML Engineer & Architect — Apr 2026 — Present

Staff AI/ML Engineer | Technical Lead — Oct 2022 — Apr 2026

Leading AI/ML architecture and technical strategy for production agentic AI systems across enterprise defense programs — guiding multiple teams and concurrent efforts, driving organization-wide AI enablement, and staying hands-on from multi-agent runtimes to MLOps frameworks and distributed data pipelines.

- Lead AI/ML architecture and technical strategy for production agentic AI systems deployed across multiple programs in secure environments, working hands-on with teams of AI, software, ML, and data engineers from design through deployment
- Serve on the engineering leadership team guiding architecture and technical direction across 8–10 concurrent AI/ML efforts, driving cross-team alignment on architecture decisions and design patterns, with continuous involvement in technical interviews and hiring across multiple teams
- Serve as a senior lead on the AI Enablement initiative within the software engineering organization at Lockheed Martin Missiles and Fire Control (MFC), driving adoption of AI tooling and agentic engineering practices reaching 200+ engineers
- Design and tune the context-engineering layer governing information flow into models each turn — RAG and retrieval pipelines, summarization and compression, short-term scratchpads, long-term persistent memory, and the ordering and pruning required to stay within context budgets — the highest-leverage surface for agent quality
- Architect the agent runtime harness powering production agents: core agent loop, tool-call orchestration, state management, retry and error recovery, planning logic, stopping conditions, and multi-agent routing, delegation, and shared-state coordination
- Define and build the action surface for production agents — integrating internal APIs, databases, ticketing systems, and program-specific platforms — standardizing on Model Context Protocol (MCP) and engineering bespoke MCP servers with model-legible, safely scoped tools; multiple production agents and MCP services running across programs, built with Pydantic AI and LangChain/LangGraph
- Stand up the evaluation, tracing, logging, and monitoring infrastructure required to operate non-deterministic AI systems with confidence, championing evals-driven development and treating observability as first-class from inception
- Architect agent guardrails, security controls, and access management — I/O filtering, prompt-injection defenses, permissioning, and scoped identity and authorization constraining agents to each user's access rights — often the gating factor for deployment clearance
- Champion team-wide adoption of agentic AI, mentoring 10+ engineers across multiple teams on agent architecture and tool-augmented reasoning
- Architected an MLOps framework spanning the full ML lifecycle — training, deployment, reproducibility, and data lineage — cutting production deployment time by roughly 85% (from ~2 weeks to under 2 days) and standardizing releases across 10+ models and pipelines; continue supporting MLOps and DataOps teams as a subject-matter expert through guidance and design reviews
- Diagnosed and eliminated training bottlenecks in PyTorch deep learning anomaly-detection models using Torch Profiler and NVIDIA Nsight; refactored the data loader and rearchitected key layers to cut training time by over 94% (from ~24 hours to under 90 minutes)
- Built distributed, functional-programming-based data pipelines with Apache Airflow and Dask, processing tens of terabytes across 15+ data sources and roughly tripling throughput for ML and analytics workflows
- Spearheaded architecture of an Apache Arrow-native data lakehouse on Amazon S3 and PostgreSQL, managing 50+ TB of time-series data, metadata, and an ML feature store serving 10+ production models with sub-second analytical queries
- Refactored and re-normalized a core PostgreSQL database, cutting query response times by over 95% (from ~10 seconds to under 500 ms) and enabling efficient scaling to ~3x higher workloads
- Deployed and operate a secure Kubernetes cluster on AWS GovCloud supporting 20+ ML and data-visualization workloads, sustaining 99.9% availability with automated scaling in a compliance-controlled environment

Applied Theta Inc
Feb 2016 — Oct 2022

Orlando, FL

ML Engineer | Software Engineer | Data Engineer | Owner

Owner-operated ML R&D practice (sister brand of Zia Technology Group) focused on ML-based algorithmic trading research and development, with selective external ML consulting.

- Founded and directed an independent ML R&D practice, defining the end-to-end technology strategy, roadmap, and agile delivery process across data, modeling, and algorithmic trading systems
- Engineered three open-source Python frameworks — Photon (deep learning lifecycle), Maxwell (market-data preprocessing and ML pipelines), and Dyson (financial modeling and algorithms) — spanning data preparation through training, evaluation, and serving
- Designed deep learning and probabilistic models (CNNs, RNNs, Transformers, Bayesian neural networks, deep ensembles) powering algorithmic trading systems across equities, options, futures, and crypto markets
- Built GPU-accelerated (CUDA) data pipelines ingesting 20+ years of trade and order-book market data at multiple time resolutions, processing hundreds of millions of data points with Dask, RAPIDS, and Apache Arrow/Parquet
- Standardized research and production codebases on TensorFlow, PyTorch, Keras, NumPy, and Pandas, eliminating duplicated implementations and streamlining the path from experimentation to live trading
- Delivered external ML consulting, including a transformer- and autoencoder-based prediction system that improved a sports-analytics client's predictive accuracy by more than 65%

Zia Technology Group Inc

Mar 2010 — Aug 2020

Orlando, FL

Data Engineer | Software Engineer | IT Consultant | Owner

External data and software consulting brand of the same owner-operator practice as Applied Theta — working with client senior leadership to shape data-management technology strategies as consultant, architect, and lead developer for businesses across industries.

- Architected data integrations and ETL pipelines connecting 25+ business applications and data sources — including MS SQL Server, Salesforce API, and QuickBooks Enterprise — automating reporting workflows and eliminating hundreds of hours of manual data handling annually
- Deployed and administered 40+ SQL/NoSQL databases with custom schema designs supporting client applications, data warehouses, and reporting systems
- Delivered full-stack web applications and reporting systems — including a client web portal unifying four backend systems in real time via Node.js and WebSockets — serving hundreds of business users
- Designed and implemented highly available cloud and hybrid co-lo/cloud virtual infrastructures for client environments, sustaining 99.9% uptime

Thinksys Incorporated

Jan 2007 — Mar 2010

Orlando, FL

Software Developer | DBA | IT Consultant | Owner

Owner-operated consulting practice collaborating with client senior leadership to align technology with business needs — serving as consultant, DBA, and lead developer on critical IT infrastructure, data-integration, and web development projects.

- Administered 30+ MS SQL Server instances and databases powering custom and third-party business applications across client environments
- Designed and developed complex data integrations connecting multiple business applications and back-end data sources across client environments
- Developed web-based interfaces providing business users self-service access to back-end data sources, reducing manual data requests
- Led a team of 4 network engineers and developers delivering client infrastructure and software projects

Biziteks Incorporated

Dec 2001 — Jan 2007

Orlando, FL

Network Engineer | Software Developer | DBA | Owner

Full-stack IT consulting practice supporting networks, servers, and applications for small and mid-sized businesses across a range of industries.

- Led a team of 15 technicians and network engineers supporting 200+ client networks and applications, improving on-time delivery and service reliability across engagements
- Deployed and supported hundreds of client networks, servers, and MS SQL Server database environments for custom and third-party applications
- Built dynamic, database-driven websites and web applications for small-business clients

SELECTED PROJECTS & OPEN SOURCE

Photon — ML Framework

`github.com/sequenzia/photon · Jun 2018 – Oct 2022`

Open-source ML framework extending TensorFlow, Keras, and PyTorch for deep learning on financial time series

- **Stack:** Python, TensorFlow, Keras, PyTorch, NumPy, Pandas, Scikit-learn, Apache Arrow, Apache Parquet, Distributed Training
- Native TensorFlow distributed-strategy support with shared inputs/outputs across networks for deep ensemble learning and dynamic learning-rate scheduling
- Real-time preprocessing — dataset splitting, normalization, scaling, aggregation, and time-series resampling — with custom batching, padding, and masking
- Simple interface for saving, serializing, and loading entire networks, including learned and hyperparameters
- Detailed parameter logging throughout the pipeline for interpretability and reproducibility

Maxwell — ML Data Pipelines

`github.com/sequenzia/maxwell · Jun 2018 – Oct 2022`

GPU-accelerated pipelines processing 20+ years of market data with Dask, RAPIDS, and Apache Arrow

- **Stack:** Python, Pandas, NumPy, Numba, RAPIDS, Dask, Apache Arrow, Apache Parquet, Scikit-learn, NVIDIA CUDA
- Automated labeling of time-series data with algorithmic trading attributes — stop losses, long/short price targets, ATR, VWAP
- Domain-driven feature engineering from technical indicators and statistical markers, with automated binning, one-hot encoding, scaling, and normalization tuned for market microstructure data
- ML pipeline orchestration tools managing data flow across modeling and algorithmic-trading workflows with fluid feature extraction
- GPU-accelerated (CUDA) processing with Dask and RAPIDS for hundreds of millions of data points

Dyson — Modeling & Algorithm R&D `github.com/sequenzia/dyson · Jun 2018 – Oct 2022`

Financial market models spanning statistical methods through deep neural networks and probabilistic reasoning

- **Stack:** TensorFlow, Keras, Scikit-learn, statsmodels, CNNs, RNNs, Transformers, Bayesian Neural Networks, ARIMA
- Model families spanning CNNs, RNNs, Transformers, Temporal Convolutional Networks, Autoencoders, Bayesian neural networks, ARIMA, and structural time-series models
- Extensive exploratory data analysis (EDA), feature engineering, scaling/normalization, resampling, and principal component analysis (PCA)
- Advanced generalization and regularization techniques that lift model performance beyond the training data, specialized for financial time series

Algorithmic Trading System `Feb 2016 – Oct 2022`

Python/C++ libraries powering live execution and backtesting across equities, options, futures, and crypto

- **Stack:** Python, C++, NumPy, SciPy, Pandas, Numba, statsmodels

EDUCATION & PROFESSIONAL DEVELOPMENT

Autodidact and lifelong learner: 20+ years of self-directed, build-first professional development spanning networking, databases, software engineering, data engineering, machine learning, and agentic AI — with expertise demonstrated through production systems, self-authored open-source ML frameworks, and a growing ecosystem of open-source agentic tools.

- Coursera (DeepLearning.AI): Machine Learning Engineering for Production (MLOps) — Oct 2022
- Google: TensorFlow Developer — Sep 2022