

Matt Oremland, PhD

Director, Data Science

mattoremland@gmail.com • (518) 560-1473 • Saratoga Springs, NY 12866
<https://www.linkedin.com/in/matthew-oremland-31a312101>

PROFESSIONAL SUMMARY

Data science leader with 9+ years building production ML systems and managing cross-functional teams in regulated, high-stakes environments. Proven track record delivering end-to-end data science products from incubation to deployment, establishing responsible AI frameworks, and driving measurable business value through AI-powered decisioning. Deep technical expertise in Python, SQL, machine learning, and LLMs with experience deploying models on cloud platforms. Strong leadership skills managing data scientists, engineers, and vendors across agile delivery cycles. Healthcare-adjacent experience includes working with HCP/HCO healthcare reference data and operating in FDA-regulated pharmaceutical environments requiring rigorous compliance and auditability standards.

PROFESSIONAL EXPERIENCE

Redica Systems

April 2025 – Present

Director, Data Strategy & Analytics

Remote

- Lead data science team delivering AI-powered regulatory intelligence products for pharmaceutical clients including Eli Lilly and Merck; manage team roadmap aligning long-term data science initiatives with product strategy and business requirements
- Built production LLM classifier using OpenAI API processing unstructured FDA inspection reports; established responsible AI framework ensuring model interpretability, explainability, and adherence to governance standards for enterprise deployment
- Develop and track product OKRs measuring value creation through metrics including model accuracy, system latency, user adoption, and business impact; communicate data science approach and results to business stakeholders in understandable terms
- Architect data science products from incubation to production including Python-based risk scoring models, reference data harmonization systems, and analytics visualization libraries; establish CI/CD practices and automated testing frameworks
- Manage cross-functional collaboration with engineering, product, and client stakeholders ensuring alignment of goals and priorities through focus on end-user experience; present technical work at industry conferences (CHPA, ISPE)
- Design feature engineering pipelines processing structured data (facility attributes, historical violations) and unstructured data (regulatory text) in Snowflake; implement data quality validation and monitoring frameworks

Tidal Wave Analytics LLC

January 2025 – Present

Consulting Data Scientist / Founder

Saratoga Springs, NY

- Deliver end-to-end data science products across multiple clients including SaaS, tax technology, and consumer products; manage vendor relationships and contractor coordination to meet project goals and timelines
- Built production random forest model for Paradigm deployed to Databricks with nightly batch processing; designed automated testing, monitoring, and alerting infrastructure ensuring production-grade system reliability
- Engineered generative AI system using RAG architecture with HuggingFace and PyTorch for tax compliance decisioning; developed data processing pipeline ingesting unstructured documents (PDFs) and implementing vector search for semantic retrieval
- Developed full-stack data product (Escalators puzzle game) demonstrating product development mindset: Python backend, React/Node.js frontend, Supabase database, Google OAuth; running in production with competitive leaderboards and daily user engagement
- Establish best practices for maintainable code, version control (Git), automated testing, and deployment workflows across client engagements; facilitate delivery in both code-centric and low-code environments

Gilead Sciences

January 2023 – March 2025

Director, Data & Analytics (Medical Affairs)

Foster City, CA

- Managed team of 5+ data scientists, Spotfire developers, analysts, and vendors delivering analytics products for Medical Affairs organization; mentored team members, conducted performance reviews, and established professional development programs
- Built centralized data lake consolidating healthcare professional (HCP) and healthcare organization (HCO) reference data from 5+ disparate sources; designed ETL pipelines processing structured and unstructured healthcare data with data quality validation
- Developed automated data processing workflows for HCP/HCO profile enrichment including entity resolution, relationship mapping, and affiliation tracking; reduced manual operations by 70% delivering measurable value to enterprise
- Established data governance framework with formal lineage tracking, quality metrics, and access controls ensuring compliance with healthcare privacy standards and regulatory requirements
- Communicated with senior management (VP/SVP level) ensuring alignment of analytics strategy with organizational goals; presented technical work and business value to executive stakeholders
- Managed multiple concurrent projects across data lake development, dashboard delivery, and vendor integrations; prioritized team resources and resolved blockers in fast-paced healthcare environment

Takeda Pharmaceuticals

May 2021 – December 2022

Associate Director, Digital & Data Science

Lexington, MA

- Led data science and statistical monitoring teams delivering predictive ML models for pharmaceutical manufacturing; managed cross-functional teams including data scientists, engineers, and quality analysts using agile methodologies
- Built production ML models using ensemble methods, PLS, and PCA deployed in manufacturing environment driving real-time quality decisions; established model KPIs tracking both functional performance (accuracy, precision) and non-functional metrics (latency, reliability)
- Designed and implemented digital dashboards for real-time process monitoring integrating sensor data streams with statistical process control algorithms; used data visualization tools to communicate insights enabling operational decisions
- Served as Spotfire system administrator managing enterprise analytics platform including user access, governance standards, and technical infrastructure; established tooling frameworks and best practices across organization
- Led semi-annual APQR (Annual Product Quality Review) process coordinating stakeholders across quality, manufacturing, and regulatory; demonstrated ability to tell stories using data and insights that drive action and change
- Engineered automated data pipelines for ETL processing manufacturing data; implemented data quality checks, error handling, and monitoring ensuring production-grade reliability

Regeneron Pharmaceuticals

May 2016 – May 2021

Senior Process Data Scientist

Rensselaer, NY

- Built end-to-end data science products including automated optimization system for bioreactor feed quantities, predictive quality models, and statistical limit calculators; delivered production systems that changed business decisions and operational planning
- Developed machine learning models using ensemble methods and PLS regression for product quality prediction; designed feature engineering pipelines, cross-validation frameworks, and performance monitoring systems
- Architected data pipelines for manufacturing data ingestion, cleaning, transformation, and storage processing both structured (batch records, sensor data) and unstructured (operator notes, process deviations) data sources
- Implemented automated testing and validation frameworks ensuring model reproducibility and regulatory compliance; established code review practices and version control standards
- Collaborated with cross-functional teams including manufacturing, quality, engineering, and IT; communicated complex ML concepts to non-technical stakeholders and demonstrated ROI through quantified process improvements
- Managed vendor relationships for data acquisition and analytics tool licensing; coordinated with contractors and external consultants on specialized modeling projects

Mathematical Biosciences Institute

2014 – 2016

Columbus, OH

Postdoctoral Research Fellow

- Conducted research applying machine learning and optimization algorithms to biological systems; published 5 peer-reviewed papers and 1 textbook chapter demonstrating thought leadership in computational methods
- Developed novel AI methodologies combining genetic algorithms with mechanistic modeling; established research rigor standards including reproducibility, documentation, and peer review

EDUCATION

Doctor of Philosophy (PhD) in Mathematics

2014

Virginia Polytechnic Institute & State University

Dissertation: Optimization and control of agent-based models in biological systems. Research focus on machine learning algorithms, optimization theory, and computational methods for complex decision-making.

SKILLS & TECHNOLOGIES

Programming & Data Engineering: Python (Expert) • SQL (Expert) • Git • ETL/ELT Pipelines • Data Quality Validation • Automated Testing

Machine Learning & AI: Scikit-Learn • TensorFlow • PyTorch • HuggingFace • OpenAI API • Ensemble Models • PLS • PCA • Random Forest • Generative AI • RAG Systems

Cloud & Data Platforms: Snowflake • Databricks • AWS • Supabase • SQL Databases • NoSQL Databases • Sigma

ML Operations & Deployment: CI/CD for ML • Model Monitoring • Production Deployment • MLOps • Infrastructure as Code Concepts • Automated Testing Frameworks

Data Visualization & Analytics: Spotfire • Custom Python Visualizations • Dashboard Development • Data Storytelling • Executive Reporting

Leadership & Delivery: Agile Methodologies • Team Management • Vendor Management • Stakeholder Communication • OKR Development • Product Management Mindset • Cross-Functional Collaboration

Domain & Governance: Healthcare Data (HCP/HCO) • FDA-Regulated Environments • Responsible AI Frameworks • Data Governance • Regulatory Compliance • Model Interpretability

PRESENTATIONS & PUBLICATIONS

- CHPA Quality and Manufacturing Meeting (2025) — Solo Presenter: AI-driven regulatory intelligence in FDA-regulated industries
- ISPE Annual Meeting & Expo (2025) — Solo Presenter: Digital transformation and AI applications in compliance
- CHPA Regulatory, Scientific & Quality Conference (2025) — Solo Presenter: The Impact of AI on FDA-Regulated Industries
- Redica Systems Webinar (2025) — Solo Presenter: From Reactive to Proactive: Leveraging AI/ML in Quality Operations

PUBLICATIONS

- "Optimization and control of agent-based models in biology: a perspective" — Bulletin of Mathematical Biology, 2017
- "A computational model of invasive aspergillosis in the lung and the role of iron" — BMC Systems Biology, 2016