Mark C. Kuzyk, PhD

mckuzyk@gmail.com · (509) 432-6667 · Spokane, WA · https://www.linkedin.com/in/mark-kuzyk-113198aa/

SUMMARY

PhD Data Scientist with a decade of experience spanning experimental physics, industrial ML, and LLM-powered data pipelines. Specialist in high-dimensional vector modeling, deep neural networks, and scalable NLP systems for complex structured and unstructured datasets. Proven track record delivering revenue-generating analytical products in patent analytics and financial data, with a PhD-level foundation in Quantum Theory enabling rigorous, first-principles approaches to modeling and inference.

WORK EXPERIENCE

Staff Data Scientist   Jun 2024 – Mar 2026

Moat Metrics · Spokane, WA

- Developed a custom deep learning architecture using PyTorch to predict patent entity metrics; implemented a modular pipeline with automated feature scaling, LeakyReLU activations, and inference clamping to ensure realistic value bounds
- Automated valuation scoring of patent portfolios with a scalable inference system capable of handling thousands of unique entities a day, utilizing GPU/CPU-agnostic device configuration and efficient Polars-based data retrieval to optimize processing speeds across compute environments (local, AWS Batch + EC2)
- Engineered LLM pipelines for large-scale patent text processing, including few-shot and chained prompting, structured output extraction, and local LLM inference via vLLM with KV caching to optimize throughput; deployed batch workloads on GCP
- Created an ETL pipeline for a Snowflake database linking patents to government contracts, including parsing and cleaning of malformed patent IDs and contract numbers from heterogeneous sources, resulting in a $250,000 external contract
- Led a high-intensity data wrangling and analysis exercise contracted by a high-profile private equity firm, delivering results that generated $200,000 in revenue above the base contract
- Collaborated on a project constructing a knowledge graph on 5 million active US patents through batch LLM parsing, local LLM calls, and Boolean satisfiability (SAT) algorithms

Research Data Scientist   Sep 2021 – Jun 2024

Aon PLC · Spokane, WA

- Developed an Adaptive Clustering System leveraging a human-in-the-loop workflow to jointly optimize a PyTorch-based deep denoising autoencoder and a multi-class classifier; engineered a training pipeline where user-corrected classifications backpropagate through the latent space, effectively warping patent embeddings to align with expert domain knowledge and improving visualization clarity via PCA and UMAP-based dimensionality reduction
- Built an entity alignment system using Elasticsearch blocking and Conditional Random Fields to probabilistically resolve patent entity identity across heterogeneous data sources
- Implemented graph traversal algorithms to reconstruct patent assignment histories across complex ownership chains

Senior Research Associate (Postdoctoral)   Jun 2019 – May 2021

Duke University · Durham, NC

- Characterized trapped ions via microwave and optical transitions; optimized PID laser stabilization and performed high-sensitivity interferometric vibration analysis of the trapping apparatus
- Mentored 3 graduate students across physics and ECE departments

Senior Research Associate (Postdoctoral)   Jun 2018 – Jun 2019

University of Oregon · Eugene, OR

- Conceived and modeled a quantum network based on diamond defect centers, characterizing vibrational resonances and band structure via COMSOL Multiphysics and building a scanning confocal microscope for defect measurement

Research Assistant (PhD Candidate)   Jun 2010 – Jun 2018

University of Oregon · Eugene, OR

- Built Python simulation libraries for light-matter interactions and designed automated data acquisition systems integrating hardware and software for real-time experiment monitoring
- Directed 4 major projects end-to-end (theory → experiment → publication); contributed to 17 peer-reviewed publications and 2 patents

EDUCATION

PhD in Physics   June 2018

University of Oregon · Eugene, OR   GPA: 3.88

Bachelor of Science in Physics   May 2009

Washington State University · Pullman, WA   GPA: 3.5

SKILLS

Programming & Tools:	Python, SQL, COMSOL Multiphysics, PyTorch, Polars, Numpy, Matplotlib, Scikit-Learn, Statsmodels, DuckDB, Postgres, Snowflake, Elasticsearch, Git, Docker, Virtual Environments and UV, AWS (EC2, S3, Batch), GCP, vLLM, Gitlab CI/CD Pipelines, Terraform (limited)
Data Science & ML:	Deep Neural Networks, LLM Prompt Engineering (zero-shot, few-shot, chaining), Vector Search and Embeddings, Approximate Nearest Neighbors (ANN), Signal Processing, Time-series Analysis, Regression, KMeans and HDBSCAN Clustering, KNN Classification, Conditional Random Fields, Human-in-the-Loop, Active Learning, Joint Optimization, Metric Learning, Manifold Learning, Bayesian Inference
Soft Skills:	Scientific writing (17 peer-reviewed publications with over 2,000 citations), cross-disciplinary collaboration, mentorship, technical communication

AWARDS & HONORS

Marthe E. Smith Memorial Science Scholarship   2012

Weiser Sr Teaching Assistant Award	2016
The Optical Society Quantum Optical Science & Technology Technical Group Outstanding Poster Presentation Award	2017