YAROSLAV RYABOV, PhD

Rockville, MD 20852 | yaroslav.ryabov@gmail.com

LinkedIn: linkedin.com/in/yaroslav-ryabov | Google Scholar: scholar.google.com/citations?user=tCdEY0sAAAAJ | H-index: 27

PROFESSIONAL SUMMARY

Lead Data Scientist with PhD in Theoretical Physics and 10 years delivering AI/ML solutions for financial regulatory and autonomous systems. Unique combination of hands-on experience with AI (LLMs, multi-agent systems, MCP/A2A protocols), fundamental background in theoretical physics and mathematics, and working knowledge of financial regulatory operations including cryptocurrency markets. Led data science for a production AI agent adopted by >95% of its ~200 users and delivering ~$2.5M in measured annual savings; lead a team building an A2A-based multi-agent platform. Expert in multi-agent architectures, prompt optimization, synthetic data generation, and LLM evaluation and monitoring. Combines proven production delivery with strong research foundation (42 papers, 6 patents, H-index: 27).

TECHNICAL SKILLS

AI/ML:	MCP (Model Context Protocol), A2A (Agent-to-Agent), Multi-Agent Systems, Large Language Models (Claude, Llama), Prompt Engineering & Optimization, DSPy, LLM Evaluation & Observability, NLP, Deep Learning, CNNs, Synthetic Data Generation
Programming:	Python, Java, C/C++, SQL, MATLAB
ML Tools:	Google ADK, Arize, HuggingFace, Prodigy, PyTorch, TensorFlow, Scikit-learn, Jupyter, AI Code Assistants
Cloud/Data:	AWS (EC2, S3, SageMaker), Google Cloud Platform, Oracle, MySQL, PostgreSQL
Development:	REST APIs, CI/CD, Git, Agile/Scrum, VS Code, IntelliJ

PROFESSIONAL EXPERIENCE

FINRA | Rockville, MD

Lead Data Scientist | 2022 – Present

Led data science for a production AI agent that analyzes ~70,000 eFOCUS financial reports annually for FINRA's systemic risk monitoring of member firms, enabling shift from selective review to comprehensive analysis of all submissions.

Launched the agent in Q2 2026 to ~200 users with >95% adoption; delivers ~$2.5M in measured annual savings, recovered its development cost in under 3 months, and saw agent-quality issues in only 0.3% of 6,600+ user sessions in the first three weeks.

Created and lead an A2A-based multi-agent platform for synthetic document generation, as technical and data science lead of a 4-person team. Specialized agents own distinct pipeline stages, generating documents from templates with custom labeling schemas for entities and their relationships. Completed MVP phase; production version in development.

Own design of evaluation templates and methodology on the Arize platform, establishing the de facto standard for evaluating and monitoring LLM- and agent-based systems built at FINRA. Earlier developed LLM evaluation framework on FINRA examination data and Cognitive Abilities Test (CogAT), providing evidence that shaped senior management commitment to AI adoption.

Built custom annotation platform on Prodigy and led team of 12 volunteer annotators to create high-quality FINRA-specific dataset from 10 complex regulatory documents (~100 pages each), producing 3,000 labeled entities and 5,000 relationship annotations validated through triple-verification process.

Built LLM-based synthetic data generation platform producing 12 FINRA-specific benchmark datasets (300-1,000 documents each) to enable unbiased evaluation and comparison of AI solution candidates.

In 2026, led 6 AI/ML projects (including eFOCUS), contributed to 4 more, served as independent data science reviewer for 9, and advised 11 run by other teams: 30 projects across FINRA. Mentor junior data scientists and support technical hiring.

Populus Group (Contractor for FINRA) | Rockville, MD

Data Scientist | 2021 – 2022

Built RoBERTa-based transformer model classification platform ensuring consistent examination quality for Written Supervisory Procedure reviews, solving variable precision problem across reviewers (especially junior staff). Developed 11 specialized classification models increasing review accuracy from 69.8% to 91.1% while accelerating throughput 30%.

Designed human evaluation framework to assess classification model performance, validating 7 of 11 models for production deployment.

Intelligent Automation, Inc. | Rockville, MD

Research Scientist | 2016 – 2021

Contributed to FAA-funded research supporting one-million-flights initiative through airspace architecture development and airport traffic optimization. Rescued stalled airport traffic optimization platform by identifying and fixing critical flaws in core MILP formulation design, delivering functional solution simulating air/ground traffic interplay across runway configurations for major hubs (NYC/NJ, Dallas-Fort Worth, DC metro).

Developed geometric multiresolution analysis methods for approximating complex high-dimensional data.

Created optimization algorithms for unmanned aerial vehicle swarm operations using space-filling fractal Peano curves.

Federal Communications Commission | Washington, DC

Senior Software Developer | 2013 – 2016

Developed and maintained enterprise platforms for commercial radio frequency licensing systems, implementing database optimizations and system enhancements to improve transaction processing efficiency. Built simulation tools enabling FCC to model and evaluate new spectrum auction variants before deployment.

Earlier Research Positions

National Institutes of Health (2007–2012) | **Purdue University** (2006–2007) | **University of Maryland** (2003–2006) | **Hebrew University** (1999–2003): Postdoctoral research in computational biophysics, protein structure determination, and theoretical physics.

EDUCATION

Kazan State University | Kazan, Russia

Ph.D. in Theoretical Physics | 1996

M.S. in Theoretical Physics | 1993

PUBLICATIONS, PATENTS & AWARDS

42 papers, H-index: 27 | 6 patents | 21 invited lectures

Preprint: “The Game of Peers: Toward a Credential Algebra for Human-AI Actor Networks” (2026), under review at Journal of Computer Security

Awards: Gates Foundation ($100K, 2011) | NIH Fellows Award (2010) | National Research Council ($230K, 2007–09) | Lady Davis Fellowship ($50K, 1999–2001)

ADDITIONAL INFORMATION

Languages: English (fluent), Russian (native) | **Citizenship:** United States