
ALEX TSANKOV

+1 720-837-8697 — tsankov.jobs@gmail.com

Summary

Self-motivated engineering professional with over 10 years of career experience in software development and site reliability engineering. My greatest strength is the ability to rapidly develop new software with minimal guidance in large institutional environments. I can collect business requirements from technical or non-technical management, architect a solution to achieve the goal, rapidly iterate based on feedback, provide technical leadership, and deploy to customers in a secure/reliable manner.

Specialties: *Distributed System Design, AI/ML Ops, Mission-critical Financial Systems, Hybrid Cloud*

Education

Columbia University - Masters of Public Administration (focus: Global Policy Studies)

University of Colorado, Boulder - B.A. Computer Science & B.A. Philosophy (Dean's List)

Skills

Programming - Python, Go, Javascript/TypeScript, Java, Bash

Systems - Linux, AWS, GCP, Jenkins (CI/CD), Terraform, Kubernetes, OpenStack (VMs), Helm, Argo CD, Docker (Containers), Content Delivery Network (CDN), IAM (Identity), Load Balancers (Nginx / Cloudflare), DNS, Networking, Git, Software Defined Networking, Cryptography, P2P Systems, Prometheus, Grafana, Istio, RBAC

Web Development - Django, Express, Next.js, MySQL, REST, NoSQL, React, Vue, Streamlit

AI & ML - AWS Bedrock/Sagemaker, Azure/OpenAI, Nvidia DGX Superpods, RAG, llama.cpp/gguf, Databricks, “llm-as-a-judge”, Benchmarking systems, Ground-truth testing, “Guardrails” safety systems, MCP Frameworks (FastMCP), AI Agents, Apache Airflow, Model Registry, vLLM, HuggingFace, ONNX

Experience

Bloomberg LP — July 2020 - Present

Senior MLOps Engineer; New York, NY (40 hrs/wk) — December 2023 - Present

Currently working on building out AI Infrastructure for the Bloomberg Law / Bloomberg Tax / Bloomberg Government production products. This involves working with multiple cloud providers for LLM access and services, as well as internal data science teams for deploying classical AI models in the public cloud. The role also involves interfacing with legal experts to evaluate the AI outputs to ensure quality and avoid hallucinations.

- Led development of an internal platform to benchmark LLM models and prompts on proprietary legal tasks, integrating “LLM-as-judge” for subjective metrics and enabling teams to assess whether new models or prompts deliver measurable customer-facing improvements as well as cost savings.
 - Core contributor to internal “LLM-Gateway” platform that enables proxying of LLM inference requests to heterogeneous upstream cloud providers (AWS, Azure, Bare-metal). This platform services live, customer-facing AI products with production SLAs and has strict uptime as well as audit requirements.
 - Rewrote authentication and authorization module, and implemented entire RBAC system for all employees in Bloomberg Industry Group (1000+ users) to dynamically access AI services.
-

- Worked on a FastMCP fork that is Bloomberg compatible (auth, internal service connectivity) out of the box and greatly simplified the process for downstream product teams to enable their services for MCP.
- Built system to enable hundreds of legal experts to perform validation and testing “Gigwork” of LLM legal offerings and aggregate results via an internal ticketing platform to enable faster product evaluations.
- Created a “Bring your own model (BYOM)” system that would allow developers to quickly deploy and scale open-source LLM models or models they trained themselves on AWS Sagemaker. This involved logging, auth, cost tracking, and scaling to multiple GPU instances. Simplified a multi-week process into an hour. Used Databricks model registry.

Senior Site Reliability Engineer (SRE); New York, NY (40 hrs/wk) — July 2020 - December 2023

Led modernization of CI/CD pipelines and spearheaded cloud transformation initiatives within a hybrid financial cloud environment. Acted as a key technical decision-maker, shaping the design and direction of bare-metal compute infrastructure, VM configuration/scaling, and high-reliability deployment practices.

- Drove technical development of “Continuous Deployment” systems for the internal data science team to perform model deployments with. This was used for large machine learning tasks in the development of BloombergGPT. Leveraged Nvidia GPUs (DGX Superpods) and internal data lakes (Hadoop, Spark, Databricks).
- Deployed large scale applications and supported many users with Kubernetes and Helm deployments (large clusters of 30+ machines running production code).
- Supported large bare-metal data centers, Openstack VMs, and interconnects between multiple public clouds (AWS, GCP, Azure, etc.) provisioned with Terraform. Used Ansible for configuration.
- Developed an internal SBOM (Software Bill Of Materials) generator. The tooling I built generates SBOMs for every code deployment, and stores them in a database for vulnerability checking.

Chainalysis — June 2019 - July 2020

Software Engineer; New York, NY (40 hrs/wk)

Chainalysis is the world's leading digital asset analytics company. It builds high-performance investigation and compliance tools for use by law enforcement. I transformed a Java monolith into public cloud micro-services.

- Developed Java modules to monitor illicit digital currency transactions (Used to analyze \$100s of millions of activity). Participated in Docker-izing of production monolithic Java software.
- Developed rapid continuous deployment system for sanctioned entity information (stolen and seized funds).

Paxos — September 2017 - July 2018

DevOps Engineer; New York, NY (40 hrs/wk)

Served the speed and security needs of large financial customers in trading/exchange environments.

- Onboarded clients using FIX connections for sub-millisecond trade execution for precious metals.
- Experience working with Hashicorp Vault for secure secret management in multiple deployments.

Workiva — May 2016 - September 2017

Cloud Infrastructure Engineer; Boulder, CO (40 hrs/wk)

Workiva maintains an online accounting system, similar to Google Sheets, for aggregating company financial data. It is used by the vast majority of the Fortune 100 for their SEC financial filings.

- Utilized serverless execution (AWS Lambda and Codebuild) for maximum scalability.

- Use of "infrastructure as code" (Cloudformation) for proper versioning and reliable deployment.

Google — September 2015 - May 2016

DevOps Apprenticeship; Denver, CO (20 hrs/wk)

- Supported the Youtube global content delivery network (CDN) by automating remote hardware replacement.

CU Boulder's "Next Generation Networks" Lab Group — January 2014 - May 2016

Student Researcher; Boulder CO (15 hrs/wk)

- Supported Prof. Eric Keller and his "Next Generation Networks" lab in the development of A software defined network (SDN) controller module, to integrate with the Bro intrusion detection system (IDS).

Certifications

CompTIA Security+ SY0-701 (2025)

AWS Certified Solutions Architect Associate SAA-C03 (2024)

Cisco Certified Network Associate CCNA (2011) - Currently Expired

CompTIA Network+ (2010) - Currently Expired

Achievements

Team Lead for Columbia Capstone Project on Semiconductor Policy with CISA (2023) - Led a group of three other students and conducted numerous interviews of international policy makers and semiconductor industry figures to publish *Rebuilding America's Semiconductor Industry: How State Governments Can Drive Domestic Capacity and Maximize the CHIPS Act*. I gained valuable expertise in critical supply chain policy and economic considerations. Once published, the team briefed officials at Cybersecurity and Infrastructure Security Agency (CISA), part of DHS, on the findings of our research.

Semi-Finalist, Cyber 9/12 Strategy Challenge (2022) - Led a team to devise a response to a hypothetical cyberwar scenario in South Asia. Hackathon format that fused cyber and strategic challenges in a competition format, where current and former high-ranking government officials scored our responses to the scenario.

Invited to Present at San Jose NFV World Congress - Association for Computing Machinery (2016) - Published: Matthew Monaco; Alex Tsankov; Eric Keller. *Taking the Surprise out of Changes to a Bro Setup*. 6th ACM Conference on Data and Applications Security and Privacy (CODASPY-2016). New Orleans, LA.

Projects

Co-Creator of farmsdirectory.com (2024) - Worked with farms in the New York Hudson Valley to build software that enables consumers to find nearby sustainable farms and restaurants who carry locally-sourced, organic food.

Created go-live (2020) - Built a fast, portable, ultra light-weight (<5MB) Go command line utility that serves a file directory over HTTP. Used for local web development, production static-site serving, or as a network file host.

Co-Created CUFCQ.com (2015) - Designed and built a popular open-source Ruby on Rails site (2,000+ users/month) that summarizes and visualizes University of Colorado professor evaluation statistics. This project involved scraping and cleaning large amounts of Excel data (XSLX) to make it accessible for visualization.
