

# SAHAJ SHUKLA

1-201-936-5159 | <https://www.linkedin.com/in/sahajshukla> | [sahajshukla@gmail.com](mailto:sahajshukla@gmail.com) | <https://github.com/sahajshukla> | [Portfolio Link](#)

---

## CAREER SUMMARY:

Audit Automation Engineer with 3+ years designing SOX-compliant data, analytics, and compliance solutions for global finance and SaaS. Expert in Spark, Databricks, cloud microservices, and AI-driven GRC platforms. Built Vantage AuditOS, a zero-trust system delivering real-time, encrypted audit automation. Proven leader in technical delivery and process modernization. Open to new opportunities and available to start immediately.

## SKILLS:

**Machine Learning & Data Engineering:** Python, SQL, Spark, Databricks, Pandas, NumPy, PyTorch, TensorFlow, OpenSQL-7B, NLP Tools  
**Audit, Finance & Analytics:** Technology Audit, SOX/ITGC Compliance, Risk & Financial Analytics, Financial Modeling – CAPM, Monte Carlo, Time Series, Black-Scholes; Dashboarding – Power BI, Tableau, Streamlit, Risk Management, GRC Auditing, Soc 2 type 2 audit testing  
**Data Systems & Storage:** PostgreSQL, Snowflake, SQL Server, MongoDB, Redis, DynamoDB, Object Stores, ETL, Schema Design, Data Models  
**DevOps & Collaboration:** Terraform, Docker, Kubernetes, Azure DevOps, Bash, Agile/Scrum, Jira, Cross-Functional Collaboration, Cronjobs  
**Cloud Tools:** AWS (S3, Lambda, EC2), Azure (Data Factory, Blob Storage), Apache Airflow, Apache Kafka, AWS Kinesis, Event-Driven ETL

## EXPERIENCE:

### Software Engineer 2 | BDIPlus | New York, New York

January 2025 - Present

- Led development of a contact extraction platform (CCMP) for AvidXchange, integrating six data sources including EDW, APIs, and Azure Blob Storage, and migrating the system to Spark Databricks, increasing supplier-contact coverage from 19 to 73 percent.
- **Built** an NLP pipeline using **Azure OpenAI** and **Gemma 3-8B**, achieving **97% accuracy** through deduplication and standardization.
- Deployed **Terraform-based Azure** infrastructure with secure Azure Vault credentialing, parallelized Spark workloads for millions of records, and optimized performance through batching and resource tuning.
- **Oversaw** delivery for a team, managing sprints, client communications, and demos with exception tracking through automated logs.
- **Developed** and maintained backend APIs for ChatCDP, supporting enterprise clients (**Morgan Stanley and American Express**), enabling schema-agnostic query generation across PostgreSQL, MySQL, and Oracle databases using the Neilus JSON-function architecture.
- **Migrated** **Afficiency's** infrastructure and fourteen production cronjobs to BDIPlus, designing a unified scheduling and failure-monitoring service that improved reliability and reduced incident turnaround time.

### Associate, Finance Audit and Data Analytics | Nomura Holding Americas | New York, New York

April 2024 – Jan 2025

- **Led** modernization of audit testing by engineering a fraud-detection model analyzing 160K+ journal entries (~\$466B USD), uncovering control anomalies through statistical and rule-based analysis and establishing data-driven review procedures across Finance Audit.
- **Initiated** cross-team data-access framework linking Finance and Risk datasets, enabling faster exception testing and improving transparency for regional audit teams.
- **Automated** end-to-end sampling and exception testing across Financial Reporting, Collateral Management, and Dodd-Frank audits, standardizing analytics logic and cutting manual effort by ~30 %.
- **Designed** interactive PowerBI dashboards consolidating five years of loss data to identify control failures and inform remediation strategy.
- **Partnered** with global audit teams to define risk metrics and automate validation workflows, improving testing and data lineage.
- **Championed** cross-functional data sharing standards between Finance and Risk functions, initiating centralized audit analytics.

### Analyst, Technology Audit | BlackRock Financial Management LTD. | New York, New York

August 2021 – January 2024

- **Automated** IT control testing with cronjobs, reducing manual effort by ~6 hours per control and validated across global audit teams.
- **Built** a real-time dashboard for audit resourcing, giving management visibility into 200+ active audits resource allocations.
- **Evaluated** design and operating effectiveness of 65+ IT controls, coordinating remediation in alignment with GRC framework.
- **Co-authored** 9 cross-domain audit reports and resolved seven high-impact issues, improving control assurance for systems.
- **Developed** quantitative risk models using AUM and incident data to prioritize annual audit coverage and guide technology audit planning.
- **Recognized for** introducing process-level automation standards and reusable Python libraries adopted across regional audit teams, enhancing testing consistency.

## PERSONAL PROJECTS:

### Vantage - AI Audit Automation Platform:

- **Built** a zero-trust audit analytics platform automating control testing, evidence generation, and SOX-compliant workpapers.
- **Designed** a FastAPI + Next.js microservice stack with PostgreSQL and 30 indexes, achieving sub-500 ms query latency.
- **Integrated** a self-hosted OpenSQL-7B model fine-tuned for deterministic, template-based SQL generation and validation.
- **Implemented** AES-256 and SHA-256 key isolation to encrypt all financial data at rest and in transit.
- **Developed** an autonomous scheduler and in-memory cache for recurring tests, alerts, and real-time monitoring.
- **Created** a multi-level approval and commenting workflow ensuring full audit-trail visibility and accountability.

## EDUCATION:

MS in Business Intelligence and Analytics | Stevens Institute of Technology | Hoboken, NJ.  
Bachelor in Electronics and Telecommunication Engineering | University of Mumbai | India.

August 2019 - June 2021  
August 2015 - May 2019