**DAVID RUBER**

**Email: davidrubera11@gmail.com Phone: (315) 625-6859**



**PROFILE SUMMARY**

* Results-driven Data Engineer with a decade of expertise in Data engineering across cloud platforms with a total of 12 years in IT.
* Extensive experience utilizing Google Cloud Platform (GCP) services, including BigQuery, Dataflow, Dataprep, and Pub/Sub, for data engineering solutions.
* Proficient in building and managing GCP data pipelines with tools like Cloud Composer and Cloud Dataflow.
* Proven ability in developing and deploying applications on Google Kubernetes Engine (GKE).
* Strong background in implementing security and compliance on GCP, ensuring data privacy and regulatory adherence.
* Track record of optimizing cost and resource usage within GCP environments.
* Skilled in AWS services such as Amazon EMR, Redshift, and Glue for efficient data processing.
* Expertise in architecting scalable, cost-effective solutions on AWS, with proficiency in configuring AWS Lambda for serverless computing.
* Adept at setting up AWS Kinesis streams to process real-time data, enhancing system responsiveness and data-driven decision-making.
* Proficient in leveraging AWS DynamoDB to create scalable, low-latency NoSQL databases for dynamic applications.
* Deep expertise in optimizing and managing Amazon Redshift data warehouses to deliver high-performance analytics and business insights.
* Experienced in integrating AWS services into CI/CD pipelines, streamlining automation for continuous integration, delivery, and deployment.
* Skilled in setting up and securing AWS Virtual Private Cloud (VPC) environments.
* Knowledgeable in Azure services, including Azure Data Lake Storage and Azure Databricks, for data storage and analytics.
* Proficient in managing Azure virtual machines (VMs) for cloud infrastructure operations.
* Extensive experience managing on-premises data infrastructure, including data warehouses and databases.
* Familiar with AWS DevOps practices for continuous integration and deployment.
* Expertise in using Git for version control in DBT projects, ensuring proper tracking and documentation of data model changes.
* Skilled in performance optimization and tuning of on-premises data systems.
* Proficient in data migration strategies between on-premises and cloud environments.
* Strong troubleshooting skills in resolving issues within on-premises data systems.
* Proven ability to maintain high availability and disaster recovery solutions in on-premises environments.
* Experienced in implementing CI/CD pipelines using tools like Jenkins and GitLab CI/CD.
* Adept in automated testing processes, including unit, integration, and regression testing.
* Skilled in gathering and analyzing project requirements to ensure alignment with business goals.
* Experienced in Agile project management, contributing to successful outcomes through data-driven analytics and collaborative teamwork.

**TECHNICAL SKILLS**

**Languages & Scripting:** Spark, Python, Java, Scala, Hive, Kafka, SQL

**Cloud Platforms:** AWS, GCP, Azure

**Python Packages:** NumPy, TensorFlow, Pandas, Scikit-Learn, Matplotlib, Seaborn

**Databases:** Cassandra, HBase, Redshift, DynamoDB, MongoDB, SQL, MySQL, Oracle, RDBMS, Amazon RDS, DynamoDB

**Data Ingestion & Streaming:** AWS Kinesis, AWS IoT Core, Azure Data Lake, Google Pub/Sub

**Data Analysis:** Data Modeling, Statistical Analysis, Sentiment Analysis, Forecasting, Predictive Modeling

**Project Management:** Jira, Kanban

**Data Warehousing & Analytics:** Redshift, Snowflake, BigQuery, Teradata, Oracle, Azure SynapsDB

**Data Integration & ETL:** AWS Glue, Aws EMR, DMS, Apache Nifi, Spark, Azure Data Factory, Data Bricks, Azure HdInsight, Google Data Flow, Google Data Proc

**Serverless Computing:** AWS Lambda, Fargate, Batch, GCP Data Fusion, Functions, GKE, Step Function

**Data Governance & Security:** AWS IAM, KMS, Secrets Manager, GCP IAM, Security Command Center

**CI/CD:** AWS Code Pipeline, Code Build, Code Deploy, Azure DevOps, Jenkins, Terraform, CloudFormation, Docker, Kubernities

**Cluster Management:** Cloudera Manager, Apache Ambari

**Data Visualization:** PowerBI, Tableau, Aws QuickSight

**Search & Indexing:** Elasticsearch, Kibana

**WORK EXPERIENCE**

**Lead AWS Data Engineer Sep 2023 – Present**

**ETSY, Dumbo, Brooklyn**

I spearheaded the design and implementation of robust data pipelines on AWS. By leveraging services like IoT Core, Data Pipeline, Redshift, Athena, and DBT, I optimized data ingestion, transformation, and analysis processes. I also developed interactive dashboards with QuickSight and utilized EMR for large-scale data processing. Additionally, I implemented scalable and secure data storage solutions using S3 and Glue Data Catalog, while ensuring efficient deployments with CodeDeploy and CodeCommit.

* Designed and implemented secure and reliable data ingestion pipelines from IoT devices using AWS IoT Core.
* Automated data movement and transformations with AWS Data Pipeline, optimizing data processing workflows.
* Leveraged Amazon Redshift for high-performance data warehousing, enabling efficient querying and analytics.
* Authored and executed queries in Amazon Athena to analyze data stored in S3, supporting data profiling and ad-hoc reporting.
* Developed DBT models to support incremental data loads, improving processing efficiency in data pipelines.
* Created interactive data analytics dashboards using Amazon QuickSight, enabling insightful visualizations.
* Used Amazon EMR for distributed data processing at scale, utilizing Hadoop and Spark.
* Designed and managed scalable and cost-effective data storage solutions using Amazon S3 for both structured and unstructured data.
* Utilized AWS Glue Data Catalog for efficient metadata management and data discovery, enhancing data governance.
* Deployed applications to compute services through AWS CodeDeploy for streamlined deployments.
* Managed source code repositories and version control with AWS CodeCommit for efficient team collaboration.
* Utilized Apache Spark with Python (PySpark) to perform distributed data processing, optimizing performance for large-scale data analytics tasks.
* Delivered scalable data solutions with Snowflake, driving data-driven insights and operational efficiency.
* Demonstrated extensive experience in cloud-based data warehousing, leveraging Snowflake to optimize performance through data modeling and query optimization.
* Maximized Snowflake’s capabilities in modern data ecosystems, enabling streamlined workflows and contributing to data-driven decision-making.
* Implemented Amazon EBS for block-level storage, ensuring reliable data management for EC2 instances.
* Leveraged Amazon EFS for scalable shared file storage, enabling seamless collaboration and data access.
* Engineered real-time data ingestion and streaming solutions using Amazon Kinesis for fast and efficient data processing.
* Orchestrated workflows using AWS Step Functions to automate ETL pipelines and integrate various AWS services.
* Developed serverless functions with AWS Lambda for event-driven automation and data processing.
* Managed containerized applications with AWS Fargate, ensuring scalability and resource optimization.
* Handled batch processing workloads with AWS Batch, optimizing resource allocation and scheduling.
* Implemented IAM policies to enforce granular access control, ensuring data security and compliance.
* Safeguarded sensitive data using AWS KMS for encryption, ensuring security across services.
* Managed secrets and credentials securely using AWS Secrets Manager, ensuring automated rotation.
* Enhanced data discovery, classification, and security compliance using AWS Macie.
* Managed and optimized relational databases with Amazon RDS, ensuring high availability and reliability.
* Implemented NoSQL solutions with Amazon DynamoDB, providing flexible, high-performance database services.
* Leveraged Amazon Aurora for high-availability relational database setups, ensuring low-latency access.
* Utilized Amazon DocumentDB for scalable NoSQL document database solutions supporting complex data models.
* Managed time-series data storage and analytics efficiently with Amazon Timestream for real-time insights.
* Designed and executed automated ETL processes with AWS Glue, enhancing data quality and transformation efficiency.
* Used AWS DataSync for seamless data migration and synchronization, ensuring efficient data transfers.
* Conducted database migrations with AWS DMS, minimizing downtime while ensuring data consistency.
* Established CI/CD pipelines with AWS CodePipeline for continuous integration and efficient code deployment.
* Automated source code compilation and builds using AWS CodeBuild.

**Lead Data Engineer Jan 2022 – Aug 2023**

**HSBC Bank, New York City, NY**

I built and maintained scalable data pipelines on GCP, leveraging services like Dataflow and DataProc. I designed efficient data storage strategies using GCS and Bigtable, ensuring data accessibility and reliability. Additionally, I developed ETL processes and integrated data from various sources using Dataflow, DataProc, and Python, ensuring data quality and accuracy.

* Built and maintained scalable data pipelines on GCP, ensuring efficient data processing, elasticity, and cost management through services like Google Cloud Dataflow and DataProc.
* Designed data storage strategies using Google Cloud Storage (GCS) and Bigtable, ensuring data accessibility, reliability, and cost-efficiency, while adhering to GCP best practices.
* Developed API integrations within GCP, utilizing Pub/Sub for data ingestion, facilitating smooth data flow, and handling increasing data volume and complexity.
* Developed ETL processes on GCP using Dataflow and DataProc, ensuring efficient data transformation, processing large datasets, and adhering to GCP’s data quality standards.
* Continuously monitored and enhanced data processing workflows on GCP for performance, scalability, and cost-effectiveness, utilizing tools like Dataflow and DataProc.
* Implemented robust security measures, including IAM policies on GCP, to safeguard sensitive data and ensure compliance with privacy regulations and GCP security standards.
* Developed and maintained scalable ETL processes using Snowflake, ensuring efficient data ingestion from various sources, including AWS S3, Azure Blob Storage, and Google Cloud Storage.
* Designed and implemented ETL (Extract, Transform, Load) processes using Python to integrate data from diverse sources, ensuring data quality and accuracy.
* Designed and optimized Snowflake data models to support analytical workloads, enhancing query performance through effective use of clustering keys and micro-partitioning.
* Established and maintained data governance practices on GCP, ensuring data quality, integrity, and consistency across platforms, using BigQuery for reporting and analysis.
* Designed and deployed data architecture solutions closely aligned with business needs, utilizing a comprehensive range of Google Cloud Platform (GCP) services such as Google Cloud Dataflow, Pub/Sub, BigQuery, Bigtable, DataProc, and IAM to optimize data workflows.
* Leveraged GCP services including Dataflow, Pub/Sub, BigQuery, Bigtable, and DataProc, combined with expertise in Hadoop, Spark, Hive, and Sqoop, to create and refine cloud-based data processing workflows tailored for GCP’s infrastructure.
* Created and optimized data systems and pipelines within GCP, maximizing performance and scalability through services like Dataflow, BigQuery, and Bigtable.
* Conducted complex data analysis leveraging GCP’s analytics services, particularly BigQuery, to generate actionable insights and support informed decision-making.
* Integrated data from various sources into a centralized platform on GCP, designed for streamlined reporting and analysis using BigQuery and Bigtable.
* Collaborated with data scientists and analysts by providing well-structured, clean datasets hosted by GCP, enabling advanced analytics and machine learning projects powered by BigQuery and Dataflow.
* Optimized cost and performance for data storage, processing, and analytics on GCP, effectively utilizing services like BigQuery, Dataflow, and Bigtable.
* Collaborated with stakeholders to understand data requirements and deliver tailored solutions within the GCP environment using appropriate services.
* Stayed up to date with the latest advancements in big data technologies, focusing on the evolving GCP ecosystem.
* Managed and prioritized multiple GCP-based data engineering projects, ensuring timely delivery and cost-effective solutions.
* Effectively communicated technical concepts and findings to both technical and non-technical stakeholders, supporting data-driven decisions within the GCP environment.
* Developed real-time data processing solutions using Pub/Sub and Dataflow within GCP to support near real-time decision-making capabilities.
* Mentored junior team members on GCP services, offering technical guidance and keeping the team updated on GCP best practices and emerging trends.
* Documented data pipelines, workflows, and GCP best practices to facilitate knowledge sharing and support across the team.
* Resolved complex data challenges and troubleshooted issues during data processing and analysis, leveraging GCP's wide array of tools and services.

**Sr. Data Engineer Sep 2020 – Dec 2021**

[**AstraZeneca**](https://www.bing.com/ck/a?!&&p=9bd7391ccf0eba06JmltdHM9MTcyOTcyODAwMCZpZ3VpZD0wYWUwZTU1ZC05NGVlLTZlZTQtM2Q3Ni1mNzFkOTUxYzZmZjcmaW5zaWQ9NjA2Mw&ptn=3&ver=2&hsh=3&fclid=0ae0e55d-94ee-6ee4-3d76-f71d951c6ff7&psq=WHERE+IS+Astra+Zeneca+LOCATED+IN+AMERICA&u=a1aHR0cHM6Ly93d3cuemlwcGlhLmNvbS9hc3RyYXplbmVjYS1jYXJlZXJzLTE1NjA5L2xvY2F0aW9ucy8&ntb=1)**, Wilmington, DE**

I designed a robust data warehouse architecture on Azure Synapse Analytics, enabling complex data analysis. I also contributed to the development of a data lake on Azure Data Lake Storage, enhancing data accessibility and compatibility. Additionally, I automated tasks and improved operational efficiency by developing Azure Functions and leveraging Azure HDInsight for big data processing.

* Designed a robust data warehouse architecture on Azure Synapse Analytics and executed complex data analysis queries within the Azure ecosystem.
* Contributed to the development and integration of a data lake on Azure Data Lake Storage, enhancing compatibility with various applications and development projects.
* Automated tasks and improved operational efficiency by developing Azure Functions using Python within the Azure cloud environment.
* Implemented Azure HDInsight to process big data across Hadoop clusters, leveraging Azure Virtual Machines and Azure Blob Storage for optimal performance.
* Created Spark jobs that seamlessly executed in HDInsight clusters using Azure Notebooks, streamlining data processing.
* Developed efficient Spark programs in Python for HDInsight clusters, optimizing data processing capabilities.
* Successfully deployed the ELK (ElasticSearch, Logstash, Kibana) stack in Azure, facilitating website log collection and analysis.
* Designed and implemented robust ETL pipelines using tools such as Apache NiFi, Talend, or Informatica, facilitating data ingestion from diverse sources, including relational databases, APIs, and flat files.
* Ensured code quality and reliability by implementing comprehensive unit tests using frameworks like PyTest.
* Architected serverless solutions using Azure API Management, Azure Functions, Azure Storage (Blob), and Azure Cosmos DB, achieving performance optimization with auto-scaling features.
* Designed schemas, cleaned input data, processed records, formulated queries, and generated output data with Azure Synapse Analytics, streamlining data management and analysis.
* Enhanced data warehousing capabilities by efficiently populating database tables using Azure Stream Analytics and Azure Synapse Analytics.
* Developed User Defined Functions (UDF) in Scala to automate business logic, improving application efficiency.
* Designed Azure Data Factory pipelines to ingest, process, and store data, integrating seamlessly with other Azure services.
* Executed Hadoop/Spark jobs on Azure HDInsight with data stored in Azure Blob Storage, enabling scalable big data processing.
* Created Azure Resource Manager (ARM) templates building custom infrastructure for pipelines, optimizing resource management.
* Implemented Azure Active Directory (Azure AD) roles, instance profiles, and policies for secure user authentication and access control, ensuring compliance.
* Leveraged Azure Data Lake Storage for scalable and available data lake architecture.
* Developed Azure Functions for serverless automation and task execution within the cloud environment.
* Utilized Azure HDInsight for efficient big data processing and analytics.
* Designed data integration workflows with Azure Data Factory, enabling seamless data movement and transformations.
* Implemented Azure AD for identity and access management, ensuring secure authentication similar to AWS IAM.
* Proficient in DBT (Data Build Tool) for transforming and modeling data, optimizing analytics workflows.
* Extensive experience using DBT to streamline data transformations and enhance data quality.
* Skilled in leveraging DBT to create structured data models, enabling more effective analysis.
* Expertise in implementing version-controlled, modular data transformations with DBT for scalable data pipelines.
* Strong understanding of DBT's role in modern data stack architecture, facilitating efficient data processing.
* Leveraged DBT for data transformations, contributing to data-driven insights and decision-making.

**Data Engineer Jan 2019 – Aug 2020**

**Edisson International, Rosemead, CA**

I integrated Business Intelligence tools like Tableau and Power BI with the data warehouse for insightful data visualization. I developed and tested Spark SQL scripts to manage and process datasets efficiently. Additionally, I utilized Kafka for real-time data processing and Sqoop for data ingestion, ensuring seamless data flow and timely analysis.

* Integrated Business Intelligence tools like Tableau and Power BI with the data warehouse for seamless data visualization.
* Developed and tested Spark SQL scripts to manage datasets, monitoring job performance through the Spark UI.
* Used Spark to filter, format, and store data in the Hive warehouse, ensuring efficient data handling.
* Created Hive tables to store and manage data from multiple sources, maintaining the Hive metastore for metadata management.
* Developed Kafka Connect-based data pipelines to integrate data from various sources into Kafka topics for real-time processing.
* Automated daily ETL processes with bash scripts and Cron jobs to streamline data ingestion.
* Utilized Kafka Streams for real-time data transformation and enrichment directly within the Kafka ecosystem.
* Imported data from local file systems and RDBMS into HDFS using Sqoop, automating workflows with shell scripts.
* Evaluated Hadoop-based data processing techniques to detect anomalies within datasets.
* Implemented a streaming job using Apache Kafka to ingest data from REST APIs.
* Employed Gradient Boosted Trees and Random Forests to establish a benchmark for accuracy in predictive models.
* Utilized DBT to transform and model data efficiently, ensuring structured and repeatable processes.
* Designed HiveQL and SQL queries to extract data from the data warehouse and create user-friendly views for consumption.
* Processed input data by defining schemas, writing UDFs, and generating output using Hive, while also cleaning and organizing the records.
* Created and optimized Snowflake schemas (both star and snowflake) to enhance query performance and simplify data retrieval.
* Worked with NoSQL databases such as MongoDB or Cassandra, using Python for efficient storage and retrieval of unstructured data.
* Leveraged Snowflake's elastic scalability to manage increasing data workloads and facilitate seamless data expansion.

**Data Engineer Mar 2016 – Dec 2018**

**Chevrolet, Detroit, Michigan**

I collaborated on an AWS data engineering project, utilizing services like Glue, Lambda, Step Functions, Python, and Java to build end-to-end data pipelines for cloud migration. I employed AWS Glue Crawlers to automatically discover and catalog metadata, and integrated AWS Fully Managed Kafka for real-time data transfer to Spark clusters in Databricks. Additionally, I leveraged AWS Redshift and Redshift Spectrum for secure cloud-based data storage, ensuring scalability and accessibility while supporting data migration.

* Collaborated on an AWS data engineering project, utilizing services such as AWS Glue, Lambda, Step Functions, Python, and Java to build end-to-end data pipelines for cloud migration.
* Employed AWS Glue Crawlers to automatically discover and catalog meta data from data sources like S3, RDS, and other data stores.
* Scheduled Glue Crawlers to periodically refresh the data catalog, ensuring up-to-date metadata management.
* Integrated AWS Fully Managed Kafka streaming solutions for real-time data transfer to Spark clusters within AWS Databricks.
* Successfully migrated data from on-premises SQL Servers to Amazon RDS and EMR Hive, optimizing data management and facilitating seamless cloud migration.
* Utilized AWS Redshift and Redshift Spectrum for secure cloud-based data storage, ensuring scalability and accessibility while supporting data migration.
* Managed AWS resources, including EC2 instances and Hadoop clusters, to optimize performance during the data migration process.
* Leveraged PySpark for efficient data ingestion from various sources, encompassing both structured and unstructured financial data.
* Engineered and maintained a Hadoop Cloudera distribution cluster on AWS EC2, enhancing data processing capabilities to support migration initiatives.
* Developed AWS Lambda functions in Python and Java to execute specific tasks within the data pipeline, such as triggering Glue jobs and monitoring pipeline health.
* Utilized Spark SQL and the DataFrames API for efficient data loading into Spark clusters, particularly for data migration projects.
* Created ETL (Extract, Transform, Load) jobs using AWS Glue ETL jobs written in Python, Java, and Scala, employing built-in transformations and custom scripts to clean and transform data as required.
* Defined workflows that integrated Glue jobs, Lambda functions, and other AWS services, optimizing the data processing pipeline.
* Demonstrated expertise in data manipulation and analysis using Python, Java, SQL, and Snowflake, essential for data migration and analysis tasks.
* Leveraged Snowflake, Snowpipe, and Redshift Spectrum for effective data processing and analysis during migration.
* Utilized PySpark libraries to build scalable, high-performance data processing applications.
* Designed and implemented AWS Lambda functions for serverless data processing, optimizing execution times, memory allocation, and concurrency settings critical to migration workflows.
* Orchestrated complex data workflows using AWS Step Functions, enhancing reliability and automation in migration processes.
* Managed end-to-end data collection, processing, and analysis using Kinesis services, supporting data migration efforts.
* Implemented real-time data streaming solutions with Amazon Kinesis, enabling timely data collection and analysis.
* Proficiently handled Amazon DynamoDB, a highly scalable NoSQL database service, to meet data migration and storage requirements.
* Demonstrated expertise in database modeling and design for DynamoDB, crucial for effective migration strategies.
* Utilized AWS CodePipeline for continuous integration and continuous deployment (CI/CD) in data migration workflows.
* Designed and optimized data warehousing solutions using AWS Redshift, leveraging its capabilities for high-performance analytics and migration tasks.
* Integrated AWS Redshift with various AWS data services, streamlining workflows, including migration pipelines.
* Implemented and optimized data transformations, aggregations, and analytics using functional programming principles in Scala.
* Demonstrated proficiency in cloud-based data warehousing using Snowflake, leveraging its multi-cluster, shared data architecture for efficient migration.
* Effectively separated storage and compute in Snowflake to enhance scalability for migration and analysis.
* Leveraged AWS CloudWatch for real-time monitoring and troubleshooting during data migration processes.
* Utilized AWS CloudFormation for automated cloud resource provisioning, ensuring efficient setting up of data migration environments.
* Implemented data transformations using DBT (Data Build Tool), essential for migration and transformation projects.

**Hadoop Engineer Jan 2014 – Feb 2016**

**Health Net LLC., Woodland Hills, Los Angeles**

I managed a diverse range of datasets within the Hadoop environment, optimizing data processing and analysis. I improved data processing speed by integrating and optimizing Hive, Sqoop, and Flume, streamlining ETL workflows. Additionally, I developed a dynamic data warehousing solution using Hive, enabling detailed analytics for client-based transit systems.

* Managed a diverse range of datasets, from unstructured to structured data, within the Hadoop environment to ensure efficient data processing and analysis.
* Improved data processing speed by integrating and optimizing Hive, Sqoop, and Flume into existing ETL workflows, streamlining extraction, transformation, and loading processes.
* Developed a dynamic data warehousing solution using Hive, enabling detailed analytics for client-based transit systems.
* Handled various data formats such as JSON, XML, CSV, and ORC, and implemented Hive partitioning and bucketing for optimized data organization and retrieval.
* Managed the full lifecycle of Hadoop clusters, including installation, node commissioning/decommissioning, high availability configuration, and capacity planning to ensure seamless operations.
* Executed cluster upgrades on staging platforms before production deployment to minimize disruptions and ensure system stability.
* Used Cassandra for processing JSON-documented data and HBase for storing region-based data, addressing diverse data needs effectively.
* Configured and managed Zookeeper and ZNodes for high availability, contributing to a fault-tolerant Hadoop infrastructure.
* Implemented Apache Ranger for access control and audits, ensuring compliance with security protocols and regulatory standards.
* Performed HDFS balancing and fine-tuning to enhance MapReduce application performance, improving data processing efficiency.
* Designed and executed data migration plans for seamless integration of new data sources into the Hadoop ecosystem, centralizing data management.
* Streamlined cluster setup and management using open-source tools like Puppet, Java, and Python for configuration and deployment.
* Enhanced security with Kerberized authentication, ensuring secure user access within the Hadoop environment.
* Customized YARN Capacity and Fair schedulers to optimize resource allocation and prioritize job execution.
* Provided insights on cluster capacity and growth planning, aiding in decisions regarding node configuration and resource allocation.
* Optimized MapReduce counters to expedite data processing and improve performance in data-intensive operations.
* Designed and implemented robust backup and disaster recovery strategies for Hadoop clusters, ensuring data resilience and business continuity.
* Executed upgrades, patches, and fixes on Hadoop clusters using rolling or express methods to minimize downtime and maintain system stability.

**Software Engineer Jan 2012 – Dec 2013**

**SecureWorks, Atlanta, GA**

I worked on the internal ticketing system, focusing on API development, feature generalization, and stability improvements. I built a pub-sub-API using RabbitMQ to enable real-time notifications for various clients and modernized the ticket distribution engine using a distributed Redis cache. Additionally, I improved system instrumentation by establishing logging and metrics standards and developing comprehensive monitoring dashboards.

* Worked on the internal ticketing system and related workflow management tools, focusing on API development, genericizing bespoke features, and improving stability.
* Built a published-subscribed API for tickets using RabbitMQ, enabling clients to be notified of events on tickets of interest. Managed over 60 subscribers across engineering for automated integrations.
* Modernized the distribution engine for routing tickets to available representatives by building a client of the pub-sub-API. This client uses user-configurable routing rules and a distributed Redis cache for horizontal scaling.
* Improved instrumentation across systems by establishing logging and metrics standards, creating actionable alarms, and developing generic Splunk log dashboards and Grafana metric dashboards for service health monitoring.
* Advocated for Python adoption within the team, provided documentation for transitioning from C++, and collaborated on establishing Python (and C++) standards for CI/CD enforcement.

**EDUCATION**

**Master of Science in Business Analytics (MSBA)**

Mory University, Goizueta Business School, Atlanta, GA

**Master of Science in Electrical and Computer Engineering**

Carnegie Mellon University, School of Engineering, Pittsburgh, Pennsylvania