




Scott A Lyons

Senior Engineer
Seattle, Washington

 scottalyons
 theslyons
 slyons

EXPERIENCE

• Spice.ai

Senior Engineer

Seattle, Washington (Remote)

September 2024 - November 2024

- Contributed to the development and enhancement of the Spice.ai open-source product, focusing on integrating key functionality to improve usability and extend product capabilities.
- Authored comprehensive technical documentation, including API references and deployment guides, to support developer adoption and streamline product onboarding.
- Created and recorded instructional video content, explaining core product features, configuration steps, and usage scenarios to enhance user experience and facilitate knowledge transfer within the developer community.
- Utilized core design principles to guide content creation, ensuring that all resources provide actionable insights and adhere to Spice.ai's strategic goals and technical standards

Technologies: Rust, Apache DataFusion, Spark, AWS, Azure, Data Catalogs, Axum, Kubernetes

• Prophecy.io

Staff Data Engineer

Seattle, Washington (Remote)

April 2022 - September 2024

- Drove the development of customer data pipelines using our product
- Contributed to the code generation aspect of the product
- Built dozens of custom reusable code components for customers and internal projects
- Built out the first and subsequent versions of the Product documentation
- Worked with customers to migrate existing legacy ETL projects to a more modern Big Data ecosystem
- Helped create best Data Engineering practices both internally and with customers

Technologies: Python, Databricks, Scala, Spark, Terraform, Kubernetes, Azure Data Factory, Azure Data Lake

• Microsoft - Consumer Software Engineering (CSE)

Software Engineer 2

Redmond, Washington

September 2018 - April 2022

- Member of the Data Tech Domain team gathering industry trends from various Customer engagements and projects
- Responsible for educating customers and field-oriented developers on industry trends related to Big Data.
- Collected feedback from field teams, triaged it and reported it back to Product groups
- Worked closely as part of the Azure Databricks feedback loop to expedite customer issues and drive product planning based on industry trends
- Converted R scripts to SparkR and encapsulated them for execution through Azure Data Factory (H&M)
- Designed Data Architecture and built the initial Spark Streaming version of near-realtime ML models (Energinet)
- Wrote benchmarking tests dealing with handling encrypted files in a Big Data pipeline (Walgreens)
- Wrote initial Spark/HBase implementation for distributed event handling through Event Grid (AT&T)
- Served as an emergency consult on scripting and validating HDInsight cluster deployments (Credit Suisse)
- Helped port Feast(<https://github.com/feast-dev/feast>) to run on Azure Databricks
- Did design reviews for large-scale data warehousing scenarios (Kroger)
- Built the infrastructure deployment automation for the CSE Kafka on-boarding event using Terraform and Kafka
- Wrote benchmarking tools to test CosmosDB's Cassandra offering

Technologies: Python, Azure Databricks, Scala, Spark, Terraform, Docker, Cassandra, Kubernetes, Azure Data Factory, EventHubs, Kafka, Azure Data Lake, Azure HDInsight, Azure CosmosDB

- **Microsoft - Customer Advisory Team (CAT/DataCAT)**

Software Engineer 2

Redmond, Washington
September 2015 - August 2018

- Worked as an OSS Data specialist
- Worked with customers to unblock technical issues, provide best practices
- Lead a team to convert a legacy SQL pipeline to Spark/Hadoop, reducing hardware cost by 80% and runtime by 96%
- Did design reviews with customers to provide optimal performance for a variety of ETL/OLTP scenarios
- Helped plan and automate data gathering for internal project that lead to key design decisions in ADLS Gen 2 (Hierarchical Namespaces)
- Wrote and iterated on Large-scale GDPR-enabled IoT Streaming project using Spark, EventHubs and Akka
- Worked closely with EventHubs product group to rewrite Spark drivers for best performance and developer experience

Technologies: Python, Scala, Azure Databricks, Hadoop, Yarn, Azure HDInsight, Azure CosmosDB, Kubernetes, EventHubs, Airflow, MongoDB.

- **Microsoft - Compute Insights Team**

Software Engineer

Redmond, Washington
March 2014 - September 2015

- Built tools to manage and observe Data Science jobs
- Built data visualization tools based on insights from low-level hardware metrics
- Wrote introductory courses on OSS Big Data concepts, notebooks and Hadoop for my team

Technologies: Python, Scala, F#, C#, Hadoop, HDInsight, D3.

TECHNICAL SKILLS

- **Programming languages:** Python, Scala, Rust, F#/.NET, JavaScript/Node
- **Cloud services:** Azure, GCP
- **Scripting/Infrastructure:** Terraform, Docker, Kubernetes, Bash, Powershell, Bicep

EDUCATION

- **Carleton University**

B.Sc in Computer Science

- Served as head of Services for the Computer Science Society
- Deployed OpenStack for students to try out cloud computing
- Worked as a software contractor part-time

Ottawa, Ontario, Canada

September 2009 - February 2014