# Scott A Lyons

Software Engineer 2 / OSS Data Engineer Redmond, Washington



#### EXPERIENCE

## Microsoft - Consumer Software Engineering (CSE)

Software Engineer 2

Redmond, Washington September 2018 - Present

- o Member of the Data Tech Domain team gathering industry trends from various Customer engagements and projects
- o Responsible for educating customers and field-oriented developers on industry trends related to Big Data.
- o 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)
- o Designed Data Architecture and built the initial Spark Streaming version of near-realtime ML models (Energinet)
- o Wrote benchmarking tests dealing with handling encrypted files in a Big Data pipeline (Walgreens)
- o 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
- o Did design reviews for large-scale data warehousing scenarios (Kroger)
- o 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

<u>Technologies:</u> 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

- $\circ \ \ \text{Worked as an OSS Data specialist}$
- 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%
- o 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)
- o Wrote and iterated on Large-scale GDPR-enabled IoT Streaming project using Spark, EventHubs and Akka
- o Worked closely with EventHubs product group to rewrite Spark drivers for best performance and developer experience

<u>Technologies:</u> 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

- o Built tools to manage and observe Data Science jobs
- o Built data visualization tools based on insights from low-level hardware metrics
- o 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
- Currently learning: Tensorflow, xgboost, Golang, Typescript

## **EDUCATION**

## • Carleton University

B.Sc in Computer Science

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

Ottawa, Ontario, Canada September 2009 - February 2014