Last updated in Apr 2026

Damon Myers

Staff Software Engineer | Data Platforms

Portland, OR | damon.shane.myers@gmail.com | linkedin.com/in/myersdamon | (432) 257-7656

Summary

Staff Software Engineer with 10 years of experience and deep expertise in building and operating data platforms on AWS. Expert in Spark, Airflow, and SQL/NoSQL databases across both batch and low-latency streaming workloads, with a track record of reducing operational costs, improving data quality, and delivering reliable pipelines at scale.

Experience

Staff Software Engineer, Data Platform, SentiLink – San Francisco, CA (Remote) Sept 2024 – Apr 2026

- Led the team through redesign and implementation of a new version of the core identity dataset data model and pipeline. Enabled entity resolution across 3+ datasets compromising 4+ billion records, reduced pipeline runtime and cost by 30%, and reduced effort to onboard new identity data sources.
- Automated complex identity dataset ETL pipeline using PySpark, Airflow, and AWS ECS, reducing monthly operational effort from 3 weeks to 2 days.
- Proposed and championed a north-star data platform architecture encompassing orchestrated ingestion pipelines, gRPC data microservices, automated data quality checks, and a self-service lakehouse; secured cross-team buy-in and roadmap commitment.
- Enforced automated data quality validation using PyDeequ across all batch pipelines, achieving zero production data quality incidents over two quarters.
- Designed and built a self-service lakehouse on AWS Glue and Athena, enabling Data Science and Product teams to self-serve their analytics needs
- Conducted performance testing to ensure a smooth migration from RDS PostgreSQL to Aurora PostgreSQL, reducing costs while improving reliability.

Staff Software Engineer, Aetion – New York, NY (Remote) Sept 2023 – July 2024

- Designed and built scalable ingestion tools in Databricks (Scala Spark, PySpark) and Airflow to transform EHR data, supporting hundreds of datasets ranging from tens of gigabytes to tens of terabytes.
- Led a team of 3 engineers building a Scala Spark test framework that uncovered critical data inconsistencies between datasets stored in Parquet and a proprietary data format, preventing corrupt data from reaching production.

Senior Software Engineer, Aetion – New York, NY (Remote) Mar 2022 – Sept 2023

- Led a team of 5 engineers developing a configuration-driven data ingestion pipeline in Airflow, PySpark, and Databricks for EHR data, reducing dataset onboarding time from 4 weeks to 1 week by eliminating custom transformation SQL.
- Optimized Elasticsearch indexing configuration and cluster scale, resulting in an annual cost reduction of $400,000.
- Implemented a framework for integration testing of Spark applications, significantly reducing production defects.

Senior Software Engineer, LegitScript – Portland, OR (Remote) Dec 2020 – Feb 2022

- Designed and implemented the proof-of-concept Data Lakehouse that became the foundation for LegitScript's production Data Lake, using CDC streaming (AWS DMS) to S3 with ingestion into Databricks and Delta Lake, plus a GraphQL abstraction layer on AWS Lambda.
- Developed a high-performance web scraping pipeline capable of crawling and ingesting 500,000 websites daily using TypeScript, AWS Step Functions, Lambda, and EventBridge.
- Led maintenance and enhancement of a horizontally scalable web crawler operating in a Kubernetes cluster on AWS EKS.

Software Engineer, LegitScript – Portland, OR Nov 2018 – Dec 2020

- Designed and maintained a low-latency streaming ETL pipeline processing 2.5 billion changes per month from MySQL to Elasticsearch, scaling the cluster to handle tens of terabytes of data.
- Trained and deployed ML models using AWS SageMaker, achieving 97% accuracy for website categorization via HTML analysis.

Junior Software Engineer, PLEXSYS Interface Products – Camas, WA Sept 2017 – Nov 2018

- Designed and implemented frontend and backend components for real-time simulation applications using C++, Qt, and Object-Oriented Design principles.

GIS Developer, City of Odessa – Odessa, TX Aug 2014 – Aug 2016

- Built web mapping interfaces for data visualization using ArcGIS JavaScript API and automated GIS database administration with Python.

Skills

Languages: Python, SQL, Go, Scala, TypeScript, C++

Data Engineering: Spark/PySpark, Databricks (Scala/PySpark), Delta Lake, Iceberg, Airflow, AWS Glue, Athena, PyDeequ

Data Stores & Processing: PostgreSQL, MySQL, Elasticsearch, DuckDB, AWS DMS (CDC)

APIs & Services: gRPC, GraphQL, microservices, REST

Cloud & Infrastructure: AWS (ECS, Lambda, EKS, DMS, SageMaker, S3), Kubernetes, Docker, Terraform

Practices & Tools: CI/CD, TDD, observability, data quality, data governance, HIPAA compliance

Education

University of Texas of the Permian Basin, BS in Computer Science – Odessa, TX June 2014 – Dec 2016

- Vice President, ACM