

# JORDAN LEWIS

JordanAllenLewis@gmail.com • (925) 819-0063 • [linkedin.com/in/JordanAllenLewis](https://www.linkedin.com/in/JordanAllenLewis) • [JordanAllenLewis.com](https://JordanAllenLewis.com) • Livermore, CA

## SUMMARY

**Data Engineer** with hands-on experience in data ecosystems, Kubernetes, APIs, and cloud technologies. Skilled in deploying and managing scalable data platforms, integrating cross-functional data systems, and automating workflows to enhance data accessibility and governance. Proven track record of leading complex projects and driving data-driven decision-making in agile environments.

## EDUCATION

**M.S. Computer Science** Graduated May 2024  
Georgia Institute of Technology, Atlanta, GA

**B.S. Computer Science** Graduated May 2019  
Arizona State University, Tempe, AZ

## EXPERIENCE

**RR Donnelley**, Remote - Livermore, CA (*Aug. 2022 – Present*)

### Data Infrastructure Engineer II

May 2024 – Present

- Serving as a full-stack, versatile member of a highly agile team, contributing to many facets of our Data Mesh initiative.
- Deployed and currently managing Dremio via Helm on an on-prem Kubernetes cluster, along with JMX metrics, Prometheus, and Grafana for real-time performance monitoring of Dremio.
- Implemented single sign-on (SSO) and LDAP role-based authorization for Dremio to enforce secure and automated user access, ensuring permissions align with team-based governance policies.
- Created MuleSoft API proxies on top of internal minion services with authentication through OneLogin and an internal auth minion service, facilitating secure data sharing across multiple acquired organizations.
- Enhanced our internal Data Mesh website by integrating MuleSoft Dremio APIs, enabling users to submit queries and retrieve results directly through the website, improving data accessibility and demonstrating the power and vision of the Data Mesh.
- Collaborating with IT teams to open networking pathways between Dremio and multiple on-prem Hadoop clusters, cloud services, and databases to ensure seamless data integration and accessibility for end users.
- Partnering closely with the DevOps, IT, and Data Platform teams to optimize deployment, integration, and maintenance of Data Mesh components, coordinating weekly troubleshooting calls with Dremio support to resolve challenges and improve system stability.
- Implementing data governance practices by working with the governance team to ensure all imported data in Collibra is properly classified, maintaining compliance and enhancing data discoverability across the organization.
- Continuing work and responsibilities listed below as a Data Infrastructure Engineer I.
- **Primary technologies:** Collibra, Docker, Dremio, Git CI/CD, Grafana, Hadoop, Helm, JMX, Jenkins, Kubernetes, LDAP, MuleSoft, Next.js, OneLogin, Prometheus, Python, REST APIs, Snowflake, Superset, Unix.

### Data Infrastructure Engineer I

Aug. 2022 – May 2024

- Creating a centralized data catalog, providing a unified view of company-wide data assets.
- Developing APIs to ensure seamless data access, contributing to a more responsive and user-friendly data ecosystem.
- Integrating existing internal websites with the Data Mesh, enhancing user experiences across the organization.
- Configuring ETL pipelines for consumers within the company, as well as Data Products on the public Snowflake Marketplace.
- Developed and launched an internal Next.js 14 web application, serving as a landing page/information hub for the Data Mesh project. Features real-time API calls to display live stats on number of data products, data sets, columns, etc.
- Authoring blog-style Data Mesh Guild posts which effectively communicate new deployments and key updates to stakeholders across the enterprise, enhancing cross-functional collaboration and driving informed decision-making.
- Assessing, deploying internally to Kubernetes/OpenStack, and creating compelling demos for technologies which have potential to enhance the Data Mesh.
- Participating in evaluations of various technologies for data virtualization, visualization, event streaming, generative AI, etc.
- **Primary technologies:** Collibra, Docker, Git CI/CD, Hadoop, Jenkins, MuleSoft, Next.js, Python, Snowflake, Unix.

**USAA**, Plano, TX (*Sept. 2019 – Jul. 2022*)

### Tech Lead - Software Engineer II

May 2022 – Jul. 2022

- Technical lead for a team of 5+ globally diverse engineers.
- Created technical designs and product roadmaps to drive reliable, reusable, and scalable solutions.
- Continued my responsibilities listed below as a Software Engineer II

### Software Engineer II

Aug. 2021 – May 2022

- Developed small batch data load and model processes using Domino, R, Python, Git CI/CD, and Airflow.
  - First team in organization to transition business developed models to IT supported infrastructure/technology, benefiting in 60% reduction in code and 95% reduction in hours worked monthly (\$84,000 per model/year).
- Programmed ETL jobs using Python, shell scripts, and IBM DataStage.
- Developed and maintained data integrity through run time controls over various data hops.
- Orchestrated data using a variety of databases such as Snowflake, Oracle, and Netezza.
- Visualized financial forecast data using React, D3, Plotly, and Tableau.
- **Primary technologies:** Airflow, Control-M, DataStage, Domino, Git CI/CD, Hadoop, Hive, Oracle, Python, R, React, Snowflake, Tableau, Unix.

### Software Engineer III

Sept. 2019 – Aug. 2021

- Created data pipelines from application databases to enterprise data warehouses.
- Developed and maintained data integrity through run time controls over various data hops.
- Programmed ETL jobs using Python, shell scripts, and IBM DataStage.
- Instructing quarterly courses to bring employees up-to-speed on IBM DataStage and ETL methodologies.