

# Patrick F Callahan

## Data Developer, Team Lead

P: +1-571-265-5098 | E: [patrick.f.callahan@gmail.com](mailto:patrick.f.callahan@gmail.com) | L: <https://www.linkedin.com/in/patrick-f-callahan/> | G: <https://github.com/sherlock-wings>

### Profile

I specialize in leading projects on the ingestion, transformation, and modeling of Big Data for Businesses, with 8 years experience in Data Engineering/Architecture and a focus on the Finances & Insurance sector. I have endless energy for taking on technical challenges, and I use my skills to make critical datasets accessible, understandable, and actionable.

### Skills

- |   |  |                                |
|---|--|--------------------------------|
| ✓ Snowflake, Snowpark                                 | ✓ dbt  | ✓ Fivetran                     |
| ✓ Python, Javascript, SQL                             | ✓ Data Architecture (RBAC, Kimball Modeling) | ✓ Asana, Jira                  |
| ✓ AI-Driven Development skills<br>(Claude, MCP, etc.) | ✓ git, DevOps                                | ✓ Client Relationship Building |
|   | ✓ AWS (S3, Lambda, IAM, etc)                 | ✓ Project Leadership           |

### Certifications

- ✓ **Snowpro Advanced Data Engineer Certification** **Expires Feb 2027**  
[View credential](#)

### Professional Experience

**Lead Consultant | Atrium AI - Bozeman, Montana (Remote)** **June 2025 - Present**  
[Atrium website](#)

*DE Lead (Project) - Asset Management Firm* *Nov 2025 - Present*

Led Architecture and Engineering efforts for an Asset Management firm who was trying to optimize their existing Snowflake implementation.

- Worked on a pilot project to simplify spaghetti code and optimize existing pipelines by migrating from Procedures to dbt
- Implemented DevOps, Disaster Recovery, Comprehensive Data Masking, and Sandbox, Dev, QA, Preprod, and Production Environments
- Refactored legacy reporting, reducing compute from ~110 minutes per refresh to 5-8 minutes per refresh
- Layed a data model foundation using Medallion and Kimball architectures to allow the firm to expand their optimization efforts

*DE Lead (Project) - Large Fortune 500 Tech Corporation* *June 2025 - Nov 2025*

Led Architecture and Engineering efforts for a large tech corporation, who was looking to build a Snowflake pipeline to power their end-user applications.

- Worked with the Client to build and end-to-end POC design
- Implemented design across rapid sprints, working with multiple other engineers on the Client and Atrium side
- Wrote and executed multithreaded scripts to produce billions of rows of fake data for testing purposes
- Conducted benchmark testing to prove out the robustness of the pipeline implementation

*DE Lead (Project) - Small Tech Firm* *June 2025 - Oct 2025*

Led Architecture and engineering efforts in an ARR Dashboard project. The goal was to bring this team out of Excel and into the modern cloud data space.

- Led the database design and modeling, including end-to-end RBAC and Kimball-type Entity Relationships
- Built pipelines to continuously ingest from S3 buckets to Snowflake for multiple data sources
- Working with another engineer, built a DAG in dbt for end-to-end transformations that harmonized ARR data from many sources
- Worked with Client and Atrium-side Analysts to construct SQL Reports based off of dbt output models

**Sr Consultant I, Data Engineering | Hakkōda, an IBM Company - Heredia, Costa Rica (Remote)** **July 2023 - June 2025**  
[Hakkōda Website](#)

*Senior Engineer (Project) - Large Agrochemical Corporation* *Feb 2025 - June 2025*

Worked with a small team to build a Genetic Relational Data Graph, also called a Pangenome, to help predict all possible haplotypes across the genomes of key species of interest.

- Built custom parsers (Snowpark, Python) for massive genetic line files
- Restructured data to fit a Graph AI tool being tested as part of a POC ([RelationalAI](#))
- Benchmarked RelationalAI performance against Client's existing inference methods

*Lead Engineer (Project) - Asset Management Firm* *Mar 2024 - Feb 2025*

Led Engineering efforts in a data migration project. Data was moved from On-Premise legacy systems into Snowflake. The goal was to provide their

Back-Office departments with improved Data Quality and faster return times on reporting so that they could perform tasks like NAV Reconciliation, Position Reconciliation, and Management Fee Analysis with greater accuracy, efficiency, and automation.

- Led team efforts in building and maintaining a production data pipeline. This pipeline transformed raw data through a staging layer, modeled layer, and reporting layer, using Kimball Methodology (Fact Tables, SCD I/II, etc).
- Data was ingested from on-premise SQL Server instances via Fivetran. Transformations and unit testing was implemented in Coalesce, with a DAG of about 400 Nodes. Version Control (i.e. git) & CI/CD handled with Azure DevOps
- Simultaneously owned production support and development of new features from sprint to sprint
- Onboarded new developers to the project, upskilling them by providing documentation and 1-1 time to help newcomers understand the project's tech stack and work policies.

**Lead Engineer (Project) - Private Software Corporation**

**Sep 2023 - Feb 2024**

Led Engineering efforts in a project to refactor and repurpose large amounts of legacy logic. The logic had been performing poorly and became difficult to maintain on their local Oracle systems. The project's goal was to improve the accuracy and efficiency of existing reporting on Customer Accounts, as well as to expose/correct inconsistencies in how business rules had been understood and applied.

- Oracle Data was migrated into a Snowflake using Airbyte for ingestion, dbt for Data Transformation & Validation, and Azure DevOps for Version Control (i.e. git) & CI/CD
- Refactored thousands of lines of SQL from stored procedures and existing Oracle models, breaking them up into reusable components, such as dimension tables, parameterized views, or mapping tables
- Brought runtime for report generation down from approximately 3.5 hours to 8 minutes
- Worked with business analysts and junior Developers to produce an interactive PowerBI Dashboard for C-Suite Employees

**Analytics Engineer | CMS Payments Intelligence, Inc - Atlanta, GA**

**Jan 2022 - May 2023**

[CMSPI Website](#)

Work on a team of Developers to serve data pipelines, SQL Reporting, client dashboards, and other kinds of data products to Consulting Teams working with Large vendor clients. This was done so clients' massive streams of customer payment data from processors such as Adyen, Chase, Fiserv, Braintree, Elavon, etc could be easily summarized, filtered, and drilled into to gain financial insights.

- Work under a product owner for specific client-based engagements
- Use Databricks for all Cloud Data Processing, heavily leveraging both Databricks SQL and the PySpark API as needed
- Process data using a medallion architecture, organizing Gold-layer objects into a business-defined logical hierarchy
- For data ingestion, leverage a wide variety of Python-based techniques, such as webscraping, PDF parsing, API calls, FTP, etc
- Build interactive PowerBI Dashboards to satisfy client requirements. Work with internal consulting team iteratively to achieve a finished product

**Data Analyst | LabArchives, LLC - San Diego, CA (Remote)**

**Feb 2017 - Dec 2021**

[LabArchives Website](#)

Function as the go-to Data person at a small WebApp company. Develop and support a catalog of SQL Reports, PowerBI Dashboards, and other Data Assets to satisfy the needs of various company Departments; e.g. for Sales, deliver reports on new potential leads and important events on existing accounts. For Customer Support, deliver reports on tickets, Support Call uptime, and other key metrics. Work with end users regularly to ensure all reporting delivers useful insights as directly as possible.

- Maintain a deep understanding of the WebApp's Data Structure in MySQL, including all data definitions and business rules
- Leverage Python to send reports out via email on a scheduled basis
- Use PowerBI to develop a catalog of Dashboards to allow users to quickly search and filter for the data they need
- Develop SQL Reports from scratch. Debug and optimize existing reporting as data grows and changes

**Product Specialist/Customer Support Representative | LabArchives, LLC (Remote)**

**Jun 2014 - Nov 2020**

Maintain high level of product knowledge and be able to train coworkers and customers on product features via online and in-person meetings Receive and handle Support tickets via phone call and email quickly, effectively, and professionally. Promote the product to strategic customers as needed.

- Construct presentation materials to lead customer webinars
- Hold both strategic in-person Presentations and regular online webinars to get customer interest and answer questions
- Use Kayako for IT Support Desk work (Emails/Tickets, Phone Call support, etc)
- Work with Development Team to QA new features; work with Customers to communicate feature requests to Development
- Maintain relationships with the Administrators of key Enterprise accounts

## Education

**George Mason University - Fairfax, VA**

**Master's in Information Technology**

**Bachelor's in Chemistry**

**Aug 2018 - May 2021**

**Aug 2010 - May 2014**

