# in Adam T. Gaddis

### Big Data Systems & Software

✉ agaddis02@live.com  ⌂ agaddis02  ☎ (214) 335 - 3200  </> Python, Go, Scala

## Experience

- **Arch Energy Partners** Dallas, TX (Hybrid)
  *Lead Data Engineer/Scientist* *March. 2023 - Present*
  - Developed robust data ingestion pipelines using Go, facilitating efficient data acquisition from various APIs. Develop Python and Scala code on Databricks, leveraging Apache Spark for processing extensive well production data, (500 Million+ Daily).
  - Created a bespoke tool for tracking and monitoring active oil wells, integrating real-time data and analytics capabilities. Developed an ELT pipeline in Python to streamline accounting data integration into the data warehouse.
  - Expertly set up and maintained a multi-cloud environment across GCP, AWS, and Azure, optimizing for cost-effectiveness and performance. Pioneered the use of BigQuery for centralized data warehousing, enhancing data accessibility and analytics.
  - Engineered multiple Google Cloud Functions to automate and refine data transformation and storage jobs, enhancing system efficiency and reliability.

  Technologies: Go, Scala, Python, GCP, Bigquery, S3, GCS, Cloud Functions, Streamlit, Databricks, Apache Spark, MongoDB

- **Bank of America** Dallas, TX (Hybrid)
  *AVP Data Engineer* *Sep. 2022 - Mar. 2023*
  - Led the establishment of SDLC practices for the team, transitioning code to a centralized Bitbucket repository and meticulously documenting all ongoing systems.
  - Overhauled real-time text-based data segments of the data warehouse, converting SQL Server scripts and SSIS packages into efficient PySpark code and Kafka streams, achieving processing of hundreds of thousands of data points in near real-time.
  - Successfully developed a Flask-based web application to revamp a report tracking system for a sister team, managing and updating over 1000 internal reports encompassing more than 500,000 data points.
  - Spearheaded the continuous integration and continuous deployment (CI/CD) processes using Bitbucket and Jenkins, ensuring seamless and efficient deployment of software updates.

  Technologies: Python, Azure, Jenkins, Git, Hadoop, MapReduce, HDFS, Sqoop, Impala, CosmosDB

- **LOOP** Dallas, TX (Remote)
  *Senior Data Engineer* *April 2022 - Sep. 2022*
  - ○ Orchestrated the deployment of an Airflow instance within a Docker container on GCP, serving as the central scheduler for complex data processing and management jobs.
  - Collaboratively developed an internal model with a 3rd Party Data Provider for Version 2 of our original quoting model. This resulted in a 75% increase in the "Sign on first Quote" metric, indicating enhanced customer engagement. The model was deployed through Vertex AI.

  Technologies: Python, C#, Docker, GCP (Cloud Storage, Vertex AI, BigQuery), Airflow, ERD, API's.

- **Southwest Airlines** Dallas, TX (Remote)
  *Data Platform Engineer Intern* *Jan. 2022 - April 2022*
  - Developed a POC Python Framework to properly parse X12 EDI Documents.
  - Orchestrated and automated Python ELT Pipeline, using Airflow, to ingest X12 EDI Invoices Daily into an On-Prem Postgres Data Warehouse.
  - Deployed and monitored pipeline using Docker and Jenkins.

Technologies: Python, Airflow, Postgres, Docker, Git, Jenkins.

- **Goldman Sachs**                                                   Dallas, TX (In Person)
  *Data Platform Engineer Intern*                                    *Jun. 2021 - Aug. 2021*
  - Developed Python Application to compare database schema's between on-premise systems and their cloud equivalents, reconciled 100% of mismatches.
  - Rewrote Enterprise Data Platform test cases to utilize pytest, testing coverage increased from 70% to 91%+.
  - Developed and Deployed a key Python Library (PAT) streamlining workflows with instant connections to 5 central internal data stores.

  Technologies: Python, JIRA, Confluence, API's, Git, Docker, Kubernetes, Tableau, PyPi.

- **University of North Texas at Dallas**                            Dallas, TX (Remote)
  *Lead Analytics Engineer*                                          *Aug. 2020 - December. 2021*
  - Oversaw a team of 5 engineers focused on Data Governance, Reporting, and Quality for the University, Rewrote existing scattered SAS pipelines into an all inclusive Scala system, decreasing time for State report generation by 250%.

  Technologies: Scala, SAS, SQL Server, Oracle, Azure, Git, API's.

- **Scout Energy Partners**                                          Dallas, TX (Hybrid)
  *Data Scientist*                                                   *Jun. 2019 - Aug. 2020*
  - Developed Python ETL pipelines to retrieve more than 10 million+ data points daily from in-field SCADA Systems.
  - Built predictive analytical models forecasting net production returns from various fracking strategies, aiding in strategic decision-making.
  - Constructed and deployed a Python machine learning model on Databricks ML platform for analyzing incoming deals on long-term PDP wells, utilizing geospatial data and reverse time-series analysis.
  - Published an internal whitepaper on the model, leading to significant operational tweaks. Collaborated closely with Production Engineers and Geologists to refine the model and enhance its applicability.

  Technologies: Python, Azure, Databricks, Github Actions, SQL Server, Geopandas, Fiona, Shapely, Scikit-learn

## SKILLS

Orchestration: Airflow, Luigi, Cloud-Based(ADF, DataFlow)      Databases: Postgres, CassandraDB, MongoDB
Python Frameworks: Flask, PySpark, FastAPI                    Scala Frameworks: Spark, Hadoop, Hive
Datawarehouse: Bigquery, Databricks, Azure Synapse            Cloud Platforms: GCP, AWS, Azure
Analytics: Arcgis, All Major BI Tools,                        ML & AI: PyTorch, Keras, OpenCV, Apache Spark MLlib

## EDUCATION

**University of North Texas at Dallas**                               Dallas, TX
*B.B.A in Analytics — GPA: 3.7/4.0*                    *Aug. 2020 – Dec. 2022 (Graduated)*
Notable Courses: *Statistical Modeling, Database Applications, Data Warehousing, Big Data Management.*

**Dallas Community College — (Richland)**                            Dallas, TX
*A.A.S in Software Development — GPA: 3.6/4.0.*        *Aug. 2016 – May. 2020 (Graduated)*
Notable Courses: *Operating Systems, C# Development, System Design, Java, Python, Networking.*

## PROJECTS

- Oil & Gas Contract: Developed a ELT process for an oil & gas client, to interact with Combocurve API v2, and ingest all data available, built using Pyspark & SQL Server
- LightningScout Web App: Developed Scala Back-end to interface with a flutter based application that is used our data collection system.
- LightningScout ELT: Using Python, Azure (Postgres), to automate the ingestion of 100k Robot Scouting Actions Annually.

- Slack Bot Scheduler: Python based, event driven slack bot, interacting with S3's, TBA API v3, and internal scripts.
- Scouting System Mentor FRC3005: I teach kids how to program crazy systems & application through Robotics in Python, Scala, and Go that professionals are just now learning.