

# Vincent DiRenzo

[GitHub](#) | [LinkedIn](#) | [Medium](#)

---

## EXPERIENCE

### Data Consulting Company - Data Engineer

2022 - Present

- Led project to build, implement, and maintain an analytics data warehouse for serving data to a Looker dashboard using GCS/BigQuery.
- Built Python Docker images for ingesting/transforming 4 data sources. Used Cloud Composer and Terraform to run Docker images daily.
- Created CI/CD pipeline using Github Actions to add linting/formatting hooks, integration tests for 3rd party API endpoints using Pytest.
- Created and maintained Cloud Composer environment using Terraform, scaled environment resources per daily ingested data size.

### Programmatic Advertising Platform - Data Engineer

2022 - 2022

- Responsible for the aggregate media data infrastructure of Omni, an analytics platform used by 1500+ OMG agencies worldwide.
- Refactored data ingestion from a 'data swamp' to a data lake with a semi-structured schema and lifecycle rules for each data source.
- Worked with 3rd party vendors to maintain access to API's, implement schema changes. Defined business needs working with analysts.
- Built Airflow ETL pipelines on AWS using PySpark, S3, and Athena for 8 distinct data sources that ingested 50k/records daily each.

### Data Consulting Company - Data Engineer

2021 - 2022

- Built, deployed, and maintained data infrastructure solutions on GCP/AWS for 10+ clients. Acted as client contact for ongoing support.
- Refactored existing Airflow codebases from version 1 to 2. Implemented DAG optimizations resulting in a >40% reduction in runtime.
- Built pipelines to extract data from 3rd party API's with Python, transform data using PySpark, and load to data warehouse solution.
- Developed data warehousing model using fact/dimension schema. Built data marts that made business metrics accessible to all teams.

### Social Media Suite - Developer Relations, Falcon.io

2020 - 2021

- Oversaw premium tier customer usage of the Falcon API. Conducted live code demonstrations of API usage with Postman and Python.
- Managed relationships with upstream data partners to identify changes in data availability and proactively change API structure.
- Shared API customer feedback with Product & Engineering teams. Helped maintain the feature roadmap and API documentation.
- Led enablement for Sales teams, developed collateral with Marketing, and reported revenue and adoption metrics to leadership.

### SaaS Focus Group Solution - Industry Lead, Public Opinion Research

2019 - 2020

- Established the Public Opinion Research team. Grew team to 3 analysts, and secured initial clients ahead of 2020 Presidential Election.
- Generated \$300,000 in revenue during the 2020 Presidential Election through professional services billed to the Public Opinion team.
- Worked with the ML team to build an ETL pipeline for processing user input on platform and training sentiment analysis model.
- Worked with clients and data vendors to scope and fulfill project needs. Produced data analysis and visualizations for presentations.

### Big 3 Public Cloud Provider - Account Manager

2016 - 2019

- Responsible for managing 100+ clients and \$1.5 million in annual run rate. Acted as the Google Cloud SME to customer end users.
- Led ongoing enablement seminars for customers, oversaw participation in EA/Beta programs, and regularly provided roadmap updates.
- Scoped and proposed GCP architectures for client use cases. Led conversations between Google technical teams and client end users.
- Performed 'live-coding' on the GCP console to enable new client end users. Used Architecture Diagramming Tool to build tutorials

### Political Media Outlet - Marketing Data Analyst

2015 - 2016

- Analyzed platform traffic and marketing data with Python and SQL to identify channels, campaigns, and content yielding highest ROI.
- Collaborated with the Engineering team to build ETL pipeline and data warehouse schema. Created data marts optimized for analysis.
- Made dashboards with PowerBI for senior leadership use. Presented sales, marketing, and platform traffic data in a clear format.
- Assisted in the creation of the Politico EU expansion proposal. Created analysis and visualizations to justify investment to ownership.

### National Security NGO - Open Source Intelligence Analyst

2014 - 2015

- Created OSINT collection tools using Python, social media API's, and Postgres to collect posts concerning specific topics or regions.
- Collaborated with a bipartisan team of staff members to prepare analysis of collected data. Wrote executive summaries for senior staff.
- Participated in discussions with DOD, Intelligence, State Dept. officials, and lawmakers about social media and intelligence collection.
- Created documentation and enablement to allow other staff members to use OSINT collection tools. Recorded training for future use.

---

## PROJECTS

### War in Ukraine ETL Pipeline & Dashboard

[Github](#), [Medium Post](#)

### Automatic Target Recognition using Deep Learning

[Github](#), [Medium Post](#)

### Asteroid Diameter Predictor Web Application

[Github](#), [Medium Post](#)

### Agile Machine Learning Prototyping with Streamlit

[Github](#), [Medium Post](#)

---

## SKILLS

- |                       |                         |                     |
|-----------------------|-------------------------|---------------------|
| • Python              | • SQL                   | • Scala             |
| • Shell Scripting     | • Atlassian Suite       | • Docker/Kubernetes |
| • Amazon Web Services | • Google Cloud Platform | • MongoDB           |
| • Airflow             | • Terraform             | • dbt               |
| • Jenkins             | • Git Version Control   | • Snowflake         |

---

## EDUCATION

### New York University - Bachelor's, Individualized Study

2010 - 2014