






# Gabriel (Runye) Wang

## Data Engineer

### Contact Info

-  343 Gold St,  
Brooklyn, NY, 11201
-  706-362-3561
-  me@gabrielwry.com
-  gabrielwry
-  <https://gabrielwry.com>

### Skills

#### Programming Languages

Python, SQL, Scala, Java

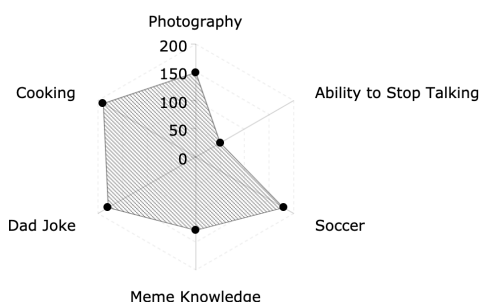
#### Frameworks

Big Data - Airflow, Spark, TensorFlow, Kafka  
Platform - AWS, GCP, Databricks, Tableau  
DevOp - UNIX, K8S, Docker, Git, Harness  
Project Management - JIRA, Agile

#### Data Science

Information Retrieval  
Anomaly Detection  
Recommending System  
Natural Language Processing

### Personal



### Objective

**Data Engineer** with **3+ years** experience of delivering **ETL pipeline, infrastructure** for complex marketing scenario, with exposure to **Machine Learning** research .

### Education

**Bachelor of Science in Computer Science**  
**Bachelor of Eng. in Engineering Physics**

August 2014 - May 2018

Emory University, Atlanta, GA

### Publication

- “**Making a Move: Geographic and Social Mobility in Late 19th Century Wedding Announcements in Turin** *La Stampa*” - SSHA 2019 Annual Meeting, November 21-24
- “**Automatic Information Extraction of the Narrative Elements Who, What, When, and Where**” - *Sociology Methodology* 2020 June Edition

### Experience

#### Data Engineer II at Tilting Point, New York, NY

Sept 2019 - Present

- Design and implement **Spark ETL pipeline** with **Airflow** to ingest revenue and marketing data from more than 50+ APIs to power Business Intelligence that achieve company-level 300% ROAS growth;
- Architect and implement **Anomaly Detection** system with **deequ** on critical data services which delivers metrics to **Prometheus, AWS Cloud Watch** and **PagerDuty**;
- Work with DevOp engineer to deploy and maintain data services on **AWS** ecosystem including **Airflow** on **EKS**, configuration database on **RDS**, access control on **VCP**, etc.
- Research on **Marketing Mixed Model** for campaign targeting while preserving privacy; work with **Data Scientist** to deploy **TensorFlow** production model to **Databricks ML** reducing cost per install by 20%.

#### Associate Data Scientist at HEED LLC, New York, NY

Sept 2018 - July 2019

- Designed and prototyped **ETL** for streaming sensor data and with Social Media feeds processed by **Information Retrieval System**; worked with Engineering team to productionize and deploy on **GCP**.
- Designed and implemented Business Intelligent tools on **Data Studio, ELK** to support creative content team with real time data analysis;
- Analyzed app user activities with **decision tree learning** and **random forest** to support marketing team strategy decision and boost 400% user growth from v1.0 to v2.0;