

## PROFESSIONAL SUMMARY

Data engineering professional with 15+ years building systems that matter. Discovered systematic demographic coding errors affecting all Black and Asian-American voters, developed geospatial ML algorithms improving classification accuracy from 23% to 64%. Built Civic Graph data warehouse processing billions of records and platforms serving thousands of analysts nationwide.

## KEY ACHIEVEMENTS AND IMPACT

Discovered systematic race coding errors affecting all Black and Asian-American voters

- Algorithm reduced mapping costs by 73.5%, saving campaigns and organizations \$4.7M
- Built redistricting platform used by thousands of analysts nationwide
- Achieved 87% prediction accuracy for voter turnout vs. industry standard of 71%

## CORE COMPETENCIES

Data Engineering • Programming

## PROFESSIONAL EXPERIENCE

### Siege Analytics | Partner - Austin, TX 2005 - Present

#### Data, Technology and Strategy Consulting

- Discovered systematic race coding errors affecting all Black and Asian-American voters, developed geospatial machine learning algorithms improving demographic classification accuracy from 23% to 64%
- Built redistricting platform used by thousands of analysts nationwide with real-time collaborative editing and Census integration
- Achieved 87% prediction accuracy for voter turnout vs. industry standard of 71%, reducing polling error margins from  $\pm 4.2\%$  to  $\pm 2.1\%$

### Helm/Murmuration | Data Products Manager - Austin, TX June 2021 - May 2023

#### Civic Graph & Civic Pulse Director

- Conceived, architected and built Civic Graph multi-tenant data warehouse processing government data from Census, Bureau of Labor Statistics, National Council of Educational Statistics
- Built multi-dimensional data warehouse measuring socio-economic changes in America at every level across attitudinal, behavioral, demographic, economic and geographical dimensions
- Managed engineering teams of 7-11 professionals while setting technical direction for data architecture

### GSD&M; | Analytics Supervisor - Austin, TX November 2019 - June 2020

#### Big Data Engineering Transformation

- Transformed small data team into big data engineering team, scaling from laptop datasets to Hadoop Clusters and Hive on AWS
- Managed accounts including United States Air Force, Southwest Airlines/Chase and Indeed

- Rewrote mission and offerings of department and drafted integration plan with strategy team

**Mautinoa Technologies | Software Engineer - Austin, TX August 2016 - February 2018**

**SimCrisis Product Owner/Engineer**

- Conceived, architected and engineered econometric simulation software for humanitarian crises intervention measurement
- Built SimCrisis GeoDjango web application using multi-agent modeling to create econometric simulations of crisis economies
- Designed modular application accepting rules extensions for ethnic strife, different crises/disasters, supply failures

**Myers Research | Senior Analyst - Austin, TX August 2012 - February 2014**

**RACSO Product Owner/Engineer**

- Designed comprehensive survey instruments for specialized voting segments and niche markets
- Co-developed RACSO web application managing all aspects of survey operations from instrument design to data analysis
- Wrote RFP and analyzed bids from 1,200 vendors for research platform development

**PCCC | Research Director - Washington, DC 2010 - 2012**

**Political Research & Data Analysis (FLEEM System)**

- Conceived, architected, and engineered FLEEM web application using Twilio API handling tens of thousands of simultaneous phone calls using emulated predictive dialer for regulated political surveys
- Developed IVR polling system for early quantitative research supporting Senators Martin Heinrich and Elizabeth Warren
- Built comprehensive tabular and graphical reporting system with Python, GeoDjango, PostGIS, and Apache webserver

**Salsa Labs | Software Engineer - Washington, DC January 2011 - August 2011**

**Geospatial CRM Development**

- Maintained and extended comprehensive geospatial analysis and reporting tools for Java-based CRM system used by tens of thousands simultaneously
- Developed custom tile server for Web Map Service (WMS) integration using GeoTools and OpenLayers
- Built advanced geospatial analysis capabilities using Java, JavaScript, MySQL, and TileMill

**Lake Research Partners | Programmer - Washington, DC April 2008 - December 2008**

**Political Research & Analytics**

- Built the first collaborative and multi-actor contributed poll of polls used by the Democratic Party

- Harmonized data from 20+ polling firms with incompatible methodologies and encoding systems
- Created comprehensive meta-analysis framework handling millions of survey responses that became the \$400M Polling Consortium Database at The Analyst Institute, now valued at \$1B+

## KEY PROJECTS

### National Redistricting Platform (2020 - 2021)

**About:** Cloud-based GeoDjango platform for redistricting analysis with real-time collaborative editing and Census integration, used by thousands of analysts nationwide

**Technologies:** GeoDjango, PostGIS, AWS, Docker, React, Python

**Impact:** Reduced mapping costs by 73.5%, saving organizations \$4.7M in operational expenses

### FLEEM Political Polling System (2010 - 2012)

**About:** Completely self-built IVR system using Twilio API that contacted tens of thousands of voters daily, replicated call center functionality to performance parity

**Technologies:** Twilio API, Python, Django, PostgreSQL, JavaScript

**Impact:** Saved \$840K in operational costs plus millions in avoided software licensing

### Geospatial Demographic Classification System (2013 - 2016)

**About:** Machine learning platform that discovered systematic coding errors and improved demographic classification accuracy from 23% to 64%

**Technologies:** Python, Scikit-learn, PostGIS, GeoPandas, TensorFlow

**Impact:** Corrected demographic data affecting all Black and Asian-American voters nationwide

## TECHNICAL SKILLS

**DATA ENGINEERING** *Apache Spark, PySpark, Dask* (Large-scale data processing and distributed computing); *dbt, Airflow* (Data transformation pipelines and workflow orchestration); *PostgreSQL/PostGIS, Snowflake* (Database design and geospatial data management); *AWS (EC2, RDS, S3), Docker* (Cloud infrastructure and containerization)

**PROGRAMMING** *Python* (NumPy, Pandas, Scikit-learn, Django, Flask (15+ years)); *SQL* (Complex queries, optimization, spatial analysis (15+ years)); *R* (Statistical modeling, ggplot2, spatial packages (12+ years)); *JavaScript* (React, D3.js, Node.js, real-time applications (10+ years))