

## 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

Built real-time FEC analysis systems using Python, Pandas and PySpark to detect likely fraud, money laundering and financial crimes across billions of records daily, performing time series analysis on trillions of records in the political spending sub-economy valued over \$2 trillion

## CORE COMPETENCIES

Data Infrastructure • Programming and Development • Cloud & DevOps

## PROFESSIONAL EXPERIENCE

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

#### Data Science & Political Analytics

- 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  $\pm 2.1\%$
- Trigonometric algorithm for boundary estimation reduced mapping costs by 73.5%, saving campaigns and organizations \$4.7M and enabling smaller nonprofits to conduct analysis
- Built real-time FEC analysis systems using Python, Pandas and PySpark to detect likely fraud, money laundering and financial crimes across billions of records daily, performing time series analysis on trillions of records in the political spending sub-economy valued over \$2 trillion

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

#### Democratic Electoral Technology

- Led design and implementation of enterprise-scale multi-tenant data warehouse for geo-referenced demographic, econometric, and electoral data
- Managed engineering team of 11 professionals while setting technical direction for data architecture
- Modernized legacy ETL processes by implementing dbt and PySpark workflows, reducing processing time by 57%

### GSD&M; | Analytics Supervisor - Austin, TX 2018 - 2019

#### Advertising Analytics

- Restructured the Decision Sciences Department to scale capabilities from small-scale data analysis to comprehensive big data operations
- Implemented spatial analysis and consumer segmentation methodologies that revealed new insights about existing customers
- Advanced Statistical and ML techniques for segmentation and behavioral clustering

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

### **Software Development**

- Conceived, architected and engineered econometric simulation software for humanitarian crises intervention measurement
- Liaised with data and engineering directors at multinational NGOs (UNICEF, IFRC)
- Geospatial analysis on populations and boundaries for impact assessment

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

### **Political Research & Analysis**

- Designed comprehensive survey instruments for specialized voting segments and niche markets
- Developed sophisticated analytical products and reports that delivered actionable insights to clients
- Co-developed a web application to manage all aspects of survey operations, from instrument design to data collection and analysis

## **PCCC | Research Director - Washington, DC August 2011 - August 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

## **KEY PROJECTS**

### **National Redistricting Platform**

**About:** Cloud-based GeoDjango platform for redistricting analysis with real-time collaborative editing and Census integration, used by thousands of analysts nationwide during 2021 redistricting cycle

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

**Impact:** Reduced mapping costs by 73.5%, saving organizations \$4.7M in operational expenses. Served 12,847 analysts across 89 organizations.

### **Geospatial Demographic Classification System**

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

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

**Impact:** Corrected demographic data affecting all Black and Asian-American voters, improved electoral prediction accuracy by 22%

### **High-Performance Geospatial Tile Server**

**About:** Custom tile server for Web Map Service integration enabling interactive visualization of CRM and Census data

**Technologies:** GeoTools, OpenLayers, Java, MySQL, TileMill, JavaScript

**Impact:** Improved contact rates by 53% and segmentation accuracy by 88% through enhanced data visualization

## TECHNICAL SKILLS

**DATA INFRASTRUCTURE** *Processing* (Apache Spark, PySpark, Dask, parallel computing, distributed systems); *Pipelines* (Airflow, dbt, ETL design, data quality monitoring, automated testing); *Storage* (Data warehousing, data lakes, columnar storage (Parquet), data modeling); *Streaming* (Kafka, Redis, real-time processing, event-driven architecture)

**PROGRAMMING AND DEVELOPMENT** *Python* (15+ years: NumPy, Pandas, Scikit-learn, TensorFlow, Django, Flask, GeoPandas, Asyncio); *R* (12+ years: Statistical modeling, ggplot2, dplyr, spatial packages (sf, sp), Shiny); *SQL/PostGIS* (15+ years: PostgreSQL/PostGIS, MySQL, complex spatial queries, optimization, database design); *JavaScript* (10+ years: React, D3.js, OpenLayers, Node.js, real-time applications, WebSockets); *Java* (8+ years: Enterprise applications, Spring framework, geospatial libraries (GeoTools)); *Other Technologies* (Shell scripting, Git, Docker, Kubernetes, infrastructure as code)

**CLOUD & DEVOPS** *AWS* (EC2, RDS, S3, Lambda, CloudFormation, Auto Scaling, Load Balancing); *Containerization* (Docker, Kubernetes, container orchestration, microservices); *Monitoring* (CloudWatch, Prometheus, Grafana, application performance monitoring); *CI/CD* (Jenkins, GitHub Actions, automated testing, deployment pipelines)