**Adam Gutcheon**

**Lead Data Engineer**

**Professional Experience:**

**Protagonist 04/16 - 01/18**

**Lead Data Engineer**

I managed the data tech platform and built tools for Natural Language Processing.

* My team and I built an almost entirely serverless data pipeline on AWS that uses cutting-edge machine learning techniques to process, extract features, analyze, and enrich a stream of millions of texts a day.
* We also built a single, unified data API as a single point of contact between the application and data layers.
* Working closely with our data scientists, we created training and test data sets for ML model creation, then adapted and operationalized the models into the pipeline.
* As team lead, I mentored junior engineers, coordinated professional development, and worked closely with
* management and the product team to groom and prioritize the product backlog, remove obstacles to progress, and facilitate clear lines of communication.
* Technologies used include Python and its math and machine learning libraries, Cython, Spark (using Scala),

PostgreSQL, ElasticSearch, Docker, Airflow, Kafka, and many AWS services (including Lambda, Step Functions, API Gateway, Kinesis, RDS, ECS, EMR, and Redshift).

**Voyager Data Services 04/12 - 04/16**

**Data Technology Consultant**

Data tech and team manager, trainer, obsessive. Some projects include:

*A Top Membership-Based Website*

Managed the company’s new BI platform team. Led data architects and developers building self-service solutions for the business, while transitioning from a Microsoft-only to a heterogeneous solution.

* Using open source technologies, built self-service frontends for business users to explore, monitor, and analyze data.
* Used new integration technologies such as Airflow and Kafka to connect disparate data sources and allow greater flexibility in choosing the right tool for each kind of analysis.
* Built and managed backend infrastructure including PostgreSQL, MSSQL, Hadoop, and AWS (Redshift, EMR).
* Introduced machine learning and AI to our BI stack, using Python and Spark.

Previously, I’d led the migration of OLTP infrastructure from disparate systems including MSSQL, MySQL, and MongoDB to PostgreSQL on VMWare and AWS (EC2 and RDS). This one-year project was completed in seven months, and no full-time DBA was needed to manage the resulting infrastructure.

* Created systems for deployment, provisioning, and monitoring which existing DevOps resources could handle, obviating the need for a full-time DBA.
* Trained developers and webops in PostgreSQL administration and internals.
* Compiled custom kernels, created debian source packages, and tuned ESXi to address existing as well as prospective performance and availability issues.

*A Start-Up in the Analytics Services Space*

Managed and maintained dozens of PostgreSQL servers for both analytic and transactional workloads involving millions of airline tickets and billions of airline fares and rules. Created data warehouse models for intensive analytics applications.

* Planned and implemented replication for load balancing and parallel query using pgpool-ii, leveraging existing hardware and allowing the company to pivot from a scale-up to scale-out strategy.
* Standardized ETL on Talend Integration Suite, which reduced DBA time spent managing processes and errors over 60%.

**Democratic National Committee 05/10 - 04/12**

***Data Technology Manager***

Managed the data warehousing, analytics, and reporting infrastructure for voter information, polling data, fund-raising, and the development of predictive models of voter behavior. Ensured availability, quality, and performance of analytics data for analysts, political scientists, developers, and vendors.

* Led the integration of spatial data into our analytic applications, allowing the creation of neighborhood models of voter behavior which were heavily used by our campaigns.
* Technical lead on the replacement of a Netezza device with Hadoop and the conversion of SQL and shell scripts to map/reduce methods (mostly in Python), which allowed significantly more experiments to be run.
* Introduced machine learning to our voter targeting process.
* Developed the election night prediction engine which called congressional and gubernatorial races with only one incorrect call.

**Area Cooperative Educational Services 01/08 - 05/10**

***Database Specialist***

Developed software, administered database systems, and taught professional development workshops for internal and external clients.

* Trained teachers, technical staff and school administrators on measures of individual and group student
* achievement, the features of student data, basic statistics and how to use their student database tools and technologies.
* Developed a Python (Django) web application to track, report and analyze behavioral interventions for special-needs
* students, resulting in a 28% increase in compliance with Positive Behavior Intervention protocols.
* Developed a web application that tracks time, expenses, outcomes and accountability measures for school improvement consulting programs, and integrated this with billing and payroll systems.

**Windsor Data Services 06/99 - 12/07**

***Principal Consultant***

Consulted with small and medium-sized businesses to create and improve data systems, including creation and refactoring of databases, forms, reports, analytical structures, and integration. Designed and administered inventory, customer, and

point-of-sale relational databases, and created forms and reports to interact with them. Installed and administered small-site network infrastructure and services including web, mail, and authentication.

**Education:**

Charter Oak State College - 2005

Bachelor of Arts, History