dheeraj.chand@gmail.com
+1 2025507110
GitHub: github.com/dheerajchand

Dheeraj Chand

[RESEARCH, ANALYSIS, ENGINEERING] → UNDERSTANDING

Austin, TX (30.2672°N, 97.7431°W)

PROFESSIONAL SUMMARY

Senior Data Engineering & Infrastructure Professional with 15+ years of expertise in building scalable data platforms, ETL/ELT pipelines, and enterprise data warehouses. Expert in cloud architecture, big data processing, and data infrastructure with proven track record designing and implementing comprehensive data solutions, managing cross-functional teams, and optimizing data processing workflows. Deep specialization in Apache Spark, Snowflake, AWS, and data pipeline optimization with experience serving major brands, organizations, and political candidates.

CORE COMPETENCIES

Software Engineering • Data Engineering • Data Analysis • Geospatial / Demographic Expertise • Research & Analytics • Programming & Development • Data Infrastructure

PROFESSIONAL EXPERIENCE

Siege Analytics, Washington, DC | PARTNER 2005 – Present

Data Engineering and Infrastructure Architecture

- Architected data infrastructure processing 15+ billion voter records to support meta-analytical voter file corrections
- Built scalable ETL pipelines enabling analysis of 50,000+ electoral boundaries across all levels of government
- Developed Python boundary estimation algorithm that reduced mapping costs by 75% for 200+ organizations
- Architect enterprise-scale cloud data warehouse solutions on AWS (EC2, RDS, S3) processing billions of records for electoral analytics and demographic analysis
- Design and implement scalable ETL pipelines using PySpark, dbt, and PostgreSQL/PostGIS for large-scale geospatial and demographic datasets
- Develop advanced analytical tools and machine learning algorithms using Python, Pandas, NumPy, and Scikit-learn for fraud detection and spatial clustering
- Manage strategic client relationships across political, nonprofit, and technology sectors using Django/GeoDjango web applications with integrated data infrastructure
- Drive technical architecture decisions for data-intensive applications using Docker, Git, and modern DevOps practices
- Built multi-tenant data infrastructure supporting concurrent access from diverse client organizations

Helm/Murmuration, Washington, DC | DATA PRODUCTS MANAGER June 2021 – May 2023

Data Platform Development and Team Leadership

- Conceived and developed comprehensive data framework using Python, Pandas, and PostgreSQL to clean, validate, and normalize government data from Census, BLS, and NCES
- Architected and built multi-tenant data warehouse and data lake using Snowflake, dbt, and AWS processing millions of records with millions of columns for longitudinal analysis across attitudinal, behavioral, demographic, economic and geographical dimensions
- Modernized legacy ETL processes by implementing dbt and PySpark workflows, reducing processing time by 57%
- Developed advanced data pipelines for machine learning applications enhancing consumer segmentation and predictive modeling capabilities
- Led training initiatives for analytical and engineering staff on open source geospatial technology (QGIS, GRASS, OSGeo) for analysis, segmentation, and visualization using Tableau and PowerBI
- Developed five-year strategic plans for data warehouse architecture using Scala, PySpark, and Apache Spark that became foundation of company's distinguishing products

dheeraj.chand@gmail.com
+1 2025507110
GitHub: github.com/dheerajchand

Dheeraj Chand

[RESEARCH, ANALYSIS, ENGINEERING] → UNDERSTANDING

Austin, TX (30.2672°N, 97.7431°W)

• Led cross-functional teams of seven to eleven engineers, designers, analysts, and external stakeholders using Agile methodologies and modern DevOps practices

GSD&M;, Austin, TX | ANALYTICS SUPERVISOR November 2019 – June 2020

Data Engineering and Big Data Infrastructure

• Transformed the small data team into a big data engineering team, going from working on small datasets on laptops to using Hadoop Clusters and Hive on AWS

• Implemented spatial analysis and consumer segmentation methodologies that revealed new insights about existing customers

• Introduced version control and Agile methodologies to the data team, improving project delivery timelines by 40%

• Managed three analysts, mentoring them in advanced market research techniques and data analysis

• Rewrote the mission and offerings of the department and drafted a plan for how it would integrate with the rest of the strategy team

• Managed accounts for United States Air Force, Southwest Airlines/Chase and Indeed with focus on data-driven insights and infrastructure

Mautinoa Technologies, Washington, DC | SOFTWARE ENGINEER August 2016 – February 2018

Data Engineering and Humanitarian Crisis Solutions

• Architected and developed SimCrisis, a GeoDjango web application using Python, PostgreSQL/PostGIS, and NetLogo for multi-agent modeling and econometric simulations of crisis economies

• Built modular application using Python, Django, and GRASS accepting rules extensions for ethnic strife, different crisis types, supply failures, and disaster scenarios

• Collaborated with senior officers from International Federation of Red Cross, UNICEF, and Chaos Communications Congress to enhance platform using Docker and Ubuntu

• Conceived and developed predictive application using Python, Pandas, and Jupyter to forecast how crisis economies respond to different humanitarian interventions

• Applied agent-based modeling, statistical analysis and machine learning systems for humanitarian impact assessment

• Developed data models and processing pipelines for sensitive humanitarian data

Myers Research, Washington, DC | SENIOR ANALYST August 2012 – February 2014

Data Engineering and Research Infrastructure

• Architected and developed RACSO, a comprehensive web application for pollsters to fully administer research including questionnaire creation, versioning, and reporting

• Led RFP process and analyzed bids from 1,200 vendors before selecting optimal implementation partner

• Built prototype in R for comprehensive polling administration and sample file management

• Provided strategic counsel to Democratic campaigns, political actors, and NGOs through quantitative and qualitative research affecting millions of dollars in campaign spending decisions

• Developed data infrastructure to support comprehensive research operations and client reporting

Progressive Change Campaign Committee, Washington, DC | RESEARCH DIRECTOR August 2011 – August 2012

Political Technology Development and Data Infrastructure

• Conceived, architected, and engineered FLEEM web application using Twilio API handling tens of thousands of calls using emulated predictive dialer for regulated political surveys

dheeraj.chand@gmail.com
+1 2025507110
GitHub: github.com/dheerajchand

Dheeraj Chand
[RESEARCH, ANALYSIS, ENGINEERING] → UNDERSTANDING
Austin, TX (30.2672°N, 97.7431°W)

- Developed IVR polling system for early quantitative research supporting Senators Martin Heinrich and Elizabeth Warren
- Built tabular and graphical reporting system with Python, GeoDjango, PostGIS, and Apache webserver
- Designed survey deployment system facilitating thousands of simultaneous phone surveys, saving PAC nearly $1 million annually in polling costs
- Managed comprehensive research operations for progressive political initiatives and candidates
- Developed data infrastructure to support large-scale political research and analysis

Salsa Labs, Inc., Washington, DC | SOFTWARE ENGINEER January 2011 – August 2011

Political Technology Development with Data Infrastructure

- Maintained and extended comprehensive geospatial analysis and reporting tools for Java-based CRM system used by tens of thousands of users 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
- Integrated mapping and visualization tools for political campaign data analysis interfacing with Government and Activism APIs
- Collaborated with political strategists to translate geospatial requirements into technical solutions
- Handled billions of records with millions of columns in high-performance CRM system with geospatial processing

The Praxis Project, Oakland, CA | INTERIM TECHNOLOGY MANAGER April 2009 – October 2009

Nonprofit Technology Integration and Data Infrastructure

- Led technology operations for multi-million dollar organization while assisting in search for full-time CTO
- Directed all technology decisions and practices for massive multinational non-governmental organization
- Developed comprehensive frameworks for internal and external technology audits
- Led training initiatives for beneficiaries on spatial and Census data analysis for public health research
- Conducted training programs for NGO staff in web development using Drupal, PHP, and MySQL
- Managed technology infrastructure supporting community health initiatives across multiple countries
- Architected and developed 25 Drupal sites to integrate with membership databases, activism CRMs and government agencies, under guidelines from Kellogg Foundation and Robert Wood Johnson Foundation

TECHNICAL SKILLS

DATA ENGINEERING & INFRASTRUCTURE *Big Data Processing* (Apache Spark, PySpark, Hadoop, Hive, Snowflake, dbt); *Data Warehousing* (Multi-tenant Architecture, ETL/ELT Pipelines, Data Lakes); *Cloud Platforms* (AWS (EC2, RDS, S3, Lambda), Google Cloud Platform, Microsoft Azure); *Streaming Data* (Real-time Processing, Kafka Integration, Event-driven Architectures); *Data Pipeline Optimization* (Performance Tuning, Scalability, Monitoring); *Data Governance* (Quality Control, Privacy Compliance, Security, Data Lineage)

PROGRAMMING & DEVELOPMENT *Python* (Pandas, NumPy, PySpark, Apache Airflow, Luigi); *Scala* (Apache Spark, Functional Programming, Big Data Processing); *SQL* (PostgreSQL/PostGIS, MySQL, Oracle, SQL Server, Snowflake SQL); *Database Design* (Schema Design, Indexing, Query Optimization); *Version Control* (Git, GitHub, BitBucket, CI/CD Pipelines); *DevOps* (Docker, Kubernetes, Terraform, Infrastructure as Code)

DATA ARCHITECTURE & SYSTEMS *Distributed Systems* (Microservices, Event Sourcing, CQRS); *Data Modeling* (Dimensional Modeling, Star Schema, Data Vault); *API Development* (RESTful APIs, GraphQL, Data APIs); *Monitoring* (Data Quality, Pipeline Monitoring, Alerting); *Security* (Data Encryption, Access Control, Compliance); *Performance* (Query Optimization, Caching, Load Balancing)