

Alex Narayan (Senior Data Engineer)

email: alex@alexandarnarayan.com | website: gigatexal.blog | LinkedIn: alexandarnarayan | Location: Berlin, Germany

I am a pragmatic and capable engineer with over 10 years of experience in data and data engineering. During my time in industry I have worn many hats: Database Administrator, Python Developer, Site Reliability Engineer, and now Data Engineer. All of my experience from those previous roles has shaped me into a well-rounded and capable engineer that seeks solutions that are manageable, cost-effective, and practical. I also enjoy bridging the communication divide that often exists between business and engineering teams by being able to communicate simply between the two. I am seeking to join a team where my experience and skills can be of value and also from whom I can learn and grow my own talents.

Experience:

Senior Data Engineer Delivery Hero GmbH June 2020 - Current

- Helped to design and operationalize an in-house replacement for Google Analytics
 - Coordinated with advertising teams, front-end and app teams, and back-end teams to define primary-key columns from the noise using the Request For Comment (RFC) approach as well as ad-hoc investigations
 - Created an optimized schema for incremental batched ingestion of app/web tracking events in compliance with GDPR
- Designed, implemented, deployed, and documented a CI/CD pipeline for AWS Lambda functions using Github Actions for a team of Data Analysts in the HR department
 - Gathered requirements via stakeholder interviews
 - Evaluated a number of solutions before deciding on GitHub Actions
 - Implemented a robust pipeline to allow the team to iterate quickly and test with confidence
 - Design was prototyped without access to any production data such that development teams were never privy to any sensitive data
- Helped to migrate existing >5TB Datawarehouse from AWS Redshift to BigQuery on GCP for Datawarehousing
 - Validated requirements and coordinated with stakeholders on updated dimension model
 - Primarily responsible for migrating an integral dimension table used by 100s of reports and stakeholders every day
 - Taking into account BigQuery's new pricing model and the timeline for the migration worked with stakeholders within and without the team to determine columns still relevant

- Conducted analysis to determine the most efficient partitioning and clustering columns
 - Migrated Airflow DAGs from AWS to GCP
- Furthered an internal culture of seeking the most performant, most cost effective approaches to Datawarehousing on BigQuery
 - Constantly keeping up-to-date on the latest developments of BigQuery in both features and pricing
 - Being rigorous in code reviews looking to understand the query-plans of SQL queries and looking for anti-patterns
 - Adding and enforcing partition filtering where applicable
- Improved the interviewing experience by replacing take-home projects with coding challenges in Python and SQL to be worked on in a pair-programming approach during a technical interview as proctored by Coderpad
 - Coding challenges gathered from problems solved in previous sprints that touch core Data Engineering concepts
 - Improved the wording of the challenges for candidates whose native language was not English
 - Created an environment where interviews are a safe place for candidates to show their skills, experience the values of DeliverHero and imagine what it would be like to work together
 - Reduced the time to an offer for a candidate from **1-2 months to 3 weeks**
- Codified a checklist of best practices for both code authors and code reviewers that included a requirement for documentation that led to **> 95%** of critical infrastructure and models being deployed with code comments and additional documentation on Confluence
- Created internal, detailed written and video guides, and conducted a live training session for how to best use Git and Github in the context of the internal mono-repo after noticing many engineers struggling with it

Site Reliability Engineer (SRE) Finleap Connect Oct 2019 - March 2020

- Assisted the senior SRE in building a culture of devops by working closely with the development teams to improve the cloud-native deployment of core infrastructure
 - Created monitoring dashboards in Prometheus and Grafana
 - Consulted internally on SLA/SLO/SLIs
 - Helped to maintain the Kubernetes infrastructure
- Developed datalake consuming streaming data from Google Pub/Sub and Google Cloud Functions using BigQuery as a storage and query engine to reduce operational overhead of legacy reporting stacks and also increase reporting capabilities **> 5X**

Python Software Engineer Finleap Connect (Formerly Figo GmbH) Nov 2018 - Oct 2019

- Primarily responsible for migrating the **> 1TB** primary database for the core product to Kubernetes by researching, evaluating, and ultimately adopting the cloud-native database CockroachDB
 - Wrote all the shell scripts and QA queries to both export the data from MySQL and validate it on arrival
 - Saved **days** worth of engineering time and headache fixing broken MySQL replication
 - Allowed for seamless scaling via Kubernetes primitives
- Developed features for the flagship Python 2 based fintech banking API

Datawarehouse Engineer CorVel Corporation Sep 2017 Apr 2018

- Ran, extended, and triaged internally designed ETL tool written in C# that leveraged Microsoft's SSIS APIs
- Participated in design and deployment of additions to existing facts and dimension tables
- Identified and fixed data quality issues

Database Administrator CorVel Corporation Aug 2013 - Sep 2017

- Assist Senior Database Administrators in configuring, securing, and performance testing new SQL Server instances
- Participated in reviewing SQL queries to enforce best-practices and fix bottlenecks
- Took ownership over all deployments of database objects for software releases and worked closely with internal Quality Assurance engineers in a waterfall model
 - Performed the deployment of database objects for 21+ databases across 3 environments (dev, qa, production) daily with a single production deployment of new versions of core products on the 3rd Friday of each month
 - Provided on-call support for any replication failures or other related issues
- Maintained the internal source control repository (Microsoft Team Foundation Server) and trained other engineers in source control best practices and mitigated merge conflicts
- Developed tools in Python and C# to aid Senior Database Engineers in handling common replication bugs

Education: BSc. Financial Economics - Brigham Young University - Idaho 2009-2013

Skills: SQL; Python; Amazon Web Services (AWS); Google Cloud Platform (GCP); S3; Airflow; ETL; ELT; Kafka; PySpark; Source-Control (Git); Shell Scripting; Linux

References: Available on request