

Ani Chandrashekhhar

Computational Linguist + Backend Engineer

ani.chandrashekhhar@gmail.com | Munich, Bavaria | American Citizen + Blue Card

EDUCATION

Bachelor of Science in Computer Science and Linguistics
University of Illinois at Urbana-Champaign

August 2015 – December 2018

LANGUAGES SPOKEN

English (Native)

German (B2)

Spanish (B2)

PROGRAMMING SKILLS

Languages: Python, GoLang, Java, Godot, OCaml

Concepts: MapReduce (Hadoop), Live Services, Streaming Architecture, Job Management (Jenkins/Airflow)

AWS Tools: Lambdas, Kafka, Kinesis, EKS, ElasticSearch, Cloudwatch

Open Source Tools: Terraform, Bash, Git, Kubernetes, Docker, Spark, CircleCI, Jenkins

EXPERIENCE

Data Engineer

Stylight GmbH - München, Bayern

October 2022 – Present

- Designed and maintained services in Python and Scala, replacing older services and frameworks to fit the continuously changing business requirements
- Managed deployment (CircleCI), maintenance (AWS Stack) and optimization of the main data pipeline in the company - combining, normalizing, and transforming large amounts of data from Kinesis, GCS, DynamoDB, Kafka, REST endpoints, and Google Sheets into cost-effective (partitioned) BigQuery tables
- Reduced costs for BigQuery by **35%** using SQL optimization, different methods of ingestion after attaching in-house monitoring tools

Privacy Engineer

Unity Technologies - San Francisco, California

March 2020 – May 2022

- Acted as a liaison for all engineering teams on Unity-Ads, guiding them to write privacy-first code and coordinated all escalation efforts across teams in the case of a legal incident (GDPR, CCPA, LGPD, PIPL)
- Investigated potential user privacy violations and security threats for over 2.5 billion daily clients
- Engineered automation pipeline in Golang that generated legal audits and compliance data for lawyers, and service metrics for other engineers

Distributed Backend Engineer

Arity - Chicago, Illinois

April 2019 – March 2020

- Rewrote services from daily batch jobs in PySpark to low-overhead stream processing jobs using Flink, resulting in a ~ **40% reduction** in memory usage
- Deployed and monitored multiple cloud applications and created Grafana dashboards with alert and escalation policies
- Wrote performant code using Apache Spark, Cassandra, S3, and Kafka to process 37 million records a day
- Created internal APIs and batch services in Go, Scala and Python to retrieve large DataFrames to/from Cassandra, Hadoop, AWS S3, and Postgres

NLP/Data Science Contractor

AbbVie Pharmaceuticals - Champaign, Illinois

November 2017 - December 2018

- Built database conversion pipeline in Python that normalized millions of pharmaceutical records automatically
- Used dynamic programming and hashing to reduce normalization time from ~**30 minutes to 10 seconds** per write
- Optimized Node.js/Express.js app using worker threads to handle large amounts of reads and writes concurrently
- Implemented privacy protocols to serve sensitive clinical data to internal clients

NLP/ML Lead Student Engineer

The National Endowment for the Humanities - Urbana, Illinois

August 2017– May 2018

- Lead team of students and professors to develop supervised learning model (Naïve Bayes/Random Forest/CounterVec/ TF-IDF) in Python and Wolfram to develop insights on millions of archival records of some of the seminal researchers in cybernetics, including John von Neumann, Heinz von Förster, and Norbert Wiener
- Coordinated team efforts with **MIT**, **Wolfram**, the **British Library**, and the **American Philosophical Society** to extract (OCR), tokenize, lemmatize, and then model large amounts of unstructured data across three decades in Python

AWARDS

- Best medical software** (website eye tracking for paralysis) created during ArchHacks2018 from *Epic Technologies*
- UIUC German Department Award**, recipient of the Robert and Elanor Flinn Scholarship for best research in 2016