

Alexander Dahl

Stockholm
✉ alex@dahl.dev
📄 dahl.dev
[Github](#)
[LinkedIn](#)

Experience

- 2023-01 **Data Engineer, Mentimeter, Stockholm, Sweden.**
Current
 - Moving Clickstream data from the Mentimeter application into Redshift
 - Scheduling dbt transformation with Prefect
- 2022-06 **Sabbatical.**
- 2023-01 Travelling and working on my side project:
 - [bostadsbussen.se](#), a website for scraping real estate listings for the Swedish market. Built with React, Go, SQLite. Hosted on Cloudflare, fly.io and GCP. [Tech breakdown on my blog](#)
- 2021-09 **Data Infrastructure Engineer (Google Cloud), ICA Gruppen, Stockholm, Sweden.**
- 2022-06
 - Set up and designed the data infrastructure stack for ML on GCP (Vertex, Composer & BigQuery) and a CI/CD pipeline with GKE using Terraform and Terragrunt.
 - Built the infrastructure to give Data Scientists a lot of autonomy. The DS team delivered a time series prediction model for inventory levels to stakeholders with minimal help from Data Engineers.
 - Created a devcontainer as a one-stop-shop for setting up an environment with the internal tools as well as GCP. New Engineers went from manually installing their environment, taking several days, to being up and running in an hour by simply pulling a docker image. Promoted this company wide through demos, which resulted in it being adopted by several teams.
 - Held workshops with data scientists about services on GCP to ease their transition from working on-prem to the cloud
- 2020-08 **Data Engineer (On-prem Hadoop cluster), ICA Gruppen, Stockholm, Sweden.**
- 2021-08
 - Developed machine learning data pipelines for personalized marketing that reaches 4 million ICA customers each week and delivers 500 Million SEK in yearly value
 - Used Spark, Airflow, Python and SQL for the above mentioned pipelines which processed the transaction history of all customers at ICA (10B+ rows)
 - Lead initiative to increase performance for critical pipelines. Decrease CPU and memory usage by as much as 50% by rewriting queries and taking advantage of Spark's features
 - Developed CI/CD pipelines for linting, automatic testing, code quality checks and better release process with Gitlab CI, Jenkins and Ansible
 - Decreased release to production cadence from once a week to several per day through a completely automated CD pipeline
 - Spread knowledge about shell and Linux to my team to the point that they were self-operating regarding basic scripting and Linux tasks
 - Revamped the Onboarding process so new developers were up and running and closing JIRA tickets within a week

Technologies

- Data Eng. Python, SQL, Go, Airflow, Spark, Hadoop, Prefect
- DevOps Terraform, Docker, Jenkins, Gitlab CI, Grafana
- GCP BigQuery, Dataflow, Composer, Vertex, IAM, Cloudrun
- General GNU/Linux, Git, bash, networks

Education

- Graduated 2020-05 **M.Sc. Computer Science & Industrial Engineering and Management,**
Linköping Institute of Technology, Linköping, Sweden.
- **Master's thesis:** Product allocation in a warehouse solved with data mining
 - **Student Union:** Was Head of IT for LARM, Linköping's largest career fair, in 2018-2019.
 - **Study abroad:** One Erasmus year at Karlsruhe Institute of Technology in Karlsruhe, Germany

Other

- Checkout my Linux-environment scripts [here!](#)
- Rode a bicycle 10 000km from New York to Costa Rica, [here is a picture of my route!](#)
- I like using emojis 😄