

EXPERIENCE

- **Briink Intelligence GmbH** Berlin, Germany
Senior AI Engineer April 2024 - Present
 - **Responsibility:** Contributed to building Briink's Agentic AI Platform for ESG analysts, enabling questionnaire pre-filling using advanced RAG and multi-step reasoning techniques. Improved system F1 score from 0.65 to 0.85+ for ESG use cases and co-led the rollout of an integrated evaluation platform and LLM observability, enhancing accuracy and transparency, supporting their seed round of 3.8M€.
 - **Tech Stack:** Python Stack(PyTorch, FastAPI etc), Docker, GCP(Cloud SQL, Cloud Run, AI studio etc), Modal, VectorDB (Elastic Search), LLM stack (OpenAI, Anthropic, Langchain, Langsmith, CrewAI, Cohere, Open source rerankers - BGE, MiniLM etc), Segment, June Analytics
- **Stealth Startup based out of NYC** remote, Berlin, Germany
Independent AI Consultant July 2022 - Feb 2024
 - **Responsibility:** Instrumental in developing company's AI Analyst using SOTA LLMs, spearheaded data pipeline creation for ingesting and embedding data for establishing surplus. Established and improved RAG-based chat system and explored model fine-tuning. Additional applications involve response clustering, insight-driven research points, enhancing our SaaS platform's user experience and efficiency. Contributed in scaling the product from zero to 45K\$ MRR.
 - **Tech Stack:** Python (PyTorch, Pandas, FastAPI etc), Docker, AWS (S3, ECS, StepFunctions, ECR, Bedrock), Modal, Heroku, VectorDB (Chroma, Qdrant), OpenAI's API Stack
- **Curalie GmbH** Berlin, Germany
Senior Data Engineer - AI Factory November 2022 - February 2024
 - **Responsibility:** Contributed to the integration of real-time data transformation and ingestion pipelines, devised alongside Data Scientists, to integrate advanced ML models (for anonymisation, trust score prediction on diagnoses, symptom search etc) into our product as part of our interoperability stack used by our clinical partners. Later, led the roll-out of an LLM RAG-based chat service for the medical sector, focusing on enhancing its performance with SOTA LLMs and other fine-tuned models.
 - **Tech Stack:** Python (Spark, PyTorch, Pandas, etc), Kotlin(Spring, HapiFHIRR, etc), Tableau, Docker, AWS (MSK (managed Kafka), EKS, Athena, StepFunctions, EMR, Bedrock etc), Terraform, ClearML, HL7 - FHIR Standard, Chroma VectorDB, OpenAI's API Stack
- **Sunshine Smile GmbH aka Plusdental** Berlin, Berlin, Germany
Data Engineer October 2020 - October 2022
 - **Responsibility:** Spearheaded the deployment of version 2 of our analytics infrastructure, involving with collaboration of other senior engineers. Conceptualized, implemented and maintained the data infrastructure for business needs. Collaborated closely with data scientists and analysts to develop state-of-the-art machine-learning systems for demand forecasting and case refinement predictions. Was part of the company until it got acquired in a 138M€ deal by Straumann.
 - **Tech Stack:** Python, Airflow, SingerIO, DBT, Looker, Metabase, Airbytes, Docker, AWS (EC2, S3, ECS, Redshift, DMS, SageMaker), GCP(BigQuery, Analytics platform) Gitlab, CI/CD, Matillion
- **Glassdollar UG** Berlin, Berlin, Germany
Data Engineer April 2020 - July 2020
 - **Responsibility:** Constructed diverse Extract, Transform, Load (ETL) pipelines enhancing our proprietary "Data Engine" deployed on a self-hosted Airflow cluster. Additionally, we addressed the critical need for temporal tracking of company features, solidifying our data infrastructure targeted at this.
 - **Tech Stack:** Python, Airflow, Pandas, NumPy, BeautifulSoup, Django, RecordLinkage, SciKit Learn, Custom Search Engines, PostgreSQL, DynamoDB, WARC Archiving, AWS (EC2, S3, Eks, RDS, Redshift), Sentry
- **Redstone Digital GmbH** Berlin, Berlin, Germany
Data Engineer October 2018 - March 2020
 - **Responsibility:** Contributed to the creation of a multi-tier proprietary data engine enabling API integration for early-stage startup data collection. Spearheaded internal API/webhook integration and assisted in ML research for investment rankings. Our strategic development of the DA engine, Sofia, catalysed the company's evolution from a VCaaS with zero to 500M€ AUM with multiple sector funds.

- **Tech Stack:** Python, Airflow, Pandas, NumPy, Beautiful Soup, Flask, Sci-Kit Learn, MySQL, MongoDB, HDFStores, AWS (EC2, S3, RDS, SNS), SPIKEs with Neo4J, GraphQL, CommonCrawl WARC formats.

• Udacity

Remote

Reviewer and Mentor

Dec 2017 - September 2018

- **Student Mentor:** Mentored and reviewed projects of hundreds of students enrolled in Udacity's Intro to Self Driving Car ND, MLND & AI programming with python ND.
- **Tech Stack:** Python, Pandas, NumPy Beautiful Soup, Flask, Sci-Kit Learn, PyTorch, TensorFlow, TensorBoard, HDFStores, AWS (EC2 (GPU Computes), S3, RDS, SNS), Neural Networks - Deep CNN architectures, DQNs, Skip-Layer architectures, Fully Convolutional Architectures for Semantic Segmentation

• Karlsruhe Institute of Technology

Karlsruhe, Germany

Student Research Assistant

Feb 2018 - July 2018

- **Detecting lightning strikes in high-speed videos:** Built an end-to-end classifier using Deep CNN Architectures to identify given any image is a lightning image or not. This model provides the basis for the smart data warehousing for the project.
- **Tech Stack:** Python, Pandas, NumPy, Flask, PyTorch, TensorBoard, MySQL, SQL Alchemy, OnPrem data centres for remote computations, Neural Networks - Deep CNN architectures, Skip-Layer architectures (For image classification purposes)

• Ketchupp, D I Info solutions

Delhi, India

Backend - Systems Engineer

Summer 2016

- **Feature Extraction and Detection:** from the text in the context of recorded food deals from various websites to analyze the nature and information about it from the data.
- **Ketchupp Food Bot:** Developed a Facebook Messenger bot using basic NLP and Decision Flow implementation. The bot allowed any Facebook user to get food media updates from the Ketchupp Blog and access many other exciting features.
- **Tech Stack:** Python, Pandas, Beautiful Soup, Flask, SQL Alchemy, NLTK, AWS (EC2, S3, RDS, SNS), Facebook Messenger Platform, REST APIs

PROJECTS

- **Deep-Q Network on flappy bird:** The project shows that the Deep-QN learning algorithm from Deepmind's pioneering paper can further generalized to the notorious Flappy Bird.
- **Lane Detection for self-driving car using semantic segmentation:** Developed a Fully Convolution Network on top of VGGNet using skip layer architecture to label the pixels of a road in a image.
- **Travlr Inc.:** Worked as a Co-Founder. Built a tool for travellers which allows them to share their experience with the world by way of Travlr Stories at a global platform.
- **Image classification on 102 category flower data set:** Developed and trained a Feed Forward Classifier Network on top of VGGNet to label the given image of a flower as one of the categories available in the 102 Category Flower Dataset.

EDUCATION

• CGC College of Engineering

Landran, Punjab, India

Bachelor of Technology in Computer Science; Percentage: 78.2%

Aug. 2014 - May 2018

LANGUAGES AND FRAMEWORKS

Python, C++, C, Scala, Javascript, SQL, NoSql, GraphQL, PyTorch, TensorFlow, Django, Flask, Airflow, Spark, Docker, DBT, Looker, Metabase, CI/CD, Cloud Services (AWS, GCP, Heroku), NodeJS, Solidity

SELECTED COURSEWORK

Machine Learning, Deep Learning, Computer Vision, Reinforcement Learning, NLP, DBMS, Design and Analysis of Algorithms, Application Level Programming (Web & Mobile), Big Data Analytics with Spark,