Andreas Schaler

- Full-Stack and ML Engineer -

andreas.schaler.cs@gmail.com
614 316 8982
https://www.linkedin.com/in/andreas-schaler-ab38b5201/
https://github.com/Andreas3333

Skills

Programming Languages - Python, Typescript, C++, Bash, Django REST Framework, FastAPI, Flask, React, Vue
Cloud and IaC - Terraform, AWS CDK, Cloudformation, SAM, boto3, EC2, ECS, EKS, Lambda, SNS, SQS, Event Bridge, S3
Machine Learning - Pytorch, Hugging Face, Tensorboard, Bert models, Supervised Fine Tuning (SFT) for domain specific NLP, Cloud and local inference serving, vllm, llama.cpp, RAG, sentence-transformers, langchain, langraph
Etcetera - Linux, Systemd, D-Bus, Kubernetes, Docker, Podman, Kustomize, Helm, Skaffold, AWS Serverless, Ansible, Packer, GItlab CI/CD, GitHub Actions, Pydantic, Open API Spec., RabbitMQ, Postgres SQL, Vite, uv, Poetry, MkDocs

Employment - 01/2023 - Present

Nimbis Services — DevSecOps Engineer

- Developed and implemented multi stage container builds and flexible entrypoint service start up scripts for managing dependencies, build configuration, and supported containerized services in development and production.
- Developed and implemented parameterized container builds pre-provisioning LLM model weights at build time, resulting in improved efficiency for LLM model artifact packaging and deployments, reducing startup time for development and production environments by 15%.
- Expanded Containerfile usage to enable local persistent container build caches and persistent CI/CD runners caches reducing build times for local and CI/CD Job environment build time by 10%.
- Developed and implemented CloudFormation templates and Lambdas to automate network resource creation in AWS saving Operations Team 1.5 hours per deployment.
- Implemented Terraform IaC to support containerized deployments of LLM services on AWS EC2 nodes.
- Built DRF REST services using Domain Driven Design (DDD) patterns implementing clean internal and external interfaces supporting modularity, and code organization along with Open API Spec.
- Co-Authored, built, packaged and deployed Vue 3 component library for traditional SPA web application and REST API deployed in AWS EKS Kubernetes cluster and deployments managed by Helm.
- Developed preemptive failure mechanism to fail jobs with missing prerequisite improving job submission by 15%.
- Implemented Library type Helm Charts used for packaging and managing service deployments including service dependencies enabling dynamic service configuration through injection capabilities for configuring sub-chart.
- Developed React based SPA utilizing the AWS Cloud Scape library.
- Designed and implemented Event Driven solution POC architecture implementing a virus scanner which leveraged Event Bridge, S3, ECS, and SQS monitoring quarantine S3 bucket containing potentially malicious files.
- Implemented max concurrency capability enabling full saturation of remote compute resources managed through an asynchronous service manager.
- Implemented Skaffold profiles and profile requirements to expand and support multiple configurations of services for local development and automated production deployment environments.

Projects - 01/2023 - Present

LLM NER REST API — Named Entity Recognition (NER) classifier API

Two colleagues and I completed data acquisition, feature engineering, model training, system design, service implementations, and containerized model artifact deployment on AWS of a multimodal RAG on document system. The NER REST API leveraged a DeBertav2 base model for token classification. This model was selected based on its benchmark performance results at the time on the CoNLL03 NER dataset.

Personal Finances App — Web application for analyzing personal spending habits leveraging fine tuned base Bert

A traditional SPA and REST API project that allows users to upload their monthly transaction data and visualize their spending for the month. The transaction data is classified by a supervised fine tuned Bert model for multi class sentence classification. The classifications by the Bert model are then used as annotations on the data. With the data annotated visualizations are created allowing the user to assess their spending habits.

Extension and Adaptation of AWS RES — Fork and extension project

This system is an extension of the AWS RES service for the company. My contributions to this project involve reverse engineering of the solution to extend the SPA web application (utilizing the AWS Cloud Scape React component library), adding functionality via serverless implementations, altering and extending Cloudformation, building, packaging and publishing of deployment artifacts, extending compute cluster node bootstrapping mechanisms, and implementing hardened automated AMI build processes for cluster node images and VDIs.

Education - 01/2020 - 12/2022

Kent State University, Bachelors — Computer Science (CS) Data Engineering Concentration

Certifications

CompTIA, Security + Certification — 03/2025