Drew Yang

214-727-9608 | drew.yang.dev@gmail.com | Houston, TX | **LinkedIn** | **Github** | **Technical**

Technical Skills

- **Code:** Python, bash, SQL, powershell, Java, C/C++
- **CI/CD:** Github Actions, Azure DevOps, Jenkins, ArgoCD
- **Cloud/IaC:** AWS(Certified), Azure, Terraform, Packer, Pulumi
- **Configuration:** cloud-init, SaltStack, Ansible
- **Container:** Docker, Podman, Kubernetes, Rancher
- **Observability:** Datadog, Grafana, Loki, Prometheus

Experience

DevOps Engineer **July 2021 - Present**

DataJoint - *Provide science operation for neuroscience research.* Houston, TX

- **DataJoint Works** - *A SaaS platform to empower scientists to design and operate data pipelines for their experiments and analysis in a more efficient, scalable, valid, and reproducible way.* [Details]
  - Developed **Terraform** modules to manage public/private network, database, ephemeral resource lifecycle, security, event-driven automation, budget and usage alert, notification, and monitoring across DataJoint's and customers' **AWS** accounts
  - Developed **CI/CD** in **Github Actions** leveraging **spot instances** as self-hosted runner for build, test and deployment.
  - Provisioned, configured and maintained managed and self-hosted **Kubernetes** clusters.
  - Integrated single sign-on, role-based access control, secret manager, and vulnerability scan for **security compliance**.
  - Introduced **OpenTelemetry** to the team and integrated observability with **CloudWatch**, **Datadog**, **Loki**, and **Grafana**.
  - Implemented admin API with **FastAPI**, **SQL**, **boto3**, and **Pulumi** to enable users to manage compute infrastructure.
  - Architected and implemented ephemeral job scheduler in **Python** as core component with **Packer** and **Terraform**.
  - Extended **Jupyterhub** with remote kernel on top of the existing architecture to reuse **idle** compute interactively.
  - Collaborated with the team in an **Agile** approach using **Jira**, also used **Github Project** for open-source projects.
- **DataJoint Core/Elements** - DataJoint Core is an open-source toolkit for defining and operating computational data pipelines. DataJoint Elements is a collection of pre-assembled modules for neuroscience pipelines. [Github]
  - Improved **docker** build consistency, efficiency and security.
  - Improved and automated Python package **releases** by conventional commit and semantic versioning.
  - Integrated **mkdocs** to improve documentation development efficiency and reader experience.
- *DataJoint Works, Core, and Elements improve the research efficiency of 10+ neuroscience labs as of this moment. My contribution technically improves DataJoint Works' robustness, flexibility, and scalability, also automates manual toil through internal and external collaboration to improve productivity in both commercial and open-source development.*

Software Engineer **May 2019 - July 2021**

dataVediK - *Optimized oil and gas operations with machine learning.* Houston, TX

- **Hyper-converged Data Analysis Platform** - *A SaaS platform integrating data management, machine learning, and data analytic services for oil and gas.* [DrillVedik]
  - Implemented **CI/CD** pipelines with **Azure DevOps** and **Jenkins** for build, test, validation, and deployment.
  - Integrated **MLflow** as a machine learning operational pipeline to improve model comparison, versioning, and serving.
  - Set up **Airflow** to automate the data processing pipeline.
  - Developed DrillVedik interactive drilling analytic dashboard with **Plotly Dash**, **Flask**, and **Redis**.
  - Architected and developed the full stack of the prediction task manager web application with **HTML**, **CSS**, **JavaScript**, **Flask**, **Celery**, **RabbitMQ**, **gunicorn**, and **nginx**.
  - Analyzed drilling pump operation data and trained multiple machine learning models to **classify** drilling status.
  - Researched and applied feature engineering on drilling data, and trained a **regression** model for drilling prediction.
- *Although this was an MVP project, I learned and practiced a variety of hands-on skills from software development and deployment, machine learning to cloud computing. It also inspired me about the importance of DevOps through collaboration.*

Education

Southern Methodist University, *Master's of Computer Science* Dallas, TX | Aug 2017 - May 2019

Qingdao University, *Bachelor's of Software Engineering* Qingdao, China | Aug 2013 - May 2017