Drew Yang

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

Technical Skills

- **Code:** Python, bash, SQL, Markdown, powershell, Java, C/C++
- **CI/CD:** Github Actions, Azure DevOps, Jenkins, ArgoCD
- **Cloud/IaC:** AWS, Azure, Terraform, Packer, Ansible
- **Config:** cloud-init, SaltStack
- **Container:** Docker, Podman, Kubernetes(kOps, kubectl, helm)
- **Observability:** OpenTelemetry, CloudWatch, Datadog, Grafana

---

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]
  ◦ Administrated DataJoint's and several other customers' **AWS** account with **Infrastructure as Code** tools like **Terraform**.
  ◦ Configured VPC, Subnet, IAM, S3, EFS, EC2, RDS, Lambda, SQS, SNS, SES, CloudWatch, Route 53 and Secrets Manager.
  ◦ Provisioned and maintained Production and QA **Kubernetes** clusters with **kOps** and **helm**.
  ◦ Architected and implemented ephemeral computational cluster with **Packer** and **Terraform** to support **CPU** and **GPU** usage.
  ◦ Developed **CI/CD** pipelines with **Github Actions** starter and reusable workflows to automate build, test and deployment.
  ◦ Integrated **Jupyter Hub** and kernel gateway as part of the **internal developer portal**.
  ◦ Implemented customer onboarding API with **Flask**, **SQL**, **bash**, **boto3** and **Terraform** to automate infrastructure provisioning.
  ◦ Introduced **OpenTelemetry** to the team and integrated observability with **CloudWatch**, **Datadog** and **LGTM stacks(Grafana)**.
  ◦ Implemented **single single-on**, **role-based access control**, **secret manager**, **vulnerability scan** for security compliance.
  ◦ Collaborated with the team in **Agile** approach using **Jira** and **Confluence**, 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]
  ◦ Collaborated with internal scientists to standardize the support of **Matlab** and **GPU** for several workflows.
  ◦ Implemented **dev container** for **open-source** repositories to allow any collaborators to work on Github Workspace.
  ◦ Integrated **mkdocs** to improve documentation development efficiency and reader experience.
- *DataJoint Works, Core and Elements improve research efficiency of 10+ neuroscience labs as of this moment. My contribution technically improves DataJoint Works' robustness, flexibility and scalability, also automated manual toil through internal and external collaboration to improve the productivity in both commercial and open-source development.*

Software Engineer(MLOps)   **May 2019 - July 2021**

dataVediK - *Optimize oil and gas operations by machine learning.*   Houston, TX

- **Hyper-converged Data Analysis Platform** - *An SaaS platform integrated 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 machine learning operational pipeline to improve model comparison, versioning and serving.
  ◦ Set up **Airflow** to automate data processing pipeline.
  ◦ Developed DrillVedik interactive drilling analytic dashboard with **Plotly Dash**, **Flask** and **Redis**.
  ◦ Architected and developed full stack of the prediction task manager web application with **HTML**, **CSS**, **JavaScript**, **Flask**, **Celery**, **RabbitMQ**, **gunicorn**, **nginx**.
  ◦ Analyzed drilling pump operation data and trained multiple machine learning models to **classify** drilling status.
  ◦ Researched and applied feature engineering on drilling sensor data, trained a **regression** model for drilling speed prediction.
- *Although this was a MVP project, I have learned and practiced varieties of hands-on skills from software development and deployment, machine learning to cloud computing. Also inspired me about the importance of DevOps through the 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