

## PROGRAMMING SKILLS

---

- **Languages:** Python, Javascript, C++, SQL, Java, PHP, SQL      **Frameworks:** Flask, Django, Node, Kafka, Laravel, Wordpress, ScikitLearn, tensorflow, Keras, GreatExpectations, PyTest      **Technologies:** AWS,GCE, OpenStack, Linux Clusters, High Performance Computing

## EXPERIENCE

---

- **RBC** Toronto, Canada  
*Data Engineer* Jan 2023 - Nov 2023
  - Worked in DMO team building ETL pipeline for Fund Accounting data assets. Improved and updated existing ETL pipelines as per business requirements. Migrated some existing pipelines to Airflow. Identified and fixed duplication issue across data assets. Migrated one pipeline to process raw csv files rather than consume data from kafka topic. Compared different libraries to pandas and identified the best to use for our specific use case.
  - **Technology Used** Openstack, Docker, Python, Airflow, GreatExpectations, Pandas and Bash.
- **Dunlap Institute of Astronomy and Astrophysics** Toronto, Canada  
*Lead Software Developer* Nov 2020 - Jan 2023
  - Built and tested pipelines to process radio astronomy data to produce data enhanced catalogs. Managed users and resources on an Openstack infrastructure provided by CADC. Built, maintained and improved web services for two projects cutouts and transient marshall. Updated data engineering pipelines with meta data specific to data being displayed on CADC. Built system for data replication between 3 centres. Created tutorials for data access and software packages developed. Part of the management team reviewing and coordinating work across different universities.
  - **Technology Used** Openstack, Docker, Python, Pandas and Bash.
- **Voices** London, Canada  
*Senior Web Developer* Jun 2020 - Oct 2020
  - Migrated whole code-base from using hard-coded values to use AWS Security Manager. Migrated some python code from 2 to 3. Fixed production bugs, and created code for A/B testing of one of the product features.
  - **Technology used** AWS (S3, EC2, Batch, lambda, KMS, RDS, DynamoDB), Node. Python, PHP, and JS.
- **Intelerad medical Systems** Toronto, Canada  
*Senior Data Scientist* Jan 2019 - Mar 2020
  - Built model to infer follow up detection from medical records using NLP, and also the estimation of follow up time frame. Built the backend and frontend. Ported the code on AWS. Used terraform to deploy on AWS. Made the whole architecture scalable so that adding a new client is easy. Worked on standardizing procedure code maps across different clients. Made a simple model available on AWS for internal people to test.
  - **Technology used** AWS (S3, EC2, Batch, lambda, Cognito, QuickSight, KMS, RDS, Secrets Manager, Sage Maker, DynamoDB), docker. Python, Java, Terraform and Cloud Formation.
- **Quest Global** Bangalore, India  
*Senior Lead Engineer* Dec 2017 - Sept 2018
  - Worked as part of Port Logistics team, which is built a portal for ports. Wrote business logic for ingestion pipeline, whitelist, container availability using pyspark to transform data in postgres DB to Elastic Search so backend and frontend could consume data. Built a few API's for getting container data. Rewrote logic for what Altova was in python to consume data from SQS. Implemented KMS encryption in some AWS Lambda jobs. Led a data science team for some projects like estimating Vessel Turnaround time, gate hour traffic etc.
  - **Technology used** PySpark, Altova, AWS (EC2, EMR, S3, KMS, Lambdas, SQS), PostgreSQL, RabbitMQ, Elasticsearch, hadoop, CDH, pytest, Robot framework, node etc
- **Chikitsak** Bangalore, Canada  
*Senior Lead Engineer* Oct 2015 - Dec 2017

- Data collected via bluetooth from various medical devices and sent to server. Set up 2 testing centres for an eye care hospital with our devices. Made a web front end for the customer to view data in a consolidated form. Dashboard for partners to view stats on diseases and field workers. Some initial trials to predict diseases from measured health parameters. The server was on AWS and had 2 components OpenMRS patient record system running on tomcat and the backend in java. There was a python middleware written in Django. This was changed later. Initially implemented in Google Compute Engine then on AWS.
- **Technology used** AWS (EC2, S3), Google compute engine, MySQL, python (Django and Flask), scikit-learn, Tensorflow, php, bokeh, JavaScript (node, D3, backbone, plotly), pytest, git, ETL scripts in python and SQL, dash.

## • Gramener

Hyderabad, India

*Senior Data Scientist*

*Mar 2015 - Oct 2015*

- The in house product, which was used for data visualization, has a steep learning curve, so to enable understanding, built tutorials on various python modules used in the product. Developed dashboards on PDB stats, compute resources, BO server, sentiment analysis of tweets for various customers.
- **Technology used** AWS, (JavaScript) D3 and backbone, python, Tornado, pytest, git

## • PepperMedia

Bangalore, India

*Senior Consultant*

*June 2014 - Feb 2015*

- Peppermedia acted as a link between the media-company Zee networks and Video on Demand Platform Company Voddler based in Sweden. Interacted with the team in Sweden for the technical issues. Interacted with the media team in Mumbai and helped streamline the admin interface for them. Resolved many issues the media team had with the backend. Old framework of python named 'Planet' was used. Base server was Nginx and deployment was done using fabric.
- Maintained and improved the backend system which was hosted on AWS. Fixed many bugs.
- **Technology used** Python, AWS (EC2, S3), Nginx, Fabric, JavaScript for frontend, jmeter, Ansible

## • SBGrid

Boston, USA

*Software Curator*

*Jun 2012 - Oct 2013*

- SBGrid manages installation and upgrade of a suite of bioinformatics and structural biology software ( 250) for many commercial and educational institutes around the world. Compiled, installed and tested software on diverse hardware for SBGrid. Set up ways to monitor specific websites so we could be aware of software updates instantly. Interacted with developers and users of software over software issues and alpha testing.
- **Technology used** compilers, python, Perl, C++, cvs, virtual machines.

## • Samsung India

Bangalore, India

*Chief Engineer*

*Dec 2009 - Mar 2011*

- Did landscaping on healthcare applications for mobile. Was Chief Engineer and technical mentor for healthcare and bioinformatics team. Set up test cases for bioinformatics and the mobile gateway team. Set up relevant bioinformatics databases for the Pharm DB project.
- **Technology used** Python, Perl, C++, R.

## • EMBL-Hamburg

Hamburg, Germany

*Technical Officer*

*Apr 2003 - Dec 2009*

- Set up remote job submission for ARP/wARP (a model building program) and AutoRickshaw (a structure determination pipeline) on a Linux cluster. Implemented PDB database server ftp mirror in house. Part of Automatic Data Collection initiative (EDNA). Made modifications to the software so it could be implemented on one of the beamlines (X13). Was in charge of installation and maintenance of structural biology and related software on diverse hardware at the beamline. Wrote scripts to automate selection of latest version of software used. Interacted with developers of crystallographic software and helped with alpha and beta testing. Was in charge of setting up required programs the M2M courses conducted in 2005, 2007, and 2009. Worked with researchers in other groups like EXAFS and SAXS.
- **Technology used** Perl, shell, python, Web Services, C++, R.

## EDUCATION

### • Rochester Institute of Technology

Rochester, NY

*Master of Science in Imaging Science; GPA: 3.5*

### • University of Cincinnati

Cincinnati, OH

*Master of Science in Physics; GPA: 3.5*

## PUBLICATIONS

### • Link to Publications