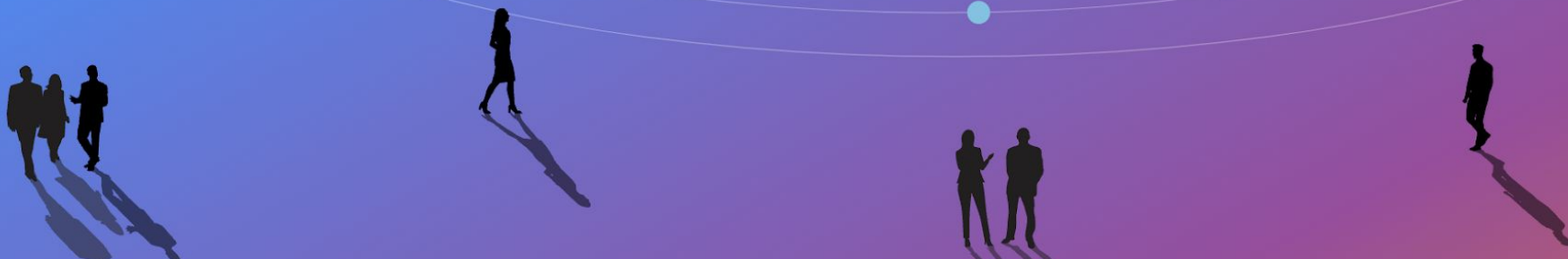


Extend. Expand. Exemplify.

The SGCI team continues to offer its services AND brings you a new Center of Excellence to Extend Access, Expand the Community, and Exemplify Good Practices for CI through Science Gateways.



Machine Learning Hub for Tapis

Dhanny Indrakusuma
SGX3 Summer Internship Report
August 3, 2023



NSF awards
1547611
2231406

Introduction



Dhanny Indrakusuma

Software Engineering Intern

M.S.I.S. Student: Applied ML & AI
School of Information, The University of Texas at Austin

ACI Team



Joe Stubbs, PhD
Research Associate, Manager



Anagha Jamthe, PhD
Research Associate

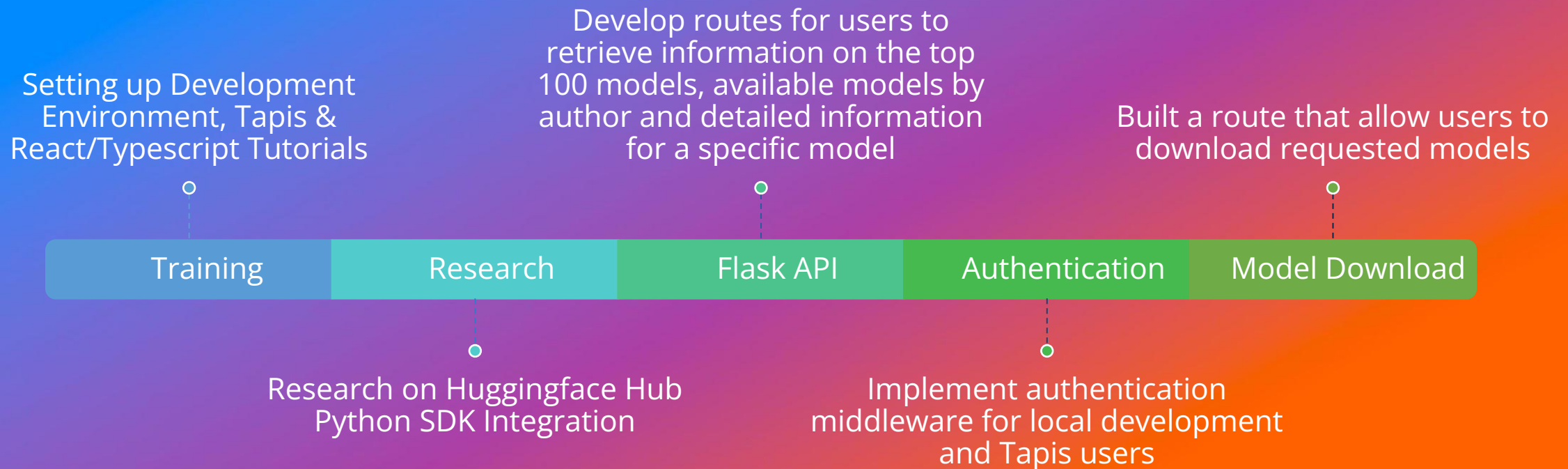


Nathan Freeman
POC, Engineering Scientist
Associate

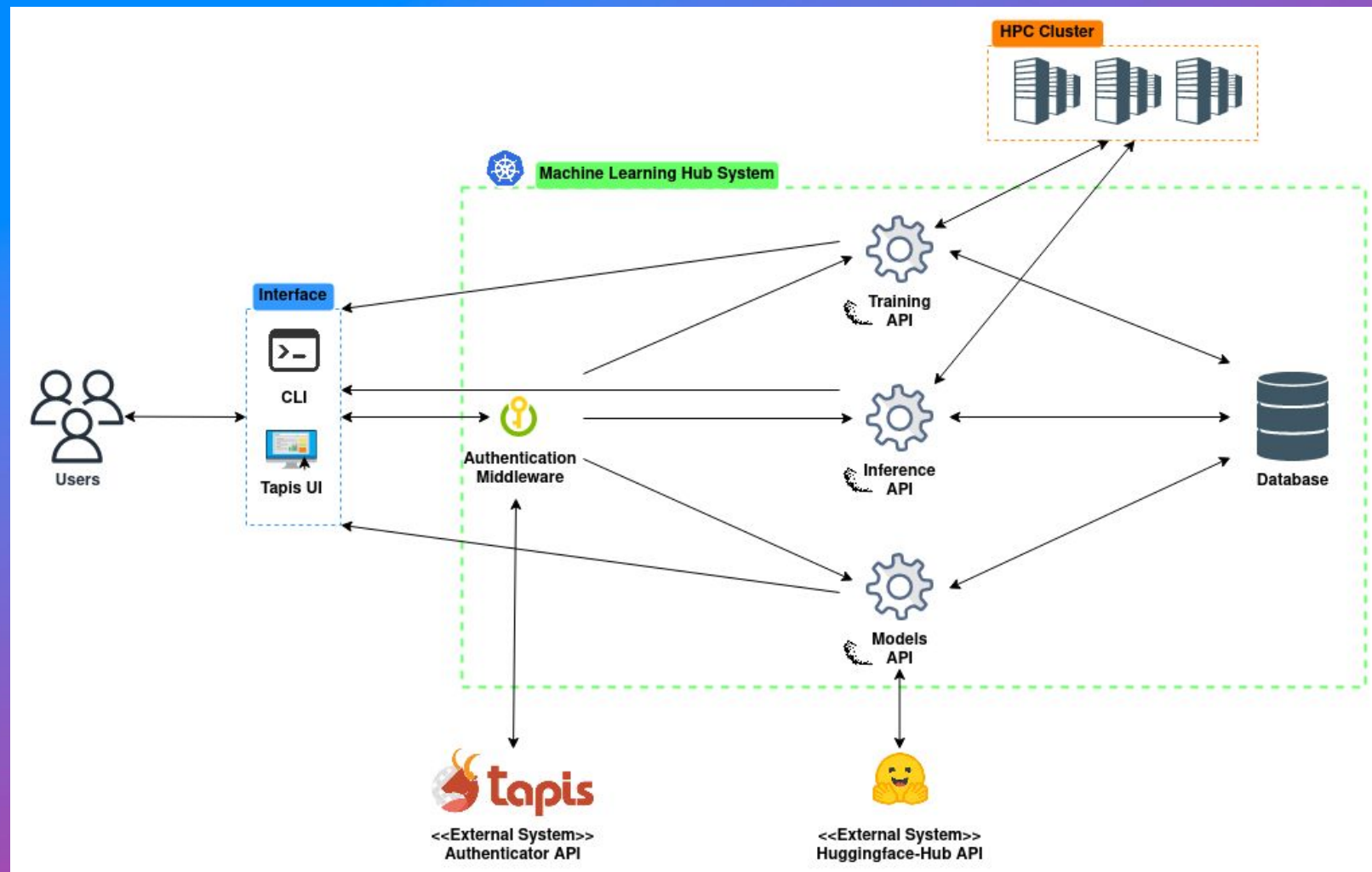
Project Goals

- Machine Learning Hub REST API as part of Tapis
- Integrate Huggingface Python SDK
- Create Models API for non-technical Tapis users
- Authenticate Tapis users for API usage

Tasks Timeline



System Architecture



Models API

- Tapis Python SDK for authentication
- Huggingface Hub Python SDK integration
- 5 endpoints:
 - local development authentication
 - ML models information
 - ML model download

Demo: Models API

The screenshot shows a GitHub repository for 'ml-hub' on the 'dev' branch. The repository contains several files related to a machine learning API, including 'middleware', 'gigignore', 'Mi-hubAPI.yaml', 'Pipfile', 'Pipfile.lock', 'README.md', 'app.py', 'requirements.txt', and 'utils.py'. The README.md file is open, displaying the following content:

Tapis Machine Learning Hub

Machine Learning Hub API within Tapis framework utilizing REST architecture.

Development

This project is under active development. It contains a Flask backend with React and TypeScript frontend, integrated into Tapis Framework. The API is designed to run in a Unix environment.

Installation

Ensure that you have `python 3.10.6` installed in your machine.

Run commands below in your terminal to begin installation:

- Clone repository: `git clone https://github.com/tapis-project/ml-hub.git`
- Run `pip install --user pipenv` to install Python virtualenv management tool
- Once Pipenv is installed, move to cloned folder: `cd ml-hub`
- Run `pipenv shell` to activate the Pipenv shell

The right sidebar shows the repository's metadata: no description, website, or topics provided; 0 stars, 5 watchers, and 0 forks; no releases published; no packages published; and two contributors: nathand and dhanayal.

Challenges

- Starting a large-scale software project from the beginning
- Integrate Tapis authentication
- REST API design
- Returning downloaded model to user

Learnings

- Best practices in structuring large software project
- Version control & code review
- Testing API using Postman
- Advanced Flask functionality & REST API concepts

“Magic” moment

- Insight into back-end development
- Get to know how software engineers work
- Being a tourist in my own city



Next Steps - Fall 2023

- Develop Inference API
- Implement Training API
- Deploy application in the Tapis Kubernetes Cluster
- Connect Flask back-end to Tapis UI's React/Typescript front-end

Thank You

Dhanny Indrakusuma
dhannywi@utexas.edu

<https://sciencegateways.org/>

Extend. Expand. Exemplify.

The SGCI team continues to offer its services AND brings you a new Center of Excellence to Extend Access, Expand the Community, and Exemplify Good Practices for CI through Science Gateways.