

Jake Li

4/30/22

IRIS-HEP Project Proposal

Introduction

ServiceX is a data extraction and delivery service. Users input datasets with specific instructions, and ServiceX opens the files in the datasets and, using the given framework, extracts the data and stores it in an easily analyzable format. ServiceX can be co-located with datasets to provide fast and efficient data reduction strategies, provide data as awkward arrays, and work as a data backend to Coffea and TRExFitter analysis tools.

Planned Work

This project will focus on developing a Jupyterhub plugin that will provide users access to a ServiceX dashboard. With this, users will be able to access ServiceX in a convenient manner outside of the main ServiceX dashboard. The goal will be to provide a possibility to monitor all the users' transforms and give access to the dashboard-like functionality that already exists on the website. A stretch goal of the project will be to add more existing features to the dashboard, such as an easier sorting of transforms or access to log forms in case data delivery goes wrong. This project will be under the mentorship of Oksana Shadura, Gordon Watts, Alexander Held, Mason Proffitt, and Benjamin Galewsky.

Software Deliverables

The languages that will be used for the project will be Python and Javascript. The workspace that the project will be contained in is Jupyterhub, and the end result of the project will be a Jupyterhub plug-in. Python will be used as the structure for the Jupyterhub plug-in, while the use

of Javascript is included for any potential web-based programming that will be done as a part of the project.

Timeline

Approximately 10 weeks from around June 6, 2022 to around August 12, 2022. Proposed total of 10 FTE-weeks with no major activities during that time.

Week 1-2: Study background of project. Become familiar with JupyterHub and ServiceX. In particular, learn about creating plug-ins in JupyterHub and ServiceX's authentication system. Create and present the introduction presentation for the project.

Weeks 3-5: Start to work on developing the Jupyterhub plug-in. Develop the framework for the plug-in using tools learned from weeks 1-2. Create and present the mid-project presentation.

Weeks 6-8: Work out any bugs that appear from initial development of the plug-in. Begin testing created plug-in with template dashboards.

Weeks 9-10: Test created plug-in with ServiceX dashboard, work out any bugs, tidy up look of the plug-in. Create and present the final presentation for the project.