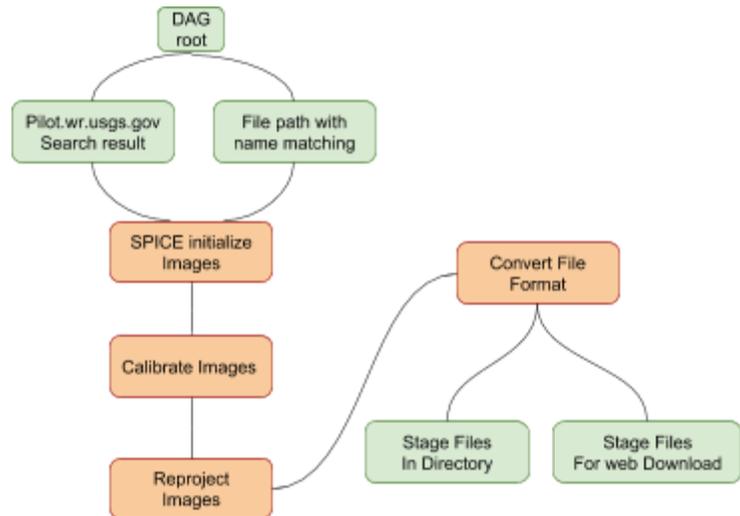


CS486C – Senior Capstone Design in Computer Science
Project Description

Project Title: AirFlow ISIS3 Processing Pipeline	
Sponsor Information:	Scott Akins IT Specialist USGS Astrogeology sakins@usgs.gov Dr. Jay Laura Research Scientist / Geographer USGS Astrogeology jlaura@usgs.gov

Project Overview:

The USGS Astrogeology Science Center develops and maintains a software toolkit for working with planetary images called the Integrated Software for Imagers and Spectrometers (ISIS <https://isis.astrogeology.usgs.gov/>). This software allows planetary scientists to use images captured by remote sensing missions such as Viking or the Mars Reconnaissance Orbiter. They can convert the images from the different formats maintained by the spacecraft to a format ISIS understands. The images can then have SPICE kernels applied



(<https://naif.jpl.nasa.gov/naif/spiceconcept.html>), which defines the spacecraft location and pointing information. After the SPICE data is added, the images can be calibrated to remove noise and other artifacts, map project the images for use in a GIS application such as ArcMap, or convert the file format. Learning how to work with the ISIS toolkit and the idiosyncrasies of the various NASA missions can be a significant challenge for members of the planetary science community who need images they can use in their own research. The Astrogeology Science Center hosts on demand image processing pipelines call Map-Projection-On-the-Web (POW) using a series of ISIS commands run on our processing cluster which converts images to the ISIS file format, initializes them with SPICE, map projects the images, and converts the finished product to the user's preferred file format (<https://astrocloud.wr.usgs.gov>). However, this service always executes all of the steps in the processing pipeline. This project is to incorporate this processing pipeline in the AirFlow

framework (<https://airflow.apache.org>) with a customizable directed acyclic graph (DAG) to allow users to select which of the processing steps to follow in the pipeline. For example, people working with the images may only want the initial step of converting the files from the original mission format to the ISIS3 cube format or to have them converted and SPICE kernels applied so that these files can be used in other workflows such as creating a digital elevation map (DEM) from images with stereo camera angles. The workflow for an end user using the POW service is to first search for the images they want to process in the Planetary Data System Imaging Node database search engine, Pilot (<https://pilot.wr.usgs.gov>). Pilot allows users to search through the remote sensing images captured by NASA images by selecting which planet, what instruments from the NASA missions, where on the planet using a latitude/longitude bounding box, and other criteria (such as the angle the between the spacecraft and the planet when the image was captured). These search results can then be submitted for processing through the POW service.

Project Requirements

Summary- the goal of this project is to take the existing POW pipeline and break down the separate processing steps into separate AirFlow DAG processes to allow users full control over the steps that are run in their processing pipeline. The web application will provide a graphical drag-n-drop interface in which users can select desired processing steps, graphically arrange them into the desired processing workflow, then edit/configure each step with various parameters (if applicable).. The capabilities this interface needs to support are:

- Specify the files to be processed either through a file path or from a CSV download of a Pilot image search result (<https://pilot.wr.usgs.gov>)
- Select which of the processing steps in the POW pipeline to apply to the list of files
- Specify the processing options for each step such as what map projection to use and what file format to convert the finished products to.
- Specify either an output directory (for USGS Astrogeology Science Center team members) or to have the finished files to be packaged for download using the standard POW distribution

Our team here at USGS can provide the existing POW solution, the steps we have taken to implement AirFlow for the processing, and a development environment for running the pipelines. We expect these overall specifications to become more precise as part of the early design and requirements process.

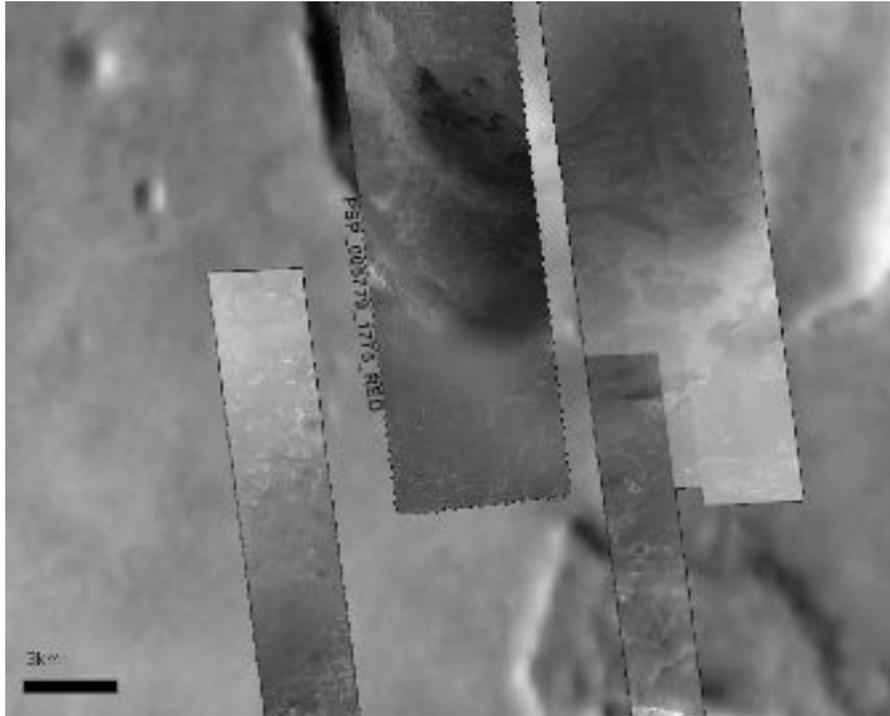
Knowledge, skills, and expertise required for this project

- Experience with Python will be essential as the POW pipeline and AirFlow are all written in Python.
- Some familiarity with the underlying architectures of distributed systems including processing clusters and using Docker Containers to deploy and manage environments.

Equipment Requirements:

- We will provide access to a development environment for running the pipelines. This development environment will have access to the PDS image files to run the processing on, our processing cluster with the ISIS software, and POW processing pipeline installed.

Github URL: <https://github.com/USGS-Astrogeology/PDS-Pipelines>



Example of images processed from POW overlain on a global Mars basemap.

Software and other Deliverables:

Basic deliverables include:

- Design / Architecture documents demonstrating the system being run in as a single instance.
- Docker container for deploying and easy setup.
- AirFlow framework with an interface to specify user customizable processing using a directed acyclic graph. The interface can be an API, web tool, or other method the team comes up with.
- Professionally documented (and tested) source code, posted to the USGS Astrogeology's Github page (we will make a repository for the source).