

Requirements Specification Document

12/04/21



GeoSTAC

Sponsor:

United States Geological Survey (USGS)
Astrogeology Science Center

Mentor: Melissa Rose

Team Members:

Jacob Cain
Zachary Kaufman
Gavin Nelson
Amy Stamile

Version 2.0

Accepted as baseline requirements for the project:

For the client: _____ *For the team:* _____
Signature *Date* *Signature* *Date*

Signature *Date*

Table of Contents

Table of Contents	1
1 Introduction	2
2 Problem Statement	3
3 Solution Vision	4
4 Project Requirements	5
4.1 Functional Requirements	5
4.2 Performance Requirements	8
4.3 Environmental Requirements	10
5 Potential Risks	11
6 Project Plan	12
7 Conclusion	13

1 Introduction

New technological advances have opened opportunities for space research and exploration¹. The majority of this exploration is through launching satellites that orbit planetary bodies, including planets, moons, and asteroids. These satellites collect large amounts of data and take images while orbiting the planetary bodies. The information is then sent back to Earth for the planetary science community to research and better understand our solar system.

As space exploration increases among both federal agencies and private citizens, there is a need for community access to accurate planetary maps and data. Mars is a target for exploration due to its proximity to Earth inside of our solar system. In order to plan for future exploration, it is vital that scientists use the data and images that our satellites and non-human space missions have gathered.

Using these resources, scientists can perform analyses of data from Mars to plan for future missions and scientific discovery. There are many tools that allow scientists to analyze and create maps from the information gathered; however, these tools are not well-developed. The planetary science community processes these images by using software that requires extensive knowledge for complex tools. In addition, the planetary science community must store terabytes of data on their own devices in order to interact with and research the images.

The United States Geological Survey (USGS) Astrogeology Science Center (ASC) in Flagstaff Arizona provides the international planetary science community with analysis ready data. These products support research in planetary cartography, geoscience, and remote sensing. The ASC also develops software for scientific and cartographic analysis of planetary data. The planetary data is distributed to the international scientific community. Finally, ASC provides the community with cartographic products and archiving data².

¹ <https://www.nasa.gov/specials/60counting/future.html>

² <https://www.usgs.gov/centers/astrogeology-science-center/about/our-mission>

2 Problem Statement

The USGS Astrogeology Science Center serves data using a community developed standard to the planetary science community to access Analysis Ready Data (ARD). ARD are data that has already been processed through the complex imaging software at USGS. This significantly reduces the magnitude of data processing and removes the need to download and prepare large amounts of data for research and analysis³.

USGS has already distributed this analysis ready data through the Spatial Temporal Asset Catalog (STAC). STAC is a standard for storing, discovering, and analyzing spatial-temporal data to describe various geospatial information. This provides better indexing and discovering of the analysis ready data⁴.

The STAC specification describes a JSON schema for machine data discovery. The STAC specification also provides an API specification for developers to write human-usable discovery tools. Using these machine-accessible specifications, USGS would like a human-usable data discovery tool that uses web mapping to help users search for and locate data.

In 2019, the USGS assigned an NAU capstone team to develop an interactive web map that supports planetary data. The map is called CartoCosmos, reflecting the capstone team's name. The CartoCosmos team developed a plugin extension for Leaflet, an open source Javascript library for interactive maps. This plugin extension was developed to support mapping of planetary data sets.⁵ The CartoCosmos web map is great for visualizing large planetary image mosaics, but has no support for visualizing the individual STAC assets or STAC catalogs the clients wish to make available.

Thus, the USGS team has tasked Team GeoSTAC with upgrading and adding new features to the web map. This includes the ability to visualize individual STAC asset locations on a map of Mars and load the associated images into the CartoCosmos web map.

³

https://www.usgs.gov/core-science-systems/nli/landsat/us-landsat-analysis-ready-data?qt-science_support_page_related_con=0#qt-science_support_page_related_con

⁴ <https://stacspect.org/>

⁵ <https://ceias.nau.edu/capstone/projects/CS/2020/CartoCosmos-S20/#/>

3 Solution Vision

To solve the problem of interacting with the USGS STAC catalog within the CartoCosmos web map, the team will use the following solutions:

- Query the STAC API endpoint for a given planetary body and render footprints onto the CartoCosmos web map.
- Implement a frontend search functionality that queries information from the STAC API and displays the information of the selected footprints.
- Render Cloud Optimized GeoTIFF (COG) images within CartoCosmos for viewing without the need to download large image data.

The overall vision for this project is to have CartoCosmos display a certain number of footprints for specified planetary bodies. These footprints will be interactive: The user will be able to individually click a footprint and display the information associated with it. The user will have the option to view the COG image associated with the footprint. In addition, the user will be able to select multiple footprints. This will display a table of available footprints of the selected area and allow the user to select a footprint based on the table. Finally, the user will be able to use a searching tool that will allow them to search within the existing rendered footprints to display only the footprints associated with the information searched.

This project will be focused on overlaying the Mars map and the Europa map with the associated STAC catalog footprints. But the goal is to provide a strong framework for other developers to easily implement additional STAC catalogs for different planetary bodies within CartoCosmos.

The following sections of this documentation will discuss the project requirements. This includes functional, performance, and environmental requirements associated with the project.

4 Project Requirements

In order for the GeoSTAC team to implement the proposed solution, requirements must be laid out in terms of domain-level requirements. The domain-level requirements are based on the user's perspective of the front end functionality.

The domain-level requirements for this project include:

- The ability to view an overlay of image footprints within the CartoCosmos web map.
- A query search functionality for the user to search data from rendered footprints.
- An interactive selection tool to select multiple footprints within the CartoCosmos web map.
- The ability to display Cloud Optimized GeoTIFF (COG) images within the web map when selecting a footprint.

The following sections of the project requirements focus on the functional, environment, and performance requirements associated with the domain-level requirements.

4.1 Functional Requirements

The functional requirements discussed below provide details for each high level functionality required for this project. Each of these sections will further discuss the low level functionality needed in order to accomplish each requirement.

4.1.1 Displaying image locations through footprints by use of STAC API dynamically.

Displaying image locations through footprints requires accessing bounding box geometry data within STAC items of a STAC catalog. These STAC catalogs are accessed by dynamically interacting with the STAC API, hosted on Amazon Web Services (AWS). Once accessing these STAC items for a particular planetary body, the bounding box geometries will be used to plot the footprints onto a layer of the Leaflet web map. The process of dynamically accessing a USGS STAC catalog and displaying image footprints can be visualized in Figure 1.

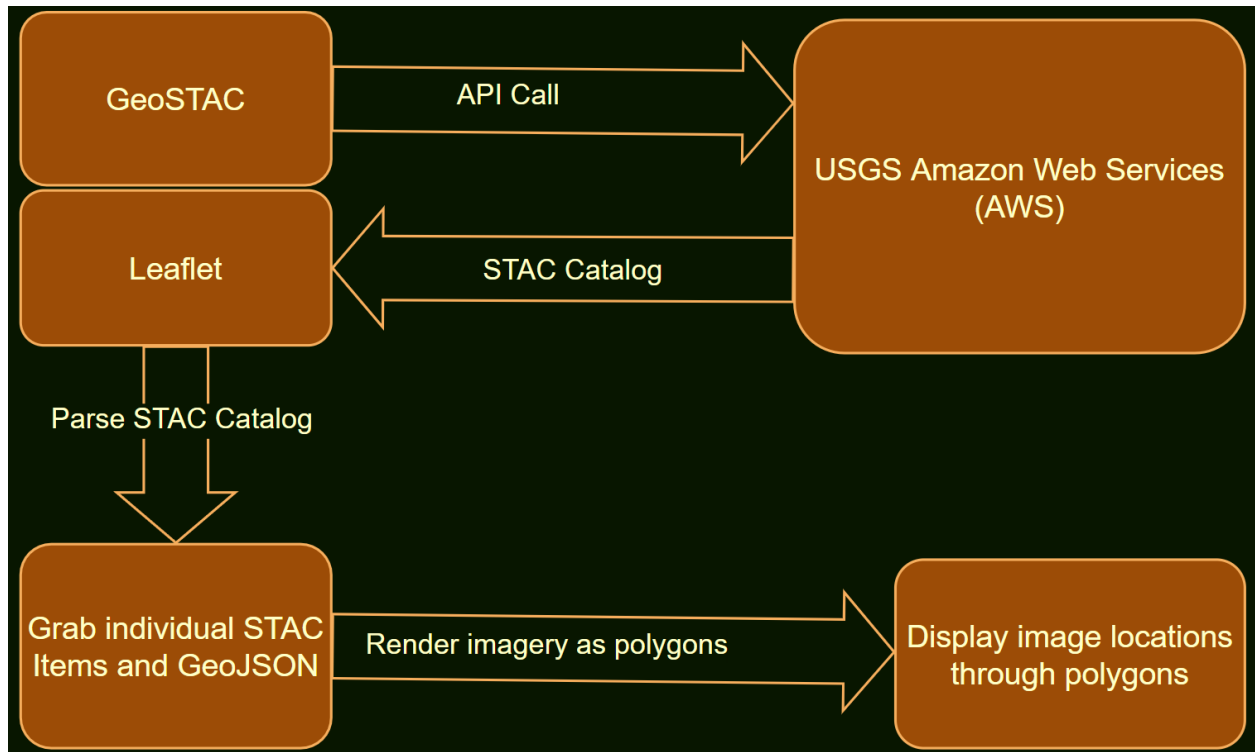


Figure 1

4.1.2 Load Cloud Optimized GeoTIFF (COG) images from a selected polygon

A requirement of this project is the ability to render Cloud Optimized GeoTIFF(COG) images within the Leaflet web map. In order to accomplish this, the Leaflet web map needs to have the functionality to load and render these images. This functionality will allow for a user to select a single polygon and then display the associated COG image. A visual representation of this process is provided in Figure 2.

To provide the functionality of loading a single COG image from a selected polygon, the team will use a Leaflet plugin known as GeoRaster layer. The GeoRaster layer provides the functionality needed for visualizing a single COG within a Leaflet web map.

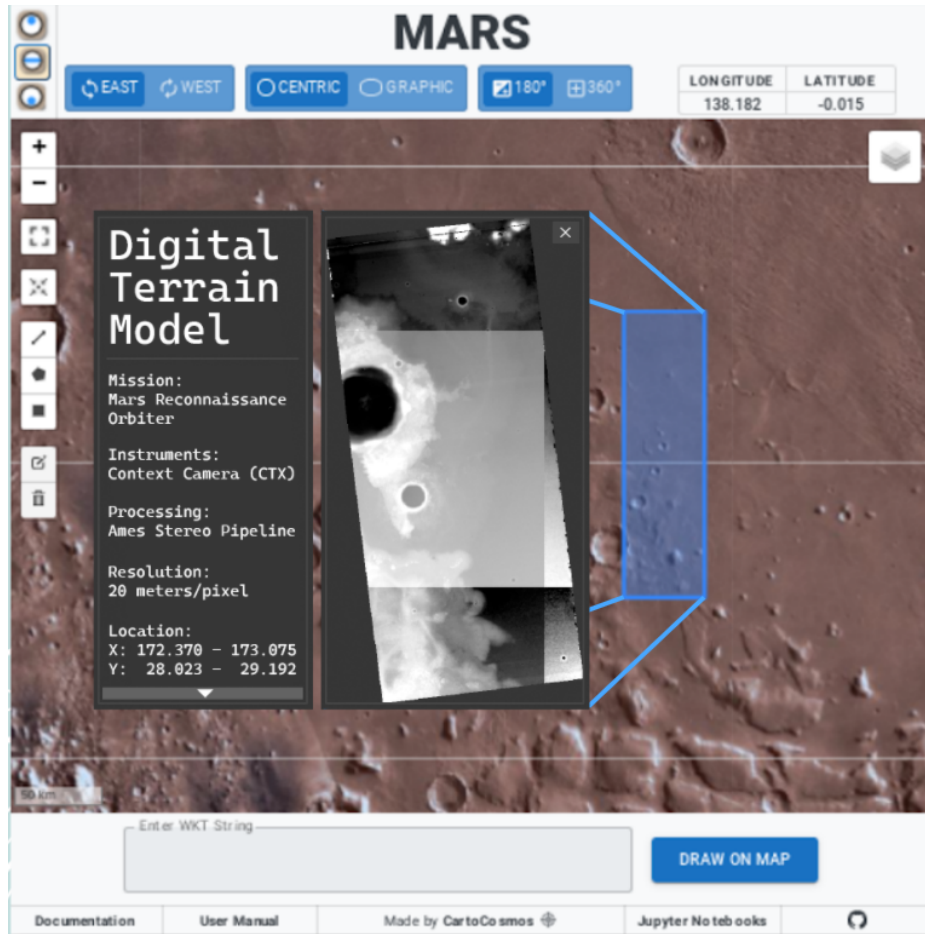


Figure 2

4.1.3 Search and selection functionality of rendered footprints.

The STAC catalogs can potentially hold thousands of footprints for a particular planetary body. In order to sift through the large amount of footprints displayed on Leaflet, searching and selection capabilities are needed to narrow down these footprints.

In order to implement searching, a search box will be added to the CartoCosmos interface. This search box will be used to narrow down the available footprints by searching specific keywords within the metadata. These keywords include the spacecraft mission, image date, and latitude/longitude information. Adding in the ability to search various metadata keywords will allow the user to reduce the displayed footprints to only footprints pertaining to their search criteria.

Selection of multiple footprints will also be implemented within the Leaflet web map. This will provide the user a table of information for all the footprints within a

particular area. The Leaflet Javascript library has various tools to provide interactive selection in the web map. One of these tools is the lasso selection tool. The user can use the lasso tool to draw a circle around the selected footprints of a particular area of the map. A table will then display short descriptions of the selected footprints. When clicking through the table, the user will be brought to a window that will display the entire metadata associated.

4.1.4 Code Documentation

The objective of GeoSTAC is to implement new features onto the CartoCosmos Leaflet GUI. CartoCosmos did a good job of documenting their code when creating their Leaflet GUI. This was to ensure future maintainability beyond their capstone project. To ensure that this project is easily maintainable after the project's completion, the team will need to provide well documented code and update the existing documentation of CartoCosmos. This includes ensuring:

- Every class, function, variables, etc. provides a clear description of it's intended use and functionality.
- The existing CartoCosmos user manual within Jupyter Notebooks is updated accordingly to provide instructions of the new features.

4.2 Performance Requirements

The performance requirements listed below are what will be tested to ensure usability and accuracy within the project. Usability and functionality is a large part of the project since the user only interacts with the graphical user interface. Data accuracy is critical for this project due to the interface being utilized by scientists who need accurate data for their research.

4.2.1 User Testing

This project is intended to be an easy to use interface for scientific research. Therefore, it's various functions must be either familiar to the average user, or easy to learn in a short amount of time. Users should be able to learn the functions intuitively from the interface or from the user documentation.

Selection and search are the two most essential features in terms of functionality the user needs to understand how to execute. Selection may need not be intuitive if the user is not familiar with the Leaflet tool icons on the graphical interface. Once a user

knows how to make a selection, they should easily be able to view or save the data images they need. The searching of metadata keywords will not be intuitive due to the specific criteria associated with the querying. That is, the user needs to know of existing keywords and have examples provided of how to search these keywords. Therefore, documentation will be needed to explain the format of the search query to the user.

In order to test the usability of these features, user testing will be done from within USGS as well as outside of USGS. This is because this mapping tool will be open to the public. The team will (1) provide the test users with objectives and (2) see if they can perform their objectives in the specified amount of time. If the test users encounter difficulties, or make comments or suggestions, their responses should be recorded and used to improve the interface. These user tests will be a useful metric for the team to use to modify any documentation or functionality within the GUI.

The following is a list of objectives that will be provided to the user tester, along with target time tasks should be finished within to consider a test success:

Function	Objectives	Max Time
Area Selection	Highlight the area around an interesting spot on the map.	2 min
Search by Filter Selection	Search for an image taken on a specific date within selection	4 min
STAC Images	View or download a STAC data image.	1 min

Table 1: User Testing Requirements

4.2.2 Data Accuracy

Since this project is a data discovery tool, the team must test that the received API data is accurate. This means testing that the API information is rendering to the correct planetary body as well as checking that the metadata to particular footprints are correct. This will be tested through contract testing. Contract testing checks that the calls/responses are correct between AWS and the application.

4.3 Environmental Requirements

The environmental requirements listed below describe in detail the client's constraints for this project. These include utilizing an existing UI platform that uses a particular language and libraries as well as use of a specific API for obtaining the necessary data to be included in this project.

4.3.1 Building on existing Leaflet CartoCosmos Web Map

Use of CartoCosmos Leaflet web map is required by the client. This project focuses on extending the capabilities of this existing interactive map. CartoCosmos is a web-based interactive map for planetary data, designed to provide astro geological data to the planetary science community. The source of CartoCosmos is available to the current team, and it was made using Leaflet.

Leaflet is an open-sourced, interactive map library for JavaScript. It can dynamically load map tiles for easy navigation and zooming over a large map. Typically, Leaflet is used for displaying Earth on map webs. The CartoCosmos project expanded on the base functionality of Leaflet to show maps of other planetary bodies. The ability to show geography of other planetary bodies is a central feature to the current project.

Leaflet is modern and extensible by a large library of plugins. The project will use Leaflet plugins to load in and display data, as well as enable users to interact with and find the data they need.

4.3.2 Utilizing The SpatioTemporal Asset Catalog (STAC) API

The team will use a SpatioTemporal Asset Catalog (STAC) to obtain the necessary imagery files and data. This is an environmental requirement due to USGS already using STAC to distribute analysis ready data.

These ARD STAC catalogs are located within the Amazon Web Services (AWS). In order to obtain the STAC catalogs, the application must dynamically interact with the API. The interaction with the API will be done through the use of Javascript fetch requests to dynamically obtain the STAC catalogs. These fetch requests will be implemented when the user displays a particular planetary body.

Dynamically accessing the API will allow the web map to immediately display additional catalogs as they are added to STAC. This is important as new data and imagery may provide a higher level of scientific accuracy, or cover new areas that prior iterations of the data did not.

5 Potential Risks

There are a few concerns regarding the potential risks associated with development tools and data that will be used. These risks require a concrete understanding of the potential problems in order to provide a satisfactory product. The risks that have been identified are discussed below.

Amazon Web Service (AWS) STAC Catalog loading times

With new STAC catalogs being added, there will potentially be tens of thousands of footprints rendering. Based on the size of STAC items for a particular planetary body, this can potentially increase the load time of the data footprints significantly. Loading a moderately sized catalog of a thousand items can take up to ten seconds. For many users, a loading time of ten seconds is enough that they might leave before the loading is finished, or at least think something is wrong. Furthermore, the client has specified that the application will need to work with larger catalogs of over thirty thousand items. Therefore, the risk of high loading times must be addressed during the development. There are a couple of potential solutions. If the application can load a small selection of the data instead of a whole catalog, it will eliminate potential lag while loading. This subset of items could be based on the area the user is zoomed into, a selection that has been drawn, or a STAC item limit set by the user. Another alternative if a large number of items are to be loaded is to show the user a progress bar to inform them the application is still working while they are waiting. Ideally, having a combination of these solutions would be optimal so that the user is not restricted to long wait times.

CartoCosmos Code Base

The risk in using the CartoCosmos code base is that team GeoSTAC will be building off of a code base that is no longer being supported or maintained. If there are any underlying bugs or problems that exist in the CartoCosmos code base, there would be no support from the previous CartoCosmos team. With this risk, the concern of not only quality but functionality of the existing code and usability of the CartoCosmos application arises. If any bugs or potential problems are found, the team will need to make the necessary fixes in order to provide a working product.

Risk	Likelihood	Severity	Solution
------	------------	----------	----------

Amazon Web Service (AWS) STAC Catalog loading times	Moderate	Moderate	Inform the user that the selected STAC Catalog may take a few minutes to fully load and become usable.
CartoCosmos code base	Low to Moderate	Moderate	Team GeoSTAC will provide necessary support to fix any bugs found.

Table 2: Risk Analysis

6 Project Plan

The plan for updating the CartoCosmo planetary mapping with STAC is split between two semesters, Fall and Spring. During the Fall semester the team has been focused on project planning. The Spring semester will be the development phase. The team is currently working on the end of the planning phase, as seen in Figure 3 below.

As shown in the figure there are two major tasks remaining in the planning phase of the project. The technological feasibility study was completed earlier this semester. The team is currently working on the technological demonstration and requirements document. In order to complete the requirements documents, USGS will approve and sign off on the new features of the CartoCosmo planetary map. This document will act as a contract between team GeoSTAC and USGS for what needs to be implemented during the development phase. The team will display key components of the project during the technological demonstration at the end of the Fall semester in early December.

As the project progresses into the Spring semester, the team will enter the development phase. As shown in Figure 3, the team will begin with the integration of the STAC catalog inside of the CartoCosmo Leaflet web map. It is crucial that the team can acquire the STAC catalogs from the API before moving forward in development. The team's plan for after the integration of the API is to render the footprints and COG assets inside of Leaflet. The team plans to use different existing plugins for Leaflet to display the footprints and COG assets that are acquired from the STAC catalog overlaying the CartoCosmo Leaflet map. The main goal is to be able to click one or more footprints to bring up the corresponding COG assets inside of Leaflet. The next step in the development stage will be implementing search functionality to allow the user to search within the rendered footprints. At this point, the team will need to assess how much time is left before implementing any additional features. Once the team has completed the potential stretch goals, GeoSTAC can begin testing the new mapping

application for the performance requirements listed above and fix any bugs before delivering USGS the best product.

GeoSTAC Project Schedule

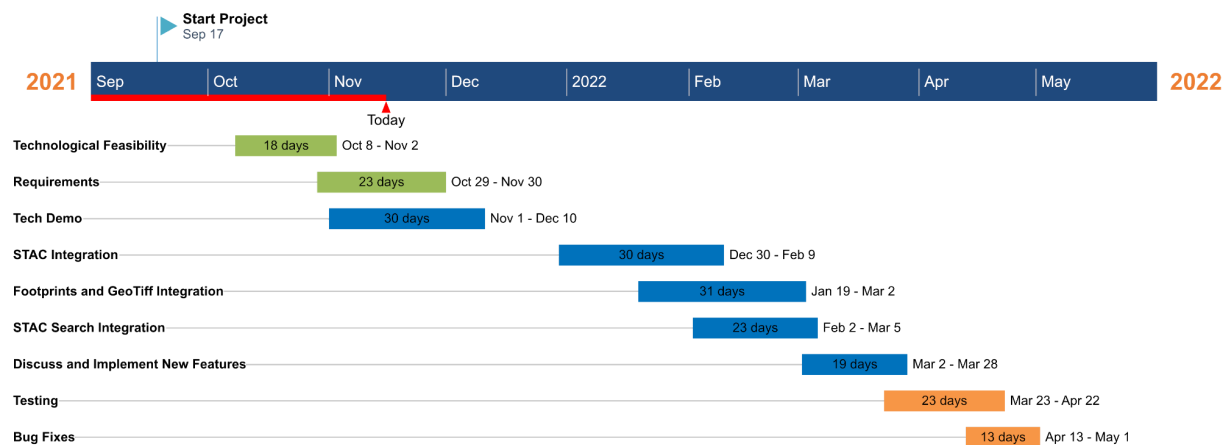


Figure 3: Team Schedule

7 Conclusion

As the need for accurate planetary data and images increases for the scientific and academic community, it is more important than ever that there is software to easily access this data. The USGS Astrogeology Science Center in Flagstaff Arizona provides the international planetary science community with new knowledge of our solar system. Without this data, new planetary exploration missions would be nearly impossible. The USGS Astrogeology Science Center needs a way for the planetary science community to interact with the analysis ready data from their STAC API.

In order to solve the need for a new interface that utilizes the USGS STAC API, team GeoSTAC will update a previous Leaflet web map created by NAU Capstone team CartoCosmos. Team GeoSTAC will be using this CartoCosmos web map to develop new features to allow users to access analysis ready data from the USGS STAC API. In order to get a better understanding of the requirements of the project, GeoSTAC created this requirements acquisition document. The requirements of the project have been broken down into 3 sections, functional, performance, and environmental. The team has investigated potential risks and provided solutions. Finally, GeoSTAC described their plan for developing the project over the next year. This document will be

used as a contract between the clients and team GeoSTAC explaining the expected outcome of the project at the end of the team's senior year.