Requirements Document

November 16, 2020

LOST EXPRES

Sponsor: Joe Llama

Mentor: Volodymyr Saruta

Team Lead: Brooke Caldwell

Architect: Jared Cox

Customer Communicator: Olivia Thoney

Release Manager: Ian McIlrath

Meeting Recorder: Austin Bacon

Client: ____________________

Team Lead: ____________________

Table of Contents

1. Introduction

The Star Wars Saga, 2001 Space Odyssey, Alien, are all examples of how people's obsession with the existence of familiar life in the universe has made its way into pop culture and affected entire generations. At Lowell Observatory, Dr. Joe Llama and his team are discovering what was once only science fiction. They are gathering millions of data points on the sun every minute of every day to get a better understanding of what makes the sun support life on Earth. Through their research, the team is working towards finding another sun-like star that will lead them to another Earth-like planet. While Earth is the only planet of its kind right now, our client is searching for the familiar worlds that have captivated the imaginations of generations.

Dr. Llama's research about the sun relies heavily on gathering radial velocities. The radial velocity of our sun is the speed at which it moves towards or away from the Earth. Astronomers can use this data to calculate how much the radial velocity of a star changes in order to detect an orbiting planet pulling on the star. Our client is using the Lowell Observatory Solar Telescope (LOST) along with Yale University's newly developed EXtreme PREcision Spectrograph (EXPRES) to record extremely accurate measurements on the sun.

Led by Dr. Llama, the research team is gathering approximately 40GB of data every day. The LOST, along with the EXPRES, is taking 90-second exposure pictures for 8 hours every day. This data is currently recorded in hard-to-read FITS files that are not easily accessible. This is the problem that our client has recruited our team, LOST EXPRES, to solve. We will create a web application for Lowell Observatory that will allow our client, his fellow astronomers, and members of the public to view comprehensive and interactive graphs of the data gathered. Our team consists of Brooke Caldwell (Team Lead), Jared Cox (Architect), Ian McIlrath (Release Manager), Olivia Thoney (Customer Contact), and Austin Bacon (Meeting Recorder). our client's website will feature role-based permission admission, fast database queries, downloadable files, and dynamic graphs of radial velocity and one dimensional and two-dimensional spectrum. All features will be implemented while maintaining an intuitive user interface that is open to the public.

The purpose of this document will be to break down and enumerate both functional and non-functional requirements for this project. The aim of this is to have the document be the contractual basis for what the team is expected to deliver; specifying features and performance details that will be the basis of acceptance for our product. If the product meets the stated requirements agreed to by the client, the project can be closed. In every project, there is a set of obstacles the team must overcome in order to deliver a viable product; the following section explains these obstacles in more detail.

2. Problem Statement

Our client's search for Earth 2.0 revolves around the information gathered from the Lowell Observatory Solar Telescope (LOST). The telescope is pointed at the sun, equipped with EXPRES every clear day. EXPRES is able to continuously capture the sun's spectra. It has an exposure time of approximately 90 seconds, this means that a complete spectrum of the sun is captured every 90 seconds for around 8 hours a day. This data is sent in Flexible Image Transport System (FITS) files from the telescope to servers at Lowell at around 8 o'clock each night. Then, through our client's own Python script, the data is being stripped and reduced for manageability by our client. This process finishes at about 10 o'clock at night. Our client is able to create and view graphs of the data through his own Jupyter Notebooks. This entire process is depicted in Figure 1.0 below. With the new addition of the LOST, the observatory is now a 24/7 observatory. The all-day cycle of data collection makes this project both unique and challenging.

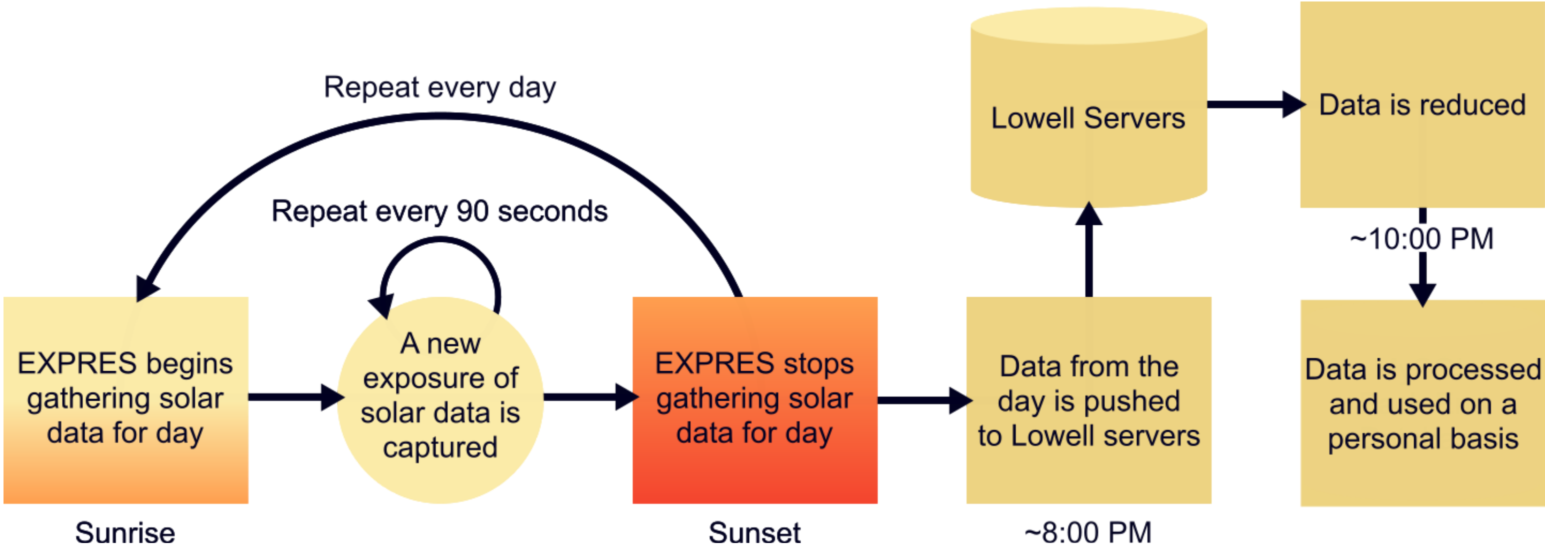

Figure 1.0: The workflow of gathering solar data each day

Currently, the workflow ends when the data is finished being reduced. At this point, only the client is able to interact with the data and any graphs he creates or analysis he conducts is private. This is where the problem arises. The client would like to extend the workflow so that reduced data is published publicly. In addition, he is not just looking for a site where astronomers can download raw data. The client is looking for a web application that will allow the public to see and interact with graphs of the data gathered from LOST. This will include three different types of graphs: radial velocity, 1-dimensional spectrum, and 2-dimensional spectrum. The user interface needs to strike a balance between being simple enough for the general public to explore, while still being detailed enough for astronomers to use in their research. After understanding what the client is asking for, the current problem with the Earth 2.0 project can be broken into three main challenges to solve.

- Lack of data visualization. The client has been able to construct his own graphs from the data using personally written Python scripts. However, these graphs are static and require an understanding of Python and refactoring the client's codebase.
- Lack of publicly accessible information. As previously mentioned, the data from LOST is currently being handled personally by the client. Lowell Observatory

values transparency and cooperation with the public. They hope to educate people about research taking place at Lowell, which includes the project involving LOST and EXPRES. Currently, there does not exist any public database or dashboard for viewing data gathered from LOST.

- Lack of process for sharing LOST data with other astronomers. Along with educating the general public, Lowell Observatory strives to collaborate with other observatories and researchers across the world. Spectra and radial velocities gathered from LOST right now are not easily accessible for download by anyone outside of Lowell.

The client's Earth 2.0 project using LOST and EXPRES is generating hundreds of GB of data each week. However, a process is not currently in place to compile, visualize, or publish this data in any form. This directly contradicts Lowell Observatory's mission, presented on its site, "to maintain quality public education and outreach programs to bring the results of astronomical research to the general public." To help the client solve the challenges presented and to assist Lowell Observatory in achieving its mission, the SOLAR EXPRES team has devised a custom software solution that addresses each problem.

3. Solution Vision

After determining which areas the client and his team are lacking in, LOST EXPRES began brainstorming effective solutions. It was important to keep in mind during this process that the client was not only our client. Instead, the solution would need to be developed with Lowell Observatory, public users, and astronomers in mind, along with our client himself. Through conversations with the client and research into similar software solutions, LOST EXPRES has a solution vision. The team plans to develop a custom web application that will present visualizations of the data gathered from LOST within a modern and intuitive user interface. Users will be able to visit a site, which will be a subdomain of Lowell Observatory's main site, and view a plot of the radial velocities of the sun for each day. For each radial velocity, users will then be able to select additional details about the data. This will include a 1-dimensional spectrum graph, a 2-dimensional spectrum graph, downloadable text files of the radial velocity and 1-dimensional spectrum data, and finally downloadable FITS files of the 2-dimensional spectrum data. Now, instead of the data processing finishing at 10:00 PM, processed data will be entered into a database. The database will serve as part of the backend which, together with the front-end, will allow users to view and request results. The updated workflow is presented in figure 1.1.

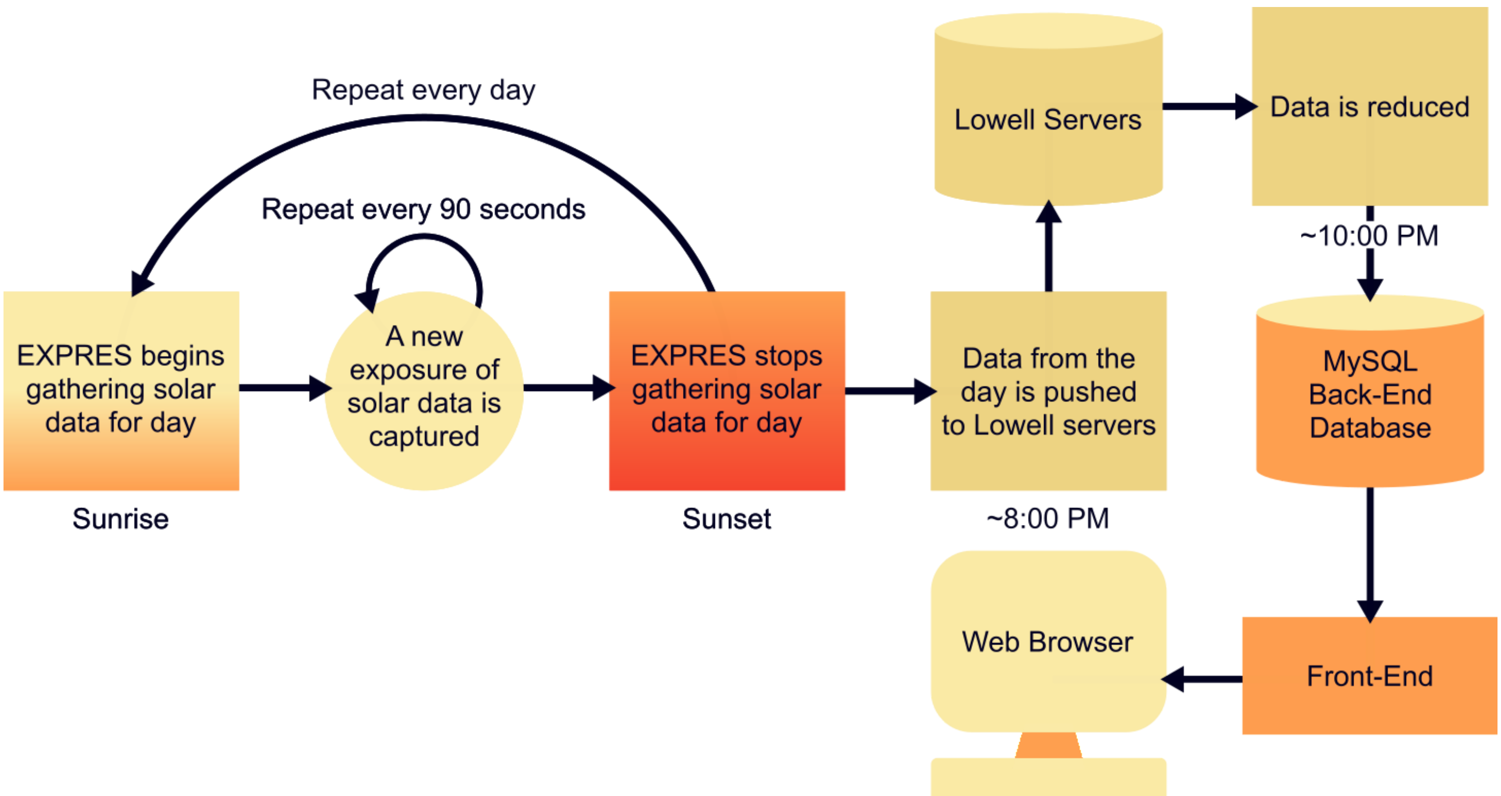

Figure 1.1: The updated workflow of gathering solar data each day to be sent to a SQL database

The key features that will solve the client's current problems are presented below.

- Publicly accessible application. Creating a web application that is available to the public is priority number one. In order to achieve the mission of Lowell Observatory to present astronomical research to the public, we need to create an application that does just this. Our application will contain relevant data from the Earth 2.0 project. Of course, the other features we implement will go above and beyond simply serving the public information. Additional features will introduce modern user interfaces that encourage interest in the general public and will support more collaboration with other astronomers.

- Interactive graphs. For the radial velocity graph, users will be able to hover over data points to see additional information. They will also be able to zoom in and out of selected areas of the graph. The 1-dimensional spectrum graphs will allow the user to customize the y-axis with the option of plotting wavelength vs. flux, or wavelength vs. signal to noise. Again, users will be able to zoom in and out in order to view the range of data they desire. The 2-dimensional spectrum graphs will have all the same features of the 1-dimensional spectrum, while also allowing users to select the order(s) that are plotted. These interactive graphs solve the client's problem of not having a convenient data visualization tool.

- Downloadable files. In addition to presenting data through graphs, our solution will offer users the opportunity to download raw data from LOST. Offering raw data to the public solves the problem of not having publicly accessible information. For other researchers not working at Lowell, this feature will be invaluable. It gives astronomers everywhere access to raw data collected from LOST. Thus, this feature provides a process for sharing LOST data with other astronomers.

- Role-based permissions. While discussing the problem and solution, two groups are being considered: the general public and other researchers. This will be reflected in our application through role-based permissions. We will implement three different roles. The most basic role is a public user. Public users will be able to access the application without logging in or providing any credentials. They will have access to all types of graphs and downloadable files. However, they will not have access to all the data. Only data that is deemed "public," will be available. This will just include all data from the time range that is marked public. Next will be the researcher role. Researchers have access to all data that has been collected from LOST. Finally, the administrator will be able to mark data as public or private and will approve new researchers. These role-based permissions ensure the unique problems of each type of user is solved. It also encourages collaboration between the public, Lowell, and other researchers. These roles along with the specific features of our solution vision are presented in Figure 1.2.

	Access Data	Visualize Data	Set Data Visibility	Upload Data	Approve Members
Administrator	✓✓ Can access and download all data	✓✓ Can visualize all available data	✓ Can assign public visibility	✓ Can manually upload data	✓ Can approve member registrations
Researcher	✓✓ Can access and download all data	✓✓ Can visualize all available data			
Public	✓ Can access and download public data	✓ Can visualize all public data			

✓: Access to most data ✓✓: Access to all data

Figure 1.2: Levels of access per users

Our solution will not only solve the current problems the client is experiencing but will implement features and designs that improve the user experience and focus on convenience. The final application will assist Lowell Observatory in its mission to provide the public with astronomical research, encourage scientific advancement through collaboration with other researchers, and ultimately prove a useful tool for our client in his search for Earth 2.0.

4. Requirements

4.1. Functional

4.1.1. Levels of Access

The website will need to have 3 levels of users: the general public, the team members, and the Admin. Each access level also builds upon the previous level. The general level will not be required to log in first, and are only permitted to view public data. The team member level will require a login and can see all data. The Admin level has the same viewing permissions as team members, but also is able to make data public and add additional team members. These 3 levels of access can also be thought of as the three different types of end users for this project.

4.1.1.1. General Public

The general public role is the default role given to users visiting the site. General public users will be directed to a default landing page. This type of user will not be able to log in, and can only see data that has been marked as public.

4.1.1.1.1. No Login Required

There will not be any logging in for the general public role. Users will simply be directed to the website's landing page, and are able to navigate through all publicly available datasets.

4.1.1.1.2. Selected Amount of Data

Users that are not logged in will not be able to see any data that has been set to private. Public users will not be able to select datasets which are private.

4.1.1.2. Team members

The administrator will have the capability to assign the team member role to chosen users. The team member role will only be given to other astronomers at Lowell or a separate astronomy organization. Before the role can be accessed, a team member must submit a registration form to the administrator. Team members will be required to log in when they access the site. Once logged in, they will be able to see both privately and publicly set data.

4.1.1.2.1. Registration and Logging In

Team members will have to register with a form before they can log in. To register, the user will navigate to the team member application page and enter their full name, the institution they belong to, and their email address. When this form is submitted, the administrator will receive an email notification, and the Admin will decide whether or not to register this person as a team member. If the team member is approved, they will receive an email notification along with their new password.

Team members must log in first in order to have access to any private datasets. Users can navigate to the team login page from the website's landing page to gain these permissions.

4.1.1.2.1.1. Personal information

In order to register on the website, the user will need to enter information that tells the administrator who they are. This information includes their full name, the institution they are employed at, and their email address.

4.1.1.2.1.2. Auto Generation of Password

After a team member has been approved, an email will be sent to the team member containing their new password. All generated passwords will be between 12-15 characters, and contain at least 1 number, uppercase letter, and lowercase letter. An algorithm will be developed to handle the generation of this password. A password will automatically be generated for a new team member when the Admin approves their request for the team member role.

4.1.1.2.2. All Data Visible

Team members will be able to see data that has not yet been made public. Private data will be highlighted differently from public data. This will allow team members to tell which data isn't public yet.

4.1.1.3. Admin Access

Just like team members, the administrator will be able to access all datasets. The Admin will have the additional capability to add more team members and set data to either public or private.

4.1.1.3.1. Registration and Logging In

The Admin will need to log in using the exact same method that team members use.

4.1.1.3.1.1. Personal Information

The administrator will have the same types of personal information tied to their account as team members.

4.1.1.3.1.2. Auto Generation of Password

A password will be generated for the Admin, with the same specifications as the Team Members. The password will be 12-15 characters, contain at least 1 number, uppercase letter, and lowercase letter. Once generated it will be sent to the Admin email address.

4.1.1.3.2. All Data Visible

Just like team members, the Admin will be able to view all data uploaded to the site.

4.1.1.3.3. Grants Team Member Permissions

People who want to be team members will have a basic form they can fill out on the website. The form will require them to enter their full first and last name, the institution that they belong to, and their email address. Once submitted, the Admin will be notified of the new application through email. Approved team members will receive an email in their inbox informing them that they have been accepted. They will also be given their auto-generated password for their account.

4.1.1.3.4. Decides Private or Public

The Admin will be able to dictate what data can be seen by everyone, and what data can only be seen by team members. By default, all data uploaded to the site will be set to private. The Admin will have access to an interface that lets them sort datasets by public or private. This interface will allow the Admin to set multiple datasets from private to public, or vice versa.

4.1.2. Uploading Data

The uploading of data is a process that needs to be handled once every 24 hours. Currently, the EXPRES machine will send all of its data to the servers at 8 pm. By 10 pm, the client's script has finished stripping the data. When this stripping process has finished, the data will begin the process of being uploaded to the website's database. This process must be completed by 8 pm the next day in order to maintain a 24 hour uploading schedule.

4.1.2.1. Automatic Uploading

The data from the EXPRES machine will be uploaded to the website's database every 24 hours. Currently, the EXPRES machine will output its data between 8 and 10 pm, and the website will add an additional step to this workflow. This step will handle the file created by the EXPRES machine and transform it into data that can be stored within a database. This step must also be completed before 8 pm the next day.

4.1.2.2. Admin Access

The Admin will have access to an interface that allows them to upload data files. The Admin will simply select a file from their computer to upload for this procedure to complete. This feature will be helpful in case the automatically posted data ran into a problem somewhere.

4.1.3. Graphing Data

In order for our client, his team, and the public to view the data collected on the sun, the data must be plotted on graphs. These graphs will consist of spectrum measurements and radial velocities.

4.1.3.1. Graphing Radial Velocity

The radial velocity data will be queried from a database, calculated using a set of functions supplied by our client. The graph will be plotted with the x-axis being the Modified Julian Date (MJD) of the reading, and the y-axis being the radial velocity point associated with the MJD.

4.1.3.1.1. Default

The default of the radial velocity graph is all of the data collected.

4.1.3.2. Graphing Spectrum

The spectrum graph can be of two types. The first being a 1D Spectrum and the other being a 2D Spectrum. The graph will be plotted with the x-axis being the Wavelength and the y-axis will be the Flux associated with the Wavelength. There will also be the option to change the y-axis to be Signal to Noise.

4.1.3.2.1. Spectrum 1D

4.1.3.2.1.1. Default

The default for the 1D spectrum will be plotting all points where the x-axis is from 480 nanometers to 500 nanometers. The y-axis will be plotting flux.

4.1.3.2.2. Spectrum 2D

4.1.3.2.2.1. Selection of Orders

A 2D spectrum graph has orders, therefore there will be a selection of which orders the user wants to view. The user will have the option to select individual orders or to select all.

4.1.3.2.2.2. Default

The default for the 2D spectrum will be plotting all points of one order. This order will not be hard coded, giving the client the option to change which order is the default. To start, this order will be set to the 50th order.

4.1.3.3. Graph Refactoring

All of the listed graphs will have the ability to be refactored. The first is being able to zoom in and out on a section of the graph. Another is manually setting the range of x and y-axis. The last is a reset button.

4.1.3.3.1. Zooming In/Out

Zooming in will come in two forms. The first in will be a button which will zoom in on the center of the graph. The second will be another button, which will allow a click and drag feature. By clicking and dragging the user will create a box. This will set the boundaries of the box to the new x and y-axis ranges.

Zooming out will also be a button to zoom out on the center of the graph.

4.1.3.3.2. Change of Axis

There will be a table which will have the option of adjusting the minimum and maximum of both the axes. This table will include four boxes and a submission button. On one side will be the labels indicating the axis, and on the other will be labels for minimum and maximum. By typing into the corresponding box, the user can adjust the view of the graph. In order to update, the user will need to hit the submission button.

4.1.3.3.3. Reset

There will be a reset button which will set the graphs back to their default settings.

4.1.3.4. Hover Over Each Point

When a user hovers their mouse pointer over a point or line of the graph, a text box will appear giving associated data on that point.

4.1.3.4.1. Radial Velocity

For the radial velocity graph the associated point will give the radial velocity number, and the option to see either version of the Spectrum graph of the associated point.

4.1.3.4.2. Spectrum

For the spectrum graph, the user will be hovering over the graph line. This will give the approximation of where the pointer versus the x-axis. The data which will be given is the wavelength, and either the flux or signal to noise value.

4.1.4. Downloading Data

The client wants to give the general public and other researchers the ability to download information on the graphs they are viewing. The data they can download is outlined below, with Radial Velocity, 1D Spectrums, and 2D Spectrums.

4.1.4.1. Radial Velocity

4.1.4.1.1. American Standard Code for Information Interchange (ASCII) File

The Radial Velocity graph data will be stored as an ASCII file. This file will contain the information listed below that is available to the given permissions level. This file will also contain all of the information, regardless if the user has refactored the graph.

4.1.4.1.2. Associated Data

The data included in the ASCII file will be the Date in MJD, the Radial Velocity, and the Error value of the associated Radial Velocity. This data will be listed in three columns with titles listing the type of data in each section.

4.1.4.2. 1D Spectrum

4.1.4.2.1. ASCII File

The 1D Spectrum graph will be stored as an ASCII file. Since permissions level will not let an unauthorized user see this graph if not permitted there is no need to restrict the download. This file will still download all data regardless of if the graph is refactored.

4.1.4.2.2. Associated Data

The data that will be included in this file will be the Flux, the associated Wavelength, and the Signal to Noise relationship. This data will be listed in three columns with titles listing the types of data in each section.

4.1.4.3. 2D Spectrum

4.1.4.3.1. Fits File

For the 2D Spectrum graph, our client will provide a Fits file which will be stored in the database alongside the other data. When downloading this Fits file will be downloaded. Since permissions level will not let an unauthorized user see this graph if not permitted there is no need to restrict the download. All of the data will be downloaded regardless of if the graph is refactored.

4.1.5. Page Organization

This website will serve more than the singular function of serving data to users. To accommodate the required functionality, the website will be organized into 7 pages each with specific functions. Some of the pages will require authentication to access.

4.1.5.1. Public Pages

Pages accessible to the public will not require a Team member account to view. All information available to the public will be adjustable from the administrator dashboard.

4.1.5.1.1. Landing Page

The landing page will be the starting point for any new visitors. It will contain a link to the main graphing section of the website, as well as access to the news page. On all three main public pages, the option for team members to sign in will be present.

4.1.5.1.2. Graphing Section

This will serve as the main content of the site for the general public, allowing the viewing and download of selected data from Lowell Observatory. It is required that the Radial Velocity chart will be initially loaded on the page so the selection of a spectrum is possible. This may be changed in the future.

4.1.5.1.3. Education Section

Serving as a secondary function of the website, Lowell will manage a news and education section for updates and interesting finds. Functionally, we will design a skeleton page to later be filled in by Lowell staff, organizing news events in a list ordered from newest to oldest. This section may be integrated into the landing page in the future.

4.1.5.1.4. Login Page

With certain functions of the website reserved for researchers and administrators, signing in is required by account holders to access some features. The Login page will have inputs for an email and password, as well as buttons to confirm login and to recover a password.

4.1.5.1.5. Recover Password

The page to recover a registered account will have an input for the recovery email and a button to submit. The server will send an email to the submitted account.

4.1.5.2. Login Required Pages

As the main purpose of this website is to share information among researchers, they should be able to access a larger set of data. By registering with Lowell, researchers at universities will be able to do just that.

4.1.5.2.1. Authorized Users Graphing Section

Researchers will have access to the same interface as a general public, but with additional exposures that the LOST team can set in the Administrator dashboard. Unreleased exposures will be denoted by a stylistic difference to highlight them.

4.1.5.2.2. Administrator Dashboard

The Administrator dashboard will have a list of each exposure organized by date and time. Selecting an exposure will display its graph, and each item in the list will have an option to set its access to public or private.

4.1.6. Stylistics

Aspects of style will be used to differentiate functions and options, while elevating the appearance of the website.

4.1.6.1. Aesthetic

Apart from design elements used to draw the user's attention, aesthetics serve to convey a sense of tone and make the website look professional. The theme will make use of a light palette, accented by blocks of color; components and sections of the site will use light borders as separation. Corners will be rounded to give a sense of nesting, and drop shadows will be used to bring elements forward.

4.1.6.2. Colors

Lowell Observatory has provided us with a set of colors that appear on their logo. They will be used to give context to the user interface.

4.1.6.2.1. LOST Red: Pantone 485

#EE3124 / R:194 G:042 B:034

4.1.6.2.2. LOST Orange: Pantone 144

#F8971D / R:248 G:152 B:029

4.1.6.2.3. LOST Yellow: Pantone Yellow C

#FFDD00 / R:255 G:233 B:000

4.1.6.2.4. LOST Green: Pantone 361

#3DAE2B / R:120 G:174 B:064

4.1.6.2.5. LOST Light Blue: Pantone Process Cyan

#00AEEF / R:000 G:174 B:239

4.1.6.2.6. LOST Dark Blue: Pantone 287

#002F87 / R:000 G:083 B:155

4.1.6.2.7. LOST Purple: Pantone 2583

#A25EB5 / R:147 G:111 B:177

4.1.6.3. Logos

Lowell Observatory has provided several versions of their logo for our use, and will be integrated into the graphics of the website.

4.1.6.4. Text

The typeface of the website has been given as a media requirement by Lowell Observatory's media team. The fallbacks for the font are from recommended

4.1.6.4.1. Montserrat Regular

The typeface Montserrat has been given as a media requirement, and will be implemented as the main text of the website

4.1.6.4.2. Deja Vu

DejaVu Sans is recommended as the primary fallback font for Montserrat.

4.1.6.4.3. Verdana

Verdana is recommended as the secondary fallback font for Montserrat.

4.1.6.4.4. Sans-serif

Sans-serif Sans is the default fallback for Montserrat

4.2. Performance

Performance requirements are the stipulations stating how some of the listed functional requirements will perform. For this project, the performance requirements will revolve around the loading of displays, and the uploading of data.

4.2.1. Loading Graphs

Users should not have to wait for long periods of time to view the graphs, but with the massive amount of data the server must send, some wait time is expected. Each of the graphs will be measured regarding the time it takes to load the default settings and the maximum amount of data to plot.

4.2.1.1. Radial Velocity Graphs

Since the default setting for the radial velocity graph is to display the complete set of data, this section will only look at the default data.

4.2.1.1.1. Default Data

Loading the default amount of data should take no more than 10 seconds from when the search is queried to displaying the data to the user.

4.2.1.2. Spectrum Graphs

The two kinds of spectrum graphs could take different amounts of time to load, due to the amount of data in each type.

4.2.1.2.1. 1D Spectrum

The 1 dimensional spectrum graph is an overall view of the spectrograph's results, made from summing each individual order, reducing file size.

4.2.1.2.1.1. Default Data

Loading the default amount of data should take no more than 5 seconds from when the search is queried to displaying the data to the user.

4.2.1.2.1.2. Maximum Data

Loading the maximum amount of data should take no more than 10 seconds from when the user requested the refactored graph to display the data.

4.2.1.2.2. 2D Spectrum

The 2-dimensional spectrum is a combined set of 86 separate orders, each made of thousands of points. This may have an impact on the load time.

4.2.1.2.2.1. Default Data

Loading the default amount of data should take no more than 5 seconds from when the search is queried to displaying the data to the user.

4.2.1.2.2.2. Maximum Data

Loading the maximum amount of data should take no more than 20 seconds from when the user requested the refactored graph to displaying the data.

4.2.2. Uploading Data

When our client submits the files to the website for upload; it will take no more than 24 hours to finish. This includes parsing the files, creation of database tables, and population of the tables.

4.3. Environmental

Environmental requirements are any technological restrictions imposed by the client. There are two environmental requirements for this project: Python and Docker. Python is required in order to access and retrieve data from the EXPRES machine. Docker is required in order for Lowell to host the webpage.

4.3.1. Python

Python is an environmental requirement for our website, because our client is currently using Python code to gather data from LOST. This incentivizes the use of Python within the backend of the website to ensure integrity of data. It also will create some internal consistency on Lowell's end and make it easier for Lowell to maintain the website after completion.

4.3.2. Docker

In order to integrate the website into Lowell Observatory's current system, the website will need to be held within a docker container. Containers are small execution environments that run separately from one another. A docker environment will allow for the website to be more easily maintained because of its standalone structure. A docker will also be easier for the client to host because dockers are portable, which makes them easier to integrate into existing systems.

5. Potential Risks

Every software project naturally has potential risks attached to it. Being prepared for a possible risk is the biggest factor in resolving it. There are a few potential risks that have been considered for the project given to us from The client. The data displayed will be based on very accurate and precise data, and may be used by our client and his team to analyze the data. In the table below, we have listed the potential risks along with the likelihood, the potential harm or severity of the risk, and the proposed solution/ mitigation of the problem.

RISK	LIKELIHOOD	SEVERITY	MITIGATION
Inaccurately displayed data	**Low**	**Moderate**: Misrepresented data could result in wasting the client's time and making the website inaccurate	Plot data exactly as given to avoid rounding numbers and losing precision
Inaccurately stored data	**Low**	**Low:** If data is misrepresented it could cause the website to lose credibility	Store data exactly as it is extracted from Lowell Telescope
Data loss due to hacking	**Low**	**High:** Losing data could cause a gap in the graphing process	After speaking to the client about this risk, we have decided not to implement any mitigation strategy
Authentication handled incorrectly	**Moderate**	**High**: Handling authentication incorrectly could allow for private data to be public, fake data could be uploaded	Client will have the ability to give privileges, you cannot become a team member without the client's permission
Website being hacked	**Moderate**	**High:** If the website is hacked and someone else gains admin privileges, they may compromise the website by deleting data or adding inaccurate data	After speaking to the client about this risk, we have decided not to implement any mitigation strategy
Data Mismanagement / duplicate entries	**Moderate**	**Moderate**: Duplicate entries or managing the data incorrectly could cause extra unnecessary memory expenses	Make sure the database is checked for duplicate entries before inserting data into the database

Figure 5.0: Outline of risks associated to the project

5.1. Inaccurately Displayed Data

There is a risk present within this project that the data received by the LOST is improperly displayed and managed. This could be a result of the plotting package being implemented incorrectly.

5.1.1. Likelihood: The likelihood of this risk becoming a real-life problem is quite low, this is because plotting and graphing the data is one of the most important aspects of the project, so accurate implementation will be held as one of the most important requirements.

5.1.2. Severity: The severity of this risk is moderate, this is because the website is rendered useless if our client isn't receiving accurate information. This could result in the website losing credibility.

5.1.3. Mitigation: To mitigate this risk, it is important to communicate with our client about how precise numbers need to be in specific cases, this way it is easy to prevent losing important data through rounding numbers. Another key to the mitigation strategy is to run unit tests on the different graphs. We will verify our graphing program works by testing it on a small sample of data. Once we confirm with our client that the data is represented correctly and as accurately as possible, we can implement the graphs on the site with all the data.

5.2. Inaccurately Stored Data

The risk being analyzed here is in the fact that data bad organization of the DBMS, or even an error in the plotting algorithm.

5.2.1. Likelihood: The likelihood of this risk becoming a problem is low, this is because our team will be storing data just as it is given to us from Lowell Observatory.

5.2.2. Severity: The severity of this risk is low. This is because the team won't be able to properly graph data if it isn't properly stored.

5.2.3. Mitigation: To mitigate this risk, the team will be putting data into the database exactly as it is recorded by the telescope, this will prevent losing accuracy and precision from rounding floating point numbers.

5.3. Mishandling Authentication

Authentication is a large part of this project. This is because the Admin (currently only Dr. Joe Llama) has all privileges, meaning that an Admin is able to delete, add, and set data to public or private.

5.3.1. Likelihood: The likelihood of this risk happening is low. Since the number of Admin accounts will be incredibly minimal, currently only one, it will be difficult for a member of the public to gain access to the account. In addition, there is little motivation for the public to try to access restricted accounts. Still, any application with accounts comes with some risk for mishandling authentication.

5.3.2. Severity: The severity of this risk is high. Even though data is uploaded to the site on the backend, a user with Admin capabilities could still change data or display false data. In the case of mishandling authentication, the worst-case scenario would be private data being released to the public and public data being set to private. This would be frustrating for our client, who would have to fix the private and public data.

5.3.3. Mitigation: To mitigate this risk, Dr. Joe Llama will be the only admin, decreasing the chance of the admin account being hijacked. He will also be the only one accepting new account creations, so each created account will be reviewed before being created. We are also using the Vue.js framework for our front-end. We chose this technology partly because of the added security it has over raw JavaScript.

5.4. Website Being Hacked

There is a fair chance that the website created could be hacked. This unfortunate scenario could be the cause of a malicious attacker and it has a higher probability due to Lowell Observatory being a world renowned institution.

5.4.1. Likelihood: The likelihood of this risk becoming a real issue is moderate due to the fact that they are a well known institution.

5.4.2. Severity: The severity of this risk is quite high, this is because if the website does have an attacker hack into the system, the attacker could potentially harm Lowell servers, or even the users of the website.

5.4.3. Mitigation: After discussing this issue with the client, we have come to the conclusion that no mitigation strategy will be implemented for this risk. In order to properly mitigate this risk, keeping the website software managed and up to date would be crucial to protect from an attack.

5.5. Data Loss Due To Hacking

The risk being analyzed here is data being lost due to hacking. This is in the case that an attacker hacks into the website and attempts to delete or remove information from the database.

5.5.1. Likelihood: The likelihood of this risk is moderate. This is because the project will branch off of the Lowell Observatory website. So the security of the website is dependent on their servers.

5.5.2. Severity: The severity of this risk is high. This is because if data is removed from the database, there will be gaps in the graphs displayed.

5.5.3. **Mitigation:** After discussing this issue with the client, we have come to the conclusion that no mitigation strategy will be implemented for this risk. In order to properly mitigate this risk, keeping the website software managed and up to date would be crucial to protecting from an attack.

5.6. Data Mismanagement and Duplicate Entries

The performance and outcome of this website will be partially dependent on the management of large amounts of data. If data is accidentally entered twice or handled inefficiently, an unnecessary strain is put on the server. If the system is unable to process all of the data that is given, the website will not function properly.

5.6.1. **Likelihood:** The likelihood of this risk happening is moderate. Even though Dr. Joe Llama reduces the data as soon as it is captured, there are still large amounts of data to sort through. Data could also be unintentionally entered twice, doubling how much needs to be stored and organized.

5.6.2. **Severity:** The severity of this risk is moderate, this is because this risk could cause the website to behave incorrectly, perform at slower speeds, or crash altogether from handling too much memory.

5.6.3. **Mitigation:** To mitigate this risk, there will be a check to make sure duplicate data isn't being inserted into the database. This will not prevent the risk of data being inserted twice, but will also keep data output accurate. The mitigation strategy for this risk is to give a unique ID to each input, and have the database do a check for an input's ID before entry into the database.

6. Project Plan

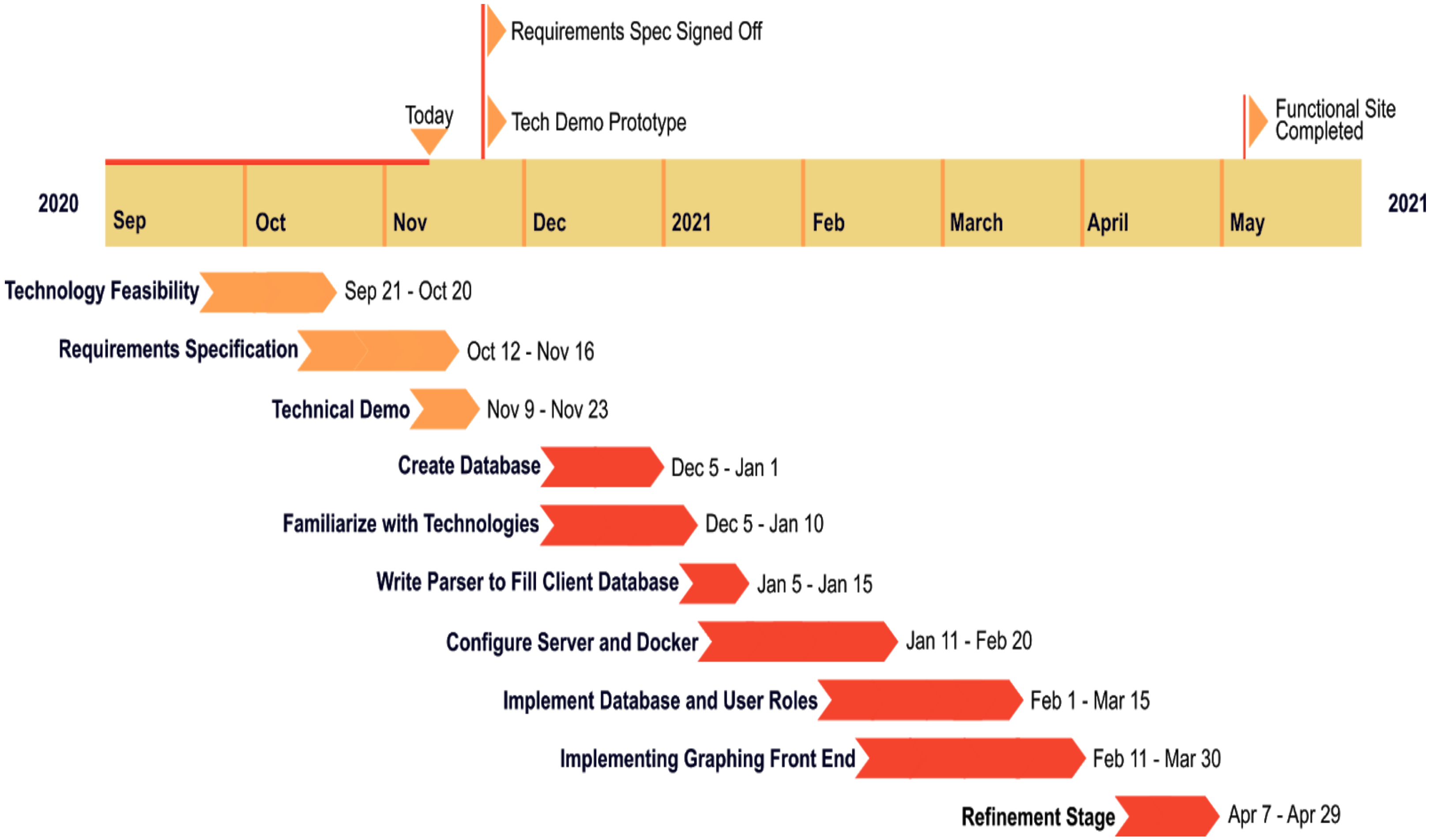

Figure 6.0: Development schedule plan

The project has been sectioned into milestones to show current progress as well as future tasks to complete. The plan is to take care of the back-end first, to confirm that all technologies will work as expected. The database will be created first, along with familiarizing the team with SQLAlchemy. Next, a parser will be written in Python to help collect data from the LOST and import it into the database. After this we will write the graphing code using Plotly. At this point the back end will be completed, which is when the front end can be refined. Refining the front-end will include testing the user interface and improving the speed of the application. This will lead to the completed product in May.

6.1. Technical Demo

November 2020

Work on the Technical Demo will run through the end of the semester. We plan to have it finished before the official due date to avoid running short on time. The Technical Demo will include graphs of sample data, in order to simulate a scaled-down, working version of the actual product. We will have a radial velocity graph rendered with Plotly and implementing all the interactive features that the client is looking for.

6.2. Create Database

January 2021

Creating the database will be done over winter break, and will run on for roughly a month. This will be a MySQL database. While creating the tables will not take much time, importing the current data into the tables will be a lengthy process simply due to the size of data being given to us.

6.3. Familiarize Team With Technology

January 2021

Also over winter break, the team will be tasked with familiarizing and learning the technologies used in the project. The technologies include Vue, SQLAlchemy, and Plotly.

6.4. Write Data Parser

February 2021

We need to write a program to parse the client’s data and import it into the database. This will be done using Python, MySQL, and SQLAlchemy This task is scheduled from January 5 until January 15. Since this is a simple parser, we suspect the majority of the ten days will be spent getting all the technologies to communicate and work properly together. Writing the code to import the data into the database will not take as long.

6.5. Configure Server and Docker

March 2021

Configuring the server and Docker environment will run from mid January through late February giving a little over a month total to complete. This will include getting our web application up and running on a server. It will also include getting the Docker environment that has been requested by the client setup.

6.6. Implement Database and User Roles

March 2021

Implementing the database and user roles goes hand-in-hand. This is due to the fact that user accounts and permission will be stored in the database, so we are unable to implement user roles without first implementing the database. Implementing the data will include connecting our database to the application and verifying the application can read from the database. It is not necessary for the application to write to the database.

6.7. Write Plotly Code

March 2021

After all other technologies are working and proven to be stable, we can finally graph the data on the app using Plotly. Since graphing with Plotly is included in our technology demo, we are optimistic that this process will be easier to implement in the spring.

6.8. Refinement Stage

April 2021

This is the final milestone of the project. Here we will be testing the application for bugs and vulnerabilities, checking usability, and finishing up on any other milestones that fell behind. We will also refine any user interface features and ensure the client is happy with the aesthetic of the site.

7. Conclusion

Since discovering that first interstellar planet, astronomers and researchers have searched the sky to find evidence of planets among the stars we look up to every night. They had believed it was an exciting new beginning for the field of astronomy; however, our window of opportunity is slowly closing. Every year, more and more satellites collect above our atmosphere, interfering with extraterrestrial observations. New constellations of satellites are reported to have a significant impact on the precision of astronomical measurements, with several mega-constellations being planned for the relatively near future. With these events in the distance, the LOST team will aid in the global collection of data; hopefully aiding in the development of more precise technologies and measurements while they can.

The web application will allow interested researchers and the general public to access and visualize extremely precise data, free of charge. Lowell Observatory has the means and will to distribute information that would otherwise be proprietary; letting anyone in the world use data they have recorded in order to further progress on spectrographic analysis. This will help Dr. Llama with his future goal of finding an Earth-like planet around a Sun-like star. The main features of the website are as follows:

- Storing Data will be the base function of the website. While it may not be the main purpose for its existence, being able to store the data taken from the telescope and spectrograph will be a critical part of running the entire website. The spectrograph takes numerous exposures, roughly 5 minutes long each, totalling to over 40 GB a day. To support this, we will set up a database that can automatically process and record a day’s data overnight.
- Graphing Capabilities will allow raw data to be visualized before the available data needs to be downloaded. This feature will be the main function of the website besides storing the data. There will be a spectrum graph for each individual exposure, aggregated by a scatter plot of each exposure's radial velocity. Each graph will have features to define a selected range for each axis of

each graph, and the spectrum graph will have the option to reveal individual orders versus the combined plot.

- Levels of Access will restrict the public from accessing certain exposures, but will allow registered researchers to access that same data for a larger pool of exposures. Administrators will be able to set what exposures are public, as well as verifying being able to register researchers and uploading data manually.

The next step we will take is to construct a semi-functional mock-up that will satisfy main domain level requirements. The reason for this is to practically test how each of our previously selected technologies will integrate with each other; demonstrating that they are a good fit for the website's given requirements. Based on our technology feasibility research we are confident that our website will be highly convenient for quickly viewing several exposures over a settable period, and maybe it will help develop sharing large sets of data between researchers. Overall, our team is very optimistic that the chosen technologies and developed requirements will help create the desired product for our client and the researchers they work with.