# Finding Hay in the Needle Stack: Doing the Right Thing at the Right Time

PROJECT PLAN

**Team Number: sddec19-09**
**Client: Collins Aerospace**
**Adviser: Srikanta Tirthapura**
**Team Members/Roles**
Austin Konsor: Software Architect/Chief Engineer
Mireille Mwiza Iradukunda: Test Engineer/Report Manager
Gooi Yin Hong: Lead Designer/Web Master
Yealim Sung: Project Leader/Meeting Scribe
Team Email: sddec19-09@iastate.edu
Team Website: sddec19-09.sd.ece.iastate.edu

Revised: 2/21/2019/Version1

# Table of Contents

# List of Figures



# List of Tables

# 1 Introductory Material

## 1.1 Acknowledgement

We are thankful and excited to be assigned this project with Collins Aerospace. Specifically we would like to thank Andy Zobro and Chauncey Becker for their continued help, resources and guidance. We hope to fulfill their needs and deliver a final product to help with their lasting success.

We would also like to thank Srikanta Tirthapura. His support and guidance will be very valuable in the coming months as we prepare to deliver an outstanding product to Collins Aerospace. Finally, we would like to thank all teaching assistants in this process, as much of it is volunteered time, it is greatly appreciated.

## 1.2 Problem Statement (2 paragraphs+)

Collins Aerospace has about 80 directories of test files that they use for display in airplanes. Within those directories, there are many copies of the same files, some with small variations or different functions that produce different results. At the moment, they have no way of knowing which tests are the best option to use, or given a file or function, is there a better program to run with the same conditions.

Collins Aerospace would like an application that, given a signature from the user, will mine through the different test files and make suggestions of a better file to be running. There are different stages in this process however. First we need to build a wrapper that will classify each of these tests and functions in order to work with later. Then, given that data, will search through and find the best option to give the user. If possible, we would like to automate this, so the program won't require a user input, instead the application will just give suggestions on what test they should be using instead. With the final product, we hope to not only make their work more efficient, but increase their work by producing the best results possible.

## 1.3 Operating Environment (one paragraph +)

The final product will be a desktop application (written in Python) that will strictly be run only on Collins Aerospace systems, or computers. This is to ensure that no data is seen outside of Collins Aerospace. Their systems just use Windows OS, so there isn't any difficulties with the desktop application.

## 1.4 Intended Users and Intended Uses (two paragraph +)

The only users of this application will be the developers at Collins Aerospace.

### 1.5 Assumptions and Limitations

Assumptions

- Mock Test Data will be sufficient to build application - Collins Aerospace is unable to give us the true data from their SQL servers to keep it private, so they will give us a mock data set to work with which will imitate a real set.
- This application will only run on a Collins Aerospace system, even for development and testing.

Limitations

- As stated above, we are only able to get a sample data set to work with, however it should be sufficient enough to build the application and work with the true data.
- We must work on the development at the ISU Research Park, there they will provide a Collins Aerospace laptop to work off of. We may code on our own devices, but all data will be stored on that laptop.
- We may need a Collins Aerospace employee present in order to gain access to the laptop, so we will need to work around all of our schedules to do so.
- There are no costs associated with this device, unless Collins Aerospace decides to provide a private repository to work off of.
- Only one of our team members have previous experience with big data and Python.

### 1.6 Expected End Product and Other Deliverables

The delivered product will be a desktop application that the developers at Collins Aerospace will run on their devices. The application will mine given data and return a suggested test file, which will save a lot of time and increase productivity. We will also need to deliver all related documents such as design, code and architecture documents in order for them to build off of this application in the future if they wish.

# 2 Proposed Approach and Statement of Work

### 2.1 Objective of the Task

The goal of this project is to deliver a software product that visualizes data and provides insights to Collins Aerospace engineers to help them in developing and testing the software. This tool should be able to assimilate and visualize data and develop signatures for tests. The resultant end product would be an application that will take several files as an input and possibly output a report and a visual.

### 2.2 Functional Requirements

**Analyze input files** Given the file and key, the program should find what the key represents. It should also identify how similar one program is to others and how many of them are using the same version.

**Visualize end result** After comparing the input file with other files, it should be able to output the result in a tabular form to help the user to clearly see the result.

**Generate final report** The summary of analyzation and visualization should be included in this final report that will be generated when the program ends.

**Quick, accurate, and reliable output** The result should help the engineers to quickly and easily determine what to do and provide deeper understanding.

**Offer suggestions** Given a file the program should see the trends and offer suggestions that will improve the result.

### 2.3 Constraints Considerations

The main constraint would be the security of the program itself and the input datas. As our client is in a corporate it is not easy to freely access their datas and resources. Team members are required to physically be in the facility in order to test the program and the input datas will be fabricated.

The team members are not familiar with Python language. Therefore, the members need to figure out how they are going to learn the language and who to find when there is a problem. The standard protocol would be to discuss internally if there is any problem, then to discuss it with the faculty advisor, and finally discuss it with the client.

### 2.4 Previous Work And Literature

Since Big Data is a rapidly growing field, there are multiple projects that has similar concept and outcome. We will add the reference and cite the sources in the future if we are going to follow the previous work.

### 2.5 Proposed Design

The table below is the possible design for the visualization of the output. C represents that the program crashed, P represents that it passed, and F represents that it failed. Regarding the information given, the user can tell which File is similar to which. The report format will be discussed in the future.

| File 1 | File 2 | File 3 | File 4 |
|---|---|---|---|
| C | C | P | F |
| C | P | P | F |

| | | | |
|---|---|---|---|
| P | P | F | F |

Table 1: Proposed design of the visualization output.

### 2.6 Technology Considerations

There could be two possible ways to design our project. The first option would be to make a desktop application and the second option would be to make a web application. If we make a desktop application, it would be easy to deliver it as a zip file where the client would install the software themselves. On the other hand, web application will need to interface via HTTP and TCP with the backend application. We are still discussing about our options.

### 2.7 Safety Considerations

Due to the nature of the project, there will not be any physical harm or risk in the process of developing our project or using our product.

### 2.8 Task Approach

Figure 1 as shown below is our current approach to our task at hand based on the design thinking model. In the Requirement Analysis phase, we will try to understand and identify the client's problem. We will list important requirements and specifications for the project. We will try gain all the requirements before developing the applications. Most of the work are done in documentation and meeting with the client in this phase. The development phase would begin in the next semester.

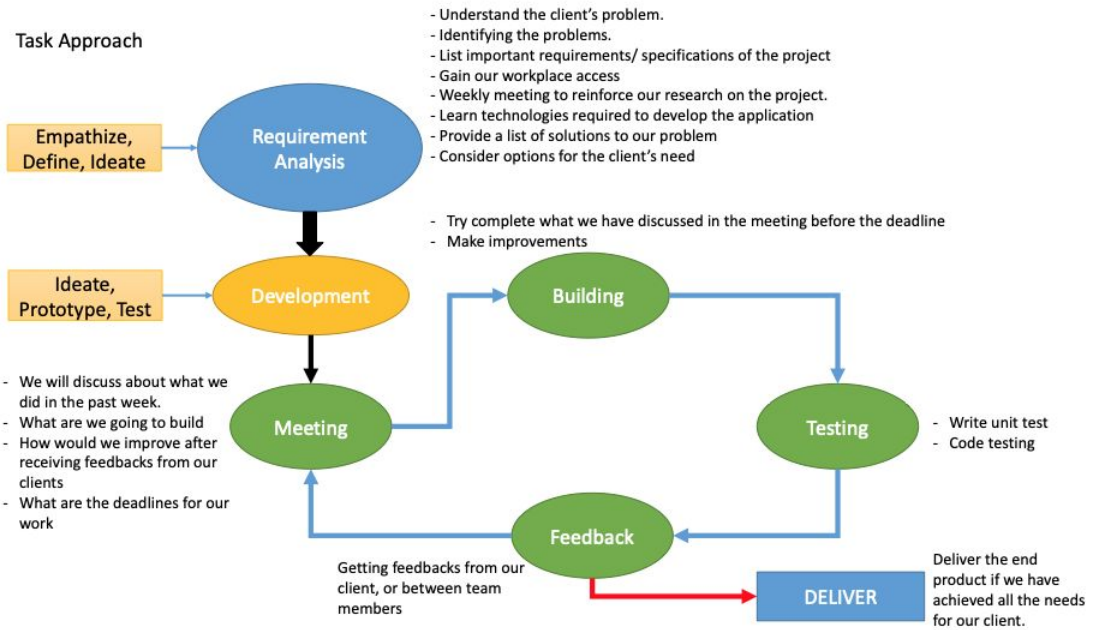

Figure 1: Project Workflow

### 2.9 Possible Risks And Risk Management

**Cost:** After discussing with our clients, there are currently zero cost for the project as they will provide us equipment, laptops for us to develop the application.

**Materials:** No materials are needed for our project as most of the part can be done using computer.

**Equipment:** Computers are provided by our clients. The possible risk is that we might not be able to develop the application remotely which might also hinder our progress due to their company's policy and security measure.

**Knowledge of area:**
There is a risk that we might be developing the system's architecture in the wrong direction, but to mitigate this problem, we have been asking a lot of questions regarding our goals and requirement for this project.

We might also lack of python experience, because most of our team members only have a basic knowledge about using python, we have been starting to learn and revise the technologies that we are going to use. Risk of our end product algorithm are not as efficient and could consume time to process because we are probably going to deal with a huge amount of data.

**Accuracy Issue:** Since we are dealing with data and also comparing signatures between them, we must be precise with our results.

### 2.10 Project Proposed Milestones and Evaluation Criteria

Key Milestones: classifying data, collect relationship between the test, develop an application for data visualization, use those test's relationships to visualize the data.

Testing: We will be testing each of the milestones.

- First we will test if the classified data are being classified correctly by using the fabricated data set that will be provided by our client.
- We would also try to build known relationships as a start to test our program.
- Develop an application that will be able to retrieve data and show the data in an easy way to understand, and we will test if the visualizations are being output correctly.

### 2.11 Project Tracking Procedures

To keep track of our progress throughout the course. In this semester we will be using Trello, writing the "Weekly Status Report", and conduct internal weekly meeting, and meeting with our clients.

In the next semester, we will continue using Trello, writing "Weekly Status Report" and probably using a private repository (company's security policy) for developing the application.

### 2.12 Expected Results and Validation

Expected Results:
Building a wrapper: The wrapper should be able to classify each of the test functions and build a mechanism that is able compute a variety of relationships between each data sets. After collecting those relationships.

Visualization:
- The application should be able to show those data accurately.
- The visualization should be able to suggest the engineers on what test they should be using.
- How the tests have been performed historically.

Validation:
- We will test each of parts separately to ensure each unit works as expected.
- We will perform integration testing, system testing, and acceptance testing.

### 2.13 Test Plan

Unit Testing:
We will be using unit testing to test each smallest part of our program. We will be using inputs to test for an expected result.

Integration Testing:
After developing more and more parts in the application, we would combine those parts and test to ensure that they work well with each other.

Acceptance Testing:
Before delivering the end product of the application, we would test if the application meet the requirements of our client.

# 3 Project Timeline, Estimated Resources, and Challenges

### 3.1 Project Timeline

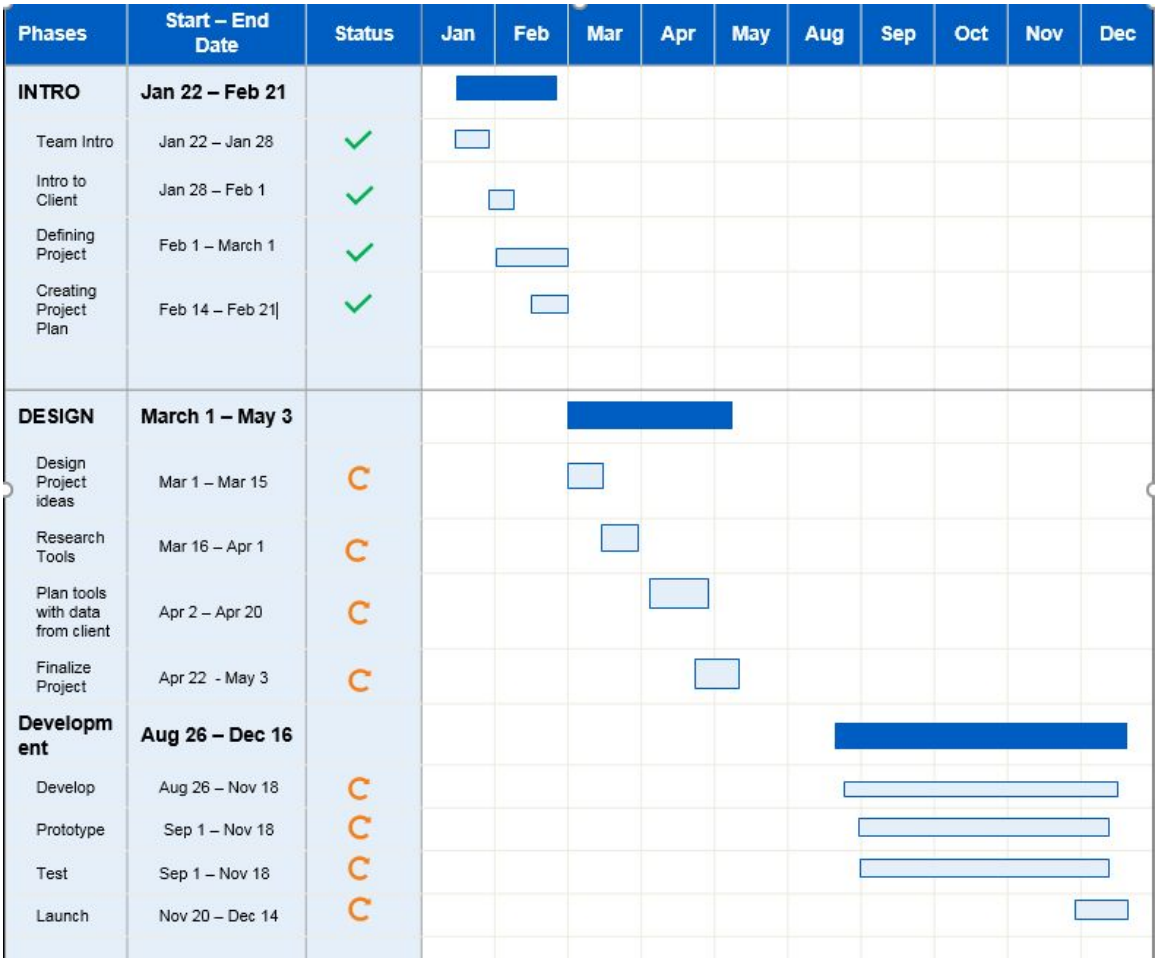


*Figure 2:* Senior Design Schedule over the course of 2 semesters.

As shown above, The project will be divided into 3 phases. 2 phases will be part of semester 1 and the last phase will be in semester 2. We are currently almost at the end of phase 1 (introductory phase). During this phase we worked on getting introduced to the team, the client and the problem statement. Over the course of 1 month we have managed to empathize with the client's problem to better understand their problem statement. The proposed timeline for this phase is 1.5 months.

Phase 2 will start on march 1st, After the empathizing step, we will start working on defining and ideating for the proposed solution for the problem statement. We will be researching in particular how we can achieve these solutions. We will be doing research both on the data to be used and the best tools we will use to work on the project. The proposed timeline for this phase is 2 months.

Phase 3 will be the last phase that will last a whole semester. It is going to be comprised of development, testing and release. We have not written the exact steps that will be

followed but that is because the decisions are going to be taken in phase 2. For now we know that we will have 3 steps and will be a cycle and that is Development, Prototyping and Testing and once the product is good enough for the client and ourselves, we will proceed to release it. The proposed timeline for this phase is 4 months.

We chose to divide these phases like this because although development and testing is very important, it is equally important for us as a team to draft and design our project, get a deep understanding of the problem before tackling it. This will ensure less mistakes that could be costly in the long run.

### 3.2 Feasibility Assessment

The only challenge we are anticipating right now is being able to access Collins Aerospace data. They are still finding a way in which we can access their data to be able to develop the tool. This is part of the overall big challenge that will be a security challenge on their side. From the meetings we have had so far, it looked like it was going to be a challenge getting privileges granted to us. We believe they are working on it and that soon we will be able to get it to facilitate our work.

### 3.3 Personnel Effort Requirements

Each member of the team currently has their roles. However when it comes to research and development, everyone is expected to do their part and cooperate to ensure that everyone is on the same page. All members are expected to give 100% of their effort in a timely manner. As we are in phase 1, the tasks are not clear yet but will be soon and the individual times will be getting reported in the weekly reports.

### 3.4 Other Resource Requirements

Collins Aerospace computers will be used in the process of testing. The extra resource will be a business GitHub.

### 3.5 Financial Requirements

The monthly payments might occur once the team decides to use a business GitHub for privacy purposes

# 4 Closure Materials

## 4.1 Conclusion

Using the right test from a huge data set can be a tiring task to do. It can be a waste of time finding the right test, and sometimes wasting resources and cost for the company.

Our main purpose of this project is to help our client Andy Zobro and his team from Collins Aerospace which has more than 80 directories of test files to develop an application that is able to help their engineers to work more efficiently. The application that we will be developing will be able to give suggestions on what test should be used, shows how the tests have been performed historically, and the relationships between them.

We conduct meetings with our client weekly and we are open to receive any feedbacks to improve the client's need, each of our team members works closely with each other and help each others' back in order to succeed.

## 4.2 References

## 4.3 Appendices

| File 1 | File 2 | File 3 | File 4 |
|---|---|---|---|
| C | C | P | F |
| C | P | P | F |
| P | P | F | F |

Table 1: Proposed design of the visualization output.

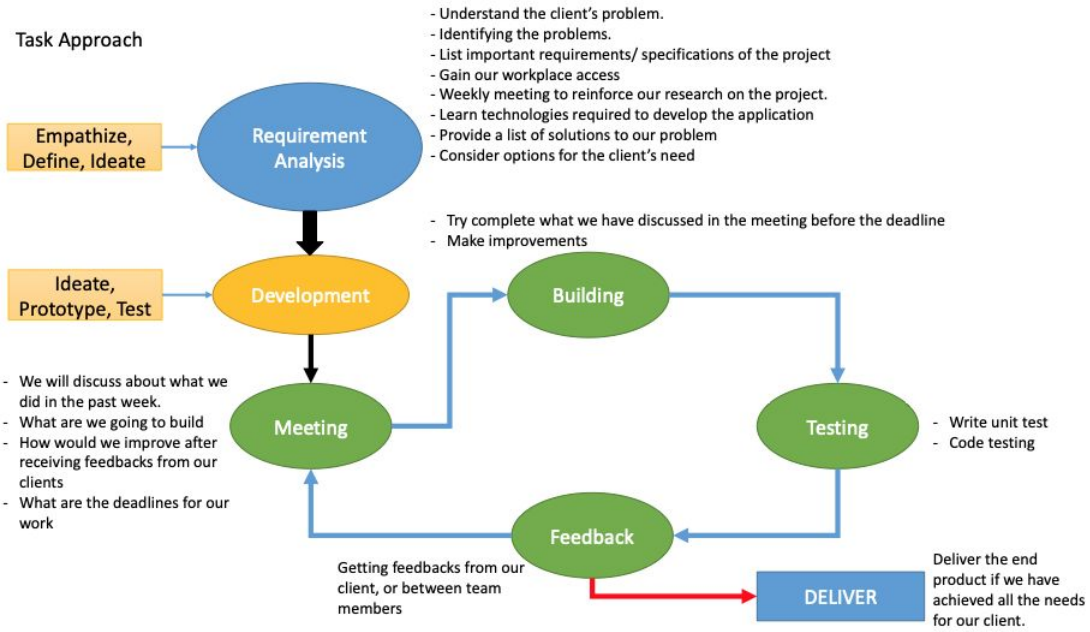


Figure 1: Project Workflow

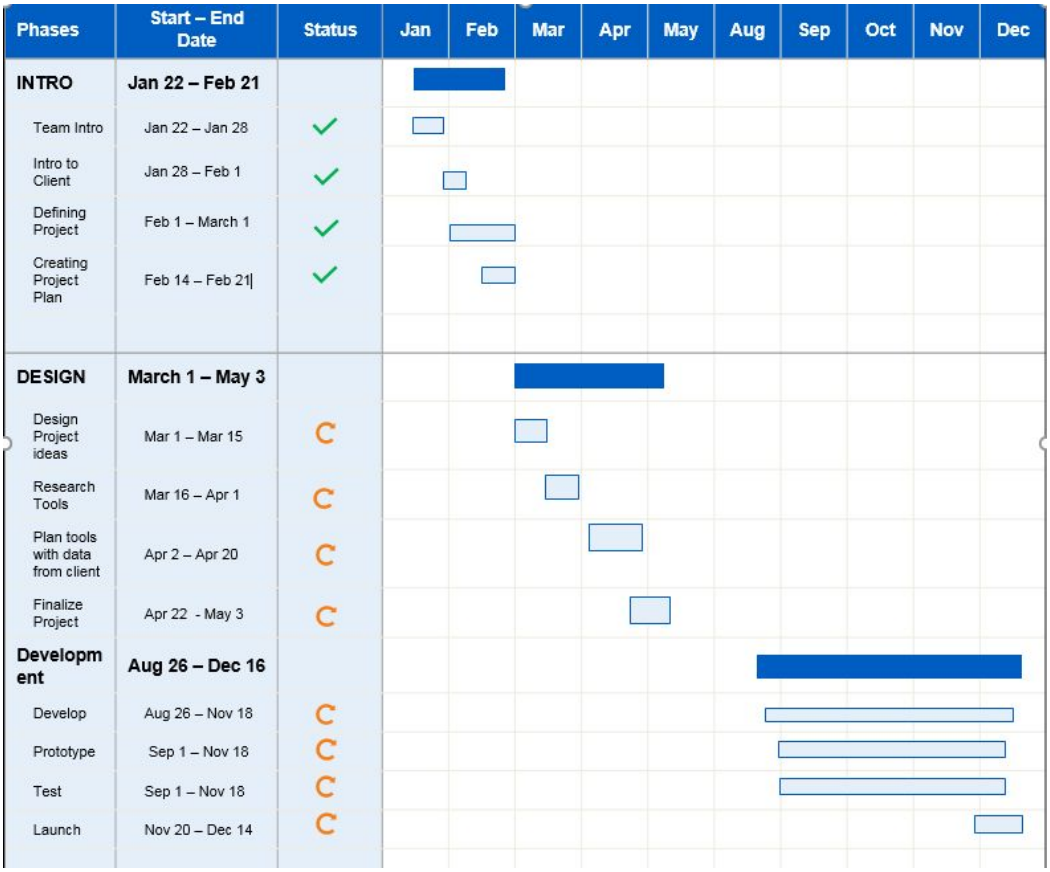


*Figure 2:* Senior Design Schedule over the course of 2 semesters.