Software Proposal Document for visualization project

Lokesh Agnihotri
Supervised by:Eric Slaats

February 9, 2023

Abstract

The main idea of this document is to learn the visualization technologies. For this project I will choose MRR git repositories and apply multiple visualization techniques to it in order to learn various plotting techniques and ways of showing the data. In the process I will be learning the use of various python libraries such as matplotlib, seaborn.

Sioux works on many projects that involve large ecosystems of software components. It can be challenging to get and keep an overview because of the complexity and size of such software stacks, especially when the whole ecosystem evolves over time. A software ecosystem or component is represented in multiple dimensions, like:

1. Architecture;
2. Repositories;
3. Quality metrics;
4. Documentation; pipelines;
5. Backlog of Featured and Issues;

It is hard to oversee the impact in all these dimensions when planning or executing development activities. Sioux aims to improve this by using the power of visualization. Visualizing the software stack in multiple views, and linking these views to each other, can have benefits like:

1. Easier assessment software or scope changes, and identifying areas of interest;
2. More efficient introduction of new team members;
3. More direct communication about and navigation within the software stack.

1 Introduction

1.1 Purpose of the project document

I will use this document to plan my learning outcomes and apply the feedback received during the study. This document will grow based on received feedback applied method to solve the issues encountered during the learning.

2 Benefit of the assignment

1. Easier assessment software or scope changes and identifying areas of interest.
2. More efficient introduction of new team members.
3. More direct communication about and navigation within the software stack.

2.1 Analysis of the assignment

Let's analyse the problem and what the are we trying to fix to get a better understanding of the assignment.

I will be visualizing the dependencies on the GitHub repositories. I will start with visualizing the two repositories to create the concept that can later be scaled to implement more than two repositories to visualize the dependencies structure.

Benefit If we Can visualize which dependencies are common to which modules in the repositories, it will be easier to see which modules are getting affected when a certain dependency is introduced, or if a certain problem occur.

2.2 Approach

I will start with finding a way to collect the push data from the GitHub for the organization. And validate the data manually, if it is correct. Next step will be to structure the data. After that I will start with analysis and visualization of the dependencies between different repositories.

2.3 Motivation

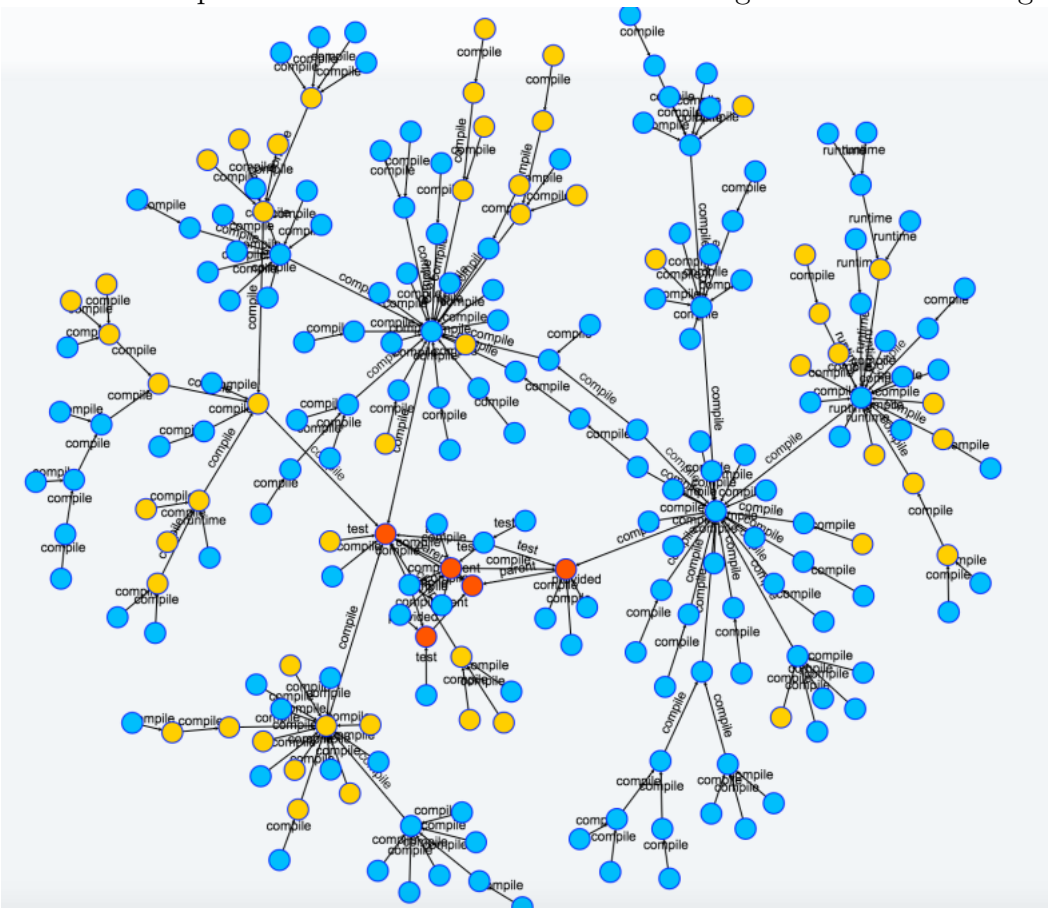
In large organizations where a team of 100 or more is working on a huge software containing 1000s of lines of code. Such visualization can give a lot of insights and key person can be identified easily.

2.4 Problem Definitions

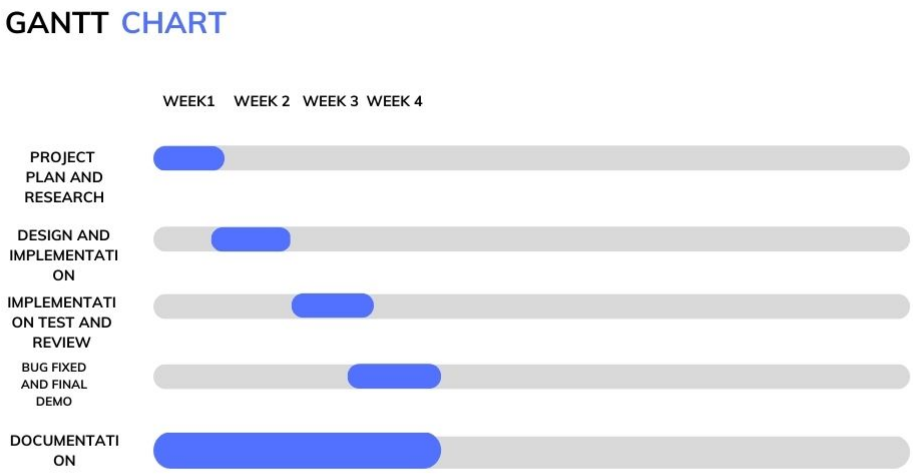
See the code health in the form of the dependencies visualization

2.5 Prototype

The user will start with running the python script from the CLI and enter the URL of the repositories he is interested in. Upon making this, he will see a visualization of the dependencies. Which will look something similar to following image.



3 Time plan



4 Deliverable

I will be delivering the following documents:

1. Requirements document
2. Design Document (shows the design of the application, the components and how they interact)
3. Source Code
4. Test cases.(A document showing the test cases and implementation)
5. Technical document
6. Guide to install and use the software.

.