

University of Pittsburgh School of Medicine

Executive Summary

Community Partner

Annerose Berndt

Student Development Team

Sanika Natu

Skylar Weaver

Akash Khanolkar

Serena Chen

Background

The organization that we are working with is the University of Pittsburgh School of Medicine. It works closely with the University of Pittsburgh Medical Center (UPMC), which is a world-renowned medical facility. In terms of the clients that we are working with specifically, they are Dr. Annerose Berndt and faculty members within the Division of Pulmonary, Allergy, and Critical Care Medicine of the University of Pittsburgh School of Medicine. Even more specifically, we will be working on our project within Dr. Berndt's lab.

Project Description

Project Opportunity

One of the most prominent problems facing our client is that the application that her lab used for its murine genome-wide association studies (GWAS) has been down for around a year. During this time there wasn't an open platform that allowed researchers in the field to share, analyze, and collaborate with their research data. The researchers can continue to complete their research, but they are required to run SNP associations manually and they can no longer share their data with other researchers. An open platform can facilitate not only collaboration but also co-creation among researchers. The reporting and analytical functionality in the previously working application allowed researchers to better assess their results through a standard process and algorithms, which now have to be run manually and individually. Pinpointing and solving the problem of malfunctioning in the current system enables researchers to better access, analyze and share the research results among peers. Finally, a well-documented and deployed application can make it easier for Dr. Berndt and her lab to maintain the application in the long run.

Project Vision

We are proposing to develop and test the current Ruby on Rails system in order to allow Berndt Lab researchers the opportunity to share, collaborate, and analyze their data. The result of this project will create a usable, well-tested and sustainable system so that researchers can easily upload their research data onto the system, perform analyses for research purposes and finally share this data. By providing various components to the system like file upload functionality to easily upload CSV data files, statistical analysis to formulate reports and graphs of the data, and social media features to facilitate collaboration, the project will ensure that researchers can increase efficiency of their

research processes. We are aiming to repair the current system and get the system to a deployable state by May.

Project Outcomes

The goal of this project was to recover the Mouse Data Research Project that lost important data and code during a major server crash. After the crash, the project was inoperable, undocumented, and had poor test coverage. We automated the deployment process by implementing bash scripts to setup the server. From start to full deployment the bash scripts will set up the server environment and deploy the project. After successfully deploying the project, we focused on fixing much of the lost code from the server crash. Finally, a very important part of the project was fixing the statistical analysis pipeline, which includes R, Sidekiq, and other software modules. After fixing these modules we focused more on problems that were useful to the project in the long-term and making it maintainable.

Project Deliverables

We are delivering a functional Ruby on Rails application and experimental data management system that is deployed on a DigitalOcean server. The code is stored on a GitHub repository that includes documentation on setup on localhost and deploying on the server.

Recommendations

Our goal was to make this project as sustainable as possible once we leave from this semester, so as long as the client follows the documentation during times of issues there shouldn't be a problem. It is also important that the client's team takes the time at least on a biweekly basis to check the status of the server. Through Monit, a server management software we implemented, they can run the simple command Monit Status on the server and see if the server is under distress. This is an important recommendation because the main reason that the project crashed is that the server was not properly being maintained when it crashed and the code was not easy to salvage.

Student Development Team

Sanika Natu served as the project manager. She is a fourth-year student majoring in Information Systems with a double major in Business Administration. She will be working at PwC beginning in August 2015 as a technology consultant specializing in IT Infrastructure.

Skylar Weaver served as the Quality Assurance and DevOps lead. He is a third-year student from Central PA and will be working for Deloitte this summer as a business technology consultant.

Serena Chen served as the Python / R lead. She is from China and is a third-year student in Information Systems with a minor in Business Administration. She will be working at KPMG HK as a technology consultant this summer.

Akash Khanolkar served as a devops lead. He is a third-year student majoring in Information Systems. He is from Princeton New Jersey. He will be doing product development this summer in India.