

|  |
| --- |
| Final  Report |
|  |
| November 2  Capstone Project - Phase 2  Authored by: WebScape Team |



|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| IntroductionProject, Context and Goals The proposed project is a project in the health and research sector focusing on the bacterial pathogen Staphylococcus Aureus (S. Aureus) and its antimicrobial resistance (AMR) genes.  The project requires developing a user-friendly, readily usable system that allows scientists and clinicians to find strains of a pathogen and its relationship to others.  The project team is required to work with the Read Lab at Emory University in Atlanta-USA that works on sequencing and computational analysis of infectious pathogens to build upon the existing work from a previous capstone team. The clients of the project being professor Timothy Read in Emory University and professor Jim Hogan in the Queensland University of Technology. They plan to achieve the business goal of offering a user-friendly system to the scientists and clinicians (customers) by implementing the proposed project.  The project scope includes improving a Capstone Project, a website called Staphbook. In order to reach the overall purpose of the project (that is to substantially enhance the reporting capabilities of the portal), it requires to improve the visualization components and enable them to scale, and to support some machine learning based exploration of the dataset from Emory University. The project scope also requires a new feature that supports researching AMR genes of S. Aureus.  The goal for the current state of the project is to introduce new features of finding and predicting antimicrobial resistance (AMR) genes in S.Aureus samples and to improve the existing features provided with the previous Capstone Project. In addition, the system will be developed allowing scientists and clinicians to find strains of a pathogen on a very large database. The project objectives are given below in brief:   * Display and visualize the relationship of AMR genes of every sample of S. Aureus with two visualizations. * Allow user interaction with the visualizations. * Enhance the system with search, filter, and showing metadata functions. * Integrate with the previous project.  Project Outcomes The project aims to deliver two key outcomes:   * It allows users to gain more information regarding S.Aureus.The previous Capstone Project lacked this feature, particularly its AMR genes and AMR sequences that resist specific antibiotics. This feature may help users to research, study or access scientific data of these genomes from Emory University’s dataset. Also, it will provide a ready and user**-**friendly system that will allow scientists and clinicians to display and visualise information of AMR genes for each sample of S. Aureus as well as the relationship of AMR genes for each sample of S. Aureus. * It eases users from accessing the overall information of every sample.This can be done by providing an easy to read data visualization of these genomes. Data visualization also includes interactive capabilities (e.g. hover mouse to see quick information of a sample, mouse click to access more details of the sample).   In this Capstone phase 2 stage, the project team has planned to enhance the capability of the Staphobia portal by creating the following:   * A portal that will display which genomes are resistant to which type of antibiotics * Added mouse hover feature to the web solution to display AMR sample information * Added global bubble chart to indicate samples that are resistant to antibiotics * Provided users with a sample AMR graph for each sample * Provided a node network graph that highlights the AMR connection between each of the samples present within the Staphobia web portal and * Provided essential features (hovering effects, interaction, filtering) for individual graphs.  Project Setup (Phase 2)Project Management Approach The project team has used the Agile-Scrum project management approach to deliver this project. It was suitable for the project mainly due to its nature, size and complexity.  Agile project management is an iterative approach to managing software development projects that focuses on continuous releases and incorporating customer feedback with every iteration. This project is a kind of software development that requires incorporating customer feedback continuously hence the Agile approach suits the scope.  Scrum is an [Agile](https://www.wrike.com/project-management-guide/glossary/#agile) project management methodology, for managing projects that allow for rapid development and testing, especially within a small team led by a Scrum Master, whose main job is to remove all obstacles to getting work done.  The project is managed by a small team involving only five members led by the Project Manager and work is done in short cycles (sprints). The project manager is responsible for the overall project delivery that includes planning day to day tasks with other team members, monitoring the progress of activities undertaken by each member of the project team, organizing meetings with the team members daily to discuss current tasks and any roadblocks that need clearing to ensure the progress and delivery outcomes. The Project Manager also organized meetings with the client and/or client’s representatives and was held fortnightly to discuss the work that has been done, any progress throughout the period and any blockers. In addition, team meetings were held with the client’s representatives if and when required to clarify the scope and if any changes were required to be introduced during the progress of the development of the system.  Project Schedule being updated and revised as per the circumstances and revisions required to the scope. A Change Register was maintained by the Project Manager with necessary assistance from other team members to record all agreed changes during the progress of the development. The project team undertook weekly Risks and Opportunities review meetings to understand any new risks and opportunities to record, and also to update the Risks and opportunities Register and to determine risk mitigation actions to implement minimizing the impact of those risks. Client Expectations The Client’s main expectation is on-time delivery of the deliverables as per the agreed schedule, also the quality of the deliverables.  Each week the project team conducted a meeting with the client’s representative - Jim Hogan from the Queensland University of Technology to provide an update of the progress and to discuss any concerns on the interim deliverables.  In these weekly meetings, the work that has been done was discussed in detail and any further updates that are required to further enhance the project presented. Each meeting, the project team discussed the progress achieved at that time and detailed planning of the project for the outstanding scope. As the main client from Emory university is based in another time zone,  the meetings were only held with Jim Hogan. Jim Hogan acted as Client’s Representative and the Product owner for the project in both phases 1 and 2.  Progress reports were prepared on a weekly basis by the project team and forwarded to Jim. In the weekly meetings, the project team presented Jim the work that the project team has completed including any demos (if available) developed.  Detailed discussions were held on the work that has been completed with the Product Owner Jim Hogan to gain his feedback and any improvements/ enhancements that are required to satisfy the client’s expectations. Clarification on anything the team has done was also given to Jim to ensure that Jim (as Client’ representative)  understood the suggested ideas of the project team and/or whether the proposed ideas are feasible.  After weekly project meetings, Jim Hogan liaised with the client from Emory University in the United States of America, Timothy Reed to ensure that the interim work that  had been completed and presented by the project team in weekly meetings satisfied the Client’s expectations. Based on the discussions with Timothy Reed, Jim provided the project team necessary feedback as to how the project team may further improve the outcomes of the project or what other scope the project team should consider in developing the proposed system to meet the client’s expectations. Team Collaboration During phase 1 the team created a social contract outlining the guidelines on how the team will work together as a team throughout both phases 1 and 2 of the project. Once the details of the social contract were finalised and how the team was going to collaborate during both phases 1 and 2 of the project, a team agreement was established outlining each of the member’s roles and what the team members agreed as a team. The roles of each team member that were established for the project are as below:   * Viet Hoang Do - Full-stack Developer * Quoc Huy Nguyen - Full-stack Developer * Trong Dat Nguyen - Back-end Developer * George Erian - Project Manager   The team had discussed and agreed on the role interfaces and how to share the responsibilities in  delivering intefcate requirements of the project scope.  Daily discussion forums were held in a group chat that was created on Facebook to clarify day to day issues and to solve any critical problems an individual member is experiencing. The forum also used to monitor the progress,  resequencing activities  and/or to provide  answers to any project related questions within the team.  All members participated in weekly meetings held with the client’s representative. Each member was given an opportunity to raise any concerns in this meeting with the client’s representative and  to show the progress of the element; they are individually responsible and gain any feedback on the work that has been produced. Communication Plan The Project communication plan is given with the below table  that shows how the project  team manages its communication with its main stakeholders, in particular the project team and the client.   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | **Communication Plan** | | | | | | | | **ID** | **What** | **Who** | **How** | **When** | **Why** | **Responsible** | | 1 | Project Launch | Client | University Web | Pre-project | Inform relevant Parties | Business Development Unit | | 2 | Daily progress | Project team | Emails | Daily | Project team is responsible for development | All members | | 3 | Scope Change | Project Manager | Change Register | Weekly | To record agreed changes to the original scope | Project Manager | | 4 | Project progress | Project Manager and  Project team | Report via email and zoom meetings | Weekly | Record weekly progress gaunt the planned | Project Manager | | 5 | Progress meetings with Client | Project Manager | Zoom meeting | Weekly | Discuss the progress and issues | Project Manager /Client Rep |   Throughout the project, meetings were held with the client Jim Hogan from the Queensland University of Technology in order to gain feedback on interim submissions or deliverables. These meetings are a vital element of communication as it provides the client's view that  the solutions the team has been implementing satisfies the client’s expectations and/ or receives further feedback on the work delivered  which is critical to enhance the project outcomes. Project Plan and RiskProject planning and progress The project has been divided into two phases over two periods. At the beginning of phase 1 of the project, the project team discussed with the client the project goals and features required to achieve the business outcomes. Accordingly, user stories, release and sprint plans, as well as artifacts, were created as part of the work in phase 1. As the scope of the project was not well defined and agreed upon at the phase 1 stage and also, the work is in progress, all the tasks that were done required refinening over the time. At the end of phase 1, the team was ready to move on to phase 2 to start implementing the features that were produced as part of phase 1.  As this project follows the SCRUM approach, a backlog of user stories was created which was done during phase 1 of the project. The user stories were prioritized during the start of phase 2 of the project with estimates provided. Along with the necessary guidance received from the client throughout the project  duration, the project team ensured that the team was on track, and provided further updates etc. as required.   Development and delivery of the artefacts that were created during Phase 1 were performed during Phase 2 of the project. The development was completed over time during iterations through multiple sprints with meetings held each week with the client Jim Hogan to receive necessary feedback on the artifacts and syatem  development content of the Staphobia Portal.  The development phase began in mid year 2021. The project is expected to start on 23rd July 2021 and planned to finish on 22nd October 2021 (approximately 13 weeks - 5 day a week calendar).  The Project team decided to adopt the SCRUM framework for the development phase 2. During each sprint iteration, the following tasks were planned to be fulfilled  Daily scrum: Each day during the daily meeting each member of the team discusses the tasks that have been done, any blockers with the tasks and what will be worked on next.  Sprint review: Meetings with the clients were held at the end of each sprint. Based on the feedback that was received from the clients any amendments to the product backlog or the product that is to be released were considered.  Sprint retrospective: The team will conduct an assessment of what has been done at the end of each sprint looking at the following:  What went well during the previous sprint  What went wrong during the previous sprint  , and how can be improved to minimise any negative impact experienced due to the issue  Any pending action items for the upcoming sprints.  During phase 1 a plan was developed for phase 2. The plan summary is given below.    Delivery Schedule  (Refer Appendix XX)  Release 1  Delivery date: 20th August  2021  Total Story Points: 13  Description: A  mouse hover feature that shows AMR information of a sample and A global bubble chart  Release 2  Delivery date: 17 September 2021  Total Story Points: 13  Description: Provide users with an AMR graph for each sample. The graph should indicate which contigs in the sample contribute to which AMR. In addition, the team will build a basic node network graph that shows the AMR connections between each of the samples/  Release 3  Delivery date: 22 October  2021  Total Story Points: 12  Description: improvements for the individual graph and node network graph. Features such as interaction, clustering and the theme will be implemented in this stage.  The project schedule activities were updated as per daily and/or weekly progress measure. The percentage of a task completed was measured proportionate to the original task duration. If any task or an activity runs over the original  planned duration, the complete project schedule was reviewed and resequenced to allow necessary additional time to complete the particular task/activity. These revisions resulted in more concurrent tasks to undertake by the project team that required  additional working hours. The project team discussed and agreed to provide necessary support to any member needing additional support in delivering a particular task if it records any time slippage. Risk management The project risk management process including identification of key risks and/or opportunities , preparing risk /opportunities register, analysis of risks and/or  risk assessment criteria and risk mitigation strategies are discussed in this section.  Project risks that may cause harm to the development of the project or barriers to the delivery of the project have been identified by the project team in phase 1 stage and this list was then reviewed and updated in phase 2 considering the scope and delivery plans of the project phase 2.  In a  risk management exercise of any project, identification of the key risks associated with the  project  is essential to prepare an effective risk management strategy. Identification of risks associated with the project and assessing  their potential negative impact are key to ensure that the project team can manage those risks if arisen.  In the assessment, it needs to assess both the likelihood (highly likely, unlikely) and impact of those risks (low, high or medium).  The Project Manager updated the Risks and opportunities register weekly with the Project team in weekly internal meetings. All listed items are reviewed and discussed to see if any changes are required to their status (likelihood and/or change to the impact. Subsequently, the risk and opportunities list is reviewed with the Client Representative in weekly meetings to provide them the status of the existing risks in delivering the project.  Risk Register was updated also to include any new risk or opportunity item, also to remove any risk or opportunity item from the list if it is certain that that risk or opportunity will not be a matter to consider any more.  A copy of the Risk Register used is included as an Appendix to this report.  The project team made special attention to the special circumstances related to COVID‐19. The client was updated regarding restrictions introduced by the Australian state and federal governments under emergency laws and their impact on project delivery specially with limited access to university facilities.  The table below provides a description of the risk, the estimated likelihood values and the impact level of the risk along with the mitigation strategies to ensure the project is well managed.  All the risk factors are also affected by the limitations and constraints ensured in the feasibility analysis.   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | Risk ID | Description | Likelihood | Impact | Response Category | Response Action | | 1 | Not enough data for the data visualization feature | Medium | High | Reduce | Find the appropriate tools to generate data or find data from the external databases | | 2 | The project may take longer time to finish than planned | Medium | High | Reduce, Share | Adapt the project and increased workload for each member to meet the demand and expectation goal of each week | | 3 | The development team encounters faults/bugs | High | Low | Reduce | Testing and debugging will be used throughout the development cycle | | 4 | Serious Conflict within the team | Low | Medium | Avoidance | Resolve the conflict with the team | | 4 | Serious Conflict with Client/ Client's expectations | Low | High | Avoidance | Resolve the conflict with the Client | | 5 | Member of the team has unexpected leave | Low | Medium | Transfer | Reevaluate the project and reassign tasks to other members | | 6 | The technical problem that affects the ability to continue the project | Low | Low | Accept, Transfer | Backup the progress every time and use backup technology or devices to continue the progress | | 7 | A planned feature is not fully implemented or is not progressed | Medium | High | Accept, Transfer | Negotiate with the client and think of new features to make up the progress | | 8 | COVID 19 related restrictions | Low | Low | Accept | Absorb to the delivery schedule and scope change |   In the risk assessment process, the project team identified a key risk raised during the first phase of the capstone project which is ‘Not enough data for the data visualization feature’ (risk ID: 1). This risk rose as there is no data for AMR genes in the Staphopia API. The team has developed a strategy to combat this challenge by implementing and making use of all the tools suggested by professor James Hogan. First of all, the team gets the sequence data from the API and makes them as input for AMRFinderPlus. AMRFinderPlus generates the results which our team can insert into our dedicated database and be used to develop AMR genes visualization. This process of preparing data is approved by the clients. Project Experience This section of the report provides some learning and experiences gained by the project team in relation to projects faced during the course of the project delivery.  Situation 1 : A weekly client’s meeting on Zoom  Task: Presenting an interim development of the system to provide how the feature agreed in last weeks’ meeting built into the works.  Action: Once the presentation commenced by the project team member on zoom it was noticed that the system does not work as planned and some critical alignment were not introduced and a trail run was not done.  Result: It was not possible to present the planned run of the feature, though it was a minor element of the presentation, the presentation was not complete.  Learning: The project team understood that the team should do a trial run prior to presenting to the client's team. The team accepted it as a team issue not an individual’s fault or error. From that day onward, the team did review and run the full presentation prior to the client’s meeting and ensured the presentation can be delivered fully.  Situation 2 : Changing the release plan and deliverables  Task: After some discussions with the client, the team agreed to change the some aspects of the project to include more features as well as the release plan   Action: We realized that in order to meet the client’s expectation, we need to move up the schedule for the last two releases, which made it really difficult since the initial release plan was designed to reduce the amount of work toward the end of the semester.  Result: Even though we were behind the schedule for a few weeks, the team still managed to make the final release on time.  Learning: We understood that the ability to adapt to change is one of the most important properties of Agile methodology. The planning phase and communication with the client is also crucial as we need to make sure what we are currently developing is closely aligned with the expected end product. Artefact DescriptionDesign In the first phase, the low-fidelity and medium-fidelity prototypes were designed to illustrate the key ideas, the main purpose, and the major functionalities that would be expected to deliver [Figure x, Figure x]. The design includes six major components. A navigation bar assists users in navigating across the web pages. Two small panels display the information of samples and antibiotics. A search panel is implemented to add or remove samples from the list for visualizing. Adobe XD was used to develop medium-fidelity prototypes. Then, they were delivered to our client to get some feedback and ensure our ideas are feasible and practical. Functionality For the backend of our web application, the functionalities are included as shown below:   * Process data that is stored in Postgres SQL from the previous Capstone team project * Store the generated AMR Gene Data in MongoDB * Make an API endpoint available for both user and frontend websites to query and extract data into the usable JSON format   In terms of the frontend of the website, it uses the formatted JSON data retrieved from the backend and visualizes those data as bubble and network graphs. The frontend also has some functionalities to search, filter, and display information related to AMR samples and antibiotics. The core functions of the frontend are:   * Produce bubble graph and network graph to display the formatted data. This feature would enable public access to the data and transform the complicated data into easily comprehensible visualizations for the target audience. * Display the formatted data about samples and antibiotics in different tables and panels. This feature would facilitate the AMR research of S.Aureus. It helps users save time researching vast amounts of resources since all useful information has already been well presented. * Provide a wide range of mouse actions to interact with data visualizations. The users could drag, click, hover, zoom in and out, and highlight the particular data. It would enhance the reporting capabilities of the portal and improve the aspect of visualizations. * Provide the tool to search and filter the list of data in users' interest. This feature would make the visualizations be scalable and support larger processed data in future development. * Allow users to navigate between AMR visualizations and other functions provided in the website of the previous team. It is aligned with the project scope by extending new features to the existing web application of the previous team.  Architecture The project employs a client-server architecture, where the server is a REST API providing formatted data to the frontend clients. The client sends the request to the server via different actions such that clicking a node, clicking a button, filtering the sample list, etc. Figure x displays the process of data acquisition, handling, and application. The implementation of our application is shown in Figure x.  The backend employs NodeJs and ExpressJs to gather data from MongoDB. After retrieving the data, it is preprocessed and extracted into a usable JSON format. Then Redis - an in-memory data structure store, is used to store the processed data and send it to the client directly upon request. Each request submitted to Redis requires a key to specify the data of interest.  In terms of the frontend, React library is used to develop the user interface and UI components. A React library 'antd' that contains a wide range of interactive and user-friendly components is also used in the development process of the frontend. As for the AMR visualizations, we use D3.js to present the data for the user. Redux library is also implemented and acts as a store to control the flow of events across the application.  For the demonstration purpose, we use Heroku to deploy the AMR visualizations from our project repository. It helps us deliver the app continuously and get feedback from our client more easily. Here is the URL of our demo AMR visualizations: <https://staphbook.herokuapp.com/>.  Finally, we integrate our project into the existing web application of the previous team. The deployment is not performed for the final product due to the huge volume and extraordinary complexity of the PostgresSQL database implemented in the previous work.   Technical DescriptionStructure Our project has three important folders. The first one is the 'express-server' folder that acts as our backend main entry. The second folder is the 'react-client' which is the root folder of the frontend. It contains App.js file which is the frontend main entry. The third one is the 'Nodewebsite' folder that contains the source code of the previous team. There are also some worth-noting files such as package.json files. They are located at the root directory of the server and client. It contains information about the dependency needed to be installed for the react application. Version Control and Continuous Delivery Git is implemented as our project version control system. To facilitate continuous delivery, the Heroku pipeline is also integrated into the development process. Whenever the update takes place, the project will automatically be deployed to Heroku. Figure x describes the feature workflow and Figure x illustrates the Heroku pipeline. A feature branch is added with each feature being created. When the development in a feature branch is finished, Heroku would generate a review app, which is used to get feedback from team members and clients. Once we all agree upon it, a pull request is made to merge the feature branch onto the main branch. Before the new feature is officially released, the app will be tested thoroughly by a wide range of acceptance test cases. Quality and MetricsAcceptance Test Below is a list of tests acquired by comparing the expected results with the actual outcome. The product has been deployed with the following domain name: <https://staphbook.herokuapp.com/>   |  |  |  |  | | --- | --- | --- | --- | | **Input** | **Expected Outcome** | **Result** | **Screenshot/s** | | Open the home page | Display the description for all AMR visualizations. | Pass |  | | Access ‘/amr’ endpoint | Show the bubble chart displaying AMR information of each sample. | Pass |  | | In the Search Sample section, click the antibiotic dropdown | Show all samples that are resistant to the antibiotic. | Pass |  | | Type in the sample ID in the search bar | Dropdown the menu, selecting all the samples matching that ID | Pass |  | | Select a sample and click Delete. | The graph ignores that sample. | Pass |  | | Select an antibiotic and click Delete | The graph ignores the AMR bubble and its samples. | Pass |  | | Click Restore button | All removed samples and antibiotics appear again in the graph | Pass |  | | Hover the pointer over the sample in the graph | Highlight the sample and show its ID. | Pass |  | | Click a sample in the graph | Showing sample information in the AMR Table, Antibiotic and Sample Information panels. | Pass |  | | Click an antibiotic in the graph | Zoom in that antibiotic and see further information about its sample | Pass |  | | Click the outer circle of the graph | Zoom out and see the whole graph | Pass |  | | Click the ‘AMR Network Graph’ tab | Change the current chart into the node network graph. | Pass |  | | Hover over the sample. | Highlight the cluster containing that sample. | Pass |  | | Drag the sample to another position | The sample follow the pointer, then go back to its initial position | Pass |  | | Scroll up and down | Zoom in and zoom out of the graph | Pass |  | | In the Search sample panel, click the ‘Minimum Spanning Tree’ tickbox | Display the minimum spaning tree of node network | Pass |  |  Burndown Chart The burndown chart is created to estimate the performance of our work during sprints: https://lh3.googleusercontent.com/0hTwtCcuthXc8RLg83uclAPesFgnYvMAsL8LILGV8HE9XDJZJ6cijx1QL6yXmzR6OtS59JjS-eKcJEZFd-SofehNnkUQckG0O67eL38_QitD1XaZ6XFmnmsSefdnSa7yxmK18G7X |