

NIST Capstone Project

1. Overview & Scope

NIST has requested a web application to process raw files to view the functionality and precision of mass spectrometers when processing samples. Our job is to create a web application that allows a user to drag and drop raw files into a browser, process the data with a pipeline and present the results in a simple and intuitive fashion

2. Problems, Value Proposition & Goals

The problem being addressed is one of precision and availability of resources. As of now, the scientific community has to rely on the mass spectrometer to produce results and assume that all aspects of the machine are functioning within the appropriate boundaries. There are tools available to determine the health of a mass spectrometer, however each of these tools are available separately. The value of a pipelined web application is that all of the resources to evaluate the health of a mass spectrometer will be available in a single place. The goal of the project is to produce a containerized pipelined web application to evaluate mass spectrometer health. The application will be available to the scientific community in a browser or privately with the download of the containerized application.

3. Goals

Goal Title	Description	Deadline
Shiny Docker Container	<ul style="list-style-type: none">A fully functional container of the shiny web server and application.	March 1
MySQL Container	<ul style="list-style-type: none">A fully functional container of a mySQL database with persistent storage	March 1
Docker Network Link	<ul style="list-style-type: none">A reliable link between mySQL and Shiny Container should be made	March 1
Drag and Drop	<ul style="list-style-type: none">The Homepage should include drag	March 1

Homepage of Web Application	<ul style="list-style-type: none"> and drop functionality for .raw file. • a drop down widget to indicate “Hela sample” • a progress bar • link to the pipeline page 	
Pipeline Web Page	<ul style="list-style-type: none"> • The second web page after the homepage shows the results of the raw file. • There should be 5 boxes: sample, LC, source, MS1, MS2. • Each box should be “clickable” or each box should have a link to their perspective pages • The results of each of the boxes should indicate a result with a color; green= good, yellow= mediocre, red = poor. • There should be an index for the user indicating what the colors represent. 	April 1
“Sample” Web Page	<ul style="list-style-type: none"> • Graphics for the sample 	April 1
“LC” Web Page	<ul style="list-style-type: none"> • Graphics for the LC 	April 1
“Source” Web Page	<ul style="list-style-type: none"> • Graphics for the source 	April 1
“MS1” Web Page	<ul style="list-style-type: none"> • Graphics for the MS1 	April 1
“MS2” Web Page	<ul style="list-style-type: none"> • Graphics for the MS2 	April 1

4. Non-Goals

Non-Goal Title	Description
Installing docker and its dependencies on other machines	<ul style="list-style-type: none"> • We expect docker to be installed on all machines that want to use our application locally..
Making Assumptions about data	<ul style="list-style-type: none"> • We will not be using any machine learning algorithms to make assumptions about the data or determine outliers

	<ul style="list-style-type: none"> • We expect important data points to be given to us by NIST representatives.
--	--

5. Risks and Mitigation

Risk	Mitigation
Pipeline tools not compatible with R	Research tools extensively and investigate compatibility .
Data doesn't persist in database	Docker volumes
Multiple users on the same IP can grab old data from other users in the same facility	Restrict access to pulling old data from database for multiple users working from the same IP