Alexia Filler
CMSI 402
4/6/2020

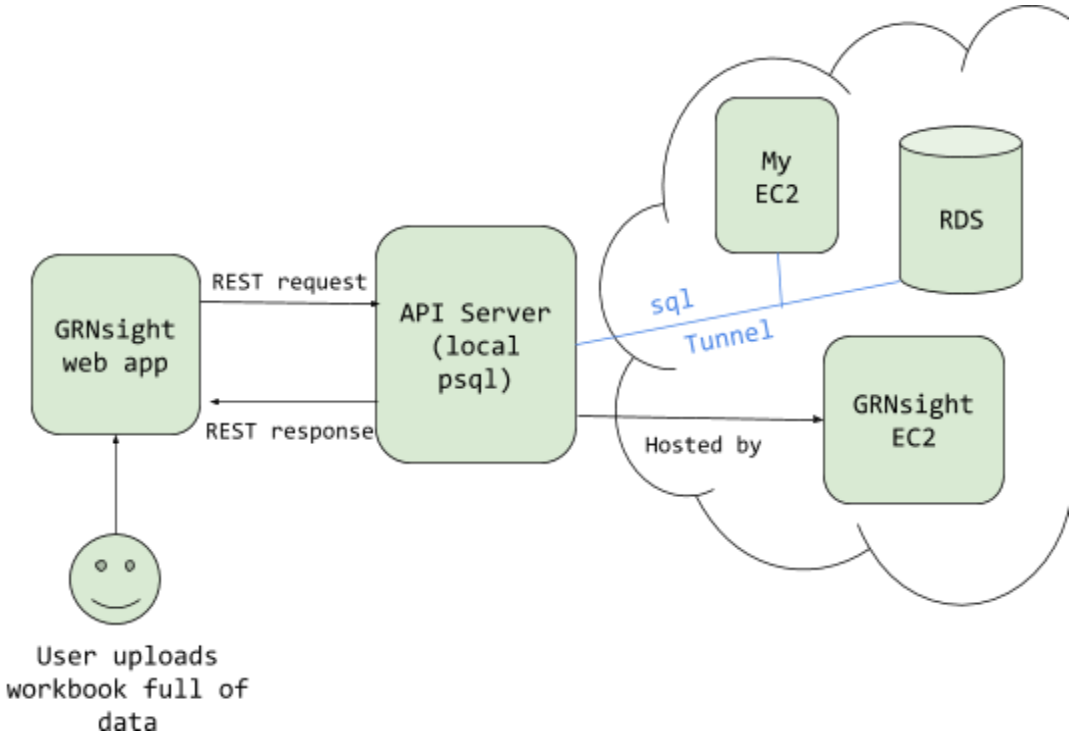
UPDATED Software Requirements Specification for GRNsight Expression Database

5.1 Introduction

5.1.1 The GRNsight Expression Database project will be a feature added to GRNsight, a gene regulatory network visualization site. The database itself will contain gene expression data so that users of the site can access and interact with pre-existing expression data without necessarily having to upload their own expression data file. This will help ensure that expression data has consistent formatting, and it will provide examples of this proper formatting for the user. Also, it will allow the user to do node coloring on their GRN data. The database itself will require use of AWS RDS (PostgreSQL engine) and an AWS EC2 instance for the sake of tunnelling to the RDS instance, and integrating the database into the site will require the use of NodeJS (particularly the sequelize package)..

5.1.2 Diagrams

5.1.2.1 High-Level UML Diagram



5.1.3 Outline of Document

The remainder of this document is structured as follows:

- 5.2 contains a list of the functional requirements of this project.
- 5.3 contains a list of the performance requirements of this project.
- 5.4 contains a list of the environment requirements of this project.

5.2 Functional Requirements

This section outlines the features that the completed system can be expected to have.

5.2.1 Expression Database

The expression database is what is at the heart of this project; it is the structure that will store the desired expression data in a normalized form so that it can be easily accessed by users of GRNsight.

5.2.1.1 Expression tables

The expression database shall contain normalized expression tables, which will contain the following fields:

- Index ← a generated index for each entry
- ID ← a unique identifier that tells which gene the data describes
- Standard name ← the standard name of the gene
- Data ← all columns with expression data from the experiments/publications

5.2.1.2 Production rates table

The expression database shall contain normalized production rate tables, which will contain the following fields:

- ID ← a unique identifier that tells which gene the data describes
- Standard name ← the standard name of the gene
- Production rates ← data that is an estimate of the production rates

5.2.1.3 Degradation rates table

The expression database shall contain normalized degradation rate tables, which will contain the following fields:

- ID ← a unique identifier that tells which gene the data describes
- Standard name ← the standard name of the gene

- Degradation rates ← degradation rate data

5.2.1.4 Metadata table

The expression database shall contain normalized metadata tables, which will contain the following fields:

- Index ← a generated index for the entry
- Control strain
- Treatment strain
- Control condition
- Treatment condition
- Concentration
- Concentration units
- Timepoints
- Time units
- Replicate number

5.2.1.5 References table

The expression database shall contain normalized reference tables, which will contain the publication information of the studies used in gathering the data in the database.

5.2.2 Graphical User Interface

The graphical user interface required will be an add-on to the existing GRNsight interface. It will provide the user with the ability to interact with the expression database by downloading or viewing the data or by seeing how the data affects the displayed GRN.

5.2.2.1 Dropdown selection

The menu of options shall exist in the Data panel on the left of the display, with dropdown options so the user can select the data set and time points within the expression database that the user would like for node coloring.

5.2.2.2 Select "Node Coloring With Expression Database"

The options panel will contain a button that the user can select to turn node coloring with expression database data on or off.

5.2.2.3 Download data option

The options panel will contain a button that the user can select to download the desired data from the expression database.

5.3 Performance Requirements

This section describes performance requirements for the completed expression database project.

5.3.1 Database access and query result return time

5.3.1.1 The database shall return the results of a valid query within 10 seconds.

5.3.2 Ability to display desired data without errors

5.3.2.1 The GUI shall display the results of a valid query without errors.

5.3.2.2 The GUI shall display appropriate error messages when provided an invalid query.

5.4 Environment Requirements

This section outlines the hardware and software that will be required to bring this project to completion.

5.4.1 Hardware requirements

5.4.1.1 Machine that will be able to support the listed software requirements

5.4.2 Software requirements

5.4.2.1 GitHub account with read/write access to GRNsight repository

5.4.2.2 AWS Educate Account

5.4.2.3 AWS RDS instance (free tier with PostgreSQL engine)

5.4.2.4 AWS EC2 instance (free tier)

5.4.2.5 NodeJS

5.4.2.6 PostgreSQL

5.4.2.7 Code editor