



From: GSoC-Idea7 Mentors

To: Candidate Applicants

Date: 28-02-2014

Subject: Candidate Evaluation Task -1

## Description:

This document describes the evaluation task that will be used to fairly evaluate the candidates.

## Objective:

Describing the inputs, deliverables, time frame and assessment rules of the evaluation task.

## Overview of GSoC-Idea 7:

We are studying the effects of mutations on protein-protein interaction (PPI) networks in cancer. The effects of over 800,000 missense mutations are being analyzed and stored in the Cancer Network Altering-Variant Database (CNA-VD), a relational database implemented in MySQL. CAN-VD stores the PPI interactions mediated by wildtype and variant protein sequences to build and compare the PPI network in the two conditions and understand the effects of mutations on the network topology and, consequently, the cellular and biological functions of the cancer system.

## Task Goals:

1. Develop a web-based interface for PPI data retrieval with a comprehensive search function.
2. Develop a visualization module using Cytoscape.js to visualize the retrieved PPI data as a network for the wildtype interactions only.

## Task Details:

### 1. Platform, Servers and language:

OS: Windows 7 64bit

Webserver: IIS 7.5 (Windows 7 default)

Database server: MySQL (V5.1 or later)

Scripting: HTML, Javascript, PHP and JSON

### 2.The Sample Database:

The data tables can be downloaded as text files from [here](#). Please download the tables and import it to a MySQL database. In the same folder you will find the database ERD (CAN-VD-ERD.png). Please use it as a guide for the relations between tables.

### 3.Cytoscape.js:

Information about how to integrate Cytoscape.js into your website can be found at [Cytoscape.js webpage](#).

### 4.Pages to be developed:

The desired web interface for CAN-VD is expected to perform three main functions that are browsing, searching and network visualization. Therefore, you are asked to developed pages that perform the three main functions.



- A. Browsing: A basic data browsing page(s) that allow browsing, listing, sorting and filtering the data retrieved from the database (please see the browsing options at the [Rice Proteogenomic Database](#), for instance).
- B. Searching: simple and advanced search function and displaying the search results in a readable and organized manner (please see the searching and results options at the [Rice Proteogenomic Database](#) and [GeneMania Database](#)).
- C. Network visualization: in the database there is a table named T\_Interactions. In this table we store the predicted interactions. In the sample database, we added some wild type interactions only. The interaction takes place between central protein that contains special features (the Domain\_EnsPID in the table) and another protein(s) (Interactions\_EnsPID). Your task is to retrieve the interactions of certain protein (based on the search results, for instance) and display the interactions as network using Cytoscape.js.

#### **5. Bonus tasks:**

You can get extra point in the assessment by doing some of the following bonus tasks.

- A. Layout: include a style sheet that helps in customizing the interface layout.
- B. Adding balloons that show up when clicking protein name, mutations or interaction and dynamically display some information in the balloon.
- C. Suggesting better design for the database or functions to be added to the web interface.

#### **6. Expected Deliverables:**

- A. The MySQL Database: as SQL file, dump file or backup file (whichever you think easier for deployment).
- B. HTML and PHP pages: including any required Javascript files for Cytoscape.js or written by you.
- C. Deployment procedures: a document describes how to deploy your project.

#### **7. Time frame:**

The time frame of the task is 4 days (March 1<sup>st</sup> ~ March 4<sup>th</sup>)

#### **Task Assessment:**

The assessment will be done by the mentors, as following:

- 1- MySQL database creation and deployment: 10 points
- 2- Functionality: total 40 points
  - a. Browsing: 10 points
  - b. Searching and results display: 15 points
  - c. Network visualization: 15 point
- 3- Code Documentation: 10 points
- 4- Timeframe: -5 points for each extra day.
- 5- Bonus tasks: 5 point per task.

We wish you all the best and looking forward to see your work.

GSoC-Idea7 Mentors:

Mohamed Helmy

Max Franz

Gary Bader

Contact: [mohamed.attishahata@utoronto.ca](mailto:mohamed.attishahata@utoronto.ca), [maxkfranz@gmail.com](mailto:maxkfranz@gmail.com), [gary.bader@utoronto.ca](mailto:gary.bader@utoronto.ca)