Design document | Premonition

**Management Summery:**  
  
Premonition was created to identify interactions between proteins and present this information in a clear, easy to digest way. This enables researchers to efficiently search for, and understand the interactions between proteins. This is done by presenting a 2D map of proteins with coloured lines connecting proteins that interact with one another. By interacting with proteins information about the protein will be displayed with referral to the proteins NCBI entry so the user can easily access the proteins more complex information or to search research papers. The development costs are around €24,000 (20 weeks of development time with a total of 24 labour hours a week for €50 per hour). Though these costs may vary (around 75-200%) depending on the development cycle. The major milestones are front-end to back-end information exchange, back-end execution of the Premonition Python program, GUI display for the linked proteins, and design of the web interface.

**Background and context:**

Premonition already exists as a Python script that gets called and executed from a system’s command terminal. Though for many bio-informatics this is fine, but it is not as accessible to the greater scientific community as would be ideal. The goal of creating a web interface for Premonition is to make its functionality easily accessible so that researchers and students can make use of the program without needing an IT background. The program will be hosted by Hanze University of Applied Sciences as a free-to-use service for the scientific community. The program will display a 2D map consisting of proteins with coloured links to each other to visualize the interaction between them. Each interaction and protein itself will also function as a way to display more in-depth information about the protein in question. The user only has to upload a file consisting of the proteins of interest, as well as a reference file containing known protein-to-protein interactions. Based on this reference the program will created a map using the highest probability options.

**Requirements Specifications:**

Top-level features:

The following are the must-have features for the Premonition web-application.

1. Web-based datafile upload system with optional parameter configuration. This feature is accessible on the surface web, so anyone in the world may use it.
2. A robust back-end python executor that can deal with data that has abnormalities. In cases the program simply cannot execute the files, this will be communicated to the user in a clear, and easy to understand way.
3. A GUI system that creates a 2D map of the proteins and their interactions based on highest probability. Each protein element in the map has a small window summary information which is called from the NCBI database. This information window will also contain a direct link to the NCBI database entry of the protein in question. The individual links between protein elements also clearly display the molecules that are being transferred.

Use cases:

|  |  |
| --- | --- |
| ID: | UC-1 |
| Title: | Identify possible interactions of a protein. |
| Description: | Researcher is working on understanding why a biological process gets interrupted. |
| Primary Actor: | Researcher. |
| Preconditions: | Researcher has loaded the web-interface in their browser. |
| Postconditions: | The researcher has access to a map of predicted protein interactions. |
| Main  Success Scenario: | 1. Researcher uploads a file with proteins of interest and a reference sheet. 2. System calculates the probability of interactions between the proteins in the query sheet. 3. System displays a 2D map of the proteins with links to other proteins they are likely to interact with.  4. Researcher clicks on a protein that can potentially disrupt the bio-process of interest, displaying a window with summarized information about the protein.  5. User continues to click on the NCBI link to read research papers about the interrupting protein. |
| Extensions: | 1A: The given protein of interest sheet does not contain the right format of proteins. 1A1: The systems displays a error message to let the user know what’s wrong and offers to try again. 1A2: The user either rectifies the issue or backs out the program. 2A: The reference sheet is not the correct format.  2A1: The systems displays a error message to let the user know what’s wrong and offers to try again. 2A2: The user either rectifies the issue or backs out the program.  5A: The system cannot link the query protein’s interactions with other proteins. 5A1: The system displays a error message. Asks the user to check the given query. 2A2: The user either rectifies the issue or backs out the program. |
| Frequency of Use: | Once in a while when the need for clear visualization arises. |
| Status: | In development. |
| Owner: | N. Mooldijk |
| Priority: | High priority. |

**Technical requirements and maintenance:**

Java v17 with Springboot framework, and Java.IO libraries.   
Python 3.8 with systems library.

The hosting server needs to be able to support Python and Java, and this includes the updating of libraries after major updates so the program may continue to work.

Git:  
<https://github.com/JoshuaTolhuis/Premontion_Web.git>