

Week 12

Work done

The week has been spent studying database design for Django and network graph representation in order to store and represent our protein information.

Django has Object-relational mapping(ORM) from Python to the underlying SQL database. This allows a programmer to work in an iterative manner with data that is set structured and accessed through declarative means.

The most suited model seems to be a Protein model containing the name of the protein and a many-to-many field called interactions mapping to other Proteins. This self referential many-to-many model can contain a through table, which maps to an SQL table with foreign keys to and from the various proteins. These connections which make up the through table have been named Relations, and each one stores the interacting proteins, as well as data pertaining to the scientific article, such the title, abstract, and html link.

On the other side of the project, representing the protein network will be done using [d3js](#), a javascript library for manipulating web documents based on data. The library syntax is similar to jquery, and the library offers plenty of examples ready to use.

The researcher has chosen a directed graph approach with labels which allows for hyperlink embedding. By clicking on a node, the user is sent directly to the desired webpage. When combined with the information for the database, this will allow a user to click on a protein node and see the article that links that protein to the original one he searched for.

The code for both of these has been uploaded under 02Website\01 Database and Network Graph.

Problems

Outside of the underperforming neural networks, nothing new.

To do for next week:

The next week and next sprint will be spent putting together the website and pubmed crawler, essentially bridging all the separate components of the project together.