Week 09

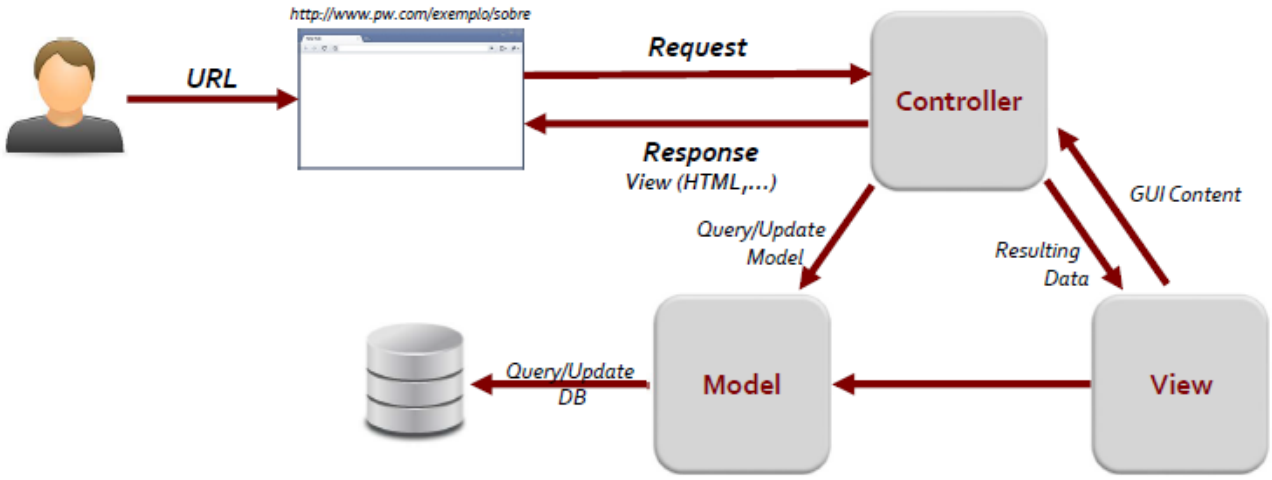
**Work done**

The week has been spent studying the Django framework and Pubmed interfaces.

Overall Django  is a Model View Controller type framework using HTML, CSS, JavaScript and Python.

Since websites are static in nature, their dynamism can only happens as a request-response interaction between a client machine and a server. When a request is sent to the server, a respective controller function handles it and sends a reply. This is generally a generated view containing classic website coding(HTML, CSS, JS).

The controller also has the responsibility of communicating with the back end database. Rather than forcing the programmer to deal with SQL queries and returns, Django uses an ORM (Object Relational Mapper). In essence, the tables in the SQL database are mapped to python objects that are easier to interact with.




The theoretical details are easy enough to understand. Getting used to the idiosyncrasies of the MVC Django implementation, however, takes practice, and learning the ins and outs takes time. A small mockup website was created to test out some basic features.

On the other front, there are multiple already implemented Pubmed retrieval interfaces in python. These query the Pubmed library and return the results as objects with accessible properties such as article.abstract or article.title. We will mainly be using two: pubmed-lookup and pymed.

**Problems**

At this point in time, the researcher has 4 tasks:

1) Entity recognition problem: Since the Pubmed articles don't come with a list of proteins, effort must be made into developing a recognizer of some sort.

2) Pubmed crawler: Since the Pubmed maintainers discourage heavy queries and concurrent lookups, a way must be developed to index and classify all the articles for a quicker lookup.

3) Familiarization and building a website in a new framework: Django

4) Writing a summary research article about the project.

Overall, the researcher doubts all these will be finished by the deadline. At best, a very rudimentary tool with poor results can be created. As such, perhaps the advisor can suggest what should be prioritized.

**To do for next week:**

Accumulating information and applications relating to the entity recognition problem and applications.