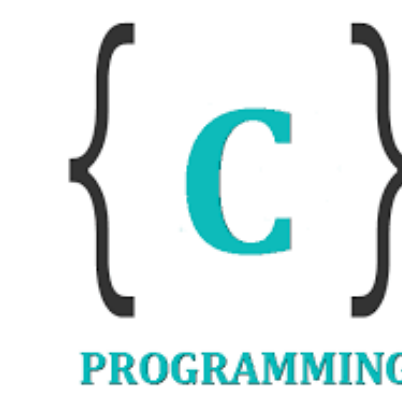




Senior Project, 2016, Spring

GenomePro 3.0

Student: Roberto Arciniegas, Florida International University
Mentor: Michael Robinson, Florida International University
Instructor: Masoud Sadjadi, Florida International University

Problem

A genome is an organism's complete set of DNA, including all of its genes. Each genome contains all of the information needed to build and maintain that organism. The work of the Human Genome Project has allowed researchers to begin to understand the blueprint for building a person. Genome files can be extensively large therefore there is a need for specific tools and methods for cleaning up those files and for sequence processing that can lead to the predictions of undiscovered genes .

Current System

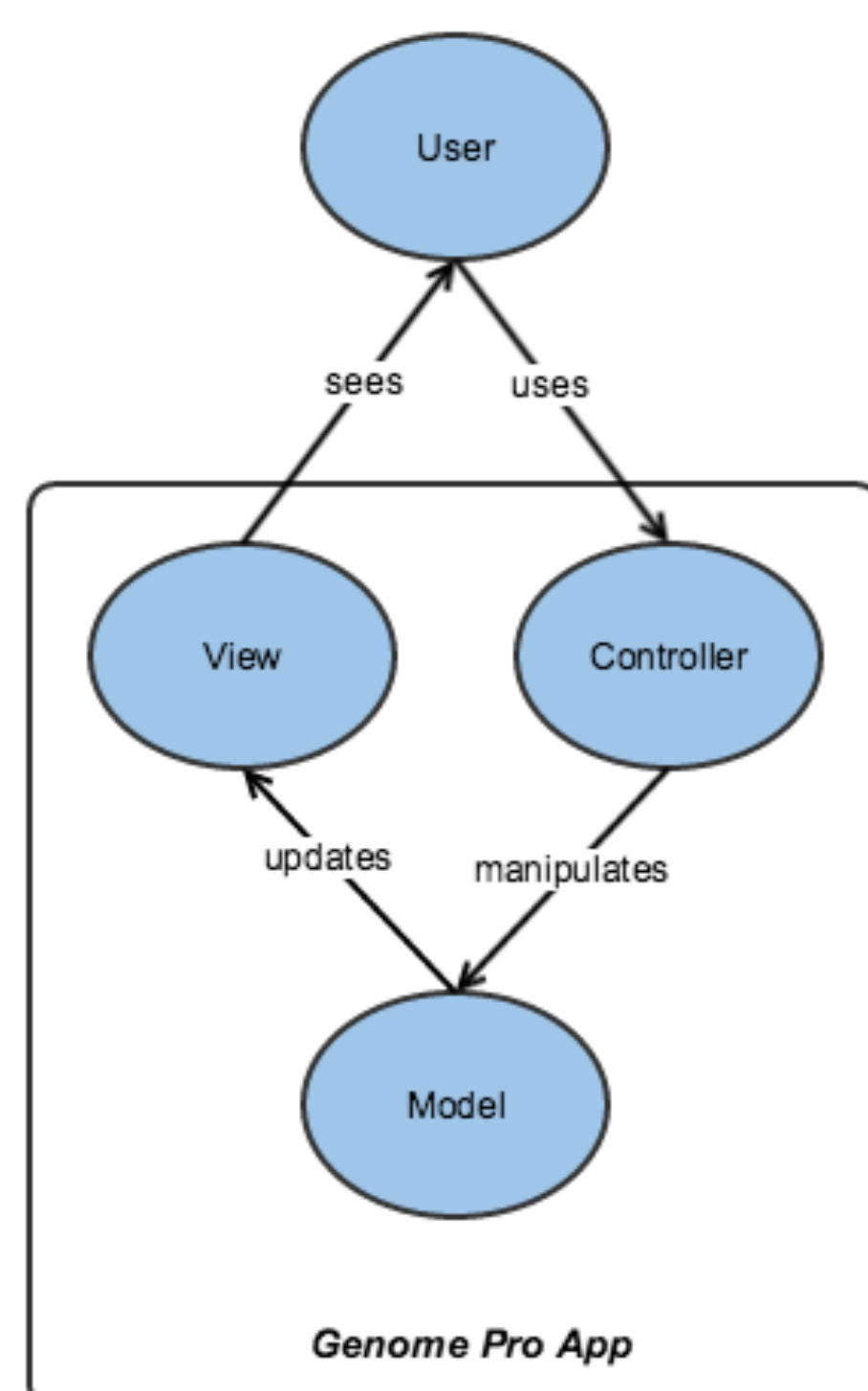
GenomePro 2.0 is a sophisticated web service that hopes to be as useful to a Biomedical Researcher as a scientific calculator is to a mathematician. It provides the ease of use and accessibility to a task that would otherwise be very tedious and time consuming. With extras such as charts generation from the data uploaded, Biomedical Researchers can see exactly where repeats in genome data happen, and compare side by side different subsequences of bases found in data types such as DNA, RNA, and Protein. All the result files are stored securely, and can be downloaded anytime, anywhere

Requirements

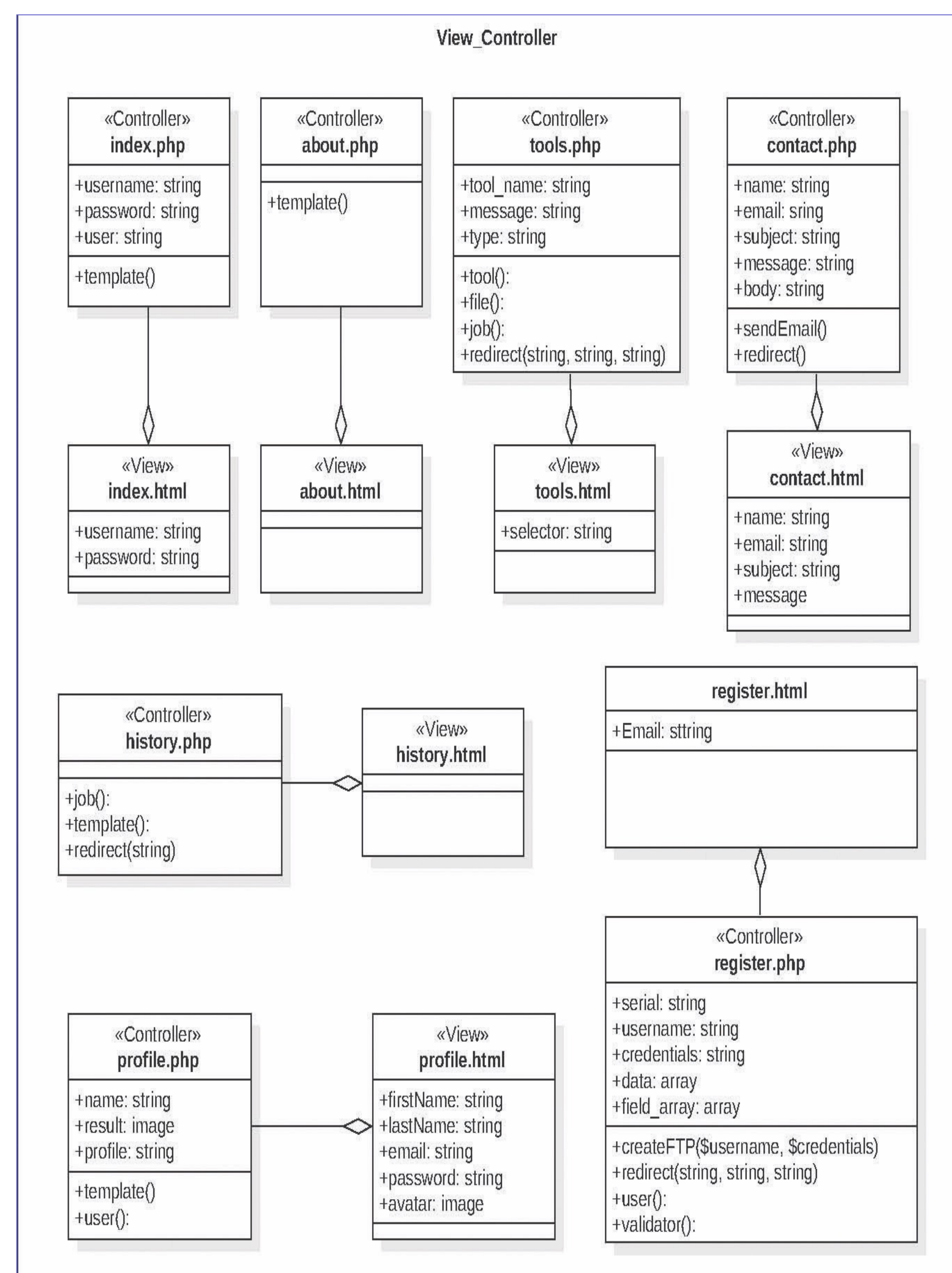
- C Program: File Validator
- Tool: Genome Signatures
- Tool: Unique Sequences
- Tool: Genome Probes
- Tool: Extract Sequences
- Tool: Sort Sequences
- Tool: Create Maps
- Tool: Generate Charts
- Queue jobs in a FCFS (FIFO) fashion.
- FTP results with encryption.
- Automate the execution of jobs.
- Save results securely.
- Allow multiple jobs and chaining.

System Design

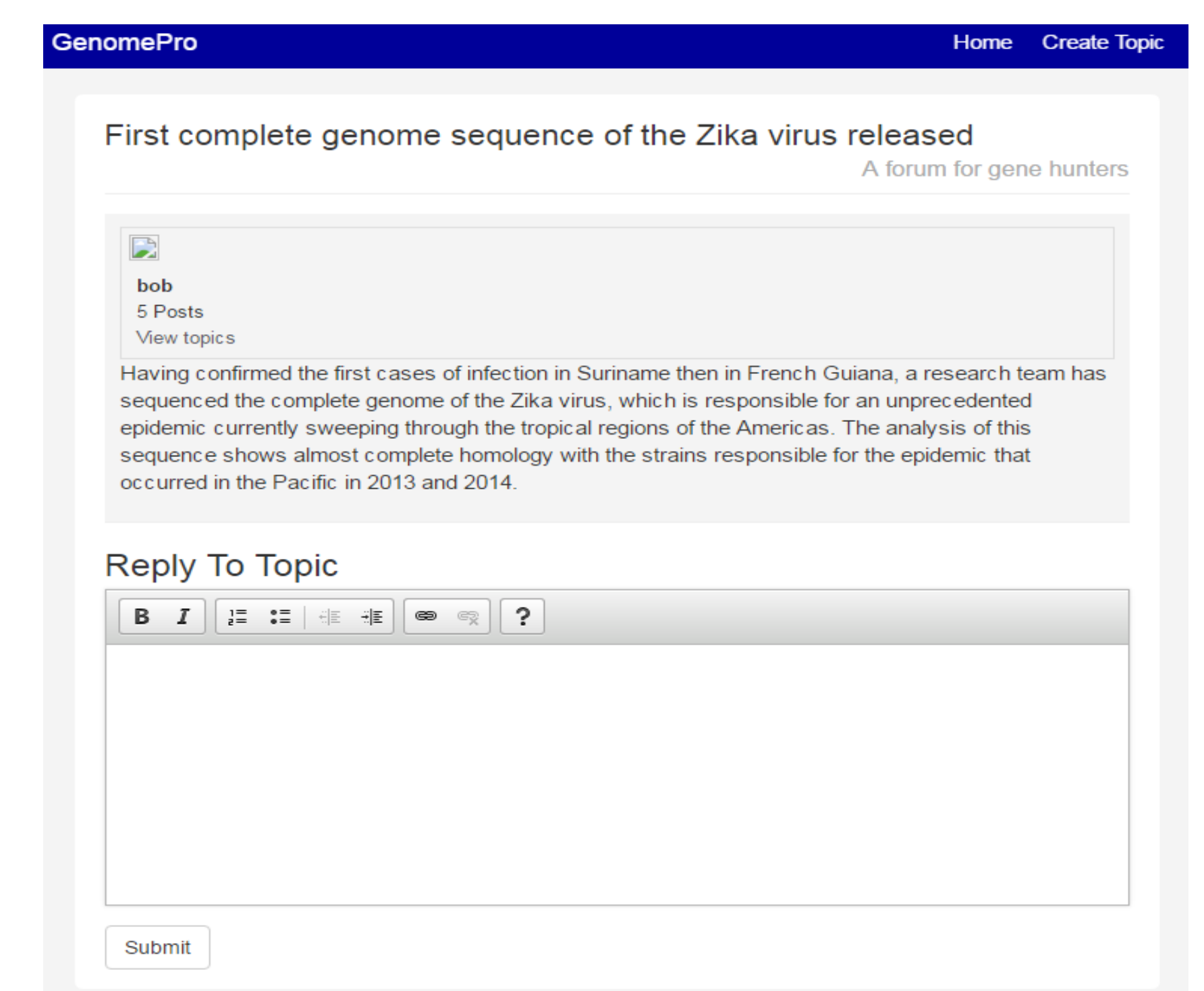
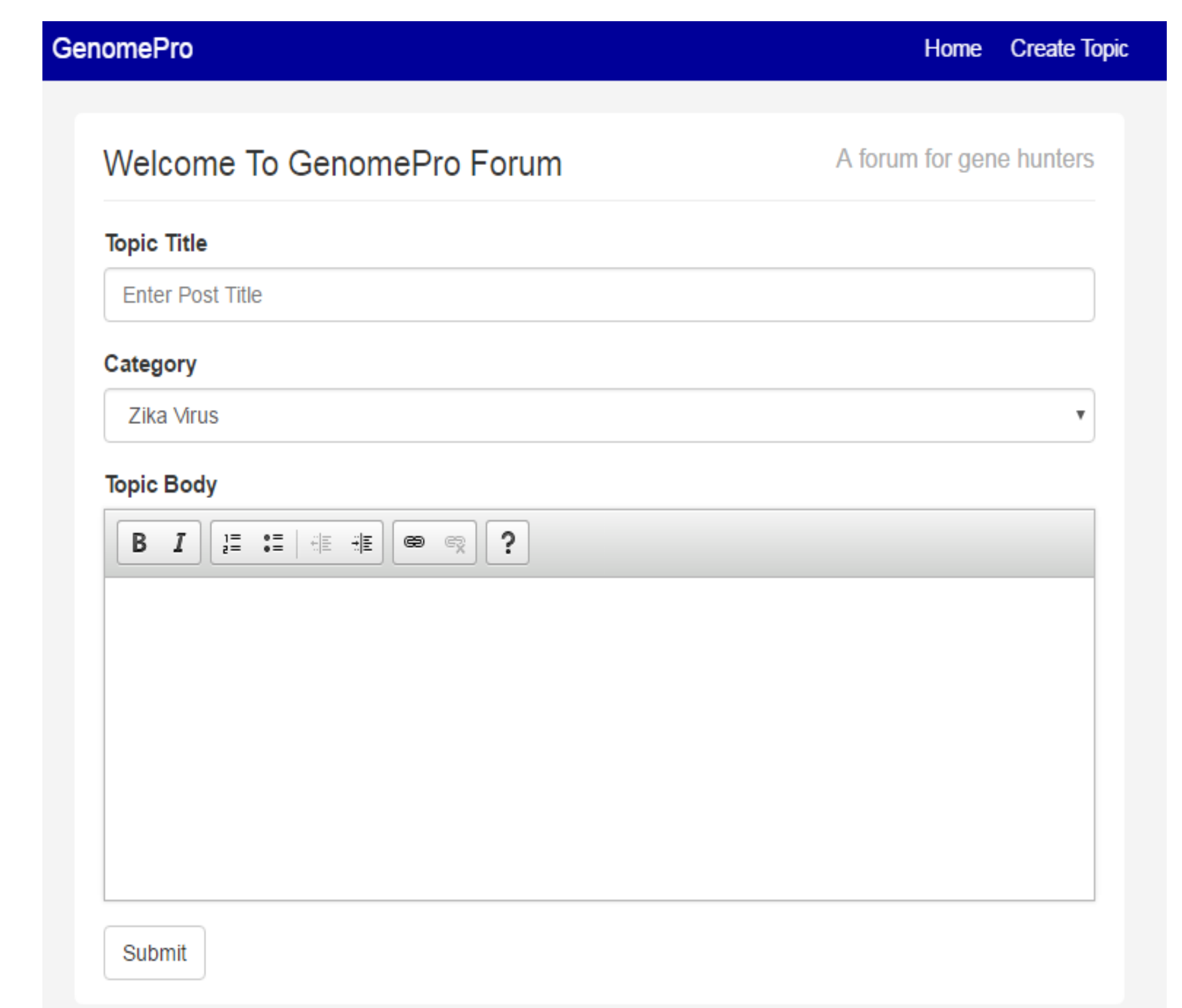
The Genomic Sequences application provides a model-view-controller (MVC) design pattern. The model represents data, business logic and rules. The view represents the presentation. Finally, the controller accepts input and provides a bridge of communication between models and views.



Object Design



Screenshots



Verification



Implementation

GenomePro 2.0 offers the speed of the ANSI C language and its Standard Library functions to deliver powerful performance for processing large volumes of genomic data.

The web site was implemented using HTML and CSS organized with JavaScript as the common client-side scripting language to handle user interactions within the pages.

The backend implementation required PHP for the server-side scripting language, which was successfully completed by Guido Ruiz.

Summary

Data sharing encourages more connection and collaboration between researchers, which can result in important new findings within the field. In a time of reduced monetary investment for science and research, data sharing is more efficient because it allows researchers to share resources. In this release of Genome Pro we have implemented a forum that allows users to share the results of their research.

Acknowledgement

The material presented in this poster is based upon the work supported by Michael Robinson. I am thankful to the help that I received from my group members.