



# CS486C – Senior Capstone Design in Computer Science

## Project Description

<b>Project Title:</b> PIMpoint Collaborator Data Entry Point	
<b>Sponsor Information:</b>  AN AFFILIATE OF  City of Hope.	Jonathon Todd, Associate Bioinformatician Translational Genomics Research Institute Pathogen and Microbiome Division (TGen North) jtodd@tgen.org 623-755-7240

### Project Overview:

Public health is an ever shifting target. Bacteria, fungi, viruses, and other pathogens are in a constant struggle to grow, evolve and spread across the globe. There is a high demand for new solutions, treatments and cures requiring the dedication of countless scientists, researchers, and specialists from all over the world. Cooperation at this global scale is no small task; errors can lead to significant increases in the time it takes to get results. For example, samples taken from Africa can spend weeks in transit before arriving at an international lab, which will then require further time for analysis. With so much time required for these steps in the process, it is crucial that all other parts of the process are streamlined to ensure quick results returned to those who need them.



TGen North works with dozens of collaborators worldwide to tackle various public health issues. This includes DNA sequencing and analysis on a variety of sample types including patient samples from clinical partners, and environmental samples like bacteria and fungi that are relevant to human health. In an attempt to organize a tidal wave of important data, we have been working on a database and tracking system called PIMpoint. The specific goals of the system are information sharing with our partners, and precise sample tracking throughout every step of the analysis process. Unfortunately, this system is not saving us as much time as it could due to a failure to clearly organize submission of samples and information about them by our external collaborators.

Currently, our diverse set of collaborators are using equally diverse methods for data sharing: mailed hardcopy print-outs, email, Microsoft Excel documents, and more. To further complicate the matter, each collaborator may have their own system for how they organize and label samples. Samples may be incorrectly, illegibly, or ambiguously labelled, and metadata may need to be hand-entered when the samples are received. Because of

this lack of standardization, too much time is wasted as we try to decipher and correct the data we receive.

To combat these issues, we envision our solution to be a secure web app that provides a clear, easy to use gateway to our internal workflow system for our external collaborators. After authenticating, collaborators submitting samples will use a form-based interface that ensures that all required information is provided and it in the proper formats. Having entered the key information about the job, users will be able to print out custom labels (e.g. on standard Avery label sheets) to use in labeling their samples, ensuring clarity and perfect correspondence with job metadata entered. Some key features we would like to see in the application:

- A secure, modern Web2.0 web application, to allow access from any modern web browser.
- User accounts and profiles (mailing address, order history/status, etc.), with secure authentication to strictly control access to the system (only known users/domains should be able to view/add data).
- A dashboard for authenticated users, allowing them to view their submitted jobs and the current status of those jobs, as well as submitting new jobs. May also provide access to analysis results for completed jobs.
- A clean, configurable forms-based interface. Based on the job characteristics selected by users, the GUI will gather precisely the information relevant for that job.
- Integration with our internal database and workflow management system, so that entered data can be easily and automatically stored.
- The ability to generate custom labels that collaborators may attach to the samples they send us, instead of using their own (haphazard) in-house labeling systems.
- Provide shipping instructions and printable shipping labels to the collaborator after they have entered their data.

Design of the system should be as flexible and configurable as possible, as it may be useful to offer/sell the system to other laboratories with similar profiles and workflows, to use in organizing their incoming job information. Thus forms and other interface elements should be modular and easily (re)configurable for other organizations, i.e. implemented as modular elements that can be easily edited for other deployments.

Having a clean, highly-usable workflow management system will reduce both errors and turn-around times in the highly time-sensitive area of genomic identification and fingerprinting, which can ultimately save lives. A properly designed product will have major positive impact not just for the TGEN lab, but could be made available as a general purpose tool for managing the intake task for similar labs around the country.

### **Knowledge, skills, and expertise required for this project:**

- Knowledge of modern Web2.0 development techniques and frameworks (html, CSS, Bootstrap, etc.)
- UI/UX Design and evaluation expertise.
- Knowledge of webserver configuration, database organization and integration, and other systems concepts.

- Team will need to develop basic understanding of the sample intake and analysis process, as well as informatics tools already embedded in our workflow.

### **Equipment Requirements:**

- There should be no equipment or software required other than a development platform and software/tools freely available online.
- We will provide a test database and schema for this project to interact with.

### **Software and other Deliverables:**

- A complete user manual for end-users of the system, preferably implemented as an online document or help system in the web app itself
- A complete systems administration manual, detailing installation and configuration of the system, including editing/creation of forms and data entry workflows. Should be simple for lab managers elsewhere to install and configure their systems
- A strong as-built report detailing the design and implementation of the product in a complete, clear and professional manner. This document should provide a strong basis for future development of the product.
- Complete professionally-documented codebase, delivered both as a repository in GitHub, BitBucket, or some other version control repository; and as a physical archive on a USB drive.