**Alexandria Persing**

**Progress Report Log**

Capstone Spring 2015

Elevator Pitch: I’m creating a search engine for text documents. Basically, it will allow you to pass in a document, then it determines the content of that document and gives back other documents with similar content. This will be extremely useful for applications such as finding books or articles similar to one you’ve read recently.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sprint** | **Rating**  **(0-4)** | **Backlog Items** | **Status** | **Student Comments** | **Faculty Comments** |
| 1 |  | **1)** Create an api for document mapping to a given content tree  **2)** Unit tests for document mapping  **3)** Create an application to load/save/edit content trees  **4)** Change the way the file io for text extraction works, so that it uses the readAllLines iterator (so as not to use overly much memory), rather than the readAllText method | 100%  100%  100%  100% | Things I have completed already by the first meeting:  **1)** Finding test data  **2)** An api for creating content trees and adding and subtracting data from them  **3)** Unit tests for the content trees  **4)** An api for extracting individual post text from a file  **5)** Unit tests for text extraction  ------------------------------------------------------  I made the api for document mapping as well as unit tests to go with them. I also created a gui application which allows you to make content trees, edit them, and load and save them to the file system. |  |
| 2 |  | **1)** Create the api for comparing documents  **2)** Assemble a set of test documents/queries  **3)** Create a content tree based on the test documents  **4)** Actually test my system against the test documents and document results  **5)** Create an application for scanning documents which also allows you to view the created data tree in some way |  |  |  |
| 3 |  |  |  |  |  |
| 4 |  |  |  |  |  |
| 5 |  |  |  |  |  |
| 6 |  |  |  |  |  |
| 7  8 |  |  |  |  |  |
| 9 |  |  |  |  |  |