Senior Projects

Spring 2022

**Individual** Weekly Report

NAME: Mason Lanham TEAM: Word Data Extractor

Period covered by this report: From April 13th, 2022 to April 20th, 2022

**Describe briefly this week’s activities (if any) and estimate the amount of time devoted to each one. MAKE SURE YOU SAVE A COPY OF EACH REPORT.**

1. Research Activities 0 Hours

**Explain:**

2. Out of class time meeting(s) 0 Hours

**Explain:**

3. Documentation Activities 0 Hours (other than code documentation)

(e.g., producing requirements document, test plan documentation, project related presentation, etc.)

**Explain:**

4. Design Activities 0 Hours

(e.g. Producing System Architecture/Design, etc.)

**Explain:**

5. Coding Activities (including code documentation) 26 Hours

**Explain:** I integrated all the work that was done by my team members together and added to it. They had developed programs that extracted text, tables, and images, but didn’t capture all the data from the document that we needed. I developed their programs to be able to capture the additional data and modified them, so they interfaced with the class structure in our design document that I implemented as “skeleton” classes a few weeks ago. The result was an unordered list of, Textual Element, Graphical Element, and Table objects. I then developed and tested a method to order those objects by comparing the text I could extract with docx2python to the unordered list of objects. I developed a method for identifying and extracting procedures from our textual elements as well as tables, however it has some issues that I will try to iron out this next week.

6. Testing Activities 0 Hours

**Explain:**

7. Other Project Related Activities 0 Hours

**Explain:**

8. Other Non-project Related Activities 6 Hours

(e.g., reading for/preparing/**non-project related** in-class presentation/report)

**Explain:** I researched and read some articles that I used to write my writing assignment.

9. My project related planned activities for the coming week are:

Ironing out the issues that I have with the procedure parser, and issues with the ordering of the elements.

Integrating the code I’ve written in the last week with the interpreter.

Writing and practicing my parts for the final presentation.