

# Leeds Status Report Feb 7-Feb 14 2021

Weekly status report for Feb 7-Feb 14 2021

Scope	Schedule	Cost	Risks	Quality
-------	----------	------	-------	---------

## Key Performance Indicators

- **Schedule:** **ON SCHEDULE**
  - Schedule Variance: 0 Days
  - Percent complete : **65%**
- **Labor:** **ON SCHEDULE**
  - Labor hours: 12 hours/person
- **Administration:** **COMPLETE**
  - Sponsor meeting, group meeting, attendance, hour log

## Summary

This week we continued to work on Moody's Manual. Each person, had their own take on how to process and create the text blobs. Some chose to use the supplied text data, and some decided to use their own models to see if their was better performance that way. Other than that, we also created a script to load all of the boundary data our sponsor had into our database so we didn't have to work with the clumsy text files.

## Work planned for this week

- Text blobs with > 1 column, CAP-75
- Hopefully continue work with CAP-72 given sponsor's agreement

## Work completed this week

### Text blobs with >1 column

✔ We found out where the mastercolumns file was kept and how to work with it. This gave us the information on if there was multiple columns per page in order to structure the output correctly.

📁 It is hard to decide which of the 99 outputs is best to keep. This is the next challenge to overcome

### Upload data from cluster to DB

✔ We wrote a script that would upload all of the word data from the cluster to the database. This makes it so that you don't have to interact with all the cumbersome files

⚠ Each page has 99 different boundary box readings. We did not want to upload all 99 versions so we chose to upload all data from the 51st brightness level. This might not be optimal for every page

## TV Corpora Analysis

✖ This task is blocked for right now as our sponsor wants us to focus on Moody's.

## Plans for next week

- Choose optimal text blob, specifically picking 1 of the 99 outputs to work with CAP-76

## Open Issues

### Deliverables and Milestones for sprint (10/6 to 10/20)

Deliverable or Milestone	WBS	Planned	Forecasted	Actual	Status
TV Corpora Analysis	CAP-72	2/14/21	2/28/21		Blocked
Text blobs > 1 column	CAP-75	2/14/21	2/14/21		Complete
Choose optimal text blob	CAP-76	2/28/21	2/28/21		In progress

## Open Change Requests

Change Request Name	Change Request Number	Requested Date	Current Status