CS410 Progress report

Netids:
td2 (Tianli Ding)
yimengh2 (Yimeng Han)
yusenw2 (Yusen Wang)

Progress made:
We were setting up the environment to make EducationalWeb work on our end and reading the source code and project structures. We have also proposed a file to parse a whole pdf file into single files. We are planning to write a scraping script to automatically download slides from several course websites and auto integrate those slides with the current system.

Remaining task:
We are planning to use beautifulsoup and simple URL requests to auto download slides from several UIUC courses' websites to scale up the current system

Challenges faced:
This is a huge system with lots of folders and files. Therefore we still need more time to understand and make more improvements.