

Project Proposal  
Darren Muliawan – darrenm2 (captain)

For the final project, I choose to work on improving the ExpertSearch System, focusing on automatically crawling faculty webpages. The ExpertSearch's data right now is coming from the MP2 submissions of the previous course offering. I am planning to improve the system by creating a service that keeps on monitoring the faculty homepages from the MP2.1 signup sheets to look for a new faculty members that was added after the MP2 submissions of the previous course offering. To show that my implementation works better, I need to show that there are actually new faculty members that is not included in the ExpertSearch's dataset. This crawler will be a separate system that will update the dataset that ExpertSearch uses. I am going to use Python to complete this project. The main tasks that need to be completed for this project includes,

1. Automatic crawling for the faculty homepages
2. Ability to detect changes of faculty members
3. Updating the dataset that ExpertSearch uses

These tasks should be able to be completed in ~20 hours.