

The tasks that has been completed:

- Given a url, it can determine whether the webpage is a faculty webpage
 - This is done by checking various attributes on the webpage, and then if the attributes match up, then the page is a faculty page
- The urls for the bios of the faculty on the webpage are saved to a txt file

The tasks that are still pending:

- Making this work on a large variety of webpages, as it is currently only tested on 1 webpage
 - General functionality exists, so just adapting to other types of pages and checking for other kinds of attributes will allow this to be done on other pages
- Ensuring that all bio urls on a page are found, and that none of the urls are links to pages that are not a faculty bio
- Write tests to verify the correctness of the results

A challenge we are facing is the broad range of webpages that exist and all of the different ways we have to account for them.