Structures and Communication

Crawler

- Starting URL
- sitemap [] Page: array of pages
- Num_pages: # pages collected
- Num_workers: # spawned workers
- Print limit: # assets per page to print
- Max_pages: #pages to crawl
- Max time: seconds to crawl
- Filename: filename to output sitemap



