Project Documentation

SoftwareCide-Squad

December 13, 2016

1 Requirements

1. Get data from HTML pages

2. Display it in the form of pages on the screen

3. Display some or all attributes of gathered data, depending on user preference

4. Ability to drag pages across the screen

5. Option to group sets of pages together

6. Option to remove unwanted pages from the screen

7. Option to save to other file types (e.g. excel spreadsheet)

2 Specifications

1. Program shall get specific types of HTML pages (ebay searches, Course participants, Tutees, Module Catalog).

2. Program shall read specific data from said pages.

3. Program shall display the data in the form of one page per entry within a window.

4. Pages shall be draggable across the screen, as a real person would move real pieces of paper on a real desk.

5. Program shall allow the user to toggle whether each attribute of the entries is hidden or shown on the screen.

6. Program shall allow the user to group pages together such that they can all be moved at once.

7. Program shall allow the user to discard chosen pages.

8. Program shall allow users to export those entries which are still being displayed to other file formats, such as excel spreadsheets.

9. Program shall not allow the user to edit the data.

3 Sequence Diagram

3.1 Web scrapers

```
GUI          Document Loader          Link Scraper          Page Scraper

 │  1: Passes search paramters  │                              │
 ├─────────────────────────────►│  1.1: Passes document of search result page
 │                              ├─────────────────────────►│
 │                              │  1.2: Returns URLs for all results
 │                              │◄ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ┤
 │                              │  1.3: Passes documents for each result page
 │                              ├──────────────────────────────────────────►│
 │  1.3.1: Returns DataDoc's for every search result
 │◄ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ┤
```