GROUP 8

NLP FOR DECOLONIZING HARM REDUCTION LITERATURE

DATA SCRAPING INTEGRATION TESTS

Evaluator: David James Gallenberger
Date: 4/12/2025
Branch: Task-363

1. URL Navigation Test Cases

Test Case #	Testing For	Procedure to Execute	Pass	Fail
1.1	Valid URL	1. Open the application. 2. Navigate to the Scrape page. 3. Enter "https://www.google.com/" into the "URL of Web Page" text bar. 4. Click on the "Submit" button. Expected result: a new pane opens, displaying the Google homepage.	✓	
1.2	Malformed URL	1. Open the application. 2. Navigate to the Scrape page. 3. Enter "http://./localhost:8008/"into the "URL of Web Page" text bar. 4. Click on the "Submit" button. Expected result: No new pane opens, and an error message appears.	✓	
1.3	Inactive URL	1. Open the application. 2. Navigate to the Scrape page. 3. Enter "http://localhost:8008/" into the "URL of Web Page" text bar. 4. Click on the "Submit" button. Expected result: No new pane opens, and an error message appears.		✓
1.4	No URL	1. Open the application. 2. Navigate to the Scrape page. 3. Click on the "Submit" button. Expected result: No new pane opens, and an error message appears.	✓	

Comments:

1.3: On submitting a valid but inactive URL, the application returns to the home page with no user-discernable explanation.

2. Manual Scrape Functionality Test Cases

Test Case #	Testing For	Procedure to Execute	Pass	Fail
2.1	Import Button Enable	1. Open the application. 2. Navigate to the Scrape page. 3. Enter "https://www.dhmo.org/" into the "URL of Web Page" text bar. 4. Click on the "Submit" button. 5. In the new window, confirm that the Selection Mode is set to "Manual" (if it is set to "Auto", click the radio button to toggle to "Manual" mode). 6. Highlight the paragraph beneath "Welcome" in the web page. Expected result: The paragraph should become highlighted for import, and the "Import Selected Data" button should become enabled.	✓	
2.2	Import Button Disable	1. Open the application. 2. Navigate to the Scrape page. 3. Enter "https://www.dhmo.org/" into the "URL of Web Page" text bar. 4. Click on the "Submit" button. 5. In the new window, confirm that the Selection Mode is set to "Manual" (if it is set to "Auto", click the radio button to toggle to "Manual" mode). 6. Highlight the paragraph beneath "Welcome" in the web page. 7. Click on the highlighted text (or anywhere within the web page that isn't a link) to deselect the text. Expected result: the "Import Selected Data" button should become disabled again.	✓	
3	Import Button Functionality	1. Open the application. 2. Navigate to the Scrape page. 3. Enter "https://www.dhmo.org/" into the "URL of Web Page" text bar. 4. Click on the "Submit" button. 5. In the new window, confirm that the Selection Mode is set to "Manual" (if it is set to "Auto", click the radio button to toggle to "Manual" mode). 6. Highlight the paragraph beneath "Welcome" in the web page. 7. Click "Import Selected Data". 8. Click "Export". Expected result: the window should close, and the highlighted text should be imported into the main application.	✓	

Comments:

3. Automatic Scrape Functionality Test Cases

Test Case #	Testing For	Procedure to Execute	Pass	Fail
3.1	Hotkey Changes	1. Open the application. 2. Navigate to the Scrape page. 3. Enter "https://www.dhmo.org/" into the "URL of Web Page" text bar. 4. Click on the "Submit" button. 5. In the new window, confirm that the Selection Mode is set to "Auto" (if it is set to "Manual", click the radio button to toggle to "Auto" mode). 6. Click the gear icon to the left of the Hotkey Indicator at the top of the window. 7. Press the "1" key. 8. Press and hold the "1" key. 9. While holding down the "1" key, click on the paragraph below "Welcome" in the web page. Expected result: The hotkey for selecting elements should change to "1", and the paragraph should be highlighted for import.	✓	
3.2	Import button enable when selecting text	1. Open the application. 2. Navigate to the Scrape page. 3. Enter "https://www.dhmo.org/" into the "URL of Web Page" text bar. 4. Click on the "Submit" button. 5. In the new window, confirm that the Selection Mode is set to "Auto" (if it is set to "Manual", click the radio button to toggle to "Auto" mode). 6. Hold down the import hotkey denoted by the Hotkey Indicator at the top of the window (by default, this is "Alt"). 7. While holding down the import hotkey, click on the paragraph below "Welcome" in the web page. Expected result: The paragraph should become highlighted for import, the "Import As Single Entry" button should become enabled, and the "Import All Separately" button should become enabled.	✓	
3.3	Import button disable when deselecting text	1. Open the application. 2. Navigate to the Scrape page. 3. Enter "https://www.dhmo.org/" into the "URL of Web Page" text bar. 4. Click on the "Submit" button. 5. In the new window, confirm that the Selection Mode is set to "Auto" (if it is set to "Manual", click the radio button to toggle to "Auto" mode). 6. Hold down the import hotkey denoted by the Hotkey Indicator at the top of the window (by default, this is "Alt"). 7. While holding down the import hotkey, click on the paragraph below "Welcome" in the web page. 8. While holding down the import hotkey, click on the paragraph below "Welcome" in the web page. Expected result: The paragraph should cease being highlighted for import, the "Import As Single Entry" button should become disabled, and the "Import All Separately" button should become disabled.	✓	
3.4	Single entry importing	1. Open the application. 2. Navigate to the Scrape page. 3. Enter "https://www.dhmo.org/" into the "URL of Web Page" text bar. 4. Click on the "Submit" button. 5. In the new window, confirm that the Selection Mode is set to "Auto" (if it is set to "Manual", click the radio button to toggle to "Auto" mode). 6. Hold down the import hotkey denoted by the Hotkey Indicator at the top of the window (by default, this is "Alt").	✓	

		7. While holding down the import hotkey, click on the paragraph below "Welcome" in the web page. 8. While holding down the import hotkey, click on the second paragraph below "Welcome" in the web page (beginning with "Explore our many Special Reports"). 9. Click on the "Import as Single Entry" button. 10. Click on the "Export" button. Expected result: the following text should be imported into the application as a single entry: Welcome to the web site for the Dihydrogen Monoxide Research Division (DMRD), currently located in Newark, Delaware. The controversy surrounding dihydrogen monoxide has never been more widely debated, and the goal of this site is to provide an unbiased data clearinghouse and a forum for public discussion.Explore our many Special Reports, including the DHMO FAQ, a definitive primer on the subject, plus reports on the environment, cancer, current research, and an insider exposé about the use of DHMO in the dairy industry.		
3.5	Separate entry importing	1. Open the application. 2. Navigate to the Scrape page. 3. Enter "https://www.dhmo.org/" into the "URL of Web Page" text bar. 4. Click on the "Submit" button. 5. In the new window, confirm that the Selection Mode is set to "Auto" (if it is set to "Manual", click the radio button to toggle to "Auto" mode). 6. Hold down the import hotkey denoted by the Hotkey Indicator at the top of the window (by default, this is "Alt"). 7. While holding down the import hotkey, click on the paragraph below "Welcome" in the web page. 8. While holding down the import hotkey, click on the second paragraph below "Welcome" in the web page (beginning with "Explore our many Special Reports"). 9. Click on the "Import All Separately" button. 10. Click on the "Export" button. Expected result: the following text should be imported into the application as two entries: Welcome to the web site for the Dihydrogen Monoxide Research Division (DMRD), currently located in Newark, Delaware. The controversy surrounding dihydrogen monoxide has never been more widely debated, and the goal of this site is to provide an unbiased data clearinghouse and a forum for public discussion. Explore our many Special Reports, including the DHMO FAQ, a definitive primer on the subject, plus reports on the environment, cancer, current research, and an insider exposé about the use of DHMO in the dairy industry.	✓	

Comments:

4. Exporting of Scraped Data to the linked Label Studio Project

Test Case #	Testing For	Procedure to Execute	Pass	Fail
5.1	Linkage with Label Studio	1. Open the application. 2. Navigate to the Annotation page. 3. Input [valid Label Studio URL] into the "URL of Label Studio Project" text bar. 4. Click on the "Submit" button. Expected result:		

Comments:
 Label studio not connected in this branch.