

Website Input - Bug #1615

Some content is not carried over from browser through proxy

11/22/2016 08:11 AM - Luke Murphey

| | | | |
|---------------------------------------|----------------|------------------------|------------|
| Status: | Closed | Start date: | 11/22/2016 |
| Priority: | Normal | Due date: | |
| Assignee: | Luke Murphey | % Done: | 100% |
| Category: | Web Controller | Estimated time: | 0.00 hour |
| Target version: | 4.0 | | |
| Description | | | |
| Things such as style tags are dropped | | | |

History

#1 - 11/22/2016 08:22 AM - Luke Murphey

Should try loading the content and seeing if document.documentElement.outerHTML includes the same data.

If not, check out other options such as

<http://stackoverflow.com/questions/7263824/get-html-source-of-webelement-in-selenium-webdriver-using-python>.

#2 - 11/23/2016 08:25 AM - Luke Murphey

The Firefox document does seem to have the expected content.

The next step is to either:

1. Execute the cleaner code against the FireFox HTML and confirm it is the cause
2. Tell load_page not to clean the HTML and see if this fixes it

#3 - 11/23/2016 06:00 PM - Luke Murphey

This is an issue with lxml. See <http://stackoverflow.com/questions/27807792/gae-lxml-cleaning-html-but-not-style-attribute> and <http://stackoverflow.com/questions/22750556/the-purpose-behind-lxml-htm-clean-to-remove-style-from-tags>

#4 - 11/23/2016 08:13 PM - Luke Murphey

- Category set to Web Controller
- Status changed from New to Closed
- Assignee set to Luke Murphey
- % Done changed from 0 to 100