

During the demo session with Ajarn, several issues were observed regarding the project outcome. Specifically, some websites could not be scraped for information. After a thorough review, it appears that the problem was not related to the extraction or parsing of content. Instead, it seems the problems are caused by a few different factors:

1. **Dynamic Content Loading:** Many modern websites, including IEEE Xplore, utilize JavaScript to load content dynamically after the initial page load. As Jsoup does not execute JavaScript, it is unable to access content that is loaded via JavaScript. Consequently, the information I was trying to scrape was loaded in this manner, so Jsoup would not detect it.
2. **Anti-Scraping Measures:** Websites often implement various anti-scraping mechanisms, such as requiring specific user agents, detecting and blocking bots, or employing CAPTCHAs. If the website identifies that requests are coming from a bot or script, it may return errors or block access. It is possible that IEEE Xplore has such measures in place, and I should have verified if the site requires a specific user agent or if other anti-scraping mechanisms are being used I would have not added it to my test case.
3. **Access Restrictions:** Some websites necessitate user authentication or subscriptions to access their content. If the website imposes restrictions on viewing content without proper credentials, my scraper would be unable to fetch the information. For instance, during the demo session, the article Ajarn asked me to visit was restricted to members, requiring sign-in details to fully access it:  
<https://ieeexplore.ieee.org/abstract/document/7724353>

I have reviewed the code and test files to troubleshoot these issues and have taken appropriate measures by implementing new test cases. These new test cases don't have the same problems as before.