**Code Review**

File Name: web.py  
Version Name: 511e2b7  
Location: team01-Project/src/web.py

Documentation:Good and extensive use of documentation - makes it easier to navigate through the code. As a documentation suggestion, it would be helpful to include what would happen if there were no URL’s found. This would make the initial summary complete.

Line 38 has the excessive ‘a’  
Lines 37-38 need more detail.  
Line 54 shifts the tense from present to past.   
Line 69 ought to specify that it is the recently added article.  
Line 102 is missing the “if” in “crawl the page, if it is not already…”  
Line 111 should include “article” instead of “one” for more clarity.

Coding Style:  
Clear coding style.   
Good use of variables.   
The code is well structured.

Correctness:  
Through inspection, the code seems to do the correct thing.

Testing  
There are only a few tests provided in test\_web\_crawler – testing different depths. Furthermore, only cnn.com seems to be tested and no testing related to adding references between articles and sources were included (despite crawl seemingly having this functionality).

Suggestions:

1. More tests can be provided to test different websites – the format of websites could vary by a large amount throughout different sites. Both local/national and international websites could be tested.
2. Specific types of invalid urls could be provided. Perhaps a page with just a video or sites that contain no links.
3. Tests involving references need to be added.

Summary

Add more detail to some comments. For example, lines 29 and 37. In the case of line 37, you can perhaps include the format of the publish date.