

Getting 1st Paragraph Only of Article

Posted by dougb - 2010/05/15 06:28

Hi Carty,

I'm only getting back the 1st paragraph of the article when using the "Proprietary Methods" option for grabbing feeds. Example

<http://superschoolonline.com/adult-dating-tips-try-these-out.html>

and

<http://www.children-and-divorce.net/children-and-divorce-blog/dating-after-divorce-women/mature-women-divorce-and-dating-again>

above only returned the 1st paragraph from the feed, should i be using a different option?

Thanks,
Doug

=====

Re: Getting 1st Paragraph Only of Article

Posted by Carty - 2010/05/15 09:46

In some websites 'Proprietary' method of extraction won't pull the full content and in these cases, you use the second method 'Between a known location' to extract full content.

For the Super School website the right setting would be, content is between:

and

Try it and tell me how it goes.

Regards,
Carty..

=====

Re: Getting 1st Paragraph Only of Article

Posted by dougb - 2010/05/15 16:36

Hi Carty,

Thanks very much for your reply. However, , i can only specify the 'Between a known location' for the entire feed which has close to a hundred postings. The "div" tags you you mention (thanks) are only applicable for this single posting, how would I set the 'between a known location' to grab content for for all others?

I guess what i'm concerned about moving forward is; the software is supposed to be able to grab the entire article, which in this case (the text) is all together with no page breaks, extra tags, etc. I'm really hoping this works, without alot of tweaking, so an entire article can be grabbed, I've found 2 on a quick review of the feeds, that were not grabbed fully, is something not working as designed?

Regards,
DougB

=====

Re:Getting 1st Paragraph Only of Article

Posted by Carty - 2010/05/16 15:15

No, the software will of course grab the

and other tags like responsible for paragraphing options so the result is not just a cramped up article.

Yes, the setting is only for the website you mentioned and it changes for every other website. You can easily find this setting by looking through the source code of the webpage. Maybe I should create a detailed video explaining how to get this setup properly.

Regards,
Carty..

=====

Re:Getting 1st Paragraph Only of Article

Posted by dougb - 2010/05/16 21:22

Hi Carty,

Thanks, I understand the software is complex, but for below article the text is simply that, just text and html tags;

Mature women, divorce and dating again&

Divorce for babyboomers is not uncommon. Many of these relationships stayed together while the children were living in the house, but now that they are grown and have families of their own, many babyboomer married couples decide that they want to explore life in opposite directions. This choice does not take away from the fact that you raised a great family and that you had a successful marriage, it just means that you have grown apart and now you are both heading in directions that are guided by different life priorities.

One of the challenges of returning the dating pool in your later years is that you have to deal with adult children of divorce. While you may think that younger children would take your divorce and re-entry into the dating world harder than adult children, this really is not the case. Your adult children have spent more time living with the reality that you and their father were a couple. You will need to talk with your adult children to explain that dating is what you need to be happy right now and that you need their support.

Once family matters have been taken care of you will be ready to face the dating world as a newly single woman. If it has been several decades since you last dated then you may be surprised at how things have changed. Women are more aggressive now in the dating world, but this does not mean that you have to pretend to be anything that you are not. Just relax, be yourself and do not be afraid to initiate a conversation with someone that you find interesting.

This entry was posted...

You said "the software will of course grab the

and other tags like ..." so If I understand correctly, in this case A-B-S should be grabbing the entire article, without having to be specifically told what tags to look for,as these are just

tags. In fact, A-B-S seems to be only grabbing the 1st paragraph, could this be a bug?

I wanted to make sure of my facts, so here's the feed in case you'd like to check:

<http://www.children-and-divorce.net/children-and-divorce-blog/feed>

Thanks and Regards,
Doug

Re:Getting 1st Paragraph Only of Article

Posted by Carty - 2010/05/17 05:32

I saw that this website is having some

codes in capital and the software ignored it. In my next version I will fix this issue so it grabs the entire article in these cases.

Meanwhile, you can try the second option "Between a known location" for this feed. Content is hidden between:

and

This entry was posted on

Regards,
Carty..

=====

Re:Getting 1st Paragraph Only of Article

Posted by dougb - 2010/05/17 13:58

Hi Carty,

Really glad to hear we know why this is not grabbing. Changing the "Between a known location" tags is limited because all the posts are different in the feeds I'm grabbing, so, it would only work for a very small number, time consuming, and not address long term moving forward.

Is there any way , please, I can fix (in the code) myself or you can forward me a 'fix' to grab the capitol characters (any any thing else that may be causing the article not to be grabbed)?

Otherwise, will I have to wait very long for the next release? Is there some way you can test (new version) on the 2 feeds I mentioned in this thread so I'm sure this will get addressed (for benefit of all) ?

www.children-and-divorce.net/children-and-divorce-blog/feed
<http://superschoolonline.com/feed>

*note: the exact post(s) may have changed.

Best Regards,
DougB

=====

Re:Getting 1st Paragraph Only of Article

Posted by Carty - 2010/05/17 21:40

Dear DougB,

Ok, I will try to get it done asap. Meanwhile, if you learn how to use the "Between a known location" method, it can never get any easier. Let me give you a short trick.

Use fire fox. Install add-on called FireBug.

Open the webpage in FireFox and right click at the start of the content in that page.

From the context menu or the right click menu that appears, choose "Inspect Element" option.

This will open FireBug and show you the code in the page source. From there you can easily find the surrounding tags.

Try it out and tell me if you find it helpful.

Regards,
Carty..

Regards,
Carty..

=====

Re:Getting 1st Paragraph Only of Article

Posted by dougb - 2010/05/17 23:04

Hi Carty,

Thanks, as a developer i'm familiar and use the DOM Inspector in FireFox, thanks.

I really appreciate you looking at the code to fix, in the mean time I will consider your suggestion but most likely inspect each post and 'eyeball' the results.

Regards,
DougB

=====

Re:Getting 1st Paragraph Only of Article

Posted by Carty - 2010/05/18 07:16

Thanks :)

Regards,
Carty..

=====