

StocksAndFisheriesKB - Bug #12420

Narrative text is messed up when in multiple xml nodes

Sep 07, 2018 05:04 PM - Aureliano Gentile

Status:	Closed	Start date:	Sep 07, 2018
Priority:	Normal	Due date:	
Assignee:	Yannis Marketakis	% Done:	100%
Category:		Estimated time:	0.00 hour
Target version:	GRSF		
Description			
The text for scientific advice and for status and trends are messed up or put in separate boxes. See this example http://data.d4science.org/ctlg/GRSF_Admin/a45df585-b8c0-3d33-99a6-36c07a02e918 and the source file http://firms.fao.org/fishery/xml/resource/10484/158284/en			
Related issues:			
Blocks StocksAndFisheriesKB - Bug #12906: Update "problematic" records in the...			Closed Nov 20, 2018

History

#1 - Nov 16, 2018 01:40 PM - Yannis Marketakis

This issue happens because the sequence of the narrative texts was not preserved (it was not defined to do so). FORTH will investigate if it is do-able without updating the mappings and re-constructing the GRSF KB.

In addition, the value for state and trend should be merged (since they have the same source, owner, ref year)

#2 - Nov 20, 2018 09:33 AM - Yannis Marketakis

- Blocks Bug #12906: Update "problematic" records in the GRSF admin catalog added

#3 - Nov 20, 2018 01:58 PM - Yannis Marketakis

- Status changed from New to Closed

- Assignee changed from Nikos Minadakis to Yannis Marketakis

- % Done changed from 0 to 100

The issue has been resolved. The narratives for the fields scientific_advice and state_and_trend are concatenated while respecting their order in the original sources. The concatenated narratives are separated using a semicolon (;). @aureliano.gentile@fao.org if you would like to change this character with something else, please let us know.

As a proof of concept, we've updated a legacy record (http://data.d4science.org/ctlg/GRSF_Admin/3a37e42b-c19e-316f-a5dc-35a6cf1f98b0) and a GRSF one (http://data.d4science.org/ctlg/GRSF_Admin/a45df585-b8c0-3d33-99a6-36c07a02e918), however, we will update all the records as soon as we resolved all the issues ([#12906](#)).

#4 - Nov 20, 2018 03:18 PM - Aureliano Gentile

Thanks a lot!

Actually multiple text nodes are generated by the FIRMS Word-to-XML converter tool for handling different paragraphs, so a semicolon is not needed. A paragraph would make it since the proper punctuation is already there. Therefore, is it possible to have a paragraph delimiter between the two blocks? (Or line breaks if easier. Or simply a space if nothing else is doable). I see that also formatting is ingored (e.g. bold, italics..) but I understand this is for a further development later on.

#5 - Nov 20, 2018 03:44 PM - Yannis Marketakis

- Status changed from Closed to Feedback

- % Done changed from 100 to 90

I think that we cannot use HTML formatting tags (i.e.
, <i>, , etc.) for the properties of the catalog. @francesco.mangiacrapa@isti.cnr.it is that correct?

If we cannot use HTML formatting tags, I will simply put a whitespace to differentiate narratives.

#6 - Nov 20, 2018 04:44 PM - Francesco Mangiacrapa

Yannis Marketakis wrote:

I think that we cannot use HTML formatting tags (i.e.
, <i>, , etc.) for the properties of the catalog. @francesco.mangiacrapa@isti.cnr.it is that correct?

Yes, It's correct. You may use simply the " " (whitespace) to separate them or by adding the ";" (semicolon and whitespace).

#7 - Nov 20, 2018 04:47 PM - Aureliano Gentile

It is (wrongly) correct, I would say ;-)

Anyway, so the best would be to add simply a space.

With thanks

#8 - Nov 21, 2018 11:38 AM - Yannis Marketakis

Thanks for your reply @francesco.mangiacrapa@isti.cnr.it .
So, I will replace ';' with whitespace.

#9 - Nov 21, 2018 11:50 AM - Yannis Marketakis

- *Status changed from Feedback to Closed*
- *% Done changed from 90 to 100*

Files

screenshot-i-marine.d4science.org-2018.09.07-16-56-06.png	187 KB	Sep 07, 2018	Aureliano Gentile
---	--------	--------------	-------------------