Morpho - Bug #1622

problems with morpho's eml parser

07/07/2004 11:00 AM - Margaret O'Brien

Status:	Resolved		**Start date:**	07/07/2004
Priority:	Immediate		**Due date:**	
Assignee:	Saurabh Garg		**% Done:**	0%
Category:	morpho - general		**Estimated time:**	0.00 hour
Target version:	1.6		**Spent time:**	0.00 hour
Bugzilla-Id:	1622			

Description

A file that is valid according to the eml parser at
http://knb.ecoinformatics.org/emlparser/index.html failed to pass the morpho
parser. Morpho cant tell me what the problem is, only that it cant save to the
network due to some bad eml. There were no apparent problems loading the file.
Also, the only place that morpho gave any eml error message is when I try to
save changes from the editor. There is no indication of a problem from the
save->network action, except that the icon in the upper right still says
"local", instead of changing to "network". The offending file is available at
the url above.

History

#1 - 07/13/2004 12:04 PM - Saurabh Garg

Hi Margaret, I tried inserting the eml document into metacat. The error that I
am getting is that Morpho is not able to insert the data files (sbclter.380 and
sbclter.415) correctly.
It would be great if you could post a sample of these data files and the error
log that is in the install directory. Thanks, Sid

#2 - 07/13/2004 01:00 PM - Margaret O'Brien

Hi Sid-
I put the two data files and the error log in with the metadata at
http://sbc.lternet.edu/external/EML/badEML/

BTW - I had this happen a few more times yesterday: first, after I got 9
packages into metacat via morpho, but it bombed on the 10th. I'll put that
package in the same directory. Then later in the day, on another computer (pc,
morpho1.5.1) another user (uid=mardian) downloaded several of those 9 successful
packages, and got the same error message every time he exited morpho's editor.
We didnt bother trying to save his edits back to metacat. I'll get ahold of this
second error log when I can.

#3 - 07/15/2004 12:09 PM - Saurabh Garg

Margaret, the thing causing trouble in eml doc #379 is the empty tag in
method/methodStep/description/section/para. When you edit it using the editor
it deletes the section/para part as it is empty. But it doesnt remove
description tag as it is a required tag. The bug is that the editor doesnt put
either section or para back in the new document as according to eml2 one of
them is also required.

If this is what is causing the trouble, you can try fixing it by putting some
description in the description tags.

#4 - 07/22/2004 05:21 PM - Margaret O'Brien

Sid-
You needed an example of morpho creating some bad eml. There are two files at
the url above:
379.26 valid input to morpho
379.27 morpho output after "tree" editor, no longer valid.

379.26 is in metacat (data.piscoweb.org), and passes the parser at knb. Scott
coombs (uid=scoombs1; he has "all" permission) picked it up with morpho, made a

small change to an address and tried to save it. 379.27 is the result. Both
morpho and the knb parser said this file was invalid. Here is the error message
from knb's parser:

Document is EML valid.
There were no EML errors found in your document.

XML-Schema Errors Found
The following errors were found:
cvc-complex-type.2.4.b: The content of element 'creator' is not complete. It
must match
'((((((((("":individualName)|("":organizationName))|("":positionName)){1-UNBOUNDED},("":address){0-UNBOUNDED}),("":phone){0-UNBOUNDED}),(""
:electronicMailAddress){0-UNBOUNDED}),("":onlineUrl){0-UNBOUNDED}),("":userId){0-UNBOUNDED})|("":references))'.

#5 - 07/22/2004 06:28 PM - Saurabh Garg

Margaret, thanks for taking the time to put the bad eml documents online. The
trouble with these is same as I mentioned in my precious comment:

" The bug is that the editor doesnt put either section or para back in the new
document as according to eml2 one of them is also required."

We will try to get this bug in the next release. Till then if you want to fix
the document number 379 - you can add the following tags in
methods/methodStep/protocol/creator
<individualName>
</individualName>

If you want to avoid this problem for the next time, a better way might be to
go into the editor - scroll down to this field and just put a '.' in there.
That way editor wont delete the tags the next time. If you want to edit with
using morpho, you can add
<individualName>
<surName.</surname>
</individualName>

The trouble here is empty tags in the document when it is opened in Morpho.
Morpho doesnt know whether you have removed the values in the editor or they
were initially not there. So it goes into the eml document and removes all the
tags which are empty - including the schema required ones. This is the bug that
we will try to fix.

#6 - 07/23/2004 08:54 AM - Matt Jones

And Sid, in fixing this I think what we should be doing is running the eml
parser on the document when a Save event is detected. If the validation fails,
we should at a minimum provide an understandable list of the places where more
metadata is needed, but even better guide them to those locations so that they
can fix them quickly and easily.

#7 - 06/20/2005 12:38 PM - Saurabh Garg

Comment from a recent commit done to fix this bug:
The changes made in this bug result in the following:
1. The user is notified if there is a field which cant be trimmed because it
is required and text has not been provided for that field.
2. The Trim button is faster as some of the redundant computation is removed
now.

```
The above changes should fix some of the problems that users face while using
the tree editor - for example, using the trim and show all button results in
invalid eml documents.

On the other hand, tree editor code is complex and the above changes might
not have fully fixed the problems. More testing is required.
```

More comments:
Not running the eml parser yet. Though The above bugfix does fix the problem
which Margaret mentioned above.

#8 - 09/28/2005 12:42 PM - Saurabh Garg

Problems with the fix of this bug is mentioned in bug# 2200. Also added another
bug# 2209 based on comment# 6 from Matt. Closing this bug.

#9 - 03/27/2013 02:17 PM - Redmine Admin

Original Bugzilla ID was 1622