

EDIT - feature request #9114

Handle malformed URIs in user type

06/30/2020 04:23 PM - Andreas Müller

Status:	Closed	Start date:	
Priority:	Priority14	Due date:	
Assignee:	Katja Luther	% Done:	70%
Category:	cdmlib	Estimated time:	0:00 hour
Target version:	Release 5.30		
Severity:	normal		
Description			
The fix for #9111 should be moved to a new URI user type wrapper that is used instead of the current URI class. Otherwise the workaround needs to be applied wherever URIs are used.			
Related issues:			
Related to EDIT - feature request #9111: Handle Malformed URIs		Rejected	

Associated revisions

Revision 26790e9b - 01/05/2021 07:45 PM - Andreas Müller

ref #9114 adapt URI to wrapper in cdmlib-app

Revision de8249cf - 01/05/2021 08:05 PM - Andreas Müller

ref #9114 add URI Wrapper and replace all java.net.URI occurrences by the wrapper in cdmlib

Revision 999006a7 - 01/05/2021 08:08 PM - Andreas Müller

ref #9114 preliminary ignore RdfViewTest

Revision 4cd4de2d - 01/05/2021 08:13 PM - Andreas Müller

ref #9114 adapt URIs to URI wrapper in vaadin

Revision dd94c99b - 01/05/2021 08:16 PM - Andreas Müller

ref #9114 adapt URI to wrapper in cdmlib-app (cont.)

Revision bf3cb991 - 01/05/2021 08:34 PM - Andreas Müller

ref #9114 adapt some classes with URI to wrapper in taxeditor

Revision 96378f2b - 01/05/2021 09:02 PM - Andreas Müller

ref #9114 adapt some classes with URI to wrapper in taxeditor (cont.)

Revision 5b2e9e6b - 01/05/2021 09:07 PM - Andreas Müller

ref #9114 add compareTo to URI wrapper

Revision f5a55599 - 01/05/2021 09:26 PM - Andreas Müller

ref #9114 adapt some classes with URI to wrapper in taxeditor (cont.)

Revision 5dfe9f9b - 01/05/2021 09:37 PM - Andreas Müller

ref #9114 adapt some classes with URI to wrapper in taxeditor (cont.)

Revision 15c5de38 - 01/05/2021 09:49 PM - Andreas Müller

ref #9114 adapt some classes with URI to wrapper in taxeditor (cont.)

Revision 7b06bd8f - 01/06/2021 12:34 AM - Andreas Müller

ref #9114 adapt some classes with URI to wrapper in taxeditor (cont.)

Revision 5acb8103 - 01/06/2021 01:30 AM - Andreas Müller

ref #9114 adapt some classes with URI to wrapper in taxeditor (cont.)

Revision 4d87641f - 01/06/2021 01:35 AM - Andreas Müller

ref #9114 add toFile() to URI to wrapper

Revision 6b33224a - 01/06/2021 10:17 AM - Andreas Müller

ref #9114 adapt javadoc for new URI wrapper user type

Revision f9c3f295 - 01/06/2021 10:18 AM - Andreas Müller

ref #9114 adapt json mapping to new URI wrapper

Revision 3154d43f - 01/06/2021 10:50 AM - Andreas Müller

ref #9114 further adapt some classes to URI wrapper in taxeditor

Revision 4ad992ed - 01/06/2021 11:00 AM - Andreas Müller

ref #9114 further adapt jsonConfigurations and add new constructor to URI wrapper

Revision c22065e1 - 01/06/2021 01:00 PM - Andreas Müller

ref #9114 fix RdfViewTest

Revision 9916da4a - 01/06/2021 02:18 PM - Andreas Müller

ref #9111, ref #9114 remove TaxEditor specific URI encoding code as encoding is handled in cdmLib now

Revision ac58825c - 02/04/2021 04:53 PM - Andreas Müller

ref #9114 fix dozer mapping by reverting URI mappings to java.net.URI

Revision a8dd847c - 01/24/2022 03:52 PM - Katja Luther

ref #9114: implement encoding of fragment part of uri

Revision 1dcdb672 - 02/07/2022 11:00 AM - Katja Luther

ref #9114: adapt parsing of fragment and add test

Revision 09da5a1b - 03/10/2022 02:35 PM - Katja Luther

fix #9114: add testcase with more than one fragment

History

#1 - 06/30/2020 04:23 PM - Andreas Müller

- *Related to feature request #9111: Handle Malformed URIs added*

#2 - 01/05/2021 06:01 PM - Andreas Müller

- *Description updated*

- *% Done changed from 0 to 40*

#3 - 01/06/2021 12:08 PM - Andreas Müller

- *Status changed from New to In Progress*

I implemented a new class eu.etaxonomy.cdm.common.URI and use this in model but also at other places like import sources.

Unfortunately, simply subclassing java.net.URI was not possible as it is final.

A method getJavaUri() exists to transform the above class to standard java.net.URI.

Currently there are still 2 open issues:

- httpinvoker: within httpinvoker there is still a ClassCastException and therefore TaxEditor fails to build. First I thought this is because CdmModelCache needs to be build anew, but running the script does not change the model cache file so this does not seem to be the reason
- RdfViewTest.testMarshalRdf does not run successfully and therefore I set it to ignore. The children type [dc:identifier/](#) do not seem to be build correctly.

#4 - 01/06/2021 02:15 PM - Andreas Müller

- Assignee changed from Andreas Müller to Katja Luther
- Target version changed from Release 5.18 to Release 5.19

Katja, can you please check if you can find out what the httpinvoker problem is. I currently can't start the TaxEditor from the IDE for debugging.

#5 - 01/06/2021 02:27 PM - Andreas Müller

In [#9111#note-1](#) AK suggests to also encode fragment and query parts. This has **not** been implemented in the original code in [#9111](#) which has been copied to cdmlib.

I think we should also implement it for fragment and query. Probably not for authority and scheme?

Also the handling of backslash needs to be discussed. In the original version backslash was encoded, too. This is probably not wanted, at least not for the path part. Currently a URISyntaxException is thrown if path contains a backslash.

#6 - 01/27/2021 04:48 PM - Andreas Müller

- Target version changed from Release 5.19 to Release 5.21

#7 - 03/11/2021 08:51 AM - Katja Luther

- Target version changed from Release 5.21 to Release 5.22

#8 - 03/18/2021 06:46 PM - Andreas Müller

- Tags set to fast

This ticket seems to be done or needs to be splitted into an open issues ticket. It can be moved back to 5.19(?) afterwards.

#9 - 04/19/2021 02:40 PM - Katja Luther

- Status changed from In Progress to Feedback
- Assignee changed from Katja Luther to Andreas Müller

Andreas Müller wrote:

In [#9111#note-1](#) AK suggests to also encode fragment and query parts. This has **not** been implemented in the original code in [#9111](#) which has been copied to cdmlib.

I think we should also implement it for fragment and query. Probably not for authority and scheme?

Also the handling of backslash needs to be discussed. In the original version backslash was encoded, too. This is probably not wanted, at least not for the path part. Currently a URISyntaxException is thrown if path contains a backslash.

Did you implement this already?

#10 - 04/19/2021 11:35 PM - Andreas Müller

- Assignee changed from Andreas Müller to Katja Luther

Katja Luther wrote:

Andreas Müller wrote:

In [#9111#note-1](#) AK suggests to also encode fragment and query parts. This has **not** been implemented in the original code in [#9111](#) which has been copied to cdmlib.

I think we should also implement it for fragment and query. Probably not for authority and scheme?

Also the handling of backslash needs to be discussed. In the original version backslash was encoded, too. This is probably not wanted, at least not for the path part. Currently a URISyntaxException is thrown if path contains a backslash.

Did you implement this already?

No, needs to be finished

#11 - 04/20/2021 12:43 PM - Katja Luther

- Assignee changed from Katja Luther to Andreas Müller

Andreas Müller wrote:

Katja Luther wrote:

Andreas Müller wrote:

In [#9111#note-1](#) AK suggests to also encode fragment and query parts. This has **not** been implemented in the original code in [#9111](#) which has been copied to cdmlib.

I think we should also implement it for fragment and query. Probably not for authority and scheme?

Also the handling of backslash needs to be discussed. In the original version backslash was encoded, too. This is probably not wanted, at least not for the path part. Currently a URISyntaxException is thrown if path contains a backslash.

Did you implement this already?

No, needs to be finished

Should I do this? Because you worked on this ticket.

#12 - 04/20/2021 01:33 PM - Andreas Müller

- Assignee changed from Andreas Müller to Katja Luther

I only moved the original algorithm to the user type. The remaining issues are about the algorithm. As you implemented the algorithm I thought you will do this.

Or at least give hints why e.g. fragments and query parts were not yet encoded and why backslash was encoded. I didn't want to change the algorithm without feedback on these issues.

Maybe also Andreas K. can give feedback on this.

#13 - 07/09/2021 02:37 PM - Andreas Müller

- Target version changed from Release 5.22 to Release 5.25

This still needs feedback so I move to current milestone. Once it is fixed we can move to 5.19 where most of the work was done or keep it in current milestone.

#14 - 07/16/2021 02:53 PM - Andreas Müller

- Target version changed from Release 5.25 to Release 5.45

#15 - 01/24/2022 03:59 PM - Katja Luther

- Assignee changed from Katja Luther to Andreas Kohlbecker

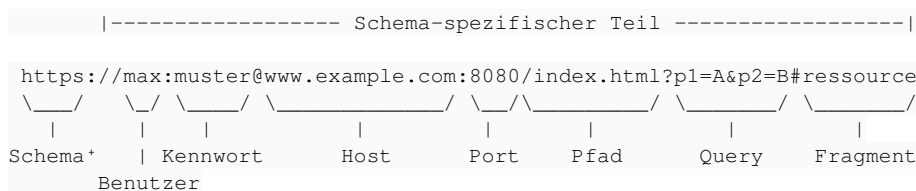
Implemented encoding of fragment part, query part was already implemented. @AK could you have a look and give feedback whether this implementation is ok?

#16 - 01/27/2022 03:14 PM - Andreas Kohlbecker

- Assignee changed from Andreas Kohlbecker to Katja Luther

According to the URI and URL Schemes defined in the corresponding RFCs:

URI = scheme ":" hier-part ["?" query] ["#" fragment]



There is always only one fragment. So it makes not much sense to handle multiple fragments as in our code. A better algorithm would be to split the path+query+fragment string at the first # character and to treat the

right part as fragment.

BTW: There is no test method in eu.etaxonomy.cdm.common.UriTest to test URI strings with fragment.

#17 - 02/07/2022 11:49 AM - Katja Luther

- Status changed from Feedback to Resolved
- Assignee changed from Katja Luther to Andreas Kohlbecker

Thanks for your comment, I adapted the code and added a test.

Please add a testcase with multiple fragments, like <https://max:muster@www.example.com:8080/index.html?p1=A&p2=B#ressource&#another-frag?>

#18 - 02/22/2022 03:44 PM - Andreas Kohlbecker

- Status changed from Resolved to Feedback
- Assignee changed from Andreas Kohlbecker to Katja Luther

#19 - 03/10/2022 02:35 PM - Katja Luther

- Status changed from Feedback to Resolved
- % Done changed from 40 to 70

Applied in changeset [cdm|lib|09da5a1bd580baae6490435bc52ddd6d2aab353](#).

#20 - 03/10/2022 02:36 PM - Katja Luther

- Status changed from Resolved to Closed

Added a testcase for more than one fragment, the rest was already reviewed, so we can close this ticket.

#21 - 03/10/2022 02:40 PM - Andreas Müller

- Target version changed from Release 5.45 to Release 5.30

This was fixed in 5.30 already