

EDIT - bug #9855

Specimen images sometimes show duplicates

11/08/2021 10:07 AM - Andreas Müller

Status:	Closed	Start date:	
Priority:	Highest	Due date:	
Assignee:	Andreas Müller	% Done:	100%
Category:	cdmlib	Estimated time:	0:00 hour
Target version:	Release 5.29	Found in Version:	
Severity:	normal		
Description			
e.g. https://campanula.e-taxonomy.net/cdm_dataportal/taxon/cdcacbdba-11c0-4cb8-a2e0-8d7abc5ae756/images			

Associated revisions

Revision ece0f9f5 - 11/08/2021 04:58 PM - Andreas Müller

fix #9855 deduplicate images in TaxonServiceImpl.listMedia()

Revision c61b7547 - 11/08/2021 04:58 PM - Andreas Müller

ref #9855 remove image duplicates for MediaSpecimen in ABCD import

Revision f767e961 - 01/20/2022 10:30 PM - Andreas Müller

ref #9855 replace deduplication by in CdmUtils by stream.distinct()

History

#1 - 11/08/2021 11:02 AM - Andreas Müller

AK:

eu.etaxonomy.cdm.remote.controller.iiif.ManifestController.doTaxonMedia()

#2 - 11/08/2021 11:18 AM - Andreas Müller

- Status changed from New to In Progress

Duplication only happens in UniversalViewer (using the ManifestController.doTaxonMedia()) code. Switching to simple viewer does not show duplicates.

#3 - 11/08/2021 11:41 AM - Andreas Kohlbecker

Andreas Müller wrote:

Duplication only happens in UniversalViewer (using the ManifestController.doTaxonMedia()) code. Switching to simple viewer does not show duplicates.

So this clearly related to eu.etaxonomy.cdm.remote.controller.iiif.ManifestController.doTaxonMedia()

#4 - 11/08/2021 04:59 PM - Andreas Müller

- Status changed from In Progress to Resolved

- % Done changed from 0 to 50

Applied in changeset [cdmlib|75539fa3b6e75f1972f2c3e3f90a3b395ea9d356](https://campanula.e-taxonomy.net/cdm_dataportal/taxon/cdm|75539fa3b6e75f1972f2c3e3f90a3b395ea9d356).

#5 - 11/08/2021 05:11 PM - Andreas Müller

- % Done changed from 50 to 0

The reason for the duplicates was that in TaxonServiceImpl.listMedia() media were not deduplicate if e.g. an image of a MediaSpecimen was also attached via image gallery to an occurrence. This is dirty data but happens due to a bug in ABCD import.

I added deduplication if duplicated images exist for whatever reason. Additionally I fixed the ABCD bug to avoid such media duplicates in future.

#6 - 11/08/2021 05:12 PM - Andreas Müller

- Assignee changed from Andreas Müller to Andreas Kohlbecker

- % Done changed from 0 to 50

AK can you please review if the result is as expected.

KL can you please review the fix for ABCD.

#7 - 01/20/2022 03:43 PM - Andreas Kohlbecker

- Assignee changed from Andreas Kohlbecker to Andreas Müller

For de-duplicating the Media in TaxonService.listMedia() I suggest to use a more modern java code style:

```
taxonMedia = taxonMedia.stream().distinct().collect(Collectors.toList());
```

this way the CdmUtils.removeIdentical() method can be removed - it is used only once in the library.

#8 - 01/20/2022 10:31 PM - Andreas Müller

- Assignee changed from Andreas Müller to Andreas Kohlbecker

Andreas Kohlbecker wrote in [#note-7](#):

For de-duplicating the Media in TaxonService.listMedia() I suggest to use a more modern java code style:

```
taxonMedia = taxonMedia.stream().distinct().collect(Collectors.toList());
```

this way the CdmUtils.removeIdentical() method can be removed - it is used only once in the library.

Good point. I adapted the code accordingly. Can you check again and then pass the ticket to KL?

#9 - 01/27/2022 02:42 PM - Andreas Kohlbecker

- Assignee changed from Andreas Kohlbecker to Katja Luther

- % Done changed from 50 to 70

Excellent!

@Katja: I's your turn now.

#10 - 01/27/2022 03:01 PM - Katja Luther

- Assignee changed from Katja Luther to Andreas Müller

The ABCD Import code is correct, too. So I think we can close this ticket.

#11 - 01/27/2022 08:08 PM - Andreas Müller

- Status changed from Resolved to Closed

- % Done changed from 70 to 100

#12 - 02/15/2022 03:40 PM - Andreas Müller

- Target version changed from Release 5.45 to Release 5.29