EDIT - bug #8533

missing collector and fieldnumber in abcd import

09/10/2019 11:28 AM - Katja Luther

Status:	Closed	**Start date:**	
Priority:	Highest	**Due date:**	
Assignee:	Katja Luther	**% Done:**	100%
Category:	cdmadapter	**Estimated time:**	0:00 hour
Target version:	Release 5.11		
Severity:	critical	**Found in Version:**	

Description

when importing specimen/dna as associated units where the gathering event information is not complete in one of the derived units the already existing value should not be overwritten by a null value.

Related issues:

Related to EDIT - feature request #8549: while abcd import problems should be...	**New**

Associated revisions

Revision a3a55498 - 09/10/2019 12:05 PM - Katja Luther

ref #8533: check for all fields in gatheringEvent whether it is set already

Revision 9b1e9a85 - 09/19/2019 11:41 AM - Katja Luther

ref #8533: minor

History

#1 - 09/17/2019 10:56 AM - Katja Luther

- Status changed from New to Resolved

- Assignee changed from Katja Luther to Patrick Plitzner

problems like this should be handled more in detail and the user should be informed about this problems. Therefore I create a new ticket and close this one (#8549).

Please review whether the problem of the missing information is fixed.

#2 - 09/17/2019 10:56 AM - Katja Luther

- % Done changed from 0 to 60

#3 - 09/17/2019 11:03 AM - Katja Luther

- Related to feature request #8549: while abcd import problems should be logged more in detail added

#4 - 09/18/2019 03:20 PM - Patrick Plitzner

- Status changed from Resolved to Feedback

- Assignee changed from Patrick Plitzner to Katja Luther

I attached the two sample file: one is the DNA unit the other one the associated tissue sample.

I tested on nightly rem_conf_am and searched for DB 40073 in the specimen import search. It imported the DnaSample and the TissueSample correctly but it did neither import the author (Hutchison) nor the collecting number (1980) which are both available in the tissue sample abcd.

#5 - 09/18/2019 03:22 PM - Patrick Plitzner

- File dna_sample.xml added

- File tissue_sample.xml added

#6 - 09/18/2019 04:34 PM - Katja Luther

- File picture953-1.png added

the collector problem is solved, but the fieldnumber is not delivered. I debugged and also looked at the output and there is no fieldnumber (see the attached file).

#7 - 09/18/2019 04:37 PM - Katja Luther

ah there is another problem, the gathering agent for the tissue is Hutchison and for the tissue sample Lautner

#8 - 09/18/2019 04:43 PM - Patrick Plitzner

Katja Luther wrote:

> the collector problem is solved, but the fieldnumber is not delivered. I debugged and also looked at the output and there is no fieldnumber (see the attached file).

The field number is in the tissue sample file (*CollectorsFieldNumber*)

```
<abcd21:Gathering>
              <abcd21:DateTime>
                <abcd21:DateText>18.08.1957</abcd21:DateText>
              </abcd21:DateTime>
              <abcd21:Agents>
                <abcd21:GatheringAgentsText>Hutchison</abcd21:GatheringAgentsText>
              </abcd21:Agents>
              <abcd21:LocalityText>
Junín. Prov. Tarma, Chanchamayotal zwischen San Ramón und La Merced an vertikalen Felsen oberhalb der Straße.
</abcd21:LocalityText>
              <abcd21:Country>
                <abcd21:Name>Peru</abcd21:Name>
                <abcd21:ISO3166Code>PE</abcd21:ISO3166Code>
              </abcd21:Country>
              <abcd21:NamedAreas>
                <abcd21:NamedArea>
                  <abcd21:AreaClass language="en">Continent</abcd21:AreaClass>
                  <abcd21:AreaName language="en">Middle and South America</abcd21:AreaName>
                </abcd21:NamedArea>
              </abcd21:NamedAreas>
              <abcd21:Notes language="en" />
            </abcd21:Gathering>
            <abcd21:CollectorsFieldNumber>1980</abcd21:CollectorsFieldNumber>
            <abcd21:RecordURI>http://herbarium.bgbm.org/object/BGT0023581</abcd21:RecordURI>
```

#9 - 09/18/2019 04:44 PM - Patrick Plitzner

Katja Luther wrote:

> ah there is another problem, the gathering agent for the tissue is Hutchison and for the tissue sample Lautner

Hmm, I guess we have to move this to another issue for discussion

#10 - 09/18/2019 04:46 PM - Katja Luther

Patrick Plitzner wrote:

> Katja Luther wrote:
>
> > the collector problem is solved, but the fieldnumber is not delivered. I debugged and also looked at the output and there is no fieldnumber (see the attached file).
>
> The field number is in the tissue sample file (*CollectorsFieldNumber*)
>
> ```
> <abcd21:Gathering>
>               <abcd21:DateTime>
>                 <abcd21:DateText>18.08.1957</abcd21:DateText>
>               </abcd21:DateTime>
>               <abcd21:Agents>
>                 <abcd21:GatheringAgentsText>Hutchison</abcd21:GatheringAgentsText>
>               </abcd21:Agents>
>               <abcd21:LocalityText>
> Junín. Prov. Tarma, Chanchamayotal zwischen San Ramón und La Merced an vertikalen Felsen oberhalb der Stra
> ße.</abcd21:LocalityText>
>               <abcd21:Country>
>                 <abcd21:Name>Peru</abcd21:Name>
> ```

```
            <abcd21:ISO3166Code>PE</abcd21:ISO3166Code>
          </abcd21:Country>
          <abcd21:NamedAreas>
            <abcd21:NamedArea>
              <abcd21:AreaClass language="en">Continent</abcd21:AreaClass>
              <abcd21:AreaName language="en">Middle and South America</abcd21:AreaName>
            </abcd21:NamedArea>
          </abcd21:NamedAreas>
          <abcd21:Notes language="en" />
        </abcd21:Gathering>
        <abcd21:CollectorsFieldNumber>1980</abcd21:CollectorsFieldNumber>
        <abcd21:RecordURI>http://herbarium.bgbm.org/object/BGT0023581</abcd21:RecordURI>
```

I know, but if you have a look at the file attached, there is only "Gathering" and then "RecordURI", no CollectorsFieldNumber

#11 - 09/18/2019 04:47 PM - Katja Luther

- Status changed from Feedback to In Progress

- Target version changed from Release 5.10 to Release 5.11

I think we have to move this to 5.11

#12 - 09/18/2019 05:04 PM - Patrick Plitzner

- Severity changed from normal to critical

ok, this seems to be a bigger problem.

If you compare your debug output with the two files it seems like it has mixed the information together.

For example

```
            <abcd21:AreaName language="en">North America</abcd21:AreaName>
```

comes from the DnaSample and

```
        <abcd21:RecordURI>http://herbarium.bgbm.org/object/BGT0023581</abcd21:RecordURI>
```

comes from the tissue sample.

Maybe the Abcd206XMLFieldGetter does not reset its state correctly when handling a new unit and instead just merges all the information it can get together.

#13 - 09/19/2019 09:42 AM - Katja Luther

- File picture715-1.png added

Patrick Plitzner wrote:

> ok, this seems to be a bigger problem.
>
> If you compare your debug output with the two files it seems like it has mixed the information together.
>
> For example
>
> ```
>             <abcd21:AreaName language="en">North America</abcd21:AreaName>
> ```
>
> comes from the DnaSample and
>
> ```
>         <abcd21:RecordURI>http://herbarium.bgbm.org/object/BGT0023581</abcd21:RecordURI>
> ```
>
> comes from the tissue sample.
>
> Maybe the Abcd206XMLFieldGetter does not reset its state correctly when handling a new unit and instead just merges all the information it can get together.

I checked the two files again and the tissue sample has UnitID B GT 0023581 and the associated unit of the dnaSample is B GT 0023612, so they do not belong together.
and the field unit of B GT 0023612 has no fieldNumber:

```
      <abcd21:Unit>
```

```
            <abcd21:UnitGUID>http://herbarium.bgbm.org/object/BGT0023612</abcd21:UnitGUID>
            <abcd21:SourceInstitutionID>B</abcd21:SourceInstitutionID>
            <abcd21:SourceID>Herbarium Berolinense</abcd21:SourceID>
            <abcd21:UnitID>B GT 0023612</abcd21:UnitID>
            <abcd21:Identifications>
              <abcd21:Identification>
                <abcd21:Result>
                  <abcd21:TaxonIdentified>
                    <abcd21:HigherTaxa>
                      <abcd21:HigherTaxon>
                        <abcd21:HigherTaxonName>CACTACEAE</abcd21:HigherTaxonName>
                        <abcd21:HigherTaxonRank>familia</abcd21:HigherTaxonRank>
                      </abcd21:HigherTaxon>
                    </abcd21:HigherTaxa>
                    <abcd21:ScientificName>
                      <abcd21:FullScientificNameString>Selenicereus dorschianus Ralf Bauer
</abcd21:FullScientificNameString>
                      <abcd21:NameAtomised>
                        <abcd21:Botanical>
                          <abcd21:GenusOrMonomial>Selenicereus</abcd21:GenusOrMonomial>
                          <abcd21:FirstEpithet>dorschianus</abcd21:FirstEpithet>
                          <abcd21:AuthorTeam>Ralf Bauer</abcd21:AuthorTeam>
                        </abcd21:Botanical>
                      </abcd21:NameAtomised>
                    </abcd21:ScientificName>
                  </abcd21:TaxonIdentified>
                </abcd21:Result>
                <abcd21:PreferredFlag>1</abcd21:PreferredFlag>
              </abcd21:Identification>
            </abcd21:Identifications>
            <abcd21:RecordBasis>PreservedSpecimen</abcd21:RecordBasis>
            <abcd21:KindOfUnit>tissue</abcd21:KindOfUnit>
            <abcd21:Gathering>
              <abcd21:DateTime>
                <abcd21:DateText>04.02.2001</abcd21:DateText>
              </abcd21:DateTime>
              <abcd21:Agents>
                <abcd21:GatheringAgentsText>Lautner</abcd21:GatheringAgentsText>
              </abcd21:Agents>
              <abcd21:LocalityText>
Jalisco. zwischen Conception de Buenas Aires und Union de Guadelupe an der Jal 410.</abcd21:LocalityText>
              <abcd21:Country>
                <abcd21:Name>Mexico</abcd21:Name>
                <abcd21:ISO3166Code>MX</abcd21:ISO3166Code>
              </abcd21:Country>
              <abcd21:NamedAreas>
                <abcd21:NamedArea>
                  <abcd21:AreaClass language="en">Continent</abcd21:AreaClass>
                  <abcd21:AreaName language="en">North America</abcd21:AreaName>
                </abcd21:NamedArea>
              </abcd21:NamedAreas>
              <abcd21:Altitude>
                <abcd21:MeasurementOrFactAtomised>
                  <abcd21:LowerValue>2150 m</abcd21:LowerValue>
                </abcd21:MeasurementOrFactAtomised>
              </abcd21:Altitude>
              <abcd21:Notes language="en" />
            </abcd21:Gathering>
            <abcd21:RecordURI>http://herbarium.bgbm.org/object/BGT0023612</abcd21:RecordURI>
          </abcd21:Unit>
```

#14 - 09/19/2019 09:45 AM - Katja Luther

- File picture607-1.png added

the imported field unit for B GT 0023581 has collector and fieldNumber see attached file (picture607-1.png)

#15 - 09/19/2019 09:45 AM - Katja Luther

- File deleted (picture715-1.png)

#16 - 09/30/2019 01:24 PM - Katja Luther

- Status changed from In Progress to Resolved

- Assignee changed from Katja Luther to Patrick Plitzner

as the used example was about two different field units I close this ticket. Please make a last review for closing the ticket.

#17 - 09/30/2019 01:41 PM - Patrick Plitzner

- File tissue_sample.xml added

#18 - 09/30/2019 01:43 PM - Patrick Plitzner

- File deleted (tissue_sample.xml)

#19 - 09/30/2019 01:47 PM - Patrick Plitzner

I uploaded the correct tissue sample file and will test again.

#20 - 11/05/2019 01:24 PM - Patrick Plitzner

- Target version changed from Release 5.11 to Release 5.12

#21 - 11/05/2019 03:12 PM - Andreas Müller

- Target version changed from Release 5.12 to Release 5.11

Why did you move it to 5.12. It has status resolved and should therefore be reviewed in 5.11 milestone. If there is a reason for moving please mention it.

#22 - 11/07/2019 09:58 AM - Patrick Plitzner

- Status changed from Resolved to Closed

- Assignee changed from Patrick Plitzner to Katja Luther

- % Done changed from 60 to 100

> Why did you move it to 5.12. It has status resolved and should therefore be reviewed in 5.11 milestone. If there is a reason for moving please mention it.

Sorry, I overlooked the content of this ticket.

Tested the import again and the FieldUnit data is now correctly imported.

Files

dna_sample.xml	16.1 KB	09/18/2019	Patrick Plitzner
picture953-1.png	32.9 KB	09/18/2019	Katja Luther
picture607-1.png	44.8 KB	09/19/2019	Katja Luther
tissue_sample.xml	12.2 KB	09/30/2019	Patrick Plitzner