

Test - Task #311

Phaidra - Primo

2012-05-21 09:03 - Rastislav Hudak

Status:	Closed	Start date:	2012-05-21
Priority:	Normal	Due date:	
Assignee:	Rastislav Hudak	% Done:	0%
Category:		Estimated time:	0.00 hour
Target version:			
Description			

History

#1 - 2012-05-21 17:39 - Rastislav Hudak

Dear Ms. Putz,

On 05/16/2012 05:55 PM, Michaela Putz wrote:

Liebe Koll*

Wir haben gestern den file splitter in Primo getestet und dafür ein OAI-Exportfile von 2010 verwendet. Weil sich seither aber vermutlich bei den Feldern in Phaidra etwas geändert hat, eine Bitte: könntet ihr uns aktuelle Daten zum Testen zur Verfügung stellen?

- möglichst alle Objekttypen
- jeweils ein Set mit Titeln, die auch in Aleph vorkommen (mit AC-Nummer) und eines mit Titeln ohne AC-Nummer

attached is a text file showing the difference between our old dublin core and a new one for an actual object.
I'll try to prepare a collection with new records containing all object types and I will send it to you.

Bei den Tests hat sich auch gezeigt, dass für das Indexieren des Volltexts der Pfad zum Volltext im XML-file vorkommen muss, weil er in Primo nicht dynamisch gebildet werden kann (einzig denkbare Transformationen s.u., leider kein Suffix möglich).

Soweit ich das sehe, müsste zB für das Dokument <http://phaidra.univie.ac.at/o:458> folgender Link vorhanden sein:
<https://fedora.phaidra.univie.ac.at/fedora/get/o:458/bdef:Content/download>

I think we can use something like <dc:source>CONTENT:https://fedora.phaidra.univie.ac.at/fedora/get/o:458/bdef:Content/download</dc:source> so that in the oai output you have a link to the file. However, this will return the binary file (pdf for pdf document, jpg for page and so on..), not the fulltext.

Thanks
Regards
Rastislav Hudak

--

Rastislav Hudak rastislav.hudak@univie.ac.at
Vienna University Computer Center Phone: +43-1-4277-140 84
Ebendorferstraße 10, A-1010 Wien, Austria Fax: +43-1-4277-143 38

primo.txt

```
#####  
OLD  
#####
```

```
<record xmlns="http://www.openarchives.org/OAI/2.0/">  
  <header>  
    <identifier>oai:univie.ac.at:o:19958</identifier>  
    <timestamp>2009-08-24T10:45:18Z</timestamp>  
  </header>  
  <metadata>
```

```
<oai_dc:dc xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.openarchives.o
```

rg/OAI/2.0/oai_dc/ http://www.openarchives.org/OAI/2.0/oai_dc.xsd" xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/">

```
<dc:identifier>o:19958</dc:identifier>
<dc:creator>Plinius Secundus, Gaius</dc:creator>
<dc:language>la</dc:language>
<dc:subject>ÖFOS 2002, NATURAL SCIENCES, Biology, Botany, Zoology, Zoology</dc:subject>
<dc:subject>ÖFOS 2002, NATURAL SCIENCES, Other and interdisciplinary Natural Sciences, History of natural sciences</dc:subject>
<dc:subject>ÖFOS 2002, NATURAL SCIENCES, Biology, Botany, Zoology, Biological anthropology</dc:subject>
<dc:subject>ÖFOS 2002, HUMAN MEDICINE, Pharmaceuticals, Pharmacology, Toxicology, Pharmaceuticals</dc:subject>
<dc:subject>ÖFOS 2002, NATURAL SCIENCES, Biology, Botany, Zoology, Botany</dc:subject>
<dc:subject>ÖFOS 2002, SOCIAL SCIENCES, Other and interdisciplinary Social Sciences, Ethnography</dc:subject>
>
<dc:subject>EuroVoc 4.2, AGRICULTURE, FORESTRY AND FISHERIES, agricultural activity, crop production, horticulture</dc:subject>
<dc:subject>ÖFOS 2002, NATURAL SCIENCES, Geography, Regional geography</dc:subject>
<dc:subject>EuroVoc 4.2, SCIENCE, natural and applied sciences, space science, astronomy, cosmology</dc:subject>
<dc:subject>ÖFOS 2002, HUMANITIES, Arts, Art history</dc:subject>
<dc:subject>ÖFOS 2002, HUMAN MEDICINE, Clinical Medicine (except Surgery and Psychiatry), Diagnosis in medicine</dc:subject>
<dc:subject>ÖFOS 2002, TECHNICAL SCIENCES, Mining, Metallurgy, Metallurgy</dc:subject>
<dc:subject>ÖFOS 2002, HUMAN MEDICINE, Medical Chemistry, Medical Physics, Physiology, General physiology</dc:subject>
<dc:subject>ÖFOS 2002, NATURAL SCIENCES, Geology, Mineralogy, Mineralogy</dc:subject>
<dc:subject>EuroVoc 4.2, SOCIAL QUESTIONS, culture and religion, arts, fine arts</dc:subject>
<dc:subject>Getty Thesaurus of Geographic Names, World, Roman Empire, Italia</dc:subject>
<dc:title>Historia naturalis: libri XXXVII</dc:title>
<dc:description>Plinius, Historia naturalis (1469) </dc:description>

</oai_dc:dc>
</metadata>
</record>
```


NEW
#####

```
<record>
  <header>
    <identifier>oai:univie.ac.at:o:19958</identifier>
    <timestamp>2012-05-21T14:01:21Z</timestamp>
  </header>
  <metadata>
```

```
<oai_dc:dc xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/ http://www.openarchives.org/OAI/2.0/oai_dc.xsd" xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/">
  <dc:language>lat</dc:language>
  <dc:creator>Plinius Secundus, G. (Gaius)</dc:creator>
  <dc:date>1469</dc:date>
  <dc:subject>ÖFOS, NATURAL SCIENCES, Other and interdisciplinary Natural Sciences, History of natural sciences</dc:subject>
  <dc:subject>EuroVoc 4.2, SCIENCE, natural and applied sciences</dc:subject>
  <dc:subject>Getty Thesaurus of Geographic Names, World, Roman Empire, Italia</dc:subject>
  <dc:subject>University of Vienna, University of Vienna, Vienna University Library, E-Books on Demand</dc:subject>
  <dc:subject>European Projects, Europeana Libraries</dc:subject>
  <dc:description>Gaius Plinius Secundus&apos; Plinius &quot;Historia Naturalis&quot; ist das älteste Buch im Bestand der Universitätsbibliothek Wien. Es handelt sich dabei um eine Inkunabel (Frühdrucke aus der Zeit 1450-1500) Plinius war ein römischer Offizier und Historiker und lebte von 23-79 n.Chr.</dc:description>
  <dc:identifier>http://phaidra.univie.ac.at/o:19958</dc:identifier>
  <dc:format>application/pdf</dc:format>
  <dc:title>Historia naturalis: libri XXXVII</dc:title>
  <dc:type>Book</dc:type>
</oai_dc:dc>
</metadata>
</record>
```

Note: The dc:type may also be <dc:type>info:eu-repo/semantics/Book</dc:type> if this value was defined in metadata. In case of o:19958 it was not, so the type Book is guessed based on content model.

#####

OpenAIRE document example
#####

```
<record>
  <header>
    <identifier>info:fedora/oai:univie.ac.at:o:1677</identifier>
    <datestamp>2012-05-21T14:24:14Z</datestamp>
    <setSpec>ec_fundedresources</setSpec>
  </header>
  <metadata>

<oai_dc:dc xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/ http://www.openarchives.org/OAI/2.0/oai_dc.xsd" xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/">
  <dc:rights>info:eu-repo/semantics/embargoedAccess</dc:rights>
  <dc:rights>http://creativecommons.org/licenses/by/2.0/at/legalcode</dc:rights>
  <dc:source>ISSN: 0090-2942</dc:source>
  <dc:source>The American journal of Chinese medicine 22(11)</dc:source>
  <dc:language>eng</dc:language>
  <dc:creator>lastname1, F. (firstname1)</dc:creator>
  <dc:creator>lastname2, F. (firstname2)</dc:creator>
  <dc:date>2012-05-01</dc:date>
  <dc:date>info:eu-repo/date/embargoEnd/2012-05-30</dc:date>
  <dc:subject>Dewey Decimal Classification, Literature & rhetoric, Literatures of Germanic languages, German letters</dc:subject>
  <dc:subject>European Projects, FP7, Openaire, Environment (including Climate Change)</dc:subject>
  <dc:subject>info:eu-repo/classification/ddc/836</dc:subject>
  <dc:subject>first keyword, second keyword, 3rd keyword</dc:subject>
  <dc:description>abstract of the paper</dc:description>
  <dc:identifier>http://phaidra-entw.univie.ac.at/o:1677</dc:identifier>
  <dc:format>application/pdf</dc:format>
  <dc:relation>info:eu-repo/grantAgreement/EC/FP7/gan1234</dc:relation>
  <dc:title>test OpenAIRE paper</dc:title>
  <dc:type>info:eu-repo/semantics/Book</dc:type>
  <dc:type>info:eu-repo/semantics/publishedVersion</dc:type>
  <dc:publisher>Vienna University Computer Center, Dienstleistungseinrichtungen, Stabstellen, etc., University of Vienna</dc:publisher>
</oai_dc:dc>
  </metadata>
</record>
```

#2 - 2012-05-24 15:15 - Rastislav Hudak

commandline/regenerateDCAllObjects.pl

run on entwickler. now I will run it on test and if everything works i can run it on prod over the weekend.

#3 - 2012-05-30 17:11 - Rastislav Hudak

- Status changed from In Progress to Feedback

- Priority changed from Urgent to High

Dear Ms. Putz,

we have updated the DC on all our objects in production so you can use the OAI provider directly to get the list of all objects (http://fedora.phaidra.univie.ac.at/oaiprovider/?verb=ListRecords&metadataPrefix=oai_dc) or to get a single object (http://fedora.phaidra.univie.ac.at/oaiprovider/?verb=GetRecord&metadataPrefix=oai_dc&identifier=oai%3Aunivie.ac.at%3Ao%3A95).

Is this enough for you or do you need an OAI set to be created to be available with '&set=xy' parameter?

We currently do not have objects with AC-Number filled in (except test objects). But even if we had, I don't think it will be available in dublin core.

Regards,
Rasta

#4 - 2012-06-06 17:22 - Rastislav Hudak

- File 20120601_PhaidraInPrimo_Protokoll.doc added

- Status changed from Feedback to In Progress

- Priority changed from High to Urgent

Dear Mr. Hudak,

Am 30.05.2012 16:35, schrieb Rastislav Hudak:

Dear Ms. Putz,

we have updated the DC on all our objects in production so you can use the OAI provider directly to get the list of all objects (http://fedora.phaidra.univie.ac.at/oaiprovider/?verb=ListRecords&metadataPrefix=oai_dc) or to get a single object (http://fedora.phaidra.univie.ac.at/oaiprovider/?verb=GetRecord&metadataPrefix=oai_dc&identifier=oai%3Aunivie.ac.at%3Ao%3A95).

Is this enough for you or do you need an OAI set to be created to be available with '&set=xy' parameter?

As you can see in the minutes attached, we decided to start our tests with items from the institutional repository (because they do not have an AC-number and therefore no duplicates in Primo). Could you please provide us with a set of documents from IR as mentioned above? Thanks!

Regards
Michaela

#5 - 2012-06-18 09:47 - Rastislav Hudak

- Status changed from *In Progress* to *Feedback*
- Priority changed from *Urgent* to *Normal*

Dear Ms. Putz,

the set of all documents in the test instance of institutional repository is in our Phaidra test instance phaidra-test.univie.ac.at (you should have access within the university). I have created a new set/collection called ir (https://phaidra-test.univie.ac.at/detail_object/o:4185). It is available via oaiprotider at

https://fedora.phaidra-test.univie.ac.at/oaiprotider/?verb=ListRecords&set=ir&metadataPrefix=oai_dc

It is a test set, some documents were imported more than once or contain invalid data.

Regards,
Rasta

#6 - 2012-11-28 11:39 - Rastislav Hudak

- Status changed from *Feedback* to *Closed*

no more feedback, closing

Files

20120601_PhaidraInPrimo_Protokoll.doc	119 KB	2012-06-06	Rastislav Hudak
---------------------------------------	--------	------------	-----------------