

StocksAndFisheriesKB - Task #11786

GRSF-services-core updates

May 17, 2018 10:42 AM - Yannis Marketakis

Status:	Closed	Start date:	May 17, 2018
Priority:	Normal	Due date:	
Assignee:	Yannis Marketakis	% Done:	100%
Category:		Estimated time:	0.00 hour
Target version:			
Description			
The following updates are required in the grsf-services-core component:			
<ul style="list-style-type: none">• apply the new types for Stocks: (Assessment Unit and Marine Resource) and Fisheries: (Fishing Unit and Other Fishery) (DONE)• create a new field called similarities_indicator that will contain the values with similar records / without similar records. (DONE)• create a new field called connections_indicator that will contain the values connected / not connected• Store the creation date (at least the year) of a legacy record and the series name (for the moment only for FIRMS records)• Construct the citation string and submit it while publishing a record			

History

#1 - May 17, 2018 10:42 AM - Yannis Marketakis

- Status changed from New to In Progress

#2 - May 21, 2018 10:56 AM - Yannis Marketakis

- % Done changed from 0 to 80

- create a new field called connections_indicator that will contain the values connected / not connected --> **DONE**
- Store the creation date (at least the year) of a legacy record and the series name (for the moment only for FIRMS records) --> **DONE**
- Construct the citation string and submit it while publishing a record --> **DONE**

#3 - May 21, 2018 12:49 PM - Yannis Marketakis

The aforementioned updates have been carried out. There are still the following pending issues, which I would like to check with CNR colleagues.

- We tried to improve the readability of species and water areas fields, by adding commas or dashes. These fields appear as tags and FAO colleagues mentioned that it is important to be readable. However, I noticed that for these cases we submit a string that contains the required separators which are removed when creating the tag. (i.e. check the species or the assessment area fields at <https://ckan-grsf-admin2.d4science.org/dataset/397db3fe-e27d-3d63-86c1-2eda3658db12> and the corresponding tags). Could these extra characters (i.e. commas and colons) be preserved in the tag?
- The citation string of a record is now submitted during publishing in the field named citation. The information that will be added from the catalog is the access date. When publishing it has the value [Date accessed and/or downloaded]. Is this OK? Do you want to change it to something else (so that you can find it and replace it with the actual access date)? Below you will find the citation of an indicative FIRMS record.

© FAO 2017. Striped marlin - Eastern Pacific. FIRMS Data Collection.
In: Fisheries and Resources Monitoring System (FIRMS) [online].
Updated: 2017-04-03. Accessed: [Date accessed and/or downloaded].
<http://firms.fao.org/firms/resource/10079/en>

- We have to build the URL of a record in the catalogue (for the citation of GRSF records). We know that the suffix of the URL is the UUID of the record. But what about the prefix? Is there a permanent URL prefix that we could use? I'm asking so because catalog URLs seem to be VRE-dependent (i.e. http://data.d4science.org/ctlg/GRSF_Admin/397db3fe-e27d-3d63-86c1-2eda3658db12)

@luca.frosini@isti.cnr.it could you take a look to the above questions ?

#4 - May 21, 2018 02:33 PM - Francesco Mangiacrapa

Hi Yannis,

I can reply to..

Yannis Marketakis wrote:

- We tried to improve the readability of species and water areas fields, by adding commas or dashes. These fields appear as tags and FAO colleagues mentioned that it is important to be readable. However, I noticed that for these cases we submit a string that contains the required separators which are removed when creating the tag. (i.e. check the species or the assessment area fields at <https://ckan-grsf-admin2.d4science.org/dataset/397db3fe-e27d-3d63-86c1-2eda3658db12> and the corresponding tags). Could these extra characters (i.e. commas and colons) be preserved in the tag?

Unfortunately, we can't preserve them. We have an engine "limitation" (that is CKAN-side). As you can see at:

http://docs.ckan.org/en/2.6/api/index.html#ckan.logic.action.create_tag_create

```
name (string) - the name for the new tag, a string between 2 and 100 characters long containing only alpha
numeric characters and -, _ and ., e.g. 'Jazz'
```

- We have to build the URL of a record in the catalogue (for the citation of GRSF records). We know that the suffix of the URL is the UUID of the record. But what about the prefix? Is there a permanent URL prefix that we could use? I'm asking so because catalog URLs seem to be VRE-dependent (i.e. http://data.d4science.org/ctlg/GRSF_Admin/397db3fe-e27d-3d63-86c1-2eda3658db12)

It works like a (Persistent Uniform Resource Locator) PURL that points to an intermediate resolution service (the Catalogue-Resolver). Instead of pointing directly to the location of the item (so the CKAN URL) We are using the resolution service (that is VRE-dependent because each item belongs to a VRE (where it was created)). The resolution service returns the URL to the client as a standard HTTP/s redirect.

#6 - May 21, 2018 05:32 PM - Luca Frosini

Hi @marketak@ics.forth.gr

it is ok from my side. I already added the support for 'citation' field in the code.

- The citation string of a record is now submitted during publishing in the field named citation. The information that will be added from the catalog is the access date. When publishing it has the value [Date accessed and/or downloaded]. Is this OK? Do you want to change it to something else (so that you can find it and replace it with the actual access date)? Below you will find the citation of an indicative FIRMS record.

[...]

@luca.frosini@isti.cnr.it could you take a look to the above questions ?

@aureliano.gentile@fao.org is it ok the text composed by @marketak@ics.forth.gr

@francesco.mangiacrapa@isti.cnr.it will take care of adding the [Date accessed and/or downloaded] on client side.

#7 - May 21, 2018 06:01 PM - Aureliano Gentile

the above example of SOURCE RECORD CITATION is fine with me.

#8 - May 22, 2018 08:12 AM - Yannis Marketakis

Thank you all for your answers.

The only pending issue, therefore, is the identification of the catalogue prefix.

When publishing a record, the publishing client is aware of the scope (the VRE), therefore I assume that I'll be able to retrieve the catalog prefix.

@francesco.mangiacrapa@isti.cnr.it and @luca.frosini@isti.cnr.it I assume that I'll be able to retrieve it using an ic-client, right ?

#9 - May 23, 2018 11:14 AM - Francesco Mangiacrapa

Yannis Marketakis wrote:

@francesco.mangiacrapa@isti.cnr.it and @luca.frosini@isti.cnr.it I assume that I'll be able to retrieve it using an ic-client, right ?

Yes, Yannis.

In order to build right prefix, you can retrieve the RuntimeResource "HTTP-URI-Resolver" to get the hostname of the URI-Resolver.

Using the scope of ROOT (/d4science.research-infrastructures.eu) or VRE (i.e /d4science.research-infrastructures.eu/FARM/GRSF_Admin), it is unique in the infrastructure:

```
<Resource version="0.4.x">

  <ID>9fe02de0-bc9e-11e2-85ae-b807e8ac0391</ID>

  <Type>RuntimeResource</Type>

  <Scopes>

    <Scope>/d4science.research-infrastructures.eu/FARM/GRSF_Admin</Scope>

    <Scope>/d4science.research-infrastructures.eu</Scope>

    ....

  <Category>Service</Category>

  <Name>HTTP-URI-Resolver</Name>

  ....

  <HostedOn>data.d4science.org</HostedOn>
```

in the resource you can read the property "HostedOn" that is "data.d4science.org". Then you must add the suffix "/ctlg"+"VRE_NAME".

Recap, you can build the right prefix in this way:

"<https://data.d4science.org>" + "/"ctlg" + "/VRE_NAME"

i.e.

This is a prefix for the GRSF_ADMIN catalogue: http://data.d4science.org/ctlg/GRF_ADMIN

This is a PURL for the BlueBridgeProject catalogue:

<http://data.d4science.org/ctlg/BlueBridgeProject/fao-aquatic-species-distribution-map-of-bathyrja-spinicauda>

and so on..

These links are generated first via HTTP Post.

You can see how at: https://wiki.gcube-system.org/gcube/URI_Resolver#CATALOGUE_Resolver. Unfortunately the wiki page is old and does not contain valid links :(I will update the page asap.

I hope this clarifies

#10 - May 25, 2018 11:33 AM - Yannis Marketakis

- Status changed from In Progress to Closed

Perfect Thanks

#11 - May 25, 2018 11:34 AM - Yannis Marketakis

- % Done changed from 80 to 100

#12 - Jul 26, 2018 02:08 PM - Luca Frosini

- Project changed from 8 to StocksAndFisheriesKB

- Category deleted (Default)

- Target version deleted (Unspritable)