

StocksAndFisheriesKB - Bug #18690

Links still pointing to BlueBRIDGE

Mar 02, 2020 05:55 PM - Aureliano Gentile

Status:	Closed	Start date:	Mar 02, 2020
Priority:	Normal	Due date:	
Assignee:	Luca Frosini	% Done:	100%
Category:		Estimated time:	0.00 hour
Target version:	GRSF		
Description			
As @anton.ellenbroek@fao.org noticed, links in exploiting fishery and similar records are still pointing to bluebridge, please see attached screenshot (from this record http://data.d4science.org/ctlg/GRSF/f3849a07-7c11-3206-bf5c-06199b703675)			

History

#1 - Mar 03, 2020 06:10 PM - Francesco Mangiacrapa

- Status changed from New to In Progress

#2 - Mar 04, 2020 10:20 AM - Francesco Mangiacrapa

- Status changed from In Progress to Feedback

- % Done changed from 0 to 100

Hi @aureliano.gentile@fao.org,

I just updated the configurations to resolve also the links of GRSF_ADMIN (e.g. this one http://data.d4science.org/ctlg/GRSF_Admin/5d609960-47b1-3d81-89c0-1373bc0fd4dc) towards the iMarine-Gateway. Let me know...

#3 - Mar 23, 2020 03:04 PM - Francesco Mangiacrapa

- Status changed from Feedback to Closed

#4 - Mar 23, 2020 03:24 PM - Aureliano Gentile

Dear Francesco, right this morning I was checking this ticket (my apologies for the delay). I confirm it is now fine in the GRSF Admin.

Regarding the public release <https://i-marine.d4science.org/web/grsf/data-catalogue> I made a search by "bluebridge" and I got this record (Makaira nigricans Atlantic)

<https://i-marine.d4science.org/web/grsf/data-catalogue?path=/dataset/e9d96077-6cb7-34c2-bf06-dc04e25dad89>

which indeed contains this url

https://bluebridge.d4science.org/group/grsf_admin/data-catalogue?path%3D%2Fdataset%2F7febfe97-2952-3392-bde2-dfc48d754ff5

not sure why...

#5 - Mar 24, 2020 12:17 PM - Francesco Mangiacrapa

- Status changed from Closed to In Progress

- % Done changed from 100 to 80

Hi Aureliano,

I'm adding @marketak@ics.forth.gr to this ticket.

Regarding the issue reported in your previous comment (at 18690#note-4), in my opinion during the publishing phase these links have been resolved on KB side from RECORD URL of kind: [http://data.d4science.org/ctlg/GRSF/\[UUID\]](http://data.d4science.org/ctlg/GRSF/[UUID]) to resolved link like: [https://bluebridge.d4science.org/group/grsf_admin/data-catalogue?path%3D%2Fdataset%2F\[UUID\]](https://bluebridge.d4science.org/group/grsf_admin/data-catalogue?path%3D%2Fdataset%2F[UUID]) and published into GRSF/GRSF_ADMIN Catalogue as link already resolved instead of RECORD URL.

Only the first type of link (the RECORD URL) must be stored into Catalogue not the second one (the link resolved).

The RECORD URL (e.g. <http://data.d4science.org/ctlg/GRSF/e9d96077-6cb7-34c2-bf06-dc04e25dad89>), is resolved dynamically by URI-Resolver

service at consuming-time (e.g. by pasting the link into browser). With URI-Resolver service we are able to resolve a RECORD URL (in general an ITEM URL published into a D4Science Catalogue) always pointing to the current gateway/catalogue that contains and provides the RECORD in the D4Science Infrastructure. In fact, just changing one configuration (on Information System-side) we have been able to switch from bluebridge.d4science.org to i-marine.d4science.org by using the (same) RECORD URLs already published in the GRSF/GRSF_ADMIN Catalogues.

Instead, the link resolved (e.g.

https://bluebridge.d4science.org/group/grsf_admin/data-catalogue?path%3D%2Fdataset%2F7febfe97-2952-3392-bde2-dfc48d754ff5) already points to destination and it is not dynamic.

Regarding the GRSF public release at <https://i-marine.d4science.org/web/grsf/data-catalogue>

Searching for "bluebridge.d4science" only 1 record is found <http://data.d4science.org/ctlq/GRSF/e9d96077-6cb7-34c2-bf06-dc04e25dad89>

and it contains a 'Similar GRSF Record' with url

https://bluebridge.d4science.org/group/grsf_admin/data-catalogue?path%3D%2Fdataset%2F7febfe97-2952-3392-bde2-dfc48d754ff5

Regarding the GRSF_ADMIN release at https://i-marine.d4science.org/group/grsf_admin/data-catalogue

Searching for "bluebridge.d4science" I got 27 records (you can see them at

https://i-marine.d4science.org/group/grsf_admin/data-catalogue?path=dataset&query=cT0iYmx1ZWJyaWRnZS5kNHNjaWVvY2Ui)

Yannis, please, could you republish them by fixing the RECORD URL?

#6 - Mar 24, 2020 02:57 PM - Yannis Marketakis

The URLs are not generated in the GRSF KB during the publishing. The actual workflow is:

1. We publish a record X (which does not have a URL yet)
2. As soon as it is published in the catalogue, the publishing client receives the catalog ID and the catalogue URL
3. The catalogue URL is added in GRSF KB
4. If necessary (e.g. if there are similar records) we update records so that their URLs are visible in the catalogue as well.

Now the reason for having URLs with two different prefixes is because some of them have been published a long time ago (when a particular prefix was used), while others are more recent (with the new prefix).

I guess I could update the prefixes of the URLs in the GRSF KB (manually) and update the affected records. Is that fine?

#7 - Mar 24, 2020 05:07 PM - Francesco Mangiacrapa

Yannis Marketakis wrote:

The URLs are not generated in the GRSF KB during the publishing. The actual workflow is:

1. We publish a record X (which does not have a URL yet)
2. As soon as it is published in the catalogue, the publishing client receives the catalog ID and the catalogue URL
3. The catalogue URL is added in GRSF KB
4. If necessary (e.g. if there are similar records) we update records so that their URLs are visible in the catalogue as well.

Many thanks @marketak@ics.forth.gr for clarifying the workflow.

Now the reason for having URLs with two different prefixes is because some of them have been published a long time ago (when a particular prefix was used), while others are more recent (with the new prefix).

I guess I could update the prefixes of the URLs in the GRSF KB (manually) and update the affected records. Is that fine?

If for updating the prefixes of the URLs in the GRSF KB (manually) ... you mean to use the RECORD URL of kind:

- for GRSF [http://data.d4science.org/ctlq/GRSF/\[THE_RECORD_UUID\]](http://data.d4science.org/ctlq/GRSF/[THE_RECORD_UUID]) (e.g. <http://data.d4science.org/ctlq/GRSF/e9d96077-6cb7-34c2-bf06-dc04e25dad89>)
- for GRSF_ADMIN [http://data.d4science.org/ctlq/GRSF_ADMIN/\[THE_RECORD_UUID\]](http://data.d4science.org/ctlq/GRSF_ADMIN/[THE_RECORD_UUID]) (e.g. http://data.d4science.org/ctlq/GRSF_Admin/255b8039-dfc8-3598-954e-2ad5fbf95f7e)

that's OK.

Otherwise, you can obtain programmatically the Record URL by performing a POST operation to <http://data.d4science.org/ctlq> (you can see how to at https://wiki.gcube-system.org/gcube/URI_Resolver#CATALOGUE_Resolver)

#8 - Mar 25, 2020 09:54 AM - Yannis Marketakis

Francesco Mangiacrapa wrote:

If for updating the prefixes of the URLs in the GRSF KB (manually) ... you mean to use the RECORD URL of kind:

- for GRSF [http://data.d4science.org/ctlq/GRSF/\[THE_RECORD_UUID\]](http://data.d4science.org/ctlq/GRSF/[THE_RECORD_UUID]) (e.g. <http://data.d4science.org/ctlq/GRSF/e9d96077-6cb7-34c2-bf06-dc04e25dad89>)
- for GRSF_ADMIN [http://data.d4science.org/ctlq/GRSF_ADMIN/\[THE_RECORD_UUID\]](http://data.d4science.org/ctlq/GRSF_ADMIN/[THE_RECORD_UUID]) (e.g.

http://data.d4science.org/ctlg/GRSF_Admin/255b8039-dfc8-3598-954e-2ad5bf95f7e

that's OK.

That's exactly what I mean. So, I'll proceed with this.

#9 - Mar 25, 2020 10:58 AM - Francesco Mangiacrapa

- Assignee changed from Francesco Mangiacrapa to Yannis Marketakis

Yannis Marketakis wrote:

Francesco Mangiacrapa wrote:

If for updating the prefixes of the URLs in the GRSF KB (manually) ... you mean to use the RECORD URL of kind:

- for GRSF [http://data.d4science.org/ctlg/GRSF/\[THE_RECORD_UUID\]](http://data.d4science.org/ctlg/GRSF/[THE_RECORD_UUID]) (e.g. <http://data.d4science.org/ctlg/GRSF/e9d96077-6cb7-34c2-bf06-dc04e25dad89>)
- for GRSF_ADMIN [http://data.d4science.org/ctlg/GRSF_ADMIN/\[THE_RECORD_UUID\]](http://data.d4science.org/ctlg/GRSF_ADMIN/[THE_RECORD_UUID]) (e.g. http://data.d4science.org/ctlg/GRSF_Admin/255b8039-dfc8-3598-954e-2ad5bf95f7e)

that's OK.

That's exactly what I mean. So, I'll proceed with this.

Many Thanks

#10 - Mar 26, 2020 08:06 AM - Yannis Marketakis

@francesco.mangiacrapa@isti.cnr.it I found 4 defected records using the old (bluebridge-like URI).

I've updated them manually, however I noticed that after updating them in the catalogue, GRSF Publisher service is still returning me the old URL (see log attached).

Is there any issue on the publisher service side? or should I just try to delete and re-publish the record (instead of updating it)?

#11 - Mar 26, 2020 09:33 AM - Yannis Marketakis

- File *updateLog.log* added

#12 - Mar 26, 2020 10:16 AM - Luca Frosini

- Assignee changed from Yannis Marketakis to Luca Frosini

#13 - Mar 26, 2020 11:08 AM - Luca Frosini

- Assignee changed from Luca Frosini to Francesco Mangiacrapa

I have found this exception:

```
2020-03-26 07:57:31,197 [Thread-40] ERROR JCRWorkspaceItem: Impossible to get short url. Long url will be returned.
2020-03-26 07:57:33,867 [Thread-52] WARN UrlShortener: Runtime Resource HTTP-URL-Shortener-DL not found
2020-03-26 07:57:33,867 [Thread-52] ERROR UrlShortener: An error occurred reading Runtime Resource for name: HTTP-URL-Shortener-DL, the scope is: /d4science.research-infrastructures.eu/FARM/GRSF_Admin
java.lang.Exception: No Runtime Resource with name: HTTP-URL-Shortener-DL is available in the scope: /d4science.research-infrastructures.eu/FARM/GRSF_Admin
    at org.gcube.portlets.user.urlshortener.UrlShortener.<init> (UrlShortener.java:97)
    at org.gcube.portlets.user.urlshortener.UrlShortener.<init> (UrlShortener.java:69)
    at org.gcube.common.homelibrary.jcr.workspace.JCRWorkspaceItem.getShortUrl (JCRWorkspaceItem.java:1189)
    at org.gcube.common.homelibrary.jcr.workspace.JCRWorkspaceItem.getPublicLink (JCRWorkspaceItem.java:1146)
    at org.gcube.data_catalogue.grsf_publish_ws.utils.csv.ManageTimeSeriesThread.manageTimeSeries (ManageTimeSeriesThread.java:215)
    at org.gcube.data_catalogue.grsf_publish_ws.utils.csv.ManageTimeSeriesThread.run (ManageTimeSeriesThread.java:90)
```

@francesco.mangiacrapa@isti.cnr.it is not working today. We have to wait for him.

#14 - Mar 26, 2020 12:53 PM - Luca Frosini

@marketak@ics.forth.gr can you please provide me with the JSON you are using to update the record so that I can made some test?

#15 - Mar 26, 2020 01:02 PM - Luca Frosini

I @marketak@ics.forth.gr

thanks to your and @francesco.mangiacrapa@isti.cnr.it suggestion I found that the Record URL is explicitly kept as is.

Is it a problem for you to delete and recreate the records?

#16 - Mar 26, 2020 01:02 PM - Luca Frosini

- Assignee changed from Francesco Mangiacrapa to Luca Frosini

#17 - Mar 26, 2020 01:06 PM - Yannis Marketakis

I guess I could try deleting and republishing.

I'll let you know.

#18 - Mar 26, 2020 04:19 PM - Yannis Marketakis

I've deleted and re-published 4 stock records. Now their URLs are OK.

In addition, I've updated several stock and fishery records that were linked to those 4 (either as similar records or as exploited resources). Now they seem OK.

#19 - Mar 26, 2020 04:21 PM - Aureliano Gentile

with many thanks, in the mean time I am checking the new data harvest, I assume all of these fixes will be inherited.

#20 - Mar 30, 2020 01:34 PM - Luca Frosini

- Status changed from In Progress to Closed

- % Done changed from 80 to 100

Files

screenshot-i-marine.d4science.org-2020.03.02-17_46_58.png	179 KB	Mar 02, 2020	Aureliano Gentile
updateLog.log	506 Bytes	Mar 26, 2020	Yannis Marketakis