StocksAndFisheriesKB - Bug #20975

GRSF catalogue: group filters are not consistent with totals

Mar 16, 2021 12:40 PM - Aureliano Gentile

Status:	Closed	**Start date:**	Mar 16, 2021
Priority:	Normal	**Due date:**	
Assignee:	Luca Frosini	**% Done:**	10%
Category:		**Estimated time:**	0.00 hour
Target version:	GRSF		

Description

Dear Colleagues, please see attached screenshots taken from the GRSF Admin VRE
https://i-marine.d4science.org/group/grsf_admin/data-catalogue.
The total RAM records indicates 1 374 records.
But the total by groups is 1 372 (51+69+85+1167=1372).
The correct amount is 1374 as confirmed by @marketak@ics.forth.gr with his competency queries.

This seems a recurring issues, how can it be fixed once for all? Or what best practice can you suggest to avoid reporting to colleagues wrong figures?

With thanks in advance
Aureliano

History

#1 - Mar 24, 2021 03:33 PM - Francesco Mangiacrapa

- Status changed from New to Feedback

- % Done changed from 0 to 100

Hi @aureliano.gentile@fao.org and @marketak@ics.forth.gr,

with a query on SOLR index I detected the 2 (1374-1372=2) records of GRSF_ADMIN with a resource named "RAM" that are not assigned to any "ram" group: "grsf-stock-ram", "grsf-stock-firms-ram", "grsf-stock-firms-fishsource-ram", "grsf-stock-fishsource-ram". The records are:

```
| name                               | title                                         | extras_GRSFUUID
          | extras_systemtype | extras_RecordURL
    | groups
                              | res_name                                |
|------------------------------------|-----------------------------------------------|------------------------
-------------|-------------------|----------------------------------------------------------------------
----|----------------------------------------------------------------------------------------------------
-------------------------------|-----------------------------------------------|
| 4d6f3906-6a74-3505-9ac4-df02a17bbb34 | Merluccius merluccius - Western Ionian    | 4d6f3906-6a74-3505-9ac4-
df02a17bbb34 | Assessment Unit   | http://data.d4science.org/ctlg/GRSF_Admin/4d6f3906-6a74-3505-9ac4-df02a17bb
b34 | abundance-level,abundance-level-firms-standard,catches,fishing-pressure-firms-standard,grsf-assessment-u
nit,grsf-stock,scientific-advice | RAM,Fishing Pressure,State and Trend,Catches |
| 4c315909-e34e-3f32-81cd-99ef3bbbc2fd | Oncorhynchus keta - Primorye chum hatchery | 4c315909-e34e-3f32-81cd-
99ef3bbbc2fd | Assessment Unit   | http://data.d4science.org/ctlg/GRSF_Admin/4c315909-e34e-3f32-81cd-99ef3bbbc
2fd | grsf-assessment-unit,grsf-stock
                              | RAM                                     |
```

so Yannis can add them to the proper group/s.

In my opinion, maybe, to limit this kind of errors, Yannis could implement a programmatic check that after a publishing stage, running the program it could verify the "counts" and the consistency between KB numbers and Catalogue numbers.
From a Catalogue point of view, it is possible to count the items published in the catalogs of "GRSF" and "GRSF_ADMIN" VRE using the "count" facility provided by gCat Service (see at https://wiki.gcube-system.org/gcube/GCat_Service) developed by @luca.frosini@isti.cnr.it

For Example:

To get the number of items belonging to group "Stock - RAM" the query is:

```
https://api.d4science.org/catalogue/items?q=groups:grsf-stock-ram&count=true&gcube-token={{GRSF_PUBLISHER_GCUB
E_TOKEN}}
```

(in the context of GRSF_ADMIN, so GRSF_PUBLISHER_GCUBE_TOKEN = the token of *grsf.publisher* user in the "GRSF_ADMIN" VRE)

the response is:

```
{
count: 1167
}
```

#2 - Mar 24, 2021 03:52 PM - Aureliano Gentile

With thanks, I understand the ticket is now addressed to @marketak@ics.forth.gr to ensure those two records (RAM assessment units) are assigned to the proper group(s).

#3 - Mar 24, 2021 04:51 PM - Pasquale Pagano

- Assignee changed from Francesco Mangiacrapa to Yannis Marketakis

#4 - Mar 24, 2021 05:19 PM - Yannis Marketakis

- File 4d6f3906-6a74-3505-9ac4-df02a17bbb34.json added

- Status changed from Feedback to In Progress

- Assignee changed from Yannis Marketakis to Francesco Mangiacrapa

- % Done changed from 100 to 90

Groups (and tags) are assigned during publishung/updating from the publisher service.
From our side, we are providing the json representation of the record (without any indication about groups or tags).
Attached you'll find the json representation for one of the "problematic" records.

So I think @francesco.mangiacrapa@isti.cnr.it should check why this json is not added in the proper group.

#5 - Mar 24, 2021 05:52 PM - Luca Frosini

- Assignee changed from Francesco Mangiacrapa to Luca Frosini

I'm going to investigate the issue

#6 - Jun 16, 2021 02:46 PM - Luca Frosini

- Status changed from In Progress to Paused

#7 - May 11, 2022 04:45 PM - Luca Frosini

- Status changed from Paused to Closed

Files

ram_total1.png	130 KB	Mar 16, 2021	Aureliano Gentile
ram_total2.png	20.2 KB	Mar 16, 2021	Aureliano Gentile
ram_records_catalogue_16march2021.png	47.3 KB	Mar 16, 2021	Aureliano Gentile
4d6f3906-6a74-3505-9ac4-df02a17bbb34.json	9.28 KB	Mar 24, 2021	Yannis Marketakis