

GRSF groups are not indicating the exact amount of records

Status:	Closed	Start date:	Nov 25, 2020
Priority:	Normal	Due date:	
Assignee:	Aureliano Gentile	% Done:	90%
Category:		Estimated time:	0.00 hour
Target version:	GRSF		
Description			
Following the refresh of the GRSF KB by @marketak@ics.forth.gr we noticed that the groups are not showing up-to-date figures. See for example the attached screenshot where the sum of total stocks and fisheries differs from the total of GRSF records. This nuance is also among other groups. This occurs both in grsf_admin and grsf VREs. FYI: @bracken.vanniekerk@fao.org			

#1 - Dec 15, 2020 01:38 PM - Francesco Mangiacrapa

- Status changed from New to In Progress
- Assignee changed from Francesco Mangiacrapa to Yannis Marketakis
- % Done changed from 0 to 50

By checking our Solr indexes, I just noticed that some records are not assigned to *grsf-stock* or *grsf-fishery* group.

- For **GRSF_ADMIN** are 8, in fact $16036+8=16044$ that is exactly the amount of records in the GRSF ADMIN organization (see at http://ckan-grsf-admin.d4science.org/organization/grsf_admin). They are:

name	title
	groups
----- -----	
----- -----	
----- -----	
0d5857fc-eda4-32b6-98b6-f589c6ffa5fb	Raja naevus - Portuguese Waters - East (Division 27.9.a)
dc5d0c9b-1a36-3c75-bee4-1273e3af275b Xiphias gladius - Pacific, Western Central - Pacific, Eastern Central - Western and Central Pacific Fisheries Commission (WCPFC) - Forum Fisheries Agency (FFA) - Vanuatu - Gear ne i	
d334e18e-3168-4076-bale-a44adbebb5a0 Plesionika martia - Ionian - Sardinia - Aegean - Levant - Balearic - Adriatic - Gulf of Lion - General Fisheries Commission for the Mediterranean (GFCM) - Malta - Single boat bott om otter trawls	
143b4ae3-blaa-4a58-a34d-846465108ac2 Pomacentridae - Pacific, Western Central - Pacific, Eastern Central - Western and Central Pacific Fisheries Commission (WCPFC) - Forum Fisheries Agency (FFA) - Solomon Islands - M echanized lines and pole-and-lines	
2a88cc94-788c-3212-a25b-dc347c5f30bb Polysteganus coeruleopunctatus - Indian Ocean, West / 51.8	
	grsf-stock-firms
3798d2d1-43ed-3fa0-9b0b-293d9e7aa7e6 Pandalus borealis - Atlantic, Northwest / 21.4.R - Atlantic, Northwes t / 21.4.T - Atlantic, Northwest / 21.4.S	
37c06b43-9179-329f-b6d7-2991688a776d Macrodon ancylodon - Guyana - French Guiana	
	abundance-level-firms-standard,fao-categories,fishing-pressure-firms-stan dard.grsf-assessment-unit.grsf-stock-firms

```
| 661b6b76-4e98-3d09-aa49-e9049586fc48 | Molva molva - Atlantic, Northeast
| abundance-level,abundance-level-firms-standard,fishing-pressure,fishing-pressure-firms-standard,grsf-assessment-unit,grsf-stock-firms |
```

- For **GRSF** is 1, in fact 1532 (GRSF stock) + 91 (GRSF Fishery) = 1623. The records are 1624 (see at <https://ckan-grsf.d4science.org/dataset>) and the missing one is the following one:

name	title	groups
-----	-----	-----
99904192-d6fd-39e1-9995-a8a2985b1177	Merluccius polli - Cunene - Cape Salinas - Cape Palmeirinhas	

I'm going to assign this ticket to @marketak@ics.forth.gr in order to perform the proper action (update or remove) the above records. The updating should add them to the proper group/s.

#2 - Dec 15, 2020 02:36 PM - Yannis Marketakis

- File Screenshot 2020-12-15 153541.png added

I first tried **updating** the record in GRSF VRE. However, the publisher service reported that the record didn't exist in GRSF Catalog. In addition, I cannot search for the record (using the search field),

For example, if I search for records using the term Merluccius, I face a server error

#3 - Dec 15, 2020 02:54 PM - Yannis Marketakis

For GRSF Admin, records were updated without any issues. The figures seem correct now (in GRSF Admin)

#4 - Dec 15, 2020 02:59 PM - Francesco Mangiacrapa

Yannis Marketakis wrote:

I first tried **updating** the record in GRSF VRE. However, the publisher service reported that the record didn't exist in GRSF Catalog. In addition, I cannot search for the record (using the search field),

For example, if I search for records using the term Merluccius, I face a server error

Uhhh... you're right, if the record is missing on KB side, it should be an "orphan" record. Let me know if I can remove it from the Solr index

#5 - Dec 15, 2020 03:00 PM - Yannis Marketakis

Francesco Mangiacrapa wrote:

Uhhh... you're right, if the record is missing on KB side, it should be an "orphan" record. Let me know if I can remove it from the Solr index

Sure. As soon as you remove it, I guess I can publish it again.

#6 - Dec 15, 2020 04:53 PM - Francesco Mangiacrapa

- % Done changed from 50 to 80

Removed from the Solr index (in the GRSF catalogue, record name 99904192-d6fd-39e1-9995-a8a2985b1177)

#7 - Dec 16, 2020 09:32 AM - Yannis Marketakis

- Status changed from In Progress to Feedback
- Assignee changed from Yannis Marketakis to Aureliano Gentile
- % Done changed from 80 to 90

The missing record was approved again and published in GRSF Public VRE. Now the figures seem OK.

@aureliano.gentile@fao.org and @bracken.vanniekerk@fao.org can you please perform a quick check?

#8 - Dec 21, 2020 04:44 PM - Aureliano Gentile

- File Annotation 2020-12-21 163701.png added

We checked the high level groups vs the competency queries https://i-marine.d4science.org/group/grsf_admin/grsf-competency-queries and figures seems now fine. (Total GRSF Stocks, total GRSF fisheries, total assessment units).

- Group Fishing Units seems missing in the CKan list of groups
- The group GRSF Assessment Unit (**2781**) is smaller than the Type Assessment Unit (**2784**) where the latter is the correct one according to the competency query (19. Count all Assessment Units)

- File GRSF_Admin list of Groups.png added

Aureliano Gentile wrote:

- Group Fishing Units seems missing in the CKan list of groups

- The group GRSF Assessment Unit (**2781**) is smaller than the Type Assessment Unit (**2784**) where the latter is the correct one according to the competency query (19. Count all Assessment Units)

For **GRSF ADMIN** are 3:

Jun 17, 2025

ishsource-ram |

For GRSF is 1:

name	title	extras_systemtype	groups
7a4bc5ae-1947-31a5-aaf2-7f3dcbf40f5c	Northern prawn - Atlantic, Northwest / 21.4.S	Assessment Unit	grsf-stock,grsf-stock-fishsource

I kindly ask to @marketak@ics.forth.gr to perform an update for them. Many thanks.

#10 - Dec 22, 2020 11:01 AM - Yannis Marketakis

Francesco Mangiacrapa wrote:

I believe the Group "Fishing Units" is missing in the list because it never created as CKAN group but only as "type" (alias system:type). In fact, the list of groups for GRSF and GRSF_ADMIN are consistent (see attachments). @marketak@ics.forth.gr, can you confirm?

I do not handle this type of information. I guess this is handled by the GRSF publisher service. I am only using a client of this service that provides a json object to be published.

Francesco Mangiacrapa wrote:

I detected them by performing a query on SOLR.

I kindly ask to @marketak@ics.forth.gr to perform an update for them. Many thanks.

Yes I will

#11 - Dec 22, 2020 11:19 AM - Francesco Mangiacrapa

Yannis Marketakis wrote:

Francesco Mangiacrapa wrote:

I believe the Group "Fishing Units" is missing in the list because it never created as CKAN group but only as "type" (alias system:type). In fact, the list of groups for GRSF and GRSF_ADMIN are consistent (see attachments). @marketak@ics.forth.gr, can you confirm?

I do not handle this type of information. I guess this is handled by the GRSF publisher service. I am only using a client of this service that provides a json object to be published.

Sure. But it is quite evident that the Group "Fishing Units" is never created because it is never required. The groups on GRSF and GRSF_ADMIN were created before submitting the records. In fact, the GRSF publisher service (at the moment) is not able to create a CKAN group on-the-fly (@luca.frosini@isti.cnr.it can confirm?) but only to add the record to appropriate groups by parsing the json object on publishing stage.

#12 - Dec 22, 2020 11:32 AM - Yannis Marketakis

Francesco Mangiacrapa wrote:

Sure. But it is quite evident that the Group "Fishing Units" is never created because it is never required.

Basically, approximately 11K records are fishing units (you can check this if you browse by fishing type). Therefore, from your response, I understand that the group was not created before actually publish the records

#13 - Dec 22, 2020 11:56 AM - Yannis Marketakis

As regards the wrong numbers for GRSF Assessment Units, I've updated the records as indicated by @francesco.mangiacrapa@isti.cnr.it and the figures in both VREs seem OK now.

#14 - Dec 22, 2020 12:36 PM - Aureliano Gentile

With thanks and ...Buon Natale!

#15 - Dec 22, 2020 12:53 PM - Francesco Mangiacrapa

- % Done changed from 90 to 100

Yannis Marketakis wrote:

Francesco Mangiacrapa wrote:

Sure. But it is quite evident that the Group "Fishing Units" is never created because it is never required.

Basically, approximately 11K records are fishing units (you can check this if you browse by fishing type). Therefore, from your response, I understand that the group was not created before actually publish the records

Exactly, you understood perfectly.

However, if we want to add the new group "Fishing Units" before we need to create it on both the catalogues (GRSF and GRSF_ADMIN), maybe update the business logic in the GRSF Publisher service (@luca.frosini@isti.cnr.it can reply on this) and then update all the 11K records of type "fishing units". If this new activity is required, please @aureliano.gentile@fao.org create a new ticket. Thanks.

... Buon Natale!

#16 - Dec 26, 2020 07:11 PM - Aureliano Gentile

Dear all, thank you for the analysis.

I noticed the missing group, hence I reported in the ticket, but indeed we would need to understand first the impact.

Before I create the new ticket I would like to understand if this improvement is too demanding or not. As matter of consistency that group should be there similarly to the other groups, that would ease also the filtering in the CKan interface. However same result can be achieved playing with types, but it is a workaround and for the public interface it is a bit confusing.

#17 - May 11, 2022 04:46 PM - Luca Frosini

- Status changed from Feedback to Closed

#18 - May 11, 2022 04:47 PM - Luca Frosini

I'm going to close this ticket because is rather old and I think the new version of the service will solve a lot of these issues.

We will open a new one if needed.

Files

screenshot-i-marine.d4science.org-2020.11.25-11_34_36.png	132 KB	Nov 25, 2020	Aureliano Gentile
Screenshot 2020-12-15 153541.png	33.8 KB	Dec 15, 2020	Yannis Marketakis
Annotation 2020-12-21 163701.png	59.8 KB	Dec 21, 2020	Aureliano Gentile
GRSF_Admin list of Groups.png	64.9 KB	Dec 22, 2020	Francesco Mangiacrapa
GRSF list of Groups.png	64 KB	Dec 22, 2020	Francesco Mangiacrapa