BlueBRIDGE External Advisory Board (EAB) - FIRMS
Joint meeting of BlueBRIDGE Technical Working Group (TWG3) and FIRMS TWG6 on the Global Record of Stocks and Fisheries (GRSF)
Rome, Italy, 7 to 9 February 2018
Data access and sharing policies for the Global Record of Stocks and Fisheries
Author: Akrivi Katifori, Eise van Maanen, Anton Ellenbroek, Aureliano Gentile

Table of Contents

Executive Summary

The Global Record of Stocks and Fisheries (GRSF) as a service needs to find its place within a growing volume of transborder data flows for which global data protection policies are not yet fully defined. The development and promotion of legal interoperability is therefore of high importance and its study and application in particular working cases like the GRSF can significantly contribute to general best practices to be applied in other cases as well.

This study defines legal interoperability as *the compatibility of legal rights, terms, and conditions of databases from two or more sources so that the data may be combined to be offered for different uses in a clear legal framework and without compromising the legal rights of any of the data sources used.*

The report proposes a data citation policy, disclaimer policy and terms of use, while discussing in its Annex current general relevant issues involved in data citation, which are an important aspect of legal interoperability along with the GRSF specific characteristics to be taken into account and the perspective of the three data providers involved in GRSF.

1. Introduction

The Global Record of Stocks and Fisheries (GRSF) as a service needs to find its place within a growing volume of transborder data flows for which global data protection policies are not yet fully defined. Data protection laws are very different in various regions of the world. A harmonisation of these rules is not expected in the near future, but the costs of non-interoperable laws in a highly networked world are high and will increase even more.

The development and promotion of legal interoperability is therefore of high importance and its study and application in particular working cases like the GRSF can significantly contribute to general best practices to be applied in other cases as well.

This study defines legal interoperability as *the compatibility of legal rights, terms, and conditions of databases from two or more sources so* that *the data may be combined to be offered for different uses in a clear legal framework and without compromising the legal rights of any of the data sources used.*

Legal interoperability occurs among multiple datasets when [6]:
- the legal use conditions are clearly and readily determinable for each of the datasets, typically through automated means;
- the legal use conditions imposed on each dataset allow creation and use of combined or derivative products; and
- users may legally access and use each dataset without seeking authorization from data rights holders on a case-by-case basis, assuming that the accumulated conditions of use for each and all of the datasets are met.

When data are combined from multiple sources, the resulting dataset will incorporate the accumulated restrictions imposed by each source. The fewest restrictions contained in parent datasets results in the fewest restrictions in derivative datasets. The simplest cases for tracking and legal interoperability occur when datasets are affirmatively identified as having no legal restrictions.

Because of the different copyright status of databases and data content, different mechanisms are required to manage each. Copyright and Terms of Use can govern the use of databases and some data content (that which is itself original), but contract law, trademarks, and other mechanisms are required to regulate factual data.

It is worth to consider that the system in which data content is stored may be operated by a different actors and offered through specific terms of use. In this case four types of policies may be involved: (i) the one agreed between the system operator and the primary data owner, (ii) the one selected for a derivative product that may differ from the one associated with primary data, (iii) the one that sets citation standards for all parties involved, and (iv) the one agreed between the system operator and the data consumer.

All these policies have to be captured by the 'terms of use' of the system, i.e., they are part of the rules a consumer must accept when using the system.

The report proposes a data citation policy, disclaimer policy and terms of use, while discussing in its Annex current general relevant issues involved in data citation, which are an important aspect of legal interoperability along with the GRSF specific characteristics to be taken into account and the perspective of the three data providers involved in GRSF.

2. Citation policy

It is important to highlight that GRSF should provide the possibility to the users to view and use a complete citation to the Record they are interested in as well as its related sources. At the moment this is not possible in GRSF. The suggested data citation is included on all relevant pages. This means that data users can easily copy the suggested relevant citation of the data.

The principles as stated in the previous section should be adhered to when applying to a practical data citation. This section proposes a citation format for the GRSF record based on:

- The current FIRMS record citation format [2]
- Guidelines from the Digital Curation Centre[*]
- Tenets of several research papers [1] [2] [4] [5] [7].

In case of citation for derivative products, as it is the case of GRSF, it is best practice to mention the originators of the data with which the derivative products have been realized. This is in fact required for any type of CC license, except for the CC0 license. Based upon the logic that any shared or public content in the iMarine - BlueBRIDGE Data e-Infrastructure goes with its own citation, the citation mechanism is foreseen to provide ways to concatenate all citations for the utilized sources (dataset, documents etc.).

Each citation should include elements from the list presented in Table 1.

Applying this approach on the GRSF record, a general citation format is proposed for a GRSF record and its accompanying data sources.

> © [Copyright holder]. [Title] ([Year]). In: Global Record of Stocks and Fisheries (GRSF). [Publisher]. Version: [Version]. Updated [Update date]. [Date accessed and/or downloaded [Access date]]. [GRSF Record URL]. Database sources: [Database sources]. Sources citations: [Citations to the database sources]

Table 1 presents the elements of this citation and Table 2 the proposed citations for the data sources.

Table 1. GRSF Record proposed citation elements

Copyright Holder	This includes the symbol © followed by the copyright holder name. In the case where FIRMS partnership will take ownership of GRSF the Copyright holder will be © FAO. If this is not the case, the copyright holder should be identified.
Title	The name of the cited resource itself, this may also include the name of a facility and the titles of the top collection and main parent subcollection (if any) of which the data set is a part. In the case of the GRSF, this is the "GRSF title" element of the GRSF record.
Year	This is the year of creation of the intellectual content. It is presented in the form of the stating reporting year and followed by a "-" and the last reporting year or the current year, in the case of an electronic resource constantly updated, inside parenthesis.

---

[*] http://www.dcc.ac.uk/

Data collection	The data collection in this case is Global Record of Stocks and Fisheries (GRSF), preceded by "In: ".
Publisher	The organization (or repository) either hosting the data or performing quality assurance. The Publisher in the case of the GRSF as part of the BlueBRIDGE project is the infrastructure organization, which is the D4Science.org Consortium.
Version	This is the version of GRSF, preceded by "Version: " in the following format: v[Version number]-[Year]
Update date	The date the document was uploaded or updated in the system in the format DD MM Year, preceded by "Updated "
Access date	The date the document was accessed in the system, in the format DD MM Year, preceded by "Date accessed and/or downloaded" and in brackets: [Date accessed and/or downloaded DD MM Year].
GRSF Record Identifier and Location	The identifier is a resolvable web identifier for the data, according to a persistent scheme whereas the location is persistent URL or UNF from which the data set is available. In this case it is the GRSF Record URL which contains also the GRSF identifier.
Database sources	At the current status of GRSF these are: FIRMS, FishSource, RAM
Database sources citations	See Table 2 for the proposed citation options for the data sources

Table 2. GRSF database sources proposed citation formats

FIRMS	The proposed FIRMS citation follows the FIRMS Information Management Policy[†] citation guidelines for the full citation:
	[© FAO] [Date created]. [Title of content]. [Series title]. [Contributors]. In: [Publisher]. Updated [Update date]. Accessed [Date accessed and/or downloaded]. [URL]
	The citation format for FIRMS will be provided by FIRMS and it may differ from this proposed format.
RAM Legacy Stock Assessment Database	After consideration of different options for the RAM Legacy Stock Assessment Database, the current approach for the GRSF is to cite the whole database, as is currently the proposed approach in the RAM website (See Annex RAM Legacy Stock Assessment Database), with the following format:
	RAM Legacy Stock Assessment Database. ([Date created]). Version: [Version]. Updated [Update date]. Accessed [Date accessed and/or downloaded]. Retrieved from [URL]
	The proposed database citation for RAM becomes
	RAM Legacy Stock Assessment Database. ([Date created]). Version: 4.25. Updated [Update date]. Accessed [Date accessed and/or downloaded]. Retrieved from http://ramlegacy.org/
FishSource	Following a similar approach as that of FIRMS, the following citation approach could be used for FishSource:
	© SFP. [Date created]. [Title of content]. [Series title]. [Contributors]. In: FishSource [online]. Updated [Update date]. Accessed [Date accessed and/or downloaded]. [URL]

The following table presents citation examples for different GRSF record with one or more associated data sources. For this example, GRSF is available under FAO.

Table 3. GRSF Record citation examples with one or multiple sources.

GRSF Record citation example with one FIRMS source	© FAO. *Carcharhinus porosus* Trinidad and Tobago (2017). In: Global Record of Stocks and Fisheries (GRSF). D4Science.org. Version: 1-2017. Updated 10 August 2017. [Date accessed and/or downloaded 28 November 2017]. http://data.d4science.org/ctlg/GRSF_Admin/f3c1ab0f-0f2c-3c03-8e15-1a62f7e73ef6
	Database sources: FIRMS
	Sources citations:
	© FAO 2000-2017. FAO WECAFC Working Party on Assessment of Marine Fishery Resources (FAO WECAFC- WPAMFR). Status of stocks and resources 1999. Smalltail shark - Coastal areas of Trinidad and Tobago. FIRMS Reports. Text by S. Kuruvilla, S. Soomai. In: Fisheries and Resources Monitoring System (FIRMS) [online]. Rome. Updated 1999.

[†] Firms Information Management Policy, pp.8-9, http://www.fao.org/3/a-ax530e.pdf

	[Cited 28 November 2017]. http://firms.fao.org/firms/resource/13198/en
GRSF Record citation example with three database sources	© FAO. *Pleuronectes platessa* Western English Channel (Division 27.7.e) (2017). In: Global Record of Stocks and Fisheries (GRSF). D4Science.org. Version: 1-2017. Updated 9 August 2017. [Date accessed and/or downloaded 4 December 2017]. http://data.d4science.org/ctlg/GRSF_Admin/30c0991f-8373-3607-ba9a-804abaf5cc9b
	Database sources: FIRMS, RAM, FishSource
	Sources citations: ● © FAO 2012-2017. International Council for the Exploration of the Sea (ICES). ICES Advice 2012. Plaice - Western Channel. FIRMS Reports. In: Fisheries and Resources Monitoring System (FIRMS) [online]. Rome. Updated 9 November 2012. [Cited 4 December 2017]. http://firms.fao.org/firms/resource/10425/en ● RAM Legacy Stock Assessment Database. ([Date created]). Version: 4.15. Updated [Update date]. Accessed [Date accessed and/or downloaded]. Retrieved from http://ramlegacy.org/ ● © SFP. European plaice Western English Channel. FishSource profiles. In: FishSource [online]. Updated 25 June 2015. Accessed [Date accessed and/or downloaded]. https://www.fishsource.org/stock_page/1755.

3. GRSF Terms of Use

General

Welcome to the Global Record for Stocks and Fisheries Catalogue (the GRSF), which gives access to a global reference set of stocks and fisheries records and their corresponding details.

These records are (1) those coming from different and authoritative sources (the Sources) and (2) those resulting from the aggregation and harmonization through international and dedicated standards of the aforementioned Sources records (the GRSF Records).

By accessing and/or using any information of the GRSF, you consent and agree to comply with these Terms and the iMarine Gateway Terms of Use (https://i-marine.d4science.org/terms-of-use), Privacy Policy (https://www.iubenda.com/privacy-policy/441050/legal) and Cookie Policy (https://i-marine.d4science.org/cookie-policy) which are incorporated herein by reference. The rights and obligations concerning the use of the GRSF in accordance with these Terms is personal to you and to members of your institution and is not transferable to any other person or entity.

Disclaimer

The GRSF has been produced with co-funding from the European Commission. The aggregated content it contains does not necessarily reflect the views of the European Commission or those of the BlueBRIDGE Consortium (the Consortium).

Content on the GRSF is provided on an "as is" and "as available" basis. The Consortium makes every effort to ensure, but does not guarantee, the accuracy, completeness and authenticity of the information on the GRSF. The Consortium reserves the right to alter, limit or discontinue any part of this service at its discretion and without prior notice. Under no circumstances shall the Consortium be liable for any loss, damage, liability or expense suffered that is claimed to result from the use of content posted on the GRSF, including, without limitation, any fault, error, omission, interruption or delay.

The hyperlink references to the non-GRSF websites of the Sources do not imply endorsement by, or responsibility on the part of, the Consortium for the opinions, ideas, data or products presented on such websites, or otherwise guarantee the validity of the content provided. The references are provided based on the premise that users are responsible for assessing their relevance, accuracy and suitability for application.

Copyright

The GRSF contains copyrighted material and/or other proprietary information and thus is protected by copyright laws and regulations worldwide. The Consortium grants the user the license to freely use, download and print the materials contained in the GRSF Records under the CC BY-SA 4.0 license and in accordance with these Terms. Material provided by the GRSF that is copyrighted by the Sources is governed by the terms and conditions of the copyright holder.

Citation

Each GRSF Record and Sources Record page in the GRSF contains a proposed citation to cite its content. You are advised to use this citation if you would like to refer to this content.

Acknowledgements

The GRSF can be accessed through the iMarine Data Catalogue (D4Science.org infrastructure operated by CNR-ISTI), and it is based on the open source CKAN software. Two Virtual Research Environment (VREs) have been developed: i) the "GRSF" VRE for the public users with validated records, and ii) the "GRSF Admin" VRE for authorized users to manage and validate the GRSF records. The GRSF Knowledge Base has been developed with MatWare (FORTH semantic web technology using ontologies).

References

[1] Altman, M. & King, G. (2007). A proposed standard for the scholarly citation of quantitative data. *D-lib magazine 13 (3/4).* Retrieved from http://dlib.org/dlib/march07/altman/03altman.html

[2] Egloff, W., Agosti, D., Kishor, P., Patterson, D., & Miller, J. (2017). Copyright and the Use of Images as Biodiversity Data. *Research Ideas and Outcomes*, *3*, e12502.

[3] Firms Information Management Policy, pp.8-9, http://www.fao.org/fishery/docs/DOCUMENT/FIGIS_FIRMS/2010/inf3e.pdf

[4] Green, R., & Awre, C. (2009). Towards a Repository-enabled Scholar's Workbench. *D-Lib Magazine*, *15*(5/6), 1082-9873.

[5] Lawrence, B., Pepler, S., Jones, C., Matthews, B., McGarva, G., & Coles, S. (2007). Linking data and publications in the environmental sciences: CLADDIER project workshop, Chilworth, Southampton, UK 15th May 2007.

[6] Legal Iinteroperability of Research Data: Principles and Iimplementation Guidelines, RDA-CODATA Legal Interoperability Interest Group, http://rd-alliance.org/sites/default/files/attachment/Legal%20Interoperability%20Principles%20and%20Implementation%20Guidelines_Final.pdf

[7] Pearson, S.H. (2012). *For attribution: Developing data attribution and citation practices and standards: Summary of an international workshop. Three legal mechanisms of sharing data.* National Research Council. National Academies Press, 2012. Pp. 71-87.

Annex 1 - Background on legal interoperability

RELEVANT LEGAL INTEROPERABILITY ISSUES

It appears that there are various data sharing issues in current practice, particularly when data is merged or aggregated by different entities. This section presents the most relevant citation issues identified after consultation of the relevant stakeholders of GRSF.

There is a difference between attribution and citation

Attribution *refers to the legally imposed requirement to attribute the rights holder when the data are copied or reused in a specified manner.*

The remedy against someone who fails to attribute is a lawsuit, either based on breach of contract or infringement of an intellectual property right, depending on the legal mechanism used to impose the attribution requirements. So being the rights holder is conditional for attribution.

Citation *is rooted in norms of scholarly communication. The purpose of citation is to support an argument with evidence, pointing to the source of this evidence.*

However, citation has also become a proxy for credit, albeit an imperfect one.

This is an important starting point. It reminds us that legal attribution requirements do not necessarily match our expectations for receiving credit, nor do they perfectly map to accepted standards of citation.

Whenever data are shared, there is a possibility they will not be properly cited upon reuse. Licenses and contracts attempt to eliminate this risk by imposing legal attribution requirements. Waivers, however, do not legally impose attribution. Instead, they rely on community norms to ensure proper citation (Pearson, 2012).

Citations should facilitate both giving scholarly credit, and legal attribution to all parties responsible for those data.

There is a difference between "data" and a "database"

In any data project, there are likely to be two components. The first is the data collected, assembled, or generated. The second component is the database‡ in which the data is stored and managed. We usually do not think of data content separate from the system in which it is stored, but the distinction is important in terms of intellectual property rights. The question is what, if anything, is protected by copyright.

Data that is merely factual in most cases has no copyright protection. But once this factual data is organized into a database with 'creative decisions', such as the way the data is organized and different data elements are combined, a database can have a thin layer of protection. In the EU and a few other countries, governments have implemented what are called sui generis ("of their own kind") database rights. These rights allow a database maker to prevent the extraction and reuse of a substantial part of the contents of a database, even if the contents are otherwise in the public domain.

Important to note is that licensing also concerns the database itself. A license can be built atop copyright or database rights or both. For example, Creative Commons ('CC') licenses are copyright licenses. If a CC license is applied to a database, it covers both the data and the database, all to the extent each is subject to copyright. Any use of the data or database that implicates copyright, requires attribution. Any use of the data that does not implicate copyright – if for example, the data are in the public domain – does not require attribution, even if it triggers database rights.

---

‡ A compilation of works, data or other materials (i.e. collection of facts) arranged in a systematic or methodical way. In other words, ordered by logical principles set up by the compiler.

There is a lack of clarity on the citation of compound datasets

Title 17 of the US Code (U.S.C. 17) refers to compound data products as works formed by the collection and assembling of pre-existing materials. Whenever an individual within the research community uses parts of a dataset, it is not always clear how to cite this part of the dataset. This happens frequently, especially in international organizations where data is sourced from many countries and different institutes.

In the science context in particular, projects often rely on data gathered from a variety of different sources. Depending on the licenses used, it is possible that it would require attributing each individual or institution that contributed any piece of data to the project. This is a problem we call attribution stacking. This raises yet another potential problem with attribution. Attribution obligations written into a license are, by their nature, inflexible. No lawyer can anticipate every situation in which the attribution requirements would be triggered and account for all of the circumstances in which they will be applied. This can create some absurd situations where, for example, a user or aggregator of data may technically be required to attribute 1000 different data providers, all in the exact manner that the rights holder has decided. Conceivably, the user could do all this and still not satisfy people's expectations for receiving credit or accepted standards of citation (Pearson, 2012).

DATA CITATION PRINCIPLES

FAO has implemented the FORCE 11 data citation principles[§] in all of its journals. This means that data users are encouraged to include data citations as part of their reference list.

The principles are grouped so as to facilitate understanding, rather than according to any perceived criteria of importance.

1. Importance

Data should be considered legitimate, citable products of research. Data citations should be accorded the same importance in the scholarly record as citations of other research objects, such as publications.

2. Credit and Attribution

Data citations should facilitate giving scholarly credit and normative and legal attribution to all contributors to the data, recognizing that a single style or mechanism of attribution may not be applicable to all data.

3. Evidence

In scholarly literature, whenever and wherever a claim relies upon data, the corresponding data should be cited.

4. Unique Identification

A data citation should include a persistent method for identification that is machine actionable, globally unique, and widely used by a community.

5. Access

Data citations should facilitate access to the data themselves and to such associated metadata, documentation, code, and other materials, as are necessary for both humans and machines to make informed use of the referenced data.

[§] https://www.force11.org/group/joint-declaration-data-citation-principles-final

6. Persistence

Unique identifiers, and metadata describing the data, and its disposition, should persist -- even beyond the lifespan of the data they describe.

7. Specificity and Verifiability

Data citations should facilitate identification of, access to, and verification of the specific data that support a claim. Citations or citation metadata should include information about provenance and fixity sufficient to facilitate verifying that the specific time slice, version and/or granular portion of data retrieved subsequently is the same as was originally cited.

8. Interoperability and Flexibility

Data citation methods should be sufficiently flexible to accommodate the variant practices among communities, but should not differ so much that they compromise interoperability of data citation practices across communities.

Annex 2 - The Global Record of Stocks and Fisheries - The legal interoperability perspective

The Global Record of Stocks and Fisheries (GRSF) service will provide the basis for computing regional or global indicators on the status of stocks, and will contribute to systematize a yearly production of FAO's Global State of World Fisheries and Aquaculture (SOFIA) resources[**].

It is expected that by 2017, the Global Record of Stocks and Fisheries and related status reports will contain 2000 stocks, enabling a regular computation of more reliable indicators on stock status. The Global Record will also contain 2000 fishery records, a realistic number given the effort needed to compile the cross-domain information. The potential to expand relies on co-funded new activities. The target number of requests for information from the GRSF is 100.000 per annum. The GRSF service will dynamically synchronize (initially) three main global references of stocks, and (later) extract additional stocks and related information from another five organizations.

The offered Information on the status and trends of stocks (the exploitation level) and fisheries (the ecological and economic sustainability) depends on a multi-stakeholder analysis of a complex web of data and information to produce a shared opinion[††].

The GRSF is developed in the EU H2020 BlueBRIDGE project. It is developed by FORTH - Greece, CNR-ISTI - Italy, and FAO of the UN. It will be maintained by FORTH for the application and data aspects, with an SLA being under preparation, and CNR-ISTI for the technology - information infrastructure.

The BlueBRIDGE Project Consortium Agreement is not clear on the future exploitation model for the GRSF, however this issue has been explored in the context of the project. The proposed exploitation approach is envisioned to rest on two communities:
- A public partnership focusing on an extension of the FiRMS partnership[‡‡] to also provide a secretariat, and legal ownership of the GRSF content
- A public private partnership to manage the access and use of the content

---

[**] http://www.fao.org/fishery/sofia/en

[††] https://i-marine.d4science.org/group/grsf_admin

[‡‡] http://firms.fao.org/firms/en and http://firms.fao.org/firms/about/en#Org-LegalFoundation

THE GRSF VRE

The objective of the GRSF VRE is to provide an on-line knowledge base on the Global Record of Stocks and Fisheries for a Blue Growth audience of ecologists, resource managers, market parties, and the general public with the long term objective to provide an evidence based information on the status of marine stocks and fisheries and promote responsible consumption.

Currently users need to be authorized to access the VRE. GRSF Admin is for managing GRSF records hence it is limited to a restricted number of people. It is planned to make the VRE open to search and retrieve records.

THE GRSF RECORD

A GRSF record is basically a web page with a number of metadata including a UUID. Accompanying information can be available in time series form in csv format, with textual content in pdf format, all downloadable.

An example of a GRSF record[§§] can be seen in Figure 1. A GRSF record has more info including Stock identity and Stock data (Figure 2) as well as other additional information like managerial information (Figure 3). Figure 4 presents an example of a resource. In this case, the user has the possibility to access the original resource at its source at the FIRMS service website.

Figure 1. Example of GRSF record with resources from FIRMS and RAM

[§§] https://i-marine.d4science.org/group/grsf_admin/data-catalogue?path=/dataset/fff30799-a269-356b-86a5-b22ccc860805 (iMarine e-infrastructure, login required)

License

Stock Identity

Field	Value
Assessment Area	Code: 18, System: gfcm_sub_area, Name: Southern Adriatic
GRSF Record URL	http://data.d4science.org/ctlg/GRSF_Admin/033126 02-10d0-309d-886d-b4233f8f7906
GRSF Type	Assessment Unit
Short Name	Anchovy - Western part of South Adriatic Sea
Species	Code: ANE, Classification System: ASFIS, Scientific Name: Engraulis encrasicolus
Stock Name	Anchovy - Western part of South Adriatic Sea

Stock Data

Field	Value
Abundance Level (FIRMS Standard)	Intermediate abundance Ref. Year 2006 and Rep. Year or Assessment Id 2007
Assessment Methods	Name: Echosurvey, Reference Year: 2006, Reporting Year or Assessment ID: 2007
Data owner	GFCM
Fao Categories	uncertain Ref. Year 2006
Fishing Pressure (FIRMS Standard)	Not applicable Ref. Year 2006 and Rep. Year or Assessment Id 2007
Scientific advice	It is reasonable not to increase the actual level of fishing effort on small pelagic fish SCSA endorsed the WG management advice and recommendations on anchovy (Engraulis encrasicolus) in GSA18.. Fishery data should be collected. (2007)
State and trend of Marine Resource	For Stock abundance, the large variability in time and space is mostly attributable to environmental factors. Ref. Year 2006

Figure 2. GRSF Record - Stock identity and stock data

Additional Info

Field	Value
GRSF Domain	Stock
GRSF UUID	03312602-10d0-309d-886d-b4233f8f7906
system:type	FIRMS

Management Info

Field	Value
Author	GRSF Publisher
Maintainer	GRSF Publisher
Version	1
Last Updated	August 2, 2017, 4:26 PM (UTC+02:00)
Created	August 2, 2017, 4:26 PM (UTC+02:00)

Figure 3: GRSF Record - Additional and Managerial information

Figure 4. GRSF Record - Resource Example

To provide a harmonized view of the collated data, GRSF offers a mapping for specific provider classifications to standard ones. This is offered as a GRSF service for:

- Species
- Areas
- Management entities

This transformation of the classification information is a responsibility of the provider. The providers are asked to map their classifications to GRSF standards. After this mapping is complete then the record can be considered GRSF content.

GRSF records undergo an approval stage by the providers before being published.

CHALLENGES IN THE GRSF

There are two main challenges within the GRSF that need to be considered when defining the most appropriate data policy for both data citation and attribution:

- Citing the original data providers;

- Classification harmonization

a)   Citing the original data providers

Firstly, an important issue is the fact that the three current GRSF providers are themselves aggregators of other data providers and not actual content producers. It has been the objective of this report to examine and clarify if the original content providers will be included in the citation policy.

In the case of GRSF, a development decision was made early on within BlueBRIDGE to cite only the three aggregators and not the original providers. In the table in Figure 5, FIRMS was initially presented as a source and not ICES, the original provider. However, as GRSF harvests this information, technically both approaches could be supported according to the selected citation approach. The latest version of GRSF presents both the database source and the data owner.

Abundance Level (FIRMS Standard)

URL: https://goo.gl/bSXeHF

Sprattus sprattus Northern North Sea (Division 27.4.a) Southern North Sea (Division 27.4.c) Central North Sea (Division 27.4.b)

⊞ Data Explorer

</> Embed

Add Filter

Grid | Graph | Map | 4 records | « | 1 – 4 | » | Q | Search data ... | Go » | Filters

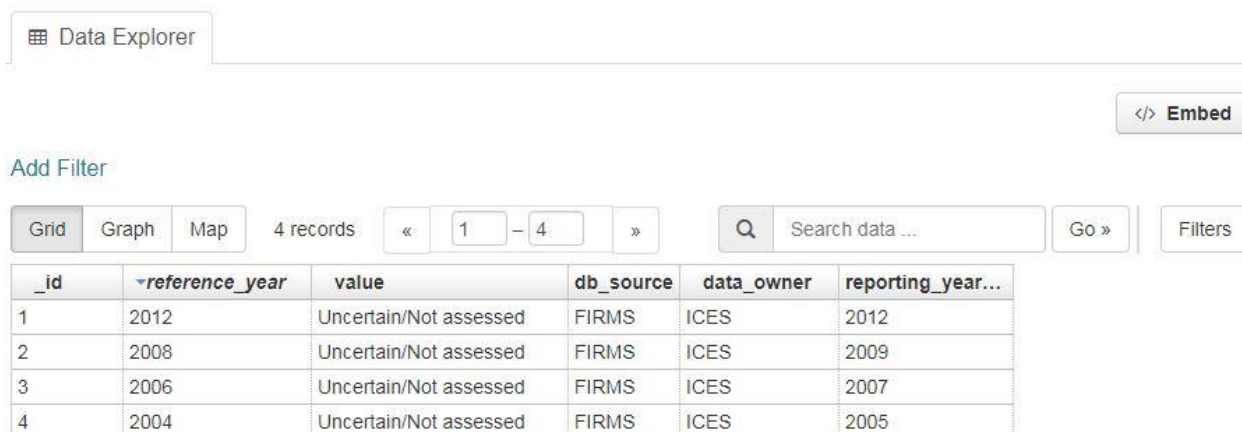| _id	▾reference_year	value	db_source	data_owner	reporting_year...
1	2012	Uncertain/Not assessed	FIRMS	ICES	2012
2	2008	Uncertain/Not assessed	FIRMS	ICES	2009
3	2006	Uncertain/Not assessed	FIRMS	ICES	2007
4	2004	Uncertain/Not assessed	FIRMS	ICES	2005

Figure 5. In this example FIRMS is identified as database source while ICES is indicated as data owner.

In this sense, the accompanying information to GRSF record creates a re-distribution issue. Data policies between the three GRSF providers and the original data providers needed to be examined and ascertain if there is need redistribution approval by them.

b)   ***Classification harmonization***

As it is discussed in the GRSF Record section, in some cases there is classification mapping for specific fields of the record for harmonization purposes. Although there is an approval stage by the providers in the workflow, it was necessary to examine if this creates a data transformation issue to be clarified in the data policy.

Annex 3 - The perspective of the GRSF Data Providers

This section presents the 3 GRSF data providers. We focus in their legal status and the provenance of the data they offer to GRSF and aim to describe the terms under which they offer these data to GRSF and also record a clear understanding of the terms under which they get these data from the individual providers.

QUESTIONNAIRE FOR THE INTERVIEWS WITH DATA PROVIDERS
1. What is the legal status of the organization?
    ● Public
    ● Private
    ● Other
    Please comment
2. Would you agree with attached citation policy, assuming that you provide data to the GRSF VRE? If no, why not? (See Section 2)
3. Do you see any practical obstacles with the implementation of the attached citation policy?
4. Would you agree with the attached terms of use policy? If no, why not?
5. Do you see any practical obstacles with the implementation of the attached terms of use policy?
6. Do you see any issue with the current collation process within GRSF?

7. Do you provide data under specific legal condition?
   If yes,
   1. Under which legal conditions for sharing and derivative products of data do your data providers provide data (e.g. CC-BY 4.0[***])?
   2. Do they restrict in any way the GRSF use case in its current pilot form?
   3. Do they restrict in any way its commercialization potential?
   If no,
   1. Do you see the lack of such licenses as a potential issue for the future exploitation of GRSF?
8. Does the original data provider have copyright over the content?
9. Have the providers waived their copyright through a CC-0 license[†††]?
10. Do you have experience with differences in legal systems and regulations on an international level throwing up a barrier for sharing and reusing of data?
11. Will the proposed business model where GRSF is offered as an FAO product, need to specify responsibilities? (E.g. the Secretariat to be responsible for compliance issues?)

FISHERIES AND RESOURCES MONITORING SYSTEM - FIRMS

The primary aim of the Fisheries and Resources Monitoring System (FIRMS)[‡‡‡] is to provide access to a wide range of high-quality information on the global monitoring and management of fishery marine resources. FIRMS provides access to information produced by a number of different Regional commissions and organizations[§§§].

The BlueBRIDGE project foresees an amendment to the FIRMS partnership to handle the issues related to GRSF.

FIRMS foresees a citation to accompany each fact sheet. An example is the following:

> © FAO 2016-2017.
>
> FAO GFCM Sub-Committee on Stock Assessment (FAO GFCM- SCSA). Stock status report 2015. Red mullet - Alboran Sea. FIRMS Reports. In: Fisheries and Resources Monitoring System (FIRMS) [online]. Rome. Updated 12 October 2016. [Cited 10 October 2017]. http://firms.fao.org/firms/resource/13769/en

RAM LEGACY STOCK ASSESSMENT DATABASE

The RAM Legacy Stock Assessment Database[****] is a compilation of stock assessment results for commercially exploited marine populations from around the world. It is a database put together by academics and government agencies, as a collaborative effort, at the moment hosted by the University of Washington.

The data provided is stock assessments that governments publish. As RAM representatives explained, approximately 90% is publicly available documents and 10% unpublished assessments

---

[***] https://creativecommons.org/licenses/by/4.0/legalcode

[†††] https://creativecommons.org/share-your-work/public-domain/cc0/

[‡‡‡] http://firms.fao.org

[§§§] http://firms.fao.org/firms/about/en

[****] http://ramlegacy.org/

from researchers, knowingly provided. These documents are provided to RAM under no specific license, according to the interviewees.

RAM users are advised to cite the database as follows in its Fair Use Policy:

Ricard, D., Minto, C., Jensen, O.P. and Baum, J.K. (2013) Evaluating the knowledge base and status of commercially exploited marine species with the RAM Legacy Stock Assessment Database. Fish and Fisheries 13 (4) 380-398. DOI: 10.1111/j.1467-2979.2011.00435.x

The RAM representatives explained however that this academic publication is outdated and it will soon be replaced by a new one.

The team considered two options:

One approach for the GRSF is to cite the whole database, as is currently the proposed approach in the RAM website , with the following format:

RAM Legacy Stock Assessment Database. ([Date created]). Version: [Version]. Updated [Update date]. Accessed [Date accessed and/or downloaded]. Retrieved from [URL]

The proposed database citation for RAM becomes

RAM Legacy Stock Assessment Database. ([Date created]). Version: [Version]. Updated [Update date]. Accessed [Date accessed and/or downloaded]. Retrieved from http://ramlegacy.org/

The second approach would be to cite the original providers:

RAM Legacy Stock Assessment Database [Date created]. [Title of content]. [Series title]. [Contributors]. In: [Publisher]. Updated [Update date]. Accessed [Date accessed and/or downloaded]. [URL]

FISHSOURCE

FishSource[††††] compiles and summarizes publicly available scientific and technical information about the status of fisheries and seafood stocks into an easily interpretable form. It aims to make information about the health of stocks, the quality of their management, and the impact of fisheries on the rest of the ecosystem more accessible to seafood buyers and help them take actions to improve the sustainability of the seafood they purchase.

FishSource is a program of the Sustainable Fisheries Partnership Foundation (SFP)[‡‡‡‡], which is US-registered non-profit. The data comes from different sources, considered as public and already available on-line. There are no agreements in place between SFP and the data providers.

The FishSource terms of use state[§§§§]: *"This Site contains copyrighted material and/or other proprietary information and thus is protected by copyright laws and regulations worldwide. SFP hereby grants you a license to freely use, download and print the materials contained in this Site for non-commercial or commercial purposes and in accordance with these Terms, provided that you do not alter or remove any copyright, trademark, or other proprietary content, notice or symbol contained on the materials or otherwise made available to you through this Site."*

FishSource does not provide a proposed citation to its users. During the interview with its representatives it has been investigated if a citation can be offered and what would be its form. An approach would be to cite the "Sources" given, which, however, in some cases are too

---

[††††] https://www.fishsource.org

[‡‡‡‡] https://www.sustainablefish.org/

[§§§§] https://www.fishsource.org/terms