

# **BlueBRIDGE External Advisory Board (EAB)**

**Technical Working Group Meeting on Global Record of Stocks and Fisheries**

**Second Session**

**(Rome, Italy, 28 February, 1 and 2 March 2017)**

## **REPORT**

## EXECUTIVE SUMMARY

The second session of the BlueBRIDGE External Advisory Board (EAB) - Technical Working Group Meeting on Global Record of Stocks and Fisheries (Rome, Italy, 28 February, 1 and 2 March 2017) offered the opportunity to discuss the on-going activities and developments of the GRSF with the EAB members, and with the representatives of the GRSF database sources (FIRMS, RAM, FishSource). Experts of the following organizations were present: Food and Agriculture Organization of the United Nations [FAO], Northwest Atlantic Fisheries Organization [NAFO], Foundation for Research and Technology - Hellas [FORTH], Consiglio Nazionale delle Ricerche [CNR], University of Washington [for RAM Legacy Stock Assessment Database], Sustainable Fisheries Partnership [SFP – for FishSource]; Mr Ramm (EAB Member - on leave from his position as Data Manager for the Commission for the Conservation of Antarctic Marine Living Resources) participated as an invited expert.

This meeting follows the completion of the compilation of the requirements (August 2016) carried out in a participatory process among the representatives of the GRSF database sources which led to the release of the GRSF prototype in January 2017.

The main objective of the meeting was to provide recommendations to finalize the development of the GRSF during the remaining year of the BlueBRIDGE project. In particular, about seventy recommendations were made for: i) standards for the fields of the GRSF records, ii) UUID and semantic identifiers iii) mappings to standards, iv) traceability, v) master data management, vi) records validation, vii) GRSF Catalogue, viii) Services/interface requirements for GRSF fisheries records, ix) requirements for reporting on stock status, x) governance and sustainability, xi) actions for the GRSF Team, and xii) a work plan for updating the GRSF knowledge base. The full list of recommendations is available in Annex 3.

The GRSF aims at providing an innovative environment supporting the collaborative production and maintenance of a comprehensive and transparent global reference set of stocks and fisheries records that will boost regional and global stocks and fisheries status and trend monitoring as well as responsible consumer practices.

FIRMS, FishSource and RAM are the three database sources from which information is harvested to form GRSF records. By collating these sources, the reporting coverage of any of these single entities is increased. The GRSF database collects information on two domains, identified as “stocks” and “fisheries”. The GRSF records are then grouped into “families” of stocks/fisheries based on specific proximity criteria.

The first GRSF prototype was presented, highlighting the advanced features to manage and browse the records. The mechanism for records generation was analysed, reaffirming the need for data harmonization through the identification and adoption of specific standards.

In particular, priority was given to international standards, while “local” standards can be accepted when they are proved to be maintained by a third party. Regarding mappings from non-standard values found in the database sources to standard values for the GRSF records, “same as” relationship were identified as the most effective approach for the detection of overlapping records. Other mapping relationships (e.g. hierarchies, intersections) are also handled. The development of reporting standards should cater for future expansion capacities.

The mechanism to build unique identifiers (UUID and semantic identifier) for stocks and fisheries was finalized to satisfy in particular the traceability and stock status summaries use cases. The unique identifiers produced by the GRSF application will be tested in the coming months by SFP-FishSource team in collaboration with selected users from the industry sector and seafood retailers. The outcome of that testing phase will further contribute to the finalization of the GRSF as well as to the formulation of the business model. The standard adopted to build unique identifiers for stocks and fisheries will be communicated at the next CWP meeting.

The FAO Master Data Management (MDM) initiative was presented and is expected to support the GRSF needs on standards and local classifications will be handled under the MDM principles. MDM data service should include a registry of available standards and mapping.

Rules for the validation and approval of GRSF records for publication were also discussed and formulated. Additional services and dedicated interfaces were envisaged in support of the data validation process and business opportunities.

User interface requirements were discussed to ensure the usability of GRSF services in relation to potential audiences and use cases. It was decided that priority for implementation should be given to the data managers' use cases, as this will ensure the full realization of GRSF potential, enabling other use cases as a consequence.

The current FAO methodology used for the preparation of The State of World Fisheries and Aquaculture report (SOFIA) was analysed in order to identify specific improvements to be implemented for the GRSF to support the reporting on stock status at regional and global level. In particular it was discussed the addition of further indicators, i.e. biomass, fishing mortality, MSY, CPUE series, and biological parameters.

The future ownership and operation of the GRSF after the project deliverable were discussed in terms of sustainability and governance. Terms of reference were compiled for the design of a GRSF business model. The consultant will start from the various options depicted during the meeting.

## ABBREVIATIONS AND ACRONYMS

ASFIS	List of Species for Fishery Statistics Purposes
COMET	COnccept Mapping Engine and Tools
CCAMLR	Commission for the Conservation of Antarctic Marine Living Resources
CNR	Consiglio Nazionale delle Ricerche
CPUE	Catch Per Unit Effort
CWP	Coordinating Working Party on Fishery Statistics
DFO	Fisheries and Oceans Canada (Department of)
DSD	Data Structure Definition
EAB	External Advisory Board
EEZ	Exclusive Economic Zone
FAO	Food and Agriculture Organization of the United Nation
FIAF	Marine and Inland Fisheries Branch
FIAS	Statistics and Information Branch
FIRMS	Fisheries and Resources Monitoring System
FORTH	Foundation for Research and Technology - Hellas
GFCM	General Fisheries Commission for the Mediterranean
GRSF	Global Record of Stocks and Fisheries
IATTC	Inter-American Tropical Tuna Commission
ICCAT	International Commission for the Conservation of Atlantic Tunas
ICES	International Council for the Exploration of the Sea
ISSCFG	International Standard Statistical Classification of Fishing Gear
MDM	Master Data Management
MSY	Maximum Sustainable Yield
NAFO	Northwest Atlantic Fisheries Organization
NetCDF	Network Common Data Form
NGO	Non-Governmental Organization
RFB	Regional Fishery Body
RFMO	Regional Fisheries Management Organisation
SDMX	Statistical Data and Metadata eXchange
SFP	Sustainable Fisheries Partnership
SOFIA	The State of World Fisheries and Aquaculture
SSOT	Single source of truth
TORs	Terms of Reference
TWG	Technical Working Group
UN/CEFACT	United Nations Centre for Trade Facilitation and Electronic Business
UUID	Universally Unique Identifier
VRE	Virtual Research Environment
WoRMS	World Register of Marine Species

## **I. WELCOME ADDRESS AND ADOPTION OF AGENDA (Agenda item 1)**

1. The TWG second session (TWG2) was held at FAO headquarters in Rome, Italy, 28 February–2 March 2017. Mr Marc Taconet, BlueBRIDGE EAB chair and FIRMS Secretary, opened the meeting at 9:00 hours on Tuesday 28 and welcomed the meeting participants. See Annex 2 for the full list of participants.
2. Mr Marc Taconet was nominated as Chair of this EAB TWG2 with David Ramm to provide support to the Chair and assisting as needed.
3. The agenda (see Annex 1) was approved by all participants and no additional items added.

## **II. THE GRSF PROTOTYPE, AN INTRODUCTION (Agenda item 2)**

4. Mr Gentile (FAO) presented the iMarine platform where two dedicated VREs (Virtual Research Environment) have been activated: the “GRSF” for the public users with approved records, and the “GRSF Admin” for authorized users enabled to manage and validate the GRSF records. Both VREs are based on the iMarine Data Catalogue powered by the open source CKAN software.
5. The GRSF Knowledge Base is built with semantic web technologies using ontologies, and feeds the GRSF Admin VRE with both source records (harvested from FIRMS, FishSource and RAM databases) and “GRSF” records (generated upon the specific GRSF requirements compiled in the BlueBRIDGE project Wiki). The GRSF records include two domains: “stocks” and “fisheries”. A GRSF stock record can be qualified as 1) "Assessment unit", whenever a stock assessment is conducted, or 2) "Resource", when there is no stock assessment available. A GRSF Fishery record can be qualified as 1) a "Fishing activity", a record bound to the fishing operation aspects, or 2) a "Fishing description" (any other approach).
6. Currently, the “GRSF Admin” contains around 14 000 records including source records and GRSF records (around 9 000) in pending status (not validated). In the first prototype, the GRSF records have been generated without any standardization of the incoming information. For this reason the GRSF prototype is not able to detect overlapping records from the different database sources. Once the standards will be implemented, the overlapping records will be merged, and the final number of GRSF records will be lower.
7. The records can be filtered by pre-defined “groups”, which divide the records in subsets, and “tags”, which are linked to specific content (e.g. based on species or areas). A full text search is also available. The searches will retrieve a list of records with titles and icons linking to accompanying “Data and Resources” files in different formats (e.g. HTML, PDF, CSV) for fishery/stock indicators, source of information, web pages etc. Mr Taconet pointed out that, in the search results list, it would be convenient to display the database sources originating the GRSF records.
8. Mr Melnychuk pointed out that there are several source records that seem to be describing the same stock but have not been merged. Mr Gentile and Mr Minadakis explained that once the standardization process has been carried out, the application will be able to detect automatically overlapping source records and similar source records to facilitate the validation process and merging which ultimately will be carried out by designated experts.
9. From the list of records, a dedicated page can be opened (clicking on one particular item) with links to database source(s), source of information and indicators associated to the stock/fishery. These are contained in CSV files which can be downloaded. The interface also provides some online tool to display data through graphs and maps. Other information (record type, species, gear, flag state, state and trend, etc.) is also displayed together with the

Universally Unique Identifier (UUID) and the semantic identifier both generated automatically by the application.

10. A managing panel can be opened to perform actions on the records. For example, the GRSF short title can be edited and the status can be changed (e.g. approved, rejected, archived). Once a record is approved, it populates the public GRSF VRE. This might be accessed by registered users. If a record is rejected, an annotation explaining the underlying reasons could also be provided.

11. Regarding indicators for fishery records, Mr Gentile explained that in FIRMS many fisheries are described as an aggregation of fishing activities targeting multiple species, with associated aggregated indicators (i.e. catches). In GRSF, these fisheries have undergone a dissection process to produce records describing single fishing activities (one gear type, one flag state, one target species), but the indicators that were available in an aggregated form cannot be dissected accordingly. In addition, some of the fishing activities resulting from the dissection might not correspond to real fisheries. In this regard, it would be important to distinguish between the “existing fishing activities”, “non-evidence fishing activities” and the “impossible combinations of gear, species and flag state”. Those records that do not correspond to any real fishing activity will be eventually archived or removed. The rationale behind the dissection process was reminded by Mr Spear, who explained that the level of resolution and detail of fishing activities is important from a traceability point of view, for example for the identification of the different impact of the single fishing gears.

12. Mr Taconet asked whether the UUIDs are generated following an international standard protocol. Mr Pagano clarified that the UUIDs are generated by a standard algorithm and they are embedded in persistent URLs. When the records will be published, the URL will be different (being in a different VRE), but associated to the same UUID. The resolver (first part of the URL) can be partially customized including the name of the organization as the entity responsible for the GRSF catalogue.

13. Mr Melnychuk asked whether UUIDs for two similar stocks would be similar. Mr Pagano answered that the UUIDs do not reflect similarities in the records content.

14. Regarding the short title displayed upfront in the record page, it was recalled that it should be user friendly and manually entered according to specific naming conventions (e.g. FIRMS standards).

### **III. GRSF STANDARDS AND GUIDELINES: FINALIZING AND IMPLEMENTATION ROAD MAP (Agenda item 3)**

15. Mr Gentile delivered a presentation regarding standards to be adopted for the different fields of the GRSF records. Approved GRSF standards must include specifications on code system, codes and titles (in multiple languages).

16. The fields constituting the records have been divided into time-independent (i.e. the fields for the identification of a stock or fishery) and time-dependent (indicators, such as catches or state and trend values). The standards for the time-independent fields are critical for the identification of the unique GRSF records.

17. Mr Melnychuk asked whether some kind of prioritization is made when there are multiple indicators available for a single field (for example landings and catches). Mr Gentile answered that no prioritization was made, as the system is collating all the available information. Mr Taconet added that specific qualifiers should be adopted to distinguish among different types of indicators. Mr Melnychuk further clarified that in some cases both landings and total catches are collected although, depending on the case, one of the two may be preferable.

18. For the field “aquatic species”, ASFIS and WoRMS (when ASFIS is not available) were accepted as standards. For those limited cases with scientific names not found within neither of the proposed standards, mapping should be provided. Mr Melnychuk confirmed that the

RAM database can have few non-standard scientific names (probably around 2 percent), and he can commit to provide mapping. The original scientific name (as provided by the source of information) will be kept within the RAM database.

19. Ms Segurado warned about the risk of creating duplicate records in case of different scientific names for the same species in ASFIS and WoRMS. It was clarified that the scientific name is taken from the sources as it is, handled as a simple text string for the overlapping detection process, and finally, experts will be responsible for validating this aspect when approving a record. Mr Ramm suggested to use WoRMS list of synonyms for mapping.

20. It was clarified that also standard names for higher taxonomic level (e.g. genus or family) are accepted in the GRSF. The GRSF ontology should capture taxonomic relationships enabling browsing the catalogue by the different taxonomic levels.

21. For the “assessment/distribution area” and “fishing area” fields, the following standards were accepted: FAO fishing area, EEZ and RFBs geographic system (e.g. ICCAT management unit, GFCM geographical sub-areas, IATTC Pacific tuna reporting area). The sources of such classifications are FAO, Marine Regions, and RFBs.

22. For “jurisdiction area” the following standards were accepted: EEZ, RFBs area of competence, and “High seas” (in case of no management authority).

23. Mr Blondel confirmed that the Marine Regions identifiers are used to reference EEZs and an internal FAO mapping is maintained for their relation with the ISO3 country codes.

24. Mr Spear explained that SFP is willing to develop a standard methodology to describe management units in relation to species, areas and management entities. The management unit (which can go finer than the jurisdiction area) is the finest level at which a fishery is managed following a consistent set of rules. Particular attention will be given to overlapping on management responsibilities, sometimes well formalized (e.g. bilateral agreements and treaties) sometimes less. There would be also funding opportunities for such task, with focus on the environmental sustainability for buyers consideration, and once completed could be used by the GRSF as well as by FAO.

25. Mr Taconet invited SFP representatives and the GRSF team to elaborate a GRSF definition for management units, in relation to jurisdiction areas and fishing areas. He also clarified that FAO cannot manage local standards and does not commit to maintaining them for GRSF. If national standards are to be considered, commitment to maintain them should be identified among partners.

26. For the “fishing gear” the following standards were accepted: the ISSCFG classification and the “SFP fishing gear classification”. Ms Segurado explained that “local” codes were added in the FishSource database to refer to particular gears which are important to differentiate gear-specific impacts (e.g. purse seine with and without FAD). Mr Spear confirmed that SFP commits to compile and maintain a list of such additional gears to be included as a GRSF standard. The SFP codes should be cross checked with existing authoritative classifications, such as tuna RFMOs code lists. The group also concluded that new gear codes should be submitted to CWP for consideration.

27. Mr Minadakis pointed out that, similarly to other fields accepting multiple standards, also the fishing gear code will be qualified with the acronym of the two recognized standards (i.e. ISSCFG, SFP).

28. For the “flag state” the following standards were accepted: ISO3 Country code (including overseas territories) and name. A discussion was held about the proper identification of the flag state regarding catch nationality. It was recommended to follow the CWP indications (i.e. the flag of the fishing vessel is the best available criterion for the assignment of nationality to catch and landings).

29. In relation to “production system type”, due to the heterogeneity of the controlled terms and underlying definitions the group concluded that the field will be kept and filled upon availability of information. However, this field is no longer mandatory, will not be used for the semantic identifier, nor it will be considered for the dissection process. Nonetheless, during the validation process, the GRSF validator should take into account this information when available, particularly to distinguish industrial fishery vs. artisanal.

30. Regarding the “state and trend indicators” for GRSF stocks no standards were retained at this stage. However, four specific fields have been created to host qualitative and quantitative information for abundance level and exploitation rate and will be populated with data as harvested from the database sources.

31. Regarding “fishery indicators” for GRSF fisheries, as previously highlighted, the group concluded that they would be maintained at source level, without being associated to the dissected GRSF fishery records. This implies that it will not be possible to filter those GRSF fisheries generated by the dissection process according to a specific catch level, although they will be accessible from the source records.

32. Regarding “management entity”, Mr Gentile reminded that so far there is no standard identified for this field, and the information is stored as a text string in all database sources. The group concluded that a roster for management entities (international and national organizations and treaties) should be created, including English acronyms and country of belonging, and should be adopted/mapped/maintained by partners. Mr Taconet pointed out that FAO has a list of management entities (initially compiled for mailing purposes to the member countries) which could be shared (upon FAO clearance) and used as a first draft of the roster. It was reminded that the “management entity” field accepts multiple values and complements the jurisdiction area information, as different management entities might have the same jurisdiction area but mandates on different marine resources. The group confirmed that the management entity value(s) will be part of the fishery semantic identifiers, while it is not needed for stocks.

33. The discussion continued on how to apply the agreed GRSF standards to the source data. Since it is not easy for the partners to update their data, at least in the short term, the group agreed that “mapping” local standards/classifications to agreed GRSF standards would be the way to harmonize the content of the GRSF knowledge base and to identify overlaps and similarities (proximities).

34. In particular, for areas, a pilot exercise was conducted in January and February 2017 mapping the RAM local area codes to international and national standards. As a general rule it was agreed that, when this kind of mapping cannot be provided, bounding box coordinates will be utilized for the detection of overlapping/similar codes and the local area code will be retained in the GRSF metadata.

35. In order to facilitate the validation process and the suggestion of possible merges, proximity rules compiled to identify similar records were presented. Species, area and gear are the fields for assessing the similarities. It was suggested that a ranking indicating the degree of the similarity could be associated to further facilitate the validation process.

36. The semantic identifiers were described. These are human readable codes and labels based on GRSF standards and associated to each GRSF record. The semantic identifier will be used to univocally describe each GRSF record (identity metadata), and will be filled with GRSF standard codes or, when not possible, with local codes.

37. Mr Taconet reminded that GRSF should aim at promoting the use of standards, therefore the international standards, when available, should be preferred to any other local or ad hoc classification.

38. Regarding traceability, the semantic identifiers will contain the codes for both fishing area and jurisdiction area. In addition, FishSource will provide the definition for the two types



of area in relation to the traceability mechanism. It was also clarified that traceability flag is only applied to fisheries, specifically to those fishing activities satisfying the following requirements: one fishing gear, one flag state, one targeted stock.

39. Regarding the proposals for implementation of GRSF standards, Mr Minadakis delivered a presentation on four different options on how to handle source data and their mapping to populate the GRSF records under the agreed standards.

40. The option of handling the mapping rules as “same as” relationship, without altering the source data, was considered the most suitable for implementation in the GRSF. The advantages of this option are that no changes in the source records are required and the “same as” relationships can be created at any time to be included in the GRSF. The disadvantages are that the absence of “same as” relationship means the total absence of mapping. In addition, the matching algorithm needs to be executed after every new insertion.

41. There could be cases in which some automatic transformations and harmonisations of the source’s data may take place to simplify the mapping compilation. The MDM initiative is also expected to provide support in this regard. Other mapping relationships (e.g. hierarchies, intersections) should be also handled to detect similarities.

42. Mr Minadakis also clarified that the “same as” relationships will only be applied when an exact correspondence among codes has been identified.

43. It was clarified that the geographic information needed to implement the proximity rules will be initially stored and handled within the GRSF knowledge base. Later on a central reference data repository could be created within the iMarine infrastructure. Mr Melnychuk warned that the bounding boxes provided by RAM were hand drawn and specifically conceived only to approximately identify location of assessment areas.

#### **IV. MASTER DATA MANAGEMENT (Agenda item 4)**

44. Mr Charef gave a presentation on Master Data Management (MDM). The presentation covered concurrent activities within the BlueBRIDGE project, the Coordinating Working Party on Fishery Statistics (CWP) and FAO in relation to the MDM concept that improve data exchange and interoperability in conformity with standards.

45. The MDM is an overarching concept that covers: - harmonization and design of global reference data as single-source-of-truth (SSOT) for international standard classifications; - formulation of data and metadata structures to facilitate interoperability between bodies/institutions; - support to standards such as NetCDF, SDMX or UN/CEFACT and assist in adoption of tools and facilities.

46. Coordinating Working Party on Fishery Statistics (since 1960) provides a mechanism to coordinate fishery statistical programmes of regional fishery bodies and other intergovernmental organizations with a remit for fishery statistics. A Reference Harmonization group has been set-up by CWP-25 with the objective to develop global DSD for catch and aquaculture production.

47. FAO is playing a leading role in handling and maintaining several CWP standards. Besides current ways of dissemination (e.g. FishStatJ, FAO Yearbook of Fishery and Aquaculture Statistics), FAO is making progress to publish data series and related metadata according to the Data Structure Definition (DSD) model. The FAO DSDs of statistical data collections will be made available in CSV format, including the full code lists of reference data. Other formats will be used, like SDMX-ML, to smoothen the data interoperability between institutions (e.g. FAO, Eurostat, ICES).

48. The MDM functional workflow manages all steps from input data to data processing and dissemination, taking into consideration the need of metadata governance and the identification of single source of truth. In the context of GRSF, the MDM workflow offers and

recommends services and tools to be used in the case of geographical (e.g DFO areas, RAM bounding box coordinates,..) and non-geographical data (e.g WORMS species list, country list,..). For example, the Bionym/COMET tool is a flexible approach to match taxon names, and this could be useful for the needs of GRSF for species codes harmonization, with particular reference to the ASFIS-WORMS mapping that will be published. The MDM efforts on harmonization of coding system (species, gears, areas) for the Tuna Atlas will also be made available to fulfil GRSF needs. Similarly, other MDM services and tools can be made ready to the GRSF use cases upon expressed needs. The workflow envisages a reference registry where the output mappings and code lists are stored to be harvested and used as input.

49. The MDM governance encompasses mainly the maintenance of the standards updates and the technology, the actors involved and their roles, the frequency of updates and the traceability of data from source of truth to end-user.

50. As for GRSF needs, Mr Taconet pointed out that the standardization of the codes at source level is still ideally the best solution. Also, mapping to national classifications should be attempted and preferred to the creation of new coding systems. However, the national/local standards should be verified and commitment to maintain and import any update should be made by the partners.

51. In response to Mr Spear about how frequently the GRSF geospatial information should be updated from the database sources, it was suggested that the iMarine infrastructure could be the possible repository to store FishSource geodata, in order to minimize the workload in case of frequent updates. In such case, Mr Pagano highlighted that the policies and procedures governing the data storage, maintenance, process and update would have to be clarified.

52. Mr Gentile asked whether the next version of RAM will use the mapping done so far to manage their area codes and align them to the standard or local codes agreed in the GRSF. Mr Melnychuk replied that most of the time, the codes they are using correspond to national standards, and where this is not the case RAM does not have the capacity to carry out that process at the moment.

## **V. DATA VALIDATION WORKFLOW - GRSF RECORDS PUBLICATION (Agenda item 5)**

53. Ms Gorelli gave a presentation on the GRSF data validation workflow for a) identification of unique stocks and fisheries, and b) validation of records content.

54. Regarding the unique identification of stocks and fisheries, different workflows were proposed for: overlapping records, unique records and similar records. These groups of records will be identified after the standardisation process (i.e. implementation of mapping rules) and the application of proximity rules.

55. As for overlapping source records, the group concluded that they should be merged into one GRSF record and that the time dependent indicators, coming from the different database sources, should be collated and associated to the record. In addition, features to allow the user to retrieve specific series of data should be made available at the interface level.

56. Mr Pagano confirmed that this option is feasible from a technical point of view and a collation of data with clear distinction of sources can be made available in a single CSV file, while enabling filtering options for users. UUIDs for the merged records will be discarded and a new UUID will be associated to the new GRSF record.

57. Ms Gorelli proposed that the collated data for multiple stocks could be also organized in a single CSV file (downloadable), with the different indicators stacked. The user would be

able to request from the application specific indicators for specific stocks to be included in the file. Mr Pagano agreed to check with CNR colleagues if this proposal can be implemented.

58. Ms Segurado pointed out that, if different series of indicators relative to different assessments are included, a new field to identify the different assessments should be added as well.

59. Regarding unique source records, i.e. source records that do not overlap with any other, the group concluded that there is no need to apply any particular procedure for the unique identification of stocks/fisheries.

60. For similar source records, i.e. records identified implementing the proximity rules, the first step in the validation implies assessing the relationship among those records. Similarities among records might be the result of the monitoring evolution of assessment units, as well as the result of imprecise mapping rules.

61. Mr Melnychuk added that partially overlapping assessment units might also result when two different agencies are performing an assessment on the same stock. He further explained that when they come across such situation in RAM they usually delete one of the two assessment units in order to avoid double counting. Mr Gutierrez suggested that, in such case, the similar records in GRSF might be flagged and accompanied with a note explaining to the user the issue and providing links to the other similar records. This way, the users would be enabled to take a decision according to the type of data/analysis they are interested in.

62. About tracking evolution of GRSF records, Mr Gentile explained that FIRMS is tracking the monitoring evolution of the stocks and the life cycle of fisheries with a specific data model. This functionality wasn't imported in GRSF since the other data providers do not offer such facility. Therefore, when records are partially overlapping because of monitoring evolution, the no longer monitored stocks will be simply flagged as "archived".

63. The issue of the records content validation was discussed by the group. Ms Gorelli recalled that the information is assumed to be correctly imported from the database sources. The validation process implies checking if: the GRSF record type has been correctly assigned; proper links to the sources of information are available; the short title is compliant with the GRSF naming convention. In addition, the application will provide a facility to annotate data inconsistencies that might arise (in the "manage item" panel). The group agreed that such inconsistencies will be first addressed internally (i.e. reporting back to the owners of the database sources). If this is not possible, a warning flag will be added to the record, warning the user on the uncertainty in the data.

64. Mr Taconet made a final remark about the operational aspect of the publication, which would require the working group to commit to validate the records and address related issues.

## **VI. UNIQUE IDENTIFIERS FOR FISHERIES – INTERFACE REQUIREMENTS TO ACCESS RELEVANT INFORMATION (Agenda item 6)**

65. Ms Segurado delivered a presentation on User Interface Requirements. A main factor in ensuring the sustainability of the GRSF will be that it meets the users' needs. Multiple and diverse types of end users are expected to be looking for the type of information that GRSF can provide. These encompass: i) data managers (which include current and future data contributors), ii) natural resource managers, iii) fisheries and environmental scientists, iv) biologists and taxonomists, and elements of the seafood industry, among them: v) consumers, vi). NGOs and traceability companies, vii) retailers/food-service companies, viii) processors/suppliers, and ix) Fishers.

66. Data managers were defined as the owners of the database sources (FIRMS, RAM legacy, FishSource) and other/future admins. They are a particular type of user being both

user and contributor. The use cases for data managers include dataset updating and retrieve/download selected information on GRSF active and archived records.

67. In analysing user groups, the meeting participants focused on what type of content each will be interested in, what type of interface will optimise its retrieval, and how to ensure that GRSF is the authoritative source for this type of information. Three different forms of user interface would seem to meet most needs: a) a non-technical User Interface mainly for returning single records, but also for submitting requests for new records, b) a feature that allows for filtering and downloading of larger datasets in a format such as a spreadsheet or database, and for consulting master lists, and c) an API that allows for connection to other systems. The potential for the services the GRSF can provide is very ample, as is the demand for these services. Decisions are needed on how much should be done by the GRSF itself and which services should be left to third parties to provide.

68. Mr Pagano expressed the need to define priorities, among use cases, as well to define incremental steps for implementation. He suggested to focus first on the GRSF data managers use case, since all the other use cases depend on this one. Although worth considering, some of the use cases seem to go beyond the GRSF system (e.g. connection to other sources, etc). Some use cases, although relevant (e.g. evolution of GRSF records across time), seem to require additional information which is not handled at present in GRSF.

69. Regarding the addition of new records, Mr Melnychuk asked what the procedure would be. It was clarified that the new records will be added through the database sources, and will go through the validation process. There will be no addition directly to GRSF, although in the future (beyond the scope of this project) it is already envisaged the possibility to suggest candidates stocks/fisheries by third parties. In this case, the new submission will follow a specific GRSF approval procedure.

70. Regarding the update of existing records, it was clarified that the new time dependent information will be attached to an existing record without altering the UUID, and during the updating process, the record will go back to pending status for validation. Once the workflow is well established and tested, the updating process for existing records could be completely automatized without the need for human validation. Mr Taconet pointed out that the different database sources have different ways to handle updates of the time dependent data. For example, in FIRMS different updates of a record are organized in different observations, and these have not been captured in the GRSF (only the most recent update).

71. Regarding use cases for end users, Mr Taconet suggested that a number of requirements could be implemented by third parties, for example the setup of a mobile application based on GRSF API, or a dedicated map viewer. In terms of action, partners should review the use cases and see what would be essential for sustainability.

72. The group agreed that the use cases for the data manager group should be the first ones to be addressed.

## **VII. UNIQUE IDENTIFIERS FOR STOCKS - TOWARDS SUPPORT FOR THE COMPUTATION OF STOCK STATUS SUMMARIES (Agenda item 7)**

73. Ms Gorelli delivered a presentation on how GRSF can support the reporting on stock status at regional and global level, explaining the current FAO methodology for the preparation of The State of World Fisheries and Aquaculture report (SOFIA). Within this, four main steps were identified: collating data, assigning a stock status to the assessment units, assigning an average status to each species at regional level (aggregating and averaging status of single assessment units, weighed by their catch), and calculating the percentage of stocks in each stock status category.

74. Regarding the data needs for stock status computation, additional types of data were considered for including in GRSF, such as biomass and fishing mortality series (especially in

absence of reference points), CPUE and MSY. Mr Melnychuk confirmed that RAM is collecting this type of data, upon availability, but not for all stocks. Ms Segurado added that in FishSource CPUE is considered as measure of abundance if other information is lacking. Biological parameters were also considered (e.g. natural mortality, age/length of 50% maturity), but are not always available in the database sources.

75. Regarding the aggregation of assessment units weighing by catch level, Mr Melnychuk explained that RAM has attempted this approach, but the result can be highly dependent on depleted stocks. Therefore, weighing by MSY might be a valid alternative. The group agreed to include MSY as GRSF field (when available).

76. Mr Blasdale pointed out that RFMOs are not users but mostly providers of this kind of data. The stock status computation at regional level does not provide any additional information to fishery managers. Nonetheless, it is useful from a communication point of view and to make comparisons among different regions and fisheries.

77. The group recommended to change the GRSF field label “reference year” to “year” and use “reference year” for the new field needed to identify different assessment versions for the same assessment unit.

78. Mr Taconet asked how the level of update of the data in GRSF could be assessed. The reference year is an indicator, but in some cases assessments are not conducted every year, so an outdated reference year does not necessarily mean that the information provided is not the most recent available. The group concluded that an attempt should be made to obtain a schedule for the future assessments from the sources of information, together with each data update. This information should be obtained using a common format from the three database sources.

## **VIII. GOVERNANCE AND SUSTAINABILITY (Agenda item 8)**

79. Mr Taconet introduced the discussion on GRSF governance and sustainability, highlighting in particular the following issues: content governance (ref. iMarine Data Access and Sharing Policy), system governance (storage, data preservation), and sustainability business model. The iMarine White Paper was recalled as a model for a public led partnership with MoUs to operationalize collaborations and with market strategies to generate incomes.

80. Mr Pagano highlighted the importance of the services provided in support of GRSF and confirmed that CNR is willing to fulfil that role ensuring the implementation of up-to-date technology in the long term. Other services targeting market/business oriented companies would be required on top of the UUID generation. CNR would be available to discuss such services with private companies.

81. Mr Ellenbroek asked about the process for a new contributor to join the GRSF. Mr Pagano replied that, from a technical viewpoint, the SLA signed with FORTH enables the use of their software to build the GRSF knowledge base, but not the human effort required for future updates. Hence the importance of standard protocols to reduce the update burden.

82. A discussion was held in regard to the GRSF standards and the handling of master data and how the FAO MDM initiative could support the GRSF needs. Partners are invited to contribute to the building, maintaining and promoting of the GRSF standards. Those standards maintained by partners may evolve into an MDM data service which should also include a registry of the available standards and mapping.

83. The group suggested to inform the Coordinating Working Party on Fishery Statistics (CWP) at its next meeting on the GRSF work on standards.

84. In terms of governance and in analogy with FIRMS, the GRSF secretariat may take the role for the validation of records (with the help of experts), while major decisions on standards and procedures may be tackled by the GRSF board, constituted by representatives

of the Partners. It was also remarked that there would not be viable options for the GRSF secretariat without tangible income from external sources.

85. On legal aspects, the legal entity as custodian of the GRSF data and the legitimate copyright holder were considered. Mr Taconet suggested that FAO could be the legal entity of the GRSF, but this possibility needs to be cross checked with FAO Legal office.

86. Mr Spear delivered a presentation highlighting further questions and issues on governance and sustainability. In particular, he pointed out that governance and sustainability of the GRSF require decisions on a wide range of issues, including content governance, system governance, data sharing policies, and a sustainable business model(s). The presentation touched briefly on those areas, and focused more on some of the outstanding questions. For governance, those questions were mainly about 'who plays what roles?' Who will 'own' the GRSF, who will maintain the standards we're setting, and who will maintain the technology/services are all key questions that need agreed answers. For sustainability, it's about 'what business model or models can be built around GRSF services?' The presentation focused on services directed at the seafood industry and service companies (e.g., traceability/technology). These services include validating that a stock/fishery exists and providing a unique UUID (e.g., machine readable) that results from a global standard. There is evidence that a business model (or models) can be built around these services to generate income to support at least part of the GRSF and its key components.

87. Along the business model, Mr Pagano suggested to designate a company to ensure the communication and marketing targeting the seafood industry and the need on traceability. Mr Taconet recapped that the business plan should include two companies, one to generate generic services and the other one in close partnership to ensure the maintenance of the services.

88. The issue of payment for accessing GRSF data that has been generated from opened and free sources was addressed. Mr Melnychuk saw a risk that the current RAM volunteer contributors might not appreciate a GRSF commercial service built upon data that were provided free of charge. It was answered that the business model should also foresee ways to return benefits to the owners of the database sources and the data providers.

89. The group concluded that more discussion is needed to explore the various options depicted in the presentation on business model. A contractor can be identified and ToRs should be prepared with the GRSF team. The ToRs should include a deliverable for a completed business plan with market analysis of the business model with costs for different options (see annex 5 for the draft ToRs).

90. The presentation on governance and sustainability provides the blueprint of possible partnership arrangement, discussion should proceed along those concepts and the discussions emerged during the meeting.

## **IX. CONSOLIDATING WORKING GROUP CONCLUSIONS (Agenda item 9)**

91. This agenda item was dedicated to review and amend the recommendations according to the comments from participants, see full list in Annex 3. Several topics triggered further discussion. The draft guidelines for mapping to GRSF standards were revised according to the outcome of the discussions, see Annex 4. ToRs for a consultant to work on GRSF business model were also drafted, see Annex 5.

92. FishSource reiterated that, in order to be eligible for traceability, one fishing activity should link exclusively to one stock. This expected situation is one of the cases covered by GRSF where one assessment unit corresponds to one fishing activity under a specific management. The traceability flag will be applied only to those GRSF Fishery records satisfying this requirement and the UUID of the corresponding GRSF Stock record will be associated. Mr Blasdale noted that it could be a weakness of the system the impossibility to

describe a multi species fishery. It was clarified that this model was conceived to satisfy the criteria for traceability. The multi species fishery can be traced back through the link to its source of information.

93. The concepts of fishing area and jurisdiction area were further discussed. It was pointed out that the fishing area might be missing in the description of a fishery. Mr Blasdale suggested that the jurisdiction area could be the default one when no other information is available. In fact if the jurisdiction area is missing that fishery would be considered as an Illegal, Unreported and Unregulated (IUU) fishing activity and therefore not eligible to be included in the GRSF. In general, it was concluded that a record should not be validated without information on management entity and its jurisdiction area. Finally the GRSF team was asked to compile a common definition of the two area types.

94. Regarding data access and sharing policy, GRSF team was asked to compile any additional requirements or specification which are not covered in the iMarine Terms of Use.

95. With regard to governance, it was reaffirmed that the GRSF long term sustainability could be modelled upon the iMarine Board structure. GRSF ownership needs to be reconciled as part of this consideration.

96. The group expressed the need for a work plan to be agreed with CNR and FORTH colleagues in order to address the implementation of the recommendations summarized in Annex 3.

#### **10. ANY OTHER BUSINESS (Agenda item 10)**

97. No additional topics were discussed.

# Annex 1

## Meeting agenda

BlueBRIDGE External Advisory Board (EAB)  
Technical Working Group Meeting on Global Record of Stocks and Fisheries  
Second Session  
Rome, Italy, 28 February, 1 and 2 March 2017  
ANNOTATED AGENDA AND TIMETABLE

### Meeting place: FAO HQ, India Room (A327)

#### Day 1

All day: 09:00 to 17:30

Coffee breaks: 10:30-11; 15:30-15:45

Lunch break: 12:30-13:45

#### Tuesday morning (09:00 – 12:30)

##### 1. Welcome address and adoption of agenda

*A brief summary of the activities of the GRSF community will be also provided. Minutes of the conference calls and the GRSF documentation are accessible in the [GRSF wiki](#).\**

##### 2. The GRSF prototype, an introduction

*During the last months, progress were made for the development of the GRSF interface and a prototype has been released in a specific iMarine VRE.*

*In this VRE we can now access and manage stocks and fishery records from the three data contributors (FIRMS partners, RAM, and FishSource).*

*The VRE was presented at the BlueBRIDGE review meeting in Brussels with positive feedback from the Horizon 2020 reviewers.*

##### 3. GRSF Standards and guidelines: finalizing and implementation road map

*Introduce the standards for the GRSF fields of the database and mapping guidelines. The [GRSF database overview](#)[\*] documents the standards for each field. The [guidelines](#)\* for the area field and the unique semantic identifier have also been drafted.*

*The GRSF prototype now contains all heterogeneous codes found in the sources. To standardize the content of the GRSF we need to adopt shared standards. In particular see also the proposal for handling [GRSF spatial references](#)\*.*

*Participants are asked to comment the proposed guidelines and standards for each of the GRSF fields and the implications these may have on their data sources.*



Tuesday afternoon (13:45 – 17:30)

## **GRSF Standards and guidelines: finalizing and implementation road map**

(cont. from morning)

### **4. Master Data Management (MDM)**

*The FAO initiative on Master Data Management can support the GRSF need of standards with mapping of local and regional classifications to global standards within FAO and iMarine platforms.*

*The available mapping and data services as well as possible developments will be presented for GRSF specific needs. The group will review the list of required code lists, and required mappings, including for the geographic dimension.*

*Participants are asked to review and amend the GRSF code lists (including geographic identifiers), and identify the mappings. Considering the proposed FAO and BlueBRIDGE facilities and data services, the group will formulate the needs which the MDM initiative could support. The group will also develop the plan for proceeding with actual mapping work.*

### Day 2

All day: 09:00 to 17:30

Coffee breaks: 10:40-11; 15:30-15:45

Lunch break: 12:30-13:45

Wednesday morning (09:00 – 12:30)

### **5. Data validation workflow – GRSF records publication**

*The GRSF data administrator needs to validate and to approve every GRSF record (stock or fishery) before it can be disseminated. The validation workflow has to guarantee due diligence to ensure that the data owners policies are respected.*

*Moreover, time dependant indicators related to the published GRSF stocks records (e.g. catch, stock status) might be available in one or more sources attached to the same record. How do we handle this information?*

*A scenario will be presented, the working group participants will review this scenario and finalize a workflow and make recommendations for its implementation.*

### **6. Unique identifiers for fisheries - interface requirements to access relevant information**

*The GRSF prototype includes the automatic creation of unique identifiers, there is a need to indicate the type of consultation/navigation services which should be accessible from the GRSF interface for users consulting the record.*

*To meet the GRSF objectives, and to ensure the sustainability of the system, an important factor is to meet the end users' needs.*

*The group will discuss a first version of these services.*

Wednesday afternoon (13:45 – 17:30)

**7. Unique identifiers for stocks – towards support for the computation of stock status summaries**

*What do we need to compute summaries of stock status (i.e. by species and area) by weighting the stock status indicators?*

*A first version of the rules and workflow that should drive the handling of such indicators will be discussed.*

**8. Governance and Sustainability**

*The future ownership and operation of the GRSF after its deliverable are open for suggestions. This includes the identification of the business model that is suitable for the stock/fishery unique ID service for seafood industry. The group will review and further elaborate the sustainability components and discuss options, including governance. These directions/options will constitute the basis for assigning a consultant during the following 12 months with the mission to depict scenarios, perform a feasibility study and cost/benefit analysis, and coordinate inputs among GRSF stakeholders.*

*Participants are asked to review the sustainability components prior to the meeting and come with suggestions. The group will agree on the road map including the consultant's TORs.*

Day 3

Half day: 09:00 to 12:30

Coffee breaks: 10:40-11

Thursday morning (09:00 – 12:30)

**9. Consolidating working group conclusions**

*This half day will be dedicated to consolidate all recommendations made by the participants addressing the various topics of the meeting. The output of the meeting will guide the further development of the GRSF and will contribute to the formulation of the exploitation plan.*

**10. Any other business**

---

[\*] Links are pointing to the GRSF Wiki with restricted access. Please use your credentials of the iMarine gateway, if you are not yet member of the GRSF VRE please contact us to be invited.

## Annex 2

### List of participants

#### **EAB Members**

##### **Food and Agriculture Organisation of the United Nations (FAO)**

Mr Marc TACONET  
BlueBRIDGE EAB Chair  
FIRMS Secretary  
Head, Statistics and Information Branch (FIAS), FAO  
Via delle Terme di Caracalla  
00153 Rome  
Italy  
Tel.: (+39) 06 57053799  
E-mail: [marc.taconet@fao.org](mailto:marc.taconet@fao.org)

##### **Northwest Atlantic Fisheries Organization (NAFO)**

Mr Tom BLASDALE  
Scientific Council Coordinator  
2 Morris Drive, P.O. Box 638  
Dartmouth, Nova Scotia  
Canada B2Y 3Y9  
Tel.: (+1-902) 468-5590 ext. 204  
E-mail: [tblasdale@nafo.int](mailto:tblasdale@nafo.int)

##### **Sustainable Fisheries Partnership (SFP) - FishSource**

Mr Braddock SPEAR  
Systems Division Director  
4348 Waialae Avenue #692, Honolulu, HI 96816 USA  
Tel.: (+1) 808-726-2582  
E-mail: [braddock.spear@sustainablefish.org](mailto:braddock.spear@sustainablefish.org)

Mr David RAMM  
Data Manager for the Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR)  
8 Acushla Court  
Sandy Bay  
Tasmania 7005  
Australia  
Tel.: +61 403 802280  
E-mail: [david.ramm@protonmail.org](mailto:david.ramm@protonmail.org)

## **GRSF Team**

### **Consiglio Nazionale delle Ricerche (CNR)**

Mr Pasquale PAGANO  
Research Staff  
Istituto di Scienza e Tecnologie dell'Informazione "Alessandro Faedo"  
Area della Ricerca CNR di Pisa  
Via G. Moruzzi 1  
56124 PISA - Italy  
Tel.: (+39) 050 621 2891  
E-mail: [pasquale.pagano@isti.cnr.it](mailto:pasquale.pagano@isti.cnr.it)

### **Food and Agriculture Organization of the United Nation (FAO)**

Mr Anton ELLENBROEK  
Fishery Officer (iMarine)  
BlueBRIDGE – Blue Assessment Work Package leader  
Statistics and Information Branch (FIAS), FAO  
Via delle Terme di Caracalla  
00153 Rome  
Italy  
Tel.: (+39) 06 57054029  
E-mail: [Anton.Ellenbroek@fao.org](mailto:Anton.Ellenbroek@fao.org)

Mr Aureliano GENTILE  
Information Manager  
Statistics and Information Branch (FIAS), FAO  
Via delle Terme di Caracalla  
00153 Rome  
Italy  
Tel.: (+39) 06 570 53754  
E-mail: [aureliano.gentile@fao.org](mailto:aureliano.gentile@fao.org)

Ms Giulia GORELLI  
Information Manager  
Statistics and Information Branch (FIAS), FAO  
Via delle Terme di Caracalla  
00153 Rome  
Italy  
Tel.: (+39) 06 57052174  
E-mail: [giulia.gorelli@fao.org](mailto:giulia.gorelli@fao.org)

### **Foundation for Research and Technology - Hellas (FORTH)**

Mr Nikos MINADAKIS  
Research and Development Engineer  
Institute of Computer Science  
N. Plastira 100, Vassilika Vouton, Greece 70013  
Iraklio, Crete, Greece  
E-mail: [minadakn@ics.forth.gr](mailto:minadakn@ics.forth.gr)

### **University of Washington - RAM Legacy Stock Assessment Database**

Mr Michael MELNYCHUK  
Research Scientist  
Hilborn lab  
School of Aquatic and Fishery Sciences  
University of Washington  
Box 355020, Seattle, WA 98195 USA  
E-mail: [mmel@u.washington.edu](mailto:mmel@u.washington.edu)

### **Sustainable Fisheries Partnership (SFP) - FishSource**

Ms Susana SEGURADO  
FishSource Director  
4348 Waialae Avenue #692, Honolulu, HI 96816 USA  
Tel.: (+1) 808-726-2582  
E-mail: [susana.segurado@sustainablefish.org](mailto:susana.segurado@sustainablefish.org)

### **Observers - Invited persons**

#### **Food and Agriculture Organization of the United Nations, Fisheries and Aquaculture Department (FAO-FI)**

Mr Emmanuel BLONDEL  
International Consultant - Geographic Information Systems in Fishery & Marine Sciences  
Statistics and Information Branch (FIAS), FAO  
Via delle Terme di Caracalla  
00153 Rome, Italy  
E-mail: [Emmanuel.Blondel@fao.org](mailto:Emmanuel.Blondel@fao.org)

Mr Aymen CHAREF  
Fishery Statistician  
Statistics and Information Branch (FIAS), FAO  
Via delle Terme di Caracalla  
00153 Rome, Italy  
Tel.: (+39) 06 57054095  
E-mail: [Aymen.Charef@fao.org](mailto:Aymen.Charef@fao.org)

Mr Nicolas L. GUTIERREZ  
Fishery Resources Officer  
Marine and Inland Fisheries Branch (FIAP), FAO  
Via delle Terme di Caracalla  
00153 Rome, Italy  
Tel.: (+39) 06 57056563  
E-mail: [Nicolas.Gutierrez@fao.org](mailto:Nicolas.Gutierrez@fao.org)

Ms Stefania SAVORE  
Fisheries Information Officer  
Statistics and Information Branch (FIAS), FAO  
Via delle Terme di Caracalla  
00153 Rome, Italy  
Tel.: (+39) 06 57056094  
E-mail: [stefania.savore@fao.org](mailto:stefania.savore@fao.org)

# Annex 3

## Recommendations by the EAB TWG2 on GRSF

### Standards for the fields of the GRSF records

#### 1. Data structure definitions

Approved GRSF standards must include specifications on code system, codes and titles (in multiple languages). Any non-standard code will be treated only as a text string and qualified as “unknown”. In case of multiple and overlapping standards, the preferred one will be that indicated by the competent authority.

#### 2. Aquatic Species

Recognized standards are: ASFIS (FAO 3Alpha code) and WoRMS (AphiaID) if ASFIS code is missing. Any other scientific name can be entered in the GRSF but this will be treated only as a text string and not qualified. The mapping between the two standards should be made available (i.e. through the iMarine TLO) for species name validation and populating the semantic identifier.

#### 3. Assessment/Distribution Area (assessment area for assessment units and distribution area for marine resources)

Recognized standards are: FAO Area (and its subdivisions), EEZ (from Marine Regions - MRGID), RFBs geographic systems (ICCAT management unit, GFCM geographical sub-areas, IATTC Pacific tuna reporting area, etc.). Bounding box coordinates are expected in absence of standard codes and mappings.

#### 4. Fishing Area

Same as Assessment/Distribution Area. When applicable, preferences for the same standards are adopted for univocally relate a fishery with an assessment unit. In case of absence of information the jurisdiction area is taken into account.

#### 5. Jurisdiction Area

Recognized standards are: EEZ (from Marine Regions - MRGID), RFBs areas of competence, and additional eligible local standards. Bounding box coordinates are expected in absence of standard codes and mappings. The team is invited to elaborate a GRSF definition for jurisdiction area, in relation to management units and fishing areas. In case of no management authority (no recognized jurisdiction area) the term “High seas” can be entered.

#### 6. Fishing Gears

Recognized standards are: ISSCFG and “SFP fishing gear classification”. The SFP codes should be cross checked with existing authoritative classifications, such as tuna RFMOs code lists.

#### 7. Flag state

Recognized standards are: ISO3 country code [Ref. to CWP Flag State nationality of the catch]

#### 8. Production System Type

This field is not mandatory, when available the information will be captured in the GRSF database but not used for GRSF title either for semantic identifier. Finally it is not considered among the default attributes to distinguish fisheries record, but GRSF admin should take into account this information when relevant (e.g for small scale fisheries).

## 9. State and trend of Marine Resource

No specific standard have been identified so far. However the GRSF record is designed to host qualitative/quantitative information in two dedicated fields for fishing pressure and abundance level. Two more fields are for standard qualitative information, so far based on the FIRMS bidimensional matrix. The GRSF should capture the time series of values in all available observations (i.e. FIRMS observations fact sheets).

## 10. Fishery data/indicators

No specific standard have been identified so far. The fishery indicators, associated to GRSF records generated by the dissection process, should not be associated to the GRSF record itself but to the source record (i.e. a FIRMS fishery fact sheet). There is the need to work out a generic standard scheme for handling fishery indicators. In addition, the term “fishery indicators” should be changed to “fishery metrics” as catches and landings are not indicators but data or estimates.

## 11. Management entity

A GRSF roster of management entities should be compiled specifying organization names, English acronyms (or local languages for national orgs), countries of belonging (as appropriate). The roster should include international, national and sub-national organizations, and treaties (and qualified in the semantic identifier). Policies to access the roster must be defined (e.g. for confidentiality issues).

## Unique Identifiers

### 12. UUIDs and Semantic Identifiers

The semantic identifier is populated with standard codes as direct entry or mapped or, where not available, with the local codes. The structure of each semantic identifier component should be made of the “standard qualifier (referencing the code system) plus the code”, i.e. “ISSCFG:.3.1.1” or “ISO3:BVI”. The qualifier for unknown code system is “NA” or “U” or equivalent (TBD).

At display level the components of the semantic identifier should be separated by a specific delimiter (e.g. “+”), empty codes should be also reported with a space or equivalent (TBD). Multiple values should be comma separated or equivalent (TBD). Further requirements specifications are in the GRSF requirements Wiki page.

For stocks: <Species> + <Assessment Area(s)>

For fisheries: <Species> + <Fishing Area(s)> + <Jurisdiction area(s)> + <Management Entity(ies)> + <Gear type> + <Flag State>

Examples:

**ASFIS:COD + FAO:21.3.M + RFB\_COMP:NAFO + GRSF-Org:INT:NAFO + ISSCFG:03.1.2 + ISO3:CAN**

**ASFIS:SLC + EZZ:BHS (or MRGID:8404) + ISO3:BHS + GRSF-Org:BHS:DMR + ISSCFG:20.0.0 + ISO3:BHS**

13. The UUID resolver (i.e. [http://data.d4science.org/uri-resolver/catalogue/GRSF\\_Admin/product/](http://data.d4science.org/uri-resolver/catalogue/GRSF_Admin/product/)) can be customized, the UUIDs persist in case of changes of the semantic identifiers (i.e. updates of a record)

14. DSDs for each standard need to be provided to the GRSF knowledge base, and must be used to align terminology in GRSF.

## Traceability

15. FishSource requirement for traceability does not accept the relationship “one fishery to many stocks”. A fishery can be flagged for traceability only if linked to one specific stock (assessment unit).
16. When this unique relationship exists, there should be the possibility to concatenate the fishery semantic identifier with the stock UUID. The GRSF Admin Management panel should give the option to flag/un-flag the record for traceability purposes and add/remove link to a specific stock UUID.

## Mapping to standards

17. **Source data harmonization** The standardization of the codes at source level is still the desirable solution. When possible, Mapping to standards is the way to handle source data when not compliant with the GRSF standards.
18. **Area codes vs. coordinates**  
Mapping for area codes are accepted for “same as relationship”, alternatively the coordinates of the bounding box should be provided if not detailed polygons coordinates.
19. **Other standards**
  - a) “Local” standards are eligible to be adopted as standard if not already covered by existing GRSF standards, and if commitment for its maintenance is provided.
  - b) The underlying local standards for RAM areas should be investigated, possibly liaising with the local agencies owning the local standard. A commitment to maintain those local classification should be ensured.
  - c) Requests to national and local agencies to provide their coding system should come from FAO. Identification of contacts to be established among partners.
  - d) Apply caution in importing too much detail for stock location in GRSF if the information is already available at source, as it can decrease user clarity.

## Master data management

20. FAO MDM initiative is welcome to support GRSF needs for standards and mapping between reference data and code lists used by GRSF.
21. CNR and FORTH are invited to confirm the proposed scenarios and to provide solutions to the unclear steps of the workflow.[Ref. to PPT presentation]
22. Under the MDM principles there should be the possibility to handle local classifications for area codes.
23. MDM data service should include a registry/catalogue to store, handle and make available the standards and code lists mappings.
24. iMarine VRE could be the appointed geoserver to store the geospatial data from FishSource. These data would be to support records validation not for dissemination.

## Records validation

25. **Overlapping source records**  
Among the various options presented, consensus was reached for the following: all the information from the different database sources is kept and collated in different fields.
26. **Unique source records**  
There is no need for particular protocols for unique stock/fishery identification.
27. **Similar source records - Proximity rules**  
In order to facilitate the validation process and identify similar records, the following



proximity rules will be implemented:

Stocks:

Same species + partially overlapping area

(Example: *Sardina pilchardus* in GSA 17 and *Sardina pilchardus* in GSA 17, 18)

Same species + adjacent area

(Example: *Sardina pilchardus* in GSA 17 and *Sardina pilchardus* in GSA 18)

Same genus + same / adjacent / partially overlapping area

Fisheries:

Same species / genus + same / partially overlapping / adjacent fishing area + same group of fishing gear.

Groups of similar records will be analysed by the validator to assess the relationship among them (e.g. monitoring evolution, partially overlapping assessment units).

Assessment units that are no longer monitored and have evolved into a different assessment unit should be flagged as archived. If assessment units are partially

overlapping, these should be annotated in a comment (available in the GRSF including the downloads) warning the user and linking to the overlapping records.

Hence the “Manage item” panel should contain an additional field for comments to users under selected categories (i.e. controlled terms: inconsistencies encountered; partial overlapping etc.).

The development team should propose a narrow/broader indication and ranking system for similarities to further facilitate the validation process (i.e. based on the percentage of area overlapping, species vs. genus, gear categories).

Therefore, in the Admin VRE, an interface highlighting groups of similar records should be made available to the validator. On the other hand, in the public GRSF VRE, a feature clearly marking partially overlapping records and archived records should be implemented. The catalogue should also offer the possibility to flag and browse records along the above proximities rules.

**28. Scientific names**

The validator should ensure that scientific names are compliant with ASFIS classification, and when not available with WoRMS.

**29. Source Content validation**

Inconsistency issues will be initially solved internally (by reporting back to the owners of the database sources), if this is not possible they will be flagged externally to users (as above described).

**30. Quality assurance**

The owners of the database sources should ensure that a quality control process is in place before submitting data to the GRSF.

**31. Handling different types of catch/landing data**

More than one series of catch/landing might be available (recreational, commercial, different estimations from different assessments) and all are potentially relevant. They should be reported adding a qualifier to distinguish between different catch series.

**32. Series of indicators from historical assessments**

Different series of indicators deriving from historical assessments can be associated to one assessment unit. These should be reported in the GRSF, since keeping track of historical data is important. A qualifier (reference year, see also following recommendation item) should be associated to identify the particular assessment associated to the different data series. Also, what RAM has identified as the “preferred values”, among those available from the different assessments should be identified.

For common usage the default display would be the most recent assessment among the multiple view options.

### 33. Downloads

Regarding the downloadable formats, the groups is asked to formulate requirements for the best structure of the file. In the file download proposal [Ref. GRSF-DataValidWorkflow-1Mar17-DownloadableDataset\_Proposal.xlsx], change “reference year” field to “year”, and use “reference year” (i.e. when the assessment has been conducted) for the new column to accommodate the qualifier to identify the different assessments for the same stock. Downloads should be accompanied with business metadata.

### GRSF Catalogue

34. The source records and the GRSF records need to be clearly distinguished in the search result page and in the record page.
35. In case of GRSF records in the results page, the sources (RAM, FishSource, FIRMS) should be acknowledged (i.e. by displaying the acronyms of the database sources)
36. The terminology between GRSF requirements document and the catalogue interface should be harmonized (see also recommendation on DSDs).
37. The GRSF ontology should capture taxonomic relationships enabling browsing the catalogue by the different taxonomic levels.

### Services/interface requirements for GRSF fisheries records

38. The data manager group should be the first one to serve and prioritization is required to indicate which use cases should be supported as priority and define incremental steps.
39. Prioritization within the class of end users should be also proposed by taking into account the business opportunities.
40. An interface to promote the GRSF content should be built.

### Services for reporting on stock status

#### 41. Additional fields needed

To meet the GRSF objective of supporting stock status summaries, it will be further elaborated with FAO colleagues on the opportunity to include further indicators, i.e. biomass, fishing mortality, MSY, CPUE series, and biological parameters.

#### 42. Aggregation

Averaging of stock status using catch data is highly dependent on depleted stocks. The possibility to weight using MSY should be assessed.

#### 43. Global reporting

The stocks used for the SOFIA assessment should be provided in the GRSF, as “marine resources” type. They should be linked clearly to the single assessment units that have been merged to create those higher level, aggregated records.

#### 44. Timeliness of updates

Owners of the database sources should identify a way to provide a schedule for each stock assessment to be used by system as qualifier for quality of report update.

### Governance and sustainability

45. The iMarine Data Access and Sharing Policy will be reviewed by GRSF partners, and complemented by GRSF policies based on specific needs and upon the GRSF services. [Ref. “Governance and Sustainability” PPT]
46. The iMarine White Paper will be reviewed by GRSF partners for suitability to a possible public driven partnership model.
47. The Business Metadata approach will be reviewed by GRSF partners to understand if it covers needs wrt Licences, copyright issues and sharing requirements.

48. The long term sustainability could be modeled upon the iMarine Board structure. GRSF ownership needs to be reconciled as part of this consideration.
49. CNR is the current service provider, and is interested to maintain that role. An SLA would be needed, also to cover the FORTH's service contribution.
50. The available SLA with FORTH covers the use of the software, but not the human effort required to maintain content and reference data. Formalizing procedures and protocols should be compiled to ease the maintenance effort.
51. MDM data service should include a registry of available standards and mapping.
52. The CWP will be informed on GRSF work on standards at its next meeting .
53. The GRSF Secretariat would have the role to decide on (meta)data validation and dissemination workflows. This may require in-kind contributions from the owners of the database sources. There would not be viable options for the GSRF secretariat without tangible income from external sources.
54. The GRSF Secretariat will also have a role to maintain the partnership (e.g. seeking new partners) and ensuring the consistency of the content.
55. A to-be-established board has to define the core business and scope of the GRSF, and the underlying services.
56. For future ownership and copyright holder, an example will have to be developed inspired by e.g. the FIRMS partnership and Applifish. (Ref. to GRSF policy)
57. The sustainability models need further discussion between BlueBRIDGE sustainability team and GRSF team. This can include the EGI initiative.
58. On legal entity, CNR can discuss data services with private companies that can sell services, market services, and maintenance.
59. On business model, more discussion is needed exploring the various options. A contractor can be identified, along the ToRs prepared with GRSF team (see Annex 5).
60. The PPT provides the blueprint of possible partnership arrangement, discussion should proceed along those concepts and the discussions emerged during the meeting.

#### **Actions for the GRSF Team**

61. Finalize draft guidelines for mapping to GRSF standards (GRSF Wiki).
62. Provide GRSF agreed definitions for fishing area and jurisdiction area.
63. About downloads, the group is asked to formulate requirements on the structure of the file.
64. In regard to quality assurance, a GRSF disclaimer for public users should be formulated.
65. Review the use cases on user interface requirements and identify what would be essential for sustainability.
66. Formulate specific GRSF VRE data access and sharing policies (e.g. citation).

#### **Work plan for updating the GRSF knowledge base**

67. A work plan needs to be elaborated to schedule the updates of the GRSF algorithm (by spring 2017), the re-generation of records (possibly 3 times by the end of the project), and the harvesting of new data from sources (possibly 3 times by the end of the project).

# Annex 4

## Draft guidelines for mapping to GRSF standards

Extracted from GRSF Wiki:

[https://support.d4science.org/projects/stocksandfisherieskb/wiki/GRSF\\_database\\_overview#Draft-guidelines-for-mapping-to-GRSF-standards](https://support.d4science.org/projects/stocksandfisherieskb/wiki/GRSF_database_overview#Draft-guidelines-for-mapping-to-GRSF-standards) (login required)

In discussing the different proposals for harmonization of data, the participants were invited to consider and keep distinguished the following processes:

- a) identification of standards at international, regional and national level;
- b) algorithm logic to facilitate the discovery of similar/overlapping records;
- c) guidelines on how to apply/implement the selected GRSF standard.

Approved GRSF standards must include specifications on code system, codes and titles (in multiple languages). Any non-standard code will be treated only as a text string and qualified as "unknown". In case of multiple and overlapping standards, the preferred one will be that indicated by the competent authority. Bounding box coordinates are expected in absence of no standard code and in absence of mappings.

### STOCKS

- Species Scientific Name - *Any other classification mapped to ASFIS (FAO 3Alpha code) or WoRMS (If ASFIS code is missing)*
- Assessment/Distribution Area - *Any other classification mapped as appropriate to EEZ (from Marine Regions - MRGID) or RFBs geographic system (ICCAT management unit, GFCM geographical sub-areas, IATTC Pacific tuna reporting area) or FAO Area (and its subdivisions).*

### FISHERIES

- Species Scientific Name - *Any other classification mapped to ASFIS or WoRMS (If ASFIS code is missing).*
- Fishing Area - *Any other classification mapped as appropriate to EEZ (from Marine Regions - MRGID) or RFBs geographic system (ICCAT management unit, GFCM geographical sub-areas, IATTC Pacific tuna reporting area) or FAO Area (and its subdivisions).*
- Jurisdiction Area - *Any other classification mapped as appropriate to EEZ (from Marine Regions - MRGID) or RFBs areas of competence.*
- Management entity - *A GRSF roster of organization names, English acronyms (or local languages for national orgs), countries of belonging (as appropriate). The roster should include international organizations, national and sub-national institutions, treaties.*
- Flag state - *any other classification mapped to ISO3 country code (plus custom solutions to manage exceptions).*
- Fishing gear - *any other classification mapped to ISSCFG classification or "SFP fishing gear classification".*

*Since detailed mapping rules are not available for all fields, the first GRSF prototype accepted in the GRSF fields any codes as provided in the sources. However, for a final version of the GRSF it is envisaged the updating or mapping of all classifications to the agreed GRSF standards above described (see paragraph below "First Implementation (Autumn 2016)").*

FIRMS, FishSource and RAM partially overlap in their data coverage. In some cases, GRSF records coming from different sources present complete or partial overlapping of information (regardless of the adopted standards). Final revision and validation of overlapping records and assignment of proximities will be done by expert.

### Detailed draft guidelines for mapping of area fields

Draft guidelines for mapping non-GRSF local area codes (at national or supra-national level) to agreed GRSF standards are hereafter specified.

**Assessment area/Fishing area:**

- For stocks identified at supra-national level (generally under RFBs), the standard should follow the geographic systems as defined by RFBs (including bilateral agreement): Statistical areas (e.g. CECAF), Geographic areas (e.g. GFCM), FMUs (e.g. ICCAT)
- In high seas, in absence of an RFB, FAO areas or coordinates for polygons/bounding box should be provided
- For stocks identified at national level, the EEZ identifier (from Marine Regions - MRGID) (e.g. WECAFC national stocks) and (if relevant) to RFB geographic systems or to FAO area (and its subdivisions)
- For stocks identified at sub-national level, the EEZ identifier or any other geographic system as defined by the concerned country, provided such system is approved as GRSF standard

**Jurisdiction area:**

- For fisheries identified at supra-national level: any local value should be mapped to the RFB competence area (when available) or 'High Seas'
- For fisheries identified at national level: any local value should be mapped to national jurisdiction (e.g. Australia EEZ , US EEZ) and (if relevant) to the concerned RFB area of competence (e.g. ICCAT, CCAMLR)

## **Annex 5**

# **Terms of Reference for the design of a GRSF business model**

- Support delivery of the GRSF policy in close collaboration with the partners, with special attention to protocols discussed and agreed by the second EAB-Working Group and consideration to compatibility with the general policy of the infrastructure and necessary extensions.
- Develop a sustainability business model. In close consultation with GRSF stakeholders and following EAB-WG2 guidance, assess feasibility of the three business model options. Support the development of a GRSF partnership arrangement.
- Develop a complete business plan, including a thorough market analysis with costs for different options. This task will inform the possibility of project funding, and it will serve to identify which end users should be prioritized, with what kind of services, and the best ways to promote the business (in terms of interface requirements).