Food and Agriculture Organization of the United Nations

FIAS/R1180 (En)

FAO
Fisheries and
Aquaculture Report

ISSN 2070-6987

Report of the

WORKSHOP ON STRATEGIC DATA POLICIES

Rome, 21–22 September 2015

Report of the

WORKSHOP ON STRATEGIC DATA POLICIES

Rome, 21–22 September 2015

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS
Rome, 2018

ISBN 978-92-5-109711-3

PREPARATION OF THIS DOCUMENT

This document is the report of the Workshop on Strategic Data Policies held in Rome from 21 to 22 September 2015.

FAO. 2018.
Report of the Workshop on Strategic Data Policies, Rome, Italy, 21–22 September 2015.
FAO Fisheries and Aquaculture Report No. 1180. Rome, Italy.

ABSTRACT

This document contains the report of the Strategic Data policy workshop held in Rome from 21 to 22 September. Originally targeting the South West Indian Ocean (SWIO) Region, it expended to a global scope with the SWIO Region as an example. As a consequence, its main objective was "Developing strategies and best practices for investments in an efficient data supply chain, of global value, building on the Indian Ocean situation used as case study. Working sessions aimed to analyze different cases though projects/initiatives presentations and discussions to identify what worked and what didn't to define general principles, identify challenges and propose strategies and best practices for cost-efficient and sustainable investments in fisheries data collection, sharing and utilization. Key principles were raised by the participants to keep tools, methodology and process in the data supply chain simple and to reuse as much as possible what already exists. Seventeen key points were identified during the discussion as common issues/common needs/general principles and were organized and developed in the Expert Document "Strategies and Best Practices".

CONTENTS

ABBREVIATIONS AND ACRONYMS

ABNJ	Areas beyond national jurisdiction
ANR	Agence nationale de la recherche
ASFIS	Aquatic Sciences and Fisheries Information System
AU	African Union
CLAV	Consolidated list of authorized vessels
CLS	Collecte localisation satellites
CNR	Consiglio Nazionale delle Ricerche
CMS	Content management system
CPUE	Catch per unit of effort
CWP	Coordinating Working Party on Fishery Statistics
DCF	Data Collection Framework
DCMAP	Data Collection Multi Annual Programme
DG MARE	EC Directorate-General for Maritime Affairs and Fisheries
EC	European Commission
EEZ	Exclusive economic zone
eRS	Electronic reporting system
ESA-IO	Eastern and Southern Africa and Indian Ocean (region)
FAO	Food and Agriculture Organization
FIRMS	Fisheries and Resources Monitoring System
FRP	Fishery Regional Portal
IFDMP	Integrated Fisheries Data Management Programme
IFREMER	Institut français de recherche pour l'exploitation de la mer
GBIF	Global Biodiversity Information Facilities
GRSF	Global Record of Stocks and Fisheries
ICCAT	International Commission for the Conservation of Atlantic Tunas
ICES	International Council for the Exploration of the Sea.
IMO	International Maritime Organization
IOC	Indian Ocean Commission
IOTC	Indian Ocean Tuna Commission
IRCS	International Radio Call Sign
IT	Information Technology(ies)
IRD	Institut de recherche pour le développement
IUU	Illegal, unreported and unregulated (fishing)
KMFRI	Kenya Marine and Fisheries Research Institute

MDM	Master Data Management
MoU	Memorandum of Understanding
NGO	Non-Governmental Organization
NSDS	National Strategy for the Development of Statistics
OGC	Open Geospatial Consortium
ORI	Oceanographic Research Institute
RDA	Research Data Alliance
RDB	Regional DataBase
REC	Regional Economic Commission
RFB	Regional Fishery Body
RFMO	Regional Fishery Management Organization
SC	Scientific Committee
SDMX	Statistical Data and Metadata eXchange
SEEA	System Environment Economic Accounting
SFA	Seychelles Fishing Authority
SIH	Système d'Information Halieutique
SOOS	*Southern Ocean Observing System*
SSOT	Single Source of Truth
SWIO	South West Indian Ocean
SWIOFC	South West Indian Ocean Fishery Commission
tRFMO	tuna Regional Fishery Management Organization
UBC	University of British Columbia
UN/CEFACT	UN Centre for Trade Facilitation and E-business
VMS	Vessel Monitoring System
WB	World Bank
VRE	Virtual Research Environment
WoRMS	World Register of Marine Species
WPFS	Working Party on Fisheries Data and Statistics
WPS	Web Processing Service
WoE	Weight of Evidence
XML	eXtensible Markup Language

EXECUTIVE SUMMARY

A Strategic Data policy workshop was organized in Rome, Italy on the 21st and 22nd of September, supported by the EU funded Indian Ocean Commission-SmartFish project.

Originally targeting the South West Indian Ocean (SWIO) Region, it expended to a global scope with the SWIO Region as an example. As a consequence, its main objective was “Developing strategies and best practices for investments in an efficient data supply chain, of global value, building on the Indian Ocean situation used as case study.”

Working sessions aimed to analyze different cases though projects/initiatives presentations and discussions to identify what worked and what didn’t to define general principles, identify challenges and propose strategies and best practices for cost-efficient and sustainable investments in fisheries data collection, sharing and utilization.

Key principles were raised by the participants to keep tools, methodology and process in the data supply chain simple and to reuse as much as possible what already exists.

Seventeen key points were identified during the discussion as common issues/common needs/general principles and were organized and developed in the Expert Document “Strategies and Best Practices”.

This document, main workshop output, will be circulated among the workshop participants for review and validation. Once completed, it will be presented at RFBs Commission meetings, reviewed and eventually endorsed as Commission’s policies.

1 INTRODUCTION

1.1 Background

The IOC-SmartFish project offers an opportunity for participating countries of Indian Ocean and related Regional Economic Commissions (RECs), Regional Fisheries Bodies (RFBs) and Non-Governmental Organizations (NGOs) to receive support from FAO for the implementation of a Regional Fishery Strategy for the Eastern and Southern Africa and Indian Ocean (ESA-IO) Region The project assists participating countries to strengthen national and regional fishery/aquaculture policy and strategy frameworks and to acquire additional knowledge on fisheries management for use in planning and monitoring. The project, now in its phase 2, is funded by the European Commission (EC) and support participating countries to fulfill international commitment towards responsible fisheries and to improve food security.

During phase 1, as a response to the SWIOFC 4th Scientific Committee (SC) meeting (December 2010) which recommended the development of an action plan towards the creation of a Regional Information System for the South West Indian Ocean region, the Chimaera portal[1] was developed[2]. This Fisheries Regional Portal (FRP) prototype is providing seamless access to information and data residing in existing international and regional systems (FIRMS, WIOFish and StatBase). It is exploiting the iMarine web semantics components and hosting solutions.

Some lessons were learned during the development of this portal:

- Chimaera pinpoints a fantastic potential for disseminating comprehensive set of knowledge on fisheries across the boundaries of source systems. The perspectives opened by Chimaera, as well as other innovative data sharing approaches, call for new data/knowledge investment strategies at the regional level: no need to duplicate efforts, instead look for complementary investment opportunities; develop partnerships aimed at pooling and sharing tools, data and expertise; design projects for jointly building comprehensive dashboards of indicators on the ecosystem approach, or the state of the marine environment.
- However few basic principles touching upon strategies and data policies should be recognized and agreed upon for realizing such potential, and ensuring sustainability of the information systems (for the benefits of the whole community);

Before further investing in Chimaera, it was decided under IOC-SmartFish (phase 2) that the first activity should be to organize a consultative meeting among regional investors and stakeholders to elaborate a regional strategy for cost-efficient and sustainable investments in fisheries data collection, sharing and utilization, for the South Western Indian Ocean. Such Strategy will be in line with the main principles defined in the broader Pan African Strategy on improvement of fisheries data collection, while extending it with a focus on data sharing and reutilization.

In order to ensure alignment with needs and expectations, this workshop should be led by a regional organization (e.g. SWIOFC, or IOTC) together with FIPS. The Workshop:

- is an opportunity to emphasize for the SWIO region what works well (best practices), what doesn't work so well (problems that need being addressed), and where are the needs in terms of regional data strategies and policies.
- will encompass issues of high level coordination among sponsors of regional data initiatives and practitioners able to influence regional outputs;

[1] http://chimaera.d4science.org

[2] An output of Result 1 of SmartFish project, component 1M3 aiming to "*improve knowledge and information in support to fisheries management processes*".

- should deal with basic/general principles, the operational part should be dealt with in separate technical (possibly global) workshop.

Given that SWIOFC couldn't attend and lead the workshop, its objectives were reviewed to take SWIO region as a case study and raise general principles with a global scope (not limited to the regional level).

1.2 Definition

- **Data**: data is understood in this document in its broadest definition, from detailed (raw) data collected in the field to processed statistics, including information (like fisheries inventories).
- **Data supply chain**: the fishery data supply chain is understood in this document as the chain of processes aiming to produce elaborated data or information from field data (mostly from the national level). It goes from data collection to data dissemination, including data collation, data storage, data processing and data exchange. It also encompasses data harmonization and data standardization concepts.
- **Blue Growth**: According to FAO Department of Fisheries and Aquaculture current working definition, Blue Growth is the sustainable growth and development emanating from economic activities in the oceans, wetlands and coastal zones, that minimize environmental degradation, biodiversity loss and unsustainable use of living aquatic resources, and maximize economic and social benefits

1.3 Workshop objectives

The main workshop objective is "developing strategies and best practices for investments in an efficient data supply chain, of global value, building on the Indian Ocean situation used as case study."

Its scope is limited to the fisheries data supply chain, including the following issues:

- data generation, data exchange (and reuse), use of data for analysis, data publishing;
- among national, regional and global levels;
- data under confidential requirements, as well as public data, including how to maximize valuation of the former into the later;
- related tools and information systems;

1.4 Workshop expected outputs and outcomes

Expected Workshops Output is (this) meeting report, including following annexes:

- List of SWIO data initiatives, positioned in the Data supply Chain (see figure 1)
- Expert Document "Strategies and Best Practices":
  - strategies and best practices for an efficient data supply chain;
  - components of data policy of global value;
  - extensions deemed necessary to existing instruments concerning the data supply chain

It should also include a roadmap for implementation and funding.

Envisaged Workshop outcomes are:

- Strategies and Best Practices expert document presented at RFBs Commission meetings, reviewed and eventually endorsed as Commission's policies
- The above process could eventually lead to extend the Pan African Strategy with implementation guidelines on Strategies and Best Practices
- Areas of collaborations identified and strengthened among stakeholders

2 WORKSHOP SESSIONS OUTCOMES

2.1 Introduction

As presented in the agenda (see appendix 1.2), the workshop was organized in 2 main phases:

- **A first phase of problem analysis to identify** what works (worked) well and what doesn't (didn't) work that well in the broad context of fisheries data in the South West Indian Ocean: this analysis was done through SWIO initiatives presentations during day 1
- **A second phase of fisheries data best practices analysis and identification** to define general principle, best practices and data policies that could be generalized to other regional contexts: that was the role of day 2 discussions enriched with presentations of past similar approaches, of innovative solutions and comparable initiatives.

2.2 Fisheries Data Supply chain in SWIO region

We are summarizing here presentations of initiative in the SWIO region, highlighting workshop main results/outputs of interest for the discussions on policies.

2.2.1 Problem setting: Chimaera

Yann Laurent presented the Chimaera implementation and the lessons learned which triggered the organization of the workshop.

Chimaera (http://chimaera.d4science.org) is the result of a SWIOFC Scientific Committee recommendation aiming to harmonize SWIO fisheries data available in 3 regional information systems, WIOFish, a fisheries inventory database (Knowledge base), FIRMS a global stock and Resource inventory with SWIO resources (knowledge base) and StatBase, a regional statistical database (statistics database).

The Regional Fisheries Portal building answered the need to create a unique entry point to these systems without creating a new centralized database. Implementing iMarine web semantics technologies and ontologies, Chimaera is based on an mechanism "à-la-Google" enhanced with the understanding of the underlying system semantics: the content of the existing systems is read the same way Google reads worldwide websites to index its content, with an additional understanding that a given term describes a certain concept (sardine is a target species for a given fisheries, but also a prey for tuna in a stock description). Once the content is ingested, Chimaera provides search tools, either simple generic searches or thematic searches exploiting the search indexes previously generated. There is no data exchange nor data storage in Chimaera.

Lessons learned during the course of the Regional Fisheries Portal development were twofold:

Chimaera doesn't own or store the data; it facilitates its accessibility by the users and increase visibility of source systems. Sustainability of such platform is deeply linked to the sustainability of the 3 source information systems: it is essential to keep visibility of the source system (through citation, sourcing in the portal, user is sent back to the portal), it implies that data are directly accessible through the web (data access and sharing policies) and that regular update of sources systems are applied, which is not the case (Cf. StatBase, endorsed by SWIOFC as regional statistical database but not updated since the end of the project in 2012). How can we ensure that new projects will enrich existing content rather than building new systems managing similar content?

Discussion:

Larry Oellermann raised the point that WIOFish team didn't understand initially that Chimaera will not store data but only search indexes. In that case, Chimaera is definitely a good complementary tool to WIOFish.

Key outcomes:

Not yet another central database

Raising visibility, knowledge of existing available SWIO fisheries data

Reusing, building on

Need to ensure sustainability of existing systems to guarantee Chimaera sustainability

Need to ensure core Chimaera management functions to ensure updates related operations

2.2.2 ORI – WIOFish

Larry Oellermann presented the WIOFish project.

WIOFish is an annotated inventory of fisheries for the western Indian Ocean. It attempts to describe each of the fisheries in as much detail as possible. Data collation is done using common template (33 questions) throughout regional yearly workshops of national experts.

Frequency is now every two years given funds constraints: no data are collected by WIOFish, information included in the database is collated from national reports/databases/local knowledge of scientists

The database was initially developed in 1999 when very little was known about the region's fisheries. It aimed at providing this information in a central database containing descriptive information for each fishery. Initially the focus was on small-scale fisheries but industrial fisheries have since been included. The system development is outsourced with a regular fee for maintenance and any changes/new features. It needs now to be migrated in something new and robust, and fee free.

It was not designed to share data which was one issue met in the Chimaera project. Knowledge is shared through the online database at www.wiofish.org and pdf reports published for each fishery.

Lessons learnt are:

- National expert turnover and need to retrain new participant in the next workshop
- Technical issues with slow internet in the countries
- Statistical data are very aggregated – lack of harmonization (in the species group) : difficult to compare
- An outsourced system is a huge technical and financial constraint

The question of historical records was after a question from Dirk Zeller. The system does not store historical data, only the last record uploaded in the last workshop is kept. The system was not designed for.

Marc Taconet highlighted WIOFish capacity to make public grey literature[3].

[3] Grey literature (or gray literature) is a type of information or research output produced by organizations, outside of commercial or academic publishing and distribution channels.

Key outcomes:

Content update only efficient during regional workshops: the key challenge is sustainable funding.

Benefit: bring grey literature buried in national offices to broad public access

Not designed to exchange with other systems – Not designed to archive historical data

Cost and constraints of an outsourced system

2.2.3 WB – SWIOFish

Xavier Vincent presented World Bank activities in the South West Indian Ocean (SWIO) Region.

The World Bank is supporting countries of the SWIO Region increase the economic, social and environmental benefits they extract from sustainable marine fisheries through the SWIOFish program. This long-term engagement seeks to improve the management of these fisheries, better harness the economic benefits, and foster regional collaboration.

Because efficient statistical systems are key to the sound management of fisheries sectors, SWIOFish supports countries in developing their statistical systems following three guiding principles: (i) Predefinition of realistic objectives and strategy; (ii) Adaptation of complexity and costs to national contexts; and (iii) Integration in national statistical systems and strategies.

SWIOFish also supports regional coordination on fisheries statistics, but focuses first on strengthening the national systems to ensure minimum capacity.

A discussion was held around the capacity/possibility to evaluate value for fisheries and the importance of the socio-economic indicator. There is a lack of standard for such assessment, Seychelles example could be extended to other cases. Showing how much money a fishery can generate could bring interest and investor to support fishery statistics and more generally speaking to the fishery sector.

Marc Taconet concluded that SWIOFISH investments are under SWIOFC governance, with national investments.

Key outcomes:

Crucial importance of National Statistical Systems: all data/information come from this level. Problem of scattered data in different national systems.

Importance of a regional approach to fisheries statistics

Underestimated importance of socio-economic indicators in National Statistical Systems (usually lacking) to assess the real importance of economic impact of fisheries sector (Economic dashboards)

2.2.4 FAO – toolbox framework facilitating integration of fishing operations related databases Including Sample based schemes / Open Artfish, and Tablets/mobile for OpenArtFish

Gertjan De Graff summarized FAO tools to support fishing operation monitoring.

Fisheries in Africa are dominated by small-scale Fisheries.

Due to its characteristics small-scale fisheries can only be monitored through sample based data collection schemes. To support and improve sample based fisheries FAO developed the following tools;

- A 10 days international training course (English and French) in fisheries statistics and data collection
- ARTFISH developed and implemented by FAO since the 90's, ARTFISH is a standard database to store and analyse data following appropriate statistical procedure
- Open ARTISH, recently developed to replace ARTFISH. Open ARTFISH is an MsAcces data base., with open source code, which can be easily adapted for local requirements
- Mobile phone application for small-scale fisheries data collection which is linked to Open ARTFISH

Discussion was held on local constraints for data collection in Africa. It was acknowledged that countries usually have limited resources with limited knowledge. Simplicity is the key to success, ArtFish and now OpenArtFish have been developed to cope with this objective. The importance to benefit from wide availability of Smartphones to develop tablet/smartphone based data collection has been agreed upon (example of Honduras/Cameroon). It still leaves the problem of data validation, GPS embedded in tablets/GPS can at least validate the location (in the field, not at home to fill-in on-line questionnaires!). SmartForms developed by iMarine pursues similar services with the prospects of cost efficiency achievable through the backing data infrastructure.

Key outcomes:

Everyone wants data, no one wants to pay for it

Fisheries statistics come from the national level.

In Africa, artisanal fisheries are dominant which monitoring requires human and financial resource. Which they don't have: keep monitoring simple.

Quick emergence of mobile data collection (OpenArtFish mobile collection, Honduras, SmartForms) with the risk linked to the tablets maintenance/replacement

2.2.5 IFREMER – SIH for Seychelles, Comoros

Emilie Leblond presented the *Système d'Informations Halieutiques* (Fisheries Information System - SIH). SIH is a French fisheries monitoring system for **scientific purposes only**, to supply data to the researchers: the tool has been designed by scientists to answer research and assessment questions asked by the social and economic issues. The system is developed using open source technologies as much as possible.

SIH aims to provide information and knowledge for fisheries monitoring, decision-making support. It covers the whole data supply chain for all French Fisheries (small-scale to industrial fisheries). SIH is built on the *Harmonie* integrated data management system, with harmonized classifications. Standard and classification are key to the SIH implementation in the complex French context.

It also implements French fisheries data access and sharing policies: all personal data are strictly confidential and can only be accessed by authorized users.

Key challenges are still sharing standard and common classification among all partners, the need to keep the system as user friendly as possible (keep interface simple to use), the traceability of data correction and the system funding

SIH for Seychelles: came from a request from the Seychelles Fishing Authority to the French Embassy. MoU for the 2007-2012 period has been signed between SFA and IFREMER. It resulted in a local implementation of the *Harmonie* database (2 IT volunteers until 2012), with the import of historical data. There has been a co-funding by IFREMER and SFA for the development of new data entry software for the catches and effort surveys on landing sites (2014-2015). And SFA has formally requested a MoU extension to re-launch SIH activities.

The main lessons learned was the need of continuity and stability of a staff dedicated to the project which was not the case.

SIH for Comoros: A request came in 2015 from the French Ministry of Foreign Affairs.
A SWOT analysis was conducted by IFREMER and independent consultants to define which countries would be the most appropriate to develop such systems.

A discussion was held first on the SIH scope: it is a system dedicated to scientific data, not routine data. Questions were asked on the need for IT skills to run such a system, especially in Comoros. The SIH is not only a tool but also a set of methodologies and standard. The system can be run on different database management systems.

Marc Taconet took the example of Comoros to discuss regional collaboration / coordination in providing support to countries support. Xavier Vincent made clear that responsibility to request support to statistical systems lies with the countries, not the World Bank. Regional coordination would be indeed needed; SWIOFC should be leading this regional coordination. WB can recommend solutions, but ultimately choice is with national authorities. Example of Seychelles shows that approach must be kept simple, past giant projects on big information systems failed.

Key outcomes:

SIH is a French national scientific data information system with a national instance tailored to Seychelles in the past years.

Standard, harmonization and classification are key to the system; it can manage different levels of classification (from national to international) but issues can raise when classifications are imposed by regional regulation but not well described (*case of ASFIS harmonization in line with DG MARE recommendations resulting with the upload of the complete ASFIS list, more than 12 000 records, in vessel e-log book system, making the system not usable by fishers. There was a lack of a regional subset of species defined by DG MARE*)

Use of open source tools

Keep user interface simple!

There are clear definitions of what is confidential or not.

Funding remains a challenge

The example of Seychelles shows that system can be tailored to a country and maintained live where there is a political commitment: the challenge is to have stable staff.

2.2.6 FAO: Global Record of Fishing Vessels, Refrigerated Transport Vessels and Supply Vessels

Dawn BORG COSTANZI presented the new Global Record programme.

The Global Record of Fishing Vessels, Refrigerated Transport Vessels and Supply Vessels is a global initiative that aims to improve information sharing on the fishing-related fleet to improve transparency and traceability, and, thus, prevent IUU fishing. The comprehensive information made publicly available will also support fisheries management and fact-based decision-making. By collecting and disseminating certified information from official authorities, as a one-stop shop for vessel, historical, authorization and MCS-related information, the Global Record will support international policy, particularly port inspectors carrying out risk analysis in implementation of the PSMA, and port and flag State authorities, amongst others.

The Global Record is a long-term programme comprising of system development, capacity development and awareness-raising, which is led by FAO Members through the Global Record Informal Open-Ended Technical and Advisory Working Group, guided, in turn, by Specialized Core Working Groups on particular technical matters. These groups are currently finalizing the requirements specifications, considering the data necessary to fight IUU fishing, in line with existing instruments, the information systems available globally, and the most simple and effective manner to facilitate data submission. In order to combat the main risk of lack of commitment and participation, the Global Record is being designed through a collaborative process, which gives members a strong sense of ownership and ensures that all needs are considered.

Strong background work on data standardization has taken place, with the Global Record collaborating with DG-MARE on XML-based customizable UN/CEFACT standards to exchange Vessel information. A system prototype, as a web application built on open-source technologies and based on the in-house VRMF, was developed by the Global Record team and presented at COFI 31 as a proof of concept, including restricted information from sample data providers. Currently at the important stage of effectively designing and developing the system, this prototype is being extended according to agreed data requirements, and to include data collection procedures, to move ahead with a pilot project that will make information from a few key member States publicly available through a web portal. The focus will then move to capacity development to increase coverage and ensure that numerous member States join the initiative, in order to close the global information gap on vessels, which currently makes control weak and creates a space for IUU fishing operations.

A quick discussion highlighted the need to make the distinction between data collection (collecting data from the field) and data collation (centralization of data collected otherwise).

Key outcomes:

Intend to collate harmonized and standardized records on vessels used for fishing and fishing-related activities: importance of using international standards, to share information with a community of users of vessel data, but also to enable simple exchange formats (csv).

Need to differentiate what is confidential from what is public.

Build on what exists for cost efficiency.

Challenge: lack of commitment and participation.

2.2.7 Overview of improvement of fisheries data collection in the SWIOFC region, including the Pan African strategy

Gertjan De Graaf summarizes the different FAO actions in improving data collection in the SWIO region.

The presentation provided an overview of developments for the improvement of data collection in the SWIOFC region and Africa.

The FAO Fisheries Department believes that working with countries is the only way to improve fishery statistics, primarily in order to meet national needs with regard to food security and fisheries management, but also to meet the needs of regional fishery bodies and FAO.

The overall framework to improve fisheries data collection is the "FAO Strategy for Improving Information on Status and Trends of Capture Fisheries (2004)". For Africa FAO supported the development the "pan-African Strategy on the improvement of fisheries and aquaculture data collection, analysis and dissemination (2014)". The major objective of the pan-African Strategy[4] is: to provide a framework and guidelines that should lead to improvements in the availability and quality of national and regional data to support fisheries management, aquaculture development and policy development in Africa. The pan-African Strategy is now a strategic document of the African Union and of its implementing agency (AU-IBAR) and should be used as a guideline for all activities related to improvement of fisheries statistics in Africa.

In collaboration with the SWIOFC Secretary, FAO supported the establishment of SWIOFC Working Party on Fisheries Data and Statistics (WPFS). WPFS organized three regional workshops (2007, 2008 and 2012). These workshops clearly identified the fisheries information needs and its gaps in the SWIOFC region.

A discussion was held on the importance of the national level for data collection from which all statistics (national/regional/international) are derived.

An important problem is that data collection is not attractive to donors.

It should be included in the National Strategy for the Development of Statistics (NSDS) as recommended by the African Union in the Pan African Strategy.

It should be routine data collection (not ad-hoc data collection funded by projects), simple and robust, aligned with the minimum core set of indicators defined in the Pan African Strategy. Agreed protocols (including selection of indicators) should be agreed upon to exchange data with other countries/regional and international organizations.

Further discussion highlighted the importance of the regional level for coordination and support: the SWIOFC Working Party on Statistics should be reactivated. SWIOFC should take responsibility and have the budget available.

[4] Available for download at www.au-ibar.org/component/jdownloads/finish/5/1958

Key outcomes:

Any fisheries statistics project in Africa should refer to the Pan African Strategy (first African reference strategy on fisheries statistics).

A minimum set of core fisheries indicators are defined in the pan-African Strategy and should serve as reference for any national statistical information system.

National fisheries statistical systems should be included in the national statistical system in the National Strategy for the Development of Statistics.

National data collection system crucial for high quality statistics

2.2.8 IOTC: CLAV

Fabio Fiorellato and Fernando Jara-Senn presented the IOTC Consolidated List of Authorized Vessels (CLAV)

A joint effort of the 5 tuna RFMOs (tRFMO), supported by the ABNJ Tuna Project: the CLAV is a successful example of how data from different sources can be automatically shared, collated and disseminated to the public audience, inherently increasing the value of the data itself once it is put in a broader, cross-organization context. As a tool to detect and deter IUU activities, the CLAV is crucial for port-state authorities, NGOs and all those officers and institutions that need a reliable instrument to cross check information through a single, constantly up-to-date entry point.

Further discussions pointed out improvements in vessels detail reporting such as IMO and IRCS.
Questions were asked about data exchanges, CLAV can expose data through web services, format obtained is XML but is not as of today UN/CEFACT compliant.
It was confirmed that EU fleet is part of the recorded vessels and that historical data are kept in the system.
Sachiko Tsuji highlighted that the same piece of software is used in others applications where it allows other vessels code addition.

Key outcomes:

CLAV is a central vessel repository for Tuna vessels with exchange protocols / data policy. It requires global/inter-agencies coordination about findings on duplicates and tRFMO action on data modification

Constraints: mostly come from absent mandatory fields, and lack of standards for some technical aspects

Need to institutionalize CLAV maintenance

2.2.9 SWIOFC – approaches to data collection and tools in support to Stock Assessment and Fishery Management (Weight of Evidence)

David Wilson presented the Weight of Evidence approach to stock assessment developed in the absence of complete assessments (Presentation created by Pedro Barros and Aubrey Harris):

- Many uncertain stocks do not have a reliable stock assessment for deriving status
- However, in many cases there are indicators and information to make a reasoned assessment of likely status
- We needed a process for making stock status determinations, in the absence of more complete assessments (data poor stocks and fisheries).
- Transparent (documented) and repeatable process
- Formalise the process and re-examine many of the 'uncertain stocks'

The WeO approach is done in 3 steps:

1. Describe the attributes of the species and fishery
2. Compile lines of evidence for status
3. Status Determination (weighing the evidence)
    - Overfished / Not overfished (Biomass)
    - Overfishing / Not overfishing (Fishing mortality)

Contribution of local experts is key to compile line of evidence. "One lesson learnt is that often, we know more than we thought."

The discussion highlighted the need for coordination of such an action at SWIOFC level.

Key outcomes:
Repeatable process, even for qualitative information.
Involve local experts to document all information available
It is a pragmatic approach: keep it simple

2.2.10 KMFRI – StatBase

Harisson Ong'anda presented the SWIO regional database StatBase: StatBase is a database centralizing the official national statistics of fisheries data from countries who are members of the South West Indian Ocean Fisheries Commission (SWIOFC). Information is voluntarily submitted with the condition that they become freely available for use. The target users include resource managers, scientists, sector professional and general public. The database was originally designed to support SWIOFP project. At the end of the project in 2013, Kenyan Marine Fisheries Research Institute (KMFRI) in Mombasa, Kenya was asked to host the activities on an interim basis. Presently the countries represented in the database are Kenya, Tanzania, Mozambique, South Africa, Seychelles, Mauritius, Comoros, France (La Reunion) and Madagascar. Two indicators may be obtained from the database, exploitation indicators (catch/effort/CPUE) and socio economic indicators (vessel register, frame survey data). FAO codes were applied to harmonize and standardize the data. The data currency is not uniform in the database due to differing frequencies of release of national statistics for respective countries. National terminologies may also sound similar but mean different things e.g. semi-industrial fisheries in one country may be similar to industrial fisheries in another country. Data entry sessions were most effective in a regional meeting set-up. There has been no funding to organize such a regional meeting since 2013.

Discussion was held on the definition harmonization of industrial Vs small-scale fisheries. It's not an easy task as definition varies from one country to another, sometime defined in the national Fishery Act. What is important is the capacity to describe it and to attach it to the data (importance of metadata).
It was acknowledged that StatBase lacks so far standard definition of indicators (each country published data the way they used to publish it nationally, not in a harmonized way. It would help to be able to run cross indicators queries in the system.

Key outcomes:

StatBase is full open data and open source technologies.

StatBase relied on workshop contribution by experts to feed the system: need for protocols to address timeliness across all contributors

StatBase collated all statistics available regardless indicator definitions. Harmonization process of national classifications to SWIOP classification (derived from international ones) was not an easy task: need for better harmonization, to readdress definition: there is a global need to a better implementation of metadata in the system.

Once funding is gone, no more update although StatBase has been endorsed as SWIOFC regional database: need to relaunch activities.

2.2.11 IOTC – integrated tuna statistical database

Fabio Fiorellato presented the new tuna statistical database.
Data are the most valuable asset for all the organizations dealing with fisheries management. But data *per-se* can lose value if they can't be shared and exchanged in a timely and flexible way. For this reason, IOTC revised the structure and functionalities of its core databases and took the necessary steps toward the adoption of an integrated solution that would encompass all the different datasets and data sources currently available, with the purpose of managing their content (according to the confidentiality policies in place at the organization) in a more harmonized way. The new IOTC Statistics Database is designed to streamline the data collection, data processing and data dissemination tasks, in order to react more quickly to the increasing needs for scientists and policy-makers to get timely and interactive access to the most current information available.

Key outcomes:

Need for integrated database given the current variety of reporting formats / existing databases in IOTC to foster data sharing and exchange

System will be developed with open source technology, no license cost

Confidential data management (implementation of data sharing policies)

Make simple user interfaces

2.2.12 FIRMS, towards the Global Record of Stock and Fisheries

Aureliano Gentile delivered a presentation on FIRMS and the coming activities towards the implementation of the Global Record of Stock and Fisheries.

Overall objective of the Global Record of Stocks and Fisheries is to provide scientists with an innovative environment supporting a collaborative, efficient and effective production of a comprehensive and shared global reference set of unique stocks and fisheries information that will boost regional and global environmental status monitoring as well as responsible consumer practices.

Within the Horizon 2020 - EU funded BlueBRIDGE project, FIRMS is leading such project component which will federate knowledge on status/trends of stocks and fisheries across various sources, and as such is expected to offer key services to stakeholders involved in "regional/global state of stocks indicators", as well as public and private actors involved in eco-labelling, traceability and sustainable fisheries. The scope is inventorying stocks and fisheries and status monitoring with focus on small scale and industrial fisheries within EEZ and in ABNJ, with a global coverage. Target audience includes scientists, policy makers, retailers, general public. Data collation and dissemination, knowledge management and sharing are the envisaged main activities. Fisheries and Resources Monitoring System (FIRMS), RAM Legacy Stock Assessment Database and FishSource are the main sources of information for the global record. The core database populated with inventories on stocks and fisheries will be complemented with information on stock assessment, fishing activities, post-harvest use, socio economic assessment, environmental assessment, management, status and trends.

The global record will be developed within the life-time of the project (30 months) and co-developed within the BlueBRIDGE partners and contributors possibly with open-source solutions. The data base will contain public data with free access and will be primarily released within the iMarine e-infrastructure. Rules agreed by partners will drive the access to data and web services. Sources of information will be always indicated.

The challenge is to constitute a critical mass of global information on status of stocks and fisheries, ensure timeliness of the updates and satisfy the target audience.

Key outcomes:

The Information management policy is developed under partners' governance.

Building on international Standard classifications

Ensuring ownership, participation, through tailoring for contributions, and return on investment

Simplicity upfront, complexity behind

2.2.13 IRD - Sardara tuna Atlas project and IRD - Observatoire de l'Océan Indien

Julien Barde presented the Sardara tuna atlas project and *Observatoire de l'Océan indien.*

In SWIO, IRD has been collecting data and conducting research studies related with fisheries since the early 1980s. Thereafter a big effort has been done on data management. Funding is mainly provided by IRD (e.g. scientific cruises), European Data Collection Framework (DCF)/ Data Collection Multi Annual Programme (DCMAP) for the monitoring of French purse seiners in SWIO, ANR (French National Research Agency, e.g. Emotion project), European Projects (iMarine, BlueBRIDGE), IOTC, and ISSF. IRD has a strong focus on tuna and tuna-like fisheries as well as on the pelagic species taken as bycatch (e.g. sharks, cetaceans, turtles,..).

Observations deal with many parameters required to support the implementation of an Ecosystem Approach to Fisheries Management: biological parameter of (by-)catch species (e.g. length, weight, stomach contents, isotopes, fatty acids, contaminants), fishing effort, drifting trajectories of Fish Aggregating Devices and fishing vessels (i.e. VMS), environmental parameters (satellite remote sensing data, acoustic, bio-geochemical and ecosystem model outputs, etc.).

Data are stored in open source Postgres/Postgis databases and mainly exploited through R scripts. Both data and processes can be accessed and combined online to calculate indicators. IRD is complying with well-known standards for (meta)data and related access protocols (OGC[5], TDWG/GBIF[6], SDMX[7]) and is keen to be part of a regional coordination network to improve data sharing in SWIO.

Two main projects are foreseen to foster coordination in the framework of the SWIO Workshop: Sardara Tuna Atlas and IRD *Grand Observatoire* projects.

Discussion was held on the availability of data collection and sharing guidelines. There are enough policy documents (pan African Strategy), enough data collection guidelines at national. However, data exchanges guidelines are missing. Generally speaking technical guidelines to implement policies are required, with the ad-hoc training and capacity building activities.

Key outcomes:

Need for guidelines from funding institutions to describe what to comply with. These Guidelines should be developed through standard setting bodies

There's a need for shared data catalogue and algorithms at regional level to avoid duplication of activities: example of metadata catalog in Australia that could be reuse in the SWIO context.

There is also a need for capacity building for describing data and catalogs

Implementation of standard (classification and tools) systems should improve data querying and sharing

2.2.14 Consequences of States fisheries operation multi reporting formats

Franck Callewaert presented the case of multi reporting format to EU and above EU.

He presented the need for national vessel activities monitoring as flag state or coastal state and the resulting wide variety of reporting formats and reporting requirements for a fishing vessel with striking figures. An attempt to make a global inventory of reporting rules was made in 2008. It stopped after having listed more than 1800 different rules.

A vessel needs to comply with all reporting requirements of all areas it is fishing in, in RFMO competence areas, in coastal states EEZ. Each change in a reporting requirement implies an update in the software used for reporting. UE fleet is 15000 vessels. If a software update cost is 1000 €, the deployment of a single requirement changes costs potentially 1.5 million euro.

This variety in reporting requirement has a direct impact on data quality: data are valid only within its context, transposition of data from one context to another usually generates poor quality data (too aggregated data as code lists aggregation level are too different, data are not comparable from year to year etc.).

[5] www.opengeospatial.org/ogc
[6] www.gbif.org
[7] http://sdmx.org

This variety poses also a threat on the capacity of fisherman to comply with all requirements with a high risk of penalties.

Conclusion

- Without coordinated approach and application of standards:
    - IT costs are blocking factor
    - In permanent change mode
    - Data quality/added value is poor
        - Very slow progress
        - Demotivating for all stakeholders

Answer is given in the next day presentation on FLUX

A discussion was held to identify some solutions for reporting harmonization. The problem is to keep track of all format changes. Francky Callewaert recommended that FAO plays an active role to define global plan for data collection and exchange. CWP should be the place to work on defining and maintaining standards in a timely manner.

Key outcomes:

The multiplication of reporting formats have a huge cost.

There is a need to implement a minimum data requirements approach: provide common classification; keep only meaningful data (no exotic element which are never exploited): keep it simple!

Stability of format is key to reduce IT costs to fishermen: need for a coordinated approach. FAO / CWP have a role to play to provide and maintain global standards.

2.2.15 UBC's catch reconstruction database

Kyrstn Zylich presented the Sea Around Us Reconstructed Catch Database.

- Focus: Global, 1950 to 2010, spatialized data, primarily catch data (effort/capacity data to come);
- Emphasis: placing fisheries data into ecologically meaningful and socio-politically feasible space;
- Catch data reconstructions cover all sectors: large-scale commercial (industrial), small-scale commercial (artisanal), and non-commercial (subsistence and recreational);
- The Sea Around Us data summaries and products are freely available for all users and aim at improving management, policy decisions and fisheries analysis;
- New website being built that makes data more accessible[8];
- Catch reconstructions utilize quantitative and qualitative information of various types. The process typically negates the assumption that "no data" means "zero" catch and reconstructs more realistic catch data for all documented sectors.

Issues and challenges are the following:

- Overcoming the concept that "no data" means "zero" catch was challenging for some external contributors and governments;
- The inherent negative bias of reported data needs addressing;

[8] www.seaaroundus.org

- ‘uncertainty’ is not the issue,, bias is;
    - Non-accessibility of existing literature (e.g. not available online);
    - Discrepancies between National and FAO official datasets;
    - Non-standard use of definitions (e.g. by-catch and discard rates);
    - The involvement of many external country based experts also required additional efforts to standardize the final datasets to Sea Around Us structure.
- Lessons learned:
    - Local knowledge is necessary;
    - Experts are better able to assess the acceptability of a number than to provide one themselves;
    - Initial standard definitions are important;
    - Reconstructing is an iterative process.

Discussion was held on the Sea Around Us transparency in the data reconstruction process to ensure processes can be reproduced outside Sea Around US. It has been acknowledged that more metadata on how reconstruction is conducted is needed. Algorithms will be soon available in GitHub[9] (open source code repository).

It was also highlighted the need for reconstruction process to match international definitions and classifications (CWP).

Key outcomes:

Lack of accessible data: Literature is not available on-line, it should be.

There should be no discrepancies between national data sets and FAO data sets.

Importance of disseminating credits, methodologies and standard used when publishing data: importance of metadata attached to data, importance of transparent and publicly available algorithms

Importance to follow international definitions (ex By Catch or global removal Vs Global Captures Production) defined by CWP

2.2.16 iMarine Regional Data Base in support to assessment and management

Anton Ellenbroek introduced the plan to develop a framework for a **Regional Data Base (RDB**) enabled throughout **iMArine Tools and Services**

The framework will enlarge the current services of the iMarine e-Infrastructure related to the management of data relevant to fisheries. The services offered already include data access, harmonization and analysis for specific work-flows, but work is ongoing to enrich these with additional free and open services based on R, Java, and WPS amongst others.

Together, these services provide the backbone of online databases, and these will be very useful at regional level, where an integrated view of status and trends in fisheries is currently difficult to obtain. The use of mobile applications for data collection is a new feature that can be exploited also in the RDB context.

[9] https://github.com

The presentation outlined how the BlueBRIDGE project will support the implementation of the ambitious plans for:

- Regional Databases for Stock Assessment; e.g. with the IRD Bluefin tuna model for ICCAT (supported by IRD), and the FAO Improved R model performance
- Global Record of Stocks and Fisheries (GRSF)
- Spatial planning and mangrove area mapping (GRID-Arendal / CLS[10])
- SmartForms Mobile Data Collection for scientific on-board sampling
- Stock Assessment Training Modules with ICES; and to
- Further Develop Open Access catalogues

  Data www.d4science.org/integrated-data-catalogue

  Applications www.gcube-system.org/catalogue-of-applications

Key outcomes:
Open source technologies: available bricks to build more complex systems. Provide a catalog of existing applications
Implementation of standard to support stock assessments
Mobile Data Collection for scientific on-board sampling

2.2.17 Integration of fisheries in broader statistical context

Sachiko Tsuji presented fisheries statistics in the broader context of global agriculture statistics

At national level, fisheries are not always integrated in the national statistical system: it proves difficult to have reliable data for national account (part of fisheries in GDP for instance); the need to have comparable data at national level is not limited to fishery.

The question of fisheries statistics inclusion on the overall national context leads to guidelines definition for such integration in agriculture census. Standard procedures are being developed in that sense.

For inland fisheries, System Environment Economic Accounting (SEEA) proved to be of interest. It is complementary to get frame data.

The most important here is to be able to identify several sources for the same information: it is a way to simplify data collection.

A discussion was held on how to ensure fisheries are included in agriculture census.

[10] www.cls.fr

Key outcomes:

Refer to standards, even across sectors: Fishery departments do not enough take into account the national economic statistical systems when developing their socio-economic systems

Find ways to utilize existing information frameworks

Refer to standards such as SEEA to more easily enable comparability

Need to have sampling frames; this can be done in different ways. Building more information by building on top of the existing frames

Need for robustness

2.3 Presentations summary and main outcomes

General principles, common issues related to the data supply chain have been identified in the SWIO context:

General principles

- General concern on the project short-term approach that should be set up within long term vision instead (data collection for instance)
- Data collection at national level is key to quality statistics at any level but is a concern (costly to monitor small-scale fisheries, problem of staff turn-over, problem of financial resources)
- Keep it simple!
- Do not reinvent the wheel! Build on what exists!

Common needs and issues have been identified:

- Fisheries Statistics need to be included in the National Strategy for the Development of Statistics; it should be compliant with the pan-African Strategy
- Any methodology / system / tool for any data supply chain level should be kept simple and robust.
- Initiative / tools / data catalog should be developed at regional level to support donor's project developments for cost efficiency (building on what existing / ensure complementary of activities / avoid duplication / ensure board and long term dissemination to the public)
- Need for guidelines at national and regional levels to support common/harmonized approaches: cf minimum core of standard indicators from the pan-African Strategy
- Need for socio economic indicators. Money rules!
- Need for best practices to implement such guidelines (policy)
- Need for a clear definition for each system of confidentiality
- Need for standard classification / definition / properly documented metadata to foster data harmonization at regional/international level: multiplication of reporting format/classification has a huge cost
- Regional coordination should be reinforced to avoid duplication, support existing initiatives
- Importance of sharing metadata together with data (credits, citation, sources, methodologies etc…)
- Importance of capacity building

2.4 Toward data policies and best practices

The four following presentations aimed to fit the previous discussions in a broader context to feed further analysis on data best practices and policies.

2.4.1 FAO's strategy to support the Blue Growth Initiative – the Global Data Framework for Blue growth

- Blue growth initiative

Marc Taconet recalled the context in which the Blue Growth (BG) initiative was built and its working definition (in section 1.2 but recalled here):

"Blue Growth is the sustainable growth and development emanating from economic activities in the oceans, wetlands and coastal zones, that minimize environmental degradation, biodiversity loss and unsustainable use of living aquatic resources, and maximize economic and social benefits"

The 3 Blue growth pillars were then presented:

1. **Environmental performance**
2. **Social and economic performance**
3. **Ecosystem services**

Answer by FAO to satisfy Blue Growth data needs is the Global Data Framework for Blue Growth. Its purpose is to promote availability and connect existing sources of knowledge. Its implementation underlies the promotion of harmonization, monitoring, collaboration. It also implies innovations and has implication on data strategies:

- Standardization and Harmonization
  - Promotion of international standards including metadata standards to fisheries
  - Master Data Management in fisheries
- Sharing , monitoring and evaluation
  - Collaborative IT infrastructure

Key outcomes:
Need for cross-sectoral data; need for responsive monitoring systems
Challenge cannot be faced by single institution, need for collaboration
Need for Harmonization, standardization

- CWP and promotion of international standards relevant to fisheries

Sachiko Tsuji presented CWP work and the path towards global depository of standards of fisheries statistics

CWP provides reference fisheries classifications and works toward an integration of fisheries in the major international classification

CWP is accessible here: www.fao.org/fishery/cwp/en

CWP will expand. CWP is considering covering non-statistical fishery data and data transmission:

- Data-transmission standards;
- Science data (e.g. observer data)
- Fisheries management data (e.g. catch reporting, MCS, Global Vessel Records)
- Catch Documentation System
- Master Data Management

Any participant interested in such discussion can request to participate to a Task Force.

Further discussion will be organized at the CWP25 (Feb. 2016)

In the following discussion, Marc Taconet insisted on the fact that this presentation reflects FAO Department of Fisheries and Aquaculture works on standards.

Emilie Leblond shared France's experience of implementing ASFIS species code list in all the French fisheries monitoring systems: log books suddenly come with tens of hundreds of new species not relevant a priori to fishing activities (sea birds for instance). She raised the need for better communication on standard use. MDM presentation (see below) intends to provide some answers.

A discussion was held on the need to provide mappings between code lists for different targets (fish name and fishery products names. How do you match them?) and define unique identifier to develop such mappings (some are obvious like the scientific name for species)

Key outcomes:

FAO's efforts to make existing global standards to be usable for the fishery sector

Towards broad availability and dissemination of fishery standards and other international standards relevant to fisheries; need to communicate well which standard for which purpose

CWP considers responding to needs for Metadata standards

Countries are eager for standards, but need for guidelines on how to use these standards (e,g this reference is relevant in this or this ocean)

Needs for mappings, e.g. marine species with WoRMS

Consider participation in the Research Data Alliance (RDA), we could set up a group for fisheries data

- Master Data Management (MDM) in Fisheries

Erik Van Ingen introduced the MDM concept to the participants.

A broad definition of MDM was provided highlighting the need and difficulty to define boundaries of what is MDM and not: "*MDM is a comprehensive method of enabling an enterprise to link all of its critical data to one file, called a master file, that provides a common point of reference*".

Given this definition, examples of MDM have been provided to the group and related tools to manage, exploit and share MDM. MDM helps managing standard, reference data and sharing these data in a large variety of formats suitable for the large variety of systems potentially consuming these data.

The deep link between MDM and open data is described with the need to share MDM as reference data for a given domain: the concept of Single Source of Truth (SSOT) is developed with the illustration of FAO reference data.

In the following discussion, Francky Callewaert reported to the group DG MARE problems with TALEND[11] (a well-known open source MDM tool). In the line of the SSOT concept, he insisted on the need for FAO to be the authoritative organization on Fisheries sector MDM, providing to other stakeholders links to reference code lists.

Key outcomes:
All data collected with public data should be available to the public
Important to work on Open data and share management with these groups
Responsibilities should be defined, with reference to the Single Source of Truth
FAO should become the single global source of authority for fisheries code lists

- The iMarine e-infrastructure: improving research in fisheries and biodiversity

GianPaolo Coro presented the iMarine infrastructure[12] to workshop's participants. He introduced the concept of Science 2.0 a new approach to science, with a key support to respect standards. iMarine infrastructure has been developed in line with Science 2.0 recommendations:

- An e-Infrastructure sustained by EU since 2004
- Supports collaborative Virtual Research Environments
- Offers hosting and management of data
- Supports Cloud computing, data integration and sharing

He then described the variety of service and bricks of software available in the e-infrastructure and its capacity to expand notably in the coming BlueBRIDGE project.

A discussion was held on the feasibility to replace RFMO's software and hardware by similar e-infrastructure. The main issue might not be on the technical side, it can indeed serve the need of RFMO with limited technical capacities. Reluctance might come on the capacity of such central tools to ensure data confidentiality. A step by step approach would be anyway required if such strategy would be approved by the Member States.

Discussion was held on the capacity to use such e-infrastructure to serve the needs of tools/data sharing and harmonization (for institutions for low technical capacities) or to deploy DG MARE software organized in packages. Francky Callewaert acknowledged the interest of such e-infrastructure but highlighted the need to be able to present it in more simple way for decision makers.

[11] www.talend.com
[12] https://i-marine.d4science.org

Key outcomes:

Pooling capacities from and for a community (data resources like taxonomic lists, species maps and software/algorithms like stock assessments, expertise in different domains).

Large potential for the e-infrastructure to serve needs of broader communities (DG MARE) but need to simplify key messages and address concerns of confidentiality, ownership…

The choice of a RFMO to move towards new IT solution must be endorsed by member states.

- Sustainability through partnerships Development

Marc Taconet presented the need to reach sustainability through partnership development in the Blue Growth context.

Partnership is key to share skills and expertise. Pooling resource together increases return on low direct investment by sharing tools, algorithms and IT resources.

Areas to develop these partnerships are listed: in capacity building, in standard developments, in monitoring systems and knowledge bases, and in the development of tools of infrastructure.

Finally keys to partnerships for data sharing success and best practices are detailed: role of contributors must be maximised by ensuring visibility of sources and sharing the ownership of initiative (provide due credit to source – provide business metadata); build on and reuse existing information assets to minimize data input and maximize exploitation of existing source documents and databases, including grey literature.

Key outcomes:

Partnership enables pooling of resources (tools, data and expertise) and implies ownership and clear understanding of benefits

Key to partnership success: contributor ownership is the most visible possible (through metadata, credit/citation, and active involvement)

Built on what exists to maximize exploitation

2.4.2 Examples of data policies

These last presentations aimed to give examples of existing activities/ innovative solutions on data management policies with possible best practices to be generalized/reused.

- iMarine data access and sharing policies

Aureliano Gentile shared with the group the iMarine experience on developing data access and sharing policies.

During the iMarine project (2012-2014) the data access and sharing policy documents were produced: iMarine Disclaimer and Terms of Use, iMarine Privacy Policy and iMarine Community Best Practices (including a Glossary of definitions and links to draft Guidelines). The Terms of Use were produced by FAO (including Legal and Communication offices) in collaboration with CNR and members of the iMarine advisory board. Key points in relation to data sharing are the Open Data and Open Access strategy, use of licences, citation policy, sharing vs. publishing and derivative works. The Business metadata logic is the envisaged way to address ownership, authorship, copyright and content description, the identity card of any shared content (what is – how to use). Adopting standard data/metadata formats will further facilitate

interoperability and data sharing. Data usage and metrics can support proper monitoring of utilization of owners data sources and can enable possible rewarding mechanisms. Adopting data access and sharing policies should contribute to trust and foster community engagement with infrastructure services, facilitate collaborations and partnerships and improve data quality and thus data analysis. In those iMarine Virtual Research Environment (VRE) where a Content Management System (CMS) is provided, the data and metadata are handled according to the policies, e.g. the Vulnerable Marine Ecosystem Database and the Tuna Atlas.

Key outcomes: Foster community engagement through trust Open Data and Open Access strategy, Common Licenses, Credits and Citations, provenance metadata, data usage metrics

- DG MARE: Integrated Fisheries Data Management Programme

Francky Callewaert presented the EC Integrated Fisheries Data Management Programme (IFDMP), formerly known as FLUX as an innovative approach to address multi reporting issues for Europe.

The IFDM programme goal is to deliver high quality data in the most effective way.

The IFDM programme has 2 aspects:

(i) first is standardization.

A single and complete fisheries data (exchange) language has been created to fit for all needed data exchanges, between any combination of parties. It is independent from any local IT solution.

The retained standardisation method was UN/CEFACT standardisation. Its main characteristics have been presented.

A unique ID is implemented for any message exchanged in the business layer: no duplication of info will occur anymore.

This business layer is the answer to the multiplication of reporting format by providing harmonization and standard.

(ii) Second aspect of IFDM programme is technology: it offers a single fisheries software suite:

- Open source, freely available: anyone can download it
- Managing all data (sending, receiving, storing, analysing)
- Simple Graphical user interface
- Compatible with UN/CEFACT standards

Organization of data flow and related technologies were presented.

The main benefit of this approach endorsed by EC is to have now a Standardised Data exchange and management platform, where IT is decoupled from business and that can grow organically. It is flexible and configurable.

The vision is to build a business community: it should enhance the UN/CEFACT standards (FAO is already contributing). It would also manage the open source software and make it evolve. The community would assist newcomers.

The discussion started on the possibility to join the Research Data Alliance working groups[13] for partners beyond. There is no restriction to join any group but it should not be science related group. IFDM programme is more about control than science

Key outcomes:

Clear separation between the information exchanged and the IT system provided for exchanges

Information exchanged is standardized as UN/CEFACT standards: reduce the number of reporting format; simplification

Each message is uniquely identified with a unique ID

A unique software is used for any type of information exchanges: cost efficiency

Software is open source and reusable by anyone (comes with guidelines)

- Building in the SWIO region on best practices for data harmonization, standards

Julien Barde illustrated with concrete examples the possibilities to build SWIO observatories and data catalogs on successful initiatives like ZOTERO, SOOS and GeoNetwork.

He recommended participating to RDA Marine Data Harmonization group for discussing how scientific format could be merged with control formats.

He highlighted the need for communities to get awareness of existing tools/initiatives and to organize themselves to exploit these capacities. Communication, collaboration and coordination would be key to success.

Indian Ocean Observatory could develop a data catalog for the SWIO region in the same way Antarctic (SOOS) has done. It can be done in an automated way if capacity building and guidelines are provided to the community to feed these catalogs.

Key outcomes:

Existing initiatives proved that data and metadata catalogs can be automatically built by experts when capacity building and guidelines are provided.

Existing open source/ free tools could be reused to build Indian Ocean Observatory including such data and metadata catalogs: it's all about developing communities to exploit data exchange capacities.

2.5 Working groups

After discussion with the participants, it was agreed upon that breaking-out into sub working groups working groups would not be practical or productive.

Break-out groups were replaced by a plenary session. The reference document was the circulated "Data strategies and best practices" draft document. One key output of the workshop being this Expert Document "Strategies and Best Practices", the plenary discussed the organization and content of such document. The draft result is found in appendix. It should be re-discussed in several rounds by email after this report validation to reach a satisfactory expert document.

[13] https://rd-alliance.org/groups

This expert document will then be presented to the regional institutions for endorsement throughout regional conference.

- Key outcomes of the discussion on the expert document “Strategies and Best Practices”:

It has been acknowledged that a vision was missing. Objectives serve a vision, activities indicates how one intend to achieve your objectives to meet the vision.

The agreed vision is:” Long term sustainable growth and development according to blue growth/economy objectives, will be achieved by having information on fisheries sector including environmental, social, economic, and aquatic sciences aspects available for facts based decision making“

From day 1 discussion and summary, Larry Oellermann extracted the below 17 key points as common issues/common needs/general principles:

1. access to data source of information
2. databases designed to interact
3. lack of local government buy in and support (staff stability)
4. incorporate alternative data sources and grey literatures
5. lack of sustainable funding, lost after development
6. need real socio economic data/money
7. Keep this simple/ core indicators
8. make sure local needs are met and provide answers to national, regional and global requirements - FAO with sense of ownership
9. ensure good quality data
10. focus on international standards and guidelines
11. information publicly available but also allowing confidentiality
12. open source
13. metadata catalogue
14. capacity building
15. transparency across systems
16. interpret fisheries information databases across larger AGRIS data system
17. community engaged through trust

These items were discussed and classified as constraints, principles, strategies in the expert document.

The key recommendations were:

- General recommendations:
    - Need for high level business data architecture with RFMO as center of excellence to defend and promote the global data architecture (identify national needs and support the development of tools to fill in gaps)
    - Make market analysis and then develop projects accordingly
    - Development of data/metadata/tools catalogs
- More regional needs:
    - Development of Chimaera Guidelines for data submission to Chimaera
    - Development of iMarine presentation for decision makers to sell it
    - Support the update of StatBase
    - Support ORI in the WIOFish software redevelopment

2.6 Conclusion

Thanks to participants' commitment and interest, the expert document has been drafted. A preliminary version of this key workshop output is available in appendix 3.

This draft needs now to be reviewed by the workshop participants in different rounds to reach a final version that will then be submitted to regional stakeholders for presentation, discussion and endorsement in a regional conference.

Additional regional activities have been identified and are summarized below on the recommendation section.

3 WORKSHOP OUTPUTS AND RECOMMENDATIONS

Recommendation 1: the expert document will be revised and a final version will be prepared with the support of FAO

Recommendation 2: a high level business data architecture, driven by the stakeholders needs is required as key document to enable national and regional stakeholders with relevant decisions when investing in data collection and sharing systems

Recommendation 3: development of iMarine presentation for a decision makers audience in order to more efficiently sell it

Recommendation 4: SWIOFC should play a central role in coordination to support data best practices and policies with the region

Recommendation 5: support regional activities

- Develop of a regional data/metadata/tool catalog
- Development of Chimaera Guidelines for data submission to Chimaera
- Support the update of StatBase –
- Support ORI in the WIOFish software redevelopment

4 APPENDIXES

4.1 Appendix 1: List of participants

1. Alejandro AGANUZZI - FAO
2. Anton ELLENBROEK - FAO
3. Aureliano GENTILE - FAO
4. Benjamin GARNAUD - World Bank
5. David WILSON - IOTC Secretary (Vice Chair of the workshop)
6. Dawn BORG COSTANZI - FAO/FIRO
7. Dirk ZELLER - Sea AroundUs UBC
8. Emilie LEBLOND - IFREMER
9. Emmanuel CHASSOT - IRD/IOTC
10. Fabio FIORELLATO - FAO IOTC Consultant
11. Fernando JARA - FAO
12. Francky CALLEWAERT - DG MARE
13. GertJan DE GRAAF - consultant FAO/WB
14. GianPaolo CORO - CNR
15. Harrison ONGANDA - Kenyan Marine Fisheries Research Institute (KMFR)
16. Julien BARDE - IRD
17. Kyrstn ZYLICH - Sea AroundUs UBC
18. Larry OELLERMANN - ORI/WIOFISH , National Oceanographic Institute South Africa
19. Luca GARIBALDI - FAO
20. Marc TACONET - FAO
21. Raul VILELA - FAO
22. Sachiko TSUJI - FAO
23. Tony JARRETT - FAO
24. Valerie HARSCOAT - Ifremer
25. Xavier VINCENT - World Bank
26. Yann LAURENT - FAO

4.2 Appendix 2: Meeting agenda

Monday 21 September 2015 (FI Meeting Room F313)

Time	Activity	Description *(Who)*
9.00 – 9.10	**1. Opening**	
9.10 – 9.20	**2. Participants introduction**	
9.20 – 9.30	**3. Introduction to the workshop adoption of the agenda**	*IOTC; FIPS*
9.30 – 9.50	**4. Problem setting** • Chimaera	Presentation of the problem identified during Chimaera FRP development (*Yann Laurent*)
9.50 - 10.35	**5. Review of initiatives supporting the data supply chain; Participants presentation (indicative list):** •	Format: A 10x5 minutes session – 10 min presentation, 5 minutes discussion. The list is indicative and reflects the organizers' knowledge; participants are called to suggest other possible initiatives that could be presented, and to confirm the titles of their presentation. Initiatives will be delivered according to following template[14]: Sector/policy objective(s), main indicators, type of information/data, approach/methods/systems/tools, considerations on data harmonization and standards, highlights of data policies, issues faced - bottleneck
	5.1. Support to countries data collection systems • ORI – WIOFish • WB – SWIOFish • FAO – toolbox framework facilitating integration of fishing operations related databases Including Sample based schemes / Open Artfish, and Tablets/mobile for OpenArtFish	Larry Oellerman Xavier Vincent Gertjan DeGraaf / Luca Garibaldi
10.30-10.45	Coffee break	
10.50-12.50	5. Participants presentation: **5.1. Support to countries data collection systems** • IFREMER – SIH for Seychelles, Comoros **5.2. Regional and Global initiatives** • FAO: Global Record of Fishing Vessels • Overview of improvement of fisheries data collection in the SWIOFC region, including the Pan African strategy • IOTC: CLAV	*Emilie Leblond, Valérie Harscoat* *Dawn BorgCostanzi* *Gertjan DeGraaf, Luca Garibaldi* *IOTC*

[14] A PPT template will be provided apart.

Time	Activity	Description *(Who)*
	• SWIOFC – approaches to data collection and tools in support to Stock Assessment and Fishery Management (Weight of Evidence) • KMFRI - StatBase • IOTC – integrated tuna statistical database • FIRMS, towards the Global Record of Stock and Fisheries • IRD - Sardara tuna Atlas project	*David Wilson* *Harrison Ong'anda* *IOTC* *Aureliano Gentile* *Julien Barde*
12.50-14.00	Lunch Break	
14.00-15.15	5. Participants presentation (contd) **5.2. Regional and Global initiatives** • IRD - Observatoire de l'Océan Indien • Consequences of States fisheries operation multi reporting formats • UBC's catch reconstruction database • iMarine Regional Data Base in support to assessment and management • Integration of fisheries in broader statistical context	*Julien Barde* *Francky Callewaert* *Kyrstn Zylich, Dirk Zeller* *Anton Ellenbroek* *Sachiko Tsuji*
15.15-15.45	**6. Wrap-up of Participants' presentations :** synthesis on data supply chain / lessons learned / Best practices / Strategies already in place / Gaps / challenges	Discussion with all participants *(Moderated by Yann Laurent)*
15.45-16.00	Tea break	
16.00-17.30	**7. Listing data initiatives/systems versus policy objectives in the region:** • Mapping national and regional needs for data to existing systems, initiatives and strategies – identify synergies • Data flow description (from national to international level) • Identify redundancies / best practices / issues, share lessons and identify where collaboration is needed to improve overall quality of data collected and processed.	Moderated working session *(Moderated by Yann Laurent)*
17.30-18.00	**8. Wrap-up:** • Outline case for a regional data policy to address challenges • Best practices to be promoted • Gaps in support to national and regional systems in the region (what's done / what's still to be done) • Areas of collaboration across institutions	

Tuesday 22nd of September 2015 (Morning in TCI Meeting Room D-532 then afternoon in AGD Meeting room B-640)

Time	Activity	Description *(Who)*
9.00 – 9.15	9. Summary of day 1 main outcomes	
9.15 – 10.15	**10. FAO's strategy to support the Blue Growth Initiative – the Global Data Framework for Blue growth:** o CWP and promotion of international standards relevant to fisheries o Master Data Management in Fisheries o The iMarine e-infrastructure: improving research in fisheries and biodiversity o Sustainability through partnerships Development	*Marc Taconet, with contributions from* o *Sachiko Tsuji,* o *Erik VanIngen,* o *Gianpaolo Coro* o Marc Taconet
10.15-10.30	Coffee break	
10.30-11.30	**11. Towards a data policy, strategies and best practices:** • Participants presentations (list to be elaborated)	Participants are expected to present: - examples of applicable principles and best practices for data sharing and access policy; - example of needs for these principles and policies Discussion and break-out groups identified according to clusters
	Data Access and Sharing Policies	Aureliano Gentile
		Francky Callewaert
		Julien Barde
	12. Identification of components and potential benefits of strategies and best practices	Discussions from a draft policy proposal
11.30-13.00	**12.a Break-out groups:**	The break-out groups will be requested to respond to a number of questions; the groups inputs will contribute to shape the high level data policy and strategy
13.00-14.00	Lunch Break	
14.00-16.00	**13. Conclusion of workshop outputs** • Reporting of break-out groups and wrap-up : • List of SWIO data initiatives, positioned in the Data supply Chain (see figure 1) • Expert Document "Strategies and Best Practices" - strategies and best practices for an efficient data supply chain;	

Time	Activity	Description *(Who)*
	- components of data policy of global value; - extensions deemed necessary to existing instruments concerning the data supply chain • roadmap for implementation and funding • Practical solutions for potential funding	
16.00-16.15	Tea break	
16.15-18.00	**13. Conclusion of workshop outputs (continued)**	

4.3 Appendix 3: List of SWIO data initiatives, positioned in the data supply chain

4.3.1 Appendix 3.1: pan-African Strategy positioning

SCOPE	National policies	Sector policies	Fisheries management in the EAF approach / stock assessment			MCS/ IUU combatting	CDS/ TCS/ traceability	Co-management ? stakeholders outreach	Research and science
			Human component	Stock assessment	Environmental component				
National Fishery Statistics Strategy (included in NSDS)	PAS	PAS	PAS	PAS	PAS	PAS	PAS	PAS	PAS
Survey Design	PAS	PAS	PAS	PAS	PAS	PAS	PAS	PAS	PAS
Data collection									
Data collation / compilation									
Data processing/ Analysis/ production of indicators									
Reporting / dissemination									
Accessibility (including confidentiality)									
Metadata availability									
Utilization									

4.3.2 Appendix 3.2: WB SWIOFISH positioning

SCOPE	National policies	Sector policies	Fisheries management in the EAF approach / stock assessment			MCS/ IUU combatting	CDS/ TCS/ traceability	Co-management ? stakeholders outreach	Research and science
			Human component	Stock assessment	Environmental component				
National Fishery Statistics Strategy (included in NSDS)	WBSW	WBSW	WBSW	WBSW	WBSW	WBSW	WBSW	WBSW	WBSW
Survey Design	WBSW	WBSW	WBSW	WBSW	WBSW	WBSW	WBSW	WBSW	WBSW
Data collection									
Data collation / compilation									
Data processing/ Analysis/ production of indicators									
Reporting / dissemination									
Accessibility (including confidentiality)									
Metadata availability									
Utilization									

4.3.3 Appendix 3.3: Chimaera positioning

SCOPE	National policies	Sector policies	Fisheries management in the EAF approach / stock assessment			MCS/ IUU combatting	CDS/ TCS/ traceability	Co-management ? stakeholders outreach	Research and science
			Human component	Stock assessment	Environmental component				
National Fishery Statistics Strategy (included in NSDS)									
Survey Design									
Data collection									
Data collation / compilation									
Data processing/ Analysis/ production of indicators									
Reporting / dissemination	CH	CH	CH	CH	CH				CH
Accessibility (including confidentiality)	CH	CH	CH	CH	CH				CH
Metadata availability									
Utilization									

4.3.4 Appendix 3.4: StatBase positioning

SCOPE	National policies	Sector policies	Fisheries management in the EAF approach / stock assessment			MCS/ IUU combatting	CDS/ TCS/ traceability	Co-management ? stakeholders outreach	Research and science
			Human component	Stock assessment	Environmental component				
National Fishery Statistics Strategy (included in NSDS)									
Survey Design									
Data collection									
Data collation / compilation					SB				SB
Data processing/ Analysis/ production of indicators					SB				SB
Reporting / dissemination					SB				SB
Accessibility (including confidentiality)									
Metadata availability									
Utilization									

4.3.5 Appendix 3.5: Open artFish positioning

SCOPE	National policies	Sector policies	Fisheries management in the EAF approach / stock assessment			MCS/ IUU combatting	CDS/ TCS/ traceability	Co-management ? stakeholders outreach	Research and science
			Human component	Stock assessment	Environmental component				
National Fishery Statistics Strategy (included in NSDS)									
Survey Design									
Data collection					OPA				OPA
Data collation / compilation					OPA				OPA
Data processing/ Analysis/ production of indicators					OPA				OPA
Reporting / dissemination					OPA				OPA
Accessibility (including confidentiality)					OPA				OPA
Metadata availability									
Utilization									

4.3.6 Appendix 3.6: WIOFish positioning

SCOPE	National policies	Sector policies	Fisheries management in the EAF approach / stock assessment			MCS/ IUU combatting	CDS/ TCS/ traceability	Co-management ? stakeholders outreach	Research and science
			Human component	Stock assessment	Environmental component				
National Fishery Statistics Strategy (included in NSDS)									
Survey Design									
Data collection									
Data collation / compilation	WF	WF	WF	WF	WF	WF	WF	WF	WF
Data processing/ Analysis/ production of indicators									
Reporting / dissemination	WF	WF	WF	WF	WF	WF	WF	WF	WF
Accessibility (including confidentiality)	WF	WF	WF	WF	WF	WF	WF	WF	WF
Metadata availability									
Utilization									

4.3.7 Appendix 3.7: Système d'information halieutique positioning

SCOPE	National policies	Sector policies	Fisheries management in the EAF approach / stock assessment			MCS/ IUU combatting	CDS/ TCS/ traceability	Co-management ? stakeholders outreach	Research and science
			Human component	Stock assessment	Environmental component				
National Fishery Statistics Strategy (included in NSDS)									
Survey Design	SIH	SIH		SIH	SIH	SIH	SIH	SIH	SIH
Data collection	SIH	SIH		SIH	SIH	SIH	SIH	SIH	SIH
Data collation / compilation									
Data processing/ Analysis/ production of indicators	SIH	SIH		SIH	SIH	SIH	SIH	SIH	SIH
Reporting / dissemination	SIH	SIH	SIH	SIH	SIH	SIH	SIH	SIH	SIH
Accessibility (including confidentiality)	SIH	SIH		SIH	SIH	SIH	SIH	SIH	SIH
Metadata availability	SIH	SIH	SIH	SIH	SIH	SIH	SIH	SIH	SIH
Utilization									

4.3.8 Appendix 3.8: ITOC – Integrated tuna statistical database

SCOPE	National policies	Sector policies	Fisheries management in the EAF approach / stock assessment			MCS/ IUU combatting	CDS/ TCS/ traceability	Co-management ? stakeholders outreach	Research and science
			Human component	Stock assessment	Environmental component				
National Fishery Statistics Strategy (included in NSDS)									
Survey Design									
Data collection									
Data collation / compilation					IOTC TDB				
Data processing/ Analysis/ production of indicators					IOTC TDB				
Reporting / dissemination					IOTC TDB	IOTC TDB			
Accessibility (including confidentiality)					IOTC TDB				
Metadata availability					IOTC TDB				
Utilization									

4.3.9 Appendix 3.9: IRD – Sardara tuna atlas

SCOPE	National policies	Sector policies	Fisheries management in the EAF approach / stock assessment			MCS/ IUU combatting	CDS/ TCS/ traceability	Co-management ? stakeholders outreach	Research and science
			Human component	Stock assessment	Environmental component				
National Fishery Statistics Strategy (included in NSDS)									
Survey Design									
Data collection									
Data collation / compilation				IRD-TA	IRD-TA				
Data processing/ Analysis/ production of indicators				IRD-TA	IRD-TA				
Reporting / dissemination				IRD-TA	IRD-TA				
Accessibility (including confidentiality)				IRD-TA	IRD-TA				
Metadata availability				IRD-TA	IRD-TA				
Utilization									

4.3.10 Appendix 3.10: iMarine RDB in support to stock assessment and management

SCOPE	National policies	Sector policies	Fisheries management in the EAF approach / stock assessment			MCS/ IUU combatting	CDS/ TCS/ traceability	Co-management ? stakeholders outreach	Research and science
			Human component	Stock assessment	Environmental component				
National Fishery Statistics Strategy (included in NSDS)									
Survey Design									
Data collection									
Data collation / compilation				iRDB	iRDB				
Data processing/ Analysis/ production of indicators				iRDB	iRDB				
Reporting / dissemination				IRDB	iRDB				
Accessibility (including confidentiality)				iRDB	IRDB				
Metadata availability				IRDB	iRDB				
Utilization									

4.3.11 Appendix 3.11: SWIOFC – WoE approach

SCOPE	National policies	Sector policies	Fisheries management in the EAF approach / stock assessment			MCS/ IUU combatting	CDS/ TCS/ traceability	Co-management ? stakeholders outreach	Research and science
			Human component	Stock assessment	Environmental component				
National Fishery Statistics Strategy (included in NSDS)									
Survey Design									
Data collection									
Data collation / compilation				WoE					
Data processing/ Analysis/ production of indicators				WoE					
Reporting / dissemination				WoE					
Accessibility (including confidentiality)									
Metadata availability									
Utilization									

4.3.12 Appendix 3.12: Sea Around Us positioning

SCOPE	National policies	Sector policies	Fisheries management in the EAF approach / stock assessment			MCS/ IUU combatting	CDS/ TCS/ traceability	Co-management ? stakeholders outreach	Research and science
			Human component	Stock assessment	Environmental component				
National Fishery Statistics Strategy (included in NSDS)									
Survey Design									
Data collection									
Data collation / compilation	SEA	SEA			SEA				SEA
Data processing/ Analysis/ production of indicators	SEA	SEA						SEA	SEA
Reporting / dissemination	SEA	SEA						SEA	SEA
Accessibility (including confidentiality)								SEA	SEA
Metadata availability									
Utilization									

4.3.13 Appendix 3.13: Global Record of Fishing Vessels

SCOPE	National policies	Sector policies	Fisheries management in the EAF approach / stock assessment			MCS/ IUU combatting	CDS/ TCS/ traceability	Co-management ? stakeholders outreach	Research and science
			Human component	Stock assessment	Environmental component				
National Fishery Statistics Strategy (included in NSDS)									
Survey Design									
Data collection									
Data collation / compilation						GFV			
Data processing/ Analysis/ production of indicators						GFV			
Reporting / dissemination									
Accessibility (including confidentiality)									
Metadata availability									
Utilization									

4.3.14 Appendix 3.14: IOTC-CLAV

SCOPE	National policies	Sector policies	Fisheries management in the EAF approach / stock assessment			MCS/ IUU combatting	CDS/ TCS/ traceability	Co-management ? stakeholders outreach	Research and science
			Human component	Stock assessment	Environmental component				
National Fishery Statistics Strategy (included in NSDS)									
Survey Design									
Data collection									
Data collation / compilation					CLAV	CLAV			
Data processing/ Analysis/ production of indicators					CLAV	CLAV			
Reporting / dissemination					CLAV	CLAV			
Accessibility (including confidentiality)					CLAV	CLAV			
Metadata availability					CLAV	CLAV			
Utilization									

4.3.15 Appendix 3.15: FIRMS

SCOPE	National policies	Sector policies	Fisheries management in the EAF approach / stock assessment			MCS/ IUU combatting	CDS/ TCS/ traceability	Co-management ? stakeholders outreach	Research and science
			Human component	Stock assessment	Environmental component				
National Fishery Statistics Strategy (included in NSDS)									
Survey Design									
Data collection									
Data collation / compilation			FIRMS	FIRMS	FIRMS				FIRMS
Data processing/ Analysis/ production of indicators			FIRMS	FIRMS	FIRMS				FIRMS
Reporting / dissemination			FIRMS	FIRMS	FIRMS			FIRMS	FIRMS
Accessibility (including confidentiality)			FIRMS	FIRMS	FIRMS			FIRMS	FIRMS
Metadata availability			FIRMS	FIRMS	FIRMS			FIRMS	FIRMS
Utilization									

4.4 Appendix 4: Expert Document "Strategies and Best Practices"

DATA STRATEGIES AND BEST PRACTICES DRAFT

1. Introduction

This expert document proposes an analysis of data supply chain in the context with blue growth objectives to propose data strategies and best practices.

2. Vision

Long term sustainable growth and development according to blue growth/economy objectives, will be achieved by having available information on fisheries sector including environmental, social, economic, and aquatic sciences aspects for facts based decision making

3. Objectives

To achieve the vision, following objectives are established:

- Enabling routine collection of data at national level for generation of basic fishery indicators
- Responding to the increasing needs for information requested by decision makers
- Optimizing investments through promoting availability for all and reuse of existing data, while recognizing and valuing sources
- Connecting existing data to generate broader knowledge

4. Definitions

The expert document will use the below terms in their proposed definitions:

- **Data**: is understood a very broad context. It covers information, statistics, data.
- **Data supply chain**: the fishery data supply chain is understood in this document as the chain of processes aiming to produce elaborated data or information from field data (mostly from the national level). It goes from data collection to data dissemination at different levels of the value chain, including data collation, data storage, data processing and data exchange. It also encompasses data harmonization and data standardization concepts.
- **Project**: a set of activities containing at least one related to fisheries data collection, processing or dissemination. A project is managed by the hosting institution.
- **Hosting institution**: the institution responsible for the implementation of the project.
- **Blue Growth**[15]: is the sustainable growth and development emanating from economic activities in the oceans, wetlands and coastal zones, that minimize environmental degradation, biodiversity loss and unsustainable use of living aquatic resources, and maximize economic and social benefits

[15] FAO Blue Growth working definition.

5. Constraints

Experts have identified from their experience the following constraints that posed problems in achieving the vision.

- Capacity of decision makers in understanding the value and potential use of information products to get government buy-in and support
- Issue of sustainable funding: everybody wants data, nobody wants to be pay for it
- National constraints: lack of human and financial resources and staff turn-over
- Lack of high level business architecture (need for definitions, with a systematic stakeholder needs driven approach)
- Lack of coordination mechanisms to ensure data long term sustainability (Accountability)
- Communication weaknesses undermining awareness and action in proper direction
- Lack of standard concepts and definitions, classifications and harmonization
- Constraint of data accessibility (Can be imposed by policies – need of system interoperability)
- Gaps in data (with a special focus on socio-economic data)
- Cost efficiency (Reuse, built on what exists, do not reinvent the wheel)
- How to access to grey literature?

6. Principles

Keep it simple

Do not reinvent the wheel / Reuse what exists

7. Strategies

The below strategies are answers to lift above constraints:

- Should enable coordination mechanism at national level, at regional level and at global level: Coordination mechanisms under formalized governance should (would/must?) be set up at national regional and/or global levels for defining data needs and policy questions, data sharing arrangement and trust, sustainable
- Should envisage principles of good governance of data collection for quality data, including inclusion of fisheries statistics in national statistics system and in National Statistical Development Strategy for long term sustainability
- Should allow to easily identify gaps - A higher focus should be put on Socio-economic indicators
- Reflect how can best practices be enforced at donors / project level –
    - e.g. mandatory to describe the data with metadata (enhance trust in data with metadata publication)
    - e.g. mandatory to define confidentiality, and different levels of confidentiality, and related rights to access and use
    - e.g. open data / semi-open data / copy left / creative commons
    - e.g. data rescue
    - e.g. open tools built on open source technologies
- Should promote data access and sharing best practices to avoid duplication: enable data comparability, enable accessibility of data in long term, of existing / unpublished / un-disseminated data.

- Should identify how to handle confidentiality, and how to mitigate confidentiality, which mechanisms (e.g. market measures) can lead data providers to provide their data
- Should identify how to unlock access to grey literature
- Should set up mechanisms to ensure proper control of data made available in the cloud
- Should foster the sharing of expertise and tools (even IOTC is not able to face all needs for data handling)
- Should identify how best to develop capacity for the above strategies to mitigate risk link to staff turn-over and avoid duplication

8. Best practices

This section aims to provide pragmatic solutions to implement above strategies along the statistics supply chain.

a. Data

- Data collection
  - Reference methodologies and tools shall be adapted to the national context: frame survey, industrial fisheries monitoring (landing at dock monitoring, or log book or e-log book, VMS etc..), artisanal fisheries monitoring (sample based survey, frame survey etc…), socio-economic statistics (important to assess impact/return on investment of projects on sector development)
  - Use International classifications (species, fishing areas, vessel type, gear type) whenever possible, as disseminated by Standard setting bodies (e.g. CWP). When practical constraints prevent the use of international classifications, harmonization strategies should be considered a priority, i.e. mapping of regional or national standards (e.g species local name mapped with scientific name, gear type at national level) with international classifications.
  - Use tools able to disseminate metadata according to international standards
  - Build on data collection tools designed for data sharing according to data sharing best practices

- Data processing
  - All algorithms should be public – promote sharing of methods and data quality controls, reproducibility of experiments
  - In processing confidential data towards more open aggregated data, confidential data should never be removed, rather aggregation resolution should be modified to allow proper inclusion/aggregation of data that cannot be published

- Data sharing
  - Metadata should be produced for data collections, data sets, and as far as possible for individual records (Metadata catalogs)
  - In producing data and Metadata, data initiatives should comply with international/regional Metadata standards: as first priority, compliance with these standards should be from a Semantic and conceptual view point (concept definitions, conceptual data model), then second priority is to set-up tools to export data/information using these regional/international physical formats. Sources should

be able to produce data according to the formats they have capacity to generate (e.g. CSV in low capacity context), and higher level infrastructures should provide data services with transformation capacities to international standards
- The recommended International standards are disseminated by main information domain on the CWP handbook website.
- Confidentiality: Metadata should describe the meaning of confidentiality for each dataset

- Data dissemination/sharing
  - Projects should feed data catalog / data repositories – provide possibility to subscribe to such catalog (capacity building to feed these catalog) – promote stability in dissemination formats and standards (Owner's responsibility)
  - Metadata should be published together with data, to allow identification of ownership, usage rights, quality levels, correct interpretation, and reuse for integration in derivative works
  - Provide detailed operational data under agreed data policies addressing confidentiality
  - Data sets which are used for an assessment should be accessible; uses DOIs (ICES example)

- Data valuation
  - A repository of metadata is needed to raise awareness of existing data initiatives (It was recommended in the Caribbean Marine Atlas initiative) so to complement the existing and avoid to duplicate;
  - In order to foster data exchange and sharing, source data initiatives should be valued when derivative work is produced building on them: due credit should be provided, provenance metadata should be preserved, data usage should be monitored and metrics provided; rewarding mechanisms either through proper recognition of intellectual property, or through contributing to increase visibility, or through revenues or credit point systems based on usage.

- Data security
  - Storage:
    - Adopt basic and simple practices of back-ups on standalone computers
    - Make best use of Cloud for rescuing data, for ensuring secured data access (safety box)
    - The hosting infrastructure is responsible for data stored especially if they break the law – verify data policies of hosting infrastructures
    - Data infrastructure managing such data repository catalogs should offer tools able to manage metadata and related metrics
  - Access:
    - Defined Confidentiality
    - Practices for data trust should be set-up to enable sharing of data in confidential context

b. Software

- Reuse of software if documentation and expertise is available
- Publish software with open source license in order to facilitate sharing and reuse
- Software are natively designed to exchange data (with policies implementation)

9. Data terms of use

The section contains proposals for general data terms of use to be recommended for data related projects

- General data policy that apply to the whole fisheries statistical chain at National and/or Regional level
  - All data and tools created in the course of the project are the properties of the hosting institution. The vendor or the project staff ensures that all material related to fisheries data are made available without any restriction to the hosting institution in electronic format including detailed data, aggregated data, tools, scripts, databases
  - In all cases, existing National or Institutional data policies prevail to the regional data policies listed below.
  - Violation of laws or third party right: if any information or data collected, processed and/or disseminated violates any laws or third party rights, the project hosting institution will request such data to be removed.
- Fisheries Supply chain policies
  - Data collection
    - Data collection developed and implemented processes should refer to regional data collection best practices (to be listed).
    - Data confidentiality should be defined at data collection level. Such definitions should be clearly disseminated as part of metadata
  - Data storing
    - Storing should not be done under a proprietary solution unless data can easily be exported in generic format (CSV) or open source language (SQL).
  - Data access
    - Recommendation is to make public general information on fisheries containing no personal information (fisheries description, stock inventories, MPA descriptions, aggregated statistics), any data containing personal information (captain name, companies names, detailed turnover figures etc.) are kept confidential.
    - Access to confidential data requires either a specific request or an agreement to comply with confidentiality for the requester or the anonymization of confidential data for exploitation by the requester (Cf. biological data) with reference to the confidentiality definition
    - Need to define clearly what is confidential or not (data access policies)

- Data sharing
  - Copyrights: Data records should be shared together with a set of metadata (‘business Metadata’) which general purpose is to describe data ownership, conditions of use and visibility of the source. Business Metadata should include licensing information which specifies use and re-use conditions, enable users to provide credit and citation, as well as traceability of the source.
  - Sharing can be in different formats: paper/electronics/flow with different policies in terms of notification when new data are available for sharing
  - License: Except where in contradiction with institutional policies, Content shall be licensed for dissemination and redistributed under the Creative Commons License.

10. Roles and responsibilities

a. International organizations
   - **FAO**: act as a key provider of international classifications through CWP **(and possibly to come**: as clearing house for metadata standards relevant to fisheries)

b. Regional organizations
   - **SWIOFC**: key coordination role to support regional data strategies and best practices and policies dissemination to regional projects
   - **RFMO**: endorse and support regional data strategies and best practices

c. National agencies
   - **National Fisheries Authorities:** responsible for data collection and statistics raising

11. Recommendation of practical solutions for potential funding

a. Complementarity among data initiatives:
   - from strategic view point
   - from data workflow view point
   - i-Marine
   - ArtFish plan of action
   - Support to CWP work

b. Areas of collaboration
   - Project drafting to avoid duplication of effort

c. Improvements required on individual initiatives
   - Chimaera guidelines to add data to the portal
   - StatBase as regional database: resources for assessment workshop to improve data and system (to be endorsed by SWIOFC)
   - Best practices for fisheries business basic activities

d. Identified gaps
   - Need for a high level architecture document: development of a global document describing the fisheries data business architecture and implementation plan for SWIO
   - Once architecture doc is available, state of the art software

12. Recommendations regarding extensions deemed necessary to existing instruments concerning the data supply chain

- Implementing data access and sharing policies:
  - handling confidentiality: having the capacity for a tool/e-infrastructure to implement specific policies (National, Regional or global) – open access to data generated through publicly-funded projects (Cf. Open Aire for FP7 / H2020 projects[16]);
  - define SLA at e-infrastructure level

[16] See www.openaire.eu/oa-policies-mandates

A Strategic Data Policy Workshop was organized in Rome, Italy on 21 and 22 September 2015, with the main objective of developing strategies and best practices for investments in an efficient data supply chain. Initially targeting the South West Indian Ocean region, the workshop was opened up to raise general principles with a global scope, using the SWIO region case study as an example to build on. Working sessions aimed to analyse different scenarios through projects/initiatives, presentations and discussions to identify what did or did not work, in order to define general principles, highlight challenges and propose strategies and best practices for cost-efficient and sustainable investments in fisheries data collection, sharing and utilization.

This report provides an overview of the main presentations given over the course of the workshop and highlights key outcomes and recommendations based on discussions and in-depth reviews. Over the course of the workshop seventeen key points were identified as common issues/common needs/general principles and were organized and developed into the Expert Document "Strategies and Best Practices" which can be found at the end of this report.

ISBN 978-92-5-109711-3 ISSN 2070-6987

9 789251 097113

I7056EN/1/01.18