IOTC-2017-WPDCS13-33

Collaboration between fisheries and computer scientists for improved data description: The case of IOTC data sets

Julien Barde[*], Emmanuel Blondel [†], Emmanuel Chassot
Taha Imzilen, Anne-Elise Nieblas[*], Paul Taconet[*]

SUMMARY

Most fisheries data sets managed by the IOTC are in the public domain. These data sets are accessible and well described but currently hard to locate outside the IOTC website and require an improved description for facilitating their use, e.g., making explicit the license rights, the data structure, the codifications used, and specifying available Web Services to access or subset the data. Metadata can be of great help to the users and also improve the citation of data management work with digital object identifiers (DOI) and data papers. Moreover, similar data sets are managed by the other Regional Fisheries Management Organizations and it would be helpful to standardize (meta)data formats and access protocols at a global scale. In this paper, we present a method based on standards for metadata and data interoperability, following FAIR principles (Findable, Accessible, Interoperable and Reusable), which enables a rich description of data sets by complying with widely used standards. We showcase how it is possible with these standards and the related applications which implement them to better describe and discover fisheries and stock assessment data as well as to build additional services like data access and visualization tools.

KEYWORDS: Metadata, data description, data discovery, stock assessment, data access, FAIR principles, catch and effort data, length frequency data, R, ISO, OGC

[*] IRD - UMR MARBEC 248, Av. Jean Monnet, 34200 Sète, France; julien.barde@ird.fr; Phone: +33 499 57 32 32 Fax: +33 499 57 32 15.

[†] FAO - Fisheries and Aquaculture Department (FI). Fisheries and Aquaculture Policy and Resources Division (FIA)

1. Introduction

Most fisheries data sets managed by the IOTC are in the public domain. These data sets are accessible and well described but currently hard to locate outside the IOTC website and require an improved description for facilitating their use, e.g., making explicit the license rights, the data structure, the codifications used, and specifying available Web Services to access or subset the data. Metadata can be of great help to the users and also improve the citation of data management work based on persistent identifiers, digital identifiers (DOI) and underlying data papers. Moreover, similar data sets are managed by the other Regional Fisheries Management Organizations and it would be helpful to standardize (meta-)data formats and access protocols at a global scale. In this paper, we present a method based on standards for metadata and data interoperability, following FAIR principles (Findable, Accessible, Interoperable and Reusable), which enables a rich description of data sets by complying with widely used standards. We showcase how it is possible with these standards and the related applications which implement them to better describe and discover fisheries and stock assessment data as well as to build additional services like data access and visualization tools.

2. IOTC data sets to be described

Among all data sets, the public domain ones should be the first to be described since they allow to establish a direct link between metadata and data. Moreover, when data can be parsed, basic chunks of codes can infer some of the key metadata elements, such as spatial and temporal extent, keywords (eg lists of species or fishing gears). It is important to keep in mind that publishing and sharing metadata does not imply publishing and sharing data. Data with restricted access can also be described.

As a start, we suggest that the following IOTC data sets listed in IOTC WebSite (list to be checked and validated) should be described:

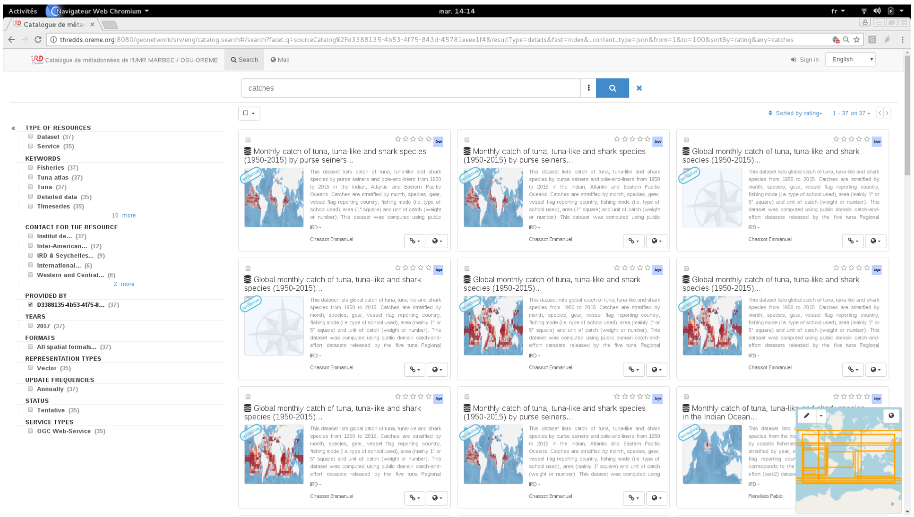
- Ongoing work
  - **Nominal catch** by species and gear, by vessel flag reporting country
  - **Catch-and-effort** by month, species and gear, by vessel flag reporting country. All CE files (CE purse seine and bait boat & CE longline & CE Other gears), including CE reference
  - **Length frequency** for tropical tunas, temperate tunas, billfish, neritic tunas, sharks and some bycatch species

   - **Fishing crafts** Fishing vessel statistics: number of vessels authorized to operate by country flag, by size category

- Future work

   - **Tagging data**: EU-funded Regional Tuna Tagging Program
   - Several other small-scale tagging projects: Electronic tagging of yellowfin and bigeye tuna, PROSPER Project Phase 2

We intend to work iteratively to enrich the metadata catalog step by step in the coming years. The current work can be browsed online (cf. Current metadata catalog used for Tuna Atlas). Figure 1 shows an example of search results and Figure 2 shows the details of a metadata sheet.



thredds.oreme.org:8080/geonetwork/srv/eng/catalog.search#/search?facet.q=sourceCatalog%2Fd3388135-4b53-4f75-843d-45781eeee1f4&resultType=details&fast=index&_content_type=json&from=1&to=100&sortBy=rating&any=catches

Catalogue de métadonnées de l'UMR MARBEC / OSU-OREME

catches

TYPE OF RESOURCES
☐ Dataset (37)
☐ Service (35)

KEYWORDS
☐ Fisheries (37)
☐ Tuna atlas (37)
☐ Tuna (37)
☐ Detailed data (35)
☐ Timeseries (35)
10 more

CONTACT FOR THE RESOURCE
☐ Institut de... (37)
☐ Inter-American... (12)
☐ IRD & Seychelles... (9)
☐ International... (6)
☐ Western and Central... (6)
2 more

PROVIDED BY
☑ D3388135-4b53-4f75-8... (37)

YEARS
☐ 2017 (37)

FORMATS
☐ All spatial formats... (37)

REPRESENTATION TYPES
☐ Vector (35)

UPDATE FREQUENCIES
☐ Annually (37)

STATUS
☐ Tentative (35)

SERVICE TYPES
☐ OGC Web-Service (35)

Monthly catch of tuna, tuna-like and shark species (1950-2015) by purse seiners...
This dataset lists catch of tuna, tuna-like and shark species by purse seiners and pole-and-liners from 1950 to 2015 in the Indian, Atlantic and Eastern Pacific Oceans. Catches are stratified by month, species, gear, vessel flag reporting country, fishing mode (i.e. type of school used), area (1° square) and unit of catch (weight or number). This dataset was computed using public
IRD -
Chassot Emmanuel

Monthly catch of tuna, tuna-like and shark species (1950-2015) by purse seiners...
This dataset lists catch of tuna, tuna-like and shark species by purse seiners and pole-and-liners from 1950 to 2015 in the Indian, Atlantic and Eastern Pacific Oceans. Catches are stratified by month, species, gear, vessel flag reporting country, fishing mode (i.e. type of school used), area (1° square) and unit of catch (weight or number). This dataset was computed using public
IRD -
Chassot Emmanuel

Global monthly catch of tuna, tuna-like and shark species (1950-2015)...
This dataset lists global catch of tuna, tuna-like and shark species from 1950 to 2015. Catches are stratified by month, species, gear, vessel flag reporting country, fishing mode (i.e. type of school used), area (mainly 1° or 5° square) and unit of catch (weight or number). This dataset was computed using public domain catch-and-effort datasets released by the five tuna Regional
IRD -
Chassot Emmanuel

Global monthly catch of tuna, tuna-like and shark species (1950-2015)...
This dataset lists global catch of tuna, tuna-like and shark species from 1950 to 2015. Catches are stratified by month, species, gear, vessel flag reporting country, fishing mode (i.e. type of school used), area (mainly 1° or 5° square) and unit of catch (weight or number). This dataset was computed using public domain catch-and-effort datasets released by the five tuna Regional
IRD -
Chassot Emmanuel

Global monthly catch of tuna, tuna-like and shark species (1950-2015)...
This dataset lists global catch of tuna, tuna-like and shark species from 1950 to 2015. Catches are stratified by month, species, gear, vessel flag reporting country, fishing mode (i.e. type of school used), area (mainly 1° or 5° square) and unit of catch (weight or number). This dataset was computed using public domain catch-and-effort datasets released by the five tuna Regional
IRD -
Chassot Emmanuel

Global monthly catch of tuna, tuna-like and shark species (1950-2015)...
This dataset lists global catch of tuna, tuna-like and shark species from 1950 to 2015. Catches are stratified by month, species, gear, vessel flag reporting country, fishing mode (i.e. type of school used), area (mainly 1° or 5° square) and unit of catch (weight or number). This dataset was computed using public domain catch-and-effort datasets released by the five tuna Regional
IRD -
Chassot Emmanuel

Global monthly catch of tuna, tuna-like and shark species (1950-2015)...
This dataset lists global catch of tuna, tuna-like and shark species from 1950 to 2015. Catches are stratified by month, species, gear, vessel flag reporting country, fishing mode (i.e. type of school used), area (mainly 1° or 5° square) and unit of catch (weight or number). This dataset was computed using public domain catch-and-effort datasets released by the five tuna Regional
IRD -
Chassot Emmanuel

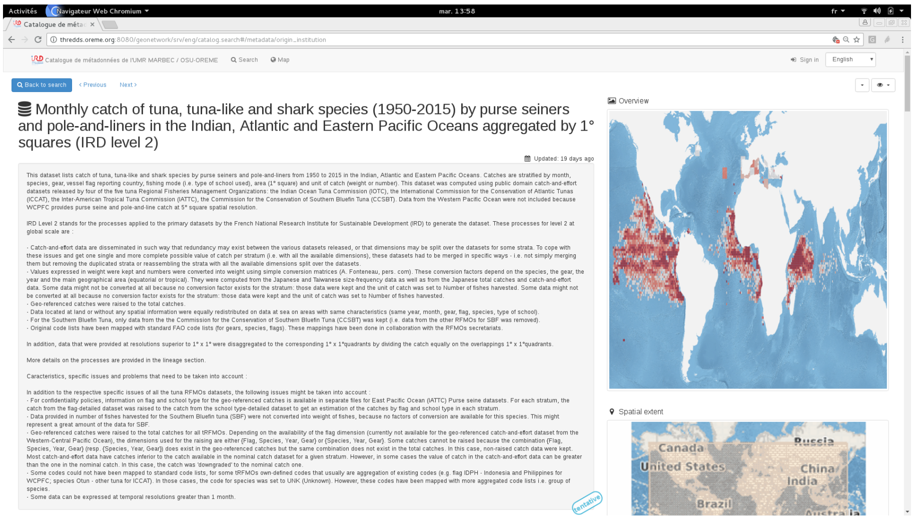
Monthly catch of tuna, tuna-like and shark species (1950-2015) by purse seiners...
This dataset lists catch of tuna, tuna-like and shark species by purse seiners and pole-and-liners from 1950 to 2015 in the Indian, Atlantic and Eastern Pacific Oceans. Catches are stratified by month, species, gear, vessel flag reporting country, fishing mode (i.e. type of school used), area (mainly 1° square) and unit of catch (weight or number). This dataset was computed using
IRD -
Chassot Emmanuel

Monthly catch of tuna, tuna-lik
in the Indian Ocean...
This dataset lists
species from the Ind
by coastal fisheries
stratified by year, m
flag reporting coun
corresponds to the
effort (task2) datase
IRD -
Fiorellato Fabio

Figure 1: Snapshot of a data search engine results.

3. Challenges for data description and data discovery

The efficiency of data discovery services highly depends on both data description efforts (i.e., the quality of metadata) and compliance with widely-used metadata standards. Though a time consuming process, there are multiple motivations to work on metadata, including increasing the number of citations of the work, enabling data discovery, and clarifying how to reuse data by describing data structure and licensing. However, the challenge is to make the metadata publication process as simple as possible when interoperability requires compliance with highly complicated metadata standards and related application programming interfaces (APIs).



Monthly catch of tuna, tuna-like and shark species (1950-2015) by purse seiners and pole-and-liners in the Indian, Atlantic and Eastern Pacific Oceans aggregated by 1° squares (IRD level 2)

Updated: 19 days ago

This dataset lists catch of tuna, tuna-like and shark species by purse seiners and pole-and-liners from 1950 to 2015 in the Indian, Atlantic and Eastern Pacific Oceans. Catches are stratified by month, species, gear, vessel flag reporting country, fishing mode (i.e. type of school used), area (1° square) and unit of catch (weight or number). This dataset was computed using public domain catch-and-effort datasets released by four of the five tuna Regional Fisheries Management Organizations: the Indian Ocean Tuna Commission (IOTC), the International Commission for the Conservation of Atlantic Tunas (ICCAT), the Inter-American Tropical Tuna Commission (IATTC), the Commission for the Conservation of Southern Bluefin Tuna (CCSBT). Data from the Western Pacific Ocean were not included because WCPFC provides purse seine and pole-and-line catch at 5° square spatial resolution.

IRD Level 2 stands for the processes applied to the primary datasets by the French National Research Institute for Sustainable Development (IRD) to generate the dataset. These processes for level 2 at global scale are :

· Catch-and-effort data are disseminated in such way that redundancy may exist between the various datasets released, or that dimensions may be split over the datasets for some strata. To cope with these issues and get one single and more complete possible value of catch per stratum (i.e. with all the available dimensions), these datasets had to be merged in specific ways - i.e. not simply merging them but removing the duplicated strata or reassembling the strata with all the available dimensions split over the datasets.
· Values expressed in weight were kept and numbers were converted into weight using simple conversion matrices (A. Fonteneau, pers. com). These conversion factors depend on the species, the gear, the year and the main geographical area (equatorial or tropical). They were computed from the Japanese and Taiwanese size-frequency data as well as from the Japanese total catches and catch-and-effort data. Some data might not be converted at all because no conversion factor exists for the stratum: those data were kept and the unit of catch was set to Number of fishes harvested. Some data might not be converted at all because no conversion factor exists for the stratum: those data were kept and the unit of catch was set to Number of fishes harvested.
· Geo-referenced catches were raised to the total catches.
· Data located at land or without any spatial information were equally redistributed on data at sea on areas with same characteristics (same year, month, gear, flag, species, type of school).
· For the Southern Bluefin Tuna, only data from the the Commission for the Conservation of Southern Bluefin Tuna (CCSBT) was kept (i.e. data from the other RFMOs for SBF was removed).
· Original code lists have been mapped with standard FAO code lists (for gears, species, flags). These mappings have been done in collaboration with the RFMOs secretariats.

In addition, data that were provided at resolutions superior to 1° x 1° were disaggregated to the corresponding 1° x 1°quadrants by dividing the catch equally on the overlappings 1° x 1°quadrants.

More details on the processes are provided in the lineage section.

Caracteristics, specific issues and problems that need to be taken into account :

In addition to the respective specific issues of all the tuna RFMOs datasets, the following issues might be taken into account :
· For confidentiality policies, information on flag and school type for the geo-referenced catches is available in separate files for East Pacific Ocean (IATTC) Purse seine datasets. For each stratum, the catch from the flag-detailed dataset was raised to the catch from the school type-detailed dataset to get an estimation of the catches by flag and school type in each stratum.
· Data provided in number of fishes harvested for the Southern Bluefin tuna (SBF) were not converted into weight of fishes, because no factors of conversion are available for this species. This might represent a great amount of the data for SBF.
· Geo-referenced catches were raised to the total catches for all tRFMOs. Depending on the availability of the flag dimension (currently not available for the geo-referenced catch-and-effort dataset from the Western-Central Pacific Ocean), the dimensions used for the raising are either {Flag, Species, Year, Gear} or {Species, Year, Gear}. Some catches cannot be raised because the combination {Flag, Species, Year, Gear} (resp. {Species, Year, Gear}) does exist in the geo-referenced catches but the same combination does not exist in the total catches. In this case, non-raised catch data were kept. Most catch-and-effort data have catches inferior to the catch available in the nominal catch dataset for a given stratum. However, in some cases the value of catch in the catch-and-effort data can be greater than the one in the nominal catch. In this case, the catch was 'downgraded' to the nominal catch one.
· Some codes could not have been mapped to standard code lists, for some tRFMOs own-defined codes that usually are aggregation of existing codes (e.g. flag IDPH - Indonesia and Philippines for WCPFC; species Otun - other tuna for ICCAT). In those cases, the code for species was set to UNK (Unknown). However, these codes have been mapped with more aggregated code lists i.e. group of species.
· Some data can be expressed at temporal resolutions greater than 1 month.

Figure 2: Snapshot of a metadata sheet.

3.1 Motivations and benefits of metadata

Metadata has increased in popularity as it has become a powerful tool to improve the citation of data collection and management, particularly through DOIs and data papers:

- DOIs: a DOI is a simple, unique identifier that enhances the ease of citing any work. Such an identifier only requires basic metadata (DataCite standard metadata elements that are basically the same as bibliographic references),

- Data papers are scientific papers which focus on the description of a data set, which has been made openly available. This is basically a literary version of metadata. There is an increasing interest for this approach, which enables data managers and authors to receive recognition for their work.

However, DOIs and data papers do not comply with rich metadata standards, which support a higher level of interoperability.

3.2 Standards for interoperability

For many years, many efforts have been made to harmonize data description by implementing metadata standards that sustain data interoperability at a global scale. Among the most widely-used standards are:

- the Dublin Core and Datacite standards (used for DOIs) that are used for basic/core or generic metadata elements (close to bibliographic references),
- Open Geospatial Consortium (OGC) standards that are dedicated to spatial information and are very rich and widely adopted worldwide,
  - general metadata about the data set (ISO 19115)
  - specific metadata describing the data structure (ISO 19110) and related services (ISO 19119)
- Ecological Metadata Language (EML) widely used for biodiversity data sharing (e.g., GBIF, OBIS) is very interesting to describe ecological aspects (data characteristics duch as species or taxa) that are not managed by OGC standards,
- Climate and Forecast conventions for NetCDF data format widely used for data related to climate and oceans (in situ sensors, remote sensing, model outputs).

Most of the time, metadata are kept separate from data but some data formats enable metadata to be embedded directly in the file (e.g., NetCDF or more common data formats such as, jpeg or png). Unfortunately, there is not a single metadata standard which fits every need. However, if possible, metadata should be written in a simple way and packaged by complying with the standards that accord to the needs of the community of users.

3.3 Mutualization of efforts and tools

Metadata are useful but not popular, and collaboration is key to achieving good descriptions. To showcase what can be achieved, we have coordinated our efforts and mutualized the tools to reach a certain level of metadata quality and quantity for the data sets that were chosen. Thereafter, any organization might choose to administer its own application if needed.

Even if the metadata standards are highly technical, their implementation should not be seen as an obstacle to describe the data sets. In the next section, we describe a method which allows keeping the description effort separated from the implementation of standards.

4. Material and Method

The method consists of using collaborative applications to facilitate the generation of metadata from data which can be managed in multiple systems:

- simple tabular data (CSV or xls) uploaded in collaborative environment (google drive, dropbox..),
- relational database management systems (SQL),
- NetCDF files usually stored on a OPeNDAP server (eg Thredds unidata server widely used by Ocean Observing Systems).

From these kinds of data sources, we created a workflow to manage the generation of rich and standardized metadata by using only simple tables as inputs.

4.1 Workflow

In our method, when data sets are extracted from data sources (like database queries or model outputs subsets) we recommend describing first the data source before its data sets since it is important to track the origin of data sets by adding a link to the related data source (called *parent identifier*).

The ignition of the workflow requires two main tables as inputs:

- description of contacts (authors),
- description of the main metadata (e.g., title, abstract, authors, keywords, spatial and temporal extent) with an additional column indicating where related data sets can be accessed.

By using simple contacts and core metadata elements, it becomes possible to fill multiple metadata elements and to skip this time-consuming task. When the related data set can be physically accessed (either by tabular data or by a SQL/OPeNDAP query or View), the data set can then be browsed and it becomes also possible to infer some other metadata elements that are common blocking points for users, such as spatial and temporal extents and lists of keywords (species, fishing gears...).

Moreover, when data and their description (metadata) are made available, it becomes possible to transform their format and manage their automated publication to access protocols like OGC WMS/WFS/WCS (for data) and OGC CSW (for metadata). Although these formatting and publishing steps are complex processes, their integration in the workflows removes this complexity and make it easy of use.

4.2 Tools to implement the workflow

Multiple tools implement widely-used metadata standards, such as OGC. APIs are available for multiple programming languages (e.g., Java and Python) and application software (e.g., Geonetwork, GeoServer or Geonode) that offer graphical user interfaces (GUIs) for those who do not need to manage data and metadata through batch automated workflows (programmatically) and can rely on manual editing and publishing tools. However, until now setting such a batch automated workflow was bound to complex programming languages and tools requiring advanced IT skills, and then not handable by a wide data management community. A methodologic objective of this work was then to overcome this huge barrier, by enabling a set of codes based on a flexible tool, less IT oriented and yet widely adopted in the data science community: the R programming language.

To achieve this, we used online tools and applications deployed in a collaborative Website (so called a Virtual Research Environment (VRE), made available by BlueBridge H2020 project). The whole work flow has been developed with R programming language. The codes are accessible online and can be compiled with an RStudio server directly accessible in the VRE Website as well as a Geonetwork metadata catalog and spatial data servers (Geoserver and Thredds)

The tables storing information about contacts and metadata about data sets are made available online either as simple tabular data (CSV files uploaded in e.g., google spreadsheets or dropbox) or directly as a dedicated table within the physical model of a SQL database (Postgres & Postgis in the case of the Tuna Atlas VRE).

4.3 Tabular data (CSV or SQL)

The suite of R packages *geometa* **?**, *geosapi* **?**, *geonapi* **?** have been used to convert metadata to comply with OGC compliant standards as well as to push the metadata in dedicated catalogs (e.g. Geonetwork). Each described data set is made available through OGC Web Services (WMS & WFS) with Geoserver (using *geosapi* R Package).

4.4 NetCDF files and Thredds server

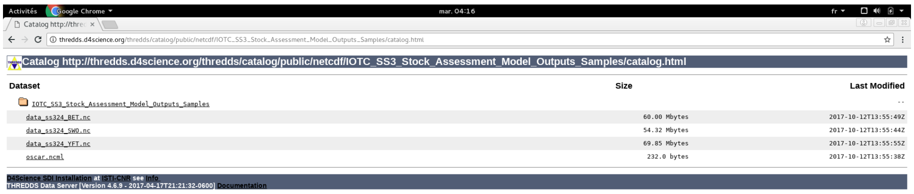
In the case of NetCDF files, we implement this workflow in the following way:

- metadata are read directy from the headers of the NetCDF files which are read remotely on a Thredds server by using the OPeNDAP access protocol (see samples),
- the metadata in the NetCDF-CF headers (compliant with Climate and

Forecast Conventions) are transformed into OGC metadata using a mapping.

- Metadata can be pushed in Geonetwork using *geonapi*. However since NetCDF files managed with Thredds are available with OPeNDAP (and WMS, WCS when relevant) we didn't use *geosapi* R package in the case of Stock Assessment model outputs (SS3).

The Figure shows an example of Thredds server giving access to some Stock Assessment model outputs which can be read remotely by our R workflow and generate metadata (as shown in Figure 4)


Activités
Google Chrome
mar. 04:16
fr
Catalog http://thre
thredds.d4science.org/thredds/catalog/public/netcdf/IOTC_SS3_Stock_Assessment_Model_Outputs_Samples/catalog.html

Catalog http://thredds.d4science.org/thredds/catalog/public/netcdf/IOTC_SS3_Stock_Assessment_Model_Outputs_Samples/catalog.html

Dataset	Size	Last Modified
IOTC_SS3_Stock_Assessment_Model_Outputs_Samples		--
data_ss324_BET.nc	60.00 Mbytes	2017-10-12T13:55:49Z
data_ss324_SWO.nc	54.32 Mbytes	2017-10-12T13:55:44Z
data_ss324_YFT.nc	69.85 Mbytes	2017-10-12T13:55:55Z
oscar.ncml	232.0 bytes	2017-10-12T13:55:38Z

D4Science SDI Installation at ISTI-CNR see Info
THREDDS Data Server [Version 4.6.9 - 2017-04-17T21:21:32-0600] Documentation

Figure 3: Snapshot of a Thredds server with stock assessment model outputs

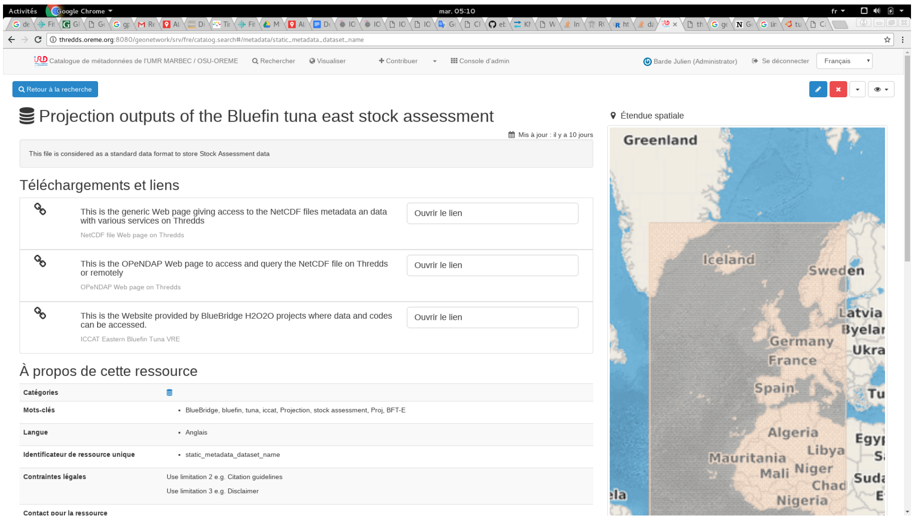
5. Results and Discussion

5.1 Metadata

This approach has been successfully tested in the following use case by implementing the method described in section 4.:

- Catch and Efforts data sets as provided by tuna RFMOs (for both data & code lists managed in the context of the Tuna Atlas project with FAO and IRD within the framework of BlueBridge H2020 project): metadata

- SS3 Stock Assessment model outputs in IOTC (coordinated with Ifremer and ICCAT VPA Stock Assessment of bluefin Tuna in the context Blue-Bridge H2020 project). As many metadata as stock assessment model



thredds.oreme.org:8080/geonetwork/srv/fre/catalog.search#/metadata/static_metadata_dataset_name

Projection outputs of the Bluefin tuna east stock assessment

Mis à jour : il y a 10 jours

This file is considered as a standard data format to store Stock Assessment data

Téléchargements et liens

This is the generic Web page giving access to the NetCDF files metadata an data with various services on Thredds
NetCDF file Web page on Thredds
Ouvrir le lien

This is the OPeNDAP Web page to access and query the NetCDF file on Thredds or remotely
OPeNDAP Web page on Thredds
Ouvrir le lien

This is the Website provided by BlueBridge H2O2O projects where data and codes can be accessed.
ICCAT Eastern Bluefin Tuna VRE
Ouvrir le lien

À propos de cette ressource

Catégories	
Mots-clés	• BlueBridge, bluefin, tuna, iccat, Projection, stock assessment, Proj, BFT-E
Langue	• Anglais
Identificateur de ressource unique	• static_metadata_dataset_name
Contraintes légales	Use limitation 2 e.g. Citation guidelines Use limitation 3 e.g. Disclaimer
Contact pour la ressource	

Figure 4: Snapshot of a metadata sheet for Stock Assessment model outputs

outputs for a given species and a given year (one per run / parameterization). This might be reduced to one metadata for all runs for a given species and a given year.

- RTTP database (ongoing): ongoing development of metadata sheets.

So far, we have been able to generate several metadata but a lot of time has been spent on the methods more than metadata themselves. Since the core of workflow is now in place, it should become easier to enrich the metadata catalog in the coming years:

- update and enrich existing metadata,
- to set up similar workflows for other data sources.

Moreover, since the catalog implements the OGC CSW standard, it can be harvested and reused by any clients (other metadata catalogs, GIS Desktop application and Web browser).

5.2 *Data servers*

Since we handle the data to generate some good quality metadata, we have been able to transform data formats and make data accessible with standardized access protocols:

- geoservergive access to data with OGC Web Services (WMS, WFS).
- Thredds

Data can be now be accessed programmatically filtered and downloaded in multiple data formats.

5.3 Visualization

Once metadata are standardized, it becomes possible to build applications on top of metadata. We built two applications:

- Shiny applications which use metadata to access the data in Stock Assessment outputs, browse data and plot some variables,
- Prototype Tuna Atlas viewer (in development) using rich OGC metadata to discover, filter and display maps of Catch and Efforts data.

6. Outlooks, future work with IOTC

For IOTC, an option to foster this approach in the coming years might be to synchronize the Working party registration process with the identification (inventory) of data sets related to submitted papers and presentations. By doing so, it would be easier to link authors with their documents and some basic metadata describing the main characteristics of underlying data sets (Title, Summary, Spatial and Temporal extent, Keywords, Species, Fishing Gears..). This might easily be done by using google spreadsheets to store and collaboratively edit details about both contacts and metadata. That would be an opportunity as well to ask the authors if they expect a DOI for the presentation and / or the related data-sets. This work can be related to the OpenAire-connect H2020 project which focus on Fisheries and Aquaculture domain and is related to Zenodo infrastructure to archive, describe (DataCite) and manage documents, data or codes with DOIs. In case of interest, this might be the follow up of this work for the next WPDCS and could be presented in similar contexts (SWIOFC, SIOFA, CCAMLR, etc.).

There is as well an open question about the need for IOTC to administrate a dedicated metadata catalog since it could be managed a higher level (e.g. CWP, FAO, IOOS . . .) or through a VRE like the Tuna Atlas. Indeed as other RFMOs manage the same type of data it would make sense for them to mutualize tools and servers and, by doing so, facilitate the (meta-) data access for users.

In the particular case of Stock Assessment model outputs a specific link with RAM legacy database team might be relevant to enable data discovery (Geonetwork metadata catalog) and data access (Thredds / NetCDF data server)

Acknowledgements

This work has received funding from the European Union's Horizon 2020 research and innovation programme under the BlueBRIDGE project (Grant agreement No 675680).