



FAO Tuna Atlas renewal

Tools and Services

Emmanuel Blondel (FAO) emmanuel.blondel@fao.org

Yann Laurent (FAO) yann.laurent@fao.org

Paul Taconet (IRD) paul.taconet@ird.fr

1. FAO Tuna Atlas – Historical background and Challenges

2. FAO Tuna Atlas Renewal – Emerging Tools & Services

2.1 Data collation / Harmonization / Import

2.2 (Meta)data publication

2.3. Data exploitation tools (query, extraction, map)



The FAO Tuna Atlas

–

1. Historical Background & Challenges

The FAO Tuna Atlas

Atlas of Tuna and Billfish Catches

The atlas of tuna and billfish catches is an interactive web-based mapping tool. It uses geographic information system (GIS) technology to enable users to map and display global catches of major tunas and billfishes by fishing gears at a 5° by 5° degrees resolution. For background information, please [read here](#).



Select tuna and billfish statistics.
Select year and quarter ranges, one more gear types and species and aggregation method. [Full instruction...](#)

Query parameters **Animations**

Select

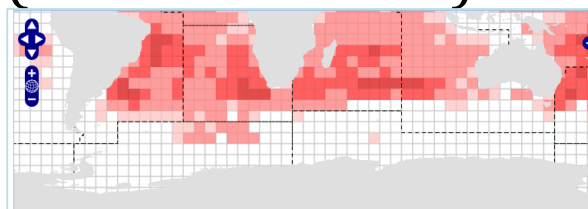
Gear Types Species

Aggregation method

- Cumulative tuna and billfish catches across years
- Average tuna and billfish catches across years
- Average tuna and billfish catches across quarters

The FAO Tuna Atlas is part of two historical activities of FAO to provide users information on tuna catches:

1. A dynamic map representation of the evolution of tuna and billfishes catches (FAO Tuna Atlas) 1950 – 2012 by 5° square



2. Global tuna Nominal catches (on-line query)



In both cases, source of information were tuna Regional Fisheries Management Organizations (t-RFMOs) statistics.

The process of data collation from t-RFMOs was:

- Based on a gentlemen agreement between FAO and t-RFMOs (no formal one)



- Done on ad-hoc basis by FAO staff (Fabio Carocci) from available data



- No transformation of data except aggregation of statistics according to the Tuna Atlas classifications:
 - TA stocks classification (ex: Albacore Indian Ocean, Skipjack tuna in the East Pacific etc...)
 - TA area classification (5x5 grid ID)
 - TA gear classification (Longline, Pole and line, Purse seine, others)
 - Country

Collation of t-RFMOs was a challenging task:

- No harmonized format for tuna catches among organizations (normalized / denormalized data)
 - No harmonized classification (difference in 5x5 square definition)
 - No notification when new data are published
- ➔ The process to compile t-RFMOs was highly manual.



A first attempt to automatize the procedure was done in 2012 with the assessment of the potential of the iMarine tool ICIS.

This assessment was also done in a time where the format of available time series from the t-RFMOs was quite stable.

The result of the assessment was recommendations for tool improvements. The Tabular Data Manager (TabMan) was created and released in 2014.

Unfortunately, Fabio Carocci left before the tool could be fully used.

There were still pending issues:

A process still manual: harmonization and curation process still requires human action



Without formal agreement on the format, there was still **a risk of changes in format** without notification





FAO Tuna Atlas Renewal

–

2. Emerging Tools & Services





Objectives

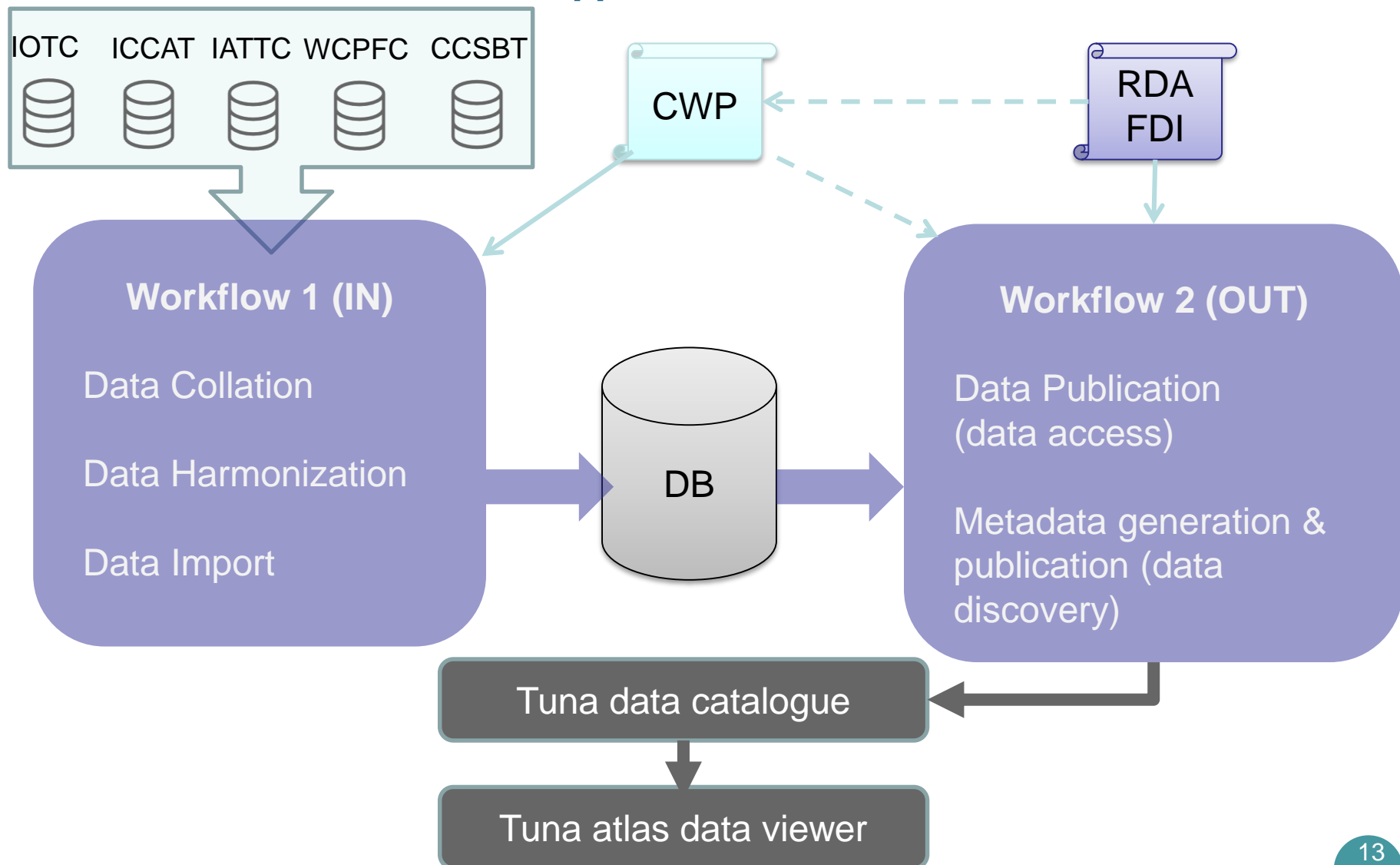
- **To renew (and possibly go beyond) the current FAO Tuna datasets & services**, drawing a pathway for an **open data portal on global tuna fisheries**
- To **foster reference data harmonization continuous process** (important role of working groups – RDA FDI and CWP) **in close interaction with Tuna RFMOs**
- **To build a partnership among institutions in order to avoid duplication of work / foster mutualization of efforts** (“why re-doing something that has been already done?”)
- To foster exchanges between key actors of Tuna data management (regional data managers, statisticians, data scientists / managers) through a single collaborative environment
- To provide a concrete technical framework for setting-up a shared Tuna Data Management Plan including
 - A **collation/harmonization flow** between the t-RFMOs and FAO, thanks to harmonization outputs (code list mappings), reproducibility, versioning, and the openness, sharing and scalability of the data harmonization processes.
 - A semi-automated **(meta)data publication flow** enabling a catalogue of Tuna fisheries data products, driven by FAIR principles and implementing internationally recognized open (meta)data standards



Objectives

- To **promote a better visibility of Global Tuna datasets and transparency of data processes** used behind, through the provision of **comprehensive metadata**
- To **facilitate and boost the use of global and regional Tuna datasets by user communities** (stakeholders, statisticians, scientists, universities, students)
- Hence, to **improve recognition of Tuna RFMOs as source data providers**
 - proper roles assignment through datasets metadata (ownership on source metadata records); citation through compiled global Tuna datasets metadata; data provenance/traceability
 - New Tuna atlas data extraction/visualization tools that can be contextualized for RFMOs. Notion of “embedded data tool” for use in RFMO websites.

Overall Approach & Architecture

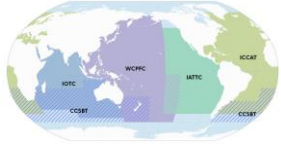




2.1 Data collation, Harmonization, Import

The Global Tuna Atlas project: general objectives

Tuna RFMOs areas of management



IOTC



ICCAT



IATTC



WCPFC



CCSBT



Tuna RFMOs
websites (public
domain datasets)

1
Collection
& Harmonization

2
Import & storing
in SQL database

Open SQL
database
storing data +
metadata

3
Compute global
datasets on tuna
fisheries

Global datasets on
tuna fisheries



4
Extraction and data
services for easy
exploitation



Discover available data
*Which datasets exist?
How they were built?*



Access the data
How to access the data?



Process the data
How to create own tuna atlas?

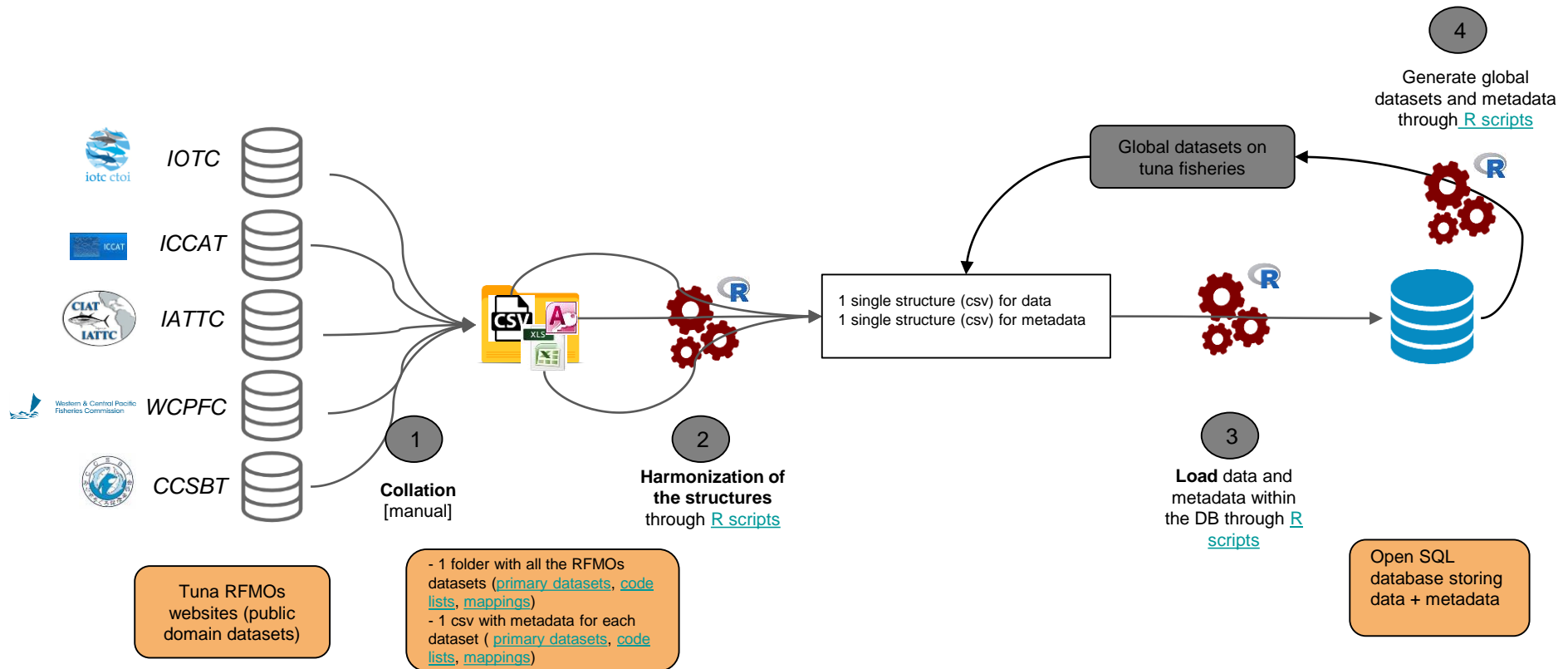


Visualize the data
How to easily create maps, plots?

Tools to exploit
the data

How it is set up

Data and metadata collation & harmonization & import





The datasets available

The datasets gathered in the database are:

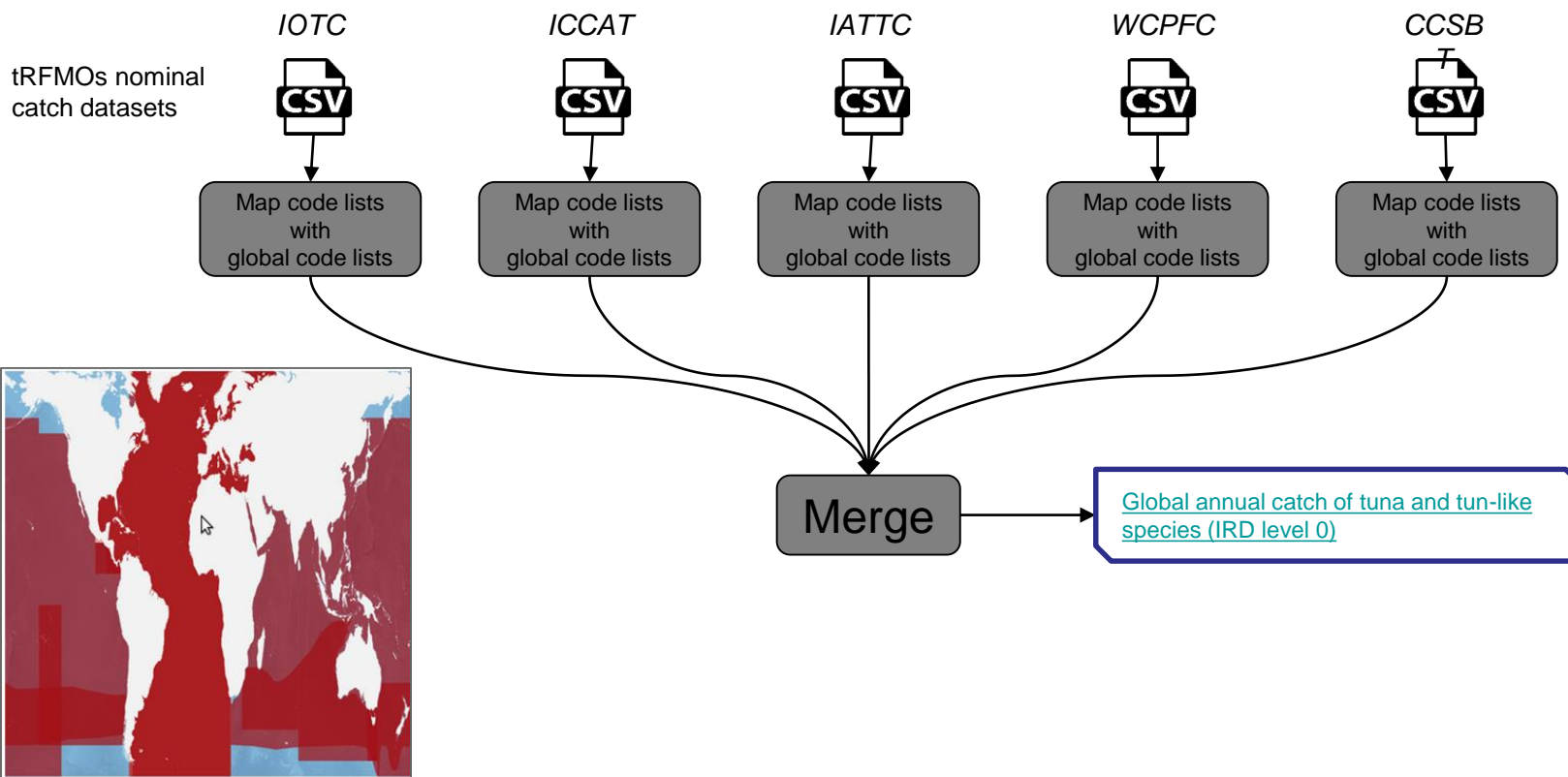
- The **public domain datasets from the 5 tuna RFMOs** (IOTC, ICCAT, IATTC, WCPFC, CCSBT) as they deliver them (i.e. without processings):
 - Nominal catch (RFMO area of competence / 1 year) (e.g. [IOTC's](#))
 - Georeferenced catch (5° / 1 month) (e.g. [ICCAT's](#))
 - Georeferenced effort (5° / 1 month)
 - Catch-at-size (5° / 1 month) ?
- **Global datasets on tuna fisheries (built by IRD)**, built by merging the regional datasets and applying some scientific corrections (e.g. conversion from number of fishes harvested to weight) to get a more pertinent overview of tuna fisheries at global scale:
 - Global nominal catch (e.g. [here](#))
 - Global georeferenced catch (e.g. [here](#))
 - Global georeferenced effort
 - Global georeferenced catch-at-size ?
- The **code lists used by the 5 tuna RFMOs** (for gears, species, fishing countries, etc.) (e.g. [ICCAT's gears](#)) + global code lists recommended by the CWP (e.g. [ASFIS](#), [ISSCFG](#))
- The **mappings between tuna RFMOs code lists and global code lists**, which are necessary to combine the datasets expressed with sparse code lists (e.g. [IOTC's gears to ISSCFG](#))

+ **Detailed metadata for each dataset**

Title, abstract, contacts,
genealogy (i.e. which source datasets were used to generate the dataset),
lineage (i.e. how the data was generated))

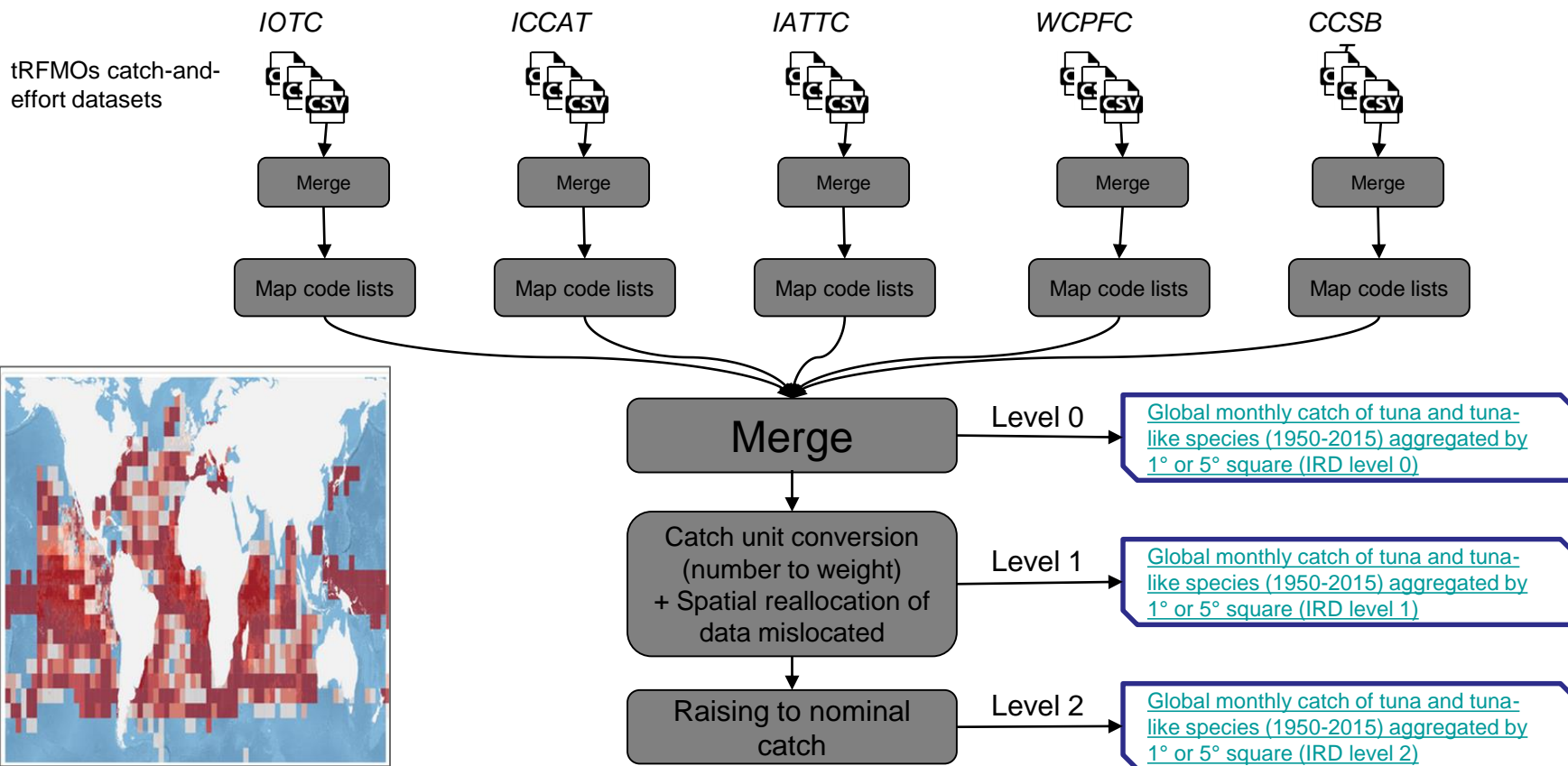
Global datasets on tuna fisheries: how they were generated (main processings)

I. Global nominal (total) catch



Global datasets on tuna fisheries: how they were generated (main processings)

I. Global geospatial catch



Create your own dataset of global catch or efforts

The screenshot shows the Dataminer web interface. The main content area is titled "Create Your Own Georeferenced Catch Tuna Atlas Dataset". Below the title, there is a description: "This algorithm allows to create own regional or global tuna atlas. It takes as input the public domain datasets of the five Tuna Regional Fisheries Management Organizations (RFMOs) (IOTC ICCAT WCPFC IATTC CCSBT) stored within the Tuna atlas database. It proposes a set of parameters to customize the computation of the tuna atlas."

The parameters section includes the following:

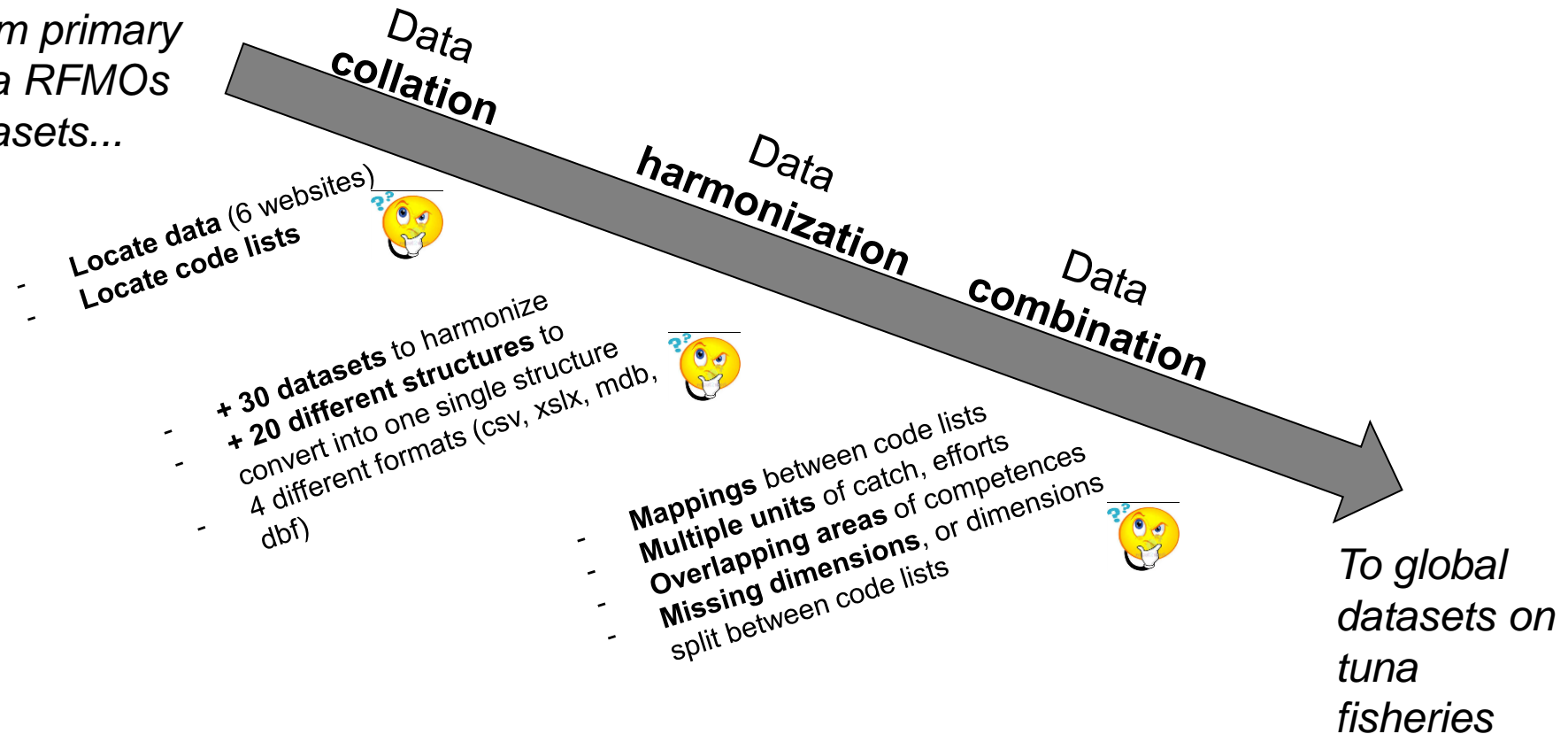
- include_IOTC:** TRUE (String Value) - Include IOTC data (Indian Ocean) in the atlas (TRUE or FALSE)
- include_ICCAT:** TRUE (String Value) - Include ICCAT data (Atlantic Ocean) in the tuna atlas?
- include_IATTC:** TRUE (String Value) - Include IATTC data (Eastern Pacific Ocean) in the tuna atlas?
- include_WCPFC:** TRUE (String Value) - Include WCPFC data (Western Pacific Ocean) in the tuna atlas?
- include_CCSBT:** TRUE (String Value) - Include CCSBT data (Southern hemisphere Oceans - only Southern Bluefin Tuna) in the atlas?
- datasets_year_release:** 2017 (String Value) - Year of release of the datasets by the RFMOs. First available year is 2017. Usually datasets released in the year Y contain the time series from the beginning of the fisheries (e.g. 1950) to year Y-2 (included). For instance 2017 will extract the datasets released in 2017 and that cover the time series from 1950 to 2015 (included)
- iccat_include_type_of_school:** TRUE (String Value) - Concerns ICCAT Purse Seine data. Use only if parameter include_ICCAT is set to TRUE. ICCAT disseminates two catch-and-efforts datasets: one that provides the detail of the type of school (Fad/Free school) for purse seine fisheries only and that starts in 1994 (called Task II catch/effort by operation mode Fad/Free school) and one that does not provide the information of the type of school and that covers all the time period (from 1950) (called Task I catch/effort). These data are redundant (i.e. the data from the dataset Task II catch/effort by operation mode are also available in the dataset Task I catch/effort but in the later the information on the type of school is not available). Combine both datasets to produce a dataset with fishing mode information (Fad / Free school)? TRUE : both datasets will be combined to produce a dataset with fishing mode information (Fad / Free school). FALSE : Only the dataset without the type of school will be used. In that case the output dataset will not have the information on fishing mode and flag for each stratum. FALSE : Only one dataset will be used (either Flag or Type of school). The parameter dimension_to_use_if_no_raising_flags_to_schooltype allows to decide
- iattc_raise_flags_to_schooltype:** TRUE (String Value) - Concerns IATTC Purse Seine data. Use only if parameter include_IATTC is set to TRUE. For confidentiality policies information on fishing country (flag) and type of school for the geo-referenced catches is available in separate files for the Eastern Pacific Ocean purse seine datasets. IATTC hence provides two public domain dataset: one with the information on the type of school and one with the information on the flag. Both datasets can be combined - using raising methods - to estimate the catches by both flag and type of school for purse seine fisheries. Combine both datasets? TRUE : both datasets (by fishing mode and by flag) will be combined - using raising methods (i.e. dataset with the information on the flag will be raised to the dataset with the information on the type of school - to have the detail on both fishing mode and flag for each stratum. FALSE : Only one dataset will be used (either Flag or Type of school). The parameter dimension_to_use_if_no_raising_flags_to_schooltype allows to decide
- iattc_dimension_to_use_if_no_raising_flags_to_schooltype:** (String Value) - Concerns IATTC Purse Seine data. Use only if parameter iattc_raise_flags_to_schooltype is set to FALSE. In the case IATTC purse seine datasets are not combined (see description of parameter iattc_raise_flags_to_schooltype) which dataset to use? Flag : use dataset with the information on flag. Information on type of school will therefore not be available. schooltype : use dataset with the information on type of school. Information on flag will therefore not be available.
- mapping_map_code_lists:** TRUE (String Value) - Map code lists (gear/species/flag/schooltype/catchtype)? When using multiple sources of data (i.e. multiple RFMOs) code lists used by the various RFMOs might be different. They should therefore be mapped to single code lists in order to be able to compare the data. TRUE : map code lists. The url to the datasets providing the code list mappings to use must be set in the parameter mapping_source_mappings. See parameter mapping_source_mappings for more details. FALSE : do not map code lists. Output data will use input codes.
- mapping_csv_mapping_datasets_url:** https://goo.gl/2hA1sq (String Value) - Use only if parameter mapping_map_code_lists is set to TRUE. Path to the CSV file containing the dimensions that must be mapped and the name of the mapping dataset for each dimension mapped. The mapping datasets must be available in Tuna atlas database. An example of this CSV can be found here: https://goo.gl/2hA1sq.

Underlying R scripts available online:

- [Create your own tuna atlas of nominal catches](#)
- [Create your own tuna atlas of georeferenced catch or efforts](#)

Challenges met in data collation / harmonization / combination

From primary
tuna RFMOs
datasets...



The key: a harmonized (standard) structure for data and metadata

Data (example of structure for catches)

flag	gear	geographic_identifier	time_start	time_end	species	catchtype	schooltype	unit	value
BLZ	PS	5402000	2009-08-01	2009-09-01	BET	C	fs	MT	9.51
BLZ	PS	5402000	2009-08-01	2009-09-01	YFT	C	fs	MT	98.58
BLZ	PS	5202006	2009-09-01	2009-10-01	BET	C	fd	MT	0.38
BLZ	PS	5202006	2009-09-01	2009-10-01	SKJ	C	fd	MT	15.76
BLZ	PS	5202006	2009-09-01	2009-10-01	YFT	C	fd	MT	2.65
...

Each dataset comes with related **metadata**

→ See example of metadata (content and structure) [here](#)



Wrap - up

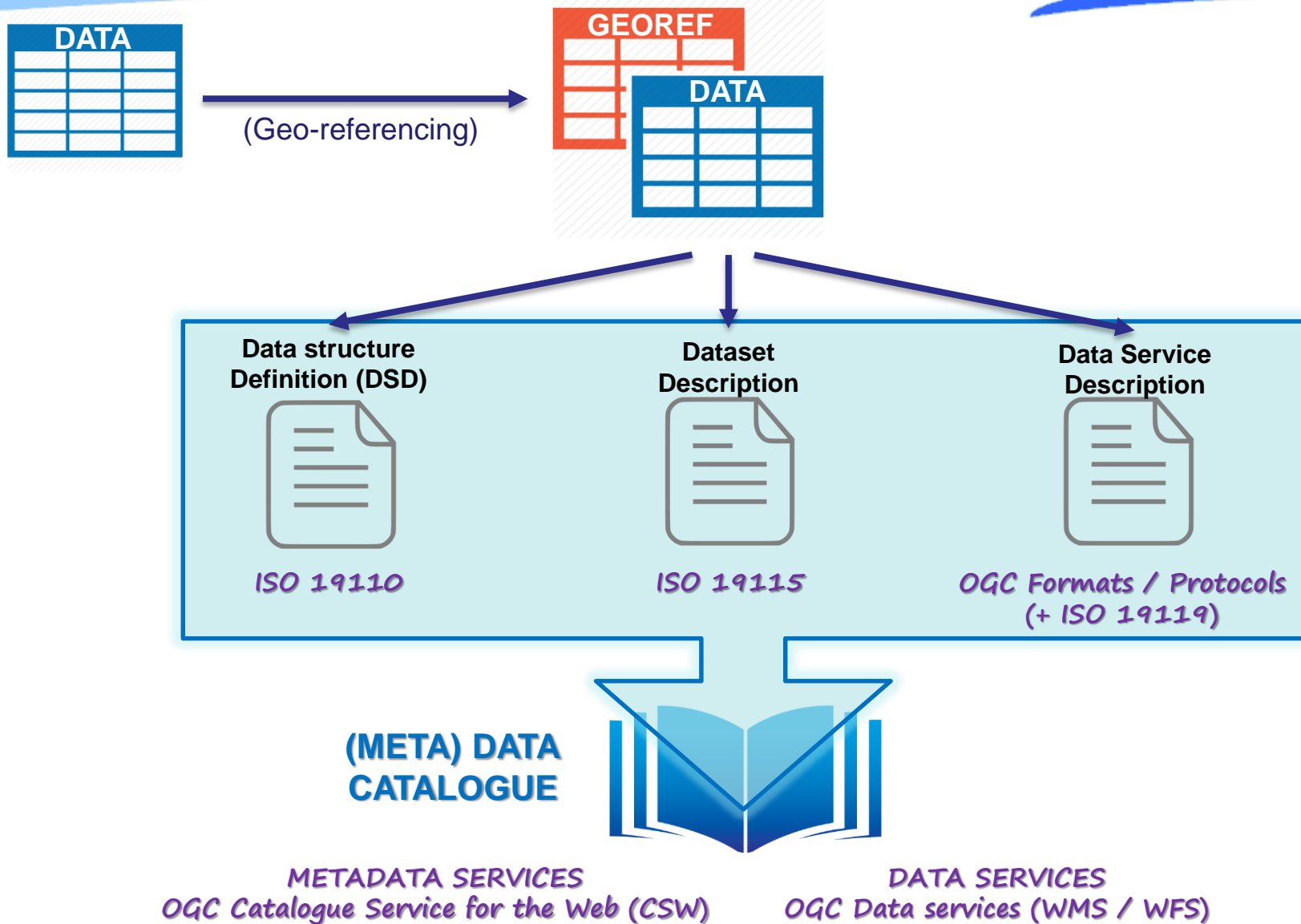
- We have created global datasets on tuna fisheries (nominal catch, geo-referenced catch and efforts)
- For this, we have i) collated and ii) harmonized the structure of the public datasets of the 5 t-RFMOs (nominal catch, catch-and-efforts, code lists, code lists mappings)
- We have stored all the data (both regional and global) + metadata in a dedicated database that is open
- We have created scripts to parameterize the computation of global datasets from regional datasets
- The whole work is reproducible yearly as long as the structure of the source data does not change OR that the t-RFMOs provide the data + metadata in the harmonized structure
- The whole work is reproducible yearly as long as the structure of the source data does not change OR that the t-RFMOs provide the data + metadata in the harmonized structure
- Once the data + metadata are in the “harmonized” structure, they can be uploaded in the DB and all the services can be derived (catalogue, viewer, R functions to process the data, etc.)

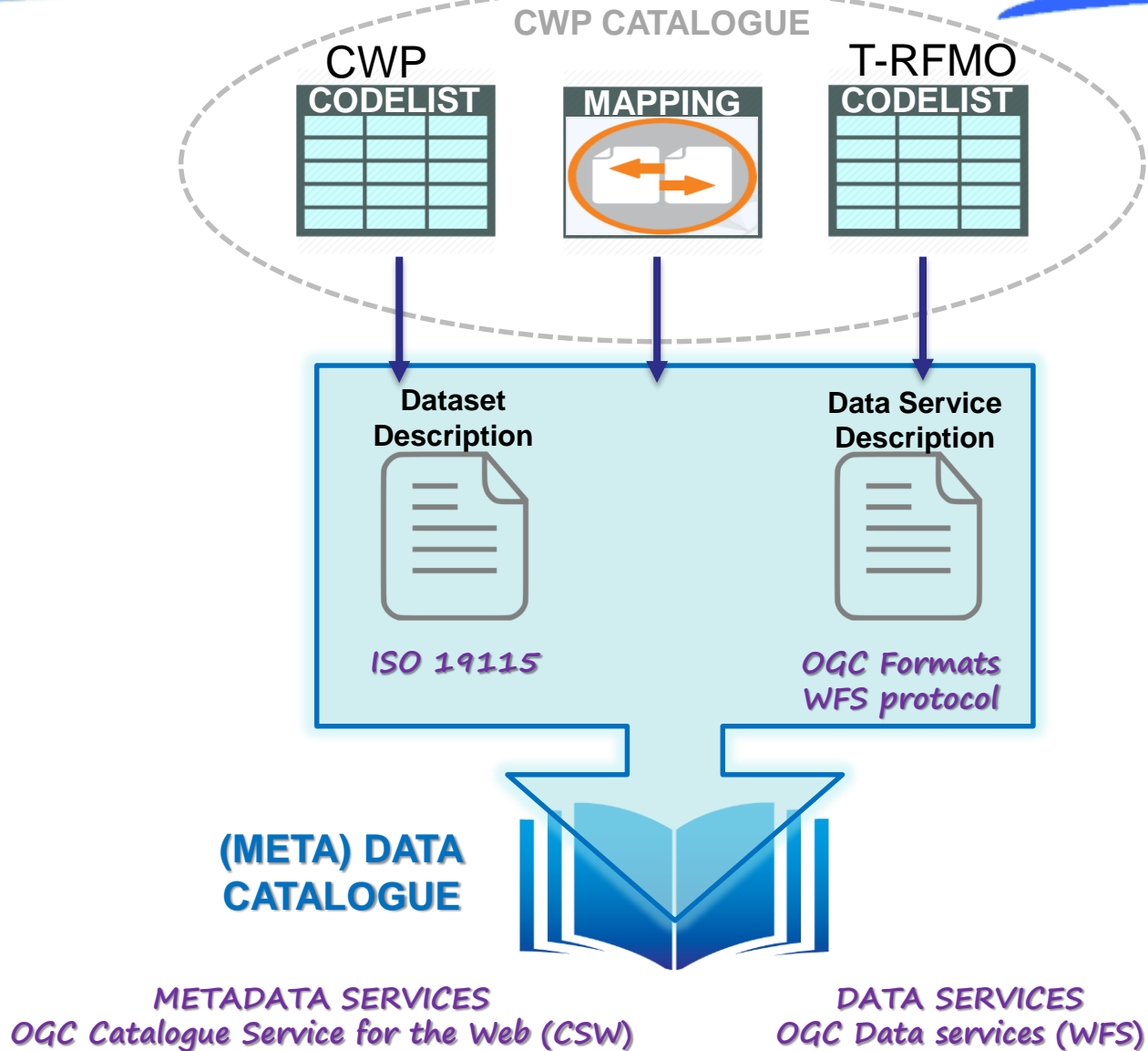


2.2 (Meta)data publication

–

Enabling the Global Tuna Atlas Catalogue





<https://tunaatlas.d4science.org/geonetwork/srv/eng/catalog.search#/home>

BlueBRIDGE Global Tuna Atlas catalogue
Q Search
Map
Sign in
English

⋮
Q
×

TYPE OF RESOURCES

- Dataset (60)
- Service (60)

KEYWORDS

- Detailed data (60)
- Timeseries (60)
- Series (60)
- Tuna (60)
- Fisheries (60)
- [10 more](#)

CONTACT FOR THE RESOURCE

- Institut de... (60)
- Inter-American... (19)
- Western and Central... (17)
- IRD & Seychelles... (7)
- Indian Ocean Tuna... (7)
- [2 more](#)

PROVIDED BY

- C6c6f858-ce96-4a02-9... (60)

YEARS

- 2018 (60)

FORMATS

- All essential formats (60)

Sorted by relevancy
1 - 20 on 60
<
>

Categories
☆☆☆☆☆ BlueBRIDGE

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 IRD -

Chassot Emmanuel

Categories
☆☆☆☆☆ BlueBRIDGE

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 (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

Categories
☆☆☆☆☆ BlueBRIDGE

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

IRD -

Categories
☆☆☆☆☆ BlueBRIDGE

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

IRD -

27

Technical workshop on global harmonization of Tuna fisheries statistics Rome, Italy – 19-22 March 2018



2.3 Data Exploitation Tools

–

Enabling Data query, extraction & maps



**(META) DATA
CATALOGUE**



*METADATA SERVICES
OGC Catalogue Service for the Web (CSW)*

*DATA SERVICES
OGC Data services (WMS / WFS)*

FINDABLE

ACCESSIBLE

INTER
OPERABLE

RE-USABLE

QUERY

EXTRACTION

MAPS

SHARING
EMBEDDING



SHARE
EMBEDD

<https://tunaatlas.d4science.org/tunaatlas>

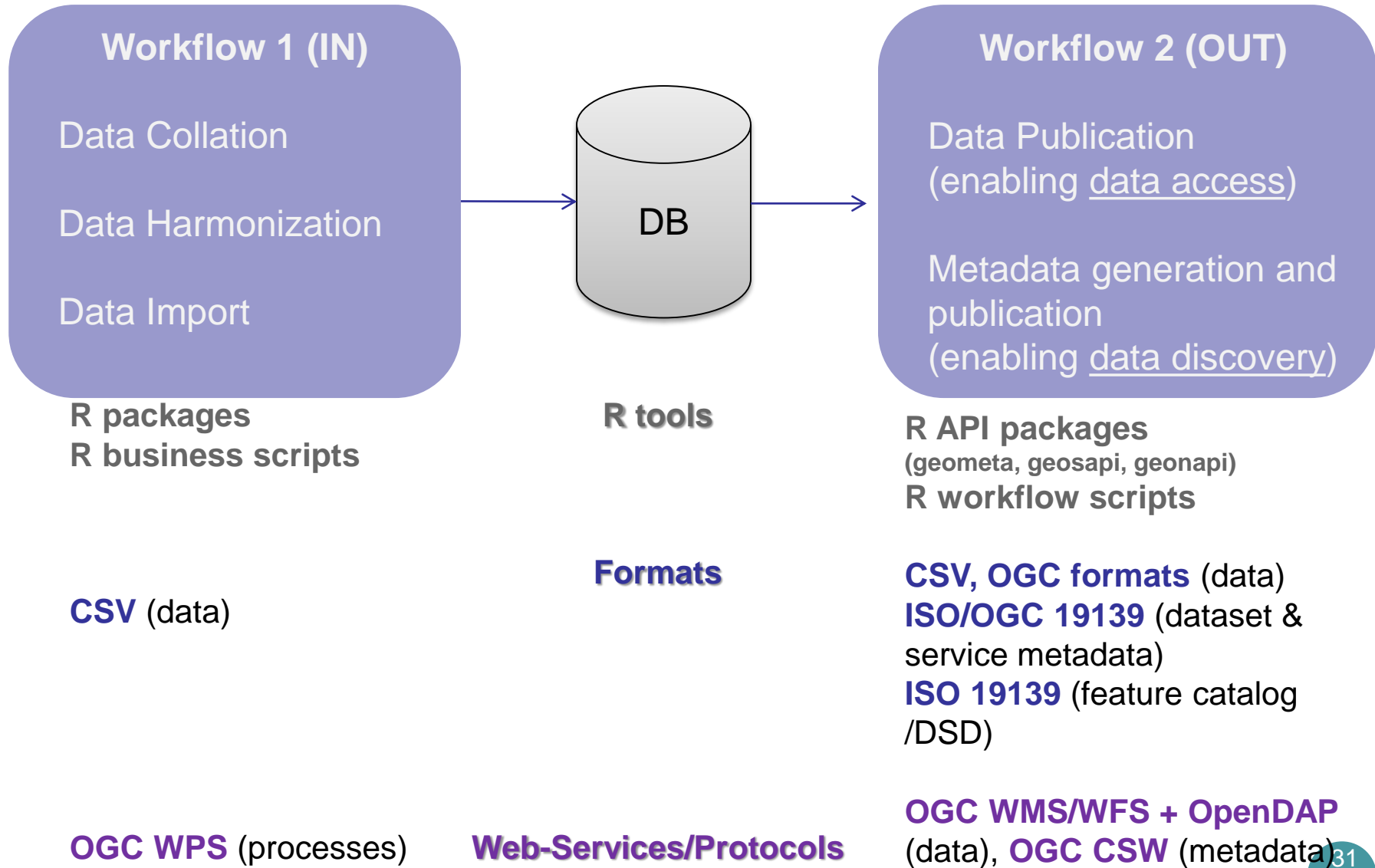
MAP

QUERY

EXTRACT

The screenshot displays the Tuna Atlas web application interface. On the left, a map of the Pacific Ocean shows tuna catch data with a color scale from light orange to dark red. A yellow arrow labeled 'MAP' points to the map area. In the center, a 'Query a dataset' panel is open, showing a search bar with the query 'global_catch_1deg_1m_ps_bb_tunaatlasIRD_level2'. Below the search bar, there are sections for 'Fishery dimensions' (including Inter-American Tropical Tuna Commission and Purse seines), 'Temporal extent' (1976-2016), 'Aggregation method' (Sum), and 'Map options' (5). A yellow arrow labeled 'QUERY' points to the fishery dimensions section. At the bottom of the query panel, there are three icons: 'AGG', 'RAW', and 'MAP'. A yellow arrow labeled 'EXTRACT' points to the 'MAP' icon. On the right, a 'Legend' panel titled 'TUNA MAPS' shows a color scale for '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)'. The legend includes a 'BASE OVERLAYS' section with options for 'Marine areas' (Grid 5x5, Grid 1x1, FAO major areas, EEZ boundaries) and a 'BASEMAPS' section with options for 'Oceans imagery', 'ESRI World Imagery', and 'ESRI - Countries'. A yellow arrow labeled 'SHARE EMBEDD' points to the top right of the interface.

Tools, Formats and Services



OGC WPS (processes)

Web-Services/Protocols

OGC WMS/WFS + OpenDAP (data), OGC CSW (metadata)



Links

- **Tuna Atlas Data viewer:** <https://tunaatlas.d4science.org/tunaatlas/index.html>
- **Tuna Atlas Data Catalogue:** <https://tunaatlas.d4science.org/geonetwork/srv/eng/catalog.search#/home>
- **Metadata** of the data available in the DB (as Gdoc spreadsheets):
 - [Code lists](#)
 - [Code lists mappings](#)
 - [Source tuna RFMOs datasets](#)
 - [Global datasets](#)
- **Data** available in the DB (as csv files):
 - [Code lists](#)
 - [Code lists mappings](#)
 - [Source tuna RFMOs datasets & Global datasets](#)
 - [Example of DOIs on IRD datasets](#)
- **Documents** and papers:
 - [Global Tuna atlas: Achievable global research and fisheries management objectives](#)
 - [Information paper on the IRD global tuna fisheries datasets \(i.e. processes applied to generate global datasets\)](#)
 - [Public domain tuna RFMOs datasets to collate yearly to generate the global tuna atlas](#)
- **Scripts:**
 - [R scripts to harmonize the tRFMOs datasets structures](#)
 - [R scripts to create global datasets \(with ability to parameterize the data combination processes\)](#)
 - R package [‘rtunaatlas’](#) to discover, access and process the data ([tutorial](#) on how to use the package)



Thank you for your attention