



Fisheries Data Interoperability

—

Geographic Information (Meta)data standards

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



Background – Research Data Alliance, FDI working group

(Meta)data information standards

FAIR (meta)data principles

**Methodology for cross-domain geo-referenced statistics
dissemination & visualization**

Use Case 1: Global Tuna Atlas

Use Case 2: Regional database – WECAFC-FIRMS Fisheries



• RDA – Research Data Alliance

- “Research Data Without barriers” – <https://www.rd-alliance.org/about-rda>
- **Vision:** *“Researchers and innovators openly sharing data across technologies, disciplines, and countries to address the grand challenges of society”*
- **Community:** 6600 members from 135 countries
- **General findings**
 - Data is fragmented worldwide, within and across disciplines and domains
 - Need of building blocks and data bridges to foster data discovery, access, exploitation including cross-disciplinary approaches



- **Fisheries Data Interoperability Working Group**
 - <https://rd-alliance.org/group/fisheries-data-interoperability-wg/>
 - Participation of FAO and IRD under EC BlueBridge project
 - **Synergy with Coordinating Working Party on Fishery Statistics**
 - Adhoc task group of reference data harmonization
 - GIS working group
 - **Activities**
 - **Inventory of (meta)data standards in support to fisheries data**
 - Statistical (meta)data exchange standards
 - Geo-referenced (meta)data exchange standards
 - Discuss a **draft standard vocabulary for defining data structures & variables, categorizing datasets**. Notion of “building blocks”
 - Draft a **generic methodology for geo-referenced cross-domain statistical data dissemination & visualization using open standards**
 - Knowledge exchange with/from other RDA interest & working groups

• **(Meta)data standards for fisheries data interoperability?**

- **Statistics-oriented standards**

- SDMX Information Model; <https://sdmx.org/>
- UN-CEFACT FLUX

- **Geospatial-enabled standards**

- ISO/TC211 Geographic information Standards
- [OGC \(Open Geospatial Consortium\)](#) Format & Service Standards

- Generic Standards: [Dublin Core](#)

- But also standards in specific domains: e.g. [EML – Ecological Metadata Language](#)

- (Meta)data standards are part of policy frameworks

Examples:

- Statistics: Commission Recommendation of 23 June 2009 on reference metadata for the European Statistical System 3 (2009/498/EC)
- Geospatial: EC INSPIRE Directive – 2007/2/CE

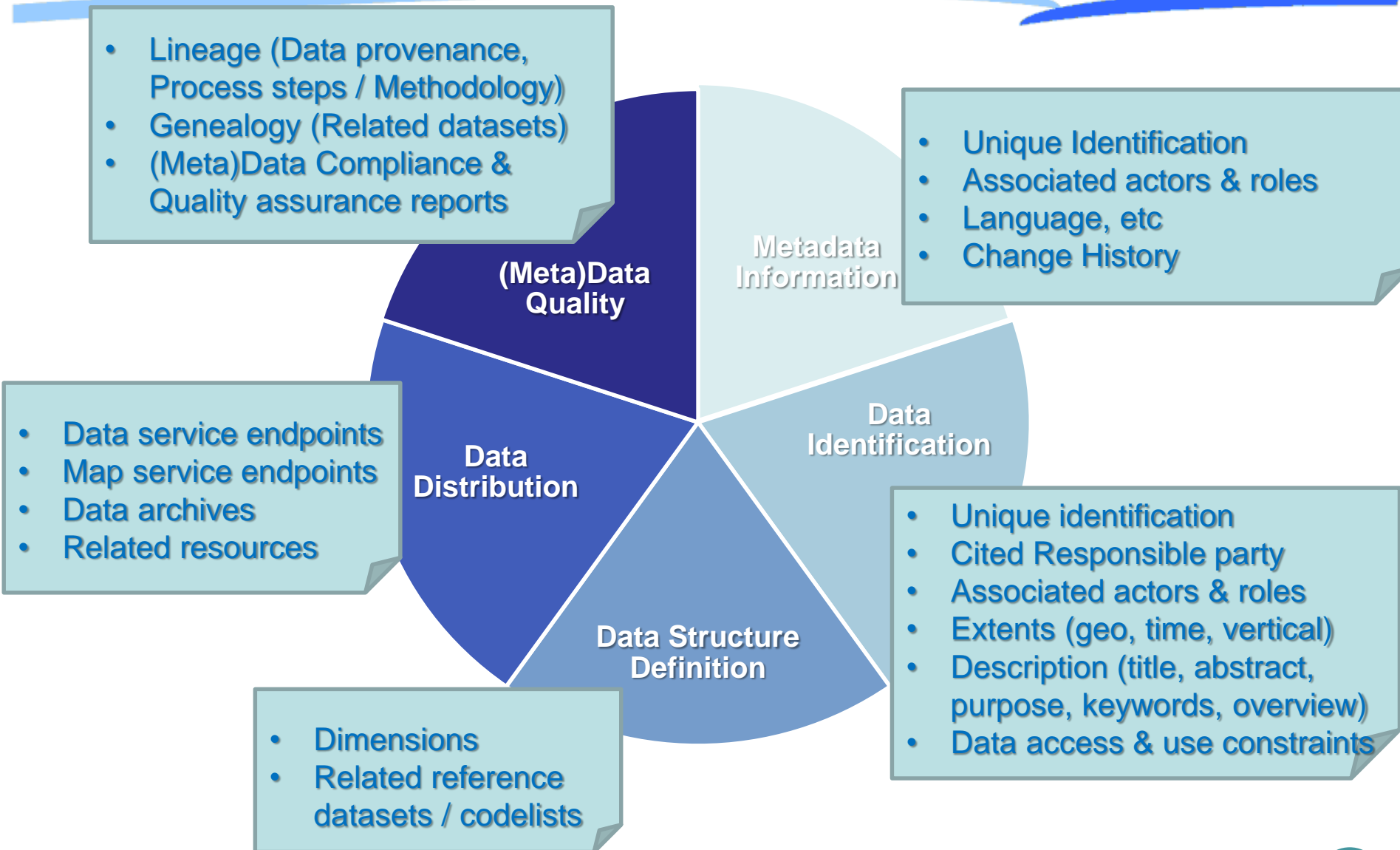
• Trends noticed through RDA

- Common pathway towards adoption of **FAIR** Data Principles (**F**indable, **A**ccessible, **I**nteroperable, **R**e-usable)
- Convergence of data domains in using:
 - Common **Data & Metadata standards**
 - Data standards:
 - Geospatial OGC formats, Unidata NetCDF
 - Metadata standards:
 - Generic: DublinCore
 - Geospatial: ISO/OGC 19115/19139, Catalogue Service for the Web (CSW),
 - Domain specific:
 - CF-Conventions, EML
 - Few mention of SDMX. Mentioned with a need of mapping from/to ISO/OGC Geospatial “world” (e.g. European Statistical framework vs. INSPIRE)
 - Common **Data access & sharing Methodologies**
 - Common Technologies (but technology is NOT the focus here)



**Data needs to have a structured description
(not only a data structure definition)**

METADATA



• FAIR Principles

FINDABLE

- F1. (meta)data are assigned a **globally unique and eternally persistent identifier**.
- F2. data are described with **rich metadata**.
- F3. (meta)data are **registered or indexed in a searchable resource**.
- F4. metadata **specify** the data identifier.

ACCESSIBLE

- A1 (meta)data are **retrievable by identifier** using a **standardized communications protocol**.
- A1.1 the **protocol** is open, free, and universally implementable.
- A1.2 the **protocol** allows for an authentication and authorization procedure, where necessary.
- A2 **metadata are accessible**, even when the data are no longer available.

INTER OPERABLE

- I1. (meta)data use a **formal, accessible, shared, and broadly applicable language** for knowledge representation.
- I2. (meta)data use **vocabularies that follow FAIR principles**.
- I3. (meta)data include **qualified references** to other (meta)data.

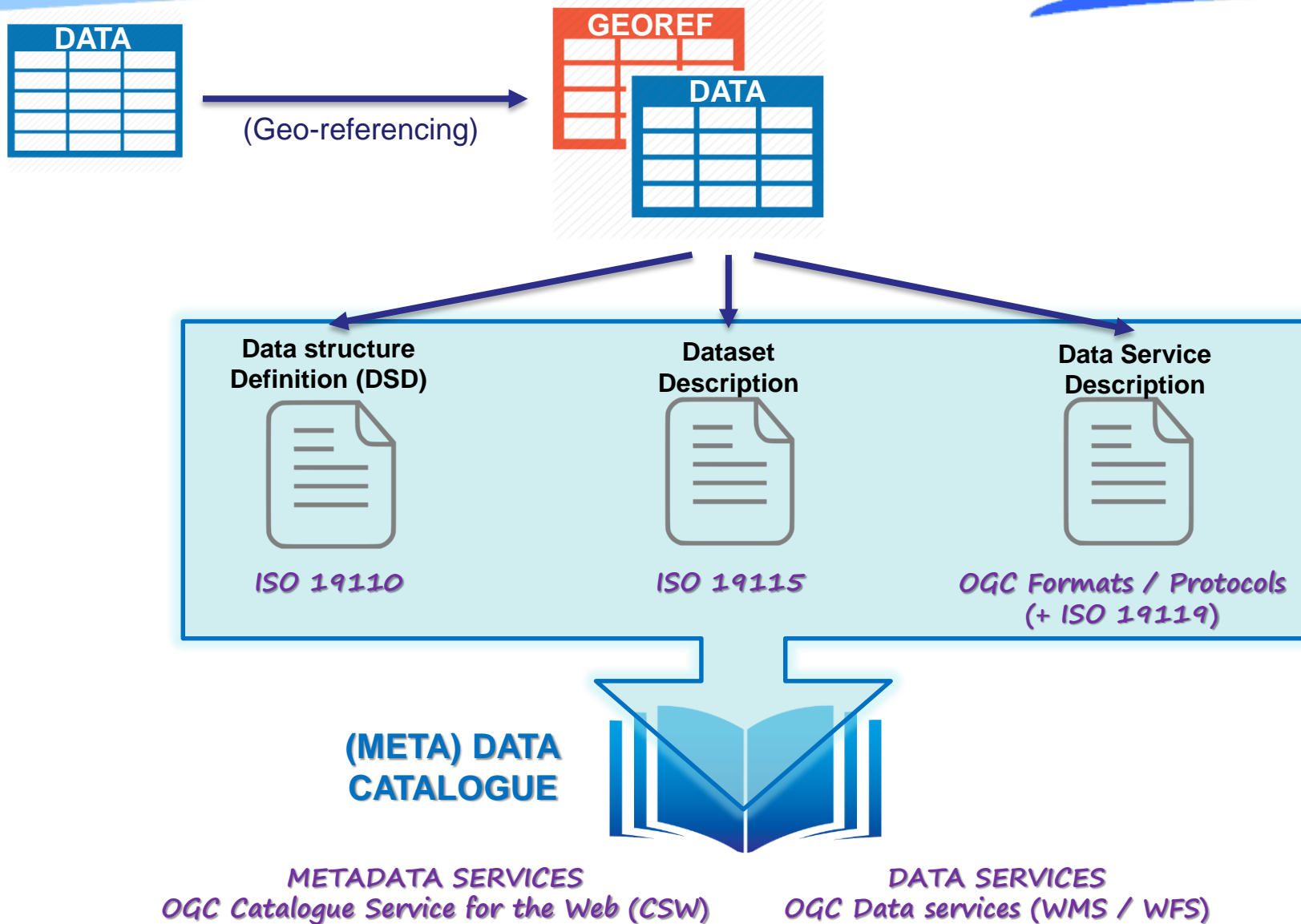
RE-USABLE

- R1. meta(data) have a **plurality of accurate and relevant attributes**.
- R1.1. (meta)data are released with a **clear and accessible data usage license**.
- R1.2. (meta)data are associated with their **provenance**.
- R1.3. (meta)data **meet domain-relevant community standards**.

<https://www.force11.org/group/fairgroup/fairprinciples>

- **Methodology for cross-domain, geo-referenced statistics dissemination & visualization – Principles**
 - **FAIR data principles**
 - **Standards**
 - Use of internationally recognized and widely adopted open standards for data, metadata, processes and services.
 - Primary choice oriented on geospatial-oriented standards: ISO/TC211 and OGC, rather than statistics-oriented standards (SDMX):
 - Wider user community and support
 - Tools availability and reliability
 - Better fitting with cross-domain / integrated vision (enabling confrontation with other data products such as stock assessment model outputs, scientific fishery surveys, vessel trajectories – VMS and AIS, etc.)
 - **Tools used to promote the approach**
 - Free and Open Source Software (FOSS)
 - Simple, Affordable and Flexible tools , ie that can be easily adopted and fine-tuned in autonomy by a wide data managers community (and not only a pure IT community)
 - Main programming language used:







**(META) DATA
CATALOGUE**



METADATA SERVICES
OGC Catalogue Service for the Web (CSW)

DATA SERVICES
OGC Data services (WMS / WFS)

FINDABLE

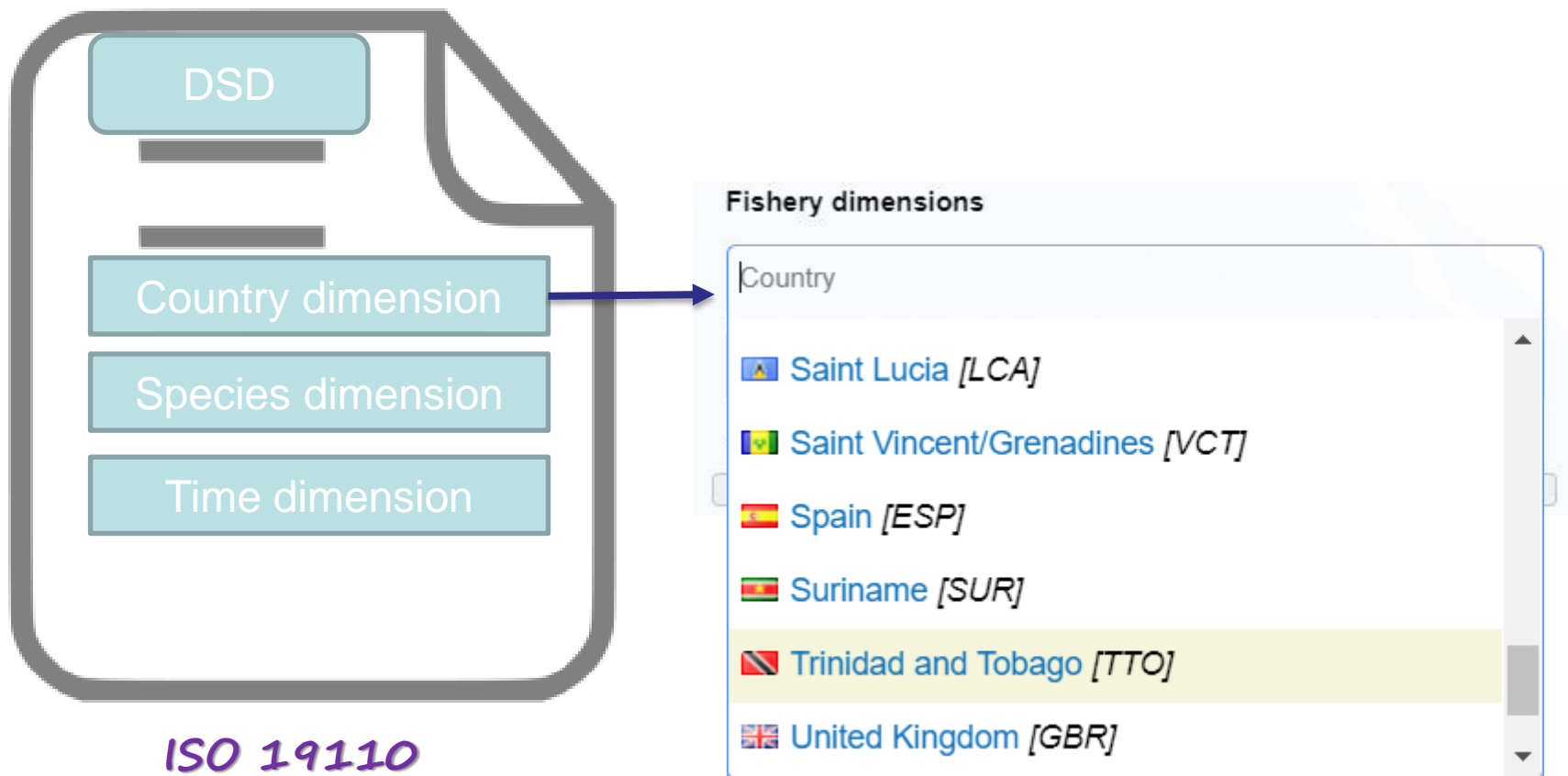
ACCESSIBLE

INTER
OPERABLE

RE-USABLE

EASILY
QUERYABLE ?

- a Data Structure Definition file is not only useful for data exchanging...
- But also for data dissemination & exploitation
(query, map, graphs, download)





Global Tuna Atlas

The screenshot displays the Global Tuna Atlas web application interface. At the top, there are navigation buttons for 'About', 'Browse', and 'Query'. Below this is a search bar with the query 'global_catch_1deg_1m_ps_bb_tunaatlasIRD_level2'. The interface is divided into several sections:

- Fishery dimensions:** Includes filters for 'Indian Ocean Tuna Commission [IOTC]' (Indian Ocean Tuna Commission) and 'Japan [JPN]' (Japan).
- Temporal extent:** Set to 1976 - 2016, with options to 'Add a Month', 'Add a Quarter', or 'Add a Year'.
- Aggregation method:** Set to 'Sum [sum]'.
- Map options:** Set to 'Ckmeans clustering' with a value of '4'.

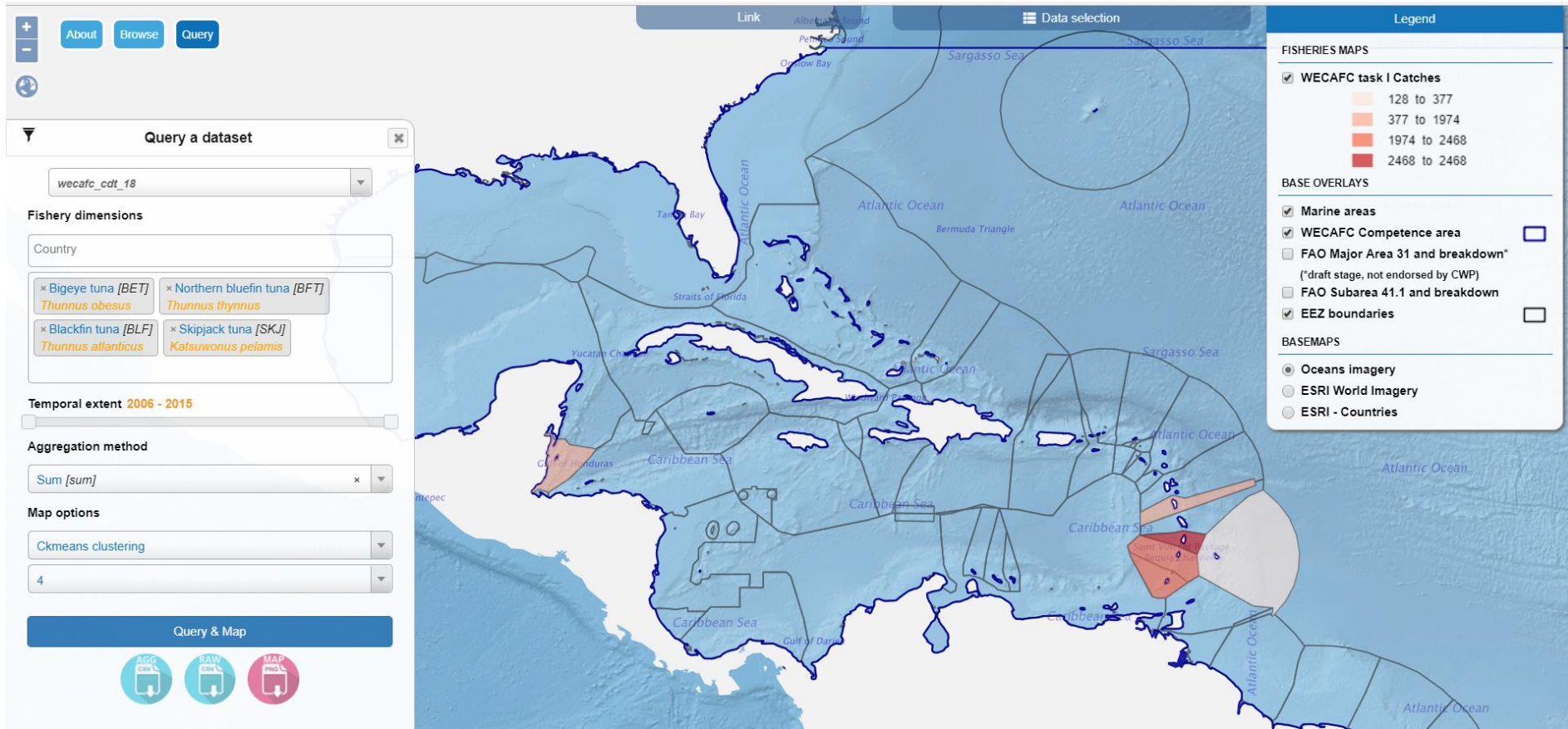
At the bottom of the filter section are three icons: 'AGG CWP', 'RAW CWP', and 'MAP CWP'. A 'Query & Map' button is also present.

The map on the right shows the Indian Ocean region with a data overlay. A legend on the right side of the map provides details:

- TUNA MAPS:**
 - 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)
 - Color scale: 1 to 252.79 (lightest), 252.79 to 600.89, 600.89 to 1090, 1090 to 2150.03 (darkest).
- BASE OVERLAYS:**
 - Marine areas
 - Grid 5x5 (CWP)
 - Grid 1x1 (CWP)
 - FAO major areas & breakdown
 - EEZ boundaries
- BASEMAPS:**
 - Oceans imagery
 - ESRI World Imagery
 - ESRI - Countries



WECAFC-FIRMS Regional Database





- **Establish mappings between (meta)data standards**
 - **Some objectives**
 - Fulfill legal (meta)data exchange obligations
 - Foster cross-disciplinary approach
 - **(Meta)data standards Mappings**
 - **Between ISO/OGC and Statistics/SDMX**
e.g. Provide reference metadata for the European Statistical System (2009/498/EC)
 - **From NetCDF-CF**
NetCDF / CF-Conventions, initially for Climate / Weather sciences
Now widely used in fields of marine sciences, oceanography & stocks assessment
 - **From/To Ecological Metadata Language**
Ecological Metadata Language (EML) used in ecology / biodiversity data communities



- **Strengthen the data extraction & visualization tools**
 - Contextualization to Regional areas/RFMOs or other criteria: already available
 - Need to fine-tune map visualization tools.
 - Additional map types: proportional maps, charts-embedded maps
 - Enable Cross-disciplines analyses
 - Capacity to interact further with statistical maps & add graphs
 - Enable visualization of confidential data, depending on use cases
 - Fine-tune data export / sharing facilities
 -



- **Work papers**

- Paper in preparation to describe the (meta)data publication & exploitation methodology and how to apply it using R software for statistical computing.
- Related use case communications:
 - Barde J., E. Blondel, E. Chassot, T. Imzilen, A-E Nieblas, P. Taconet, 2017. [Collaboration between fisheries and computer scientists for improved data description: The case of IOTC data sets.](#)
 - Nieblas A-E, S. Bonhommeau , T. Imzilen , D. Fu, F. Fiorellato, J. Barde, 2017. [Standardization of data formats, metadata, access protocols and statistical visualization of SS3 stock assessment outputs.](#)
 - Barde J., E. Blondel, P. Taconet, 2018. **Using R to manage a Spatial Data Infrastructure (SDI) in the South West Indian Ocean (SWIO).** Submitted to FOSS4G 2018 Conference <http://2018.foss4g.org>



Thank you for your attention