



Food and Agriculture
Organization of the
United Nations



Fisheries Data Interoperability Working Group (FDI WG)

Standards for fisheries data dissemination and exchange

Aymen Charef (FAO)

FAO HQ, Rome, Italy, 19-22 March 2018

The RDA; RESEARCH DATA ALLIANCE

SHARING WITHOUT BARRIERS

- RDA is a community-driven organization of over 6750 members from 136 countries, it includes **researchers, scientists and data science professionals** working in multiple disciplines, domains and thematic fields
www.rd-alliance.org
- Launched in 2013 by the European Commission, the United States National Science Foundation and National Institute of Standards and Technology, and the Australian Government's Department of Innovation
- Build the social and technical bridges that **enable open sharing** of data and tackles numerous data infrastructure challenges
- Enables data shared across barriers through focused Working Groups and Interest Groups



RDA Interest (IG) & Working Groups (WG) by Focus

Total 91 groups:
33 Working Groups & 58 Interest Groups

Domain Science - focused

- Agrisemantics WG
- BioSharing Registry WG
- Blockchain Applications in Health WG
- Capacity Development for Agriculture Data WG
- Fisheries Data Interoperability WG
- On-Farm Data Sharing (OFDS) WG
- Rice Data Interoperability WG
- Wheat Data Interoperability WG
- Agricultural Data IG (IGAD)
- Biodiversity Data Integration IG
- Chemistry Research Data IG
- Digital Practices in History and Ethnography IG
- Earth, Space, and Environmental Sciences IG

- Geospatial IG
- Global Water Information IG
- Health Data IG
- Linguistics Data Interest Group
- Mapping the Landscape IG
- Marine Data Harmonization IG
- Quality of Urban Life IG
- RDA/CODATA Materials Data, Infrastructure & Interoperability IG
- Research data needs of the Photon and Neutron Science community IG
- Small Unmanned Aircraft Systems' Data IG
- Structural Biology IG
- Weather, Climate and air quality IG

Partnership Groups

- RDA / TDWG Metadata Standards for attribution of physical and digital collections stewardship WG
- RDA/WDS Scholarly Link Exchange Working Group
- ELIXIR Bridging Force IG

- RDA/NISO Privacy Implications of Research Data Sets IG
- RDA/WDS Publishing Data IG



Mission

The FDI WG is a domain experts with a focus on **fisheries statistics** and **geospatial data** exchange <https://www.rd-alliance.org/groups/fisheries-data-interoperability-wg>

Synergy between FID WG and CWP; while CWP focuses on **conceptual aspects** of fishery and aquaculture statistical standards, the RDA focuses on the **operationalization** of standards for data and reference data hosting and interoperability

- Overall aims: First objective to foster better interoperability of statistical data; the second objective to better represent fisheries data in geospatial data formats.
- Provide guidance on legal interoperability issues (e.g copyright, Confidentiality & privacy, ownership) and to improve the discovery, access and analysis of fisheries.
- Review the capability of **standard** data model to host, incorporate and exchange the CWP reference data, global data structures and codelists

SDMX

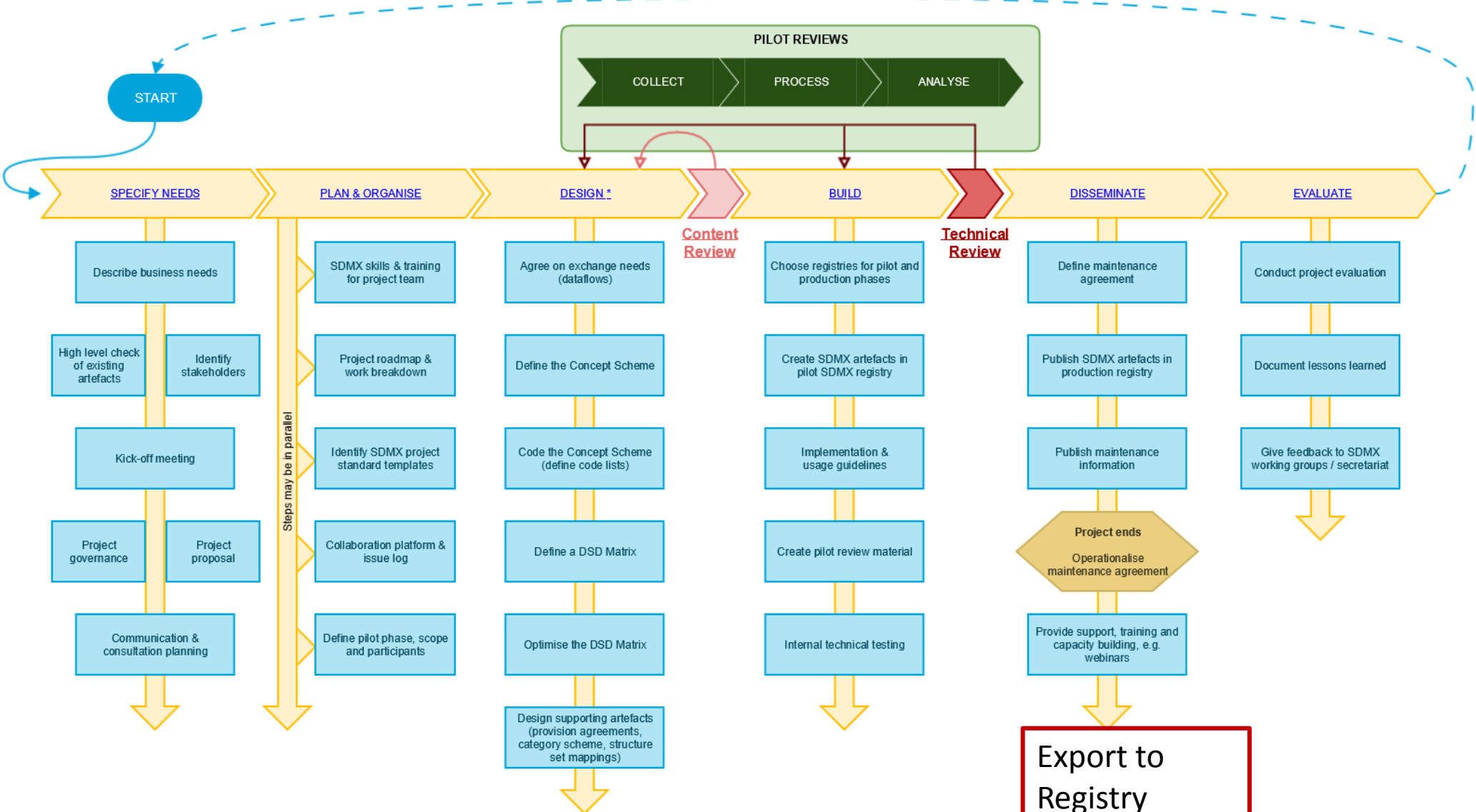
Statistical Data and Metadata eXchange



<https://sdmx.org/>

- An international cooperation initiative aimed at developing and employing more efficient processes for exchange and sharing of statistical data and metadata among international organizations and their member countries.
- The SDMX initiative is established in 2008 and sponsored by a set of institutions (UNSD, IMF, WB, OECD, EUROSTAT..)
- The community developed guidelines applicable to many statistical domains, tools. software to manage: the originator (agency), the versioning, the maintenance of SDMX artefacts,
- SDMX artefacts are reusable and published in SDMX registry (local or global): e.g Code lists, Concept schemes and DSDs

For a specific data domain (e.g DSD catch and Effort (Task II)), an SDMX project starts by creating a concept scheme that describes this domain and the data flows (e.g Country sends dataset to an organization).



SDMX implementation in Eurostat (CWP party)

[About](#) | [Legal notice](#) | [Cookies](#) | [Contact](#)

[About](#) | [Legal notice](#) | [Cookies](#) | [Contact](#)

[Help](#) | [Log in](#)



Euro SDMX Registry

European Commission > Eurostat > Euro SDMX Registry > Detailed View

Repository Administration

Include remote registries

Tree view Close en

Detailed View: Codelist ESTAT+CL_SPECIES+1.5

Artefact Information

ID: *

AgencyID: *

Version: *

Name (en): *

Description (en):

Access Level:

Is Final:

Valid From:

Valid To:

(*) Mandatory fields

Please select the export format...

Structures

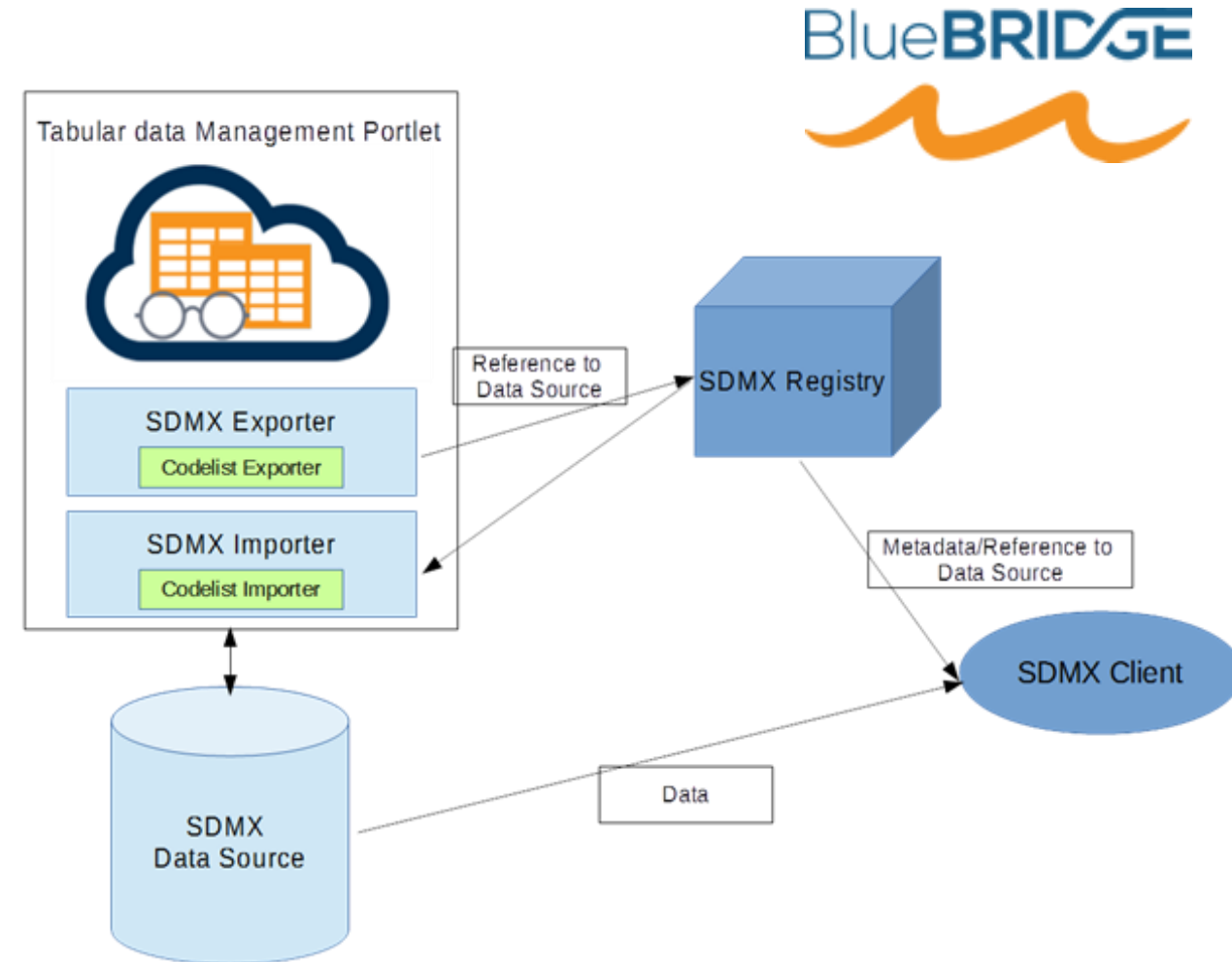
- SDMX v2.1
- SDMX v2.0
- EDI
- CSV

Download Close

	Description	Parent	Actions
GDS	Apistogrammoides pucallpae		
GDR	Gymnodraco acuticeps Ploughfish		
GDQ	Goeldiella eques		

SDMX implementation in FAO and BB

- The plan is to publish FAO reference data in SDMX format
- As intermediate solution, DSD and codelists of the statistical data collections are made available in ONE package with all attributes (CSV, information in flat format).



FAO Capture Data Structure Definition

DSD for Global Capture Production

Order	Concept_id	Role/Type	Codelist_id	Codelist_Code_id	Code system	Description
A	COUNTRY	Dimension	CL_FI_COUNTRY_GROUPS	UN_CODE	UN CODE	Country information
B	FISHING_AREA	Dimension	CL_FI_AREA_GROUPS	CODE	Gegraphical	FAO major catch areas
C	SPECIES	Dimension	CL_FI_SPECIES_GROUPS	3ALPHA_CODE	Species	ASFIS species list
D	YEAR	Time Dimension				Reference year
E	OBS_VALUE	Measure				Quantity of Capture
F	OBS_STATUS	Attribute	CL_FI_SYMBOL	CODE		FAO symbols
G	UNIT	Attribute	CL_FI_UNIT	CODE		Quantity unit

All the Codelists in this DSD are made available in ONE csv package in the FAO website

Conclusion

- SDMX offers the capability to have visible structure of the contents with explanations, metadata and versioning.
- The usual format is XML (SDMX-ML), which makes it a good option for exchange of fisheries statistical data sets and accompanied metadata.
- Building DSD would require dissociating the global CWP DSD to multiple DSD containing each 1 measure (event observation)
- CWP global DSD requires SDMX registry to create and disseminate SDMX artefacts (codelists, DSD). It could then be limited to dissemination use and not data submission

FLUX Fisheries Language for Universal eXchange

- It aims at defining a universal data exchange "language" compatible with regulations and international requirements. It is formulated according to the UN/CEFACT standards (Centre for Trade Facilitation and Electronic Business)
- A FLUX **team of specialists** was established in 2017 to promote, facilitate and support the implementation of the FLUX standard on a global scale. A FLUX User Community was also set up to exchange best practices and lessons learned.
- FLUX provides a harmonized message standard that allows RFMOs, control and enforcement authorities to automatically access the electronic data from fishing vessels, such as vessel and trip identification, fishing operations (daily catch or haul-by-haul) or fishing data (catch area, species and quantity, date and time,..)

FLUX Standard - v1 package

- *P1000-1 General principles domain (BRS v2.1.2).*
- *P1000-2 Vessel domain (BRS v3.1.0)*
- *P1000-3 Fishing Activity (FA) domain (BRS v1.0.1)*
- *P1000-5 Sales domain (BRS v1.0.1)*
- *P1000-7 Vessel Position domain (BRS v2.0.3)*
- *P1000-9 Fishing License; Authorizations & Permit (FLAP) domain (BRS v1.0.1)*
- *P1000-10 Master Data Management (MDM) domain (BRS v2.0.2)*
- *P1000-12 Aggregated Catch Data Report (ACDR) domain (BRS v2.0.2)*



Each domain is identified and numbered as follows:

P1000 – 1; **General Principles** (this document should be read prior to all others!): describes generalities common to all other specific sub-domains.

P1000 – 2; **Vessel**: The vessel domain contains all information related to the identification and the description of the vessel itself. Coordination with UIV - Global Record Project

P1000 – 3; **Fishing Activity**: Contains all information related to a fishing trip. This domain includes all data concerning landings of fish. A transshipment is considered as a special “landing” involving a second vessel and is also included in this domain.

P1000 – 5; **Sales**: The sales domain includes all transmission of information concerning first sales of fish since a landing operation, as well as all information concerning transportation of fish between landing and first sale.

P1000 – 7; **Vessel Position**: Contains all information exchanges related to the position of the vessel.

P1000 – 9; **Fishing Licence; Authorisation & Permit**: Includes national as well as international fishing licences, authorisations and permits.

P1000 – 10; **Master Data Management**: Domain for exchanging information about list of code.

P1000 – 12; **Aggregated Catch Data Report**: This domain describes the catch data, which was originally gathered from vessels by Flag States, which is aggregated and exchanged with International parties.

Implementation

- FLUX is not a system but a standard language, FLUX is only the messenger.
- **Business layer:** consists of **Business Requirement Specifications** (BRS) document, which needs to be defined/customized based on requirement for data exchange (bilateral) agreement.

Example : implementation guide for Fishing Activities (domain) message exchange in EU contains more than 100 pages.

- **Transportation layer** provides description for:
(very relevant to frequent exchanges of messages)
 - The **FLUX Envelope**, one single yet universal message format that can encapsulate any business-specific message
 - The **FLUX Protocol**, a mechanism describing how to deliver the FLUX Envelopes to their destination and without human intervention, leveraging existing advanced technologies (SOAP Web Services)

FLUX Build-in message flexibility

- **Optionality of elements:** Exactly define optionality
- **No-predefined codes:** Define codelists and characteristics to use

First two are standard UNCEFACT practices. FLUX does not contain pre-defined codes for fishing operations (because everybody has different operations...)

- **Characteristics concepts** Flexible and extensible mechanism to include items that the standard does not contain, e.g. gear properties and research data.
- **Notification/declaration** Every message (where relevant) can be a message of intention (notification) or a message of executed activity (declaration)
- **Push, pull or both** Message can be send on the senders initiative (push) or on request of the receiver (pull).
- **Corrections and deletions** FLUX includes corrections and deletions of messages

Last, Define technical details (e.g transmission, network). **They are not part of the standard.** Network topology (“spoke and hub”, network,), transmission system (e-mail, FTP,)

Conclusions

- FLUX team will interact with CWP to ensure the use of CWP Standards and concepts identified under the global DSDs to exchange fisheries statistics. Exchange of geo-information is a challenge and not yet addressed properly
- FLUX being based on XML can be widely used and relatively straightforward language whose power can be harnessed to: define message XSD (XML schema definition), delete, modify..
- Implementation to be specified at CWP party basis but NOT to be underestimated. Implementation document includes exchange agreement and requires to sit down with business-, IT- and legal experts of all the partners involved in the exchange of messages and take out all the flexibility and ambiguity.
- **FLUX Transport Layer** is an alternative way to send and receive FLUX messages. It does not have the disadvantages of e-mail, FTP etc. It requires specific IT infrastructure but it makes it very easy to send and receive FLUX messages in very frequent basis (e-logbooks).

Thus it could not really be useful for CWP Parties that exchange statistics (yearly basis)

Thank you