



联合国
粮食及
农业组织

Food and Agriculture
Organization of the
United Nations

Organisation des Nations
Unies pour l'alimentation
et l'agriculture

Продовольственная и
сельскохозяйственная организация
Объединенных Наций

Organización de las
Naciones Unidas para la
Alimentación y la Agricultura

منظمة
الأمم المتحدة
للزراعة
والإغذية

COORDINATING WORKING PARTY ON FISHERY STATISTICS

Intersessional Meetings of Aquaculture and Fisheries Subject Groups – Joint Session

Seventh Meeting of the Aquaculture Subject Group (AS) and Twenty-eighth meeting of the Fisheries Subject (FS)

2 – 5 November 2021

Guidelines for the implementation of the CWP standard for reference data harmonization

**Author: CWP ad-hoc Task Group
on reference harmonization (RH2)**

Executive Summary

The CWP 26th session in 2019 endorsed three data structures of the CWP standard for Reference Harmonization and established the ad-hoc Task Group on Reference Harmonization (TG-RH2) for the work on extension on reference harmonization through a second phase. The overall goal of this standard is to minimize time and costs of mapping data elements and facilitate data interoperability through harmonized terminology and international concepts and classifications. It intends to lay basis for data sharing agreements in fisheries statistics domain to improve multilateral exchanges and promotes the consistency and efficiency of statistical systems at all official levels.

This document presents guidance for data producers and users on the implementation of the CWP standard for reference harmonization through a standardized framework. The guidelines put forward core rules and best practices to fulfil the most typical scenarios in various contexts. The framework puts into practice the use of statistical concepts and classifications of CWP handbook on fisheries statistics and other international standards and ways to accommodate specific needs. It also specifies the syntax and the semantics for the exchange and dissemination of statistical data.

This document also presents standards and mechanisms, which are increasingly coming into use, for the collection, warehousing and dissemination of statistical data and metadata. Use cases are described such as the CWP catalogue. An earlier version of this document was presented at the CWP 26th session and feedback from CWP members was included.

The current document is intended to be a living document that will evolve over time as more is taken from requirements of the CWP members, and lessons learned from pilot implementation at national, regional and global levels.

CWP members are invited to provide feedback and input to the proposed guidelines. The final version of the guidelines will be presented for endorsement at the CWP 27th session.

1. Background

The objective of the CWP ad-hoc Task Group on Reference Harmonization is to lay basis for establishing data-sharing agreement as practical work arrangements between agencies involved in a data workflow. This is expected to reduce data reporting burden for data producers and to improve data quality by mainstreaming the cross-checking and reconciliation of information from national sources. The CWP 26th session meeting held in FAO headquarters (Rome, 2019), endorsed the CWP standard for reference harmonization which consists of a set of data structures and harmonized statistical concepts relevant to capture fisheries: Global Capture Production, Catch, and Catch and Effort ([Appendix 5 – Annex1](#) of the CWP 26 session report).

The data structures cater concepts and classifications that are agreed by the United Nations (UN) Statistical Commission which is the highest decision-making body for international statistical activities especially the setting of statistical standards, the development of concepts and methods and their implementation at the national and international level. Other concepts and classifications are supplemented by CWP handbook and international standards wherever possible.

The CWP 26th session also endorsed the work extension of the CWP standard for Reference Harmonization through a second phase and established the ad-hoc Task Group on Reference Harmonization (TG-RH2). Related to it, the session also established two ad-hoc task groups to focus on effort and catch concepts.

2. Scope and objectives

The present guidelines put forward the rules and specifications to be followed when implementing the CWP standard for reference harmonization in various contexts; national, regional and global. The focus of these guidelines is on the three data structures endorsed by the CWP 26th session (Annex 1 and 2) to be used for data collection, dissemination and exchange by statistics producers and custodians (e.g. CWP members, regional bodies).

This implementation guidelines could be used in conjunction with data sharing agreement or any other formal agreement that exist between the partners. It sets the core rules applicable at national, regional and global organizations and ensure the production of interoperable data in the domain of fisheries statistics. This will minimize time and costs of mapping data elements to standard terminology and will improve multilateral exchange.

The implementation guidelines present a framework based on the Generic Statistical Business Process Model (GSBPM)¹ which is a reference framework to define and describe processes with the main objective of designing and implementing a standardized system to support business statistics. The guidelines are also developed based on a compilation of lessons learned from pilot implementation of the CWP standard for reference harmonization into a specific Data Structure Definition (DSD). The pilot use cases are:

¹ <https://statswiki.unece.org/display/GSBPM/GSBPM+v5.1>

- Regional database of the Fisheries Committee for the West Central Gulf of Guinea ([FCWC](#)), which has been recently developed to collate and maintain catch and effort statistics at fishing units' level of six countries in the region.
- Global Atlas of Tuna and Tuna-like species ([FIRMS Tuna Atlas](#)), that compiles catch and effort statistics from five CWP members responsible of tuna fisheries, namely: CCSBT, IATTC, ICCAT, IOTC and WCPFC (and SPC).
- Data Collection Reference Framework (DCRF) of the Western Central Atlantic Fishery Commission (WECAFC), including main themes of data collection (tasks) and definition of variables.
- CWP catalogue that disseminate the global CWP data structures, CWP classifications and code lists used by CWP parties.
- Regional Commission for Fisheries ([RECOFI](#)) database, which is being designed to collate catch and effort statistics of eight members of the Commission.

3. Audience

These guidelines serve as a reference for the general statistical audience, staff responsible for producing statistics at national, regional and CWP members' levels who wish to share and / or process data created according to the specification of the CWP standard for reference harmonization. The language used in the document is not IT focused, and targets audience of mainly statisticians and also officers responsible of implementation of management applications (e.g. database managers) and not necessarily IT platform and system developers (e.g. users of PostgreSQL, Oracle).

4. Levels of implementation

The implementation guidelines set out a framework to enforce business logic at the schematic and the database levels. At the first level, a sequence of steps, borrowed from the GSBPM, is described to build a Data Structure Definition that fulfil the needs including particularities (see Annex 3). Whereas, at the semantic level (see Annex 4), technical recommendations are provided to implement a flat (denormalized or row-focused) data structure that is measurement-dependent. This would make the structure human-readable and useful essentially to users that have to manual data imputation and editing.

4.1. Schematic (structure) level

The steps to define the structure of the data to be used for data collection, dissemination or exchange are:

- **Specify needs:** identify scope or domain of the data collection (e.g. Nominal, catch, catch and effort.). This step includes statistical needs (current and future) of national and international statistical organizations.

- **Establish output objectives:** identify the statistical output objectives that are required to meet the user needs. This step includes agreeing the suitability of the proposed outputs and their quality measures with users. If there is confidentiality related Legal frameworks (e.g. relating to confidentiality) need to be considered at this step.
- **Identify concepts:** select the statistical concepts to be used (see Annex 3). Additional concepts could be added to the CWP data structure to accommodate the requirement for the data exchange (e.g. adding the concepts “fishing ground”, “sampling area”, “fishing unit”).
- **Design variables description:** define the variables-concepts and associated statistical standard (i.e. classifications) to be collected. It is expected that definitions from existing international references will be followed wherever possible (i.e. from CWP handbook). It is highly recommended that code lists are consistent, to the largest extent possible, with internationally agreed standards, whenever they exist. The following order of priority is suggested when considering the use of existing code lists:
 - international standard classifications or code lists (e.g. standards of the CWP);
 - international classifications or code lists supplemented by other international and/or regional institutions (e.g. International Organization for Standardization ISO);
 - standardised classifications or code lists used by individual international institutions e.g. RFMOs.
- It is expected that a particular classification may not meet all the needs (for example, levels or categories are too broad or too narrow). This can be the case of classifications that may include extractions from statistical, administrative, geospatial and other non-statistical registers and databases, and information about practices in other RFBs and statistical organizations. A number of alternatives may be created to meet these needs. The original categories can be split or regrouped to provide context-specific additions or alternatives to the standard aggregation structure. These are called classifications variants². If standard classifications are used for defining the code lists, totals, aggregates and other additional codes should be added following the recommendations of the issuing organisation for adding new elements (e.g. the addition of country aggregates to the ISO 3166 alpha 2 country list; or species aggregates to the ASFIS List).

Annex 5 provides details on handling the variants and levels of granularity.

- **Design the structure:** define the sequence of dimensions in the data structure and how they can better fit the representation of the data. Structure may also be designed in partnership with other interested bodies, particularly if they are joint outputs, or they will be disseminated by another organization (e.g. in the case of FIRMS Tuna Atlas).

4.2 Semantic (Business) level

The data structure and its contents should be defined to be understood by primarily humans and then machines. Criteria for the implementation at national, regional and CWP party levels are presented where main statistical concepts are described with related codes and constraints (Annex 4). At this level, the metadata should be specified so for instance code lists, datasets can be discoverable and interoperable with other information systems.

² Generic Statistical Information Model (GSIM): Statistical Classifications Model
<https://statswiki.unece.org/display/gsim/Statistical+Classification+Model>

The basic criteria should be respected when developing data structure definition and its code lists are:

- When implementing the CWP structures or designing code lists accordingly, it may not always be possible to accommodate all needs from the start due to the legacy systems (e.g. linkages to report or dissemination systems). In such cases, some transition period will often be necessary before reaching convergence.
- Identifiers of code lists take values from uppercase A to Z, 0 to 9 and "_" only. No other characters should be used. Underscore ("_") is generally used for the combination of codes (whether consecutive or not).
- It is recommended not to use lower case characters in order to avoid possible confusion with upper case characters. For instance, mixing lowercase and uppercase might be understood as it has a distinct meaning or in the case where the dash ("-") could be considered as an operator for subtraction.
- Multiple hierarchies (hierarchical code lists) can be defined on top of a flat code list. The code list should be flexible in terms of allowing different possible hierarchies and be extendible by additional codes that may disaggregate or aggregate codes that are already in the list as well as by codes that extend the coverage of the code list. Each possible hierarchy may use all codes from the flat code list or just a subset. The flat code list provides the pool of codes for the hierarchies and is common user-friendly.
- Hierarchical code list should be presented in the flat (denormalized) format to allow human readability wherever possible (e.g. each catch type is provided in column). This will not hamper machine readability and normalization of the database. Details are provided in the next section for each concept.
- Concept names are in general defined in English; other language versions may be added. Code names should be between 1 and 254 characters. The characters used should belong to the UTF-8-character set.
- Certain exceptions can be considered in the case of codes which are normally not translated (e.g. regions or territories for which the national designation could be used, agencies for which a national and English label could be used, etc.).
- Descriptions could be used if more details on the contents or on the code descriptions are needed. Multilingual representations are possible.

5. Concepts, classifications and constraints

This chapter describes definition, code list and constraints recommended for use for each concept and associated statistical standard.

5.1 Entity

It identifies the administrative or political entity to whom the data refers where the fish was caught, produced or landed. The FAO statistical standard for the FAO Standard Country or Area Codes for Statistical Use is used as reference. It is based on the *"Standard Country or Area Codes for Statistical Use"* which is owned by the UN Statistics Division (UNSD) and commonly referred to as the M49 standard³.

³ <https://unstats.un.org/unsd/methodology/m49/>

The three-digit M49 code list is available at the CWP handbook, in the list of countries or areas⁴, where it is mapped to the two- and three-letter codes (ISO-Alpha2 and ISO-Alpha3) code list presented by ISO 3166.

The FAO standard presented in Annex 6 is composed of M49 active country or area codes and a few others added by FAO in agreement with UNSD.

In the catch data structure, the country, area, or entity is designated by one code from the three-digit numerical of M49 code list.

Use case: In the context of tuna RFMOs, the entity is represented as fishing fleet. Its definition was established and adopted under the framework of FIRMS Tuna Atlas in the twelfth session of the FIRMS Steering Committee (FSC12) steering committee (October 2021) and is submitted to the CWP for consideration and adoption.

5.2 Fishing area

It refers to the geographic area which could be presented according to one of the following standard classifications.

Use case: For specific purposes, two or more classifications can be presented in the same flat data structure.

5.2.1 FAO Major Fishing Areas for statistical purposes and their breakdown⁵

They are arbitrary areas and historically determined coinciding to the greatest extent possible to the areas of competence of other fishery commissions when existing. This classification system facilitates comparison of data and improves the possibilities of cooperation in statistical matters.

The codes are selected from this code list: <http://www.fao.org/3/bt979e/bt979e.pdf>.

Use case: In a flat data structure, each level of this hierarchical classification can be presented in a separate column: “area” “subarea” “division”. On the other hand, the three levels can be aggregated into one column Division level e.g. fishing_area 34.3.3 and 34.4.4.

5.2.2 EEZ from Marine Regions

It is managed by VLIZ the Flanders Marine Institute⁶ and presents the relational classification list of geographic names, coupled with information and maps of the geographic location of these features. It is considered the standard reference for Maritime Boundaries representing the Exclusive Economic Zone (EEZ) of the world.

The codes correspond to the **Marine Regions Geographic Identifier (MRGI)**:
<http://www.marineregions.org/mrgid.php>.

Use case: This classification can be represented in a data structure in one column with these two or more codes i.e. “withinEEZ”, “outsideEEZ”.

5.2.3 RFB competence area

⁴ <http://www.fao.org/3/bt978e/bt978e.pdf>

⁵ <http://www.fao.org/3/bt979e/bt979e.pdf>

⁶ <http://www.marineregions.org/eez.php>

The codes correspond to the acronyms of the Regional Fisheries Bodies (RFBs).
<http://www.fao.org/fishery/rfb/en> (See also Annex 7).

The boundaries of their area of competence are presented here:

<http://www.fao.org/figis/geoserver/factsheets/rfbs.html>

5.2.4 CWP Areal Grid system

This grid map system is essentially used for the context of Tuna fisheries. Various maps, resolutions and square identifiers are provided here: <http://www.fao.org/cwp-on-fishery-statistics/handbook/general-concepts/major-fishing-areas-general/en/>

5.3 Time

For national and regional fishery statistics, it is recommended that date and time is recorded in ISO 8601 format using the Gregorian calendar and 24-hour time keeping system. The general format is to record date as YYYY-MM-DD (or YYYYMMDD), time as hh:mm:ss (or hhmss), and date time as YYYY-MM-DDThh:mm:ss (or YYYYMMDDThhmss); seconds may be provided optionally.

For example: UTC date time 2018-10-31T15:07Z (where Z denotes a zero UTC offset).

For the “Catch data structure”, the time unit corresponds to the Year, an integer, recorded as 4 digits. For example: 2019, 2018.

In particular cases, the time unit is made in two rather than one single column, in order to model the START / END of the time period that the measurement refers to.

The ISO 8601-2:2019 defines a set of standardised extensions to the ISO 8601 date and time formats. Extended Date/Time Format (EDTF)⁷ provides sub-year groupings (Quarter1 – Quarter 4).

In other use cases, Time is submitted in number of months, with “n” that can be anything from 2 to 11: as the number of months cannot be split into the quarter of the year.

5.4 Fishing gear

The International Standard Statistical Classification of Fishing Gear (ISSCFG) is the reference standard. The codes should be chosen from corresponding to the standard abbreviation of gear category⁸.

The ISSCFG presents codes of major categories and sub-categories of gear. For this, CWP members or other producers of statistics often use their list of additional codes according to their context. It is recommended to have a mapping for their groupings with the ISSCFG categories.

5.5 Fishing mode

This statistical concept is reported in conjunction with the fishing gear to designate the fishing practice and to enhance the fishing effort definition.

Use case: The codes commonly used by tuna RFMOs and the WECAFC Data collection framework are presented in Annex 8.

⁷ "Extended Date/Time Format (EDTF) Specification". The library of congress.

<https://www.loc.gov/standards/datetime/>

⁸ ISSCFG <http://www.fao.org/3/a-bt988e.pdf>

5.6 Fishing Effort

Definitions and code list to be established for fishing effort types based on the outcomes of the CWP ad-hoc Task Group on fishing effort concepts. The TG outcomes will be submitted to the CWP 27th session for endorsement.

5.7 Fleet segment

The combination of a group of fishing vessels of the same size and using the same gear for more than 50 percent of the time at sea during a year. To accommodate both regional and international fisheries organizations classifications, an organization can adopt a «Fleet segment » classification and coding system.

Use case: The WECAFC Fleet segment classification is defined as the combination of a Vessel type classification derived from the ISSCFV and length classes derived from International and ICCAT classifications. For mapping national vessels to the vessel type classification, the notions of either using gear exclusively, or using gear predominantly, or using gear with no-predominance, will be applied.

5.8 Aquatic species

The classification recommended for primary use is the ASFIS List of Species for Fishery Statistics Purposes which is maintained by FAO⁹. Accordingly, the mapping of other coding systems, used for various purposes, to the ASFIS system is essential, to the extent possible.

The code is selected from the code list “3A_CODE”, a unique identifier made of three letters and widely used for the exchange of fisheries statistics. The 3-alpha code corresponds to a scientific name, author(s), species, genus, family or higher taxonomic levels. It is also matched to a taxonomic code and ISSCAAP group¹⁰.

5.9 Catch Type

It corresponds to the CWP catch type concepts as described in the catch concept diagram of the CWP handbook¹¹.

Code list to be established for catch types based on the outcomes of the CWP ad-hoc Task Group on catch concepts.

Use case: In a flat data structure, it is recommended to present each type in a column (dimension).

5.10 Unit

From the FAO statistical standard series for Unit.

In the data structure Unit of measure are usually tonnes or number of individuals.

5.11 Observation status

It is recommended to use the SDMX Code list for Observation Status (CL_OBS_STATUS)¹².

⁹ <http://www.fao.org/fishery/collection/asfis/en>

¹⁰ http://www.fao.org/fishery/static/ASFIS/ASFIS_Structure.pdf

¹¹ Catch concept diagram <http://www.fao.org/3/bt981t/bt981t.pdf>

¹² Possible Ways of Implementing the Observation Status Concept, version 2.0

https://sdmx.org/?sdmx_news=possible-ways-of-implementing-the-observation-status-concept-version-2-0

Use case: FAO established its specific statistical standard for Observation status flags based on the SDMX code list (Annex 9)¹³.

5.12 Coverage

It refers to the proportion that the amount of fish (in number or weight) or fishing effort that is monitored (sampled) makes out of the total (number or weight) of fish or fishing effort estimated in the stratum concerned¹⁴. The present coverage rate presented at the Annex 10 is taken from the IOTC guidelines for the reporting of fisheries statistics. It needs to be standardized with the coding system of other CWP members (essentially tuna RFMOs).

The user should one of the codes in the Annex 10 and not input as decimal digits.

Use case: coverage rate could be “Less than 5percent of the boats covered” or “Statistics raised; coverage unknown”.

5.13 Confidentiality

Information about the sensitivity and confidentiality status of the data. The classification system and code list used corresponds to the SDMX Cross-Domain code list for Confidentiality Status (CONF_STATUS)¹⁵. The codes and their descriptions (and annotations) are available in Annex 11.

For example: F: free for publication, N: Not for publication, restricted for internal use only.

6. Data exchange formats and mechanisms

Data format provides the content and the structure of the document sent over the data network. There are several formats and standards of dissemination and exchange which can be used to implement the global CWP standard for reference harmonization and related data structures. Data exchange options were a topic of discussion during the intersessional meeting. The TG recognized the importance of defining and recommending formats and standards for data exchange. Options put forward should cover the varying capabilities and requirements of CWP Parties, as mechanisms need to be aligned to enable data sharing agreement.

Alternatives should be evaluated based on ease of implementation and operation and the following criteria:

- It should be widespread throughout the CWP parties to minimize compatibility issues.
- It needs to be readable for human and machine, complexity should therefore be kept at an acceptable level.
- Harmonization of structures needs to be possible. Structure must be standardized, and the file format must support an open standard.

¹³ SDMX “Observation Status List” (2016) https://sdmx.org/wp-content/uploads/CL_OBS_STATUS_v2_1.docx

¹⁴ http://www.iotc.org/sites/default/files/Guidelines_Data_Reporting_IOTC1.pdf

¹⁵ https://sdmx.org/?page_id=3215

As for best practices to be pointed out, the writing convention or format could be recommended when exchanging data with specific coding system (e.g. FAO areas breakdown). In this case, the easy digitalization of the codes should be considered to facilitate data interoperability and exchange.

6.1 Comma Separated Values CSV

CSV file format is widely used among CWP parties for dissemination of datasets and metadata. The readability of CSV files is acceptable and facilitates the interaction of human user. It remains the best format to exchange the global data structures, the reference data and metadata.

The main advantage of using CSV files lies in the fact that such format can be accessed through common spreadsheet software, making them easily managed manually and a useful option to accommodate data providers without information systems that can generate the data files for transmission automatically.

As an example, FAO made available data structures definitions for data domains namely global capture production¹⁶, and global aquaculture production, in a packaged format comprising Data Structure Definitions and code lists in CSV files and related metadata in text file. They will be aligned with the endorsed CWP standard for reference harmonization following the present guidelines.

6.2 Statistical Data and Metadata eXchange SDMX/SDMX-ML

[SDMX](#) is an international initiative that aims at standardizing and modernizing the mechanisms and processes for the exchange of statistical data and metadata among international organizations and their member countries. The organizations involved in the SDMX initiative developed [guidelines](#) applicable to several statistical domains. Furthermore, the community made available [software tools](#) and a registry to host reusable SDMX artefacts [7].

SDMX is not just a technical standard but offers guidelines such as a [Checklist for Design Projects](#) and [Modelling Guidelines](#) which are relevant for establishing an SDMX project for a data domain. For a specific data domain (e.g. capture data for dissemination purpose), 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). The design and creation of SDMX artefacts and the management of such a project are detailed in this standard project workflow¹⁷. The structure of this checklist is based, to the largest extent possible, on the GSBPM.

SDMX is being adopted as the data collection format for fisheries in Eurostat, in-line with policy for all statistical domains covered by the European Statistical System. FAO is making progress in the implementation of SDMX principles and acquisition of necessary tools.

Technically, SDMX standard offers an information model which describes statistical data sets and the structural metadata needed to exchange them in a standard fashion. The content of SDMX files have visible structure with explanations what is stored where in the file. The usual format in SDMX information model is XML (SDMX-ML) which makes it a good option for exchange of fisheries statistical data sets and accompanied metadata.

In the CWP context, there has been an attempt to apply SDMX principles to fisheries statistics and in particular to create the global catch DSD, the context of a joint-project SEIF that stands for SDMX

¹⁶ https://www.fao.org/fishery/static/Data/GlobalProduction_2021.1.2.zip

¹⁷ Checklist for SDMX Design Projects https://sdmx.org/?sdmx_news=checklist-for-sdmx-design-projects

for Eurostat (and DG MARE) and FAO¹⁸. The initiative aimed at the alignment and the exchange of code lists between both organizations and then other CWP Parties (ICES and NAFO).

The present guidelines are one-step further to the creation of SDMX artefacts for fisheries statistics domain i.e. DSD, concepts, concepts ID and definition (see Annex 3). It remains essential to evaluate the ability of SDMX data model to incorporate the proposed multilingual reference data and the global data structures that can be expanded with other code lists and enriched with hierarchical code lists.

6.3 Fisheries Language for Universal eXchange FLUX

FLUX standard, developed and maintained by the Centre for Trade Facilitation and e-Business (UN/CEFACT), provides a harmonized message standard allowing Fishery Management Organizations to automatically access the electronic data needed for stock management, such as vessel and trip identification, fishing operation (daily catch or haul-by-haul), fishing data (fishing area, species, date and time, and gear used), landing and sales information.

FLUX contains two distinct but related parts:

- The FLUX business layer
- The FLUX transportation layer

The core of the FLUX business layer is the detailed and standardized description of each, and any data element needed. For the FLUX business layer, standardization of the data elements and formats is based upon the UN/CEFACT approach of Business Requirements Specification (BRS).

[UN/CEFCAT BRS](#) have been defined and endorsed for the following FLUX domains:

- **Vessel Domain:** aims to standardize the exchange of fishing fleet data, and more specifically the information directly related to fishing vessels and vessels supporting fishing operations.
- **Fishing Activities Domain:** is related to data exchanges in the context of fishing activities performed by vessels during a fishing voyage. Fishing activities include all activities of vessels, related to a fishing trip. The domain contains reports related to the fishing trip: departure, arrival, entry and exit from zones, fishing operation, etc.
- **Vessel positions Domain:** provides a standard for the communication of vessel position information (e.g. VMS or AIS) between monitoring centers.
- **Fishing licenses, authorizations and permits:** to standardize the exchange of data between stakeholders in the context of request for fishing license, authorization or permit.
- **Aggregated Catch Data (ACDR):** provides standard to exchange aggregated catch data between stakeholders.
- **Master Data Management (MDM):** encompasses exchanges from a Master Data Register to any requester of Fisheries information registered in it.

The focus of the TG on reference harmonization should be directed to both [Fishing Activities and MDM](#) domains that are particularly relevant to the present proposals of global data structures covering global capture, nominal catch, catch and effort and logbook.

Technically speaking, FLUX is a language and not a system. It is a messenger that offers a protocol to create a secure and configurable network between different parties IT systems. UN/CEFACT provides a standardized schema for business process “[XML schemas](#)” and a standardized content called “Core

¹⁸ SDMX project name: EUROSTAT & FAO Joint Catch DSD https://sdmx.org/?page_id=4810

Components”. The components are harmonized and regularly published in UN/CEFACT [Core Component Library](#).

FLUX offers several advantages, including free, open and global standard to automate the collection and dissemination of the fishery catch data. It provides a common approach towards electronic logbooks for fishing vessels, interoperability between IT systems, and relatively easy exchange of data between parties. FLUX is strongly tied to XML as a data format.

United Nations Economic Commission for Europe (UNECE) established a [Team of Specialists](#) on Sustainable Fisheries to promote, facilitate and support the implementation of sustainable fisheries standards on a global scale and particularly the UN/FLUX. The project started as a small group initiative of European Union member states and later developed into a UN/CEFACT project which engaged experts from all regions of the world and made it a global standard.

Notwithstanding these advantages, implementation of FLUX is to be further explored for the purpose of interoperability of the CWP global data structures. Considering the growing importance of UN/FLUX in handling fisheries data, it is strategically important that the CWP standard for reference harmonization and embedded CWP standards and metadata are communicated to the UN/FLUX Team of Specialists so that this output be up taken in mainstream considerations by UN/FLUX for fisheries data exchange.

7. Registry and Catalogue

Upon its endorsement by CWP, the global CWP standard for reference harmonization and global data structures will be published through the CWP website, under a dedicated section.

Two alternatives of CWP catalogue were presented at the CWP 26th session and FAO was given guidance to work towards a centralized dissemination repository. FAO presented a proof of concept of a collaborative [CWP catalogue](#)¹⁹ based on CKAN, an open-source management system, which is a user-friendly discovery hub to enable easily pull and push of reference data and metadata. The CWP registry would be the index of data structures and reference data and mappings.

CWP classifications are hosted and maintained in the FAO Master Data Management tool. They are automatically pushed to populate CWP catalogue.

The CWP catalogue will disseminate:

- CWP global data structures definitions
- CWP international standard classifications and their Code lists.
- Code lists used by the CWP parties
- Mappings of the Parties' code lists.

¹⁹ The proof of concept of CWP Catalogue is not open to the public. Access requires authorization request.
https://bluebridge.d4science.org/web/cwp_secretariat/home?p_p_state=maximized&p_p_mode=view&saveLastPath=false&_58_struts_action=%2Flogin%2Flogin&p_p_id=58&p_p_lifecycle=0&_58_redirect=%2Fgroup%2Fcwp_secretariat%2Fcwp_secretariat

An implementation instance of ISSCFV in catalogue is available here²⁰.

7.1 Metadata

The metadata kept in the catalogue should be periodically checked for validity and links to the resources are still valid and responding. Metadata is also maintained in the FAO Master Data Management tool.

The metadata is published according to a standard model vocabulary based on DCAT which is an RDF vocabulary designed to facilitate interoperability between data catalogs published on the Web.

The following is an example of the metadata fields published through the CWP catalogue for each item:

Field	Value
Author	Coordinating Working Party on Fishery Statistics (CWP) Secretariat
Author email	CWP-Secretariat@fao.org
Maintainer	CWP Secretary - Stefania Vannuccini. Senior Fisheries Officer
Maintainer email	CWP-Secretariat@fao.org
Version	2020.1.0
Custodian	Food and Agriculture Organization of the United Nations
First endorsed	1980
Last update/revision	2016
Owner	Food and Agriculture Organization of the United Nations

²⁰ CWP catalogue <https://data.apps.fao.org/catalog/dataset/the-international-standard-statistical-classification-of-fishing-gear-isscfg>

Annex 1: CWP reference harmonization standard: Data structure of catch

MODULE/CONCEPT		CLASSIFICATION SYSTEM			CODE LIST			CODE LIST ID			DESCRIPTION			M/R/O
ADMINISTRATIVE/ POLITICAL ENTITY	FLAG ENTITY	UN Standard country or area codes for statistical use (M49)			M49 code list			M49_CODE			The M49 is presented in the CWP handbook where it is mapped to ISO Alpha2 and ISO Alpha3 list of countries and areas			M
GEOGRAPHIC AREA	FISHING AREA	FAO Major Fishing Areas for statistical purposes	Countries EEZ, continental seas and inland waters	RFB competence area	FAO Fishing Areas	Marine Regions	FAO RFB competence area	FAO WATER AREA GROUPS	EEZ MARINE REGIONS	RFB COMPETENCE AREAS	FAO Major Fishing Areas for statistical purposes and their breakdown http://www.fao.org/3/bt979e/bt979e.pdf	Marine Regions list of maritime boundaries and marine areas and locations. This register classification is managed by VLIZ Belgium Marine Institute http://www.marineregions.org/eez.php	RFB competence area codes are based on the RFB acronym (in English)	M
	TIME	Gregorian civil calendar according to ISO 8601			Time units			TIME_UNIT			The ISO 8601 provides coding conventions of these time resolution units e.g. 2019, Q1, 05/2019, days, ...			M
FISHING PRACTICE	FISHING GEAR	The International Standard Statistical Classification of Fishing Gear (ISSCFG)			Gear Type Code (standard abbreviation)			GEAR_A_CODE			Gear type code corresponding to the standard abbreviation of gear category http://www.fao.org/3/a-bt988e.pdf			M
	FISHING MODE	Gear Practice Qualifier			Free school/associated schools			FISH_MODE			Fishing mode utilized for catching tuna			O
CATCH	AQUATIC SPECIES	ASFIS List of Species for Fishery Statistics Purposes			Inter-agency 3-alpha code			3A_CODE			Species reference			M
	CATCH TYPE	CWP Catch type concepts (defined in the CWP handbook)			Catch type			CATCH_TYPE			Catch types presented in the catch diagram of the CWP handbook (gross catch, retained catch, landings, discards) http://www.fao.org/3/bt981t/bt981t.pdf			M
	OBS_MEASURE										Amount or quantity of the observation measure (a positive integer number)			M
	UNIT	Unified Code for Units of Measure (UCUM)			Units of measure			UNIT			Unit of measure (e.g. tonnes or number of individuals)			M
	OBS_STATUS	FAO statistical standard for Observation status flags			Observation Status Flag			FAO_FLAG_STATUS			FAO Observation status flagging codes (e.g. "E"Estimated value, "R"Revised, "U" unknown, "... "negligible, Official)			M
	COVERAGE				Coverage status			COVERAGE_STATUS			The degree of coverage of catch data for the fishing operations (not mandatory)			O
	CONFIDENTIALITY	SDMX cross-domain code lists			Confidentiality status			CONF_STATUS			Information about the sensitivity and confidentiality status of the data. (e.g. F: free for publication, N: Not for publication, restricted for internal use only) (not mandatory)			O

Annex 2: CWP reference harmonization standard: Data structure of Catch and Effort

MODULE/CONCEPT		CLASSIFICATION SYSTEM				CODE LIST				CODE LIST ID				DESCRIPTION				M/R/O
ADMINISTRATIVE/POLITICAL ENTITY	FLAG ENTITY	UN Standard country or area codes for statistical use (M49)				M49 code list				M49_CODE				The M49 is presented in the CWP handbook where it is mapped to ISO Alpha2 and ISO Alpha3 list of countries and areas				M
GEOGRAPHIC AREA	FISHING AREA	FAO Major Fishing Areas for statistical purposes	CWP Areal Grid System	Countries EEZ	RFB competence area	FAO Fishing Areas	CWP Areal Grid System	Marine Regions	FAO RFB competence area	FAO_WATERRA_GROUPS	SQUARE_ID	EEZ_MARINEREGIONS	RFB_COMPA REAS	FAO Major Fishing Areas for statistical purposes and their breakdown http://www.fao.org/3/bt979e/bt979e.pdf	CWP areal Grid system http://www.fao.org/cwp-on-fishery-statistics/handbook/general-concepts/major-fishing-areas-general/en/	Marine Regions list of maritime boundaries and marine areas and locations. This register classification is managed by VLIZ Belgium Marine Institute http://www.marineregions.org/ceez.php	RFB competence area codes are based on the RFB acronym (in English)	M
	TIME	Gregorian civil calendar according to ISO 8601				Time units				TIME_UNIT				The ISO 8601 provides coding conventions of these time resolution units e.g. 2019, Q1, 05/2019, days, ...				R
FISHING PRACTICE	FISHING GEAR	The International Standard Statistical Classification of Fishing Gear (ISSCFG)				Gear Type Code (standard abbreviation)				GEAR_A_CODE				Gear type code corresponding to the standard abbreviation of gear category http://www.fao.org/3/a-bt988e.pdf				M
	FISHING MODE	Gear Practice Qualifier				Free school /associated schools				FISH_MODE				Fishing mode (e.g. utilized for catching tuna) details in Annex8.				O
FLEET SEGMENT	FISHING VESSEL	International Standard Statistical Classification of Fishery Vessels by Vessel Types (ISSCFV-Vessel Type)				Code of Vessel Type (standard abbreviation)				FISH_VESSEL				The standard abbreviation of vessel type http://www.fao.org/3/a-bt983e.pdf				O
	LENGTH CLASSES	International Standard Statistical Classification of Vessels (ISSCFV - Length Classes) by Length Classes				Code Length Overall Classes				L_O_A_CLASS				Vessel Size by L.o.A. Classes (meters) http://www.fao.org/3/a-bt985e.pdf				O
EFFORT	OBS_MEASURE													The amount of fishing effort of a specific gear type over a certain period of time				M
	FISHING EFFORT UNIT	CWP Effort concepts definitions (For Fishing effort measures)				Effort measure descriptor				EFFORT_DESCRIPTOR				Measures of effort for each fishing gear (e.g. number of sets, number of hours fished, ...) http://www.fao.org/3/BS245E/bs245e.pdf				O
CATCH	AQUATIC SPECIES	ASFIS List of Species for Fishery Statistics Purposes				Inter-agency 3-alpha code				3A_CODE				Species reference				M
	CATCH TYPE	CWP Catch type concepts (defined in the CWP handbook)				Catch type				CATCH_TYPE				Catch types presented in the catch diagram of the CWP handbook (gross catch, retained catch, landings, discards) http://www.fao.org/3/bt981t/bt981t.pdf				R
	OBS_MEASURE													Amount or quantity of the observation measure (a positive integer number)				M
	UNIT	Unified Code for Units of Measure (UCUM)				Units of measure				UNIT				Unit of measure (e.g. tonnes or number of individuals)				R
	OBS_STATUS	FAO statistical standard for Observation status flags				Observation Status Flag				FAO_FLAG_STATUS				FAO observation status flagging codes (e.g. "E"Estimated value, "R"Revised, "U" unknown, "..." negligible, Official)				R
	COVERAGE					Coverage status				COVERAGE_STATUS				The degree of coverage of catch data for the fishing operations (not mandatory)				O
	CONFIDENTIALITY	SDMX cross-domain code lists				Confidentiality status				CONF_STATUS				Information about the sensitivity and confidentiality status of the data. (e.g. F: free for publication, N: Not for publication, restricted for internal use only) (not mandatory)				O

Annex 3: List of concepts and their definitions for implementation at schematic level. A classification system is provided for each concept.

(*) designates broad concepts (modules) with possible way of implementation by similar concepts depending on the context.

CONCEPT	DEFINITION (and context)	CLASSIFICATION SYSTEM
1. ADMINISTRATIVE/ POLITICAL ENTITY *	Basic concept (module) identifies the administrative or political entity to whom the data refers where the fish was caught, produced or landed.	<p>The "Standard Country or Area Codes for Statistical Use" which is owned by the United Nations Statistics Division (UNSD) and commonly referred to as the M49 standard.</p> <p>The M49 is presented in the CWP handbook where it is mapped to ISO Alpha2 and ISO Alpha3 list of countries and areas. (See Annex 6).</p> <p>It is noted that the actual occurrences of the "Fishing fleet" concept do not necessarily refer or correspond to a recognized country (e.g.: EUR - European Union, FRAT – French territories), nor to a distinct member / contracting party / cooperating, non-contracting party of a t-RFMO (e.g.: EU, ESP - EU (Spain), TWN – Chinese Taipei / Taiwan province of China – for some t-RFMOs). The proposed list of fishing fleet codes also includes a generic reference that applies to fishing operations and catches from unidentified sources (e.g.: NEI - not elsewhere identified).</p>
1.1. COUNTRY	It defines the assignment of nationality to catch and landings.	
1.2. FISHING FLEET	In the context of tuna RFMOs, it is a group of fishing vessels authorized to operate in a tuna RFMO convention area / area of competence, and whose fishing operations and catches of tuna and tuna-like species are responsibility of, and accounted for, by a political entity or sub-entity recognized by the corresponding tuna RFMO.	
1.3. FLAG ENTITY	In the context of fisheries operations using flags. It assigns nationalities to catch or landing.	
2. GEOGRAPHIC AREA /FISHING AREA *	Broad concept defines the geographic area which could be presented according to one of the following standard classifications.	FAO Major Fishing Areas for statistical purposes Countries EEZ, continental seas and inland waters
2.1. FAO MAJOR FISHING AREAS FOR STATISTICAL PURPOSES	They are arbitrary areas and historically determined coinciding to the greatest extent possible to the areas of competence of other fishery commissions when existing. This classification system facilitates comparison of data and improves the possibilities of cooperation in statistical matters.	The "FAO Major Fishing Areas for statistical purposes" is a hierarchical classification. It comprises 27 major fishing areas and breakdown. Each Major area is divided into subarea, division and subdivision. http://www.fao.org/3/bt979e/bt979e.pdf https://www.fao.org/cwp-on-fishery-statistics/handbook/general-concepts/fishing-areas-for-statistical-purposes/en/
2.2. CWP AREAL GRID SYSTEM	It is an areal breakdown that is essentially used for the context of Tuna fisheries (e.g. for gridded catches in the FIRMS Tuna Atlas)	Various maps, resolutions and square identifiers are provided here: http://www.fao.org/cwp-on-fishery-statistics/handbook/general-concepts/major-fishing-areas-general/en/
2.3. COUNTRY EEZ	It presents the Exclusive Economic Zone of countries as maritime boundaries for many countries. They are important for many applications such as delimitation of area if the sea in which sovereign state has special rights regarding the use of marine resources.	The Exclusive Economic Zone of countries as maritime boundaries is the relational classification list of geographic names, coupled with information and maps of the geographic location of these features. It is considered the standard reference for Maritime Boundaries representing the Exclusive Economic Zone of the world. https://www.marineregions.org/eez.php
2.4. RFB COMPETENCE AREA	It corresponds to the Regional Fisheries Bodies (RFBs).	The codes correspond to the acronyms of the Regional Fisheries Bodies (RFBs). The boundaries of their area of competence are presented here: http://www.fao.org/figis/geoserver/factsheets/rfbs.html http://www.fao.org/fishery/rfb/en
3. TIME*	It defines the time unit normally used in statistics.	
3.1. YEAR	The calendar (or Civil) year i.e., the period between 1 January and 31 December.	The ISO 8601 format using the Gregorian calendar.
3.2. TIME START/END	It defines the period in which fishing activity was performed, between time start and end.	ISO 8601 format using the Gregorian calendar and 24-hour time keeping system.
3.3. QUARTER	Quarter of calendar year (Time is referred to Q1 – Q4).	The ISO 8601-2:2019 defines a set of standardised extensions to the ISO 8601 date and time formats. Extended Date/Time Format (EDTF) provides sub-year groupings (Quarter1 – Quarter 4). https://www.loc.gov/standards/datetime/
4. AQUATIC SPECIES	It corresponds to any taxon that lives in water for most or all of its lifetime.	The "ASFIS List of Species for Fishery Statistics Purpose" https://www.fao.org/cwp-on-fishery-statistics/handbook/general-concepts/identifiers-for-aquatic-animals-and-plants/en/
5. CATCH TYPE	It defines the catch types (gross catch, retained catch, landings, discards).	CWP catch diagram of the CWP handbook http://www.fao.org/3/bt981t/bt981t.pdf

6. FISHING PRACTICE*	It is a broad concept that comprises the fishing gear and the fishing mode.	
7. FISHING GEAR	It defines the tool and equipment used to capture marine/aquatic resources.	The “International Standard Statistical Classification of Fishing Gear” (ISSCFG) is the reference standard. The codes should be chosen from corresponding to the standard abbreviation of gear category. The ISSCFG presents codes of major categories and sub-categories of gear. For this, CWP members or other producers of statistics often use their list of additional codes according to their context. It is recommended to have a mapping for their groupings with the ISSCFG categories.
8. FISHING MODE	It is associated to the fishing gear under one broad concept as “fishing practice” to enhance the fishing effort definition. It is commonly used in the context of tuna fisheries.	No standard available. See Annex 8 which presents the code list used by WECAFC and tuna RFMOs.
9. FISHING EFFORT UNIT	CWP Effort concepts definitions (For Fishing effort measures)	Measures of effort for each fishing gear (e.g. number of sets, number of hours fished, ...) http://www.fao.org/3/BS245E/bs245e.pdf
10.FLEET SEGMENT*	The combination of a group of fishing vessels of the same size and using the same gear for more than 50 percent of the time at sea during a year. To accommodate both regional and international fisheries organizations classifications, an organization can adopt a Fleet segment classification and coding system.	
11.FISHING VESSEL	It defines the type of vessel used.	The “International Standard Statistical Classification of Fishery Vessels by Vessel Types” (ISSCFV-Vessel Type). http://www.fao.org/3/a-bt983e.pdf
12.LENGTH CLASSES	It provides the “Length overall” classes for the vessel size.	The “International Standard Statistical Classification of Vessels by Length Classes” (ISSCFV - Length Classes). http://www.fao.org/3/a-bt985e.pdf
13.UNIT	It defines the unit of measure (e.g. tonnes or number of individuals).	The Unified Code for Units of Measure (UCUM). Exclusively using tonnes for catches: “t” (Metric tons), “no” (Number of fishes).
14.OBSERVATION STATUS	It provides information about the status and quality of an observation (i.e. value or an unusual or missing value). It is commonly used for dissemination purposes.	The FAO statistical standard for observation status flags (Annex 9) is based on the SDMX Code list for Observation Status (CL_OBS_STATUS) ²¹ .
15. CONFIDENTIALITY	Information about the sensitivity and confidentiality status of the data.	The classification system and code list used corresponds to the SDMX Cross-Domain code list for Confidentiality Status (CONF STATUS) ²² . The codes and their descriptions (and annotations) are available in Annex 11.
16. OBSERVATION MEASURE	Amount or quantity of the observation measure (a positive integer number).	

²¹ SDMX “Observation Status List” (2016) https://sdmx.org/wp-content/uploads/CL_OBS_STATUS_v2_1.docx

²² https://sdmx.org/?page_id=3215

Annex 4: List of concepts and their implementation at semantic level. For each concept, code list, code list identifier and its semantic representation are provided.

CONCEPT	CONCEPT ID (Column name)	CODE LIST ID	TARGET VALUE	REPRESENTATION
1. ADMINISTRATIVE/ POLITICAL ENTITY *				
1. COUNTRY	COUNTRY	M49_CODE	ISO-alpha3 code and/or ISO-alpha2 code and/or	Name in English or any other language can be added in another column. the code list can be extended i.e. Annex 6
2. FISHING FLEET	FISHING_FLEET			
3. FLAG ENTITY	FLAG_ENTITY			
2. GEOGRAPHIC AREA /FISHING AREA *	FISHING_AREA	CL_FAO_WATERAREA		
1. FAO MAJOR FISHING AREAS FOR STATISTICAL PURPOSES	AREA	CL_FAO_WATERAREA	Code in the classification	It can be represented as one column where codes are composed of the classification breakdown (e.g. 27_5_B_1_B corresponds to Atlantic Northeast 27.5.b.1.b)
	MAJOR_AREA SUBAREA DIVISION SUBDIVISION	CL_FAO_WATERAREA	Code in the classification	In flat structure, each level is represented in one column (e.g. MAJOR_AREA: 27; SUBAREA: 5). These levels can still be combined/aggregated to be used in a normalized database
2. CWP AREAL GRID SYSTEM	SQUARE_ID	CL_SQUARE_ID	GRID_CODE	To be confirmed
3. COUNTRY EEZ	COUNTRY_EEZ	EEZ_MARINEREGIONS	MRGID (MARINE REGION ID in the classification)	TO BE CONFIRMED: Cl_MRGID as code list
4. RFB COMPETENCE AREA	RFB_AREA	RFB_AREA	Code in Annex 7	
3. TIME *	TIME		YYYY in ISO 8601	
1. YEAR	YEAR		YYYY in ISO 8601	There is no specific code list but ISO 8601 is used for coding. The year can be extended to YYYY-MM if needed
2. TIME START/END	TIME_START TIME_END		YYYY-MM-DD in ISO 8601	“PERIOD” is a similar concept that can be used to express time interval. There are other ways to express the time interval: Start and end or Start and duration
3. QUARTER	QUARTER		Sub-grouping year, Quarter in ISO 8601-2:2019	Time is referred to Quarter 1 – Quarter 4.
4. AQUATIC SPECIES	SPECIES	CL_SPECIES	3A_CODE in ASFIS list	Additional columns from the same classification can be added in the same data structure (e.g. Scientific Name, Name En)
5. CATCH TYPE	CATCH_TYPE	CL_CATCH_TYPE		TO BE CONFIRMED: List of catch types: discards, nominal catches
6. FISHING PRACTICE*				

7. FISHING GEAR	GEAR_TYPE	GEAR_A_CODE	Code corresponding to the standard abbreviation of gear category in ISSCFG (e.g. PS, LHP)	
8. FISHING MODE	FISHING_MODE	FISHING_MODE	Code in Annex 8	
9. FISHING EFFORT UNIT	EFFORT_UNIT	EFFORT_DESCRIPTOR		TO BE CONFIRMED: List of fishing effort units
10. FLEET SEGMENT*	FLEET_SEGMENT	FLEET_SEGMENT		
11. FISHING VESSEL	FISHING_VESSEL	FISHING_VESSEL	Code of Vessel Type (standard abbreviation)	
12. LENGTH CLASSES	L_O_A_CLASS	L_O_A_CLASS	Code of Length Overall Classes	
13. UNIT	UNIT	CL_UNIT		It could be represented in the same column of the data structure (tonnes and/or number)
14. OBSERVATION STATUS	OBS_STATUS	CL_OBS_STATUS	Code in Annex 9	
15. CONFIDENTIALITY	CONF_STATUS	CONF_STATUS		

Annex 5: Alternatives to handle levels of granularity

Level of granularity defines the sub-classification used by the user (e.g. CWP party) for their specific requirements, the level of details included in the data structure respectively when building on the classification system. The decision of the granularity level resides in the choice of the user who wants to report the data. For instance, within or “under” the classification system “FAO Major Marine Fishing Areas”, the data structure can include breakdowns: Subarea, Division or Subdivision. In the same order of ideas, ICES subareas would be considered as areas at lower level of granularity within the major Fishing Area 27.

Another example would be the aggregations of 3-Alpha code species from the ASFIS classification where aggregation of species is shaped by “building on top”/based on the classification system. ISSCAAP groups are an obvious example of grouping of ASFIS codes used as part of the ASFIS classification system. Another example is encountered in the case of species groupings used by tuna RFMOs, where a tuna RFMO specific classification system is built on top of the ASFIS species codes.

In all these cases of whether higher or lower level of aggregation, mapping against standard classification codes is crucial to be integrated in the data structure.

In general, these classification items do not need to be all at the same level in the base statistical classification. To these classification items, one or more new levels may be added. This can include extending the base statistical classification with one or several new levels at the bottom of its base, creating a new lowest level. To meet specific needs, a number of alternatives may be created, in which the original categories are split or regrouped to provide context-specific additions or alternatives to the standard aggregation structure. These are called classification variants.

Variants are commonly of three kinds. These have been named **extension variants**, **aggregate variants** or **regrouping variants**. There could exist other types of variants. A particular variant could include elements from more than one of these variant types.

Extension variant: An extension variant is a statistical classification that extends the base statistical classification with one or several new levels at the bottom, creating a new lowest level. An extension variant thus adds new lower levels to the base classification but does not otherwise alter its original structure.

Aggregate variant: An aggregate variant is a classification that groups the categories of a linear statistical classification to create one or several aggregate level(s), thus creating a hierarchy.

Regrouping variant: A regrouping variant is a classification that introduces additional or alternative aggregate levels by regrouping categories of the base statistical classification. Two types of regrouping variants have been identified:

- a) Regrouping variants which do not violate the structure of the base statistical classification: This type of regrouping variant introduces a new level or new levels on top of, or in between existing levels of a hierarchical statistical classification without otherwise altering the original structure of the hierarchy. This regrouping variant consists of all classification Levels of the base classification plus the new variant level(s). The parent level (if any) of the new variant level can be either another variant level or a level from the base statistical classification.
- b) Regrouping variants which violate the structure of the base statistical classification: This type of regrouping variant introduces a new level or new levels on top of any but the topmost level of a hierarchical statistical classification by regrouping categories of the base statistical classification in a way which violates its original order and structure. This regrouping variant consists of all

classification levels of the base statistical classification below the new variant Level(s) plus the new variant level(s). In such a regrouping variant, a new variant level cannot have a base statistical classification level as parent level.

In all variants except regrouping variants which violate the structure of the base statistical classification, all levels of the base statistical classification are retained and one or more new Levels are inserted. In regrouping variants which violate the structure of the base statistical classification, one or more new levels are inserted and only the base statistical classification levels below the new variant levels are retained.

It is sometimes debated whether a classification database should be descriptive or prescriptive, the idea being that a prescriptive database will contain only standard classifications, whereas a descriptive database will also contain non-standard variants. The demarcation between standard and non-standard classifications or between these and more loosely structured groupings is not clear. It seems, therefore, that the criterion for inclusion in the database cannot be formal status only, but just as much the usefulness and commonality of the information provided. Most of the time the departures from the norm are legitimate, made to meet specific producer requirements or user needs. In any case alternative groupings exist and must be documented. Indeed, listing the non-standard variants used in a statistical office may be a first and necessary step towards reducing their numbers.

Annex 6: Code list of M49 country or area, active as of October 2021 (M49_CODE)

A NOTE OF CAUTION: The designations employed and the presentation of material of this list do not imply the expression of any opinion whatsoever on the part of FAO concerning the legal status of any country, territory, or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

ISO – alpha2 Code	ISO - alpha 3 Code	Name
004	AFG	Afghanistan
248	ALA	Åland Islands
008	ALB	Albania
012	DZA	Algeria
016	ASM	American Samoa
020	AND	Andorra
024	AGO	Angola
660	AIA	Anguilla
010	ATA	Antarctica
028	ATG	Antigua and Barbuda
032	ARG	Argentina
051	ARM	Armenia
533	ABW	Aruba
036	AUS	Australia
040	AUT	Austria
031	AZE	Azerbaijan
044	BHS	Bahamas
048	BHR	Bahrain
050	BGD	Bangladesh
052	BRB	Barbados
112	BLR	Belarus
056	BEL	Belgium
084	BLZ	Belize
204	BEN	Benin
060	BMU	Bermuda
064	BTN	Bhutan
068	BOL	Bolivia (Plurinational State of)
535	BES	Bonaire, Sint Eustatius and Saba
070	BIH	Bosnia and Herzegovina
072	BWA	Botswana
074	BVT	Bouvet Island
076	BRA	Brazil
086	IOT	Chagos Archipelago
092	VGB	British Virgin Islands
096	BRN	Brunei Darussalam
100	BGR	Bulgaria
854	BFA	Burkina Faso
108	BDI	Burundi
132	CPV	Cabo Verde
116	KHM	Cambodia
120	CMR	Cameroon

124	CAN	Canada
136	CYM	Cayman Islands
140	CAF	Central African Republic
148	TCD	Chad
152	CHL	Chile
156	CHN	China
344	HKG	China, Hong Kong SAR
446	MAC	China, Macao
162	CXR	Christmas Island
166	CCK	Cocos (Keeling) Islands
170	COL	Colombia
174	COM	Comoros
178	COG	Congo
184	COK	Cook Islands
188	CRI	Costa Rica
384	CIV	Côte d'Ivoire
191	HRV	Croatia
192	CUB	Cuba
531	CUW	Curaçao
196	CYP	Cyprus
203	CZE	Czechia
408	PRK	Democratic People's Republic of Korea
180	COD	Democratic Republic of the Congo
208	DNK	Denmark
262	DJI	Djibouti
212	DMA	Dominica
214	DOM	Dominican Republic
218	ECU	Ecuador
818	EGY	Egypt
222	SLV	El Salvador
226	GNQ	Equatorial Guinea
232	ERI	Eritrea
233	EST	Estonia
748	SWZ	Eswatini
231	ETH	Ethiopia
238	FLK	Falkland Islands (Malvinas)
234	FRO	Faroe Islands
242	FJI	Fiji
246	FIN	Finland
250	FRA	France
254	GUF	French Guiana
258	PYF	French Polynesia
260	ATF	French Southern Territories
266	GAB	Gabon

270	GMB	Gambia
268	GEO	Georgia
276	DEU	Germany
288	GHA	Ghana
292	GIB	Gibraltar
300	GRC	Greece
304	GRL	Greenland
308	GRD	Grenada
312	GLP	Guadeloupe
316	GUM	Guam
320	GTM	Guatemala
831	GGY	Guernsey
324	GIN	Guinea
624	GNB	Guinea-Bissau
328	GUY	Guyana
332	HTI	Haiti
334	HMD	Heard Island and McDonald Islands
336	VAT	Holy See
340	HND	Honduras
348	HUN	Hungary
352	ISL	Iceland
356	IND	India
360	IDN	Indonesia
364	IRN	Iran (Islamic Republic of)
368	IRQ	Iraq
372	IRL	Ireland
833	IMN	Isle of Man
376	ISR	Israel
380	ITA	Italy
388	JAM	Jamaica
392	JPN	Japan
832	JEY	Jersey
400	JOR	Jordan
398	KAZ	Kazakhstan
404	KEN	Kenya
296	KIR	Kiribati
412		Kosovo (Serbia)
414	KWT	Kuwait
417	KGZ	Kyrgyzstan
418	LAO	Lao People's Democratic Republic
428	LVA	Latvia
422	LBN	Lebanon
426	LSO	Lesotho
430	LBR	Liberia
434	LBY	Libya
438	LIE	Liechtenstein
440	LTU	Lithuania
442	LUX	Luxembourg
450	MDG	Madagascar
454	MWI	Malawi

458	MYS	Malaysia
462	MDV	Maldives
466	MLI	Mali
470	MLT	Malta
584	MHL	Marshall Islands
474	MTQ	Martinique
478	MRT	Mauritania
480	MUS	Mauritius
175	MYT	Mayotte
484	MEX	Mexico
583	FSM	Micronesia (Federated States of)
492	MCO	Monaco
496	MNG	Mongolia
499	MNE	Montenegro
500	MSR	Montserrat
504	MAR	Morocco
508	MOZ	Mozambique
104	MMR	Myanmar
516	NAM	Namibia
520	NRU	Nauru
524	NPL	Nepal
528	NLD	Netherlands
540	NCL	New Caledonia
554	NZL	New Zealand
558	NIC	Nicaragua
562	NER	Niger
566	NGA	Nigeria
570	NIU	Niue
574	NFK	Norfolk Island
807	MKD	North Macedonia
580	MNP	Northern Mariana Islands
578	NOR	Norway
512	OMN	Oman
586	PAK	Pakistan
585	PLW	Palau
275	PSE	Palestine
591	PAN	Panama
598	PNG	Papua New Guinea
600	PRY	Paraguay
604	PER	Peru
608	PHL	Philippines
612	PCN	Pitcairn
616	POL	Poland
620	PRT	Portugal
630	PRI	Puerto Rico
634	QAT	Qatar
410	KOR	Republic of Korea
498	MDA	Republic of Moldova
638	REU	Réunion
642	ROU	Romania
643	RUS	Russian Federation

646	RWA	Rwanda
652	BLM	Saint Barthélemy
654	SHN	Saint Helena
659	KNA	Saint Kitts and Nevis
662	LCA	Saint Lucia
663	MAF	Saint Martin (French Part)
666	SPM	Saint Pierre and Miquelon
670	VCT	Saint Vincent and the Grenadines
882	WSM	Samoa
674	SMR	San Marino
678	STP	Sao Tome and Principe
680	XSQ	Sark
682	SAU	Saudi Arabia
686	SEN	Senegal
688	SRB	Serbia
690	SYC	Seychelles
694	SLE	Sierra Leone
702	SGP	Singapore
534	SXM	Sint Maarten (Dutch part)
703	SVK	Slovakia
705	SVN	Slovenia
090	SLB	Solomon Islands
706	SOM	Somalia
710	ZAF	South Africa
239	SGS	South Georgia and the South Sandwich Islands
728	SSD	South Sudan
724	ESP	Spain
144	LKA	Sri Lanka
729	SDN	Sudan
740	SUR	Suriname
744	SJM	Svalbard and Jan Mayen Islands
752	SWE	Sweden
756	CHE	Switzerland
760	SYR	Syrian Arab Republic
158	TWN	Taiwan, Province of China
762	TJK	Tajikistan
764	THA	Thailand
626	TLS	Timor-Leste
768	TGO	Togo
772	TKL	Tokelau
776	TON	Tonga
780	TTO	Trinidad and Tobago
788	TUN	Tunisia
792	TUR	Turkey
795	TKM	Turkmenistan
796	TCA	Turks and Caicos Islands
798	TUV	Tuvalu
800	UGA	Uganda
804	UKR	Ukraine

784	ARE	United Arab Emirates
826	GBR	United Kingdom of Great Britain and Northern Ireland
834	TZA	United Republic of Tanzania
836		United Republic of Tanzania, Zanzibar
581	UMI	United States Minor Outlying Islands
840	USA	United States of America
850	VIR	United States Virgin Islands
858	URY	Uruguay
860	UZB	Uzbekistan
548	VUT	Vanuatu
862	VEN	Venezuela (Bolivarian Republic of)
704	VNM	Viet Nam
876	WLF	Wallis and Futuna Islands
732	ESH	Western Sahara
887	YEM	Yemen
894	ZMB	Zambia
716	ZWE	Zimbabwe

Annex 7: Code list of RFB area of competence (RFB_AREA)

Code	Full name
ACAP	Agreement on the Conservation of Albatrosses and Petrels
ACFR	ACFR area of competence
APFIC	Asia-Pacific Fishery Commission
BOBP-IGO	BOBP-IGO area of competence
CACFISH	CACFish area of competence
CARPAS	CARPAS area of competence
CBLT	CBLT area of competence
CCAMLR	CCAMLR area of competence
CCBSP	CCBSP area of competence
CCSBT	CCSBT area of competence
CECAF	Fishery Committee for the Eastern Central Atlantic
CIFA	CIFA area of competence
CIFAA	CIFAA area of competence
COMHAFAT-ATLAFCO	COMHAFAT/ATLAFCO area of competence
COPPESAALC	COPPESAALC area of competence
COREP	COREP area of competence
CPPS	CPPS area of competence
CRFM	CRFM area of competence
CTMFM	CTMFM area of competence
EIFAAC	EIFAAC area of competence
FCWC	FCWC area of competence
FFA	FFA area of competence
GFCM	GFCM area of competence
IATTC	IATTC area of competence
IBSFC	IBSFC area of competence
ICCAT	ICCAT area of competence
ICES	ICES area of competence
ICSEAF	ICSEAF area of competence
IOTC	IOTC area of competence
IPHC	IPHC area of competence
IWC	IWC area of competence
JOINTFISH	Joint Fish area of competence
LTA	LTA area of competence
LVFO	LVFO area of competence
MHLC	MHLC area of competence
MRC	MRC area of competence
NACA	NACA area of competence
NAFO	NAFO area of competence
NAMMCO	NAMMCO area of competence
NASCO	NASCO area of competence
NEAFC	NEAFC Convention Area
NPAFC	NPAFC area of competence

NPFC	NPFC area of competence
OAPO	OAPO area of competence
OLDEPESCA	OLDEPESCA area of competence
OSPESCA	OSPESCA area of competence
PERSGA	PERSGA area of competence
PICES	PICES area of competence
PSC	PSC area of competence
RAA	RAA area of competence
RCF	RCF area of competence
RECOFI	RECOFI area of competence
SEAFDEC	SEAFDEC area of competence
SEAFO	SEAFO area of competence
SIOFA	SIOFA area of competence
SPC	SPC area of competence
SPRFMO	SPRFMO area of competence
SRCF	SRCF area of competence
SRFC	SRFC area of competence
SWIOFC	Western Central Atlantic Fishery Commission
WCPFC	WCPFC area of competence
WECAFC	Western Central Atlantic Fishery Commission
WIOTO	WIOTO area of competence

Annex 8: Code list for fishing mode used by WECAFC (FISHING_MODE)

Code	Name	Description
N/A	n-a	Not applicable
ALL	All	All fishing modes reported together
FREE	Free	Fishing on free school, no FAD use
ASSO	Associated	Fishing on FAD associated school
DIVE	Diving	Fishing with one or more divers

Annex 9: Code list of observation status flags for dissemination (FAO_FLAG_STATUS)

http://intranet.fao.org/fileadmin/user_upload/scp/Standards_for_quality_compliance/SSS1_Flags.pdf

https://sdmx.org/wp-content/uploads/CL_OBS_STATUS_implementation_20-10-2014.pdf

Code	Description	Annotation
A	Normal value	To be used as default value if no value is provided or when no special coded qualification is assumed. Usually, it can be assumed that the source agency assigns sufficient confidence to the provided observation and/or the value is not expected to be dramatically revised.
B	Time series break	Observations are characterized as such when different content exists or a different methodology has been applied to this observation as compared with the preceding one (the one given for the previous period).
D	Definition differs	Used to indicate slight deviations from the established methodology (footnote-type information); these divergences do not imply a break in time series.
E	Estimated value	Observation obtained through an estimation methodology (e.g. to produce back-casts) or based on the use of a limited amount of data or ad hoc sampling and through additional calculations (e.g. to produce a value at an early stage of the production stage while not all data are available). It may also be used in case of experimental data (e.g. in the context of a pilot ahead of a full-scale production process) or in case of data of (anticipated/assessed) low quality. If needed, additional information can be provided through free text using the COMMENT_OBS attribute at the observation level or at a higher level.
F	Forecast value	Value deemed to assess the magnitude which a quantity will assume at some future point of time (as distinct from "estimated value" which attempts to assess the magnitude of an already existent quantity).
G	Experimental value	Data collected on the basis of definitions or (alternative) collection methods under development. Data not of guaranteed quality as normally expected from provider.
I	Imputed value (CCSA definition)	Observation imputed by international organizations to replace or fill gaps in national data series, in line with the recommendations of the United Nations Committee for the Coordination of Statistical Activities (CCSA).
K	Data included in another category	This code is used when data for a given category are missing and are included in another category, sub-total or total. Generally, where code "K" is used there should be a corresponding code "W - Includes data from another category" assigned to the over-covered category. Implementers and data reporters should use the COMMENT_OBS observation-level attribute to specify under which category the data are included.
W	Includes data from another category	This code is used when data include another category or go beyond the scope of the data collection and are therefore over-covered. Generally, where code "W" is used there should be a corresponding code "K - Data included in another category" assigned to the category which is under-covered. Implementers and data reporters should use the COMMENT_OBS observation-level attribute to specify which additional data are included.

O	Missing value	This code is to be used when no breakdown is made between the reasons why data is missing. Data can be missing due to many reasons: data cannot exist, data exist but are not collected (e.g. because they are below a certain threshold or subject to a derogation clause), data are unreliable, etc.
M	Missing value: data cannot exist	Used to denote empty cells resulting from the impossibility to collect a statistical value (e.g. a particular education level or type of institution may not be applicable to a given country's education system). ²³
P	Provisional value	An observation is characterized as "provisional" when the source agency – while it bases its calculations on its standard production methodology – considers that the data, almost certainly, are expected to be revised.
S	Strike and other special events	Special circumstances (e.g. strike) affecting the observation or causing a missing value.
L	Missing value: data exist but were not collected	Used, for example, when some data are not reported/disseminated because they are below a certain threshold.
H	Missing value; holiday or weekend	Used in some daily data flows.
Q	Missing value; suppressed	Used, for example, when data are suppressed due to statistical confidentiality considerations.
J	Derogation	Clause in an agreement (e.g. legal act, gentlemen's agreement) stating that some provisions in the agreement are not to be implemented by designated parties; these derogations may affect the observation or cause a missing value. In general, derogations are limited in time.
N	Not significant	Used to indicate a value which is not a "real" zero (e.g. a result of 0.0004 rounded to zero).
U	Low reliability	This indicates existing observations, but for which the user should also be aware of the low quality assigned.
V	Unvalidated value	Observation as received from the respondent without further evaluation of data quality.

²³ This concept is sometimes referred to as "not applicable". However, it is important to note that "not applicable" as meaning "data cannot exist" is different from the concept of "not applicable" (represented with code "_Z") as described in the list of SDMX generic codes presented in the "Guidelines for the Creation and Management of SDMX Cross-Domain Code Lists". Code "M – Missing value; data cannot exist" as mentioned here is used to characterize the impossibility for a statistical value to exist. As such it is a quality indicator used as an attribute to a statistical value. _Z is a residual category in a code list to which statistical values are attached.

Annex 10: Code list of coverage rate proposed by IOTC (COVERAGE_STATUS)

This proposal needs to be standardized to accommodate CWP members' requirements.

Code	Description
UP	Statistics partially raised; coverage unknown
UR	Statistics raised; coverage unknown
US	Statistics not raised; coverage unknown
UT	Total enumeration
UU	Not sampled
B0	Less than 5% of the boats covered
B1	Between 5%-9% of the boats covered
B3	Between 10%-29% of the boats covered
B7	Between 30%-69% of the boats covered
B9	70% or more of the boats covered
T0	Less than 5% of the trips covered
T1	Between 5%-9% of the trips covered
T3	Between 10%-29% of the trips covered
T7	Between 30%-69% of the trips covered
T9	70% or more of the trips covered
N0	Less than 5% of the fish sampled (in number)
N1	Between 5%-9% of the fish sampled (in number)
N3	Between 10%-29% of the fish sampled (in number)
N7	Between 30%-69% of the fish sampled (in number)
N9	70% or more of the fish sampled (in number)
W0	Less than 5% of the fish sampled (in weight)
W1	Between 5%-9% of the fish sampled (in weight)
W3	Between 10%-29% of the fish sampled (in weight)
W7	Between 30%-69% of the fish sampled (in weight)
W9	70% or more of the fish sampled (in weight)

Annex 11: Code list for Confidentiality Status (CONF_STATUS)

Code	Recommended code description	Description
F	Free (free for publication)	Used for observations without any special sensitivity considerations and which can thus be freely shared. Usually, source organizations provide information and guidance on general requirements for re-dissemination (like mentioning the source) either on their websites or in their paper publications. In some institutional environments the term "unclassified" is used in a sense that still denotes implied restrictions in the circulation of information. If this is the case, the organizations concerned may probably consider that "free" (value F) is not the appropriate tag for this kind of "unclassified" category and that "Not for publication, restricted for internal use only" (value N) may be more appropriate.
N	Not for publication, restricted for internal use only	Used to denote observations that are restricted for internal use only within organizations. This code may be accompanied with an additional observation-level attribute: CONF_REDIST which defines the secondary recipient(s) to whom the sender allows the primary recipient to forward confidential data.
C	Confidential statistical information	Confidential statistical information (primary confidentiality) due to identifiable respondents. Measures also should be taken to prevent not only direct access, but also indirect deduction or calculation by other users and parties, probably by considering and treating additional observations as "confidential" (secondary confidentiality management).
D	Secondary confidentiality set by the sender, not for publication	Used by the sender of the data to flag (beyond the confidential statistical information) additional observations in the dataset so that the receiver knows that he/she should suppress these observations in subsequent stages of processing (especially dissemination) in order to prevent third parties to indirectly deduct the observations that are genuinely flagged with "C".
S	Secondary confidentiality set and managed by the receiver, not for publication	If senders do not manage the secondary confidentiality in their data and/or there are also other countries' data involved (with the intention to eventually compile a regional-wide aggregate that is going to be published), the value "S" is used by the receiver to flag additional suppressed observations (within sender's data and/or within the datasets of other senders) in subsequent stages of processing (especially, dissemination) in order to prevent third parties to indirectly deduct the observations that were genuinely flagged with "C" by the sender.
A	Primary confidentiality due to small counts	A cell is flagged as confidential if less than m units ("too few units") contribute to the total of that cell. The limits of what constitutes "small counts" can vary across statistical domains, countries, etc.
O	Primary confidentiality due to dominance by one unit	Used when one unit accounts for more than x % of the total of a cell. The value of x can vary across statistical domains or countries, be influenced by legislation, etc.
T	Primary confidentiality due to dominance by two units	Used when two units account for more than x % of the total of a cell. The value of x can vary across statistical domains or countries, be influenced by legislation, etc.
G	Primary confidentiality due to dominance by one or two units	Used when one or two units account(s) for more than x % of the total of a cell. The value of x can vary across statistical domains or countries, be influenced by legislation, etc.
M	Primary confidentiality due to data declared confidential based on other measures of concentration	Cells declared confidential using mathematical definitions of sensitive cells, e.g. p-percent, p/q or (n,k) rules.

E	Not for publication, restricted for internal use only (equivalent to the code N) until the embargo time elapses; Free for publication (equivalent to the code F) after the embargo time elapses.	Used for embargoed data. The embargo time has to be specified in the EMBARGO_TIME attribute. This code may be accompanied with an additional observation-level attribute: CONF_REDIST which defines the secondary recipient(s) to whom the sender allows the primary recipient to forward confidential data.
---	--	---