EVERYFISH

EVERYFISH-COCO FORMAT SPECIFICATIONS

INNOVATION FOR SUSTAINABLE FISHERIES

Project number: 101059892
Project duration: 1 Jan 2023 – 31 Dec 2026
Website: www.EVERYFISH.eu

Deliverable ID:	**Due month:**	**Published date:**
NA	NA	

Title: EVERYFISH-COCO format specifications

Lead beneficiary: UEA, IMR

Prepared by: IMR

Abstract

This document describes the folder structure and file format used to document the image annotations developed and used in EVERYFISH. It is based on the COCO (Common Objects in COntext) format, with additional attributes to describe fisheries-related aspects, using standard vocabularies commonly used in marine science.

Authorship information

Contributing partners	Hang Zhou (UEA), Michal Mackiewicz (UEA), Iñaki Quincoces (AZTI), Manuel Cordova Neira (WR), Rocío Castaño Primo (IMR), Gert Kootstra (WR), John Reidar Mathisen (SINTEF), Rebecca Skirrow (CEFAS), Serkan Kartal (ÇU), Ahmad Kamal (AZTI), Helge Sagen (IMR).

Version history

Version number	Date	Description of changes
1.0	2025-12-15	First version of the EVERYFISH-COCO format specifications

Suggested citation. Zhou H., Mackiewicz, M., Quincoces, I., Cordova Neira, M., Castaño Primo, R., Kootstra, G., Mathisen, J.R., Skirrow, R., Kartal, S., Kamal, A., Sagen, H. (2026). EVERYFISH-COCO format specifications. Version 1.0. https://doi.org/10.21335/NMDC-450800267

Table of contents

1. Introduction

1.1. EVERYFISH HEU motivation and background

The EVERYFISH project is a close collaboration between fisheries research, academic, industry, and governance actors, as well as technology providers across the EU and other European countries. These technological innovations will enable the fisheries sector to optimise automatic data collection and ensure correct reporting of the catch in terms of size, weight, and species. The provision of verifiable catch information and improved evidence of compliance with fisheries regulations will streamline compliance activities for both fisheries and regulators, as well as generating knowledge and data for future conservation, research, and governance activities.

EVERYFISH builds on and exploits the work done in multiple previous projects and further develop key technologies to increase their technology readiness level (TRL). In total, EVERYFISH will develop 10 products and apps for automatic catch recording and reporting, as well as guidelines for implementation and integration of EVERYFISH technologies in commercial fisheries in Europe to optimize resource efficiency, improve automatic data collection, provide evidence of compliance with fishery regulations, and reduce the ecological impact of the sector on the marine environment (Figure 1).

The methods and technologies developed in EVERYFISH will be ready for implementation in commercial fisheries. We assess their TRL in terms of the maturity of the underlying technologies in a fisheries context as several of them, e.g. AI, are already quite mature and used extensively in other sectors but still immature and relatively little used in fisheries management. All the products and apps developed in EVERYFISH are useful as standalone systems to provide automatic catch registration on vessels and by fishers that use them. However, their greatest potential will be realised when the data is standardized and automatically transmitted via a distributed ledger, to be used by stock assessors, fisheries monitoring and management agencies, certification companies, and others.

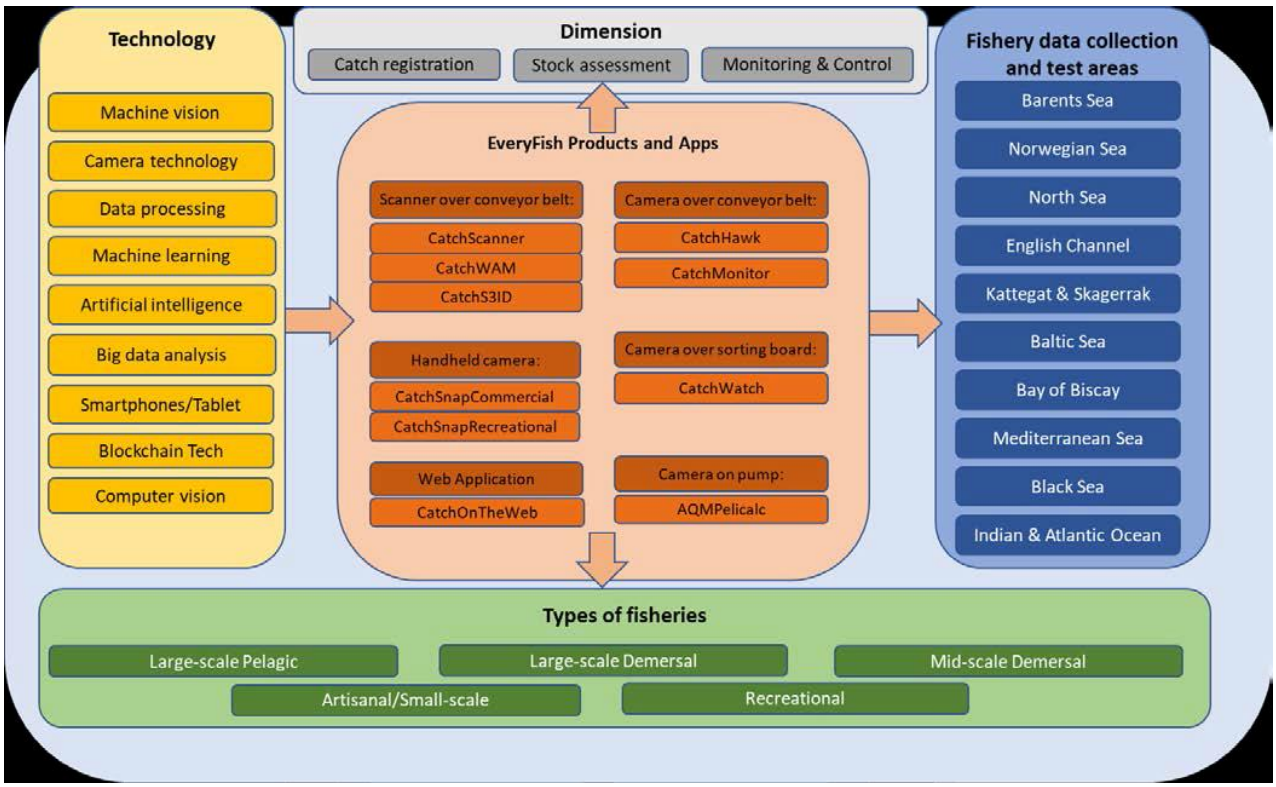


Figure 1: The conceptual structure of EVERYFISH HEU.

2. Main overview of the EVERYFISH-COCO format

During the EVERYFISH project (https://doi.org/10.3030/101059892) the consortium has defined a standard for the annotation and storage of image data targeted for fish detection. Defining a common standard is essential to ensure that data collected (images and video) can be shared, interpreted, and used easily within the consortium. To support this interoperability, a unified annotation standard, named EVERYFISH-COCO, was established. The goal is to maximise the usability of the collected datasets and enable the different AI models created within EVERYFISH to benefit from it as much as possible. Outside and after the project, the format could become a standard in the annotation and processing of image-based fish data.

The EVERYFISH-COCO data structure defines properties related to images, videos, licenses, categories, and metadata containing contextual information about the fishing activity. Information described in this structure is based on the widely used COCO (Common Objects in Context) annotation format[1], which uses json files to store the annotations and includes different kind of ground-truth information for different computer-vision tasks: object detection, keypoint detection, semantic/instance/panoptic segmentation. The COCO format contains metadata that describes both the content and the context of the annotated dataset.

The EVERYFISH-COCO data format preserves full compatibility with the COCO format using the same attribute (key) names defined in COCO, and providing additional information like fisheries-related attributes to capture domain relevant information required for fisheries monitoring. This metadata provides essential context for interpreting images/videos and supports later automated processes. Furthermore, the format can record detailed information for each individual fish, including fish length, weight, and occlusion type.

Figure 2 depicts the folder and file structure a dataset should follow.

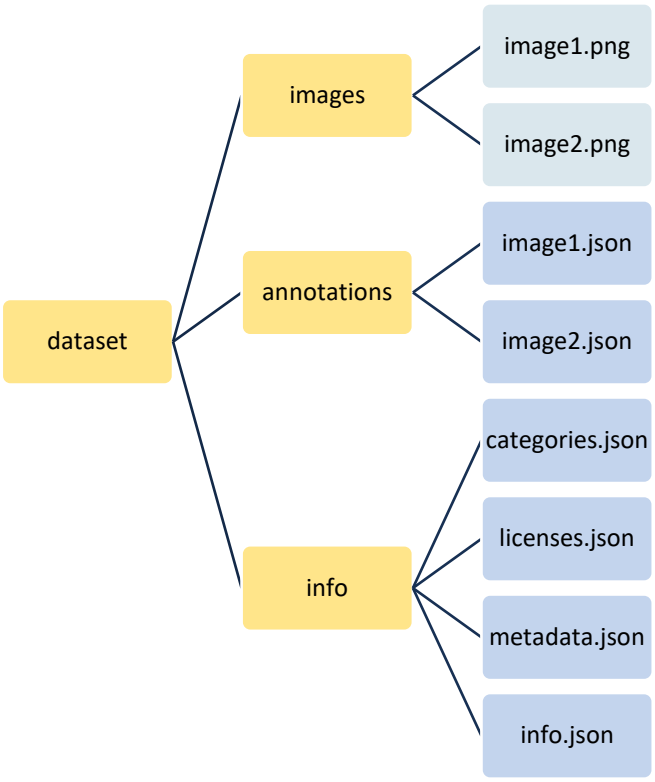


Figure 2: EVERYFISH-COCO dataset structure. Folders in yellow, files in shades of blue.

3. images folder

The *images* folder contains the images that comprise the dataset. The image format is not enforced as long as it follows open standards that don't require proprietary software; .png and .jpeg are preferred. File names must be unique within the dataset.

4. annotations folder

The *annotations* folder contains one .json file per image, with the same file name. They describe the individual images, both the metadata of the image itself (image size, date captured, etc…) as well as the annotations at object level, including the annotations and annotator type. Additionally, in the same json file, annotations at object level should be stored. A list of object annotations must be included, one per fish instance/object. For each instance, its corresponding annotation file is stored like below.

```
“image”: {
          ‘license’: 1,
          ‘file_name’: ‘image1.png’,
          ‘width’: 100,
          ‘height’: 100,
          ‘date_captured’: ‘2025-01-01’,
          ‘image_id’: 1,
          ‘annotations_type’: [‘detection’, ‘segmentation’, ‘weight_estimation’, ‘length_estimation’]
}

“annotation”: [
          {
          ‘id’: 1,
          ‘image_id’: 1,
          ‘category_id’: 1,
          ‘bbox’: [1, 1, 10, 10],
          ‘segmentation’: [[1,1,2,2,…],[2,2,3,3,…],[]…],
```

```
        'area': 10.0,
        'fish_length': 3.0,
        'fish_weight': 0.001,
        'occlusion_type': 0,
        'annotator_type': 'unknown'
        }
]
```

"image":{}		
license	int	license id from the licenses.json file
file_name	str	name of the image file, with extension.
width	int	width of the image in pixels
height	int	height of the image in pixels
date_captured	str YYYY-MM-DD	date the image was captured in ISO8601
image_id	int	assigned image_id. Unique for the dataset.
annotations_type	str	"detection", "segmentation", "weight_estimation", "length_estimation". Can contain multiple types in an array.
"annotation":{}		
id	int	object's ID. Unique per dataset
image_id	int	image_id described above
category_id	int	category_id from the categories.json file
bbox	int [x,y,width, height]	x and y coordinates from the top left image corner
segmentation	int [[x1,y1,x2,y2,…], [x1,y1,x2,y2,…]…]	list of polygons; each polygon is a list of vertices as x,y pixel positions
area	float	area of the polygon in pixels
fish_length_cm	float	estimated length of the object in cm
fish_weight_kg	float	estimated weight of the object in kg
occlusion_type	int	0 (fully visible), 1 (partially visible; partially occluded, but both ends of the fish are visible), 2 (immeasurable; at least one of the fish ends is occluded)
annotator_type	str	"human", "machine", "unknown"

5. info folder

The *info* folder contains information on dataset level: description, contributor, dataset creation date, license(s), categories used, and other discovery metadata relevant to the dataset. It enables proper evaluation of fitness-for-purpose and reuse by describing the dataset and stating the terms of use (licenses)

5.1. info.json

```
“info”: {
          ‘year’: ‘2025’,
          ‘version’: ‘1.0’,
          ‘title’: ‘Dataset title’,
          ‘description’: ‘Annotated images from Fish commercial operation at Sea in winter 2025.’,
          ‘contributor_name’: ‘Contributor McContribution’,
          ‘contributor_email’: ‘contributor@workplace.uni’,
          ‘contributor_institution’: ‘Workplace University (WPU)’,
          ‘url’: ‘https://doi.org/0000.000/dataset-000’,
          ‘date_created’: ‘2025-01-01’,
          ‘project’: ‘EVERYFISH’,
          ‘project_id’: ‘https://doi.org/10.3030/101059892’
}
```

year	str YYYY	year of the annotation
version	str	dataset's version
title	str	title of the dataset
description	str	short description of the dataset
contributor_name	str	allows for multiple, organized in an array
contributor_email	str	contributor's contact email address
contributor_institution	str	contributor's affiliation. Preferably full name and acronym in parenthesis
url	str	URL where the dataset is available
date_created	str YYYY-MM-DD	date when dataset was created, in ISO8601
project	str	project name. Optional
project_id	str	project number and funding agency or DOI. Optional

5.2. licenses.json

```
“licenses”: {
          ‘id’: 1,
          ‘name’: ‘Creative Commons Attribution 4.0 International’, OR
          ‘name’: ‘CC-BY 4.0’,
          ‘url’: ‘https://creativecommons.org/licenses/by/4.0/’
}
```

id	int	a numeric id that represents the license. Unique for the dataset
name	str	name or abbreviation of the license. For Creative Commons, see section 6.2
url	str	url of the license. For Creative Commons, see section 6.2

5.3. categories.json

```
“categories”: {
          ‘id’: 1,
          ‘name’: ‘Octopus vulgaris’,
          ‘Alpha3_code’: ‘OCC’
}
```

id	int	a numeric id that represents the species or taxa. Unique for the dataset
name	str	Latin name (species, some taxa) or human-readable name (e.g. fish, crabs, etc). If it has a corresponding alpha 3 code, use the latin name.
Alpha3_code	str	species or taxa as FAO alpha3 code ASFIS List of Species for Fishery Statistics Purposes [2] An auxiliary csv file is provided with the valid values (FAO-alpha3.csv). Must correspond with the name.

5.4. metadata.json

```
“metadata”: {
          ‘image_names’: [‘image1.png’,‘image2.png’,…],
          ‘gear_type’: ‘DIV’,
          ‘mesh_selectivity_device’: ‘0’,
          ‘targeted_fish’: ‘CEP’,
          ‘fish_region’: [‘FAO-27’, ‘ICES-9a’]
}
```

image_names	[str]	a list of images that share the same metadata. Use [":"] if all images in the dataset share the same metadata.
gear_type	str	First 3-letter code from RECO (ICES REference COdes) Metier6_FishingActivity ICES Reference Codes - RECO.[3] An auxiliary csv file is provided with the valid values (RECO.csv). See below for further clarification.
mesh_selectivity_device	str	Last element from RECO (ICES REference COdes) Metier6_FishingActivity ICES Reference Codes - RECO. An auxiliary csv file is provided with the valid values (RECO.csv). See below for further clarification.
targeted_fish	str	Second 3-letter code from RECO (ICES REference COdes) Metier6_FishingActivity ICES Reference Codes - RECO. An auxiliary csv file is provided with the valid values (RECO.csv). See below for further clarification.
fish_region	str	RFMO (Regional Fisheries Management Organization) and division acronym as *RFMO-Division*. See section 6.1. It's possible to provide a list of several values, belonging to the same or different RFMOs.

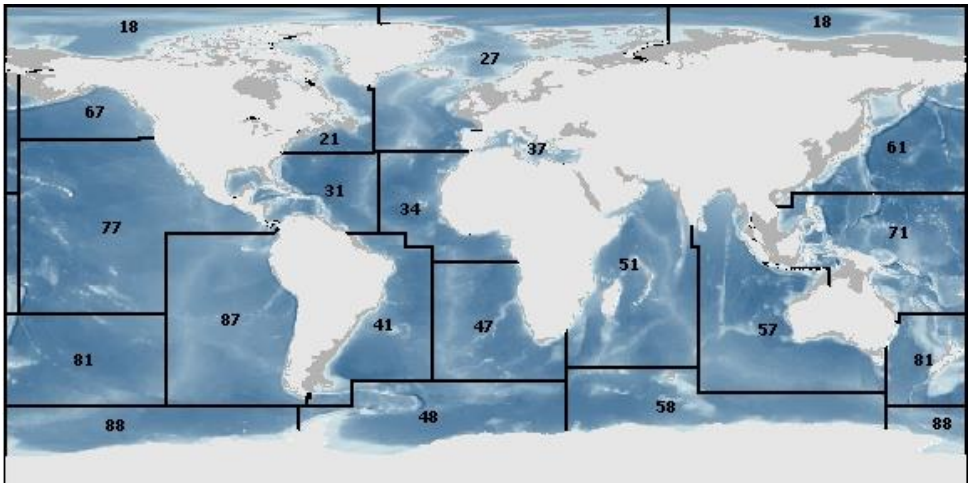
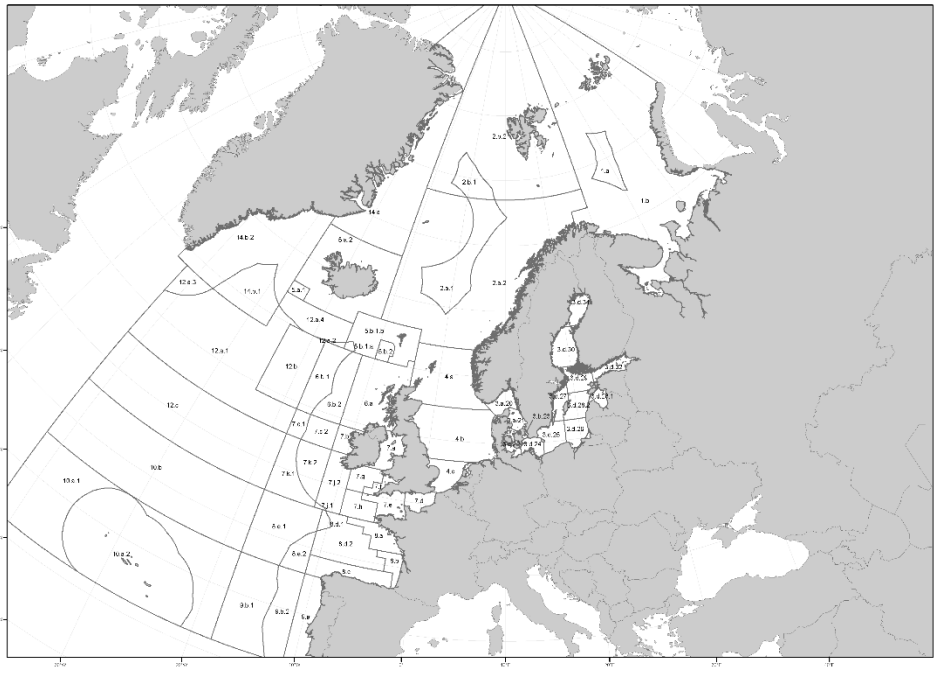
A RECO Metier6 code provides information of the gear type, target fisheries, mesh size and selectivity device (if any) in the form of a series of 3-letter codes and digits. For example, for **DIV_CEP_0_0_0**:

- gear_type: DIV (diving)
- target fishery: CEP (cephalopods)
- mesh_selectivity_device: 0_0_0:
  - 0: mesh size not applicable. Otherwise, mesh size in mm
  - 0: no selectivity device. 1 for exit window/panel, 2 for grid, 3 for square meshes.
  - 0: mesh size of the selectivity device not applicable. Otherwise, mesh size in mm

6. Auxiliary information

6.1. RFMOs and Divisions

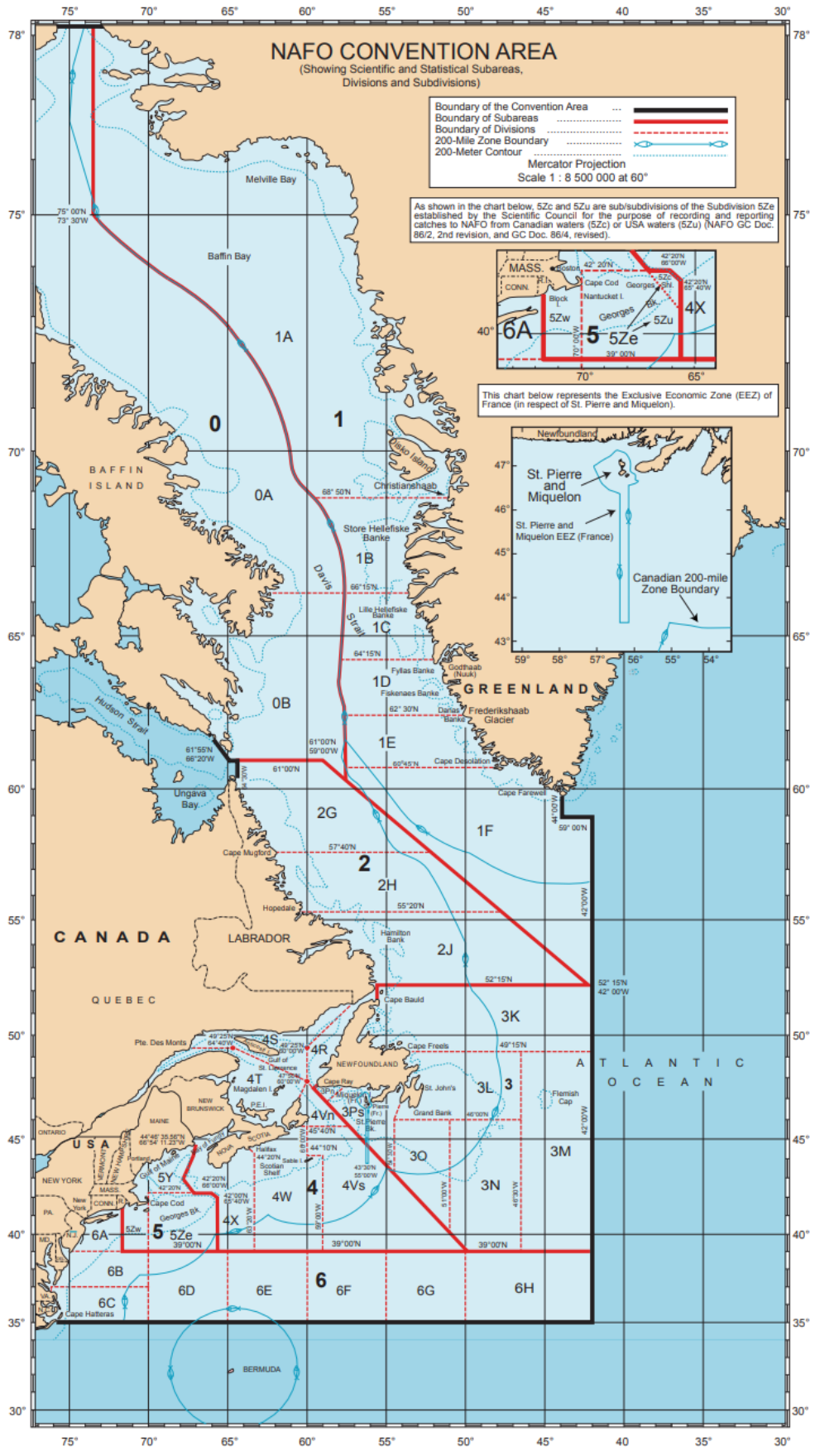
This is not an exhaustive list of the existing RFMOs, nor of their subdivision types or subareas within a type (e.g. ICES has areas, ecoregions and statistical rectangles), but a selection, common in European fisheries.

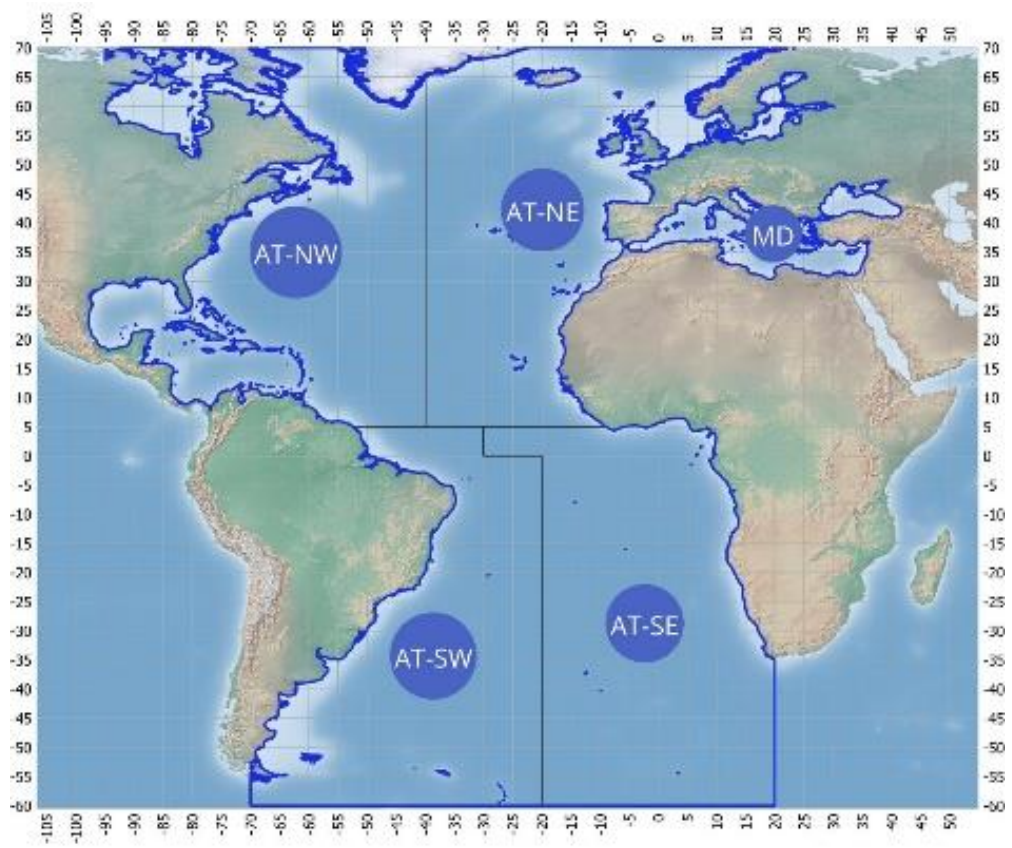
FAO Fishing Areas Fisheries and Agriculture Organization. www.fao.org (e.g. FAO-34)	
ICES International Council for the Exploration of the Seas. www.ices.dk Pre-2016 areas used roman numerals. (e.g. ICES-7.a or ICES-VII.a)	

NAFO

Northwest Atlantic Fisheries Association.
www.nato.int

(e.g. NAFO-1F)



ICCAT International Commission for the Conservation of Atlantic Tunas (e.g. ICCAT-AT-NE)	
OSPAR OSlo-PARis convention commission. www.ospar.org (e.g. OSPAR-IV)	

6.2. Licenses

This is not an exhaustive list of licenses that can be applied to datasets, and do not include licenses custom-made for institutions, projects or data publishers. It covers some of the most common data licenses and the preferred values for name and url.

name (abbreviation)	**url**
Creative Commons CC0 1.0 Universal (CC0)	https://creativecommons.org/publicdomain/zero/1.0/
Creative Commons Attribution 4.0 International (CC-BY 4.0)	https://creativecommons.org/licenses/by/4.0/
Creative Commons Attribution-ShareAlike 4.0 International (CC-SA 4.0)	http://creativecommons.org/licenses/by-sa/2.0/
Creative Commons Attribution-NonCommercial 4.0 International (CC-BY-NC 4.0)	https://creativecommons.org/licenses/by-nc/4.0/
Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC-BY-NC-SA 4.0)	https://creativecommons.org/licenses/by-nc-sa/4.0/
Creative Commons Attribution-NoDerivatives 4.0 International (CC-ND 4.0)	https://creativecommons.org/licenses/by-nd/4.0/
Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC-NC-ND 4.0)	https://creativecommons.org/licenses/by-nc-nd/4.0/

7. References

1 COCO Dataset. Data format. Retrieved 9 December 2025, from https://cocodataset.org/#format-data

2 FAO. ASFIS List of Species for Fishery Statistical Purposes. Food and Agriculture Organization of the United Nations. Retrieved 9 December 2025, from https://www.fao.org/fishery/collection/asfis/en

3 ICES REference COdes Valid Level 6 metier, . Retrieved 15 December 2025 from https://vocab.ices.dk/?ref=1647

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement number 101059892.