



(<https://www.globalseafood.org>).

---



 Fisheries

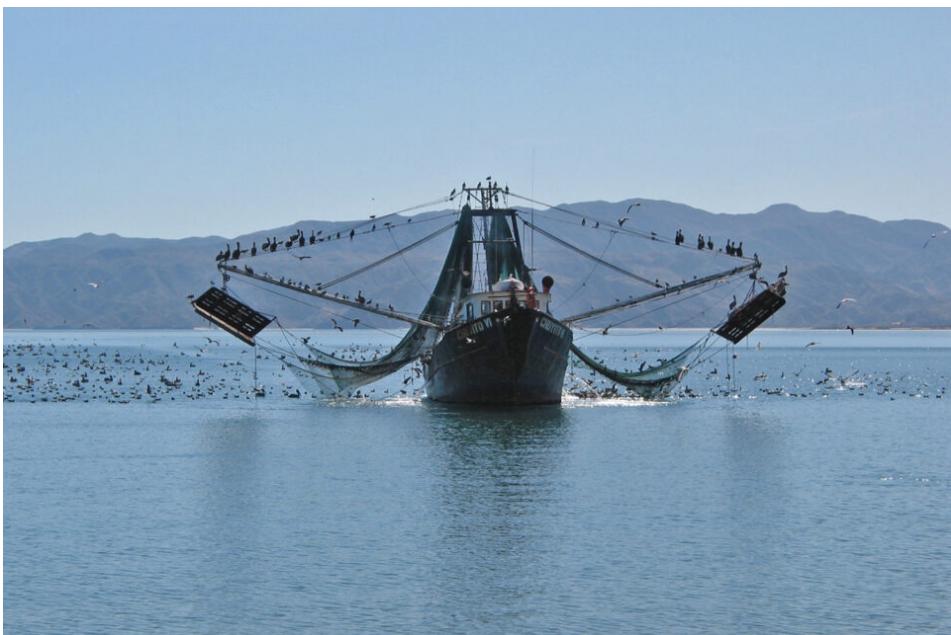
---

# Developing a harmonized global fisheries resource knowledge database

12 May 2025

By Yannis Marketadis, Yannis Tzitikas, Aureliano Gentile, Anton Ellenbroek and Marc Taconet

**A comprehensive knowledge base for aquatic resource identifiers can support the complex process of fisheries management**



Study describes the development of a comprehensive, global fisheries resource knowledge database for aquatic resource identifiers that can contribute to supporting the complex process of fisheries management. It assembled a knowledge base of aquatic resources from publicly available and known data sources, focusing on aquatic species, water areas, and fishing gear accessible online. Photo of fishing trawler in Gonsaga Bay (Baja California, Mexico) by Bengt Nyman (CC BY 3.0 <https://creativecommons.org/licenses/by/3.0/>, via Wikimedia Commons).

Fisheries management plays a crucial role in **preventing overfishing** (<https://doi.org/10.1126/science.1098222>) by regulating the harvesting and utilization of fish stocks to ensure their sustainable exploitation while preserving the marine ecosystem. It involves a range of activities aimed at maintaining the balance between the extraction of fishery resources and the conservation of aquatic ecosystems. One of the key components of efficient fisheries management is the so-called stock assessment. This involves monitoring fish populations to determine their abundance, distribution, and health. It is, therefore, of crucial importance that the description of relevant information (e.g., species) is accurate and complete.

There are many different ways of referring to a species, including a variety of common names that vary depending on the region. The scientific community uses the so-called scientific name: the genus and the species. In addition, there are alpha-numeric identifiers used that are provided from different data sources or registries and are widely used, particularly for data exchange and interoperable mechanisms. The problem is that there is no common guideline adopted by the existing fisheries management authorities and other institutions spread around the world. So, practically all of them are being used nowadays, and there are cases in which it can become rather cumbersome to analyze fishery reports produced by different authorities.

This article – [summarized](https://doi.org/10.3390/proceedings2025117004) (<https://doi.org/10.3390/proceedings2025117004>) from the [original publication](https://creativecommons.org/licenses/by/4.0/) (<https://creativecommons.org/licenses/by/4.0/>) (Marketadis, Y. et al. 2025. Building a Global Aquatic Resource Knowledge Base for Fisheries. *Proceedings* 2025, 117(1), 4) – reports on a study to implement a process that relies on the semantic web by collecting information from different data sources and constructing a single knowledge base with key species taxonomic information. By using an appropriate ontology, as the conceptual model, we managed to semantically integrate information coming from different data sources and describe them in a homogeneous manner. This process managed to interconnect the identifiers of the same species to support the provision of complementary information, which would not be possible without semantically integrating them.

Semantic data integration involves the harmonization of heterogeneous data sources by understanding the underlying semantics, relationships, and meanings within the data. It goes beyond syntactic matching to interpret the semantics of data elements, resolving the semantic heterogeneity that arises from differences in terminologies and concepts across sources. This process employs [ontologies](https://link.springer.com/chapter/10.1007/978-3-319-61893-7_11) ([https://link.springer.com/chapter/10.1007/978-3-319-61893-7\\_11](https://link.springer.com/chapter/10.1007/978-3-319-61893-7_11)), vocabularies and definitions of schema mappings to establish common semantic interpretations across unrelated datasets. We used the marine domain ontology [MarineTLO](http://dx.doi.org/10.1007/978-3-319-03437-9_29) ([http://dx.doi.org/10.1007/978-3-319-03437-9\\_29](http://dx.doi.org/10.1007/978-3-319-03437-9_29)) because it contains all the necessary information.



(<https://link.ctbl.com/aquapod>).

## Marine species data sources

For the construction of the knowledge base for marine species, the following well-known and actively used data sources were used:

1. [FAO ASFIS List of Species for Fishery Statistics Purposes](https://www.fao.org/fishery/en/collection/asfis) (<https://www.fao.org/fishery/en/collection/asfis>) provides a code list of marine species with several identifiers, such as 3-alpha code, taxonomic code, and ISSCAAP code [International Standard Statistical Classification of Aquatic Animals and Plants (ISSCAAP)]. The 3-alpha codes are made of three characters that uniquely identify the species, complemented with scientific names, taxonomic details, and common names in several languages.
2. [World Register of Marine Species](https://www.marinespecies.org/) (<https://www.marinespecies.org/>) (WoRMS) is an authoritative database that provides a comprehensive and up-to-date inventory of all known species globally. The main identifier in WoRMS is AphiaID which is a numeric code. Moreover, it contains taxonomic information, synonyms, distribution maps, and bibliographic references for each species listed in the database.
3. [FishBase](https://www.fishbase.se/) (<https://www.fishbase.se/>) is a global biodiversity information system on fishes that provides detailed information about species regarding taxonomy, morphology, ecology, distribution, behavior, and fisheries-related data. It is maintained and continuously updated by an international consortium of scientists, with support from various organizations.
4. [Integrated Taxonomic Information System](https://www.itis.gov/) (<https://www.itis.gov/>) (ITIS) is an authoritative database that provides taxonomic information including the Taxonomic Serial Number (TSN), scientific names, and taxonomic hierarchies. The database is reviewed and updated periodically.

to ensure high quality with valid classifications, revisions, and additions on newly described species.

## New database revealing salmon patterns at sea aims to curb IUU fishing



A new analysis of where salmon are concentrated at sea could aid with fisheries management and help curb IUU fishing.



Global Seafood Alliance

## Aquatic species knowledge base construction process

How information was collected from the different data sources is depicted in Fig.1 As illustrated, the resources collected from each source appeared in different formats, so it was necessary to proceed with a syntax normalization phase before actually transforming them into instances of the top-level ontology MarineTLO, and so the contents from the different data sources were gradually integrated into the knowledge base.

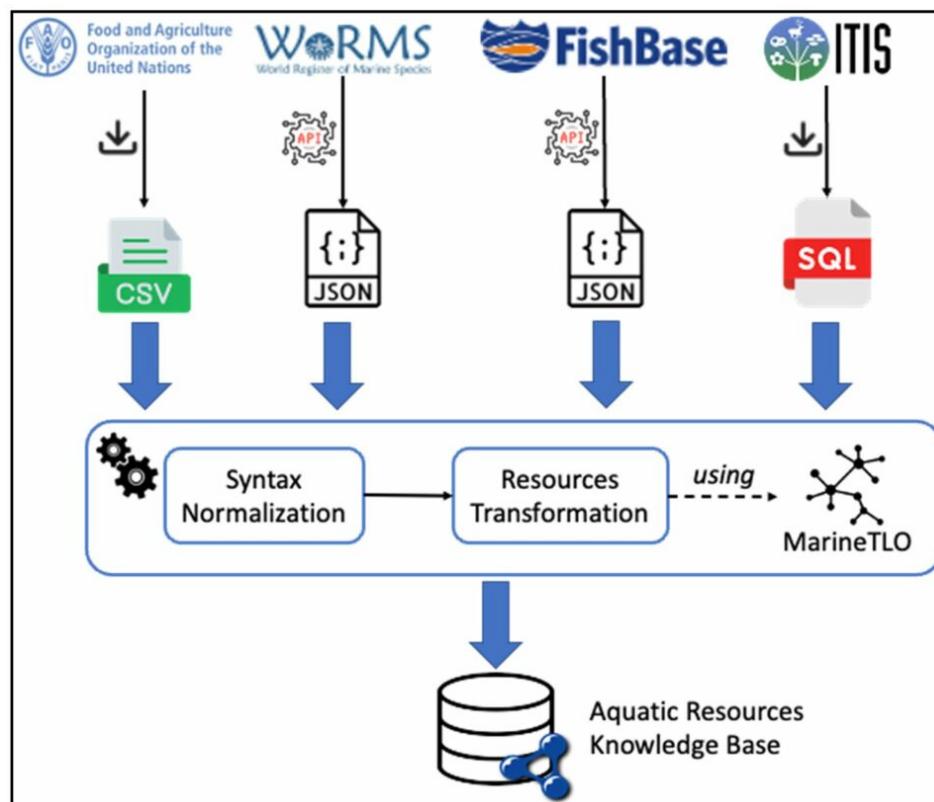


Fig. 1. The process of collecting species resources from external data sources, preparing, adapting, and ingesting them into the knowledge base.

During the construction of the knowledge base about species, we encountered several mismatches regarding the scientific names of species. To resolve those issues, we used FishBase as the source of truth, which contains a list of synonyms for each species describing if a synonym is valid or not. For the few cases that the issue could not be resolved that way, the corresponding data source owners were informed about them. After communicating with them, some of those issues were resolved.

Fig. 2 describes in a diagrammatic manner how the information is stored in the knowledge base by relying on the proper classes of MarineTLO. In this example, we report all the information that refers to the sample species used, the red mullet, with the binomial "*Mullus barbatus*" (the scientific name); the left group of information reports the different identifiers, the one in the middle reports some of the names, and the right one reports the taxonomic information of the species. Note that all the species-related information is presented in a uniform manner, although it has been collected from different data sources.

Fig. 2. Detailed information (i.e., identifiers, scientific names, common names in English, UK; Italian, IT; Greek, GR; and Chinese, and taxonomic information) about aquatic species in the knowledge base.

## Aquatic resource knowledge base

Fisheries management is not just about aquatic species, but also includes more essential information, like fishing assessment or management of water areas and fishing gear. So, we applied the same methodology for constructing knowledge bases about those entities as well. With regard to areas, we mainly used [FAO major fishing areas](https://www.fao.org/fishery/en/area/search) (<https://www.fao.org/fishery/en/area/search>) for statistical purposes as well as [Large Marine Ecosystems](https://doi.org/10.1577/1548-8446(1999)024%3C0015:LME%3E2.0.CO;2) ([https://doi.org/10.1577/1548-8446\(1999\)024%3C0015:LME%3E2.0.CO;2](https://doi.org/10.1577/1548-8446(1999)024%3C0015:LME%3E2.0.CO;2)) (LME), [Marine Regions](https://www.marineregions.org/) (<https://www.marineregions.org/>) and several national jurisdiction areas and pertinent ISO codes. For

fishery gear, we relied on the different versions of the **FAO International Standard Statistical Classification of Fishing Gear** (<https://www.fao.org/cwp-on-fishery-statistics/handbook/capture-fisheries-statistics/fishing-gear-classification/en/>) (ISSCFG) standard.

Overall, we implemented a process that built a knowledge base of aquatic resources from publicly available, well-known data sources, focusing on aquatic species, water areas, and fishing gear. The knowledge base can be browsed either through a SPARQL endpoint or through a **dedicated web application** (<https://isl.ics.forth.gr/grsf/grsf-ir>). It contains information about 40,564 marine species, 3316 water areas, and 88 fishing gear resources.

The aquatic resource knowledge base is actively used for monitoring the status of fisheries. In particular, it is used by the **Global Record of Stocks and Fisheries** ([https://doi.org/10.1007/978-3-030-71903-6\\_26](https://doi.org/10.1007/978-3-030-71903-6_26)) for harmonizing and updating, if necessary, the information on the stocks and fisheries it collects.

## Conclusions

This article describes the process for the construction of a comprehensive, standardized knowledge base for aquatic resource identifiers to support the complex process of fisheries management. Although we focused on semantically integrating marine species, water areas, and fishing gears, our methodology and tools can be applied to other entities or resources as well (e.g., identification of countries, fishery management authorities, and others).

The knowledge base can also be used as a reference for harmonizing fishery management resources, and we also report its current usage on this. Potential extensions include the addition of more data sources and entities, as well as the automation of the construction process so that the knowledge base remains in sync with the contents of the data sources that are used.

## Authors

---



### YANNIS MARKETADIS

Corresponding author

Institute of Computer Science, Foundation of Research and Technology-Hellas, 71110 Heraklion, Greece

[marketak@ics.forth.gr](mailto:marketak@ics.forth.gr) (<mailto:marketak@ics.forth.gr>).



### YANNIS TZITIKAS

Institute of Computer Science, Foundation of Research and Technology-Hellas, 71110 Heraklion, Greece



**AURELIANO GENTILE**

Food and Agriculture Organization of the United Nations, 00153 Rome, Italy



**ANTON ELLENBROEK**

Food and Agriculture Organization of the United Nations, 00153 Rome, Italy



**MARC TACONET**

Food and Agriculture Organization of the United Nations, 00153 Rome, Italy

Copyright © 2025 Global Seafood Alliance

All rights reserved.