# Marine Regions: An interoperable standard for georeferenced marine place names

Salvador J. Fernández Bejarano (salvador.fernandez@vliz.be), Britt Lonneville (britt.lonneville@vliz.be), Lawrence Whatle (lawrence.whatley@vliz.be), Bart Vanhoorne (bart.vanhoorne@vliz.be), Lennert Schepers (lennert.schepers@vliz.be)

VLIZ, Flanders Marine Institute (Belgium)

Marine Regions is a global geographic database containing (i) a gazetteer of marine georeferenced place names and (ii) maritime boundaries. The Marine Regions Gazetteer includes over 60,000 places integrating national, regional as well as global marine gazetteers. In addition, the Marine Regions Gazetteer integrates the maritime boundaries dataset that has currently its 12th release [Flanders Marine Institute, 2023]. This is a digital representation of the administrative areas defined by the United Nations Convention on the Law of the Sea of 1982 such as Exclusive Economic Zones (EEZs) or Areas Beyond National Jurisdiction. The Marine Regions dataset is accessible through a web portal (marineregions.org) and an API. Machine-to-Machine interactions via its API allow integration into larger knowledge networks such as the World Registed of Marine Species (WoRMS), Global Fishing Watch and MarineTraffic. Here, we will discuss the Marine Regions back-end, the new Linked Data Model and two improved access methods based on the linked data model - Linked Data Event Streams and a new R client. Finally, as an endorsed Action of the UN Ocean Decade under the OceanData2030 programme, we explain how we want to expand our gazetteer and impact by involving various existing and new stakeholders.

**The Marine Regions back-end: structure and rationale**

The Marine Regions gazetteer is composed of a relational database in which geographical entities are characterized by their unique resource identifier (URI), the Marine Regions Geographic Identifier (MRGID). This identifier allows users to unambiguously refer to a Marine Regions entity and is persistent and resolvable at <http://marineregions.org/mrgid/{id}>. Geographical entities are defined both by a geometry and a place type. While the spatial features are represented as different vector data types, for example, point, multipoint, line, or polygon, a place type provides contextual information on the geographic entities, for example, a sea, bay, ridge, sandbank, or oceanic trench. Not only physical place types are considered but also administrative place types, such as countries, EEZs, territorial seas, fishing zones, or marine protected areas. In addition, one or more names can be stored for every gazetteer entity, allowing us to deal with synonyms or names in different languages. The gazetteer also provides a hierarchy between the different entities, based on a parent-child relation (partOf). Other relations to Marine Regions entities can be described in the Marine Regions gazetteer, these can cover topological (adjacentTo, streamsThrough, ... ) and non-topological (administrativePartOf, influencedBy, ... ) links.

Until recently, these data were accessible through the Marine Regions website, through OGC web services and through a JSON HTTP API. In order to achieve semantic interoperability, we aligned the Web APIs using a common Linked Data model [Lonneville B. et al., 2021].

**A new Marine Regions Linked Data model**

In this section, we explain the core aspects of the Linked Data Model and elucidate certain encountered challenges. Gazetteer entities are naturally both a description of a place and the location of the place itself. Consequently, each geographical entity becomes both a

<skos:Concept> and a <dct:Location>. The whole Marine Regions dataset is a SKOS concept scheme with the geospatial extent defined as instances of the OGC's SimpleFeatures vocabulary <sf:geometry>. The relationships offered by traditional SKOS predicates such as <skos:narrower> cannot accurately capture the hierarchical nature of the Marine Regions Gazetteer: consider one gazetteer entity that may be spatially contained by another, but only partially overlaps with a third one. To address these relations, we defined <mr:isRelatedTo> along with nine subproperties (e.g., <mr:isPartOf>) and their respective inverses (e.g., <mr:contains>). They are detailed in the Marine Regions ontology at <http://marineregions.org/ns/ontology>.
The documents describing each URI (e.g., <http://marineregions.org/mrgid/5686.ttl>) are currently available through content negotiation in two RDF serializations, Turtle and JSON-LD, together with existing JSON, XML and HTML options. Nevertheless the model presented a performace challenge: How to efficiently handle documents containing large data, including those with detailed geometries or with numerous relationships? Our solution is to provide geometries with an additional link within the document, thereby enabling data consumers to retrieve geometries on demand. In cases where the number of relationships is extensive, the document incorporates the <hydra:next> predicate, directing users to subsequent pages that each contain a maximum of 40 relations.

**Retrieving Marine Regions using Linked Data Event Streams**

The Marine Regions semantic model provides the basis for replicating the database via Linked Data Event Stream (LDES). It allows data consumers to capture changed entities as immutable objects with unique URIs [Van Lancker et al. 2021]. The pagination strategy is crucial in LDES. Here we opted for fragmenting using <dc:modified> timestamps, allowing to retrieve the full change history from one single feed endpoint at <http://marineregions.org/feed>. This method simplifies the process of keeping up with updates, enabling users to set up services like (Geo)SPARQL endpoints or document stores for full-text searches. The LDES, with its layered architecture, enhances accessibility and flexibility, making it easier for third parties to integrate and utilize the Marine Regions data in various applications.

**Bringing the Marine Regions API into an R client: mregions2**

As an example of the possibilities of this semantic API we developed mregions2, a client written in the R programming language. mregions2 offers a streamlined interface to access data from Marine Regions in R for researchers, marine scientists, and geospatial analysts seeking marine geographical information. It follows the same rationale behind the API as calls to the Marine Regions gazetteer do not provide geometries or relationships by default. Instead, these can be accessed on demand as the data consumer may need. mregions2 has been peer-reviewed by rOpenSci and will be published on CRAN, ensuring that Marine Regions data is available programatically for R users. Installation instructions are available at:
https://github.com/lifewatch/mregions2

**References**

Flanders Marine Institute, (2023). *Maritime Boundaries Geodatabase, version 12*. Available online at https://www.marineregions.org/ - https://doi.org/10.14284/628

Lonneville B. et al., (2021). *Publishing the Marine Regions Gazetteer as a Linked Data Event Stream*. S4BioDiv 2021. http://ceur-ws.org/Vol-2969/paper8-s4biodiv.pdf

Van Lancker D. et al., (2021). *Publishing base registries as linked data event streams*, in: Proceedings of the 21th International Conference on Web Engineering.