- ORIGINAL ARTICLE -

ODP-DASHBOARD: Enhancing Marine Species Conservation in the South Atlantic through Linked Open Data Integration

ODP-DASHBOARD: Mejorar la conservación de especies marinas en el Atlántico Sur mediante la integración de datos abiertos vinculados

Gustavo Nuñez[1,2], Carlos Buckle[2], and Marcos Zárate[1,2]

[1]*Centro para el Estudio de Sistemas Marinos, Centro Nacional Patagónico (CESIMAR-CENPAT-CONICET), Puerto Madryn, ARG.*

[2]*Laboratorio de Investigaciones en Informática (LINVI-UNPSJB), Puerto Madryn, ARG.*

gnunez@ing.unp.edu.ar

Abstract

This paper presents the development of a Cloud-based dashboard called ODP-DASHBOARD (Ocean Data Platform), to manage information from the Marine Biodiversity domains and Oceanography. In this context, the Linked Open Data (LOD) paradigm has emerged as a set of best practices for connecting, sharing and exposing data and knowledge. A central part of this paradigm are ontologies, which allow the definition of shared vocabularies and conceptual models that help integrate this information. These considerations provide a strong motivation to formulate a system that takes into account geospatial characteristics that can provide answers to questions such as the following: (i) *How can we define spatial regions for our studies?* (ii) *How are species distributed in a given region?* (iii) *Given a particular georeference, what geographic region does it belong to?* (iv) *How to relate species occurrences to environmental variables within a specific region?*. This system is composed of (i) Architecture; (ii) Conceptual model and (iii) Linked data set for exploitation through a SPARQL endpoint. The evaluation was carried out from two approaches, the first of which consists of validating the dashboard using real data extracted from Marine Biodiversity and Oceanography repositories and then validating the proposed conceptual model using competency questions. The second approach has to do with validation through case studies defined together with researchers from the Scientific and Technological Center (CENPAT-CONICET).

Keywords: Ocean sensors, SSN/SOSA ontology, RDFLib, Aqualink.

Resumen

En este artículo se presenta el desarrollo de un tablero de mando basado en la nube llamado ODP-DASHBOARD (Ocean Data Platform), para gestionar información de los dominios de Biodiversidad Marina y Oceanografía. En este contexto, el paradigma Linked Open Data (LOD) ha surgido como un conjunto de mejores prácticas para conectar, compartir y exponer datos y conocimiento. Una parte central de este paradigma son las ontologías, que permiten la definición de vocabularios compartidos y modelos conceptuales que ayudan a integrar esta información. Estas consideraciones proporcionan una fuerte motivación para formular un sistema que tenga en cuenta las características geoespaciales que puedan proporcionar respuestas a preguntas como las siguientes: *(i) ¿Cómo podemos definir regiones espaciales para nuestros estudios?* (ii) *¿Cómo se distribuyen las especies en una región determinada?* (iii) *Dada una georeferencia particular, ¿a qué región geográfica pertenece?* (iv) *¿Cómo relacionar las ocurrencias de especies con las variables ambientales dentro de una región específica?*. Este sistema se compone de (i) Arquitectura; (ii) Modelo conceptual y (iii) Conjunto de datos enlazados para su explotación a través de un endpoint SPARQL. La evaluación se realizó desde dos enfoques, el primero de ellos consiste en validar el tablero de mando utilizando datos reales extraídos de repositorios de Biodiversidad Marina y Oceanografía y luego validar el modelo conceptual propuesto mediante preguntas de competencias. El segundo enfoque tiene que ver con la validación a través de estudios de casos definidos en conjunto con investigadores del Centro Científico y Tecnológico (CENPAT-CONICET).

Palabras claves: Sensores Oceánicos, Ontología SSN/SOSA, RDFLib, Aqualink.

1 Introduction

The Southwestern Atlantic Ocean, especially its subpolar and polar coastal areas, remains relatively understudied compared to other regions. Despite some scientific research, much of it has been limited due to the remote and harsh environmental conditions of these areas, such as harsh weather, strong winds, and cold temperatures, making extended research challenging [1]. There is an urgent need to increase data acquisition efforts in these regions to better understand the effects of climate change on marine ecosystems. Studies combining data loggers and satellite remote sensors on marine vertebrates have significantly enhanced our understanding of the ocean, revealing intricate associations between the physical environment, productivity, and foraging ecology of key migratory predators [2]. Various species like giant bluefin tuna, penguins, and marine mammals have been used to explore large marine ecosystems firsthand, providing insights ranging from depth temperature profiles to oceanographic features characterization [3, 4]. Foraging Southern Elephant Seals (SES) are particularly noteworthy for their deep diving and long-distance travel, covering thousands of kilometers across various oceanographic regions for extended periods [5]. The research group called Grupo de Estudios de Macroecología Marina (GEMMA-CONICET) focuses on SES as a model for biological, oceanographic, and ecological studies [6]. To address these challenges, a LOD-based dashboard has been developed. It aims to manage information in marine biodiversity and oceanography domains, featuring an architecture, conceptual model, and linked dataset for exploitation through a SPARQL endpoint. Validation of the ODP-DASHBOARD involved real data from marine biodiversity and oceanography repositories, along with competency questions. Additionally, case studies in collaboration with researchers from the Scientific and Technological Center (CENPAT-CONICET) were conducted to validate the proposed conceptual model. Documentation for ODP-DASHBOARD is available online at `https://github.com/gustavomarcelonunez/odp-dashboard`, and the linked dataset is accessible via DOI 10.17632/5nv5c7575w.3. The rest of the paper is organized as follows: In Section 2, we present an overview of related related vocabularies and applications that were the input to ODP-DASHBOARD. In Section 3, we describe the system, which is composed of architecture, data transformation process and integrated dataset. Section 4, discusses how ODP-DASHBOARD was evaluated following a specific protocol, and presents the use case. Finally, in Sections 5 we discuss the results and present our final considerations.

2 Related vocabularies and applications

This section shows the underlying vocabularies and ontologies that were used in the development and later we will briefly mention the applications that offer similar features to ODP-DASHBOARD.

2.1 Vocabularies and ontologies

The Natural Environment Research Council (NERC) Vocabulary Server [7] offers standardized term lists spanning various disciplines relevant to the oceanographic and broader communities. Utilizing controlled vocabularies in metadata and data tags resolves ambiguity issues in data markup, enabling computer interpretation and facilitating computer-assisted management, distribution, and long-term reuse of datasets. Geolink [8] introduces an Ontology Design Pattern (ODP) for oceanographic cruises using OWL, which combines and reuses existing patterns such as trajectory, event, and information object, making it suitable for defining relationships and classes in datasets.

Two relevant national initiatives include the National Biological Data System (SNDB) developed by the Ministry of Science, Technology, and Productive Innovation, and the National Sea Data System (SNDM) developed by the International Oceanographic Data and Information Exchange (IODE) for marine data visualization. However, these portals have ceased due to funding issues [9, 10].

The Semantic Sensor Network Ontology (SSN/SOSA) [11] has gained prominence in marine sciences, providing a standardized framework for describing sensors, observations, and metadata. SSN/SOSA facilitates seamless data integration and interoperability across various sources, allowing machine-readable descriptions of sensors' capabilities and observations' characteristics. In ocean sciences, SSN/SOSA has been applied to oceanographic monitoring, marine biology, and climate science, enabling better understanding of oceanographic processes by integrating data from various sensors like ocean gliders, buoys, and satellites.

2.2 Applications

According to our literature review, we did not find specific Southern Elephant Seal (SES) data repositories or data management systems that meet Linked Open Data (LOD) criteria [12]. We explored the Marine Mammals Exploring the Oceans Pole to Pole (MEOP) initiative and Burton et al. (2017) [13, 14], but their data does not comply with LOD principles. Movebank, managed by the Movebank Data Repository, allows sharing and analyzing animal movement data but also does not adhere to LOD principles [15]. Regarding LOD-based systems, the OpenBiodiv Knowledge Graph and Ozymandias Knowledge Graph are notable, while Biologer lacks LOD principles [16, 17, 18, 19]. We discarded Biologer due to its lack of LOD principles, which are essential for interoperability and compliance with FAIR principles [20]. However, Ozymandias and OpenBiodiv Knowledge Graphs do not provide visual components like maps.

3 System description

The main objective of our dashboard development is to provide satisfactory visual analysis to users, ensuring data reusability and interpretation by other applications. To achieve this, compliance with the FAIR principles [20], at least partially, is essential. Initially, data in CSV format extracted from a MySQL relational database served as the data source. However, due to semantic and reuse limitations, we opted to convert these data to Linked Open Data (LOD) (see section 3.2 for the data transformation process). Subsequently, the links obtained from the conversion were integrated into the dashboard, enabling redirection to the machine-readable version of each displayed data.

3.1 Architecture

Our application architecture follows the client/server model [21], where the client makes requests processed by the server to provide responses. The server side is developed using Shiny [22], an R package enabling the creation of interactive web applications. Shiny applications consist of two components: the front-end (ui) responsible for building the user interface displayed to users, and the back-end (server) generating application content, such as functions or data management and analysis.

3.2 Data transformation process

To perform the data transformation, the D2RQ Platform is used ([23]). This consists of: 1) D2RQ Mapping Language, which is used to write mappings between database tables and RDF vocabularies or OWL ontologies ([24]), 2) D2RQ Engine, a SPARQL-to-SQL rewriter that can evaluate SPARQL queries over your mapped database, and 3) D2R Server, a web application that provides access to the database via the SPARQL protocol, as Linked Data, and via a simple HTML interface. D2RQ runs at http://linkeddata.cenpat-conicet.gob.ar in back-end to browse structured data. The SES dashboard dataset can be downloaded from 10.17632/5nv5c7575w.3, navigated and queried using a SPARQL endpoint http://linkeddata.cenpat-conicet.gob.ar/snorql/, and it is published under Creative Commons Universal Public Domain Dedication (CC0 1.0) License.

3.3 Dataset

LOD [25], is an idea that arises from the Semantic Web [26], with the goal of ensuring that data published on the Web is reusable, discoverable, and, most importantly, that data published by different entities can work together. The LOD principles are summarized as follows:

- Use URIs as resource names. Each instance is uniquely identifiable through an HTTP URI. For example, the result of a measurement of the average depth of the water column measured by an instrument can be identified as http://linkeddata.cenpat-conicet.gob.ar/data/result/id-233/avg depth. All instance identifiers follow this same scheme.

- Leverage the HTTP of the URI so that people can locate and query (i.e., dereference) these resources. In accordance with Linked Data principles, dereferenceable HTTP URIs were used for our resources. For example, for the URI of the average depth shown above, the human-readable version of the URI is: http://linkeddata.cenpat-conicet.gob.ar/page/result/id-233/avg depth which was defined to dereference the previous URI.

- When a URI is dereferenced, it offers valuable information about the resource. Dataset resources form a knowledge graph and are linked to external databases using properties such as `skos:exactMatch`, `owl:sameAs` and `skos:broader`, and . These identifiers include ORCID, Wikidata entities, DBPedia resources, and NERC vocabulary server ID. For example, average water temperature is associated with the NERC identifier `SDN:P02::TEMP` through the `skos:broader` property. (See: observable-property/id-233/avg-temp).

- Include links to other URIs related to the data contained in the resource, thus enhancing information discovery on the Web. Users can easily explore additional relevant information and resources, thus enriching their understanding and enabling deeper insights into the topic.

3.4 ODP-DASHBOARD overview

The ODP-DASHBOARD platform was created to query statistical data from marine species dives, travel routes, and oceanic data from a buoy. We used the flexdashboard package in R to generate web pages via R Markdown. SPARQLWrapper package facilitated querying different endpoints by importing SPARQL SELECT query results into R as a dataframe from Python code blocks. Subsequent sections detail each module of the dashboard.

3.4.1 Dive Statistics

This module captures diving data by programming instruments to sample at intervals. After retrieval, the data undergoes preprocessing to detect anomalies and is stored in a database or CSV files. Users can analyze dives individually or apply basic statistics. It summarizes diving statistics for all ODP individuals, including maximum depths, dive counts, temperatures, and platforms. Detailed sensor information is provided, with a summary at the top, a sensor table at the bottom left, and a dive count bar chart at the bottom right. The ggplot library was used for bar charts. Platforms and sensors are linked to LOD descriptions in both HTML and RDF formats. See Figures 1 and 2 for illustrations.

3.4.2 Dive Analysis

The diving analysis module is divided into two parts: on the left, it displays the variables of interest recorded by the platform during the dives. Temperatures, depths, and duration can be shown. On the right, it shows the routes taken by the selected platform. Figure 3 shows the analysis report screen of dives performed by the ODP encoded as ANIF, with different variables captured by the instruments (e.g., maximum depth, duration, surface temperature, etc.) plotted against the number of dives on the X-axis. The line chart on this screen was built using the Plotly library [27]. On the other hand, the map showing the route taken by the platform was built using the leaflet library [28].

3.4.3 Buoy Data

This module of the application utilizes SPARQL queries to retrieve data from the oceanographic platform [29]. A plugin on the left sidebar allows users to select a date range, with data plotted in the main section of the tab. Below this, a plugin developed using the leaflet library [28] displays the geographical location of the data-collecting platform.

Figure 4 in the main section presents information captured by the platform within the selected date range. Data includes surface temperature, depth temperature, and the average of these two calculated from R functions. The center chart displays data on significant wave heights in the surveyed area, while the bottom chart shows wind speed data.

3.4.4 Cluster Analysis

This module uses the DBSCAN package for spatial cluster analysis, employing Density-Based Spatial Clustering of Applications with Noise. Parameters chosen include a neighborhood radius (ε) of 0.15 and a minimum number of points (MinPts) of 12. The algorithm forms clusters based on neighboring points within the radius, recursively expanding them if they meet the MinPts requirement. Points with fewer neighbors are classified as noise. Figure 5 illustrates this geospatial analysis.

3.4.5 Census Statistics

This module visualizes basic statistics of primary data collected during censuses spanning 1990 to 2017. The census records only visually countable individuals, without individual identification of ODPs. Land censuses occur during the breeding season peak (first week of October), when approximately 96% of breeding females are ashore [30]. Foot censuses cover nearly 200 km of open coast in Peninsula Valdés (PV) and approximately 150 km outside PV. However, censuses may vary in date and sectors studied across years. A histogram, created using the ggplot module, illustrates annual elephant seal counts during the breeding season peak, categorized by age (see Figure 6).

3.4.6 Linked Data

In this section of the application, the sidebar presents a list of marine species obtained through SPARQL queries to the endpoint called *BiGe-Onto* [31]. The user can select a species, triggering a series of queries to retrieve detailed information. The first table displays the Wikidata item, scientific name, common name, status, habitat region image, and, if available, data such as length and life expectancy. The second table (bottom left section) provides bibliographic links to various linked open data databases. The third tab (bottom right section) displays relevant information about species attributes and their associated values. (See Figure 7)

3.4.7 More

This module consists of two tabs. The first tab is dedicated to maintaining an updated record of scientific production concerning southern elephant seals by CESIMAR-CONICET researchers. It features a table listing all publications, including scientific articles, book chapters, or reports. Users can search using various parameters such as author, title or abstract text, keywords, DOI, year, etc. Figure 8 illustrates these options. It's crucial to monitor which ODP individuals are referenced in a publication, either for data analysis or map representation. The ODP Dashboard facilitates linking ODPs to their associated publications. The second tab provides information about the application.

4 Evaluation

To evaluate the underlying data model of *ODP-DASHBOARD*, data-driven validation [32] was employed to accurately represent real-world scenarios and validate the developed system. Specifically, data from oceanographic buoys provided by the Aqualink project were used to create relevant examples and relationships.

To verify the proposed model, various datasets extracted from Aqualink were instantiated, including instances of its concepts and relationships. The detailed instantiation process is described in Section 3.2. Figure 9 illustrates a record from the CSV file in the dataset titled "Isla Leones, Argentina" which contains data from an oceanographic buoy. This buoy records surface temperature, bottom temperature, wind speed,

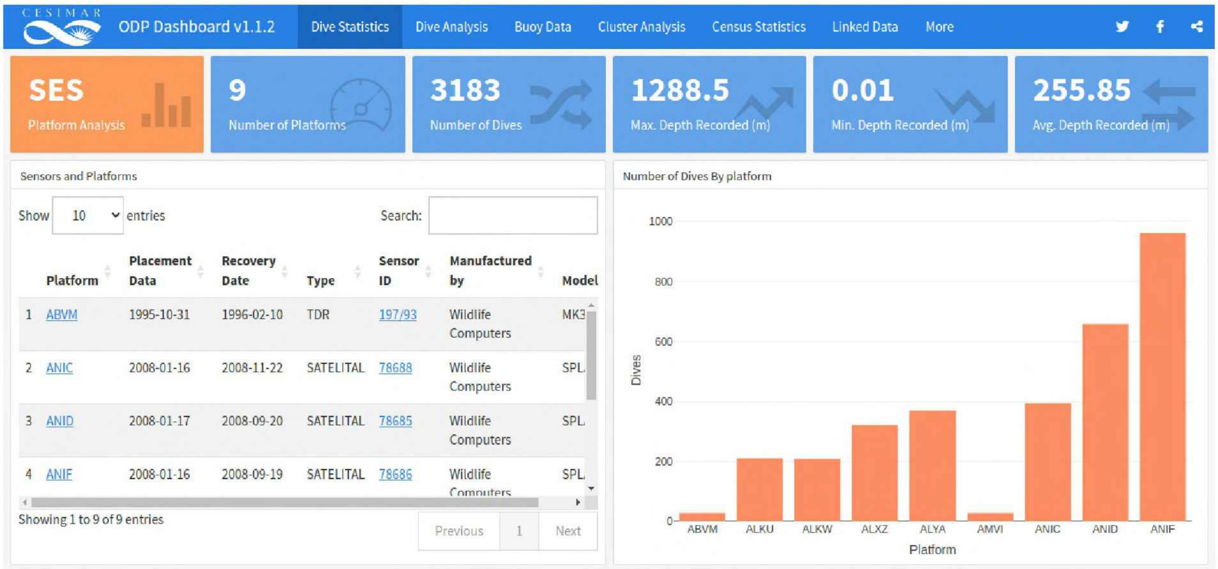

Fig. 1. Main screen of *ODP-DASHBOARD* featuring a view of diving data reports and censuses.

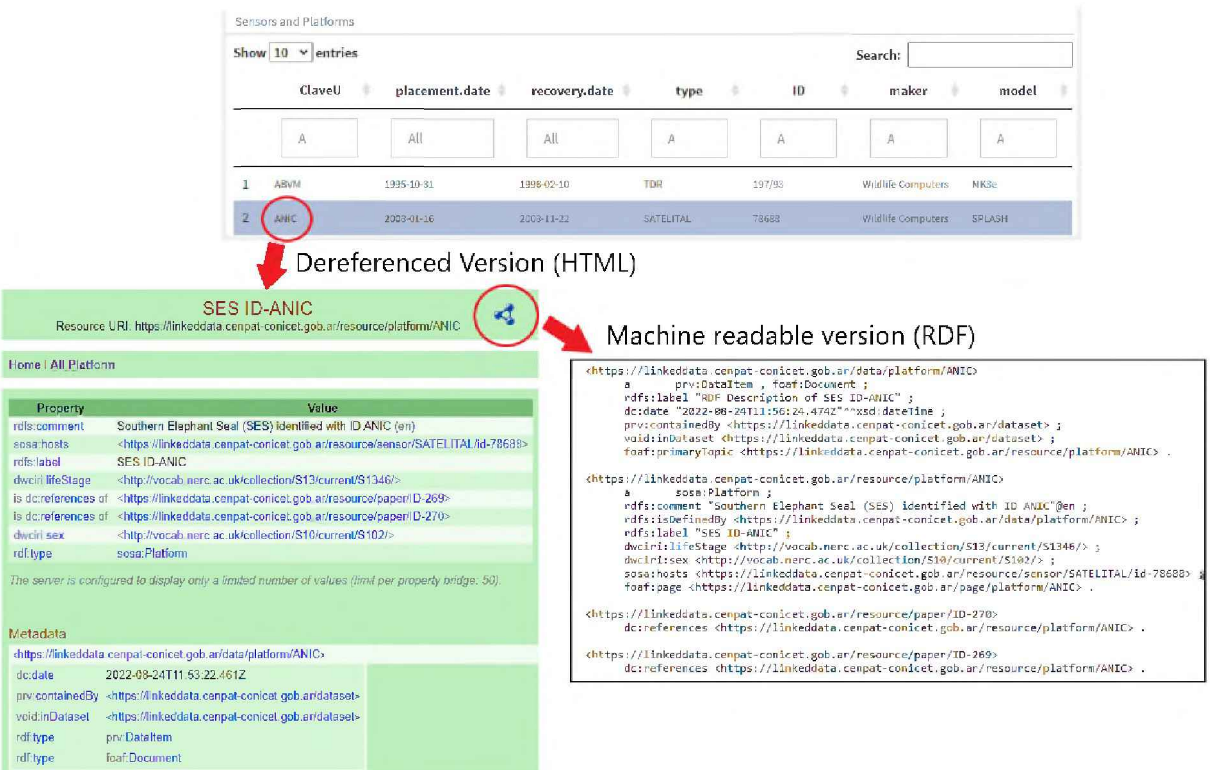

Fig. 2. *ODP-DASHBOARD* allows viewing the RDF version of the most important concepts; in this case, the information of the subject encoded as ANIC can be accessed in HTML version (dereferenced version) and RDF version (machine-readable version).

and wave height via its sensors. The dataset is available at the following link[1].

As illustrated in Figure 9, we mapped properties and sensor metadata to classes and properties of the SSN/SOSA ontology [11], including the following:

- `sosa:Platform`: Represents a physical or virtual entity that carries one or more sensors and can act as a host for observing and measuring the environment.
- `sosa:Sensor`: Represents a physical device that can measure, observe, or estimate properties of the environment.
- `sosa:FeatureOfInterest`: Represents the en-

[1] `https://aqualink.org/sites/1136`

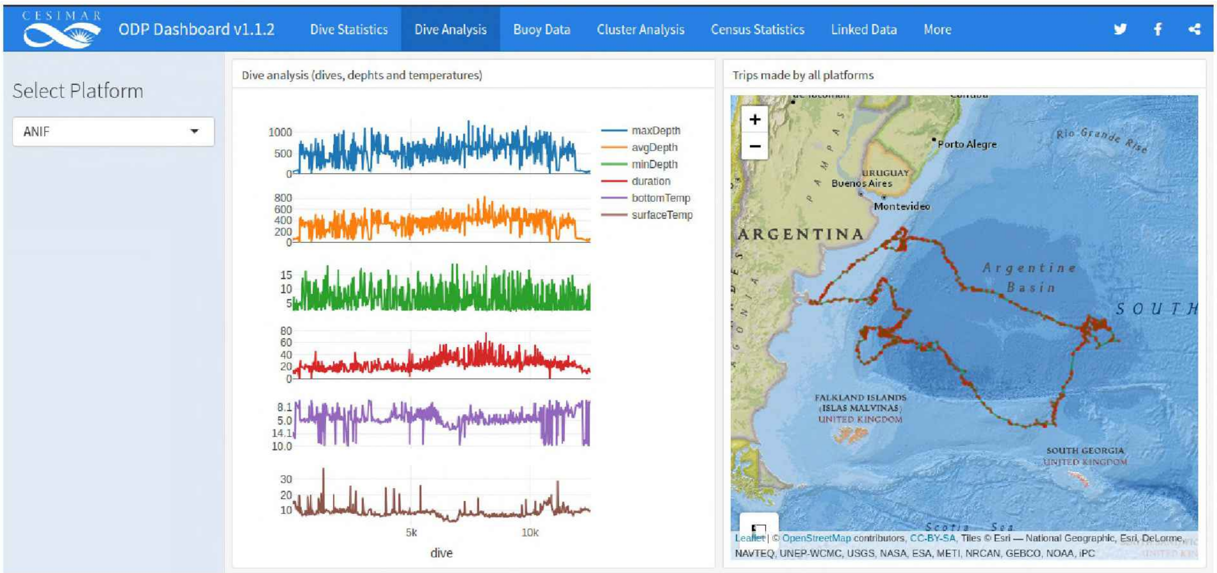

Fig. 3. 'Dive Analysis' screen overview.

Fig. 4. Users can select a date range on the left side of the 'Buoy Data' tab. The displayed data comprise surface temperature, depth temperature, and their average (top chart), wave height (middle chart), and wind speed (bottom chart).

tity or phenomenon being observed or measured by a sensor. It can be a point or area in space, a sample or specimen, or an event or process.

- `sosa:Observation`: Represents a record of an act of observing or measuring a property of a feature of interest by a sensor and captures the values of one or more observable properties at a specific time and place.
- `sosa:Result`: Represents the value or values of an observable property resulting from an observation.
- `sosa:ObservableProperty`: Represents a property or phenomenon that can be observed or measured by a sensor. It can be a physical, chemical, biological, or environmental property, among others.

Temporal entities were modeled using the W3C Time Ontology [33], while spatial entities were represented using the GeoSPARQL ontology [34]. Units of measurement were incorporated through the Quantities, Units, Dimensions, and Types (QUDT) ontology [35].

In the ontology-based testing approach, competency questions (CQs) serve as user-oriented interrogatives for exploring the ontology, implemented as SPARQL queries. SPARQL was chosen for its ability to query RDF schema and OWL models to filter individuals

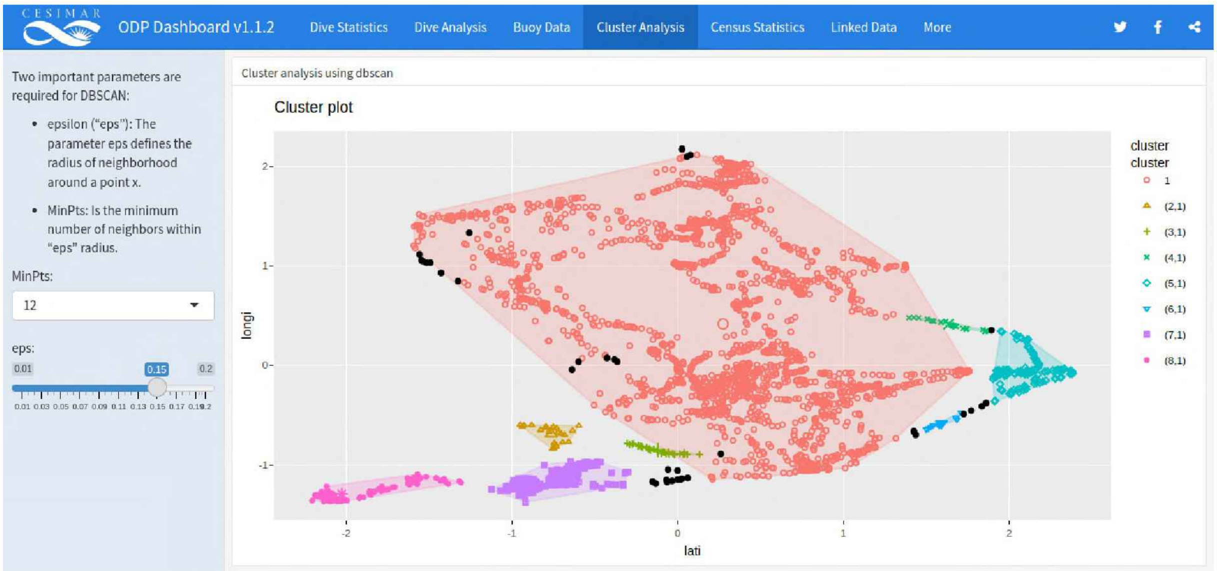

Fig. 5. ”Cluster Analysis” module.

Fig. 6. Census statistics of the different ODPs.

based on specific characteristics. Following interviews with domain experts, key objectives of the ontology model were established.

CQ01: What samples/species are collected by a particular agent?
CQ02: How many occurrences of a certain species are there?
CQ03: How many males of a species are in the dataset?
CQ04: Which specimens were collected on a specific date?
CQ05: What is the taxonomic classification of the collected object?
CQ06: Where are the mammals spatially located in the dataset?
CQ07: Which locations are associated with a certain species?
CQ08: What is the specific nature of the collected object?
CQ09: Which country does an occurrence belong to?
CQ10: Which marine region does a geographic coordinate belong to?
CQ11: What occurrences exist within a bounding box?
CQ12: Which species coexist in a certain marine region?
CQ13: Which institutions work with a certain species in Argentina?

This competency-based questioning approach has the following advantages: (i) rich semantic expressiveness, which cannot be achieved using only the graph-

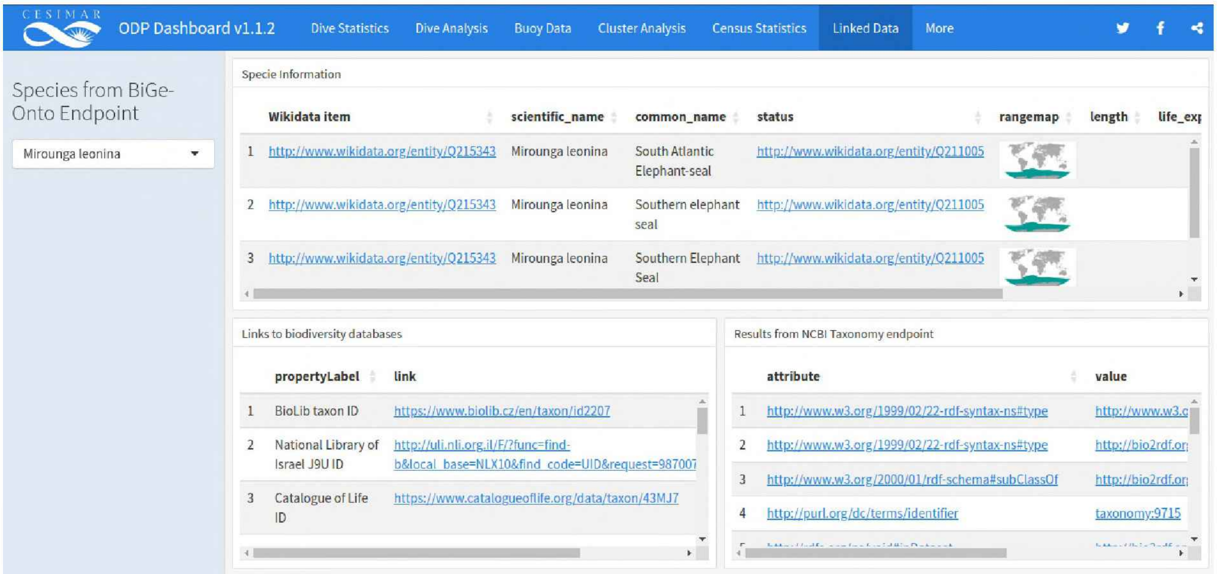

Fig. 7. "Linked Data" tab displaying information, in this case, related to the species "Mirounga leonina".

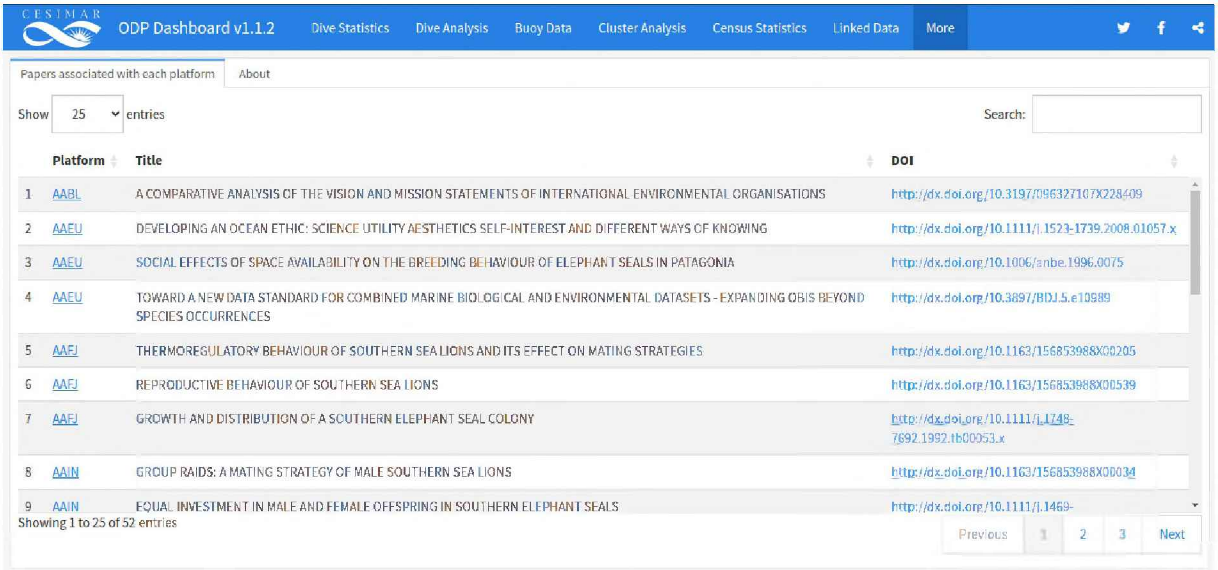

Fig. 8. Tab of the module that records the scientific production on southern elephant seals by GEMMA researchers. It includes a table with publications and search options.

ical model, (ii) evaluation of the reliability of the proposed model, and (iii) identification of inconsistencies. For each specific question, we developed a set of test cases, implementing the questions as SPARQL queries. Table 1 includes some of the developed test cases. Finally, after executing a test case, we compare the returned results with the expected results to determine its success. If the results match, the data model passes the test case. Otherwise, we analyze whether the issue lies in the conceptual model, its implementation, or even in the formulation/implementation of the competency questions. Test cases were executed using GraphDB[2]. Figure 10 demonstrates an example of test case T002 execution. As shown, the test case passed upon contrasting the actual result with the expected result. Overall, all test cases were successful, but minor issues were detected, mostly related to our model's implementation and some to the test cases themselves. Upon issue detection, necessary changes were made, and the test cases were re-executed. For instance, in test case T013, searching for the species *Otaria flavescens* in *Argentina*, the country code *ARG* was mistakenly inputted instead of the full country name, resulting in no results. To resolve this, consensus was reached to establish rules for country and species names. After adjustments, the test case was re-executed, yielding the expected results.

[2]`https://graphdb.ontotext.com/`

Fig. 9. Example showcasing how a record from a CSV file extracted from AquaLink is instantiated.

Table 1: Test cases developed

Case	CQ	Input	Expected result
T001	CQ01	Lewis, Mirtha	Otaria flavescens Shaw, 1800
T002	CQ02	Otaria flavescens	4418
T004	CQ04	28/01/1994	Otaria flavescens Shaw, 1800
T006	CQ06	Mammalia	8108 positions
T009	CQ09	Argentina	204761 occurrences
T013	CQ013	Otaria flavescens, Argentina	”CENPAT-CONICET”@en ANS.INST.PAT@en FPN-WCS@en CENPAT-FPN@en UNIV. AUSTRAL CHILE@en

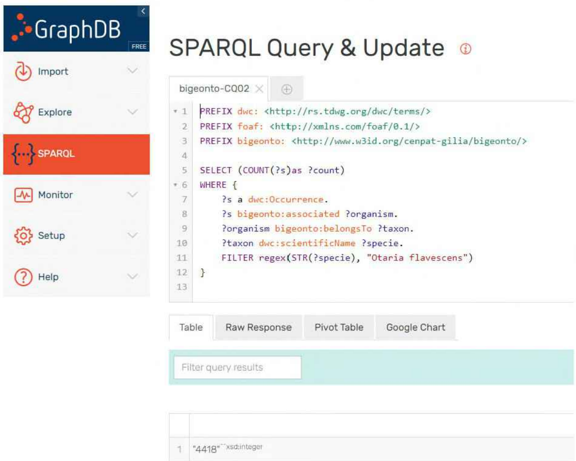

Fig. 10. Execution in GraphDB of test case T02.

5 Conclusions

The amount of data on the Web has increased enormously in the last two decades, specifically in the field of marine sciences. This amount of information makes it necessary to include semantics so that it can be processed by machines. Linked Data technology emerged in 2008 and can be defined as a set of best practices for sharing, exposing, and connecting information, data, and knowledge using RDF and OWL standards. This technology, supported by the W3C community, has been applied to various categories of data, such as Government, Publications, Social Web, Cross-domain, Geographic, Media, User-generated, and Life Sciences. As explained by [36], there are various problems in the integrated management of information in the field of life sciences, particularly in Marine Biodiversity and Oceanography. In this regard, only a few efforts have been presented and analyzed, but none of them covers all the requirements defined at the beginning. For this reason, in this paper, we present the development of a system oriented towards scientific visualization to manage information associated with Marine Biodiversity and Oceanography, using standards such as RDF, SOSA, and GeoSPARQL.

Based on these contributions, the initial steps were taken in creating an oceanographic dataset from the Aqualink initiative [37], where entities such as scientific publications, people, places, specimens, environmental variables, and institutions form part of a single shared knowledge space. In this article, we describe the modeling and publication process and the current and future uses of the dataset.

The evaluation of *ODP-DASHBOARD* shows that it meets the initial purpose of this paper: *to visually and integrally manage Marine Biodiversity and Oceanography information*, allowing us to answer the scientific questions that motivated our proposal. As shown in section 2.2, the main distinguishing feature of *ODP-DASHBOARD* compared to other systems is that, to our knowledge, it is the first attempt to cover aspects that are common in different areas of marine sciences. For example, *ODP-DASHBOARD* uses GeoSPARQL to include a standard way of providing spatial reasoning, such as topological relationships or overlaps between spatial entities. Regarding the evaluation of the *ODP-DASHBOARD* dataset, we provide a SPARQL access point so that other applications can consume and reuse our information. We present a set of user-defined SPARQL queries to demonstrate how we can manage marine science information.

Competing interests

The authors have declared that no competing interests exist.

Funding

PICT-2021-I-INVI-00245: *Extracción y explotación de conocimiento para la gestión en línea de datos en ciencias del mar.*

Authors' contribution

- Conceptualization (GN, MZ, CB)
- Software (GN)
- Supervision (MZ)
- Redaction - original manuscript (GN, MZ, CB)
- Redaction - review and editing (GN, MZ)

All the authors have read and approved the final version.

References

[1] V. Hatje, R. L. Andrade, C. C. d. Oliveira, A. Polejack, and T. Gxaba, "Pollutants in the south atlantic ocean: sources, knowledge gaps and perspectives for the decade of ocean science," *Frontiers in Marine Science*, vol. 8, p. 644569, 2021.

[2] R. P. Wilson, Y. Ropert-Coudert, and A. Kato, "Rush and grab strategies in foraging marine endotherms: the case for haste in penguins," *Animal Behaviour*, vol. 63, no. 1, pp. 85–95, 2002.

[3] B. A. Block, S. L. Teo, A. Walli, A. Boustany, M. J. Stokesbury, C. J. Farwell, K. C. Weng, H. Dewar, and T. D. Williams, "Electronic tagging and population structure of atlantic bluefin tuna," *Nature*, vol. 434, no. 7037, pp. 1121–1127, 2005.

[4] J.-B. Charrassin, Y.-H. Park, Y. L. Maho, and C.-A. Bost, "Penguins as oceanographers unravel hidden mechanisms of marine productivity," *Ecology Letters*, vol. 5, no. 3, pp. 317–319, 2002.

[5] B. J. Le Boeuf and R. M. Laws, *Elephant seals: population ecology, behavior, and physiology*. Univ of California Press, 1994.

[6] M. A. Ferrari, C. Campagna, R. Condit, and M. N. Lewis, "The founding of a southern elephant seal colony," *Marine Mammal Science*, vol. 29, no. 3, pp. 407–423, 2013.

[7] A. Leadbetter, R. Lowry, and D. Clements, "The nerc vocabulary server: version 2.0," in *Geophysical Research Abstracts*, vol. 14, 2012.

[8] A. A. Krisnadhi, R. Arko, S. Carbotte, C. Chandler, M. Cheatham, T. Finin, P. Hitzler, K. Janowicz, T. Narock, L. Raymond *et al.*, "An ontology pattern for oceanograhic cruises: Towards an oceanograhper's dream of integrated knowledge discovery," 2014.

[9] T. Robertson, M. Döring, R. Guralnick, D. Bloom, J. Wieczorek, K. Braak, J. Otegui, L. Russell, and P. Desmet, "The gbif integrated publishing toolkit: facilitating the efficient publishing of biodiversity data on the internet," *PloS one*, vol. 9, no. 8, p. e102623, 2014.

[10] J. Wieczorek, D. Bloom, R. Guralnick, S. Blum, M. Döring, R. Giovanni, T. Robertson, and D. Vieglais, "Darwin core: an evolving community-developed biodiversity data standard," *PloS one*, vol. 7, no. 1, p. e29715, 2012.

[11] A. Haller, K. Janowicz, S. J. Cox, M. Lefrançois, K. Taylor, D. Le Phuoc, J. Lieberman, R. García-Castro, R. Atkinson, and C. Stadler, "The modular ssn ontology: A joint w3c and ogc standard specifying the semantics of sensors, observations, sampling, and actuation," *Semantic Web*, vol. 10, no. 1, pp. 9–32, 2019.

[12] A. Hogan, E. Blomqvist, M. Cochez, C. d'Amato, G. D. Melo, C. Gutierrez, S. Kirrane, J. E. L. Gayo, R. Navigli, S. Neumaier *et al.*, "Knowledge graphs," *ACM Computing Surveys (Csur)*, vol. 54, no. 4, pp. 1–37, 2021.

[13] A. M. Treasure, F. Roquet, I. J. Ansorge, M. N. Bester, L. Boehme, H. Bornemann, J.-B. Charrassin, D. Chevallier, D. P. Costa, M. A. Fedak *et al.*, "Marine mammals exploring the oceans pole to pole: a review of the meop consortium," *Oceanography*, vol. 30, no. 2, pp. 132–138, 2017.

[14] M. C. Burton, H. and J. van den Hoff, "Subset of data from the tags database of known age seals - elephant seals, ver. 2," http://dx.doi.org/doi:10.4225/15/59fa97f242641, 2017, [Online; accessed 2023-05-03].

[15] R. Kays, S. C. Davidson, M. Berger, G. Bohrer, W. Fiedler, A. Flack, J. Hirt, C. Hahn, D. Gauggel, B. Russell *et al.*, "The movebank system for studying global animal movement and demography," *Methods in Ecology and Evolution*, vol. 13, no. 2, pp. 419–431, 2022.

[16] K. Janowicz, P. Hitzler, B. Adams, D. Kolas, I. Vardeman *et al.*, "Five stars of linked data vocabulary use," *Semantic Web*, vol. 5, no. 3, pp. 173–176, 2014.

[17] L. Penev, M. Dimitrova, V. Senderov, G. Zhelezov, T. Georgiev, P. Stoev, and K. Simov, "Openbiodiv: a knowledge graph for literature-extracted linked open

data in biodiversity science," *Publications*, vol. 7, no. 2, p. 38, 2019.

[18] R. D. Page, "Ozymandias: A biodiversity knowledge graph," *PeerJ*, vol. 7, p. e6739, 2019.

[19] M. Popović, N. Vasić, T. Koren, I. Burić, N. Živanović, D. Kulijer, and A. Golubović, "Biologer: an open platform for collecting biodiversity data," *Biodiversity Data Journal*, vol. 8, 2020.

[20] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne *et al.*, "The fair guiding principles for scientific data management and stewardship," *Scientific data*, vol. 3, no. 1, pp. 1–9, 2016.

[21] R. M. Adler, "Distributed coordination models for client/server computing," *Computer*, vol. 28, no. 4, pp. 14–22, 1995.

[22] C. Beeley, *Web application development with R using Shiny*. Packt Publishing Ltd, 2013.

[23] R. Oldakowski, "D2rq platform–treating non-rdf databases as virtual rdf graphs," *Nature Precedings*, pp. 1–1, 2011.

[24] P. Hitzler, M. Krötzsch, B. Parsia, P. F. Patel-Schneider, S. Rudolph *et al.*, "Owl 2 web ontology language primer," *W3C recommendation*, vol. 27, no. 1, p. 123, 2009.

[25] C. Bizer, T. Heath, and T. Berners-Lee, "Linked data: The story so far," in *Semantic services, interoperability and web applications: emerging concepts*. IGI global, 2011, pp. 205–227.

[26] T. Berners-Lee, J. Hendler, and O. Lassila, "A new form of web content that is meaningful to computers will unleash a revolution of new possibilities," *Scientific american*, vol. 284, no. 5, pp. 34–43, 2001.

[27] R. Iannone, J. Allaire, B. Borges *et al.*, "flexdashboard: R markdown format for flexible dashboards," *R package version 0.5*, vol. 1, 2018.

[28] "Package "leaflet"," https://cran.r-project.org/web/packages/leaflet/leaflet.pdf, 2023, [Online; accessed 8-Nov-2023].

[29] "Aqualink," https://aqualink.org/sites/1136, 2023, [Online; accessed 06-Nov-2023].

[30] C. Campagna and M. Lewis, "Growth and distribution of a southern elephant seal colony," *Marine Mammal Science*, vol. 8, no. 4, pp. 387–396, 1992.

[31] M. Zárate, G. Braun, P. R. Fillottrani, C. Delrieux, and M. Lewis, "Bige-onto: An ontology-based system for managing biodiversity and biogeography data," 2019, accepted paper.

[32] J. Brank, M. Grobelnik, and D. Mladenić, "A survey of ontology evaluation techniques," 2005.

[33] S. Cox and C. Little, "Time ontology in owl, w3c candidate recommendation 26 march 2020," [Online; accessed 14-Apr-2023]. [Online]. Available: https://www.w3.org/TR/owl-time/

[34] M. Perry and J. Herring, "Ogc geosparql-a geographic query language for rdf data," *OGC implementation standard*, 2012.

[35] R. Hodgson, P. J. Keller, J. Hodges, and J. Spivak, "Qudt-quantities, units, dimensions and data types ontologies," 2014, [Online; accessed 14-Apr-2023]. [Online]. Available: http://qudt.org/

[36] B. M. Good and M. D. Wilkinson, "The life sciences semantic web is full of creeps!" *Briefings in bioinformatics*, vol. 7, no. 3, pp. 275–286, 2006.

[37] G. Nuñez, C. E. Buckle, and M. D. Zárate, "Integrating oceanographic sensor data using ssn/sosa ontology," in *XI Jornadas de Cloud Computing, Big Data & Emerging Topics (La Plata, 27 al 29 de junio de 2023)*, 2023.

Citation: G. Nuñez, C. Buckle and M. Zárate. *ODP-DASHBOARD: Enhancing Marine Species Conservation in the South Atlantic through Linked Open Data Integration*. Journal of Computer Science & Technology, vol. 24, no. 2, pp. 182-192, 2024.
DOI: 10.24215/16666038.24.e17.
Received: April 15, 2024 **Accepted:** August 6, 2024.