

PREPRINT

Author-formatted, not peer-reviewed document posted on 17/08/2023

DOI: <https://doi.org/10.3897/aphapreprints.e111241>

SES Dashboard: A linked data platform to manage data on Southern Elephant Seals

 **Marcos Zárate, Elena Eder, Claudio Delrieux, Mirtha Lewis, Gustavo Nuñez, Dario Ceballos**

SES Dashboard: A linked data platform to manage data on Southern Elephant Seals

Marcos Zárate[‡], Elena Eder[‡], Claudio Delrieux[§], Mirtha Lewis[‡], Gustavo Nuñez[|], Dario Ceballos[‡]

[‡] Center for the Study of Marine Systems, National Council for Scientific and Technical Research, (CESIMAR-CONICET), Puerto Madryn, Argentina

[§] Institute of Computer Science and Engineering, Bahía Blanca Scientific and Technological Center, (ICIC-CONICET), Bahía Blanca, Argentina

[|] Computer Research Laboratory, National University of Patagonia San Juan Bosco, (LINVI-UNPSJB), Puerto Madryn, Argentina

Corresponding author: Marcos Zárate (zarate@cenpat-conicet.gob.ar)

Abstract

Background

In this paper we present SES Dashboard, a linked data platform developed to manage information of Southern Elephant Seals (*Mirounga leonina* Linnaeus, 1758) from Península Valdés colony (Patagonia, Argentina), collected over more than two decades of research. Southern elephant seals spend most of their time foraging at sea, alternating with two short and highly synchronized haul-outs ashore for breeding and molting. On land, the distribution of the colony extends along ≈350 km, while at sea individuals travel thousands of kilometers. The information recorded is mainly based on georeferenced data from censuses, bioregisters (trajectories at sea and hydrographic profiles), and bibliography based on the above data.

New information

SES Dashboard was released in 2021 and published following the principles of Linked Open Data to improve its visibility, to make it shareable, extensible and easily reusable through a SPARQL endpoint.

Keywords

Interoperability, Linked Open Data, *Mirounga leonina*, Pinnipedia, Phocidae, SPARQL

Introduction

The Southwestern Atlantic Ocean, particularly its subpolar and polar coastal environments, is relatively understudied compared to other regions of the planet. While there has been some scientific research conducted in this area, much of it has been limited in scope and duration. Several factors contribute to the limited study of the Southwestern Atlantic Ocean. One of the main reasons is its remote location and challenging environmental conditions. The subpolar and polar coastal environments are characterized by harsh weather, strong winds, and cold temperatures, making it difficult for researchers to access and operate in these areas for extended periods. This represents a serious gap in our knowledge for the construction of global models in predicting future climate scenarios among other purposes (Hatje et al. 2021). In this context, there is an enormous challenge and an urgent need to increase the data acquisition efforts in these latitudes, to further understand climate change effects on marine ecosystems.

Studies combining data loggers and satellite remote sensors deployed on marine vertebrates have revolutionized our view of the ocean and cleared up intricate associations between the physical environment, productivity, and foraging ecology of the main migratory predators (e.g. Hatje et al. 2021, Polovina et al. 2004, Hindell et al. 2020). Large marine ecosystem have been explored *in situ* using various species of pelagic predators as oceanographic platform, such as giant bluefin tuna (Block et al. 2005), penguins (Charrassin et al. 2002), and marine mammals (e.g. Fadely et al. 2005, Treasure et al. 2017). Results range from detailed depth temperature profiles (for example, Boyd et al. 2001) to mesoscale characterization of oceanographic features, such as eddies and meanders (Sokolov et al. 2006). Among the wide range of species involved, foraging Southern Elephant Seals (SES) combine deep diving with long distance travel, covering thousands of kilometers across a wide range of oceanographic regions for up to 7 months at sea (Le Boeuf and Laws 1995). The GEMMA research group (Grupo de Estudios de Macroecología Marina-CESIMAR-CENPAT-CONICET) addresses different aspects of the marine ecosystem, taking SES as model for biological, oceanographic and ecological studies (Retana 2013, Muelbert et al. 2013, Ferrari et al. 2012). The study area covers a variety of habitats and environmental gradients in sectors of the Argentine coast and the Southwestern Atlantic Ocean, including the continental shelf and slope and adjacent ocean basin. The methodology includes satellite technology and miniaturized devices for spatial ecology, animal movement and environmental data recording as well as direct biological data sampling. All these surveys generated a database with years of information and requires a web platform for its friendly use. SES Dashboard is currently designed for researchers who work specifically with *Mirounga leonina* from the Valdés Peninsula (PV) in the Chubut Province (Argentina), although the model is sufficiently generic to be adapted to other colonies and other species of marine mammals. Also, the hidrographic profiles recorded by the satellite instruments describe oceanographic features in the mesoscale adequate to analyze the behavior of different species in a useful way. Specifically, 9 adult females were used as sampling platforms, counting 3183 dives which were carried out

between 1994 and 2008. The data from terrestrial censuses were carried out between 1990 and 2017. This platform provides support for decision-making related to the conservation of the species inhabiting the Southwestern Atlantic Ocean. Also, the 2021 version was published following Linked Open Data (LOD) principles (Janowicz et al. 2014), which allows increasing the intended audience by reusing of data, so the range of possibilities is broadened. In the following sections we present the main features of data management, the underlying technologies, and the way to publish the data as LOD, and finally we discuss the potential use and reuse of data.

State of the art

According to a thorough literature review, we did not find specific SES data repositories or data management systems that meets LOD criteria.

However, we surveyed the following: Marine Mammals Exploring the Oceans Pole to Pole (MEOP; Treasure et al. 2017) initiative, which is a consortium of international researchers dedicated to sharing animal-derived data and knowledge about the polar oceans. Collectively, MEOP demonstrates the power and cost-effectiveness of using marine mammals as data-collection platforms that can dramatically improve the ocean-observing system for biological and physical oceanographers. In this case, the published data does not comply with LOD principles, which limits its exploitation by machines.

Burton et al. (2017), this database is a compendium of histories of known age seals (Southern elephant) from observations across the Southern Ocean but focussed on Macquarie Island, Marion Island, Heard Island, Mawson and the Vestfold Hills. Another initiative that is important to mention is Movebank (Kays et al. 2022), a free online platform for storing, sharing, and analyzing animal movement data. It was developed by the Max Planck Institute of Animal Behavior and is now managed by the Movebank Data Repository, which is a collaborative project of multiple research institutions. The platform allows researchers to upload and share data from GPS tracking devices, radio transmitters, and other sensors attached to animals. The data can then be visualized, analyzed, and used to understand animal behavior, migration patterns, and other ecological and conservation-related research questions. Movebank has become a valuable resource for the scientific community, providing a centralized database of animal movement data from all over the world. In this case, the published data does not comply with LOD principles, which limits its exploitation by machines.

Regarding systems that base their data model on LOD principles (Janowicz et al. 2014), we can mention: OpenBiodiv Knowledge Graph (Penev et al. 2019), presented at the annual conference of the Biodiversity Information Standards (TDWG) in 2016, is based mainly on data from [Pensoft](#) and [Plazi](#). This means that the coverage of this work mainly focused on the bibliography associated with the taxa.

Ozymandias Knowledge Graph (Page 2018), winner of the GBIF Ebbe Nielsen Challenge 2018, is based on data from the [Atlas of Living Australia](#) (ALA) and [Australian Faunal Directory](#) (AFD). This knowledge graph combines information about taxa, publications, and

authors along with their interrelationships. The third system, Biologer (Popović et al. 2020), is a simple and free-access system designed to collect data on biological diversity of Eastern Europe. This system does not use LOD principles and neither provides bibliographic information associated with species.

All projects manage taxa, persons, and institutions and have practical applications. Nevertheless, only Page (2018) and Penev et al. (2019) make use of LOD principles. For this reason, we discarded (Popović et al. 2020), because developing a system that does not use LOD as a backend can be considered a disadvantage when it comes to interoperating and complying with the FAIR principles (findable, accessible, interoperable, reusable; Wilkinson et al. 2016). However, in contrast to our approach, Ozymandias Knowledge Graph and OpenBiodiv Knowledge Graph do not provide visual components (such as maps) for studies involving species and environmental variables.

Project description

Title: SES-Dashboard

Study area description: Animal behavior

Proposed dashboard

The main goal to develop our dashboard is to provide satisfactory visual analytics to users in a way such that data can be reused and interpreted by other applications. For this we must comply with the FAIR, at least partially. Initially this development used data in CSV format extracted from a MySQL relational database as data source. Due to the semantic and reuse limitations of this approach, we decided to carry out a conversion process of these data to LOD (see the Data transformation process subsection below). After this step, the links obtained from the conversion process were incorporated into the dashboard so that the application had the possibility of redirecting to the machine-readable version of each data displayed.

Achitecture overview

The architecture of our application is based on the client/server model (Adler 1995), where the first side (client) performs requests that are processed by the second part (server) to provide suitable responses. The server side was developed using Shiny (Chang et al. 2023) which is a powerful R package that allows to create interactive web applications using the R programming language. Shiny applications have two components. The front-end component builds the web page that is actually shown to the user and includes layout, appearance, and design features. In Shiny terminology, this is called the **ui**, which stands for user interface. The **ui** file contains R functions that are then translated into an HTML file. The other component is the back-end, which includes the code for producing the app's contents (e.g. functions or data import, management, and analysis). Here, we create the objects that are later shown on the front-end. In Shiny terminology this is called **server**.

Dashboard development

SES dashboard was developed to querying the statistics of the dives and the trip routes. We used the R flexdashboard package (Sievert et al. 2022) to generate web pages based on an R Markdown document (Xie et al. 2018). To query our endpoint we use the [SPARQL Package](#) that allows to directly import results of SPARQL SELECT queries into the statistical environment of R as a dataframe. The following section describes each module of the dashboard.

1. **Diving statistics:** When an instrument is attached to record diving data (depth, time, water temperature, etc.) the device is programmed to sample at certain time intervals. After being recovered, the data is downloaded, pre-processed to find anomalies and then stored in a relational database or csv files so that the user can apply basic statistics that describe the dives or analyze each dive in particular (Sala et al. 2011). This module summarizes the diving statistics made by all SES (maximum depths recorded, number of dives, maximum temperatures and the number of platforms, i.e. SES individuals), and the detailed information for each of the used sensors. The view is shown for nine SES individuals in Fig. 1. The dive summary is detailed at the top side, sensors and instruments used are tabulated at the bottom left side, and the number of dives made by each platform are plotted in a bar graph at the bottom right side. The ggplot library (Villanueva and Chen 2019) was used for bar charts. The platforms and their sensors have links to their description in LOD format, i.e. the dereferenced version of the sensor in [HTML format](#), and the machine readable [RDF version](#). These representations are shown in Fig. 2 for the sensor of individual coded ANIC.
2. **Dive Analysis:** The dive analysis module shows the variables of interest registered by platform during dives. Temperatures and depths, as well as duration can be displayed. Fig. 3 shows the screen of analysis reports of the dives made by the SES coded ANID, 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 for this screen was built using plotly library (Sievert et al. 2022).
3. **Trip analysis:** This module retrieve the trips made by each platform and displays them on a map generated with [leaflet library](#). At each point of the trip, dive information and data associated with depth, temperature and duration are also displayed. A tool is provided to measure distances and areas in km. Fig. 4 shows the view of the trip analysis for a specific individual (coded ANIF), and the icon to measure distances and areas at the bottom left side of the screen.
4. **Cluster Analysis:** This module provides a spatial cluster analysis using dbscan package (Hahsler et al. 2019), as a tool to understand the distribution of SES based on Density Based Spatial Clustering of Applications with Noise (DBSCAN; Ram et al. 2010, Sander et al. 1998). In this case, an ϵ value (i.e., the neighborhood radius around a data point) of 0.11 and a $MinPts$ (i.e., the amount of datapoints within each radio) of 10 were chosen. The algorithm starts with an arbitrary point in the dataset. If the neighbor points within the radius are more than

10, then the algorithm considers them to be part of a cluster. Then, it recursively expands the cluster by checking if all the new points have more than the selected *MinPts* neighbor points within the radius. Eventually, it runs out of points to add to the cluster. Then, by choosing a new arbitrary point, the process is repeated. If a data point has fewer neighbors than *MinPts*, it will be considered as a "noise". After this process is completed, all points are either in a cluster or considered noise. In Fig. 5 a view with this specific geospatial analysis is shown.

5. **Census statistics:** This module allows the visualization of basic statistics of the primary data collected during censuses from 1990 to 2017. During the census, only the number of individuals that can be counted visually is recorded, this means that SES are not individualized. Ground censuses were carried out at the peak of the breeding season (first week of October), when approximately 96% of breeding females were ashore (Campagna and Lewis 1992). Surveys on foot encompass the nearly 200 km of open sea shoreline front occupied by breeding seals in PV and additional sectors software (\approx 150 km), outside PV. However, censuses may not have been consistent regarding date and surveyed sectors along the years. A chart was developed with the ggplot module to plot an histogram of the annual numbers of SES present at the peak of the breeding season of each surveyed year, discriminated by age category (see Fig. 6).

6. **Bibliography:** The purpose of this module is to register and maintain an updated record of the scientific production related to southern elephant seals generated by GEMMA researchers. A table is presented with the list of all publications as scientific papers, books chapter, or reports. It is possible to search using parameters such as author, text included in the title or abstract, keywords, DOI, year, etc. Fig. 7 shows these options. It is important to track which SES individuals were included in a publication, either because their data is used for some analysis, plot on a map, etc. SES Dashboard allows also to create links between the SES and its associated publications.

Web location (URLs)

Homepage: <https://cesimar.shinyapps.io/SESDashboard/>

Download mirror: <https://github.com/marcosdzarate/ses-dashboard/archive/refs/heads/master.zip>

Technical specification

Platform: Web (shinyapps.io)

Programming language: R

Service endpoint: <https://linkeddata.cenpat-conicet.gob.ar/snorql/> (Snorql interface)

Repository

Type: github

Browse URI: <https://github.com/marcosdzarate/ses-dashboard>

Usage licence

Usage licence: Creative Commons Public Domain Waiver (CC-Zero)

Additional information

Data transformation process

To perform the data transformation, the D2RQ Platform is used (Oldakowski 2011). This consists of: **1) D2RQ Mapping Language**, which is used to write mappings between database tables and RDF vocabularies or OWL ontologies (Hitzler et al. 2009), **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 [DOI:10.17632/5nv5c7575w.3](https://doi.org/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.

Discussion

In this section we discuss the aspects related to fulfillment of LOD principles and aspects related to the underlying semantic model used by SES dashboard. LOD (Bizer et al. 2011) is an idea stemming from the Semantic Web (Berners-Lee et al. 2001), aiming to ensure that data published on the Web is reusable, discoverable, and more importantly, that data published by different entities can work together. LOD principles are summarized in:

1. The use URIs as names for resources.
2. The Use HTTP URIs so people can look up these resources.
3. The supply of useful information using the standards (RDF, SPARQL) when using a URI.
4. The inclusion of links to other URIs to help users to discover more resources.

Usage of URIs as resource identifiers

Each instance is uniquely identifiable by 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 schema.

Usage of HTTP URIs and dereferencing

According with linked data principles, dereferenceable HTTP URIs for our resources was used. For example for the average depth URI shown above, a human-readable version of the dereferenced version using the URL: http://linkeddata.cenpat-conicet.gob.ar/page/result/id-233/avg_depth was generated to dereference the URI.

Linking to other resources

First, all resources in the dataset form a Knowledge Graph (there are no disconnected parts) (Hogan et al. 2021), following a data model discussed in the next subsection. Second, resources are linked to external databases via properties like **owl:sameAs**, **skos:broader** and **skos:exactMatch**. These identifiers can be: ORCIDs, Wikidata entities, DBPedia resources, NERC vocabulary server ID's, etc. We have created links between people and their ORCID records, publications and their OpenCitations records, and the environmental variables were related to the identifier in NERC. For example, average water temperature corresponds to the NERC identifier **SDN:P02::TEMP** through **skos:broader** property. (See: http://linkeddata.cenpat-conicet.gob.ar/resource/observableProperty/id-233/avg_temp).

Availability, sustainability, and licensing

The SES dataset can be downloaded, navigated and queried using a SPARQL endpoint, and it can be published under Creative Commons Universal Public Domain Dedication (CC0 1.0) License. All the criteria for five star Linked Data are met, as defined in (Janowicz et al. 2014). A description of the data online, an RDF version of the data, many links to structured vocabularies, and metadata about the collection is all made available. Our dataset characteristics are listed in Table 1. To explore the dataset using the SPARQL endpoint, we have developed a set of queries to answer the most common questions that researchers need to answer. For example, number of dives, trips, and values of certain environmental variables. A summary of the developed queries and their corresponding links to the endpoint are presented in Table 2.

FAIR compliance

The previous version launched in 2018 of SES Dashboard is not 100% compatible with the FAIR principles, which makes it difficult to reuse. Due to this limitation, we decided to promote the publication of the data as LOD in 2021. The LOD dataset was a conversion from the data contained in SES database into Resource Description Framework (RDF). Berners-Lee suggested LOD principles for judging data quality by its accessibility (open data access), format and structures, and interoperability with other data sources. The FAIR data principles have been introduced for similar reasons with a greater emphasis on

achieving reuse. LOD gives a clear mandate to the opening of the data, while FAIR requires an established licence for access and therefore includes the concept of reuse under consideration in the licence agreement. In addition, FAIR makes a strong reference to contextual information required to improve data reuse. In accordance with LOD principles, such metadata would be considered interoperable data as well, however the requirement to augment the data with metadata indicates that FAIR is an extension of the LOD (Hasnain and Rebholz-Schuhmann 2018).

Conclusion

We have presented SES Dashboard, a web platform capable of exploring the data collected by GEMMA research group. Each module contributes directly or indirectly to the knowledge of the temporal and spatial variation of the SES colony. This platform sorts and synthesizes information and allows quick searches to answer research queries. Also, the data life cycle can be traced back to the origin using the unique key in each SES recorded. It can be replicated for other research groups of SES with long-term and systematic studies, and it is also adaptable for the monitoring of other marine or terrestrial species that can be tagged/marked/captured and resighted/recaptured. SES Dashboard guarantee the traceability throughout the whole data life cycle. Even though SES Dashboard has not full compliance with the FAIR principles, the publication of the LOD version implicitly allows a significant fulfillment, and all criteria are met for five-star linked data. In addition, a description of the data online, an RDF version of the data, many links to structured vocabularies, and metadata about the collection, are all available. Among the approaches planned for future works, the reengineering of the platform to allow the recovery and presentation of information using our SPARQL endpoint, which enable to create links to other linked datasets to supplement the information is considered.

References

- Adler RM (1995) Distributed coordination models for client/server computing. Computer 28 (4): 14-22. <https://doi.org/10.1109/2.375173>
- Berners-Lee T, Hendler J, Lassila O (2001) The semantic web. Scientific American 284 (5): 34-43. <https://doi.org/10.1038/scientificamerican0501-34>
- Bizer C, Heath T, Berners-Lee T (2011) Linked data: The story so far. Semantic services, interoperability and web applications: emerging concepts 205-227. <https://doi.org/10.4018/978-1-60960-593-3.ch008>
- Block B, Teo SH, Walli A, Boustany A, Stokesbury MW, Farwell C, Weng K, Dewar H, Williams T (2005) Electronic tagging and population structure of Atlantic bluefin tuna. Nature 434 (7037): 1121-1127. <https://doi.org/10.1038/nature03463>
- Boyd IL, Hawker EJ, Brandon MA, Staniland IJ (2001) Measurement of ocean temperatures using instruments carried by Antarctic fur seals. Journal of Marine Systems 27 (4): 277-288. [https://doi.org/10.1016/s0924-7963\(00\)00073-7](https://doi.org/10.1016/s0924-7963(00)00073-7)
- Burton H, McMahon C, van den Hoff J (2017) Subset of data from the TAGS database of known age seals - elephant seals. <http://dx.doi.org/doi:10.4225/15/59fa97f242641>. Accessed on: 2023-8-02.

- Campagna C, Lewis M (1992) Growth and distribution of a southern elephant seal colony. *Marine Mammal Science* 8 (4): 387-396. <https://doi.org/10.1111/j.1748-7692.1992.tb00053.x>
- Chang W, Cheng J, Allaire J, Sievert C, Schloerke B, Xie Y, Allen J, McPherson J, Dipert A, Borges B (2023) shiny: Web Application Framework for R. R package version 1.7.4.9002.
- Charrassin J, Park Y, Maho YL, Bost C (2002) Penguins as oceanographers unravel hidden mechanisms of marine productivity. *Ecology Letters* 5 (3): 317-319. <https://doi.org/10.1046/j.1461-0248.2002.00341.x>
- Fadely B, Robson B, Sterlyng J, Greig A, Call K (2005) Immature Steller sea lion (*Eumetopias jubatus*) dive activity in relation to habitat features of the eastern Aleutian Islands. *Fisheries Oceanography* 14: 243-258. <https://doi.org/10.1111/j.1365-2419.2005.00379.x>
- Ferrari M, Campagna C, Condit R, Lewis M (2012) The founding of a southern elephant seal colony. *Marine Mammal Science* 29 (3): 407-423. <https://doi.org/10.1111/j.1748-7692.2012.00585.x>
- Hahsler M, Piekenbrock M, Doran D (2019) dbscan: Fast density-based clustering with R. *Journal of Statistical Software* 91: 1-30. <https://doi.org/10.18637/jss.v091.i01>
- Hasnain A, Rebholz-Schuhmann D (2018) Assessing FAIR Data Principles against the 5-Star Open Data Principles. *Lecture Notes in Computer Science* 469-477. https://doi.org/10.1007/978-3-319-98192-5_60
- Hatje V, Andrade R, Oliveira C, Polejack A, Gxaba T (2021) Pollutants in the South Atlantic Ocean: Sources, knowledge gaps and perspectives for the decade of ocean science. *Frontiers in Marine Science* 8. <https://doi.org/10.3389/fmars.2021.644569>
- Hindell MA, Reisinger RR, Ropert-Coudert Y, Hückstädt LA, Trathan PN, Bornemann H, Charrassin J, Chown SL, Costa DP, Danis B, Lea M, Thompson D, Torres LG, Van de Putte AP, Alderman R, Andrews-Goff V, Arthur B, Ballard G, Bengtson J, Bester MN, Blix AS, Boehme L, Bost C, Boveng P, Cleeland J, Constantine R, Corney S, Crawford RJM, Dalla Rosa L, de Bruyn PJN, Delord K, Descamps S, Double M, Emmerson L, Fedak M, Friedlaender A, Gales N, Goebel ME, Goetz KT, Guinet C, Goldsworthy SD, Harcourt R, Hinke JT, Jerosch K, Kato A, Kerry KR, Kirkwood R, Kooyman GL, Kovacs KM, Lawton K, Lowther AD, Lydersen C, Lyver PO, Makhado AB, Márquez MEI, McDonald BI, McMahon CR, Muelbert M, Nachtsheim D, Nicholls KW, Nordøy ES, Olmastroni S, Phillips RA, Pistorius P, Plötz J, Pütz K, Ratcliffe N, Ryan PG, Santos M, Southwell C, Staniland I, Takahashi A, Tarroux A, Trivelpiece W, Wakefield E, Weimerskirch H, Wienecke B, Xavier JC, Wotherspoon S, Jonsen ID, Raymond B (2020) Tracking of marine predators to protect Southern Ocean ecosystems. *Nature* 580 (7801): 87-92. <https://doi.org/10.1038/s41586-020-2126-y>
- Hitzler P, Krötzsch M, Parsia B, Patel-Schneider P, Rudolph S (2009) OWL 2 web ontology language primer. W3C recommendation URL: <https://www.w3.org/TR/2009/PR-owl2-primer-20090922/all.pdf>
- Hogan A, Blomqvist E, Cochez M, D'amato C, De Melo G, Gutierrez C, Kirrane S, Labra Gayo J, Navigli R, Neumaier S, Polleres A, Rashid S, Rula A, Schmelzeisen L (2021) Knowledge graph. *ACM Computing Surveys (Csur)* 54 (4): 1-37. URL: <https://dl.acm.org/doi/pdf/10.1145/3447772>
- Janowicz K, Hitzler P, Adams B, Kolas D, Vardeman II C (2014) Five stars of Linked Data vocabulary use. *Semantic Web* 5 (3): 173-176. <https://doi.org/10.3233/sw-140135>

- Kays R, Davidson SC, Berger M, Bohrer G, Fiedler W, Flack A, Wikelski M (2022) The Movebank system for studying global animal movement and demography. Methods in Ecology and Evolution 13 (2): 419-431. <https://doi.org/10.1111/2041-210X.13767>
- Le Boeuf BJ, Laws RM (1995) Elephant seals: population ecology, behavior, and physiology. Choice Reviews Online 32 (06): 32-3308. <https://doi.org/10.5860/choice.32-3308>
- Muelbert MC, de Souza R, Lewis M, Hindell M (2013) Foraging habitats of southern elephant seals, *Mirounga leonina*, from the Northern Antarctic Peninsula. Deep Sea Research Part II: Topical Studies in Oceanography 88-89: 47-60. <https://doi.org/10.1016/j.dsr2.2012.07.009>
- Oldakowski R (2011) D2RQ Platform - Treating Non-RDF Databases as Virtual RDF Graphs. Nat Prec <https://doi.org/10.1038/npre.2011.5660.1>
- Page RM (2018) Ozymandias: A biodiversity knowledge graph. bioRxiv <https://doi.org/10.1101/485854>
- Penev L, Dimitrova M, Senderov V, Zhelezov G, Georgiev T, Stoev P, Simov K (2019) OpenBiodiv: A knowledge graph for literature-extracted linked open data in biodiversity science. Publications 7 (2): 38. <https://doi.org/10.3390/publications7020038>
- Polovina J, Balazs G, Howell E, Parker D, Seki M, Dutton P (2004) Forage and migration habitat of loggerhead (*Caretta caretta*) and olive ridley (*Lepidochelys olivacea*) sea turtles in the central North Pacific Ocean. Fisheries Oceanography 13 (1): 36-51. <https://doi.org/10.1046/j.1365-2419.2003.00270.x>
- Popović M, Vasić N, Koren T, Burić I, Živanović N, Kulijer D, Golubović A (2020) Biologer: an open platform for collecting biodiversity data. Biodiversity Data Journal 8: e53014. <https://doi.org/10.3897/bdj.8.e53014>
- Ram A, Jalal S, Jalal A, Kumar M (2010) A density based algorithm for discovering density varied clusters in large spatial databases. International Journal of Computer Applications 3 (6): 1-4. <https://doi.org/10.5120/739-1038>
- Retana M (2013) Fine-scale breeding habitat preferences of southern elephant seals (*Mirounga leonina*) on Península Valdés. Aquatic Mammals 39 (3): 233-240. <https://doi.org/10.1578/am.39.3.2013.233>
- Sala JE, Quintana F, Wilson R, Dignani J, Lewis M, Campagna C (2011) Pitching a new angle on elephant seal dive patterns. Polar Biology 34 (8): 1197-1209. <https://doi.org/10.1007/s00300-011-0981-6>
- Sander J, Ester M, Kriegel H-, Xu X (1998) Density-based clustering in spatial databases: The algorithm gdbcscan and its applications. Data Mining and Knowledge Discovery 2 (2).
- Sievert C, Iannone R, Allaire J, Borges B (2022) flexdashboard: R Markdown format for flexible dashboards. 0.6.0.
- Sokolov S, Rintoul S, Wienecke B (2006) Tracking the Polar Front south of New Zealand using penguin dive data. Deep Sea Research Part I: Oceanographic Research Papers 53 (4): 591-607. <https://doi.org/10.1016/j.dsr.2005.12.012>
- Treasure A, Roquet F, Ansorge I, Bester M, Boehme L, Bornemann H, Charrassin J-B, Chevallier D, Costa D, Fedak M, Guinet C, Hammill M, Harcourt R, Hindell M, Kovacs K, Lea M-A, Lovell P, Lowther A, Lydersen C, McIntyre T, McMahon C, Muelbert M, Nicholls K, Picard B, Reverdin G, Trites A, Williams G, de Bruyn P J N (2017) Marine mammals exploring the oceans pole to pole: A review of the MEOP Consortium. Oceanography 30 (2): 132-138. <https://doi.org/10.5670/oceanog.2017.234>

- Villanueva R, Chen ZJ (2019) ggplot2: Elegant graphics for data analysis (2nd ed.). Measurement: Interdisciplinary Research and Perspectives 17 (3): 160-167. <https://doi.org/10.1080/15366367.2019.1565254>
- Wilkinson M, Dumontier M, Aalbersberg I (2016) The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 3: 160018. <https://doi.org/10.1038/sdata.2016.18>
- Xie Y, Allaire JJ, Grolemund G (2018) R Markdown; The definitive guide. 1st Edition. Chapman and 360 Hall/CRC, New York, 338 pp. [ISBN 9781138359444] <https://doi.org/10.1201/9781138359444>

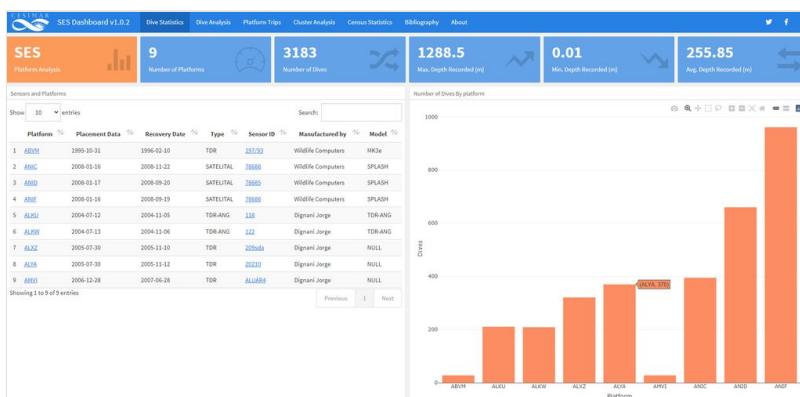


Figure 1.

SES dashboard main screen with a view of the reports of diving and census data.

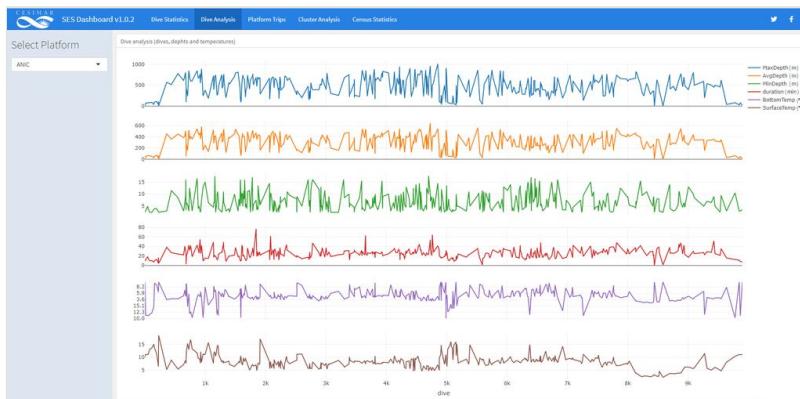


Figure 3.
Diving analysis view.

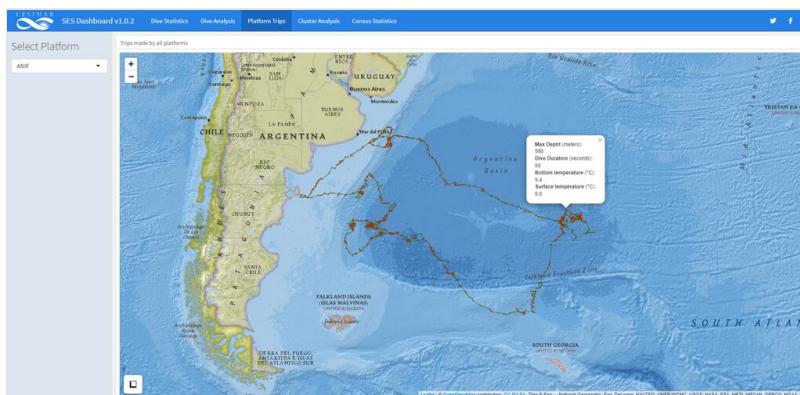


Figure 4.

Trip analysis view and the associated information at a particular point of the trip.

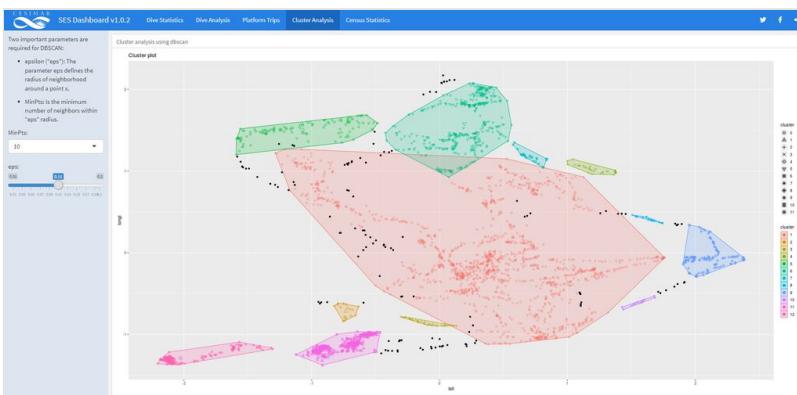


Figure 5.

View for spatial cluster analysis. All SES's trajectories are used.



Figure 6.

Census view. The bar chart shows the total SES counted grouped by categories. On the rightside, all the categories are observed, and the user can choose which category to show to facilitate interpretation.

Papers associated with each platform		Search:	
Show:	entries	DOI	%
Platform	Title		
1 ABL	A COMPARATIVE ANALYSIS OF THE VISION AND MISSION STATEMENTS OF INTERNATIONAL ENVIRONMENTAL ORGANISATIONS	https://doi.org/10.21373/10903727/922869	
2 AMU	DEVELOPING AN OCEAN ETHIC: SCIENCE, UTILITY, AESTHETICS, SELF-INTEREST AND DIFFERENT WAYS OF KNOWING	https://doi.org/10.11111/1520-1779.2008.01057.x	
3 AMU	SOCIAL EFFECTS OF SPACE AVAILABILITY ON THE BREEDING BEHAVIOR OF ELEPHANT SEALS IN PATAGONIA	https://doi.org/10.1002/ade.2395.0073	
4 AMU	TOWARD A NEW DATA STANDARD FOR COMBINED MARINE BIOLOGICAL AND ENVIRONMENTAL DATASETS: EXPANDING OBIS BEYOND SPECIES OCCURRENCES	https://doi.org/10.3897/BL1.e10988	
5 AMU	THERMOREGULATORY BEHAVIOR OF SOUTHERN SEA LIONS AND ITS EFFECT ON MATING STRATEGIES	https://doi.org/10.1183/16853889000205	
6 AMU	REPRODUCTIVE BEHAVIOR OF SOUTHERN SEALS	https://doi.org/10.1161/1560839800001029	
7 AMU	GROWTH AND DISTRIBUTION OF A SOUTHERN ELEPHANT SEAL COLONY	https://doi.org/10.11111/1748-7929.10053.x	
8 AMU	GROUP RAIDS: A MATING STRATEGY OF MALE SOUTHERN SEA LIONS	https://doi.org/10.1002/ade.2395.0054	
9 AMU	EQUAL INVESTMENT IN MALE AND FEMALE OFFSPRING IN SOUTHERN ELEPHANT SEALS	https://doi.org/10.11111/1469-7929.12070.00749.x	
10 AMU	GROUP BREEDING IN SEA LIONS: PUPS SURVIVE BETTER IN COLONIES	https://doi.org/10.1002/ade.2395.0014.0	
11 AMU	BREEDING BIOLOGY OF SOUTHERN ELEPHANT SEAL IN PATAGONIA	https://doi.org/10.11111/1560-8398.00424.x	
12 AMU	DIVING BEHAVIOR AND FORAGING LOCATION OF FEMALE SOUTHERN ELEPHANT SEALS FROM PATAGONIA	https://doi.org/10.2307/1381315	
13 AMU	POST-BREEDING DISTRIBUTION AND DIVING BEHAVIOR OF ADULT MALE SOUTHERN ELEPHANT SEALS FROM PATAGONIA	https://doi.org/10.1002/ade.2395.0003.x	
14 AMU	TEMPERATURE AND DEPTH PROFILES RECORDED DURING DIVES OF ELEPHANT SEALS REFLECT DISTINCT OCEAN ENVIRONMENTS	https://doi.org/10.1101/2021.07.05.500033.3	
15 AMU	MOVEMENTS AND LOCATION AT SEA OF SOUTH AMERICAN SEA LIONS (OTARIA FLAVESCENS)	https://doi.org/10.1007/s00383-000-0285	
16 AMU	DETECTING BIOLUMINESCENCE WITH AN IRIDIANCE TIME-DEPTH RECORDER DEPLOYED ON SOUTHERN ELEPHANT SEALS	https://doi.org/10.11111/1469-7929.2003.00281.x	
17 AMU	SOUTHERN ELEPHANT SEAL TRAJECTORIES FRONTS AND EDDIES IN THE BRAZIL-MALVINAS CONFLUENCE	https://doi.org/10.1004/1.2006.08.015	
18 AMU	DEEP DIVERS IN SHALLOW SEAS: SOUTHERN ELEPHANT SEALS ON THE PATAGONIAN SHELF	https://doi.org/10.1002/ade.2395.00008	
19 AMU	ENTANGLEMENT OF SOUTHERN ELEPHANT SEALS IN SQUID FISHING GEAR	https://doi.org/10.11111/1748-7929.2007.00105.x	

Figure 7.

Documents related to SES. The links are established through the DOI that identifies the publication.

Table 1.

Technical details.

URL	http://linkeddata.cenpat-conicet.gob.ar
SPARQL	http://linkeddata.cenpat-conicet.gob.ar/sparql
SNORQL	http://linkeddata.cenpat-conicet.gob.ar/snorql
Dump Data	http://dx.doi.org/10.17632/5nv5c7575w.3
VoID	http://linkeddata.cenpat-conicet.gob.ar/dataset
Licensing	CC0 1.0

Table 2.

Predefined queries to explore the data set.

Query	Link
Sensors associated with each platform	Query 1
Number of dives made by each platform	Query 2
Temperatures sampled by the ABVM platform	Query 3
Locations, depths and dates registered by ABVM	Query 4
Papers associated with each platform	Query 5
1990 census grouped by SES categories	Query 6