Hidrografico + . Portuguese Hydrographic Institute Marine Spatial Data Infrastructure

Hidrográfico +. Infraestructura de Datos Espaciales Marinos del Instituto Hidrográfico Portugués

Paulo Nunes, Sara Ameida, Sérgio Saraiva

REVISTA **MAPPING**
Vol. 29, 202, 14-21
septiembre-octubre 2021
ISSN: 1131-9100

Abstract

We are now beginning a new decade with a great number of challenges for marine geospatial data producers. At global scale The United Nations identify the needs for "Adaptation strategies and science-informed policy responses to global change" and elected the ocean knowledge a priority for next years. Geospatial Data is extremely important to understand the real world, for decision-making and to measure policy implementation goals.

The Portuguese Hydrographic Institute is a public organization with several roles: national hydrographic office, marine observation and marine technology national laboratory and a Portuguese Navy unity. As marine data and knowledge producer his internal data management processes are kept in line with national and European data policies and information sharing legal requirements. The Hydrographic Institute is in all aspects a data driven organization. To address the geospatial data needs for improvement the Portuguese Hydrographic Institute starts to build a new Marine Spatial Data Infrastructure through the Hidrográfico + project who granted funding from SAMA2020 program (POCI-02-0550-FEDER-035422).

This paper presents the Hidrografico + Marine Spatial Data Infrastructure building process. This asset is a big step forward and will be an important tool for the next ocean knowledge decade.

Resumen

En el año 2021 comenzará una nueva década con una gran cantidad de desafíos para los productores de datos geoespaciales marinos. A una escala global Las Naciones Unidas identifican las necesidades de «Estrategias de adaptación y respuestas políticas fundamentadas en la ciencia al cambio global» y eligieron el conocimiento del océano como una prioridad para los próximos años. Los datos geoespaciales son importantes para comprender el mundo real, para la toma de decisiones y para medir los objetivos de implementación de políticas.

El Instituto Hidrográfico Portugués es una organización pública con varias funciones: servicio hidrográfico nacional, laboratorio nacional de observación y tecnología marina y una unidad de la Armada Portuguesa. Como productor de datos y conocimiento marino, sus procesos internos de gestión de datos se mantienen en consonancia con las políticas de datos nacionales y europeas y los requisitos legales de intercambio de información. El Instituto Hidrográfico es en todos los aspectos una organización basada en datos. Para responder a las necesidades de mejora de datos geoespaciales, el Instituto Hidrográfico Portugués comenzó el desarrollo de una nueva Infraestructura de Datos Espaciales Marinos - el proyecto Hidrográfico + que otorgó fondos del programa SAMA2020 (POCI-02-0550-FEDER-035422).

Este artículo presenta el proceso de construcción de la Infraestructura de Datos Espaciales Marinos Hidrográfico +. Este sistema representa un gran paso adelante y será una herramienta importante para la próxima década del conocimiento de los océanos.

Keywords: Geographic Information Systems, Geomatics, Geospatial Webservices, Marine Spatial Data Infrastructure, INSPIRE.

Palabras clave: Sistemas de Información Geográfica, Geomática, Servicios web geoespaciales, Infraestructura de Datos Espaciales Marinos, INSPIRE.

Instituto Hidrográfico
antunes.nunes@hidrografico.pt
sara.almeida@hidrografico.pt
DEIMOS Engenharia
sergio.saraiva@deimos.com.pt

Recepción 12/12/2019
Aprobación 23/12/2019

1. INTRODUCTION

We are now beginning a new decade with lots of challenges for marine geospatial data producers. At global scale United Nations (UN) identify the needs for "Adaptation strategies and science-informed policy responses to global change" and elected the ocean knowledge a priority for next years.

Geospatial Data is extremely important to understand the real world, for decision-making and to measure policy implementation. The European Union still strongly commit to build a common Strategy for Data to thrust the data economy. This strategy will define the ways and means to build a digital single market and a data-driven society.

After the Industrial era we are now living on the Digital Era. This is a process with deep impacts on our way of live and our own coexistence with technology and other human beings. Technology and digital data are shaping our world and rising new societal and economic models.

The data, information and knowledge presents a new horizon of opportunities in several domains: economy, environment sustainability, policy, decision making, etc. The OCDE Rethinking Innovation for a Sustainable Ocean Economy affirms - the ocean observational data for sure returns great economic and society outcomes however the true impacts estimation is very difficult to quantify (OCDE, 2019, p. 163).

The UN identify a set of goals for a sustainable future – Sustainable Development Goals (SDG). With this strategic approach human kind recognize how far and deep is important grant access to high quality validated data to measure the current anthropogenic impacts and policy driven changes/outcomes.

The recently approved Open Data Directive (DIRECTIVE (EU) 2019/1024) is a legal instrument to foster the re-use of public sector data by sharing the information free and openly in machine-readable format through standardized Application Programming Interfaces (APIs). INSPIRE directive is designed to support environment sustainability and management at regional level. Those regulation and geospatial standards are the main guidelines for Public Sector Spatial Data Infrastructures (SDI) Implementation.

The Portuguese Hydrographic Institute (IH) is a public organization with the following roles: national hydrographic office, marine observation and marine technology national laboratory and a Portuguese Navy unity. As marine data producer his internal data management processes are kept in line with national and European data policies for data sharing.

As national marine research laboratory is involved in several marine data sharing projects to build and keep European marine data infrastructures. Runs in-situ observation sensors network and keep update long data time series for oceanographic variables.

Data should flow in this complex ecosystem of multiple sources and clients. For achieve this goal, data needs to be stored, manage, harmonized and shared in common standards. In all aspects IH is a data driven organization.

The stakeholder needs and the external geospatial data governance/regulation trends forced an internal data management chain status analysis. This thinking process help us to answer if the organization is fit for a future in a data sharing environment and help organization to find and clearly identify geospatial data management weaknesses.

To address the geospatial data needs for improvement IH starts to build a Marine Spatial Data Infrastructure (MSDI) through Hidrográfico + project who granted funding from SAMA2020 program (POCI-02-0550-FEDER-035422). This project, is a software development project with a single overall goal – build an integrated MSDI. SW development team combine multiple commercial and open source applications (e.g. Oracle, PortgreSQL + Postgis, ArcGIS Server, GeoServer, ncWMS Server and Geonetwork) to connect multiple data users to IH geospatial data through a controlled, centralized and integrated environment. The MSDI has a central geoportal where users can search, visualized and access the geospatial datasets. Humans and machines users could access directly to data services by linking their applications to the OGC webservices endpoints. IH wants to provide a good level of service to their MSDI users. The level of service is continuously monitor by a built in analytics and metrics system. The MSDI are now online with several data layers accessible and several capabilities still under development.

2. PORTUGUESE HYDROGRAPHIC INSTITUTE MARINE DATA MANAGEMENT

Data and Information management is an organizational activity with the aim to keep information available for users in a timely way. This resource management should be optimized as good as all other organization resources processes: finance, human resources, infrastructures, etc.

IH recognizes the importance of their marine data as an asset for the national future development.

Mapping ocean geophysics is an expensive activity dependent from ships, advanced oceanographic sensors and infrastructures. This is one and maybe the strongest reason for ocean still the last frontier of knowledge. All data collected in ocean are valuable and all are an important piece in the puzzle for a more general understanding of the ocean and its role in the balance of the Earth system. We live in a blue planet and the need for deeper ocean understanding are pushing us forward. The First global integrated marine assessment (FGIMA) (United Nations, 2015) published in 2015 presents the knowledge gaps concept – "much is known about much of the ocean, but nowhere do we have the detailed knowledge desirable for the future management of human use of ocean" (United Nations, 2015, p. 52). The Mission Starfish 2030: Restore our Ocean and Waters (European Comission, 2020, p. 19) show up the same issue. The document rises up this new old problem – the society still have "significant gaps in our understanding and knowledge" about ocean services and systems. This could be summarized in one single sentence - "no measurement, no prediction, no management".

Environmental wealth management requires continuous data streams. Real time access to data offers a clear view and awareness about real phenomenon and allow us to predict the future events. For ocean systems modelling and marine ecosystems monitoring blue data is a keystone. However, scientists and decision makers keep referring the lack of ocean knowledge.

This conclusion has been the motivation for the Ocean Science Decade (2021-2030) (United Nations, 2019)

Naturally societal goals for marine sustainability relies on accurately data, information and knowledge. The way for a better future is by "increase scientific knowledge, develop research capacity and transfer marine technology [...] in order to improve ocean health (United Nations, s.d.)".

During last years a significant investment are made off in marine data infrastructures, however the "knowledge gap" keeps rising as a problem. This seems to be a paradox and rises several questions. One of the most important and old one is - How can we reduce the knowledge gap and optimize the re-use of data for society future benefit? The Spatial Data Infrastructures are normally elected as one of the main tool for better geospatial data exploration.

The Marine Strategy Framework Directive (MSFD) have been publish in 2008. That legal document gives a strong recommendation for all state members invest in the development of their own blues strategies oriented to sustainable marine environment principles. The MSFD recognize the importance of the ocean environment as the building blocks for the blue economy.

The 2020 Blue Economy report presents the 2018 main stats. On that year blue economy in EU countries represents 5 million of employees and a Gross Value Added (GVA) of 218.3 billion euros (Comissão Europeia, 2020a, p. 15; 2020b).

The OECD Report – The Ocean Economy in 2030 estimates a generalized growth in all blue sectors (OCDE, 2016, p. 32). Measurement and management of blue domain will require data. Naturally hydrographic and geospatial data (blue data) are now and will keep their role as keystone for marine spatial planning and marine uses and needs management (Figure1).

The International Hydrographic Organization (IHO) elects the Hydrographic information driving marine knowledge as 2019's motto. The IH gets its current organizational configuration in 1960, since then IH collected data from lots of scientific campaigns in national maritime space of interest (Figure 2). All data stored in IH's databases are a valuable resource for scientific projects with a high value and economic potential for privates and public.

IH runs a national observation network (Figure3). This coastal network has several in-situ sensors: waveriders, multiparametric buoys, high frequency radars for surface currents and tide gauges stations. This system of sensors is the infrastructure part of the MONIZEE - National Exclusive Economic Zone monitoring pro-

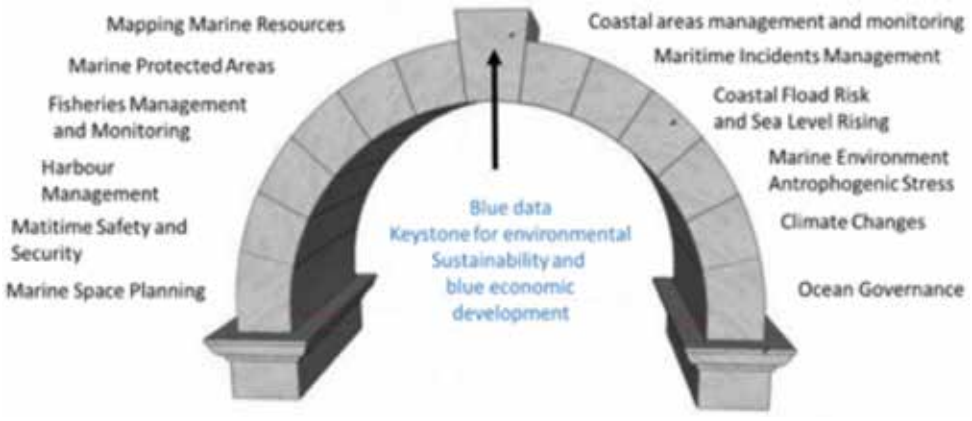

Figure 1. Marine Geospatial Data as keystone for blue economy (adapted from Harper (2016))

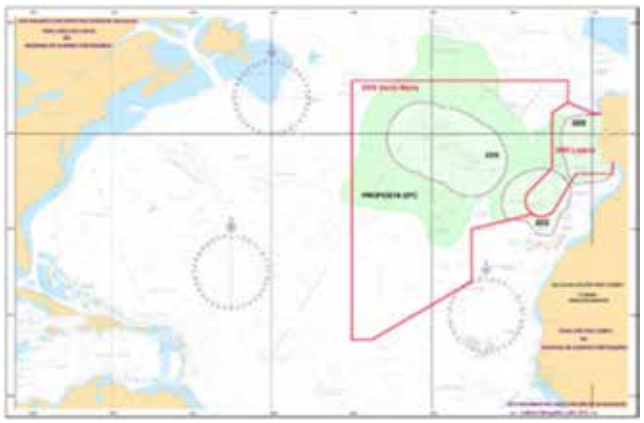

Figure 2. Portuguese maritime legal spaces

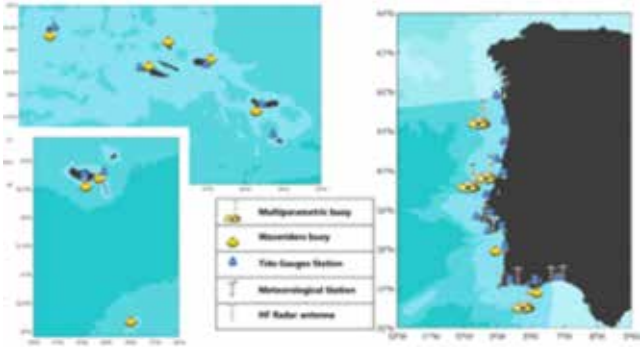

Figure 3. IH monitoring network (MONIZEE)

Figure 4. IH Strategic Map (Soares, 2020, p. 90)

Figure 5. Universal Hydrographic Data Model S-100 (OHI, 2020)

gram. This networks generates a continuous flux of near real time data arriving to IH staging areas where it is processed before data store ingestion. This in-situ data are very important to numeric forecast models calibration and to understand oceanographic coastal environment. Data is continuously analysed by quality check processes to find outliers and correct them. Intelligent systems, analytical algorithms and artificial intelligence are computational techniques with lots of potential for big data analysis, pattern detection and classification. For this reason, analytic approaches are one of the IH research field.

IH is an active geospatial data producer. Since the very beginning is involved in the main European marine data sharing infrastructures development projects: INSPIRE, SeaDataNet, EMOdnet, etc. At global level IH follows the IHO principles and the IOC/IODE data management best practices. IH continuously seeks efficiency, interoperability and high quality standards for data management.

Figure 4 shows the IH strategic plan (Soares, 2020, pp. 77-94) where we can see all the main programs align in order to support institute legal mission accomplishment. The IH strategic map are organized in the core projects inside mapping, observation and forecast programs. Those projects are supported by a transversal data program: IDAMAR. IDAMAR (*Infraestrutura de Dados do Ambiente Marinho*) supports others by offering them geospatial data management services and geospatial data systems development. The third one is the Sensors and Ocean Technology program – IHSENSORTECH.

Data sharing are one of the marine spatial data infrastructures driven. This requirement combines with the OGC web services and application programming interfaces (API) will increase machine to machine data share. The IHO are developing the S-100 hydrographic data model (Figure 5) to support the maritime data services and eNavigation initiative. (OHI, 2020)

Figure 6. SDI pillars (IHO MSDI-WG, 2017)

Aware of the importance of marine data resource, IH is strongly committed to make metadata and data reusable and shareable under FAIR principles (Wilkinson, Dumontier, Aalbersberg, & et. al., 2016) – Findability, Accessibility, Interoperability e Reusability. In order to make data available IH are developing a Data Policy align with European and National legal obligations and best practices requirements. The data policy should be a strategic document for the use of information resource who will balance the open as possible principle with the need for a sustainability of marine data infrastructures.

Figure 7. Hidrografico+ project development plan

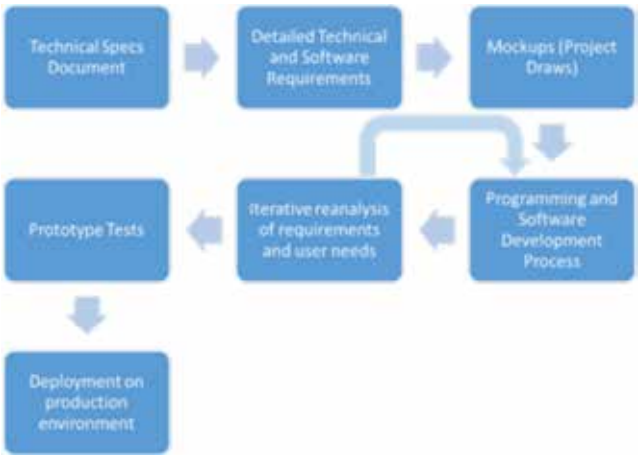

Figure 8. Project management model

3. METHODOLOGY

Conceptually a Marine Spatial Data Infrastructure (MSDI) implements the main pillars of spatial data infrastructures (Figure 6), all principles for interoperability and "FAIRness".

MSDI could be seen as a subdomain of spatial data infrastructures specialized in blue geospatial data management. The Hidrográfico + project have been organized as an information systems project (Figure 7). The goal for the project was defined at the very begging - "get a single access point for all geospatial information produced by IH". For accomplish that goal the project team starts by collect the stakeholders needs, legal requirements and data sources.

After the requirements definition cycle the IH data management team collects all technical requirements for the Hidrográfico + MDSI and transposed them to the MSDI Technical Specification Document.

All single system blocks are developed under the same framework. The IT development team uses the technical specs document as starting point to define detailed software development requirements. After that produce some mockups about interfaces and functionalities, this project draws were present to client in a brainstorming way trying to find lacks and needs. After mockups approval the development stage begins. The software project team implements an agile workflow with permanent reanalysis and revisit of user needs. After some sprints the software component is mature and ready for tests and to be deployed in the production environment (Figure 8).

So far this model of project management seems to be fit for our purpose and allow project team to align the stakeholders needs, legal obligations and data sources specific formats to serve geospatial data in one single access point. Until now the project integrates

Figure 9. Hidrográfico + MSDI geoportal interfaces

Figure 10. Functionalities of Hidrográfico + - MSDI

several data layers, some served as OGC web services, others using specific purpose API. One of the main obstacles has been the dimension of some datasets as time series and how to interact with them in an efficient way.

4. RESULTS

The Hidrográfico + MSDI started the software development phase in 12 of December 2019. Since first of August 2020 the MSDI portal is online accessible at URL - https://geomar.hidrografico.pt. Figure 9 present some of the data interfaces already available at the geoportal.

The geoportal is still under development with periodic deployment of new functionalities and data layers. So far the users recognize the potential of this new tool as an organizational asset capable of support diverse scientific projects and clients – Human and machines.

Until now it was possible to implement the functionalities presented in the Figure 10.

One of the main purposes of this project was publish validated datasets online findable and accessible. In the next stage, the goal will be the harmonization process to convert the current dataset models to INSPIRE data models. This will require specific project but the most important is get datasets well documented available and accessible since now. The MSDI is based on open source technology and have been developed in a modular approach. This requirement potential the evolution of the infrastructure and is possible to insert or remove data services in an easily way.

The Hidrografico + platform can be extended with thematic modules, built on top of the data services available. These thematic modules can have a scientific, public or commercial nature, and be integrated inside the platform, or available in outside applications, developed by third party partners, leveraging the importance and value of the data services provided by IH, by creating new consumers of these data. Examples under development include modules for tides, scientific campaigns, and sea trip planning, to be available shortly, inside the Hidrografico + platform. Other modules can be added in the future, as the platform grows in terms of data and functionality.

The Hidrografico + platform currently serves free data, for visualization, and download. Some of the data is not available for download, because in the future will require a paid access to download that data.

The data sharing capability of the platform among the geo and EO (Earth Observation) community is also underway. Currently, there are plans to incorporate the IH free data into the European Comission (EC) funded

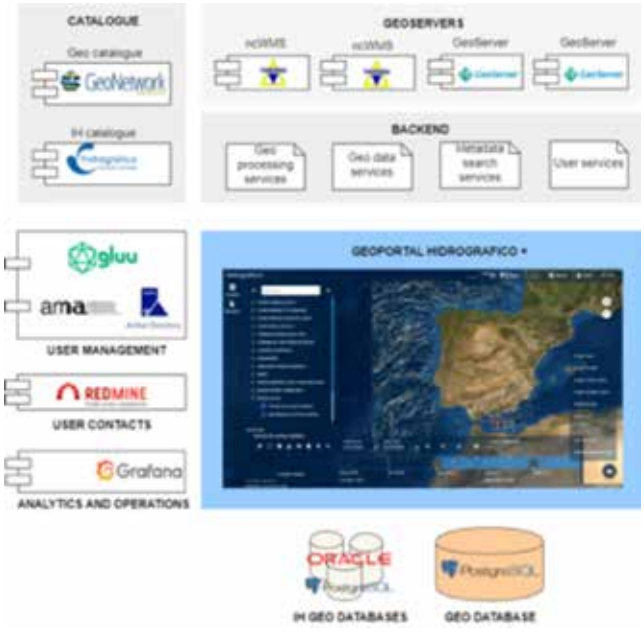

Figure 11. Hidrográfico + MSDI Archicteture

NextGEOSS catalogue (https://nextgeoss.eu/), a European contribution to GEOSS (Global Earth Observation System of Systems), which consists of a next generation European data hub and cloud platform, for EO data, where the users can connect to access data and deploy EO-based applications. The concept revolves around providing the data and resources to the user communities, together with cloud resources, seamlessly connected to provide an integrated ecosystem for supporting applications.

Another platform consuming the Hidrografico + data services, starting by free data, will be the European Space Agency (ESA) funded ECOMI (E–COMmerce platform for MIcro geoservices) platform, based on the store4EO platform (https://www.store4eo.com/), which aims to became a marketplace connecting EO service providers and users of such services. The platform aims to facilitate the delivery of innovative geo-services to various industries, the public sector, and the general public, while reinforcing that the use of EO services.

Putting all this together, the Hidrografico + is rapidly becoming a consolidated platform of geographic ocean data, by making easily available, the value of the data produced by IH, to a number of different community of users, and also and by integrating into a network of other geospatial European initiatives, where all those stakeholders can work together for the benefit of all.

The MSDI geoportal is support for integrated infrastructure of services and application servers as shown in the Figure 11.

Data search is implemented through the MSDI metadata catalogue developed using GeoNetwork application server. The geoportal data layers are dynamically updated using a real time backend service who continuously harvests the metadata catalogue to link the user to all data services available endpoints. This catalogue unifies the data servers (the several GeoServer and ncWMS), proving a consolidated metadata profile to the several OGC data services delivered by the servers, and also implementing the search capabilities to the data published in the GeoPortal.

The OGC geospatial services standards implementation is done using multiple GeoServer application servers. Time series netCDF data services are implemented using ncWMS servers, which are also linked to the GeoServer servers, to allow serving the data by multiple OGC protocols The platform is designed to easily add new specialized GeoServers to the ecosystem, to easily scale the data publishing capabilities.

Several data layers are maintained on a PostgreSQL database, taking advantage of the PostGIS extension, which adds support for geographic objects to the PostgreSQL object-relational database, providing a native support for storing and processing geographic data.

To support the MSDI management the Development team needs to implement Gluu server authentication processes based on external public administration authentication services and Internal Active Directory. This user authentication and authorization server, also provides the mechanisms to control the access to the data, because, some data or details of the data, might be only available for a certain group of users, and this server provides these functionalities.

For helpdesk/service desk the Redmine was the option taken. This server centralizes all the interactions between IH and the users, in terms of contact requests, issues, messages to be shown to the users from IH.

The Grafana application server is used for analytics and infrastructure monitoring indexes calculation. With this server, the platform becomes manageable on real-time, offering several dashboards and alerts about the activity of the entire infrastructure and applicational usage (for example, which services are requested more often, which have slower response times, etc.)

Finally, all these application servers, are aggregated into the GeoPortal, which is the main point of access to the users access the data and services.

5. FINAL REMARKS

Hidrográfico + MSDI components are used to support internal technical and scientific geospatial data management processes and to feed external clients with near real time data. This MSDI supports the organization mission as marine national laboratory, hydrographic chart authority. As navy unity supports maritime operations extremely depend of geospatial data.

Hidrográfico + since the internet deployment have been proved their value. The IH develops this project taking in account marine user needs. This MSDI should be mature at the Ocean Decade Beginning and we have great expectations about their role for decrease the ocean knowledge gap at National, European and Global level. This MSDI implements all identified requirements needed for interoperability. For sure will potentiate the access to IH blue geospatial data and have potential for private sector applications development. By this way IH will manage one infrastructure ready to be used in the Open Data Directive umbrella and align with INSPIRE requirements. This is one of the IH contributions for Ocean Decade, sustainable development goals and for national blue economy development.

The MSDI development is a never endless process. Digital era is a living ecosystems with rapidly and continuous development. For sure new requirements for digital data should show up soon. Yet, at this moment our main goals have been achieved.

ACKNOWLEDGMENT

The "Hidrográfico +" project gets funding from FEDER -POCI-02-0550-FEDER-035422.

Authors acknowledge to all the Portuguese Hydrographic Institute and Deimos colleagues which by one way or another give a contribution to the Hidrográfico + MSDI Project.

REFERENCES

Comissão Europeia. (2020a). *2020 Blue Economy Report: Blue sectors contribute to the recovery and pave way for EU Green Deal*. Retrieved from https://ec.europa.eu/commission/presscorner/detail/en/ip_20_986

Comissão Europeia. (2020b). Blue indicators online dashboard. Retrieved from https://blueindicators.ec.europa.eu/access-online-dashboard_en

European Comission. (2020). *Mission Starfish 2030: Restore our Ocean and Waters*. Retrieved from https://ec.europa.eu/info/publications/mission-starfish-2030-restore-our-ocean-and-waters_en

Harper, S. (2016). [UKHO International Hydrographic Development Projects].

IHO MSDI-WG. (2017). Publication C-17 - Spatial Data Infrastructures "The Marine Dimension" - Guidance for Hydrographic Offices". In: OHI.

OCDE. (2016). *The Ocean Economy in 2030*. Retrieved from https://www.oecd.org/environment/the-ocean-economy-in-2030-9789264251724-en.htm

OCDE. (2019). *Rethinking Innovation for a Sustainable Ocean Economy*. Retrieved from http://www.oecd.org/publications/rethinking-innovation-for-a-sustainable-ocean-economy-9789264311053-en.htm

OHI. (2020). S-100 Universal Hydrographic Data Model Retrieved from https://iho.int/en/s-100-universal-hydrographic-data-model

Soares, C.-a. C. V. (2020). *O Conhecimento científico do Oceano. Instituto Hidrográfico, conhecer o mar para que todos o possam usar*. Cadernos Navais, (57).

United Nations. (2015). *Summary of the first global integrated marine assessment*. (A/70/112). Retrieved from https://undocs.org/A/70/112

United Nations. (s.d.). SDG 14 - Conserve and sustainably use the oceans, seas and marine resources for sustainable development Retrieved from https://sdgs.un.org/goals/goal14

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., & et. al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*.

About the authors

Paulo Nunes
Navy officer and hydrographer at the Portuguese Hydrographic Institute. With more than 10 years of experience in hydrographic geospatial information management. Is the head of technical data management department responsible for Marine Spatial Data Infrastructure and scientific data life cycle management. Paulo has a Master in Geographic Engineering by Science Faculty of Lisbon University and an Hydrographic Category A Certificatied by International Hydrographic Organization.

Sara Almeida
Responsible for Portuguese Hydrographic Institute data management life cycle organizational standards and procedures. She has more than thirty five years of experience in marine data and information management, statistical analysis and programming processing algorithms for oceanographic data. Sara has a M.Sc. degree in Applied Mathematics (Operational Research), 1993 by Instituto Superior de Economia e Gestão of the Technical University of Lisbon, Portugal.

Sérgio Saraiva
Agile Coach and Technical Manager at DEIMOS, with nearly 20 years of experience in successfully delivering IT projects, in several countries, technical domains and industries, like space, aeronautics, maritime, energy, and others. Throughout his career, Sergio achieved several relevant project awards, including in the Geographic Information Systems space. Sergio is graduated in Computer Science by the Lisbon Technical University and a Project Management and Agile Practitioner certifications by the Project Management Institute.