Hydrographic open data for society

Datos hidrográficos abertos para a sociedade

REVISTA **MAPPING**
Vol.32, 211, 34-48
2023
ISSN: 1131-9100

Paulo Nunes, João Vicente, Ana Leonor Veiga, Cristina Monteiro,
Telmo Dias, Carla Palma, Miguel Neto

Abstract

Hydrography is defined as: "The branch of applied sciences which deals with the measurement and description of the physical features of oceans, seas, coastal areas, lakes and rivers, as well as with the prediction of their change over time, for the primary purpose of safety of navigation and in support of all other marine activities, including economic development, security and defence, scientific research, and environmental protection" (Publication S-32). Due to their broad concept hydrographic data and information provide the foundational bases for marine space management, fisheries, coastal environment, policy decisions, shipping, energy and oil industry, etc.

Portuguese Hydrographic Institute (IH) manages geospatial datasets from several scientific and technical domains. Data management has been on daily agenda and always has an internal priority. Facing the digital transformation tsunami and rapid evolution of society data requirements is the main driven for developing an internal sustainable open data strategy aligned with findable, accessible, interoperable and reusable (FAIR) principles and as open as possible.

The European Open Data and Open Science strategies combined with the need to fill the ocean knowledge gaps are changing the way how data producers deal with geospatial information.

This article presents several IH projects to increase sharing and reuse of hydrographic data by society.

Resumen

La hidrografía se define como: «la rama de la ciencia aplicada que trata sobre la medición y la descripción de las características físicas de los océanos, los mares, las zonas costeras, los lagos y los ríos, y también sobre la predicción de sus cambios en el tiempo, teniendo como objetivo principal la seguridad de la navegación y en apoyo de todas las otras actividades marítimas, incluyendo el desarrollo económico, la seguridad y la defensa, la investigación científica y la protección ambiental». (Publicación S-32). Debido a su concepto amplio, los datos y la información hidrográficos proporcionan las bases fundamentales para la gestión del espacio marino, la pesca, el medio ambiente costero, las decisiones políticas, el transporte marítimo, la industria energética y petrolera, etc.

El Instituto Hidrográfico Portugués (IH) gestiona conjuntos de datos geoespaciales de varios dominios científicos y técnicos. La gestión de datos ha estado en la agenda diaria y siempre tiene una prioridad interna. Enfrentar el tsunami de transformación digital y la rápida evolución de los requisitos de datos de la sociedad es el principal impulso para desarrollar una estrategia interna de datos abiertos sostenibles alineada con los principios FAIR y lo más abierta posible.

Las estrategias europeas de datos abiertos y ciencia abierta, combinadas con la necesidad de colmar las lagunas de conocimiento sobre los océanos, están cambiando la forma en que los productores de datos tratan la información geoespacial.

Este artículo presenta varios proyectos de IH para aumentar el intercambio y la reutilización de datos hidrográficos por la sociedad.

Keywords: *FAIR Principles, Hydrography, Marine Spatial Data Infrastructures, Ocean Decade, Open Data*

Palabras claves: Datos abiertos, Hidrografía, Infraestruturas de Dados Espaciales Marinos, Principios FAIR, Ocean Decade

Paulo Nunes. Instituto Hidrográfico.
antunes.nunes@hidrografico.pt
João Vicente. Instituto Hidrográfico.
delgado.vicente@hidrografico.pt
Ana Leonor Veiga. Instituto Hidrográfico.
leonor.veiga@hidrografico.pt
Cristina Monteiro. Instituto Hidrográfico.
cristina.monteiro@hidrografico.pt
Telmo Dias. Instituto Hidrográfico.
geraldes.dias@hidrografico.pt
Carla Palma. Instituto Hidrográfico.
carla.palma@hidrografico.pt
Miguel Neto. NOVA IMS.
mneto@noveims.unl.pt

DOI: https://doi.org/10.59192/mapping.393
Recepción 19/12/2022
Aprobación 27/01/2023

1. INTRODUCTION

Marine data stakeholders need to redefine their way of thinking in the geospatial data sharing process. The open movement is gathering major attention because of its promising benefits for society. However, marine high-quality in-situ data is still expensive to acquire. For this reason, it is critical to ensure the sustainability of data producers and mapping authorities due to their significant role as reference data keepers.

New sensors and sensing techniques are increasing the amount of data, but data produced by full automatic processes raises questions about quality, confidence and metadata completeness.

Marine data and information have wide applications for knowledge and economic development, e.g.: fisheries stock assessment and management, environmental monitoring and protection, navigation and trade, renewable energies and other resource development, infrastructure construction, security and defence, search and rescue, and scientific research and development.

It is widely recognized that data is a key ingredient for enhancing knowledge about the surrounding world. Thus, the frontiers of technology and data processing have always been pushed forward. In a digital data environment, remote sensing data, internet of things and machine learning techniques are all promising tools. But, none of them work completely without the human factor and they require highly skilled human resources.

Digital data market should be fair for consumers and producers as in any other exchange. This fact is obvious in business-to-business relations. However, in government-to-business relations, the users demand free and open data access, based on a taxpayer rights. This is a strong argument to increase public and open data access, but the European countries have diverse finance models, data sovereignties and copyrights models. In many countries, the state data producers still remain the only reference data producers in several domains However, they still depend on data products and services revenue to balance their budget and to keep their finances sustainable. Decision-makers need to keep in mind that data and information always have a cost to collect and make it accessible. Nothing is given for free, not even the web services with hiding business models that appear to be… free.

The critical question is - how to make public data accessible, even by law, and keep finances sustainability of public data producers without full budget support from public funds? This is a tricky question with multiple dimensions. Until this scenario changes, lots of public data producers need to keep some data products and services with full copyright property protection.

Quality Assurance and Quality control, guided by the best practices, continue to be the way to ensure the best reference data and require a significant human effort. Easy access to data induces the wrong feeling about the easiness for acquiring and processing those data. For sure, newly data analytics, algorithms, and data management systems came to make data analyst life easier. The amount of data is increasing exponentially day by day, the management and processing of big data is still a challenge, even for the most modern high-capacity systems. Keep data management on trail requires regular investment in personnel education and training, as in equipment. No matter how much data access seems easy, data collection, information and knowledge production keep their costs and long term sustainability of data infrastructures is still a real issue to deal with.

To take advantage of the digital revolution, to accelerate research and to engage the power of machine analysis at scale, while ensuring transparency, reproducibility and societal utility, data and other digital objects created by and used for research, and in all other activities, need to follow FAIR principles (see Figure 1) - findable, accessible, interoperable and reusable (FAIR) (European Commission, 2018).

FAIR principles are gathering attention from data communities, in all scientific and society domains, since the publication of the FAIR Guiding Principles for scientific data management and stewardship (Wilkinson, Mark D.; Dumontier, Michel; Aalbersberg, Jan IJsbrand; et. al., 2016) in 2016.

- Findable: users only use and reuse data if they can find them. The key elements to find data is the metadata (or rich metadata: completeness and the right descriptors use are fundamental), adopt persistent identifiers like digital object identifiers and implementing indexing mechanisms, all of those are simple principles to make data visible for users.

Figure 1. FAIR data approach (SciencesPo, n.d.)

- Accessible: Once the user finds the required data, they need to know how it can be accessed, possibly including authentication and authorisation. The licence and data access protocols description in metadata is the rule.
- Interoperable: The data usually needs to be integrated and combined with other data. Also, data needs to be interoperable between applications and systems. The use of interoperable formats and standard vocabularies are key elements.
- Reusable: The ultimate outcome for FAIR principles adoption is to optimise the reuse of data. To achieve this goal, metadata and data should be well-described. Data processing workflows should be referenced to community standards and best procedures. The provenance of data should be clear for users and the completeness of metadata should give no space for misinterpretation by users.

As noticed from previous definitions, FAIR data and open data are different concepts, although they have similarities and are dependent from each other. Open data is related with the license for use or reuse some data property – data belonging to someone, normally a data producer and the owner may exercise property rights over those assets or allow the use of those data by third parties.

FAIR data intends to potentiate the reuse of data by adopting the best principles to optimize search, finding, interoperability and reuse. This means that data can be FAIR, even if it stays with restrictive license or when it is accessible by everyone with an open licence to use and reuse (open data). It depends completely on the purpose of the data, where the data currently is in its lifecycle, and the end-usage of the data.

Making FAIR data a reality requires investment, but it is an investment with significant scientific benefits and economic returns. Numerous studies demonstrate the economic benefit and very strong value proposition of data repositories and data services (European Commission, 2018). Many authors focus primarily in the value of data openness (adopt open licences is different from provide data with no costs) but the truly value comes from implementation of the FAIRness of data combined with the use of the appropriate license model.

According to data.europa.eu - "Making data available as open data across the EU Member States is vital to leverage its potential for the European society and economy, for example, to enrich research, inform decision making, or develop new products and services. The impact of open data is mainly realised through application and depends on factors like costs, quality of the data and its documentation, or the modality of access. To further increase the impact of open data and reduce market entry barriers for start-ups and small and medium-sized enterprises (SME), these factors need to be addressed. To achieve this most effectively, efforts should target those datasets that have the biggest potential for society and the economy" – the High-Value Datasets (HVD). (European Commission, 2020).

HVD is one of the most important concepts defined in the Open Data Directive (ODD). It is expected to generate important benefits for society by open HVD data themes. Those impacts will affect the environment and the economy. The Open Data Directive currently defines six categories of HVDs: geospatial, earth observation and environment, meteorological, statistics, companies and company ownership, mobility (European Parliament, 2019).

The European strategy for data intends to leverage a single market for data that will ensure Europe's global competitiveness and data sovereignty. Common European data spaces will ensure that more data becomes available for users, while keeping the companies and individuals who generate the data in control (EU, 2020). Between 19 February and 31 May of 2020, the European Commission launched an online consultation about European Strategy for Data. The report and findings can be accessed online (European Commission, 2020). The report highlights several conclusions and potentials responses, some of which were transposed for Open Data and the re-use of public sector information also known shortly as ODD. The ODD is a legislative act that sets out goals that all Member States must achieve, while it does not provide specific obligations on how to reach those goals.

It is up to the Member States to transpose the Directive into national law in order to make the objectives, requirements and deadlines directly applicable.

The ODD mandates the Commission to develop an implementing act laying down a list of specific HVD that Member States will be obliged to publish.

However, till October 2022 the implementing act is still to be published and the official list of HVD has not been released. The identification of HVDs is a challenge by itself, as explained in the Analytical Report 15 – High-value datasets: understanding the perspective of data providers (European Commission, 2020).

The ODD introduces the following principles in its legal text (Government of Ireland, 2020):

1. Dynamic data and use of APIs;
2. Introduction of Implementing Acts to improve access to HVD;
3. Scope expanded to include public undertakings;
4. Amendment to charging principles;
5. Prohibition of exclusive arrangements.

HVD is the most important concept in ODD and the open movement is growing day by day. In all countries the public data producers are working to correspond at market demands and the Portuguese Hydrographic Institute developed a data policy in line with all these questions.

Amongst the benefits resulting from open data initiatives we can refer four main types of impact (Young & Verhulst, 2016): (i) create economic opportunity, by enabling business creation, job creation, new forms of innovation, and stimulating economic growth; (ii) help to solve complex public problems, by improving situational awareness and preparedness, increasing knowledge and expertise to bear on public problems, and by allowing policymakers, civil society, and citizens to better target interventions and track impact; (iii) improve governance, by enhancing transparency and accountability, improving service delivery and efficiency gains, and increasing information sharing within and outside city domains; and (iv) empower citizens, by improving their participatory capacity, and by acting as a catalyst for social mobilization.

The European Open Data and Open Science strategies are changing the way as Data, Information, Knowledge and Wisdom (DKIW) producers deal with geospatial information. We are living a geospatial data revolution. The open data principle is becoming an increasingly important part of the data revolution, which is recognized worldwide as a key engine for achieving the post-2015 UN Sustainable Development Goals (Petrov, Gurin, & Manley, 2016).Large quantities of data are collected and stored by various groups of public and private agencies all over the world for a wide variety of purposes (Contarinis, Pallikaris, & Nakos, 2020).

The acquisition of marine data is an example of this. The IHO Data Centre for Digital Bathymetry (DCDB), EMODNet, SeaDataNet, Framework of Ocean Observing (FOO), Global Ocean Observing System (GOOS) are examples of European and international systems designed to store and share marine data. Due to their importance, marine data should be properly managed and made available in formats useful for all scope of potential uses and users.

Marine spatial data theme is used to reference: hydrographic data, oceanographic data, land and coast data, meteorology and climate data, etc. (Contarinis, Pallikaris, & Nakos, 2020).

United Nations announced the Decade of Ocean Science for Sustainable Development to mobilize the scientific community, policy-makers, business and civil society around a programme of joint research and technological innovation. A fundamental goal of this initiative should be to enable governments to make informed policy decisions about their marine resources.

Hydrographic data and information are essential for the safe, efficient and sustainable conduct of every human activity that takes place in, on or under the sea. Hydrography is inherent to the three dimensions of the sustainable development of the oceans, ensuring that the marine environment is respected and that no adverse economic or social impact is incurred (IHO, 2019).

As every human activity conducted on sea environment relies on knowing the depth and the nature of the seafloor, equally important are tides and currents data, hydrography is an essential foundation to the development of the Blue Economy (IHO, 2019). As it is claimed by Loew, T. (2019) "Hydrographic information drives marine knowledge".

Traditionally, the use of hydrographic information was restricted to the production of nautical charts. Currently, mapping the seabed is not only important for safety of navigation but also for other domains, such as: scientific research, sustainable management, decision support, among others (Dias, 2021).

In the end, hydrographic data contributes to the development of ocean and coastal economic activities, and geographic information systems provide the means to extract the required value. Hydrographic data is the foundation for building a maritime data management system, in the framework of a Marine Spatial Data Infrastructure (MSDI) for broader use (Ponce, 2014).

2. MARINE SPATIAL DATA INFRASTRUCTURE (MSDI)

Spatial Data Infrastructures (SDI) are defined by SDI Cookbook as "the relevant base collection of technologies, policies, and institutional arrangements that facilitate the availability of and access to spatial data". Inside the SDI definition we find the requirements to implement the Findable, Accessible, Interoperable and Reusable (FAIR) principles. Spatial Data Infrastructure (SDI) provide a framework for organising geographic data, metadata, tools and users with rules, relationships and standards. A Marine SDI (MSDI) is not meant to be separated from other SDIs but is complementary in the coastal zone and oceans. Hydrographic data and information is the basis of MSDIs (Ponce, 2014).

The publication C-17 from International Hydrographic Organization (IHO) defines MSDI as the element of an SDI that focuses on the marine input in terms of governance, standards, Information Communication and Technology (ICT) and content.

MSDI allows "wider appreciation in terms of the way a variety of data types might be combined for efficient analysis by a wide range of disciplines, such as spatial planning, environmental management and emergency response. This requires data to be held in a generic way, rather than for a particular product for a limited user group or for a specific purpose. An MSDI is not a collection of hydrographic products, but an infrastructure that promotes interoperability of data at all levels". (IHO, 2017).

MSDI are vital for the spatial management of coastal and marine areas, as they contain about 40 % of the world's population (within 100 kilometers of the coast) (CIESIN, 2012) and 90 % of the catches from marine fisheries (Barbier, 2017). Coastal zones also reveal higher rates of population growth and urbanization (Neumann, et al., 2015).

The value of geospatial information increases as it becomes discoverable and usable, hence helping national and local economies. In short, MSDIs (Natural Resources Canada, 2018):

- Provide a simple, efficient, expandable and transparent governance solution;
- Facilitate provision of targeted marine applications, tools and services;
- Promote access, visibility, and delivery of value-added products;
- Leverage existing federal and international initiatives.

Hydrography has a vital role in MSDI in providing core "reference" data (such as bathymetry, maritime boundaries, coastline and geographic areas and names). After all, Hydrography is the branch of applied science which deals with the measurement and description of the physical features of oceans, seas, coastal areas, lakes and rivers, as well as with the prediction of their change over time. It does this firstly for the purpose of safety of navigation but also plays a crucial role in the support, through its data and information resources, of all other marine activities, including economic development, security and defense, scientific research, and environmental protection (IHO, 2017).

Geospatial information should be integrated with any other meaningful data to solve societal and environmental problems, and act as a catalyst for economic growth and opportunities, and to understand and take benefit from a nation's development priorities and the Sustainable Development Goals. United Nations recognizes its importance at international level. The United Nations Integrated Geospatial Information Framework (UN-IGIF) defines a strategy plan based on 9 pathways to help world countries to maximize the economic and social impact of geospatial data. The UN-IGIF offers a strategic guidance documentation with the basis and guide for developing, integrating, strengthening and maximizing geospatial information management and related resources in all countries.

The IGIF (UN-GGIM, n.d.) comprises three parts:

Part 1 is an Overarching Strategic Framework: sets the context of 'why' geospatial information management needs to be strengthened and why it is a critical element of national social, economic and environmental development. It focusses on the role of geospatial information in the digital age and how that information is critical to government functions at all levels.

Part 2 is an Implementation Guide: describes 'what' specific guidance, options and actions can be undertaken by Member States to strengthen their geospatial information management. Expanding on each of the nine strategic pathways, the Implementation Guide provides the roadmap for implementing the IGIF.

Part 3 is a Country-level Action Plan: is specific to each country and details 'how' the guidance, options, and actions recommended in the Implementation Guide will be carried out, when and by whom. IGIF develops the processes, resource materials, templates and examples that are available and that are helpful to first develop a national action plan.

3. DATA POLICY OF PORTUGUESE HYDROGRAPHIC INSTITUTE

The Portuguese Hydrographic Institute (IH) has been an organization ahead of its time. Since the beginning, it has accompanied the conceptualization of data infrastructures and has progressively been updating its processes to meet the needs of the users of its products: the Navy, navigators, researchers and citizens in general (IH, 2021). The combination of marine spatial data (Figure 2 shows some examples of marine data themes manage by IH) for efficient analysis also supports some of the world's current major challenges: development of a sustainable blue economy, e-navigation, emergency planning and response, climate change and its impacts on sea level rise and ocean acidification, and Marine Spatial Planning (MSP).

The IH as a National Hydrographic Service and State Laboratory, implements processes, computer algorithms and information systems for data management that make it possible to store, process, analyse, preserve and share this data with third parties (IH, 2022).

The sharing process is ensured through data portals, by the publication and dissemination of scientific articles, as well as through the data transfer service, complying with the quality standards of the various scientific disciplines and with the main national and international standards and good practices (Nunes, 2022).

The IH participates in the effort and benefits from the regional and international infrastructures for sharing marine geospatial information, which have been assuming a leading role with the technological evolution operated on the internet. The growing sharing of geospatial information, following open formats and licenses, contributes directly and indirectly to the economic growth of the country and to national, regional and global scientific development (IH, 2022).

The IH has closely followed the evolution of SDIs. To address the geospatial data needs for improvement, IH started to build a new MSDI through the Hidrográfico+ project, which was designed following the structuring principles of the SDIs and integrates, in its architecture, the guiding principles of the International Hydrographic Organization (IHO) and the European directives: INSPIRE and the ODD (Nunes, Paulo; Saraiva, Sérgio; Almeida, Sara; Veiga, Leonor, et. al., 2022).

The MSDI frontend is a centralized webGIS – Hidrografico +[1]. In the portal, users will find several marine and hydrographic datasets: environment observations at sea, forecasts, nautical charts and hydrographic information. The MSDI implements several functionalities for users (see Figure 3).

From the systems architecture point of view the IH MSDI implements several open source technologies. This allows IH to be compliant with INSPIRE and IHO MSDI requirements and to publish diverse data sources and formats through web services (IHO, 2022).

The main challenges will keep the MSDI aligned with digital data strategies at different levels. This is a digital environment with a continuous evolution who requires a rapidly adaptation to new clients and stakeholders. Data harmonization and development of S-100 based web services will be for sure a challenge for future (IHO, 2022).

The OGC services are the ground base for share and reusing data among several systems. IH invested in the OGC web services portfolio and new web applications for data access.

Next sections present several examples about services and products for society aligned with the open data movement. They are examples of projects in which the IH is involved, which portray its commitment in sharing data and information for the generation of knowledge in the field of marine sciences.

3.1 Electronic Nautical Chart Web Map Service

The Hidrografico+ marine geospatial data and information infrastructure aims to be the only web access point to IH data. At the beginning of 2022, a new Web Map Service (WMS) (see Figure 4) was implemented which allows the publication of nautical cartography available in the S-57 (IHO, 2000) format with the symbology associated

[1] https://geomar.hidrografico.pt

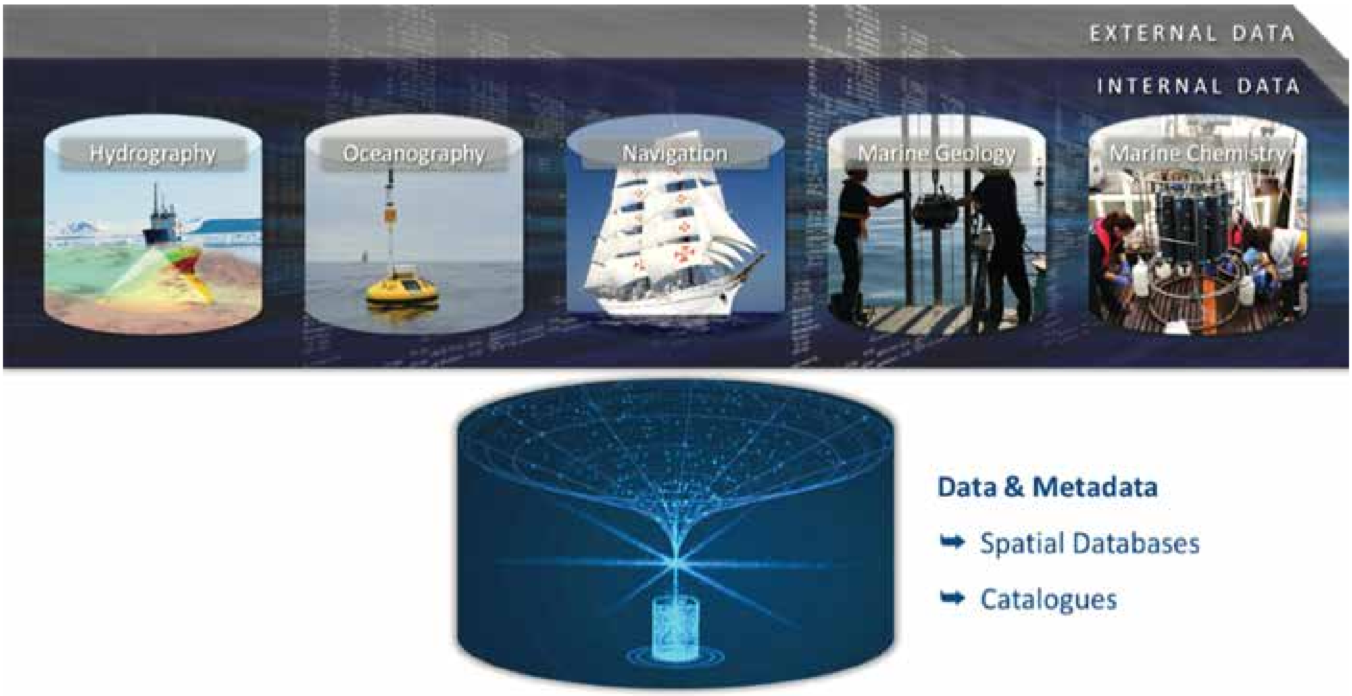

Figure 2. Data themes managed by Portuguese Hydrographic Institute

Figure 3. Hidrografico+ MSDI functionalities and components

with the electronic navigation chart - S-52 (IHO, 2014).

By accessing to this resource[2], users access the electronic cartography basemap with national coverage, through standardized formats of the Open Geospatial Consortium (OGC) WMS. The Electronic Nautical Charts (ENC) and Inland Electronic Nautical Charts (IENC) are the most updated chart information for all Portuguese maritime waters, also available for Douro and Guadiana inland waterways. These charts are accessible free of charges for non-commercial purposes by public and private users. The access could be done in a seamless way only by uploading a WMS url in GIS web or desktop systems. This basemap is of great interest for marine spatial planning activities, for planning the installation of infrastructures at sea, for hydraulic studies and many others applications.

3.2 SEAMAP 2030 – Bathymetric Open Data

The SEAMAP 2030 program[3], Mapping the Portuguese Sea, aims to provide bathymetric coverage of the seabed in high resolution (using multibeam sounders) from 30 meters deep to the edge of the EEZ or to the outer edge of the extended continental shelf (legal) when and under the terms in which the Portuguese claim is recognized by the United Nations (Soares, 2020).

This program aims to contribute to the worldwide knowledge of the seabed morphology, completing the high-resolution mapping of the maritime spaces by 2030 and is part of a global context that recognizes and understands the importance of the oceans and their increasing value in the future (Dias, 2021).

[2] The server is accessible with DNS enc.hidrografico.pt and the service properties can be viewed through a GetCapabilities request in the browser using the url: https://enc.hidrografico.pt/?SERVICE=WMS&VERSION=1.3.0&REQUEST =GetCapabilities.

[3] https://www.hidrografico.pt/iprojeto/16

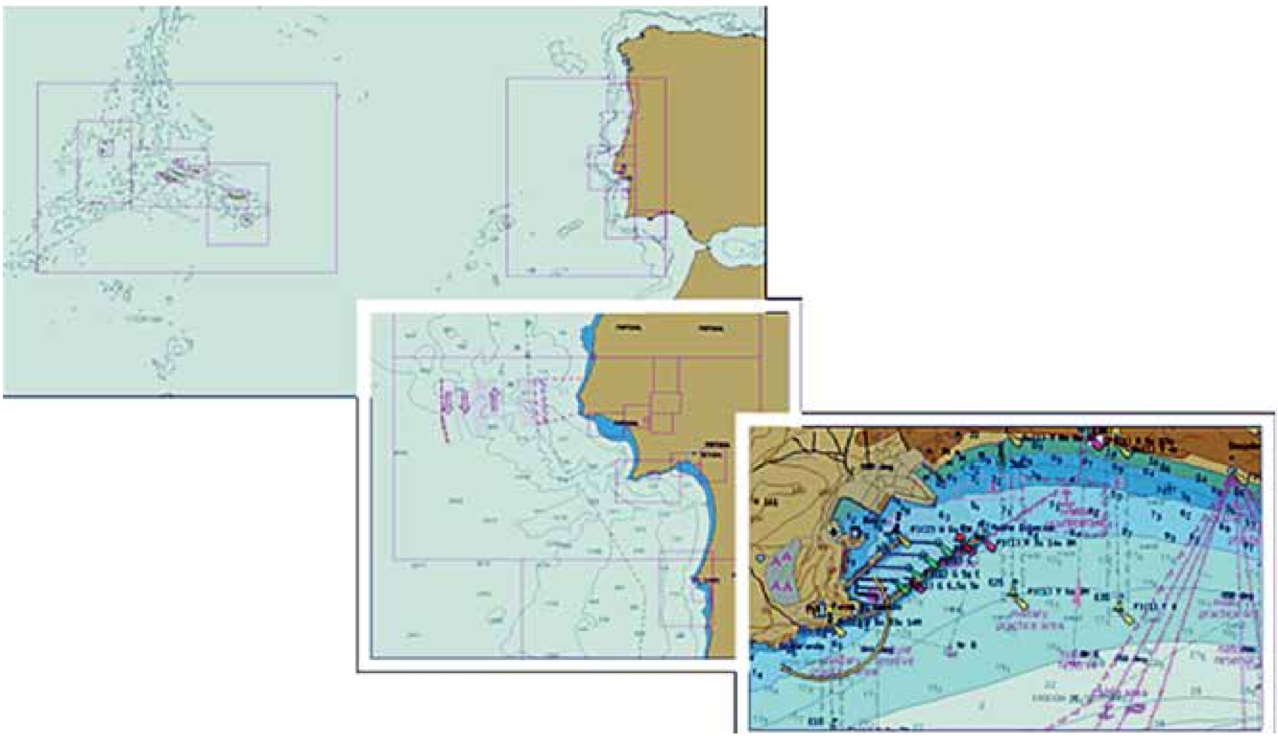
Figure 4. ENC Web Map Service

High-resolution seafloor mapping campaigns in the Portuguese waters started in 2005, in an effort to acquire the necessary data to support the continental shelf extension claim. In the following years the hydrographic surveys were continued with less frequency and in 2017, aligned with the Seabed 2030 project[4], the SEAMAP 2030 program was launched (Dias, 2021).

More recently, the IH has been requesting bathymetric data from other national and foreign institutions, such as data acquired by scientific cruises in national waters. These data, after being validated, also become part of the SEAMAP 2030 database, contributing to the increase in national bathymetric coverage.

The SEAMAP 2030 program has revealed to be more than just hydrographic surveys to complete seafloor mapping (Dias, 2021). Specific products have been developed, such as bathymetry grids to be made available free of charge through the geospatial data and information infrastructure Hidrográfico+. This service[5] provides bathymetric grids for the Atlantic sea in an area specified by the Portuguese jurisdictional waters.

Through this service, the IH offers gridded bathymetric data for users interested in the topography of the seabed. This data provides the depth of the seafloor in meters and is downloadable in predefined areas (see Figure 5). Bathymetric grids were built with resolutions according to Table I which are updated with new information at least once a year.

Although the bathymetric grids available are based on the best available bathymetry and every effort is made with respect to the quality control, the data products available from this service are not to be used for navigation or for any other purpose related with the safety of navigation.

In a very simple and friendly way, the user has access to the depth grids by choosing the specific tile and by filling out a form requesting data (see Figure 6).

Then the user will receive the data by email, as well as the associated metadata (see Figure 7).

In addition to this functionality, this service gives access to a layer with the main structures of the submarine relief. It is possible to have a 3D visualization of the main submarine canyons, hydrothermal vents, banks, seamounts, cliffs, plains and other underwater structures in the Portuguese sea.

3.3 GUAD20 Project – Guadiana Natural Navigable Heritage

The GUAD20 project was launched in 2017 with funding from the European Regional Development Fund (FEDER) within the scope of the Portugal-Spain Cross-border Cooperation Operational Program (POCTEP INTERREG V-A 2014-2020), and in addition to the Portuguese Hydrographic Institute had as partners the Public Agency of Puertos de Andalucía (APPA - Promoting Entity) and the General Directorate of Natural Resources, Security and Maritime Services (DGRM).

Interreg Europe's aim was to help regional and local governments across Europe to develop and deliver better policy. The Interreg program is one of the main instruments for economic, social and territorial cohesion in the European Union. Created as a community initiative in 1990, Interreg became a formal objective of Regional Policy in 2000, identified as European Territorial Cooperation (ETC) whose overall objective is to promote harmonious economic, social and territorial development throughout the Union[6]. The ETC began with the promotion of cooperation between cross-border regions of Member States, later extended to transnational and interregional cooperation.

[4] https://seabed2030.gebco.net. Seabed 2030 is a collaborative project between the Nippon Foundation and GEBCO with the goal of the complete mapping of the world's oceans by 2030. The effort is also compiling all bathymetric data into the freely available GEBCO Ocean Map.

[5] https://gridmar.hidrografico.pt/

[6] https://eurocid.mne.gov.pt/interreg

Table I – Resolution in meters according to depth range

Depth Range (m)	Resolution (m)
50-250	32
250-1000	64
1000-2000	128
2000-4000	256
4000	512

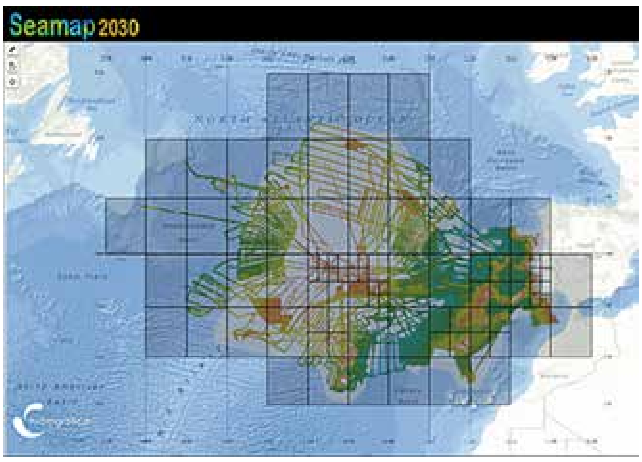

Figure 5. Gridded bathymetric data available at https://gridmar.hidrografico.pt/

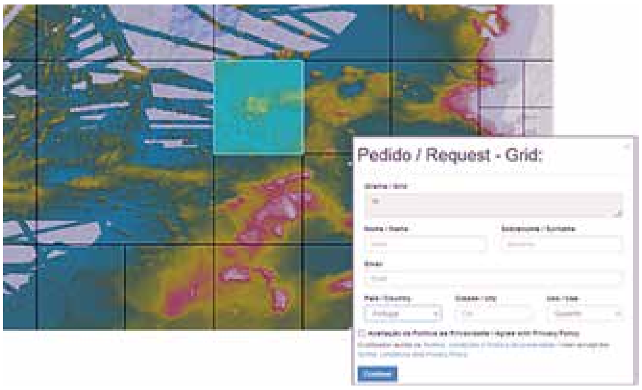

Figure 6. Access and Request depth grids

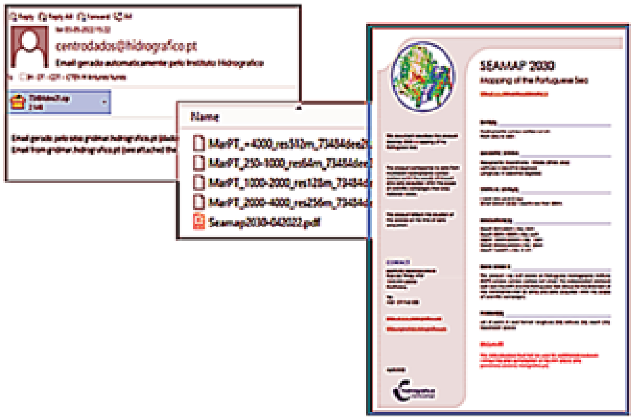

Figure 7. Received data and metadata

The GUAD20 project aimed to rehabilitate the navigability of the Guadiana River on the international section between Vila Real de Santo António and Pomarão (see Figure 9), in safety conditions, as well as to rehabilitate some existing port infrastructures.

Both countries are focused on promoting this natural and cultural asset, which must be boosted with safety and control requirements, while also preserving the environmental values present throughout the Guadiana valley.

The interventions already carried out and the ones planned for the Guadiana River, have strengthened the relations between Portugal and Spain. They also can stimulate economic development and job creation in the surrounding municipalities, by encouraging recreational and tourism nautical activities[7].

Within the scope of this project, several hydrographic surveys were carried out by the IH in the entire international section, while the physical and chemical characterization of sediments has been conducted between Alcoutim and Pomarão.

[7] https://www.sulinformacao.pt/2021/07/dgrm-participou-nas-jornadas-do-projeto-guad20-guadiana-patrimonio-natural-navegavel/

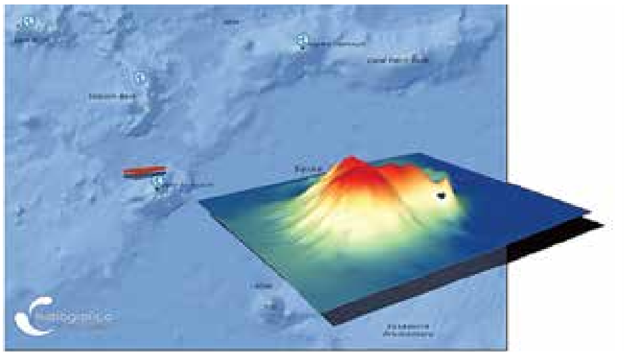

Figure 8. Submarine relief – 3D visualization

DGRM boosted the regularization of funds in the silting areas that were a danger to navigation, and from the combination of all the bathymetric information with the updated hydrographic data in the intervened areas, it was possible to readjust and redefine the route of the navigation channel (Alcoutim to Pomarão) and implement the respective signalling through the placement of beacons, lights and targets.

APPA proceeded with the construction and improvement of several existing port infrastructures in the international section. The project culminated with the publication of nautical cartography of the river series (paper and digital). Inland Electronic Navigational Charts (IENC), S-57 format, for the Guadiana Waterway are available on the IH website and can be downloaded for free[8].

3.4 Data analytics - from sensing data to smart data

Ocean decade identified the urgent need for better and broader scientific knowledge to support and drive effective policies and actions. Data Science field offers new tools to analyse ocean data in an efficient way. Algorithms based on supervised and unsupervised classification, for example, help data scientists to process large volumes of data. Ocean is a large space and an extreme environment for sensors. As explained before, the complete measurement of this space is impossible. This is why data analytics is so important for ocean knowledge. Advanced 3D interpolation based on geostatistics, like Empirical Bayesian Kriging 3D (EBK3D) (see Figure 5), gives new opportunities for ocean studies. With these statistical tools it is possible to compute complete 3D models of large ocean areas from in-situ point measurements (Figure 3 shows examples of analytical techniques use for ocean modelling).

Portuguese Hydrographic Office in assotiation with several research institutions (University of Lisbon and Politécnico de Leiria) developed the project

[8] https://www.hidrografico.pt/cart.guadiana

Figure 9. Guadiana river basin (adapted from APA (2015))

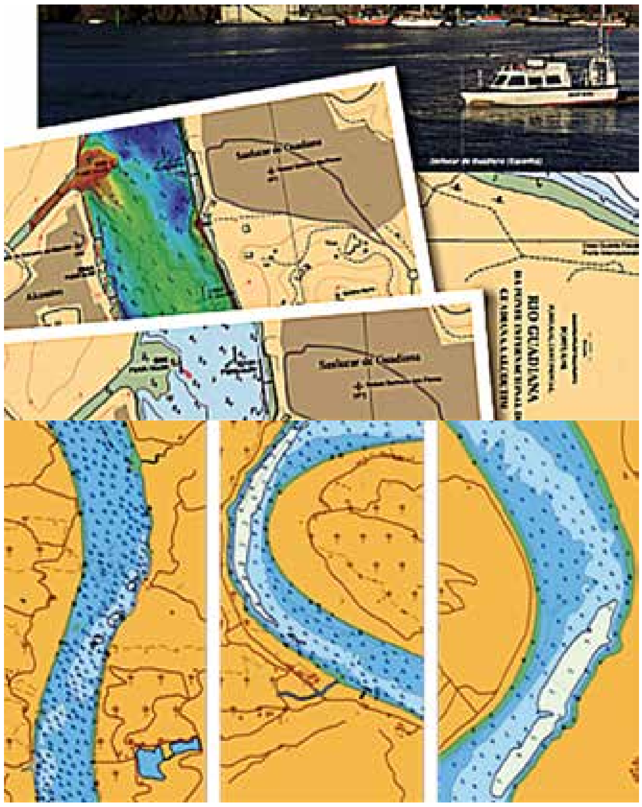

Figure 10. Extracts from Nautical Cartography (paper and electronic charts)

AQUIMAR – Caraterização geral das áreas aquícolas para estabelecimento de culturas marinhas funded by MAR 2020 program (Project MAR2020 nº MAR-02.01.01-FEAMP-0107 -https://aquimar.hidrografico.pt/).

This project has been designed to study areas designated for aquaculture in order to define their potential for different kinds of fish farms establishments. Data has been collected using profiles with CTD-Rosette system. This results in profile sampling grids. With this collections of 3D points the Portuguese Hydrographic Institute developed a methodology to compute a 3D continuous model, based on Empirical Bayesian Kriging 3D (see Figure 13). Models will be available by Web Services (OGC WMS and JSON) to be consumed by external users. This is an example of the benefits from using of the analytical tools.

From the close cooperation between the Portuguese Hydrographic Institute and the NOVA IMS – Information Management School, a national reference school in information management and analytics, results the project MarIA (https://mar-ia.pt) - Plataforma Colaborativa de Modelos de Inteligência Artificial para o Mar (funded by SAMA 2020 - POCI-05-5762-FSE-000400). This project is an innovation centric project, designed to explore new innovative approaches to build data/information products and services for blue economy sectors (from analytical models to augmented reality mobile apps, as show in Figure 14, one of the project expected outcome). Amongst the MarIA project expected outcomes is a set of analytical algorithms to solve real world problems.

The Portuguese Hydrographic Institute is organizing digital routing information to build an augmented reality mobile application for mariners. Also an hackathon is being organized in the context of the project to bring together researchers and promote innovation by inviting data scientists to solve marine related problems affecting the society.

These are examples on how to use the national potential of data centric innovation at the service of the blue economic growth. MarIA will also reduce the innovation gap between the scientific community and the common users.

As last example, the High Performance Computing (HPC), namely the European Network offers resources to make possible to run high-resolution numeric models. This kind of infrastructure makes possible to improve earth and ocean modelling, simulate scenarios and apply Artificial Intelligence to very huge datasets. This improves the quality and resolution of models outputs and opens the pathway for the truly Digital Twin of the Ocean.

3.5 From digital ocean to digital twin of the ocean

Digital Twin of the Ocean (DTO) was on central stage at International Hydrographic Office 2022 Council and was described as a strategic path by the European

Figure 11. Data and analytical tools examples in Marine GIS domain

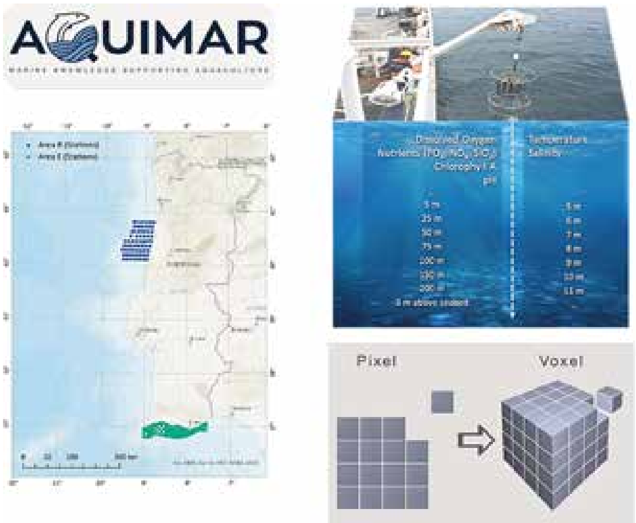

Figure 12. AQUIMAR Project data collection with CTD sampler

Commission. In the last years, we are experiencing a transformation from the concept of Digital Models of world features to true Digital Twins. This has been possible because digital transformation brings us new technological developments and new capabilities to process huge amounts of data, over large extensions, at high resolutions, like the High Performance Computing.

Digital Twins are an old concept with a new life in the ocean domain. The technological evolution gives us new capabilities to manage and process big data, to compute numeric models, to simulate multiple what-if scenarios, to develop decision support models and to predict future states based on artificial intelligence. Those information technologies resources and algorithms are tools to better observe, model and predict the ocean systems and environment dynamics. Now it is possible to get new insights about the impacts at multi scales and fill some of the ocean knowledge gaps.

What is a Digital Twin of the Ocean? The DTO is a system of systems designed to integrate a wide range of data sources in order to transform data into knowledge (see Figure 15). This complex system relies in the Marine Spatial Data Infrastructures to feed complex modelling techniques. The quality of the digital twin depends on the highly accurate real data and the accuracy of modelling solutions from mathematical, physical, statistics models.

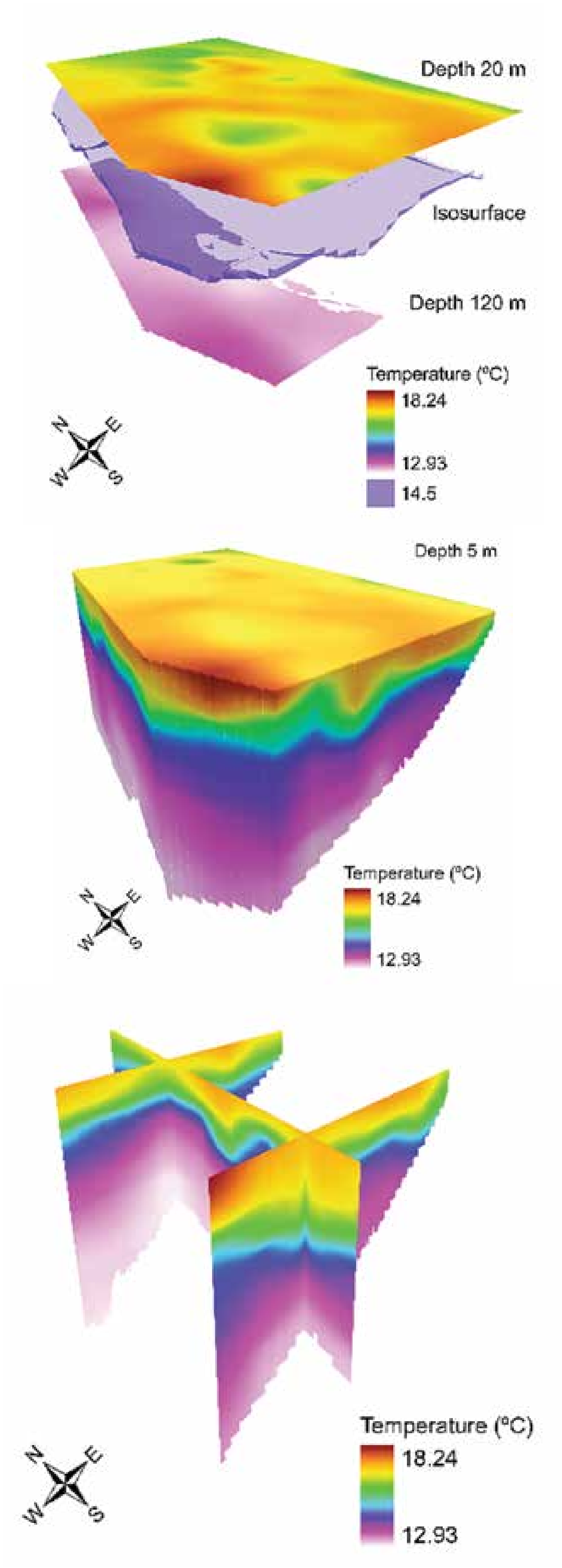

Figure 13. 3D ocean variables models from Empirical Bayesian Kriging 3D

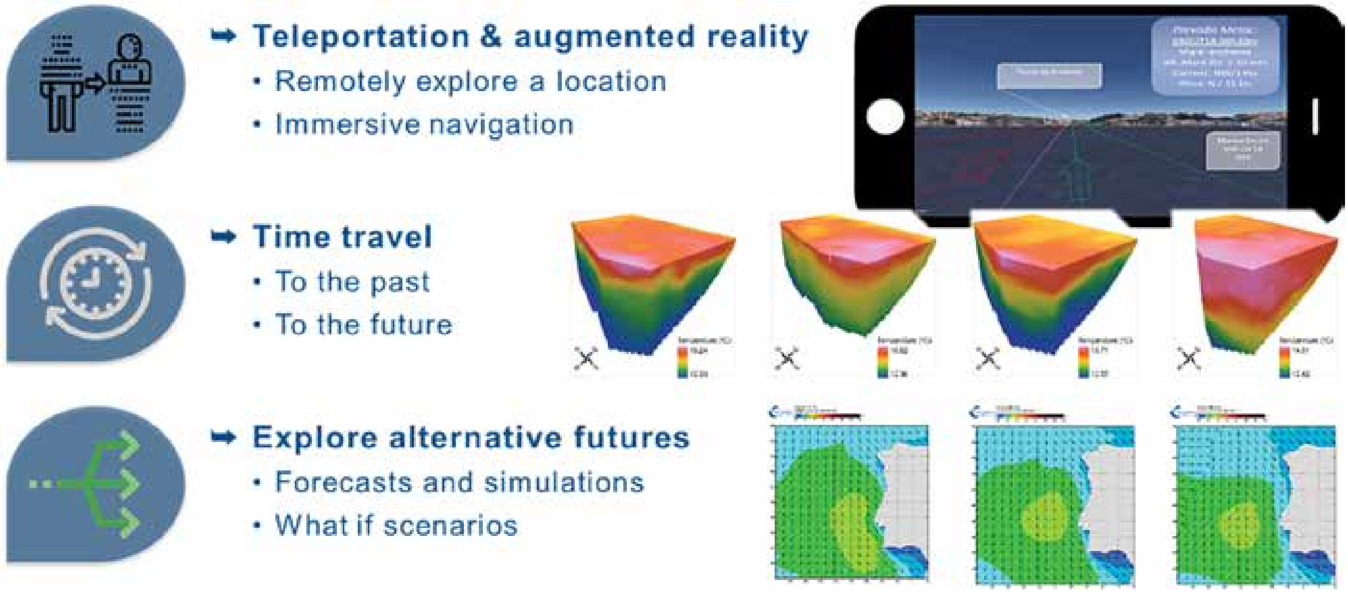

Figure 14. MarIA project analytical tools vision

The DTO needs to be as similar as possible to its real twin. This replica of the ocean is fed by continuous observations from sensors deployed over the world's oceans and from satellites flying over the Earth. DTO has also advance visualization tools and interactive decision support models to help users and policy makers to assess the impact of man made induced changes or resources consumption.

For hydrographic professionals and for the Hydrographic Offices around the world the Digital Twin of the Ocean offers an opportunity to improve the services available. This is especially important for the S-100 Universal Hydrographic Data Model implementation. For example, using the S-111 standard to gather surface current data for the development of dynamic hydrographic products could help eliminate the barrier posed by the use of different formats. This is the opinion expressed by Pierre Bahurel, director general of Mercator Ocean International at IHO Council (IHO, 2022) S-100 model is designed to transfer static and dynamic data. The S-111 is only one example of all kinds of data supported by S-100 who can benefit from DTO models.

IH has started to develop initial actions regarding DTO establishment, in order to improve its digital data services portfolio.

4 CONCLUSIONS

The world-wide community recognizes that data is a key ingredient to improve knowledge about the surrounding world. We are living the geospatial digital transformation with new data analytics algorithms and data management systems. Yet, the data management complexity does not stop to increase. The amount of data is increasing exponentially day by day, the management and processing of big data is still a challenge, even for the most modern high-capacity systems. Keep data management on trail requires regular investment in personnel skills and equipment. No matter how much data access seems easy, data collection, information and knowledge production keep their costs and long term sustainability of data infrastructures is still a real issue to deal with.

The "open data" principle is becoming the new reality in geospatial data world. The open movement is recognized worldwide as a key engine for achieving the post-2015 UN Sustainable Development Goals. In strait connection with the open movement we assist the growing importance of data FAIRness. This two core concepts are the main drivers for the Open Data Directive and INSPIRE Directive implementation.

Marine Spatial Data Infrastructures (MSDI), the branch of Spatial Data Infrastructures design to deal with marine data, has a significant economic and societal potential for coastal nations and the world in general. The Ocean is a worldwide ecosystem and the human actions in one side of the world could generate an impact on the other side. The river flow pollution and plastic debris released through rivers basins are accumulating in international waters – territory belonging to all human kind. The MSDI future is driven by the evolution of the International Hydrographic Organization's (IHO) S-100 data model for facilitating marine domain interoperability and the World Wide Web Consortium's (W3C) best practices for spatial data publishing on the Web.

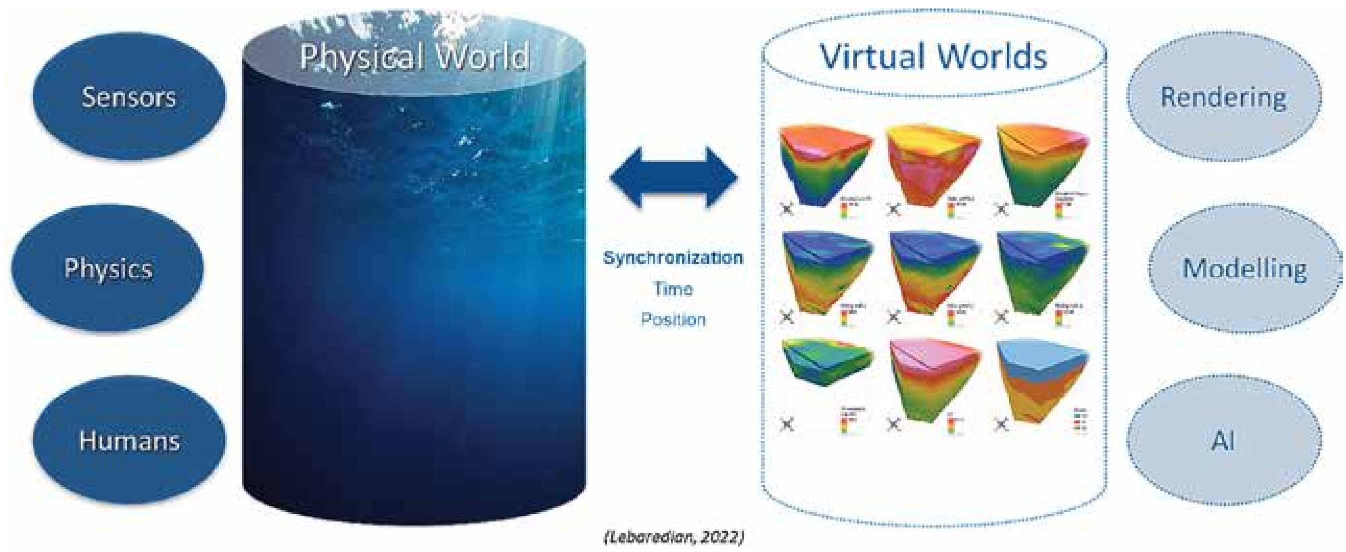

Figure 15. DTO concept in practice (adapted from Peckham (2022))

Hydrography and Hydrographic Offices have a vital role in MSDI in providing core "reference" data (such as bathymetry, maritime boundaries, coastline and geographic areas and names). A well designed MSDI deals with much more data than the hydrography. For example, the Portuguese Hydrographic Institute develops several projects in multiple scientific domains. All this valuable information needs to be proper managed and should be available in the right formats for the right users. Long goes the time when Hydrographic Data users were only mariners and nautical chart producers. Today, marine and hydrographic data are more than purpose-built products and services. Data and information have wider purposes, as they can be re-used, which represents a shift in the common thinking, inducing a data-centric model.

New analytical tools combined with information and technology tools like HPC open new fields of work in data management, information modelling, digital twins of the ocean, etc.

Portuguese Hydrographic Office (IH) has been developing a transversal program IDAMAR – Infraestrutura de Dados do Ambiente Marinho with the purpose to address the new reality and the 21st Century user's data demands. The program SEAMAP 2030, a national project aligned with the SEABED 2030 international program, is one example of the IH effort to unleash data in an open format. Making available for download the Inland Electronic Nautical Charts for Douro and Guadiana waterways are other examples on how IH is working to improve safety of navigation by unleashing free and open data.

However, as explained in this article, at this moment it is impossible to share all data with an open licence. IH depends on data products and services revenue to balance its budget and to keep its finances wealthy and sustainable. Data and information always have a cost to obtain. Highly advanced analytical tools are promising, but, none of them works without the human factor and they require highly skilled human resources. For sure, nothing is given for free, neither the web services with hiding business models and appearing to be free.

REFERENCES

APA. (2015). Caracterização e Diagnóstico da Região Hidrográfica do Guadiana. Plano de Gestão de Região Hidrográfica. Retrieved November of 2022, from https://apambiente.pt/sites/default/files/_Agua/DRH/ParticipacaoPublica/PGRH/2016-2021/3_Fase/PGRH_2_RH7_Parte2.pdf

Barbier, E. B. (2017). Marine ecosystem services. Current Biology, 27, R507-R510.

CIESIN. (2012). National Aggregates of Geospatial Data Collection: Population, Landscape, And Climate Estimates, Version 3 (PLACE III). Retrieved September 12, 2022, from https://bit.ly/3vIOH3H

Contarinis, S., Pallikaris, A., & Nakos, B. (2020). The Value of Marine Spatial Open Data Infrastructures—Potentials of IHO S-100 Standard to Become the Universal Marine Data Model. Journal of Marine Science and Engineering. doi:https://doi.org/10.3390/jmse8080564

Dias, T. (2021). PROGRAM SEAMAP 2030 – 100% of the Portuguese maritime spaces mapped by 2030. The International Hydrographic Review. Retrieved November of 2022, from https://ihr.iho.int/articles/program-seamap-2030-100-of-the-portuguese-maritime-spaces-mapof-the-portuguese-maritime-spaces-mapped-by-2030/

European Commission. (2018). Turning FAIR into reality - Final Report and Action Plan from the European Comission Expert Group on FAIR Data. Directorate General for Research and Innovation, Directorate B - Open Innovation and Open science, Brussels. Retrieved from https://ec.europa.eu/info/sites/default/files/turning_fair_into_reality_0.pdf

European Commission. (2020). Analytical Report 15 - High-value datasets: understanding the perspective of data providers. Directorate General for Communications Networks, Content and Technology, Unit G.1 Data Policy and Innovation. Retrieved November of 2022, from https://data.europa.eu/sites/default/files/analytical_report_15_high_value_datasets.pdf

European Commission. (2020). High-value datasets. Retrieved November of 2022, from https://data.europa.eu/en/publications/datastories/high-value-datasets

European Commission. (2020). Summary report of the public consultation on the European strategy for data. Brussels. Retrieved November of 2022, from https://digital-strategy.ec.europa.eu/en/library/summary-report-public-consultation-european-strategy-data

European Parliament. (2019). Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information. Official Journal of the European Union. Retrieved November of 2022, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32019L1024

Government of Ireland. (2020). The new Open Data Directive? What will it mean for public bodies? Retrieved November of 2022, from IRELAND'S OPEN DATA PORTAL: https://data.gov.ie/blog/the-new-open-data-directive-what-will-it-mean-for-public-bodies

IH. (2021). Hidrográfico+ - Nova Infraestrutura de Dados e Informação Geoespacial. Revista da Armada, 27-29.

IH. (2022). Política de Dados e Informação Técnico-Científica do Instituto Hidrográfico. Instituto Hidrográfico.

IHO. (2000). IHO Transfer Standard for Digital Hydrographic Data - Edition 3.1. International Hydrographic Organization. Retrieved November of 2022, from https://iho.int/uploads/user/pubs/standards/s-57/31Main.pdf

IHO. (2014). Specifications for Chart content and display aspects of ECDIS - Edition 6.1(.1). International Hydrographic Organization. Retrieved November of 2022, from https://iho.int/uploads/user/pubs/standards/s-52/S-52%20Edition%206.1.1%20-%20June%202015.pdf

IHO. (2017). Spatial Data Infrastructures "The Marine Dimension" - Guidance for Hydrographic Offices. Publication C-17. International Hydrographic Organization. Retrieved November of 2022, from https://iho.int/iho_pubs/CB/C-17_Ed2.0.0_EN.pdf

IHO. (2019). IHO Input to the Report of the UN Secretary General on Oceans and Law of the Sea. Retrieved November of 2022, from https://www.un.org/Depts/los/general_assembly/contributions_2019/IHOEN.pdf

IHO. (2022). IHO Eastern Atlantic Hydrographic Commision – National Report.

IHO. (2022). Transition to digital data services and other priorities discussed at IHO Council. Retrieved November of 2022, from International Hydrographic Organization: https://iho.int/en/transition-to-digital-data-services-and-other-priorities-discussed-at-iho-council

Lowe, T. (2019, July 4). IHO - Hydrographic Commission on Antarctica. Seminar on the status and the impact of hydrography in Antarctic waters. Item 2.3 - Hydrographic information drives marine knowledge. Czech Republic. Retrieved from https://iho.int/uploads/user/Inter-Regional%20Coordination/HCA/HCA16/HCA_ATCM_XLII_Item2.3_v1.1.pdf

Natural Rsources Canada. (2018). Marine Spatial Data Infrastructure (MSDI) & The Marine Cadastre Application. Retrieved November of 2022, from https://na.eventscloud.com/file_uploads/7752d75b94f61c4b76ec33bb2e25a2cb_7JoseMBalaMSDIL0MCApplication_March27LJ2018.pdf

Neumann, B., Vafeidis, A. T., Zimmermann, J., & Nicholls, R. J. (2015). Future Coastal Population Growth and Exposure to Sea-Level Rise and Coastal Flooding - A Global Assessment. PLoS One, 10(3), e0118571.

Nunes, Paulo; Saraiva, Sérgio; Almeida, Sara; Veiga, Leonor, et. al. (2022). HIDROGRÁFICO +: A INFRAESTRUTURA DE DADOS E INFORMAÇÃO GEOESPACIAL MARINHA DO INSTITUTO HIDROGRÁFICO. 10as Jornadas de Engenharia Costeira e Portuária. PIANC.

Peckham, O. (2022). ISC Keynote: Digital Twins Aren't About Making Pretty Pictures. Retrieved from https://www.hpcwire.com/2022/06/01/isc-keynote-digital-twins-arent-about-making-pretty-pictures/

Petrov, O., Gurin, J., & Manley, L. (2016). Open Data for Sustainable Development. (W. Bank, Ed.) Connections. Retrieved from https://openknowledge.worldbank.org/handle/10986/24017

Ponce, R. (2014). The New Role of Hydrography in the 21st Century. Hydro International. Retrieved November of 2022, from https://www.hydro-international.com/content/article/the-new-role-of-hydrography-in-the-21st-century

SciencesPo. (n.d.). Open Science at Sciences Po. Retrieved from SciencesPo: https://www.sciencespo.fr/recherche/en/content/open-science-sciences-po.html

Soares, C. V. (2020). O Conhecimento Científico do Oceano. Instituto Hidrográfico, Conhecer o Mar para que todos o possam usar (Vol. Cadernos Navais n.º 57). Centro de Estudos Estratégicos da Marinha. Retrieved from https://www.marinha.pt/pt/a-marinha/estudos-e-reflexoes/cadernos-navais/Documents/cadernonaval_57.pdf

UN-GGIM. (n.d.). United Nations Integrated Geospatial Information Framework (UN-IGIF). Retrieved from UN-GGIM: https://ggim.un.org/IGIF/overview/

Wilkinson, Mark D.; Dumontier, Michel; Aalbersberg, Jan IJsbrand; et. al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3(160018). doi:https://doi.org/10.1038/sdata.2016.18

Young, A., & Verhulst, S. (2016). The Global Impact of Data: Key Findings from Detailed Case Studies Around the World. O'Reilly Media, Inc. Retrieved November of 2022, from https://www.oreilly.com/library/view/the-global-impact/9781492042785/

Sobre los autores

Paulo Nunes

Is a navy officer since 2004. Works in Geospatial Data and Information Management, Marine Spatial Data Infrastructures development and management since 2016. Is certified with a Cat. A Hydrographer and a MEng in Geographic Engineering from the Faculty of Science of University of Lisbon

João Vicente

Is a Navy officer since 1995. He is the Technical-director of the Portugueses Hydrographic Institute. He has a degree in Hydrography (FIG/IHO/ICA Cat. A) and a MSc in Geographic Information Systems by Lisbon University.

Ana Leonor Veiga

Diploma in Geographic Engineering by the Faculty of Sciences of University of Lisbon and member of the Engineering College. She works as assistant of the Head of the Hydrography Division since 1993. From 1993 to 2005 she worked in cartographic production. She is responsible of the Sector of Bathymetric Data Management since 2005.

Cristina Monteiro

Diploma in Geographic Engineering by the Faculty of Sciences of University of Lisbon and has a Specialization in Hydrography Category A certified by IHO, FIG and ICA and she is a member of the Engineering College. From 2008 to 2021, she had been working as a researcher in Hydrographic Surveys, at Hidrográfico´s Hydrographic Division. Currently, she works at the Sector of Bathymetric Data Management, and she has been dedicated to external data quality control and data analysis for nautical charts updating. Researcher and/or team coordinator in EU Framework Programme I&D Projects (Guad20).

Telmo Dias

Is a Navy officer since 2006. He works at the Technical-Scientific Data Management Centre of the Portuguese Hydrographic Institute. He has a degree in Hydrography (FIG/IHO/ICA Cat. A) and a MSc in Geographic Information Systems and Science by NOVA Information Management School.

Carla Palma

Is the Head of the Chemistry and Pollution of the Marine Environment department at Instituto Hidrográfico, she is a chemical engineer specialized in marine geochemistry. She has a Master degree in Technological Organic Chemistry and a PhD in Chemistry from University of Aveiro in 2014, related with the study of metals in seawater and sediments of the Azores Platform. Since 1996 she has been working at IH, developing studies and works on the chemical properties of seawater and marine pollution, particularly in the field of trace metals in the marine environment. She has a large experience with laboratory methodologies in the field of marine pollution characterization, planning and implementation of monitoring programs in coastal and estuarine areas and processing data of pollutants. She is also lecturer of Marine Chemistry at the School of Hydrography and Oceanography at IH.

Miguel de Castro Neto

Is Dean of NOVA Information Management School (NOVA IMS) at Universidade Nova de Lisboa, where he created NOVA Cidade – Urban Analytics Lab, dedicated to smart cities and urban analytics. He was Secretary of State for Spatial Planning and Nature Conservation in the 19th and 20th Governments and Smart Cities Personality of the Year 2017 (Green Business Week / AIP Foundation). Founder of the Data Science Portuguese Association and member of the Platform for Sustainable Growth. He develops his research and teaching work in the area of Business Intelligence and Smart Cities, with emphasis on the creation of the Master in Knowledge Management and Business Intelligence at NOVA IMS, currently classified as the "World Best Master Program in Business Intelligence" by the Eduniversal international ranking, where he teaches the Business Intelligence course units.