Digital Twins of the Ocean

DITTO

Digital Twins of the Ocean

2021 2030 United Nations Decade of Ocean Science for Sustainable Development

DITTO Programme White Paper

July 2023

Pierre Bahurel[1], Ute Brönner[2], Pier-Luigi Buttigieg[3], Fei Chai[4], Eric Chassignet[5], Colin Devey[6], Enrique Alvarez Fanjul[1], Katherine Hill[7], Sung Yong Kim[8], Joana Kollert[6], Marjolaine Krug[9], Jing Li[10], David Millar[11], Mônica M.C. Muelbert[12], Jay Pearlman[13], Lisa Rolls[14], Jörn Schmidt[15], John Siddorn[7], Jeroen Steenbeek[16], Toste Tanhua[6], Dawn Wright[17], Ann-Christine Zinkann[18] & Martin Visbeck[6]

---

[1] Mercator Ocean International, France
[2] SINTEF Ocean, Norway
[3] Alfred-Wegener-Institute, Bremen
[4] Xiamen University, China
[5] Florida State University, USA
[6] GEOMAR Helmholtz Centre for Ocean Research, Kiel
[7] National Oceanography Centre, UK
[8] Korea Advanced Institute of Science and Technology
[9] Department of Forestry, Fisheries, and the Environment South Africa
[10] International CLIVAR Project Office, China
[11] Fugro
[12] IMar, Universidade Federal De São Paulo, Brasil
[13] IEEE, France
[14] United Nations Environment Programme
[15] International Council for the Exploration of the Sea (ICES)
[16] Ecopath International Initiative, Barcelona, Spain
[17] Esri
[18] National Oceanic and Atmospheric Administration and University Corporation for Atmospheric Research, USA

Scope

This white paper provides an outline of the concept of digital twins of the ocean and lays out the challenges, opportunities for action and suggestions for their implementation as defined by the UN Decade Action "Digital Twins of the Ocean" (DITTO).

Table of contents

Executive Summary

Digital twins are virtualisations or representations of real-world objects or systems, with two-way connectivity between the real and virtual world. They have been widely used in engineering and have recently been adapted to Earth system applications. The marine community is envisioning Digital Twins of the Ocean (DTO) as an opportunity to make informed operational, scientific, management, and policy decisions about the real-world ocean.

Digital twins are powered by continuous data streams from observing systems focused on their physical twins, combined with a suite of models to accurately represent actual and hypothetical ("what if") behaviours users wish to simulate. The connection between the digital and physical twins requires a well-formulated interface between complex constellations of software, interoperable data lakes, and the user. Vitally, user experience and interaction is deeply connected to all other components in a digital twin environment, requiring far greater alignment of user needs with technology implementation.

Thus, DTOs enable users to address 'what if' questions based on shared and relevant data, models, and knowledge. They empower ocean professionals, citizen scientists, policymakers, and the general public alike to visualise and explore the ocean and its possible futures.

The UN Ocean Decade Program DITTO (Digital Twins of the Ocean) is establishing a framework to realise the value of digital twins by promoting co-design of twins, harmonising methods across communities, enabling interoperability and promoting education and uptake. Importantly, DITTO will make sure that the technological and capacity development surrounding DTOs is shared by all nations equally, ensuring inclusiveness particularly for those that are traditionally underrepresented and under sampled. The development of DTOs will work towards closing existing gaps and operate under the "no one left behind" mantra of the UN Decade of Ocean Science.

Digital Twins explained

Characteristics of Digital Twins

Digital twins (DTs) are virtualisations or representations of physical objects and systems - their physical twin. They track their physical twin as it changes, and - based on a deep understanding of the physical twin's composition and properties - also allow users to simulate how the digitised system would react to hypothetical ("what if") scenarios. Digital twins have been widely applied in the engineering realm for tasks such as engine optimisation and port management, helping bring data to life and accelerating development. This connection between the digital twin and their real counterparts requires a well-formulated interface between the digital twin, environmental data, and the user. User interaction is therefore an essential function that is embedded in the design of digital twins, including visualisation, user-driven data transformation and data-science tools.

The fundamental building blocks

Digital twins are data hungry and require an **observing system** around their physical twin to keep up to date. Such observing systems must be co-designed by both the users and developers of the digital twins, to ensure relevance and accuracy. As the link between the digital twin and the observing system evolves, they can begin to guide each other's development and optimisation.

A **data analytics and prediction engine** maximises the understanding and value from these observational data through predictive modelling, emulation, artificial intelligence (AI), and machine learning (ML). Digital twins incorporate the additional capacity for the user to modify the prediction engine to explore options, "what if" scenarios, and consequences.

Complementing the data streams from observing systems, digital twins will draw from a dynamic, distributed, and diverse pool of data, software, models, and other digital assets (including other twins). Thus, a robust and **open digital exchange system** must be in place to support efficient data and information transfer. The more standardised and interoperable each class of assets are, the greater their discoverability and potential for re-use by multiple digital twins.

Once data and other digital assets from an observing system and/or digital exchange system have been gathered, most digital twins will store the data in a **local repository** for rapid access and efficient use. Depending on the needs of the digital twin, this repository may include data lakes, data warehouses, data spaces or other forms of data storage solutions, alongside repositories for other assets (e.g., software, code). Data stored in a digital twin's local repository will often be curated or transformed to improve performance.

To realise the value of the data gathered, a digital twin's **information management framework** is another essential building block. By defining common standards of collecting, storing, and sharing data, intercomparison and interconnectivity between different digital twin applications is ensured.

An **interactive and provisioning** layer allows users to visualise, interact with and tailor the data, scenarios, and models to meet their needs. They are a powerful interface to the information and tools in the data engine that are easy to adapt and use and represent one of the characteristic features of digital twinning. These provisioning layers are tailored towards human users and are often visually pleasing front-end interfaces with easy, intuitive access. Machine-to-machine provisioning is also common, in which case the provisioning layer is tailored to enable federation with other twins or systems.

Moreover, developing digital twins will require an **outreach and training capability** to enable full participation of developers, experts and users in digital twins' environments respecting the capabilities and realities of the diverse international communities. In some areas, the uptake of the concept of DTs is still incipient, possibly due to limited regional technical capacities combined with deficient sustainable access to technologies and research platforms.

Digitally twinning our oceans

Why digital twins of the ocean?

A well-constructed digital twin of the ocean will enable a wide range of users to interact with ocean data and information to improve understanding and inform decisions, and to enhance ocean literacy. Manipulating the twin to address different “'what if”' scenarios can provide information for decision-making, enabling users to explore ways in which the ocean will respond to a changing set of conditions. DTOs will provide ocean researchers, professionals, citizen scientists, educators, policymakers, and the general public alike with the capability to visualise and explore ocean knowledge, data, models, and forecasts.

A global programme for global twinning

The Digital Twins of the Ocean (DITTO) Programme is part of the UN Decade of Ocean Science for Sustainable Development and brings together actions from around the world. DITTO directly answers to **Challenge Number 8** set by the UN Ocean Decade, to “*create a digital representation of the ocean*”.

DITTO’s **vision** is a world where digital twins are used to support ocean science, ocean protection, ocean governance and a sustainable ocean economy. DITTO’s **mission** is to develop and share a common, globally inclusive understanding of digital twins of the ocean, to establish best practices in their development, and to advance a digital framework to empower the global ocean community to effectively use digital twins.

DITTO will promote the co-design of twins with targeted end-users, raise awareness of their uses and applications, and demonstrate their potential in enhancing decision making processes across multiple sectors, including ocean governance.

DITTO Building Blocks

There is a common understanding in the scientific community that there are specific building blocks that comprise Digital Twins of the Ocean:

1) (near) Real-time data from observations (in-situ sensors, satellite) feed into the twin and ensure that it changes together with its twin environment.

2) Data repositories contain digital twin relevant data and are updated frequently. They are accompanied by metadata repositories for user and machine-to-machine interaction.

3) Data processing capabilities include (but are not limited to) ocean models (hydrodynamic, global to coastal dynamics, ecosystem dynamics) for the physical, chemical, and biological representation of the ocean, AI analysis, and processing capabilities that fuse model data with data from observations. Furthermore, AI

predictive models extend environmental predictions beyond their current limitations.

4) Computing infrastructure to host data processing and re-runs of models (reanalyses, what-if scenarios). These may range from edge computing (on the sensor carrier) to the cloud and may include high-performance computing (HPC) clusters.

5) A user interaction layer which might be as simple as a dashboard, as sophisticated as immersive visualisations in virtual or augmented realities, or anything in between including feedback into the twin to run what-if scenarios or trigger actuators and alerts.

6)Education & capacity building in order to spread the know-how necessary to develop and use DTOs in a sustainable and equitable manner.

In order to address these building blocks, DITTO has organised **six** distinct **Working Groups** (WGs), which are outlined in the sections below.

Systems perspectives and needs

Digital twins are not new. However, their application to the Earth system and Ocean domains is relatively recent. Support to develop international capacity and capabilities is needed to ensure that observations and other data are accessible through appropriate data spaces, that models and analytics systems are available and usable to all, and that DTO information can be developed and applied to serve diverse international communities' user needs. This offers the potential for a wider participation of the general public (e.g., local actors) and policy makers, and supports the development of accessible language and science literacy.

Comparisons with existing and planned twin architectures, interoperability standards, sharing of good and best practices and community-based learning are all necessary steps towards defining a systematic DTO framework. We must secure investments in long-term sustainable observations and promote development of capacities, which will raise the accuracy of our models and predictions.

WG 1: Supportive ocean observations and data delivery

SUMMARY

A fit-for-purpose and increasingly accurate ocean observing system co-designed between users and developers of the observing networks needed for digital twins of the ocean will create a positive and continual feedback loop between both where information from the digital twin can be used to inform and optimise the observing network whilst benefiting from it.

Possible actions include:

- Partner with nations to identify gaps in ocean observing data sharing and establish strategies on how to address them.
- Establish links with actors from the maritime sector including offshore energy, shipping, mariculture, and tourism to explore the provision of additional data.
- Identify additional information needed other than essential ocean variables (EOVs); ocean use, topography, traffic, marine spatial planning, human impact and pressure, and social aspects, particularly ITLK (indigenous, traditional, and local knowledge) input.
- Strong links between data storage, data interoperability, data analytics and end users.

CHALLENGE

Digital twins have the ambition to provide an accurate representation of the ocean system. Thus, they require an observing system that can describe with adequate accuracy the current state of the ocean. The ocean remains under-observed in certain regions of the world, particularly in the very dynamic and complex shelf (Exclusive Economic Zone) and coastal regions, where most human interactions with the ocean occur. Observations are limited by spatial and temporal gaps. The establishment of digital twins will thus put increasing demands on accuracy and coverage of ocean observing systems. For many twin challenges there is simply not enough data to adequately initiate or verify the digital twin. While many digital twins will operate regionally, they often require input to their open boundaries from a more global system. Thus, we expect a productive co-design process to advance ocean observing systems.

A key challenge is to secure the required resources for infrastructure, personnel (capacity development) and observation networks. These networks must be sustainably grown and maintain an inclusive and fit-for-purpose observing system framework for digital twins to operate properly.

An equally important key challenge is ensuring inclusivity and the need to engage local communities, society and scientists in ocean observing activities that support DITTO. Part of this challenge is the limited access to marine technologies and research platforms in a sustainable way, lack of know-how and easy to operate instrumentation and platforms. Local actors (scientists, the public, stakeholders, ITLK holders and social media) are better equipped to act as knowledge brokers within their local social-political contexts, thus helping co-design and fit-for-purpose scenarios and solution generation.

A further challenge lies in the accessibility of existent observational data. Not all observations are readily available for digital twin use. By implementing FAIR (Findable, Accessible, Interoperable, Reusable) and CARE (Collective benefit, Authority to control, Responsibility, and Ethics) principles, we can promote and ensure open access to more data and thus narrow down observational gaps. Another challenge here is to share best practices to handle observations and to conduct data quality assurance and control appropriately, and to supply meta data of sufficient granularity following accepted standards. We need to grow and foster a culture of data sharing in ocean observing. We also need to create incentives to make data available, for instance data-papers in peer-reviewed journals like Earth System Science Data (ESSD) and create structures that make it easy and desirable to share data - technically and also politically. This can be a barrier in many nations, since some data are considered sensitive or withheld from private actors to secure exclusive economic benefits.

AREAS FOR ACTION AND PARTNERSHIPS

We might find a place to have integrated ocean observations to make a testbed/ pilot for the digital twin. We need to identify the most mature/observed EOVs as good candidates for an early digital twin pilot. We can also identify the least observed EOVs and thus identify priorities for further development.

Thus, there is an opportunity to network with identifying priority/focus areas based on gap analysis in view of the digital twin. Sustainability of observations needs to be ensured. Today the majority of *in-situ* ocean observations are made on short-term science funding. Moreover, latency of data delivery is sometimes an issue for different delivery areas.

A number of key partners for activities and connections have been identified. First and foremost, national agencies with an ocean observing mandate will be very important partners to grow and sustain the ocean observing systems, in particular within national Exclusive Economic Zones and their coastal regimes. We also acknowledge ITLK as an extremely valuable source of information.

On a global level, the international ocean observing supporting frameworks and networks of the Global Ocean Observing

System (GOOS) and their projects and regional substructures will be a key partner, as well as remote sensing capabilities coordinated by the Committee on Earth Observation Satellites (CEOS) and the national and regional space agencies.

Working with the private maritime sector and civil society to explore new streams of observational data from their respective platforms and activities is a further opportunity.

The UN Ocean Decade and its Decade Collaboration centres (OceanPrediction DCC, DCO on data, a possible DCO on ocean observing) as well as other UN decade endorsed programmes: Observing Co-Design, CoastPredict, etc. offer important areas of partnership and collaboration.

IMPLEMENTATION

There is an opportunity to:

Partner with nations to identify gaps in ocean observing data sharing and establish strategies on how to address them. During those discussions one should: Make a case for the utility of data to serve many users (measure once, use many times); make a case to national and regional governments to observe EOVs and share data; make it mandatory to share data in order to get funding; make a case for certain kinds of data (non-sensitive) to be shared.

Establish links with actors from the maritime sector including offshore energy, shipping, mariculture, and tourism to explore the provision of additional data.

Identify additional information needed other than EOVs, ocean use, topography, traffic, marine spatial planning, human impact, and pressure, etc.

Strongly link to the data lake and interoperability, and data analytics, if there is overlap and need to communicate and link (cross-reference).

WG 2: Data analytics and prediction engines

SUMMARY

Data Analytics and Prediction Engines maximise the understanding and value from data and provide means to add value to ocean observation through predictive modelling, emulation, and ML and AI to create, manipulate and analyse marine information. Digital twins incorporate the additional capacity for the user to modify prediction engines to explore ‘what-if’ scenarios and their likely consequences.

Possible actions include:

- Develop high-resolution model setups for the coastal regions that aid in decision making for climate-change related changes that affect coastal infrastructures and islands.
- Pre-defined quality controlled ‘what-if’ scenarios that explore different outcomes with applications in climate change, marine spatial planning and ecosystem-based decision making.
- Develop tailored data analytics and predictive capabilities based on ML and AI that provide the digital twin with its necessary functionality (monitoring vs. alerts vs. decision making).

CHALLENGE

The challenge with hydrodynamic and biological ocean modelling is the time needed (performance) and requirement to data storage when run operationally and in high resolution (time and space). However, recent developments in the field of ML and AI deliver promising means to address these challenges.

As digital twins for coastal areas will have different requirements to spatial and time resolution than digital twins for the open ocean, digital twins for disaster response will have other requirements to time resolution and prediction accuracy than climate change scenarios. The purpose and scope of a DTO will challenge the development of data analytics and prediction engines, as well as the intended interoperability with other DTOs and other ocean areas. This aspect will also include the separation of data and data analytics with the opportunity to bring data processing to the data, and to leverage newest technology.

AREAS FOR ACTION

Thus, there is an opportunity to network with ocean modellers and data scientists from all over the world to gather the best available technology and opportunities in the field of data analytics and predictions. Key elements to be developed during the Decade to support the worldwide implementation of the architectures include:

- Data sources, with the objective of linking/facilitating access to the existing entry points.
- Software Tools: Understood as a software component adopted by the community to be a part of the value chain, including AI elements.
- Best practices: methodologies that repeatedly produced superior results relative to other methodologies with the same objective.
- Standards: standards have the same objectives as best practices, but the difference is that they may serve as benchmarks for evaluation in addition to being processes.
- Establish operational readiness levels (ORL) indices to formalise the milestones for service evolution and describe the degree of operational maturity, considering all aspects of an operational system.
- Identification of required improvements in the value chain to improve the application of models and data sciences in the decision-making process.

AREAS FOR PARTNERSHIPS

Partnerships are critical to achieve the Data Analytics and Prediction Engine objectives. Such collaboration include:

- Close dialogues with all aspects of DTOS, but in particular pertaining to:

- Private sector, to understand their needs and expectations.
    - Data lakes and interoperability, to understand the findings of this team and include it on the architecture considerations.
    - Architecture, design, and implementation, to understand the flow of data and information in Digital Twins
    - Expert Team on Operational Ocean Forecast Systems (ETOOFS), to understand its role on the final architecture and how this body will contribute to the certification processes.
- Other Decade programmes, especially with:
    - BestPractices, to ensure coordination on the definition of Standards and Best Practices
    - Foresea, to exploit complementary visions and properly align the work.
    - CoastPredict, to ensure the link with the coastal activities.
    - OceanPredict and its Task Teams, specially Intercomparison and Validation TT and Coastal TT, to incorporate the proper scientific background.
- Decade DCC and DCU different from OceanPrediction DCC, to ensure alignment towards decade objectives.
- User experience and communication experts, to identify how models and data sciences can be applied for decision making and explore the limits existing in doing so.

IMPLEMENTATION

There is an opportunity to establish close collaboration between the different UN Decade programs that have similar requirements to data analytics and predictions.

WG 3: Data lakes and interoperability

SUMMARY

Data Repositories (centralised or distributed) provide access to ocean observations and the output from predictive systems through data communication and management in a timely way, following common data principles. These principles include an information management framework, a forum for setting data standards and protocols, and a system for data mapping architectures and data lineage to track data transformations. The data repository will be connected to the data analytics and prediction engine.

Possible actions include:

- Definition and guidelines to describe the required data repository for each Digital Twin.
- Define and describe interoperability aspects for different data from a variety of sources.
- Close collaboration with different data strategies and groups that address these strategies (Dataspaces, distributed data repositories, edge computing).

CHALLENGE

The challenge with ocean data is its spatial and temporal dimension which might differ between observations, modelled data and other (unstructured) data like video or images or indigenous/citizen data. This variety affects interoperability of different data and requires specific pre-processing of data before they might be used within a Digital Twin. The amount of available data is increasing at a significantly higher pace than capacity for data transfer which means that data needs to be handled, accessed, and processed close to its source and data repositories move from being centralised to being distributed. There are several well-established data curation activities across the globe that have started addressing these challenges.

AREAS FOR ACTION AND PARTNERSHIPS

Thus, there is an opportunity to establish a dedicated network with partners that contribute to developing a DTO area as part of DITTO, and discuss opportunities and guidelines for data observation, data repositories and best practices. Together with the UN Decade Data Coordination Group and similar collaboration activities with the private sector, partnerships for tailored Ocean Twins can be developed accordingly. National governments can join forces with neighbouring states to develop digital twins of their seas and oceans. Initiatives like this might build on interregional activities or the EU Mission which bring together partners from the different ocean basins.

IMPLEMENTATION

There is an opportunity to establish a global collaboration framework on ocean data to discuss and address requirements to data acquisition, storage and handling that result from different policies and sources or data.

The development of this framework needs to include participants from science, governments, and the private sector to account for the different perspectives on data from these stakeholders.

WG 4: Interactive layers and visualisations

SUMMARY

The Interactive and Provisioning layer, a characteristic feature of digital twinning, is a powerful, adaptable, and easy-to-use interface that allows users to visualise, interact with and tailor the data, scenarios, and models to meet their needs. Provisioning layers are catered to human users and are often visually pleasing front-end interfaces with easy, intuitive access. Machine-to-machine provisioning is also fundamental, tailoring the provisioning layers to enable federating with other twins or systems.

Possible actions include:

- Take inventory of existing visualisation strategies in sciences, serious gaming, infographics, and the Internet.
- Make inventory of existing DTOs and how specific audiences are defined and integrated.
- Dynamically fine-tune provisioning layers to cater to the needs of specific audiences.
- Seek close collaboration with groups that develop these strategies.

CHALLENGE

A digital twin must illustrate data and provision access to analytical tools that cater to the needs of wide-ranging audiences, with wide ranges of objectives and interests. The challenge here is to provision access to data and analyses with just enough detail to be useful. This equally applies to visualisation strategies, where data images and data patterns must meet the often-implicit expectations of audiences. Whereas the ability to discover and disseminate data has seen great advances through the establishment of metadata standards, such advances have not yet been made to commonly describe the abilities of analytical tools and visualisations, which inherently complicates the ability of automated systems to cater to even less easily quantified policy objectives and user expectations. As such, the major challenge to interactions and representations of a DTO will be to serve up relevant data, using accurate visualisations, from relevant models and/or prediction engines, to provide users and machines with a meaningful connection to the DTO.

AREAS FOR ACTION AND PARTNERSHIPS

Thus, there is an opportunity to establish a dedicated network with partners in science, the (gaming) industry and policy makers to discuss visualisation and data provisioning strategies with the aim to establish case-specific best practice guidelines. In separation, scientific visual representations and serious gaming approaches can be unified to develop generic, feature-rich, and flexible provisioning layers for the construction of specific DTOs. Especially here lies an opportunity to seek collaboration between DTO implementing networks, UN Ocean Decade projects, serious game developers, and the global Marine Spatial Planning community.

IMPLEMENTATION

As a first step, collaborations should be sought with existing activities that are in the process of conceptualising and constructing DTOs, and to seek dialogue to establish standards and best practices to allow homogenization and standardisation of visualisation engines, data provisioning, and interaction with analytical models. Here, in particular, DITTO can play a key role.

WG 5: Framework - architecture, design, and implementation

SUMMARY

The true impact of Digital Twins of the Ocean will depend on their **complementarity**, i.e. complementarity between oceans, models, and data which in turn depends on **interoperability** between these aspects. Interoperability can be achieved through defining the meaningful **building blocks of Digital Twins of the Ocean** as well as a functional and flexible **architecture** that accounts for existing systems, recognised architecture models and **technology** developments that are relevant for digital twins. Such an architecture needs to acknowledge related strategies on data sharing and twinning, like e.g., Destination Earth (DestinE) that aims at creating a digital replica of our planet where DTOs will play a crucial role.
DTOs will emerge for different areas, in different qualities and for different purposes. A co-created architecture framework for these twins will ensure interoperability and thus complementarity.

Possible actions include:

- Establish an operational working group and contact point for DTO developers.
- Conduct a series of workshops to establish definitions and cornerstones for digital twin architectures with a focus on interoperability and federation.
- Participate in relevant fora on Ocean data and digital (Ocean) twins to promote the working group and its results.

CHALLENGE

Interoperability: DTOs will be systems of Digital Ocean Twins that have been developed a) for a specific ocean area and b) for a specific purpose. These purposes may range from monitoring of environmental parameters or habitats to climate change assessments for a wider area of the global ocean. That means that there will be a variety of Digital Ocean Twins that should be capable of interoperating.

Dissimilarity: Digital twins or similar systems have already emerged or are being developed. The approaches and technology stacks of these twins can be very different with respect to quality, resolution in time and space and sophistication and ease of access to the twin.

Architecture framework: Rapid technology development in the field of sensor technology, edge computing and software for user interaction with data offer unprecedented possibilities and opportunities for DTOs but come with the challenge to define a robust and accepted DTO architecture framework that secures interoperability of different twins on one hand, while allowing to account for future innovations and variety at the same time. Importantly, this rapid technological development must occur under the principles of inclusiveness and capacity building for all; particularly those nations who are (traditionally) underrepresented and under sampled.

AREAS FOR ACTION AND PARTNERSHIPS

Currently, there are a number of initiatives that work toward or in support of a DTO, e.g., the EU strategy for the Digital Twin Ocean, NOAA's National Centre for Environmental Information, the IOC Ocean Data and Information System ODIS, the IOC Ocean Best Practice System OBPS, the Ocean Data Action Coalition and the UN Data Coordination Group, only to name few.

Thus, there is an opportunity to network towards one of DITTO's priority activities to '*articulate and advance a common understanding of Digital Twins of the Ocean architecture*'.

IMPLEMENTATION

Partners from the above-mentioned and other initiatives have joined forces to discuss different architectures that support interoperable Digital Twins of the Ocean. There is opportunity and motivation to coordinate ongoing initiatives and to interlink ocean-oriented digital twins to one another, as well as to existing and emerging data systems, cloud capacities, and other digital resources.

Specific topics that will be addressed include:

- to apply and extend existing interoperability architectures to support coordinated development and operations of digital twins in the ocean domain.
- to ensure that digital twins are able to interoperate with diverse ocean data systems, to both acquire and deliver data and information products, models, and other digital assets.
- to combine data, models and information across different Seas, domains, and stakeholders through a system of systems to accommodate different existing assets and digital ocean twins.
- to accompany and support the development with an interoperability dataspace and best practices for ocean data acquisition and sharing.
- to identify where cross domain digital twin interoperability is required.

WG 6: Education, training, and capacity development

SUMMARY

DTOs take advantage of the increasing number of ocean observations from a wide range of sources, the ability to curate and share those data in public data repositories, to fuse the data with ocean prediction systems and the ability to build decision support systems. The design, operation and deriving benefits from those systems require a digitally literate ocean community of practice around the globe and across sectors and disciplines. This Working Group will take advantage of existing and implement new opportunities by working with appropriate networks, programs, and projects to facilitate education and training in the area of digital ocean technology and in that sense build the needed capacity globally.

Possible actions include:

- Establish a pilot program for a professional master's degree program in "Digital Ocean Data Science and Engineering" that would be offered by a combination of digital lecture modules and on-site practicals at a network of universities.
- Liaise with other Ocean Decade efforts and promote summer schools, lab rotations and other international opportunities.
- Consult with the ECOP community and discuss opportunities of global mutual learning and exchange of expertise using peer-to-peer networks.

CHALLENGE

The challenge is that there is a rapidly increasing number of digital platforms and structures that provide access to meta-data and data holding in the marine domain. Depending on the type of data, these holdings are well organised, easy-to-access or at times more poorly documented and harder to access. At the same time, potential users of this information often lack the training and knowledge to search for these holdings and access the information within. The existence of uneven capacity globally reflects different gaps worldwide, particularly in traditionally underrepresented (global south) regions. In the global north there tend to be more people who in principle have the background and means to access the information, but often lack the knowledge on where to find it. In the global south, due to lacking infrastructure, personnel, coverage and technology, digital literacy is often less well-established, which widens the gap. Inclusiveness, that is, providing an equal access to the opportunities and resources of DTOs, is pivotal in closing this gap. Elsewhere, application sectors are somewhere in the middle.

AREAS FOR ACTION AND PARTNERSHIPS

Thus, there is an opportunity to network with other actors in this space who are in the process of developing training opportunities and providing capabilities. For example, there are a set of actors promoting the advancement, operation and delivery of the ocean observing system associated with programs such as GOOS, GEO Blue Planet and POGO. There are others in the ocean prediction sector that advance the design, operation, and use of ocean prediction systems in association with Ocean Predict and related efforts. Finally, oceanographic data systems are advanced by global programs such as IODE and ODIS but also data holding from NGO and the private sector.

Translation of scientific knowledge to local languages/realities is certainly essential for a more equitable extraction of the information in a more palatable way to the public in general, which does not require information as provided in peer-reviewed journals (using evidence, agreement, and uncertainty/confidence language).

IMPLEMENTATION

There is an opportunity to establish a network of universities and training facilities with the ambition to establish an international master program in "Digital Ocean Data Science and Engineering". Early discussions are underway and there is a need to find a funding mechanism for such a program. The cornerstones could be a repository of digital lectures that would be available globally. Those lectures would be supplemented by practical assignments offered at a local university with in-presence attendance. Finally, one could imagine that master projects could be done in a different location in order to promote mobility and cross sectorial and cultural experiences. A second opportunity is to link amongst different programs to share existing opportunities for summer schools and related advance training opportunities. One could also imagine that some of the more advanced Digital Twin of the Ocean projects could be encouraged to develop such offerings. Finally, one could explore informal peer-to-peer learning and engagement opportunities using existing networks such as the Ocean Decade ECOPs to encourage bi-lateral interactions and exchange.

Summary

Digital twins are virtual representations of objects and systems - the real twin (in this case the Ocean system or a part of it). The concept of digital twins has been widely used in engineering and has more recently been adapted to Earth system applications. Digital twins provide the ability to make informed operational, management, and policy decisions about the real twin. Digital twins build on an observing system and predictive processes or data-driven models that users can interact with to support their needs.

The connection between the digital twin and the real one requires a well-formulated interface between the digital twin, environmental data, and the user. User interaction is therefore an essential function that is embedded in the design of digital twins, including visualisation, user-driven data transformation and data-science tools. Working groups within the DITTO programme are developing actions needed for developing observation, a data analytics and prediction engine, data respiratory, visualisation, building architecture, improving education and capacity building:

An adequate and increasingly accurate **Ocean Observing System** co-designed between users and developers of the observing networks needed for digital twins of the ocean will create a positive and continual feedback loop between both where information from the digital twin can be used to inform and optimise the observing network whilst benefiting from it.

A **Data Analytics and Prediction Engine** to maximise the understanding and value from these data, that provides tools to add value to ocean observation through predictive modelling, emulation, and artificial intelligence/machine learning to create, manipulate and analyse marine information. Digital twins incorporate the additional capacity for the user to modify the prediction engine to explore options, scenarios, and consequences.

A **Data Repository** (centralised or distributed) will provide access to ocean observations and the output from predictive systems through data communication and management in a timely way following common data principles. These principles include an information management framework, a forum for setting data standards and protocols, and a system for data mapping architectures and data lineage to track data transformations. The data repository will be connected to the data analytics and prediction engine.

To set the capabilities of digital twins into value, an **Interactive and Provisioning layer** allows users to visualise, interact with and tailor the data, scenarios, and models to meet their needs as a powerful interface to the information and tools in the data engine that is easy to adapt and use, and represents one of the characteristic features of digital twinning. These provisioning layers are tailored towards users and are often visually pleasing front-end interfaces with easy, intuitive access. Machine-to-machine

provisioning is also common, in which case the provisioning layer is tailored to enable federating with other twins or systems.

The true impact of DTOs will depend on their complementarity, i.e. complementarity between oceans, models, and data which in turn depends on interoperability between these aspects. **Interoperability** can be achieved through defining the meaningful building blocks of DTOs as well as a functional and flexible architecture that accounts for existing systems, recognised architecture models and technology developments that are relevant for digital twins. Such an architecture needs to acknowledge related strategies on data sharing and twinning, such as the Destination Earth (DestinE) that aims at creating a digital replica of our planet where DTOs will play a crucial role. Digital Twins of the Ocean will emerge for different areas, in different qualities and for different purposes. A co-created architecture framework for these twins will ensure interoperability and thus complementarity.

The design, operation and deriving benefits of DTOs require a **digitally literate ocean community** of practice around the globe and across sectors and disciplines. This can be achieved by taking advantage of existing and implementing new opportunities by working with appropriate networks, programs, and projects to facilitate education and training in the area of digital ocean technology, and in that sense build the needed capacity globally.

Through all this, digital twins will enable users to address 'what if' questions based on shared data, models, and knowledge. They empower ocean professionals, citizen scientists, policymakers, and the general public alike to visualise and explore ocean knowledge, data, models, and forecasts. The use of digital twins is rapidly developing, spanning a wide range of use cases from engineering to policy to science to operational services.

The UN Ocean Decade Program DITTO is establishing a framework for DTOs by promoting co-design of twins, establishing, and sharing good and best practice, enabling interoperability, and promoting education and uptake to demonstrate and realise the value of Digital Twins of the Oceans.