Blue-Cloud
Piloting innovative services for Marine Research & the Blue Economy

D6.2 Strategic Roadmap (Release 1)

Work Package	WP6 – Blue Cloud Roadmap, Exploitation & Sustainability Measures, and trans-European Liaisons
Lead Partner	Seascape Belgium (SSBE)
Lead Author (Org)	SSBE
Contributing Author(s)	Julia Vera, Kate Larkin, Xiaoyu Fang (SSBE), Sara Pittonet (Trust-IT), Dick Schaap (MARIS), Pasquale Pagano (IST-CNR)
Reviewers	Cécile Nys (IFREMER), Nathalie Tonné (SSBE), Silvana Muscella, Sara Pittonet, Federico Drago (Trust-IT), Anton Ellenbroek, Marc Taconet (FAO), Dick M.A. Schaap (MARIS), Pasquale Pagano (ISTI-CNR)
Due Date	31.03.2021, M18
Submission Date	02.04.2021
Version	1.0

Dissemination Level

X	PU: Public
	PP: Restricted to other programme participants (including the Commission)
	RE: Restricted to a group specified by the consortium (including the Commission)
	CO: Confidential, only for members of the consortium (including the Commission)

Blue-Cloud - Piloting Innovative services for Marine Research & the Blue Economy - has received funding from the European Union's Horizon programme call BG-07-2019-2020, topic: [A] 2019 - Blue Cloud services, Grant Agreement n. 862409.

DISCLAIMER

"Blue-Cloud, Piloting Innovative services for Marine Research & the Blue Economy" has received funding from the European Union's Horizon programme call BG-07-2019-2020, topic: [A] 2019 - Blue Cloud services, Grant Agreement n.862409.

This document contains information on Blue-Cloud core activities. Any reference to content in this document should clearly indicate the authors, source, organisation, and publication date.

The document has been produced with the funding of the European Commission. The content of this publication is the sole responsibility of the Blue-Cloud Consortium, and it cannot be considered to reflect the views of the European Commission. The authors of this document have taken any available measure in order for its content to be accurate, consistent and lawful. However, neither the project consortium as a whole nor the individual partners that implicitly or explicitly participated in the creation and publication of this document hold any sort of responsibility that might occur as a result of using its content.

COPYRIGHT NOTICE

VERSIONING AND CONTRIBUTION HISTORY

Version	*Date*	*Authors*	*Notes*
0.1	10.12.2019	Kate Larkin, Xiaoyu Fang (SSBE)	TOC 1st draft
0.2	31.07.2020	Kate Larkin (SSBE)	TOC 2nd draft
0.3	05.02.2021	Julia Vera, Kate Larkin (SSBE)	TOC 3rd draft
0.4	01.03.2021	Julia Vera, Kate Larkin (SSBE)	Sections 0, 1, 2, 3, 5, 6, 7
0.5	05.03.2021	Sara Pittonet (Trust-IT)	Section 6
0.5	11.03.2021	Cecily Nys (Ifremer)	General edits, suggestions
0.5	11.03.2021	Nathalie Tonné (SSBE)	General edits
0.5	15.03.2021	Anton Ellenbroek (FAO)	General edits, suggestions
0.5	15.03.2021	Marc Taconet (FAO)	General edits, suggestions
0.6	16.03.2021	Silvana Muscella (Trust-IT)	General edits, suggestions
0.7	29.03.2021	Dick Schaap (MARIS)	Section 4, general review
0.8	30.03.2021	Pasquale Pagano (ISTI-CNR)	Section 4, general review
0.9	31.03.2021	Kate Larkin, Julia Vera (SSBE)	Final review
1.0	31.03.2021	Sara Pittonet (Trust-IT)	Final review and submission

Table of Contents

List of Figures

List of Tables

Glossary

Word	Definition
AI	Artificial Intelligence
Argo	International programme on ocean observation using Argo floats
API	Application Programming Interface
B-C	Blue-Cloud
BlueBridge	H2020 project Supporting Blue Growth with innovative applications based on EU e-infrastructures
C3S	Copernicus Climate Change Service. CFP
CAMS	Copernicus Atmosphere Monitoring Service
CMCC	Euro-Mediterranean Centre on Climate Change
CMEMS	Copernicus Marine Environment Monitoring Service
CNR	Italian National Research Council
D4Science Infrastructure	Data Infrastructure promoting Open Science (managed by CNR)
DG	Directorate-General (of the EC)
DG DEFIS	Directorate-General for Defence Industry and Space (formerly DG GROW)
DG RTD	Directorate-General for Research and Innovation
DG MARE	Directorate-General for Maritime Affairs and Fisheries
DG GROW	Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs (succeeded by DG DEFIS)
DG ENV	Directorate-General for Environment
DG CONNECT	Directorate-General for Communications Networks, Content and Technology
DG DEVCO	Directorate-General for International Cooperation and Development
DIAS	Data and Information Access Service (funded by EC COPERNICUS programme)
DTO	Digital Twin of the Ocean
EASME	Executive Agency for Small and Medium-sized Enterprises
EC	European Commission
EcoTaxa	Web application dedicated to the visual exploration and the taxonomic annotation of images focused on planktonic biodiversity
EMBL	European Molecular Biology Laboratory
EMBL-EBI	European Bioinformatics Institute
EMBRC	European Marine Biological Resource Centre
EMODnet	European Marine Observation and Data Network
ENA	European Nucleotide Archive
EOSC	European Open Science Cloud
ESEB	External Stakeholder & Expert Board (of Blue-Cloud)
EU	European Union
EUDAT	Pan-European network and collaborative data infrastructure, consisting of a network of academic computing centres
EuroBioimaging	European Research Infrastructure for Imaging Technologies in Biological and Biomedical Sciences
EurOBIS	European Ocean Biodiversity Information System
FAIR	Findable, Accessible, Interoperable and Reusable

FAO	Fisheries and Agriculture Organisation of the United Nations
G7 FSOI	G7 Future of Seas and Oceans Initiative
GPU	Graphics Processing Unit
H2020	Horizon 2020 EU Framework Programme
HPC	High-Performance Computing
ICOS	Integrated Carbon Observation System
ICT	Information and communications technology
Ifremer	The French Research Institute for Exploitation of the Sea
iMarine	Data e-Infrastructure Initiative for Fisheries Management and Conservation of Marine Living Resources
ISC	International Science Council
IoT	Internet of Things
JRC	Joint Research Centre
KERs	Key Exploitable Results
KPIs	Key Performance Indicators
MARIS	SME expert in European marine data management infrastructures (Blue-Cloud technical coordinator)
ML	Machine Learning
MOi	Mercator Ocean International
MSFD	Marine Strategy Framework Directive
MSP	Marine Spatial Planning
MSPD	Marine Spatial Planning Directive
NOAA	National Oceanic and Atmosphere Administration (United States of America)
OGC	Open Geospatial Consortium
SDG	Sustainable Development Goals
SeaDataNet	A pan-European infrastructure to ease the access to marine data measured by the countries bordering the European seas (network of NODCs)
SeaDataCloud	EU HORIZON 2020 project aiming at considerably advancing SeaDataNet Services and increasing their usage, adopting cloud and High Performance Computing technology for better performance.
SME	Small-Medium Enterprises
SSBE	SME specialised in marine policy (Roadmap coordinator) - Seascape Belgium
Trust-IT	SME specialised in market and technical research analyses in the field of ICT (Blue-Cloud coordinator)
UN	United Nations
UN 2030	The 2030 Agenda for Sustainable Development
VLIZ	Flanders Marine Institute
VRE	Virtual Research Environment
Web-based Open Science	An approach to the scientific process that focuses on spreading knowledge as soon as it is available using web-based digital, collaborative technologies. In this document, “web-based Open Science” and “Open Science” are considered synonyms, but “web-based Open Science” is used to highlight the key role of web-based technologies in enabling Open Science.
WEkEO DIAS	WEkEO is one of the 5 Copernicus DIAS, bringing in the CMEMS, C3S and CAMS

Executive summary

The ***draft* B-C Roadmap to 2030** that is launched for consultation in June 2021 will be structured around two separate documents, to cater to readers according to their level of understanding of the wider landscape surrounding the Blue-Cloud efforts:

- **The (v1) *draft* B-C Roadmap to 2030:** A short (20-25-page) document introducing a tentative, framing approach towards developing **policy** and **strategy recommendations** for the future evolution of the Blue-Cloud efforts. It will include a summary description of **Blue-Cloud added value** (feeding from **Sections 3**, **4** and **5**) and a general **Vision to 2030** (feeding from **Section 7**), as well as outlining key questions to guide contributions from key stakeholder communities towards shaping a shared list of policy and strategy recommendations.
- **A "Readers' Guide" to the B-C Roadmap to 2030:** A longer document including a detailed description of the **policy context**, **landscape** and **current developments** shaping the **Blue-Cloud project**, including the **strategic approach** towards its future evolution included in the **B-C Roadmap to 2030** is based (**Sections 1, 2, 3, 4, 5, 6** and a brief introduction to **Section 7**).

The final version of the **B-C Roadmap to 2030** (July 2022) will be an evolved, improved and conclusive version of (v1) B-C Roadmap to 2030, including prescriptive **policy recommendations to 2030** based on feedback gathered through the **consultation** and the conclusions set out by the (final) **Blue-Cloud Service Exploitation & Sustainability Plan** (March 2022).

The **Horizon 2020 Blue-Cloud (B-C) project** launched in October 2019, aiming to **demonstrate the potential of web-based Open Science in the marine domain**. To deliver on this objective, it is **piloting** the development of a **web-based cyber platform** that will provide marine scientists with enhanced **analytical capabilities**. It will facilitate their engagement in **collaborative research** and will provide them with access to powerful **cloud-computing resources**, a range of **analytical tools** and **simplified access to multi-disciplinary data** from *in situ* and satellite-derived observations to model outputs. B-C is co-designed by, and builds on, existing European capability, including trusted data services EMODnet, CMEMS and other key research and data infrastructures and e-infrastructures. In the **short-term**, the project is building this cyber platform by means of a smart federation of selected, multidisciplinary data repositories, analytical tools and computing facilities. The added-value of web-based Open Science will be demonstrated by five specific, multidisciplinary "demonstrators" or use-cases. In the **medium-** and **long-term future**, B-C aspires to upscale this cyber platform, its resources, services and applications, together with a thriving community of Open Science service providers and users. B-C will evolve to further align with wider developments at European level to catalyse **transformative solutions** to priority **societal challenges** and unravel **new opportunities** for **innovation**, in support of the **EU Green Deal** and **UN Agenda 2030**. To guide the long-term capitalization and further development of this ambition into the future, the Blue-Cloud Project is producing a **roadmap** to **2030**, which is being developed as a **co-designed**, **community-oriented policy document** with substantial stakeholder consultation and input. This document is a first **draft Blue-Cloud Roadmap to 2030,** that will be further updated and summarised in Spring 2021, ready for public consultation in June 2021. For this reason, this document should not be seen as a **preliminary blueprint**, but rather as an intermediary step towards gathering **input**, **feedback** and **insight** from the **B-C Community** towards its evolution and grounding, benefiting from wide stakeholder consultation.

Sections 1 and **2** introduce the process followed towards the development of the roadmap, as well as the policy context, opportunities and challenges that motivate the **Blue-Cloud**'s efforts, exploring the emergence of **Open Science** in the context of the digital age and how it can contribute to **support** the **European Green Deal** and the **United Nations Agenda 2030 for Sustainable Development**. It analyses relevant developments shaping Europe's **marine knowledge value chain** and reflects on some of the **outstanding challenges** currently hindering its potential to leverage on Open Science to deliver on the objectives of these policy frameworks. **Section 3** describes the **added value** that **Blue-Cloud** will bring to this landscape by **2022**, and the overarching, **strategic objectives** guiding its realization. **Sections 4** and **5** look deeper into the practical, technological and demonstrative **assets** that the project will deliver. **Section 6** reflects on the **community** of "**early practitioners**" of **Open Science** that is emerging around these efforts in the **marine domain**, as another key asset into the future. **Section 7** reflects on the overarching vision that could guide the future capitalization and further development of these results, as well as strategic, high-level actions towards achieving a shared vision.

As the **first early draft "roadmap"**, this document is extensive and **does not yet include policy recommendations**, which will be the core content of the final roadmap. The final version of the **B-C Roadmap to 2030** (see box below) will be **more concise**, focusing on translating the road ahead in practical terms of **policy actions and recommendations**, bringing its **key messages** forward. This current format is used for a better understanding of the Blue-Cloud efforts, **inviting feedback** and **contributions** from the **B-C Project Consortium**, the **B-C External Stakeholder Expert Board**, the myriad of **related projects** and **initiatives** with whom the project is in dialogue, but also from organizations and professionals **not yet directly engaged** in its efforts, but who could be interested in joining in the future. The **policy recommendations** for the final B-C Roadmap to 2030 will be drafted from community responses to the public consultation, in particular to the following key questions:

- How should **Blue-Cloud** evolve to support a **thriving environment** for web-based **Open Science** and **Open data** in the **marine domain**?
- What **governance mechanisms** are most appropriate and desirable to ensure broad **engagement** of the **marine community** and to ensure widely accepted **rules of participation**?
- What **applications** of **Open Science** in **the marine domain** could have a higher probability of success, given current availability of data, models and actors willing to engage in collaborative science, across a broad range of topics? Which of such applications should be prioritized towards addressing current **user needs** and delivering **highest societal impact**?
- How should B-C's Open Science environment and services evolve to be fit-for-use not only for scientists, but also for other Open Science users such as **policy makers** and **blue economy SMEs** and **industry**? What needs do these users have that B-C could evolve to address? What **incentives** can contribute to bring Open Science practitioners on board?
- How can B-C evolve to **further connect** with **marine data infrastructures** and **research infrastructures** to deliver full interoperability of marine data through the **B-C Data Discovery & Access Service**, aligning and in collaboration with other international efforts?
- What actions would be required to enable **B-C's Catalogue** of **analytical methods**, **algorithms** and **applications** to **be deployed in EOSC,** but also in other infrastructures **-closer to data-** or across **supercomputing platforms** in Europe?
- How should **B-C's assets** evolve to align with future **EU pilot DTO** and **DestinE** developments?

1. About the Blue-Cloud Roadmap to 2030

The **Blue-Cloud Roadmap to 2030** is a key legacy output of the **Horizon 2020 pilot Blue-Cloud project**. In this section, we introduce the **background** and **motivations** behind its conception, as well as defining its **scope** and the process developed towards its elaboration and final delivery, planned for **July 2022**. The Roadmap is being developed through a **participative**, **co-design process.** This is the **first draft** released for public discussion and contribution from the Blue-Cloud community and it is intended as a first step to gather input, feedback and additional insight towards its evolution and grounding, benefiting from wide stakeholder consultation.

1.1 Background, scope and motivations of the B-C Roadmap

The **Horizon 2020 pilot Blue-Cloud project**[1] is the component of the "**The Future of Seas and Oceans Flagship Initiative"** aiming to demonstrate the potential of web-based **Open Science in the marine domain.** The project is **piloting** the development of a **web-based cyber platform** to provide marine researchers with enhanced **analytical capabilities** to enable **Open Science**. It will provide them with powerful **computing facilities**, a range of **analytical tools** and **simplified access to data** from in-situ and satellite observations, data products and model outputs available across different blue data infrastructures and marine domains. Blue-Cloud will form a **marine thematic community** interacting and integrating with developments of the **European Open Science Cloud** (EOSC).

The Blue-Cloud project will deliver **innovative, demonstrative assets**, which will have to be taken up, sustained and evolved to capitalize on its results and wider potential. The **Blue-Cloud Roadmap to 2030** seeks to provide guidance for the future evolution of these efforts, towards further upscaling and innovating its assets and building a thriving B-C community. It will provide a **vision** towards their development beyond project-end and into the future (2030). Such a vision is built around a **mission statement** and a set of **policy** and **strategy recommendations** towards **2030**, potentially guiding a series of EU Calls for Proposals for its implementation. The Roadmap will be the result of a **collective reflection** on how the results of the Blue-Cloud project could be exploited and evolved into the future to maximize their impact, catering to a much wider user base and bringing about **disruptive change** along the **marine knowledge value chain**, aligning with wider developments at European and global level to unravel new opportunities for **Ocean**, **science-based innovation** in support of the **EU Green Deal** and the **UN Agenda 2030**.

1.2 Co-creating the Blue-Cloud Roadmap to 2030

The **Blue-Cloud Roadmap to 2030** is being developed as a **co-designed**, **community-oriented policy document**, specifically aimed at further, long-term strategic development of the Blue-Cloud's efforts. It is being developed over the course of the three-year pilot Blue-Cloud project, with substantial stakeholder consultation and input. Its facilitation and development are led by **Seascape Belgium**, in close collaboration with **Trust-IT** - as Coordinator of the Blue-Cloud project - and **MARIS** - as technical coordinator of the project- together with **CNR**, **FAO** and **Ifremer**, bringing in wide input from the Blue-

[1] https://www.blue-cloud.org/

Cloud partnership and engaging them for dialogue across key communities of stakeholders (as described in Section 3).

In 2020, building from an initial Concept Note[2], the project saw the start of an **initial phase** of **stakeholder dialogue** that extended into early 2021. This initial phase was geared at establishing targeted conversations with closer members of the Blue-Cloud Community, including Project Partners[3], the **B-C External Stakeholder & Expert Board[4]** (ESEB), representatives of the **European Commission** and key stakeholder communities (marine researchers, marine data infrastructures, research infrastructures and e-infrastructures). This is part of a 5-step framing process (as described in Figure 1) that will invite extensive and diverse stakeholder engagement, including targeted dialogues, workshops, surveys and a public consultation.

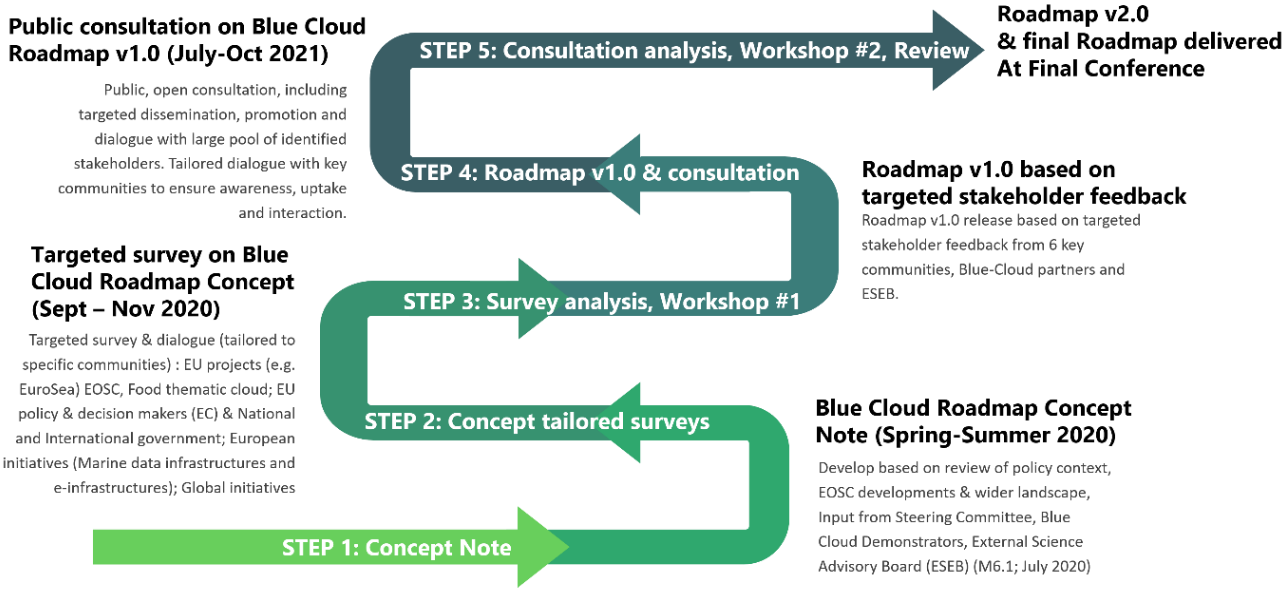


Figure 1: Development of the "Blue-Cloud Roadmap to 2030": 5 steps from Concept Note to Final Roadmap, with extensive and diverse stakeholder consultation including targeted surveys and dialogue, public consultation and workshops. **Credit:** Seascape Belgium

This is the first **draft release** of the **Blue-Cloud Roadmap to 2030**. However, it should not be seen as a **preliminary blueprint**, but rather as an intermediary step intended to gather **additional input**, **feedback** and **insight** from the **Blue-Cloud Community** towards its evolution and grounding, benefiting from wide stakeholder consultation before full completion. This draft release will open for public consultation in **June 2021** and contributions will be welcomed until **October 2021**. This consultation will aim to gather as many contributions as possible from the **wider marine & maritime community** in Europe in internationally and other actors across all **key stakeholder groups** conforming the **Blue-Cloud Community** (as described in **Section 3**).

[2] Blue-Cloud strategic Roadmap Concept Note (Milestone MS41)
[3] https://www.blue-cloud.org/partners
[4] https://www.blue-cloud.org/eseb

2. Unlocking web-based Open Science in the marine domain: Policy context, opportunities & challenges

The motivations and goals behind the **Blue-Cloud**'s efforts are shaped by the context of our times, driven by a confluence of **global change** -large scale, Planetary changes in the Earth system and in society- and the **digital revolution**. In this section, we explore the emergence of **Open Science** in the context of the digital age and how it can contribute to **support key policy responses** to global change, namely the **European Green Deal** and the **United Nations Agenda 2030 for Sustainable Development**. We then analyse relevant developments shaping Europe's **marine knowledge value chain** and reflect on some of the **outstanding challenges** currently hindering the potential to leverage on Open Science to deliver on the objectives of these European and international policy frameworks.

2.1 Open science: An emerging opportunity catalysed by the digital age

Our capacity to gather, store and analyse data derived from observations has increasingly grown and supported human progress through time, but has exploded since the dawn of the **digital age**. Today, **data is shaping the future of humanity**. It is facilitating **scientific research**, boosting **innovation**, helping to improve **public policy** and the delivery of public services -including transparency and accountability- and making societies more productive. The digital age is also bringing about profound structural changes in the global economy. The five largest firms in the world in terms of market valuation -Apple, Amazon, Facebook, Google and Microsoft- are actors in the data economy with a combined market value of nearly $4 trillion in 2018. In 2008, only one among the five largest firms was a data firm, supporting the idea of an economic transition from an economic system dominated by traditional "brick and mortar" industries to a digital economy[5]. In the coming decade, it is expected that data will completely reshape the way we produce, consume and live, becoming the lifeblood of economic development[6].

The evolution of ICT, and specially web-based technology, has brought unprecedented changes and possibilities as it enables not only **sharing** and **processing** large amounts of data, but also the **results** of **its analysis**, feeding a virtuous circle of knowledge development. As a result, and just like any other sector in society, **science** is also undergoing a profound transformation, as **scientific data**, **methods** and **results** are shared across teams, geographical borders and disciplines, opening unprecedented, new collaboration opportunities underpinned by game changers like **cloud computing**, **Artificial Intelligence (AI), big data analysis** and **machine learning**. The associated concept of "**Open Science**" is shifting the prime focus of researchers away from "publishing" towards **knowledge sharing**[7]. This shift is timely and aligns with society's expectations on how to govern our collective use of data. In a recent consultation launched by the European Commission on the European Strategy for Data[8], 91.5%

[5] H.Rashid et al. (2019) *Data Economy: Radical transformation or dystopia*?

[6] European Commission (2020) *A European Strategy for Data* https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020_en.pdf

[7] J.C Burgelman et al (2019). *Open Science, Open Data, and Open Scholarship: European Policies to Make Science Fit for the Twenty-First Century* https://www.frontiersin.org/articles/10.3389/fdata.2019.00043/full

[8] European Commission (2020) *Summary Report of the Public Consultation on the European Strategy for Data* https://ec.europa.eu/digital-single-market/en/news/summary-report-public-consultation-european-strategy-data

of respondents agreed that "*more data should be available for the common good, for example for improving mobility, delivering personalised medicine, reducing energy consumption and making our society greener*", signalling broad societal consensus that data should be used for tackling societal, climate and environment-related challenges, contributing to healthier, more prosperous and more sustainable societies. One immediate way in which data can contribute to more sustainable societies is supporting the generation of **knowledge** for evidence-based societal decision-making and solutions through **science-based** (top-down) policy making processes and (bottom-up) innovative applications. The **scientific community** plays a key role in the process of transforming data into knowledge to guide societal action.

In 2015, the European Commission initiated the process towards the creation of the **European Open Science Cloud** (EOSC) as an "**internet for science**" or an environment for hosting and processing research data to support EU science. Based on principles of minimal governance, maximum freedom of implementation and global interoperability and accessibility, **EOSC** intends to make it possible for 1.7 million researchers in Europe to store, share and re-use research digital objects (like publications, data, workflows and software) across nations and scientific disciplines, following the FAIR[9] principles for data stewardship and management (making digital objects **F**indable, **A**ccessible, **I**nteroperable and **R**eusable). These efforts are geared at triggering new insights and innovations, higher research productivity and improved reproducibility in science.

The **European Strategy for Data** launched in 2020 confirmed that the EU will continue to facilitate **discovery**, **sharing** of, **access** to and **reuse** of **open data** and **services** by researchers through EOSC, in support of **Open Science**. Building on this experience, it will also support the establishment of nine common European data spaces, including a "**Green Deal Data Space**". The EU is determined to unlock the full benefits of the digital age to support its **societal goals**, which have been recently encompassed in the "**European Green Deal"**, an ambitious political agenda to **improve the well-being of citizens, securing a healthy planet for generations to come**. Central to achieving this objective is the need for **open data** and **Open Science**, as the cornerstone of the knowledge value chain fuelling science-based evidence, solutions and innovation across all economic sectors and policy making processes. One key initiative connected to the European Strategy for Data is "**Destination Earth**", which will bring together European scientific and industrial excellence to develop a very high precision digital model of the Earth. This ground-breaking initiative will offer a digital modelling platform to visualize, monitor and forecast natural and human activity on the planet. The digital twin of the Earth will be constructed progressively, starting in 2021. Contributing to Destination Earth, a **Digital Twin** of the **Ocean** (DTO) is also planned for development, as described in the next section. It will empower citizens, governments and industries to collectively share the responsibility to monitor, preserve and enhance healthy and productive marine and coastal habitats, while supporting a sustainable blue economy. A cultural and technological shift has begun enabling **Open Science** that can lead to **breakthrough developments** along the **marine knowledge value chain**, as a driving force of **innovation** with a positive scientific, technological, economic and social spillover effect. This momentum needs to be further nurtured and catalysed, raising awareness about the **opportunities** it opens, identifying the **challenges** that need to be overcome to realize their materialization and signposting **concrete steps** towards that end.

[9] M.Wilkinson et al (2016). *The FAIR Guiding Principles for scientific data management and stewardship* https://www.nature.com/articles/sdata201618

2.2 Embracing the opportunity: How web-based Open Science can contribute to the European Green Deal and the UN Agenda 2030

The EU Green Deal

The **EU Green Deal** sets out clear objectives for Europe to become **climate neutral** by **2050**, supporting an **inclusive economic** and **ecological transition** to position European industry as a global leader in the "green" economy. It intends to use the major potential of **data** in support of the **Green Deal priority actions**, namely:

- Increasing the EU's climate ambition for 2030 and 2050.
- Supplying clean, affordable and secure energy.
- Mobilising industry for a clean and circular economy.
- Building and renovating in an energy and resource efficient way.
- From 'Farm to Fork': designing a fair, healthy and environmentally friendly food system.
- Preserving and restoring ecosystems and biodiversity.
- A zero-pollution ambition for a toxic-free environment.

The EU Green Deal recognizes the vital role of the **Ocean** across all of these priority actions. To preserve the Ocean's capacity to contribute to the wellbeing of citizens, it has launched **Mission Starfish**[10], a blueprint of 5 overarching goals and 17 targets guiding an articulated portfolio of actions -such as research projects, policy measures and legislative initiatives- that will be activated as integral part of the Horizon Europe framework programme beginning in 2021 to deliver on the objective of sustaining a **healthy hydrosphere** for our societies and the planet. As the **Mission Starfish** report highlights, healthy oceans, seas, coastal and inland waters are vital for our future. They are the lungs of the Earth, producing half of the oxygen we breathe. They are a source of healthy food, contributing 16% of the animal protein we eat. They are the planet's largest carbon sink and have absorbed 26% of anthropogenic carbon dioxide emissions since the beginning of the Industrial Revolution. They are home to the richest biodiversity on our planet. They are the source of all life on Earth and our planet's life-support system. They supply freshwater, renewable energy and provide benefits associated with our well-being, cultural values, tourism, trade, and transport. Mission Starfish provides a roadmap towards restoring the EU oceans and waters by 2030, in support of the **EU Green Deal** with interactions across all of its policy priorities, acting around key overarching objectives:

- Filling the knowledge and emotional gap to create an ocean literate society.
- Regenerating marine and water ecosystems by directly protecting and revitalizing them.
- Halting pollution of marine, coastal and freshwater ecosystems.
- Decarbonising our ocean, seas and water.
- Revamping Ocean and water governance.

There are different ways in which **web-based Open Science in the marine domain** can contribute to support the inclusive economic and ecological transition pursued by the EU Green Deal and the associated, systemic and transformative solutions for "healthy oceans, seas, coastal and inland waters" proposed by Mission Starfish, opening opportunities to improve our **knowledge** and

[10] https://ec.europa.eu/info/horizon-europe/missions-horizon-europe/healthy-oceans-seas-coastal-and-inland-waters_en

understanding of the **Ocean**, informing better **science-based policies**, **triggering innovation** across the blue economy and supporting the development of an **Ocean citizenship**:

- **Web-based Open Science** can contribute towards **more efficient, effective and cost-saving** environmental **monitoring systems** to **measure progress and compliance** with EU **policy targets** and to better **informed, science-based policies** in support of the EU Green Deal: The European Union has established a range of Directives to assess and to counteract the harmful effects of multiple drivers of change, such as for example the Marine Strategy Framework Directive (MSFD) [11] and Water Framework Directive (WFD) [12]. Member States have to determine the current environmental status of their marine ecosystems, identify threats and establish programs of measures to restore or maintain ecosystem health. These and other European Directives and policies underpinning the EU Green Deal, like the **Marine Spatial Planning Directive** (MSPD) or the **Common Fisheries Policy** (CFP), require **trusted information** and **actionable knowledge** based on and derived from **marine data**, including environmental, ecosystem and human activity. All Directives require agreed monitoring and assessment programmes that deliver regional scale outputs that are subject to intense scrutiny and maybe subject to legal challenge in the European court. Many of the evidence-based decisions are based on a highly complex set of multi- and, increasingly, trans-disciplinary datasets. **Open science** can deliver **innovative solutions** for effective and efficient **monitoring** (e.g. augmenting and complementing national monitoring data with free access to other existing data, building new remote monitoring capabilities through sensors, Internet of Things or satellite, etc.), and in particular for streamlining and optimising the assessment process (e.g. development of shared monitoring/data tools for rapid analysis, assessment of complex multidisciplinary data and co-location of satellite and in-situ observations). Open science can not only contribute to an improved and transparent evidence base comprised of better information and improved knowledge for these Directives, but also across other policy areas, such as to the wider implementation of the **Integrated Maritime Policy** or to sending better **price signals** to **carbon markets**, as a result of better information on the impact of climate change mitigation efforts.
- **Web-based Open Science** can **enhance Europe's marine knowledge value chain** by improving its environmental prediction capabilities, in support of timely, adaptative crisis management: Europe's pledge to be climate neutral by 2050 will require a huge shift in public opinion, understanding and appreciation of our inextricable link with the ocean, and its crucial role in shaping our natural- and human-centric world. Achieving this goal will require a quantum leap in our capacity to observe, understand and predict complex and interconnected natural and anthropogenic processes occurring at different spatial and temporal scales. This is especially important at a time when complex interactions and feedback among different components of the Earth system and human activities have reached a scale where abrupt global environmental change can no longer be excluded. Novel and emerging technologies -including **cloud** and **High-Performance Computing** (HPC), **Artificial Intelligence**, novel in-situ data gathering capabilities (such as **Internet of Things** or **citizen science**) or advanced and **high-**

[11] https:// ec.europa.eu/environment/marine/eu-coast-and-marine-policy/marine-strategy-framework-directive/index_en.htm
[12] https:// ec.europa.eu/environment/water/water-framework/index_en.html

resolution modelling technologies- are empowering scientists in their capacity to understand and predict such complex interactions. Earth system science, together with latest developments in Earth Observations and the aforementioned technologies can play a major role in unravelling the complexities of our Earth system and human activities, as well as in developing Europe's environmental prediction and crisis management capabilities. The EU is responding to this opportunity by proposing a "**Digital Twin of the Ocean**", as a high precision digital model that represents a consistent high-resolution, multi-dimensional and (nearly) real-time description of the marine realm. It plans to build on the EU existing forecast and climate modelling infrastructures and services and pool together all available assets related to seas and oceans (data, models, physical ocean observatories at sea) with digital technologies (cloud, super HPC capacities, AI and data analytics) to develop an open-source, easily accessible, usable and understandable interactive platform with simulation capabilities, which will allow testing scenarios that deal with different evolutions of the ocean environment. **Open Science**, underpinned by **open data**, will be the driving force behind this Digital Twin Ocean, which will rely on **analytical models** built by **thematic experts** in the **marine domain** and open the window to new, richer interactions with other communities and stakeholders, including ICT innovators, industry, NGOs and citizens. One of its valuable outputs will be advanced **science-based decision support capabilities**, including enhanced **predictive** and **simulation capacity**, at resolutions and accuracies compatible with the needs to respond to the urgent challenges and targets addressed by the Green Deal. It will also facilitate and trigger **innovation**, with huge, potential spillover effects across the policy objectives of the **EU Green Deal** (e.g. decarbonization) and the **blue economy**.

- **Web-based Open Science can expand our current body of Ocean knowledge by enabling and facilitating transboundary and interdisciplinary science:** Building Open Science capabilities in the marine domain will prepare the marine community to engage in collaborative research and science across a wealth of other disciplines, capitalizing on the opportunities brought by the single European data space, introduced by the European Data Strategy. The **European Data Strategy** recognises **EOSC** as the nucleus for a science, research and innovation data space that will become articulated with the **9 sectoral data spaces** foreseen by the strategy (see Figure 2), building a competitive data and knowledge economy in Europe endorsing Open Science[13]. By federating existing research data infrastructures, EOSC leverages national investments and adds value in terms of scale, interdisciplinarity and faster innovation, offering the possibility to access and reuse all publicly funded research data in Europe, across scientific disciplines and countries. This will allow the marine community to build on previous strategies of **Marine Knowledge 2020**[14], a key pillar of the 2011 Commission Communication on Blue Growth[15] and to further research and broaden the body of evidence of our interdependencies with the Ocean, for example advancing knowledge in the meta-discipline of **Ocean and Human Health (OHH)**[16], linking different priorities across the **EU Green Deal.**

[13] https://ec.europa.eu/digital-single-market/en/european-open-science-cloud
[14] https://ec.europa.eu/maritimeaffairs/sites/maritimeaffairs/files/docs/body/marine-knowledge-2020-green-paper_en.pdf
[15] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=COM:2014:254:REV1&from=EN
[16] https://sophie2020.eu/

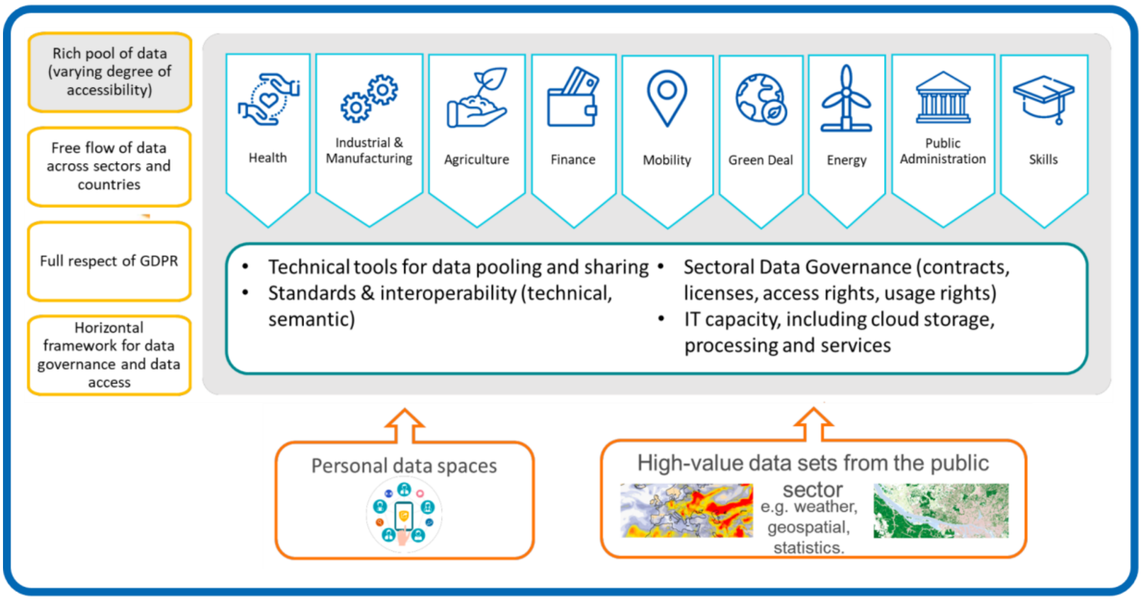


Figure 2: The European Data Spaces

Credit: Building the Green Data Space. INSPIRE Online Conference 2020

- **Web-based Open Science can contribute to establishing closer links and collaboration between marine researchers and the maritime industry for a cleaner and circular economy:** The Blue Economy Report 2020[17] recognises that the European Green Deal and the European Strategy for Data will necessitate reliable, accurate and accessible data for its initiatives. Whilst data initially originated from academia and the scientific community, increasing partnerships between trusted and long-term European marine data services e.g., EMODnet[18] and CMEMS and maritime industry are already leading to win-win collaborations, particularly in terms of data sharing. Improving knowledge, reducing operational risks or adding value to data through integration to deliver new, added value data products tailored to business needs are some of the obvious advantages that are beginning to trigger a cultural shift in data sharing amongst industry players. According to the aforementioned Report, the seven established sectors of the EU Blue Economy (marine living resources; marine non-living resources; marine renewable energy; port activities; shipbuilding and repair; maritime transport and coastal tourism) generated a gross value added (GVA) of €218.3 billion in 2018 and a total turnover of €749.7 billion, contributing 1.5% in terms of GVA and 2.2% in terms of employment to the EU-28 economy. Developing **synergies** in data sharing and **joint efforts** in **Open Science** could not only contribute to the **Marine Knowledge 2020 Strategy** goal of **improving data coverage** of European waters, but also to trigger **new data applications** and **Ocean solutions**, in support of a more sustainable and productive use of marine resources and space. In addition to these established sectors, the Report also notes other emerging and innovative sectors that could also contribute to and benefit from such win-win collaborations, including:
  - Ocean energy;
  - Blue bioeconomy and biotechnology;
  - Marine minerals;

[17] https://blueindicators.ec.europa.eu/published-reports_en
[18] https://www.emodnet.eu/emodnet-business-brochure

- Desalination;
- Maritime defense;
- Submarine cables.

The EU is currently a global leader in terms of marine renewable energy capabilities, hosting 70 % of global ocean energy (wave and tidal) in its waters. Others, including the Maritime Defence and algae sector, account for high job creation and turnover, which is only expected to increase. Some marine data infrastructures (e.g. EMODnet) already have growing collaborations and partnerships with offshore and coastal maritime industry in terms of data sharing and as users of data and data products for operations at sea. Large-scale databases also provide harmonised datasets to assess human activities (e.g. EMODnet Human Activities) and human impacts (for example, the EMODnet Chemistry thematic area for marine pollution). Demonstrating the potential of **Open Science** in the marine domain can contribute to important steps forward in advancing **open data**, for the benefit not only of large players but also SMEs across a whole range of industrial clusters and new industrial value chains, opening new opportunities for SMEs as the backbone of European Industry.

- **Web-based Open Science** can **boost Ocean literacy** and lead to **better informed, attentive and engaged citizens in Ocean** matters, contributing to **develop** a sense of **Ocean citizenship:** Web-based Open Science contributes to breaking down barriers to data access, which is the first step towards data democratization. But it also unfolds new, creative ways of using marine data to support Ocean literacy efforts, leveraging on new technologies. As the **European Atlas of the Seas**[19] shows, data can be used to develop new, interactive maps and educational resources that bring the Ocean closer to citizens. Algorithms, models and data products developed through Open Science efforts can be used to build virtual environments that visually replicate real-life marine habitats and ecosystems. Learning from the gaming industry, they can be set up for multiple purposes, from showcasing how both landscape diversity and species richness respond to various interventions planned as part of the EU Green Deal, to how building-with-nature solutions affect the marine landscape and its habitats, to delivering fully immersive experiences for educational or recreational purposes. The potential for developing closer, emotional connections between people and the marine realm opens a promising avenue towards winning societal support for sustaining a healthy Ocean, through increased awareness and understanding of the dynamics, interactions and evolution of seas and oceans and their role in our well-being and survival. Another way in which Open Science can contribute to Ocean citizenship is by further promoting and empowering **citizen science**, which in turn can contribute to feed the marine knowledge value chain by encouraging and enabling the generation of “non-scientific” data streams that complement and refine “traditional” methods of observation. Citizen science often leads to increased citizen engagement and participation in bottom-up actions, contributing to develop a sense of **Ocean citizenship**.

[19] https://ec.europa.eu/maritimeaffairs/atlas/maritime_atlas/

The UN Agenda 2030 & Ocean Decade of Ocean Science for Sustainable Development

A similar context of opportunity exists in the international arena closely related to the implementation of the **United Nations (UN) Agenda for Sustainable Development 2030**, a commitment to eradicate poverty and achieve sustainable development by 2030 world-wide. The Agenda was adopted in 2015 by Heads of State and Government across the World and encompasses 17 Sustainable Development Goals (SDGs) and 169 targets, providing a shared global vision towards **sustainable development** for all. All 17 SDGs are directly or indirectly connected to the Ocean. **SDG 14** "**Life Below Water**" is geared at **conserving and sustainably using the oceans, seas and marine resources**, but other SDGs (like **SDG 2** "**Zero Hunger**" or **SDG 13** "**Climate Action**") are also closely linked with global Ocean action. In 2017, the United Nations declared that a **Decade of Ocean Science for Sustainable Development** would be held from **2021** to **2030**. The Ocean Decade provides a common framework to ensure that **ocean science** can fully support countries to achieve the 2030 Agenda for Sustainable Development. The Decade is a transformative vision to deliver the "*science we need for the future we want*", geared at delivering the following outcomes:

- A **clean ocean** where sources of pollution are identified and reduced or removed.
- A **healthy and resilient ocean** where marine ecosystems are understood, protected, restored and managed.
- A **productive ocean** supporting sustainable food supply and a sustainable ocean economy.
- A **predicted ocean** where society understands and can respond to changing ocean conditions.
- A **safe ocean** where life and livelihoods are protected from ocean-related hazards.
- A **transparent** and **accessible ocean** with open and equitable access to data, information and technology and innovation.
- An **inspiring and engaging ocean** where society understands and values the ocean in relation to human wellbeing and sustainable development.

The Decade will provide a framework for addressing **key challenges**, including understanding and beating marine pollution; protecting and restoring ecosystems and biodiversity; sustainably feeding the global population; developing a sustainable and equitable ocean economy; unlocking ocean-based solutions to climate change; increasing community resilience to ocean hazards and changing humanity's relationship with the ocean. It specifically will address expanding the **Global Ocean Observing System**, creating a digital representation of the ocean; and delivering data, knowledge and technology to all, strengthening our capacities to produce and transfer knowledge applied to decision-making. Delivering these outcomes will require **open data** and **Open Science** to drive a **transparent** and **accessible Ocean**, together with underpinning all the thematic areas of the UN Decade, including facilitating Ocean literacy through democratising access to data, information and knowledge. **Open science** has the potential to **contribute** to key goals of the UN Decade to:

- Increase scientific knowledge, develop research capacity and transfer marine technology, taking into account the Intergovernmental Oceanographic Commission Criteria and Guidelines on the Transfer of Marine Technology, in order to improve ocean health (UN SDG 14).
- Improve forecasting of climate change, weather and ocean conditions to protect human activities in support of UN SDG 14 and other relevant goals, and of the objectives of related Conventions (for example, on biodiversity).

- Shorten the time span between research and innovation and foster economic value in the blue economy.
- Improve the professional skills and competences of those working and being trained to work within the blue economy and in the context of open data sharing.

The UN Agenda 2030 has already triggered a response from the global Ocean observation community, in recognition for the need to step up efforts towards a comprehensive understanding of the current and projected state of our ocean, seas and coasts, and the need for effectively monitoring the impact of new policies and management actions, in support of its implementation. The **Global Ocean Observing System** (GOOS), led by the **Intergovernmental Oceanographic Commission** of **UNESCO** (UNESCO-IOC) and co-sponsored by the **World Meteorological Organization** (WMO), the **United Nations Environment Programme** (UNEP) and the **International Science Council** (ISC) has launched a new **2030 Strategy** to advance international cooperation towards establishing an **operational ocean observation capability for all of the world's nations** that delivers the essential information needed for global sustainable development, safety, wellbeing and prosperity. UNESCO-IOC's "**International Oceanographic Data and Information Exchange**" (IODE) programme -which through decades has been able to collect, control the quality of, and archive millions of ocean observations- is closely collaborating in this effort through a network of over 80 oceanographic data centres in as many countries around the world, the **Ocean Biodiversity Information System** (OBIS) and the **Ocean Data Information Sources** (ODIS) catalogue. Complementing these efforts, the **Group on Earth Observations** (GEO) is building the **Global Earth Observation System of Systems** (GEOSS) and convening members of its global community to engage on the UN Agenda 2030 by working with the **UN Statistics Division** (UNSD), the **UN Inter-Agency Expert Group** on the SDG (IAEG-SDG) and the **UN SDG Custodian Agencies** for the **SDG Targets** and **Indicators**[20]. GEOSS is a set of coordinated, independent Earth observation (EO), information and processing systems that interact and provide access to diverse information for a broad range of users in both public and private sectors, sharing of environmental data and information collected from the large array of observing systems.

In 2016, the **G7** launched the **Tsukuba Communiqué**[21], recognizing the importance of **Open Science** and **ocean observations** to support the objectives of the **UN Agenda 2030**. In the framework of its **Future of the Seas and Oceans Initiative**, a dedicated working group of marine scientists and government science ministries from across the G7 agreed on action plans for 5 Action Areas, namely:

- Support the development of a global initiative for enhanced, global, sustained sea and ocean observing system.
- Support an enhanced system of ocean assessment through the UN Regular Process for global reporting and assessment of the state of the marine environment.
- Promote Open Science and improvement of the global data sharing infrastructure.
- Strengthen collaborative approaches to encourage the development of regional observing capabilities and knowledge networks, including supporting capacity building of developing countries.

[20] Anderson et al (2017) Earth observation in service of the 2030 Agenda for Sustainable Development https://www.tandfonline.com/doi/full/10.1080/10095020.2017.1333230
[21] https://www8.cao.go.jp/cstp/english/others/20160517communique.pdf

- Promote increased G7 political cooperation by identifying additional actions needed to enhance future routine ocean observations.

In its communication, the G7 acknowledged that **Open Science** can change the way research and development (R&D) is undertaken, with emerging findings leading to far greater global collaboration and encouraging a much broader range of participants and stakeholders.

The EU's long-term trusted marine data services **EMODnet** and **CMEMS** are already contributing key open access data and data products that can be used to address the UN 2030 Agenda and Sustainable Development Goals. In addition, the **EMOD-PACE** project supports better ocean data cooperation with China[22] and the international ocean governance initiative **EU4OceanObs** contributes to coordinating European efforts to global initiatives, including **G7**.

As showcased in **Section 5**, the Blue-Cloud's efforts will further enhance the EU contribution to the UN Agenda 2030 by demonstrating the potential of Open Science to contribute to SDGs 3, 13 and 14, as well as by creating new, future capabilities to advance web-based, collaborative Open Science in the marine domain, aligning with key international initiatives.

Web-based Open Science: Delivering scientific, economic, technical, technological and cultural impact in support of European and global policy objectives

To capture the main policy objectives surrounding the developments of the Blue-Cloud project, Table 1 summarizes the main priorities of the **EU Green Deal** and the **UN Agenda 2030** from an **Ocean policy perspective**, feeding from the two guiding instruments that will guide their implementation: **Mission Starfish** and the **UN Decade for Ocean Science for Sustainable Development**, respectively. Selected targets connected to implementation of these policies have also been included in this table, some for informative purposes only, but others because they are relevant from the perspective of the Blue-Cloud "territory" (i.e., the marine knowledge value chain, which is described in the next section).

As it can be seen in Table 1, the policy priorities outlined in the **EU Green Deal** and the challenges that the **UN Decade of Ocean Science for Sustainable Development** will seek to address are **well aligned** and provide a **common, clear framework for action** at European level that also delivers on a global scale. The table also summarizes how **Open Science** could **impact these policy objectives**, delivering results along five distinctive dimensions or "bottom lines": **scientific**, **economic**, **technological**, **technical** and **cultural**.

The window of opportunity that **web-based Open Science** brings towards delivering on the societal objectives laid out by the **EU Green Deal** and the **UN Agenda 2030** has a clear frame and it is wide open. The question is, are we ready to seize this window of opportunity? In the next section, we look deeper into **Europe**'s **marine knowledge value chain** to signal some of the challenges ahead.

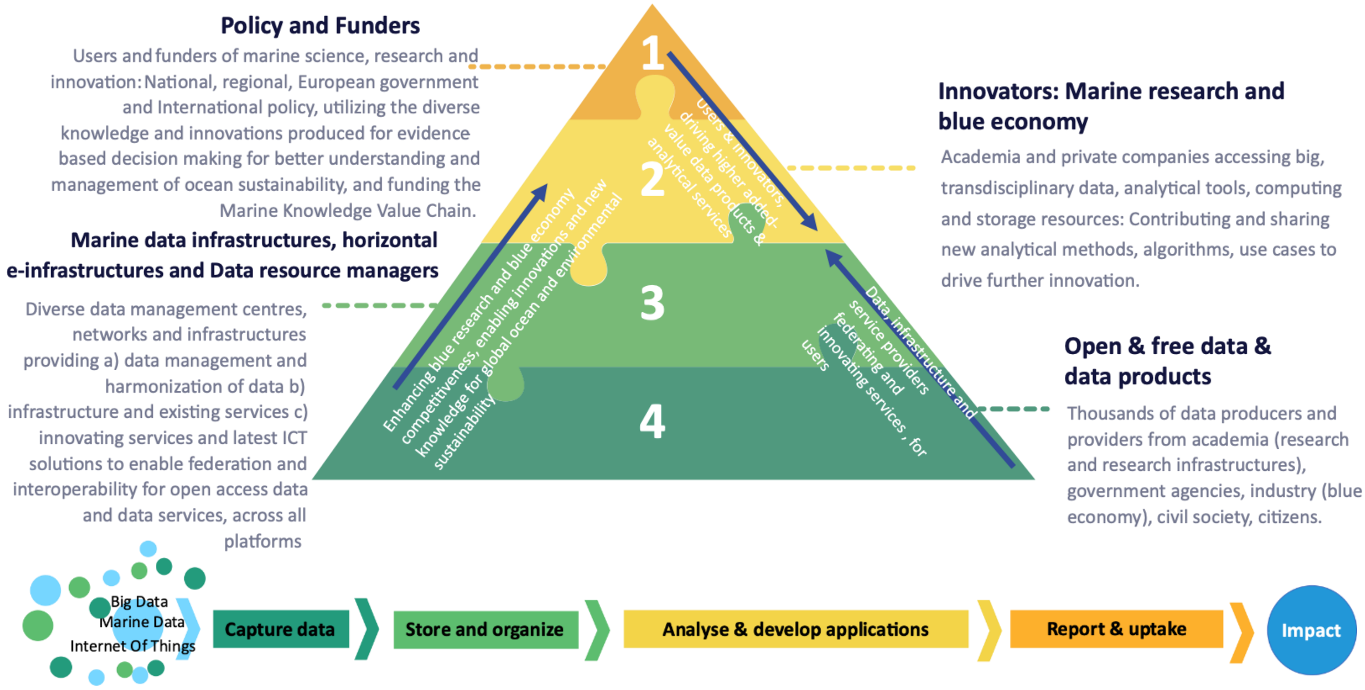
[22] https://www.emodnet.eu/en/emod-pace

Table 1: 2030 Policy Priorities and/or Challenges Identified Across Key Policy Frameworks. How Open Science in the marine domain can contribute to relevant EU and global policy objectives

2030 Policy Priorities and/or Challenges Identified Across Key Policy Frameworks EU Green Deal & Mission Starfish UN Agenda 2030 & Decade of Ocean Science for Sustainable Development						
Policy priority:	**Cross-cutting Themes** **Pollution** · **Biodiversity** · **Food** · **Climate Change** · **Ocean Hazards**	**Blue Economy**	**Ocean Observation**	**Ocean Modelling & Forecasting**	**Data & Knowledge Access**	**Ocean Citizenship**
EU Green Deal (policy priorities 2030) **& Mission Starfish** (selected policy targets 2025-2030)	• A zero-pollution ambition for a toxic-free environment (zero spill and plastic litter generation; underwater acoustic emissions are reduced by at least 50% by 2030). • Preserving and restoring ecosystems and biodiversity (active regeneration of 20% of degraded habitats; 30% of EU waters are highly to fully protected; 50% of DNA of life in ocean & waters is fully sequenced and publicly available by 2030). • From 'Farm to Fork': designing a fair, healthy and environmentally friendly food system (end overfishing; zero carbon aquaculture by 2030). • Increasing the EU's climate ambition for 2030 and 2050.	Mobilising industry for a clean, circular economy. Supplying clean, affordable & secure energy. Building and renovating in an energy & resource efficient way.	20% of data collections from citizen science. North Atlantic seabed fully mapped by 2025. European seabed fully & coherently mapped in high-resolution by 2030.	European digital twin pilot of European oceans and waters (DTO) is operational by 2025. Global digital twin of all oceans and waters operational by 2030.	All (marine & freshwater) data collected by EU Member States pooled centrally & accessible to all by 2025. Global marine & freshwater observation streamlined: All global data pooled centrally & accessible to all.	Activating education and training. Each European is a citizen of our ocean and waters.
UN Decade of Ocean Science (challenges 2030)	• Understand and beat marine pollution. • Protect and restore ecosystems and biodiversity. • Sustainably feed the global population. • Unlock ocean-based solutions to climate change. • Increase community resilience to ocean hazards.	Develop a sustainable and equitable ocean economy.	Expand the Global Ocean Observing System.	Create a digital representation of the ocean.	Deliver data, knowledge and technology to all.	Change humanity's relationship with the ocean.
How web-based Open Science in the marine domain can support these policy objectives	Open science can contribute to: • Improve our knowledge and understanding of marine ecosystems, including their status. underlying processes and interconnections. • Monitor trends and evolution around cross-cutting themes, predicting future developments around different, simulated scenarios. • Evaluate and guide policies, strategies and plans, providing scientific evidence to support and/or improve and/or update them.	Trigger innovative, marketable science-based applications and solutions to challenges. Shortening the time span between research and innovation in frontier fields.	Develop new observation technologies and capabilities through science-based innovation.	Develop models for application in "digital twins". Train future scientists, researchers in data modelling. Train future users of DTOs. Develop scenarios & policy options.	Support federation of catalogues of EU blue data infrastructures and align FAIR standards towards global capability. Promote data sharing. Train future scientists & researchers in FAIR data management.	Develop user-friendly interfaces to bring Ocean science closer to citizens.
Bottom line	**Scientific**	**Economic**	**Technological** **Economic**	**Scientific** **Economic**	**Technical** **Cultural**	**Cultural**

2.3 The European marine knowledge value chain

Over the past decades, Europe has developed an impressive capability for marine observation, anticipating the need to support and advance our knowledge of the Ocean. These efforts have contributed to the emergence of an advanced **European "marine knowledge value chain"** which brings together multiple European actors and assets (see Figure 3).



The Marine Knowledge Value Chain

Figure 3: The European marine knowledge value chain: Data, infrastructure and service providers and users, including marine research and blue economy, policy makers and funders. Credit: MARIS and Seascape Belgium

Europe's efforts to support increased knowledge of our seas[23], aligned with action at national level, has led to a remarkable increase of abundance of **marine data**, stemming from academia, major Research & Development projects and initiatives, Research Infrastructures (RIs), industry and civil society. The emergence of citizen science and technological developments such as the "Internet of Things (IoT)" anticipates unprecedented potential to exponentially grow such abundance. A resilient, existing European network of distributed **infrastructures** for data gathering, handling and sharing is making it possible to **manage**, **enhance** and **channel** this wealth of data into an additional layer of actors in the marine and maritime research and blue economy communities. Marine data infrastructures and services, such as the **European Marine Observation and Data Network** (**EMODnet**) and the **Copernicus Marine Environmental Monitoring Service** (**CMEMS**), together with **Research Infrastructures** and **horizontal e-infrastructures**, play a key, strategic role in making data **discoverable**, **accessible**, **interoperable** and **reusable** to **marine data user communities**. Through their efforts, these communities are able to apply new **analytical** methods to **transform available data into knowledge**, adding modelling and forecasting capabilities and opening new opportunities to drive **innovation**. The resulting knowledge is key to inform better **science-based policies**, using existing

[23] COM(2014) 254 final/2 Innovation in the Blue Economy: Realising the potential of our seas and oceans for jobs and growth https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=COM:2014:254:REV1&from=EN

evidence to drive a **sustainable use and conservation of the Ocean** through sound management and governance.

Further insight into how this **marine knowledge value chain** has shaped throughout the years, the **progress** achieved so far and the **challenges** that it is currently facing, is key to understanding the **added value** and **momentum** that the Blue-Cloud's efforts are striving to deliver.

In 30 years, Europe has developed an impressive Ocean observing capability

Fit for purpose and sustained ocean observations are an essential part of worldwide efforts to understand marine social-ecological systems. Observations can be samples collected on ships, measurements from instruments on fixed platforms, autonomous and drifting systems, submersible platforms, ships at sea or remote observing systems such as satellites and aircrafts. In Europe, oceanographic and marine data are collected by several thousands of **research institutes**, **governmental organizations**, and **private companies**. Various heterogeneous observing sensors are installed on research vessels, submarines, aircraft, and moorings, drifting buoys, gliders, floats, fixed platforms, and satellites. These sensors measure physical, chemical, biological, geological and geophysical parameters, with further data resulting from the analysis of water and sediment samples for a wide variety of parameters. Governments deploy monitoring programs at national and regional level to support national marine governance goals and European policy commitments. Research institutes in Europe operate research vessels and many other platforms and instruments for gathering in-situ observations and samples in European waters, but also on a global scale. Marine data are also collected by private organizations in support of economic activities. In the past 30 years, the **EU** has also developed a **world-leading position** in **ocean observation** from **space**. In 2014, the EU launched **Copernicus**, the European Union's Earth observation and monitoring programme, served by a set of dedicated satellites and in-situ systems. Today, it produces a wealth of data and information on the Earth's sub-systems (land, atmosphere, oceans and inland waters) and cross-cutting processes (climate change, disaster management and security). It is fair to say that thanks to these efforts and through close collaboration with its Member States, the **European Union** boasts an unrivalled **marine observation** and **forecasting infrastructure** that is increasing our understanding of the oceans.

Europe has made huge progress advancing FAIR and open data in the marine community

Only a few decades ago, marine data from all those observations described above were difficult to find, hardly accessible and extremely cumbersome to put together because of different standards, nomenclature and baselines. As a response to the need to develop a more coherent and interconnected system towards exploiting data from observations, since the early nineties, a wide range of initiatives funded and/or supported by the EU have not only contributed to further develop Europe's capabilities for collecting and managing marine in-situ and remote sensing data, but also to develop and advance **data centres** and **data management systems** servicing the marine domain. As a result, great progress has been achieved in terms of developing **data standards** and **data management services**, through the establishment of dedicated infrastructures.

EMODnet, CMEMS and the Data Collection Framework collectively implement the **Marine Knowledge 2020 Strategy**. Over the past decade there has been significant development towards **FAIR** and **open data** in the **marine domain** integrating data from various sources, harmonising and standardising these to EU e.g., INSPIRE (and increasingly international) standards and making more information freely available as interoperable data layers and data products following the principle of "*measure once, use many times*".

These EU data brokers and services have many operational partnerships with other, well established, European **marine data management infrastructures.** For instance, the EU infrastructure SeaDataNet is a network of national ocean data centres that are a central part of the EMODnet data flow for a number of thematic areas. And EurOBIS (marine biodiversity focus) is the EU counterpart of the global OBIS data infrastructure, with very close links and data flow with EMODnet Biology. Other thematic research and distributed infrastructures include **ELIXIR-ENA** (life sciences) and **ICOS-Ocean** (ocean carbon component of ICOS), amongst many others, are funded by, and operated by EU Member States.

These blue infrastructures have already started working together to continue supporting, expanding and mainstreaming the application of FAIR data management practices, collaborating through EU funded initiatives such as **ENVRI-FAIR**, which is working towards better analysing and improving progress towards FAIR in four environmental subdomains, including the marine domain, as well as aligning requirements towards **EOSC**. Adding to these collaborations, blue infrastructures are working in close interaction with international initiatives led by the **United Nations Intergovernmental Oceanographic Commission** (IOC), the **World Meteorological Organization** (WMO), the **Food and Agricultural Organization** (FAO), the **Group on Earth Observations** (GEO) and the **International Council for the Exploration of the Sea** (ICES), amongst others.

Thanks to all these efforts, marine data is now recognized as a **public good** in Europe and multi-resolution maps of all Europe's seas and oceans are available, together with large catalogues of **open** and **free data products** that span over multiple disciplinary themes. European efforts to promote and apply the principles of free and **open access** and **interoperability** in the marine domain are demonstrating clear value, through long-term data brokers and services e.g., **EMODnet**, and community research and development funded through **European Research Framework Programmes**, most recently **Horizon 2020** and the upcoming **Horizon Europe.**

The marine community has responded positively to pilot analytical frameworks seeking to support web-based Open Science in the marine domain

Efforts along the marine knowledge value chain have also extended into applying technology towards enhancing the **analytical capabilities** of the marine community. Data availability is not a sufficient condition to support systemic and transformative solutions for healthy oceans, seas, coastal and inland waters. As **Mission Starfish** points out, "*while the effects of pollution, human activities and climate change are documented and observable, significant gaps in our understanding and knowledge of the hydrosphere and the challenges it faces remain*". Gaps in our understanding of how to address these challenges effectively and holistically, also remain. As mentioned before, the emergence of IoT and citizen science may bring unprecedented potential to exponentially grow the abundance of Ocean data. The question is, are we ready to make the most out of this wealth of data?

Through its Research and Innovation (R&I) programme, the EU has supported the development of **e-infrastructures** as the pathway towards integrating **data services** with **analytical tools** in a seamless way, increasingly at large scale. Piloting initiatives such as the i-Marine project, the BlueBridge project, Marine Analyst or the SeaDataCloud VRE development have shown how comprehensive data management solutions can support the application of a science-based ecosystem approach to specific domains (e.g. fisheries or marine spatial planning), as well as providing attractive capacity building environments for interdisciplinary research communities involved in increasing our knowledge of the marine environment. Thanks to these efforts, a growing **catalogue** of **virtual laboratories** ("**virtual labs**") is making **Open Science** more **accessible** to researchers in the marine community, providing them with greater **computational resources** and **tools**. The e-infrastructure D4science[24] has serviced more than 2,500 users in the marine domain through its over 60 "blue" virtual labs, as a solid proof of the growing awareness and interest from the marine research community of the opportunities to thrive with the use of new digital technologies. D4Science currently services 7,000 user from 44 countries.

What is a virtual lab?

A virtual laboratory is a web-based, problem solving environment that enables users across different physical locations and expertise to collaborate efficiently in an ongoing way, providing them with tools and services to share data, analytical methods and results towards solving complex problems.

Blue data infrastructures are also increasingly involved in generating models and new, knowledge-oriented data products which are run by their teams or made available as services for external users from research, government and industry. In 2018, the EU launched **WEkEO**[25] as one of the five Copernicus Discovery and Access Service (DIAS) to provide a single access point to all Copernicus **data** and information, alongside **processing resources**, **tools** and other relevant data. The overarching objective of DIAS is to enhance access to Copernicus data and information for further use in an efficient **computing environment** implementing the paradigm of "bringing the user to the data", as one condition for unlocking the potential value of Copernicus for innovation, science, entrepreneurship, business and economic growth and science-based policies. WEkEO is the service for **marine environmental data**, **virtual environments** for data processing, and skilled **user support**.

But a number of challenges are yet impeding progress to unlock the full potential of web-based Open Science in the marine domain, in support of the EU Green Deal and the UN Agenda 2030

In spite of this extraordinary progress, a number of barriers are still impeding the marine community from embracing **Open Science** and unlocking its full potential in the marine domain, in support of wider societal objectives.

In 2020, the Blue-Cloud project launched an initial phase of dialogue and consultations with different stakeholders across the Blue-Cloud community[26]. Although the dialogue was focused on better identifying their needs and expectations with regard to the Blue-Cloud project, it was also instrumental to gain insight into some of the **challenges** that are currently hindering Open Science in the marine domain, and therefore a more productive marine knowledge value chain in Europe:

- Through the stakeholder consultation in 2020, a survey was launched towards identifying the most pressing "pain-points" for the community in its dealings with marine data. A **large majority** of respondents who identified as **users of marine data** signalled as "significant" or "very

[24] https://services.d4science.org/explore
[25] https://wekeo.eu/web/guest/home
[26] Summary of Key Messages & Recommendations of the 1st Phase of Consultations with the Blue-Cloud Community (Seascape Belgium, 2020)

significant" problems: "*harmonizing data from different sources*" (86%); "*performing computing and analytical processes across data sets*" (78%); "*having to search for data in different portals*" (77%); "*sharing and processing large data sets from observations*" (75%) and "*finding the data for which I am searching*" (72%). A **majority** of respondents who identified as **producers of marine data** signalled as "significant" or "very significant" problem "*transforming datasets into higher added value products with specific applications*" (69%). Some of these results come to support **Mission Starfish**'s claim that additional effort is required to further federate data infrastructures, to continue improving the **discovery**, **access** and **interoperability** of Ocean related data. Initial stakeholder dialogue with the **Blue-Cloud External Stakeholder & Expert Board** (ESEB)[27] has further identified the lack of **standards** for **metadata** as one of the main challenges blocking progress towards FAIR in the marine community. A recent report of the EOSC Secretariat supports the idea that **properly structured metadata** to aid findability, along with provision of services via uniform and compatible encodings using community-adopted standards, will be required to support better discovery and access to data but also machine-based processing of data flows in the marine domain, in support of interoperability[28]. The Blue-Cloud project, as described in the next section, is seeking to contribute with effective steps towards larger on-going efforts for better discovery, access and interoperability in the marine domain, facilitating dialogue amongst leading blue data infrastructures, followed by concrete actions, as explained in **Section 3**.

- However, pooling data centrally will not solve some of the most compelling needs of the marine community regarding their need to **process large data sets** from observations and performing **computing** and **analytical processes** across data sets. Nor will it immediately inspire them to transform datasets into higher added value products. Transforming **data** into **knowledge** and into **innovative applications**, **products** and ultimately **solutions** to current societal challenges will demand **computing resources**, **stronger analytical frameworks** for Ocean related data and **closer public-private collaboration** between scientists, researchers, innovators, entrepreneurs and business developers. Providing broader access to **user-friendly collaborative tools** and **services** towards **transforming Ocean data** into **societal knowledge** is one of the actions proposed by Mission Starfish towards the objective of "healthy oceans, seas, coastal and inland waters". While the cost of computing may have been a limiting factor for the success of Open Science environments in the past, the emergence of new, **"pay-on-demand" services** for computing resources (allowing users to only pay for the computing resources they use, rather than having to pay for more costly monthly or yearly subscriptions or to invest in computing infrastructure themselves) opens new opportunities for the development of collaborative tools and schemes.
- For such collaboration schemes to deliver results, however, they require engaging a **critical mass of players**. At the moment, as initial dialogue with the Blue-Cloud ESEB has highlighted, Open Science is not yet mainstreamed in the marine domain -neither across public nor private organizations. In general, and in spite of the early signals of a significant cultural change, **uptake** of sound (FAIR) **data management** practices -one of the underlying enablers of Open Science at

[27] https://blue-cloud.org/eseb
[28] European Commission - Directorate-General for Research and Innovation (2020) Six Recommendations For Implementation of FAIR Practice https://ec.europa.eu/info/publications/six-recommendations-implementation-fair-practice_en

a broad scale- is still low across the academic sector. So is the uptake of **Open Science tools** and **environments**. With 1.7 million researchers in Europe, and given the scale of the challenges ahead, a penetration of 7,000 users in a successful Virtual Research Environment (VRE) signals a long road ahead for the scientific community. As the aforementioned report of the EOSC Secretariat also points out, it is widely seen that researchers do not see sufficient benefits of **FAIR** data, and therefore are not willing to put in the efforts in implementing FAIR data management practices. A similar reasoning could be made for **Open Science**. The report points at the need to **fund** FAIR **awareness-raising**, **training** and **education** in academic and research institutions, as well as **funding**, **rewarding** and **recognising improvements** of **FAIR practice** and developing and **monitoring** adequate policies for FAIR data and research objects. It also highlights that while a successful strategy towards mainstreaming wide adoption of FAIR will benefit from top-down policies, FAIR seems to evolve naturally and more effectively when motivated by **need** and **common benefits**. The report signals "communities" as being key to advance towards agreed formats for data, common vocabularies, metadata standards and accepted procedures for how, when and where data will be shared.

Therefore, making the case for both **FAIR data management** and **Open Science** in the **marine domain** by **showcasing** its **benefits** is key to **achieve buy-in** from the marine, academic, scientific and research communities. Building an initial community of **early Open Science advocates and practitioners** can be a powerful contributor/catalysator to that objective. **Virtual Research Environments** provide a good ground for researchers to experiment with available **tools** and **facilities** and to grasp the benefits of Open Science, acting as a **bottom-up community building tool**. However, through its Report, the EOSC Secretariat also alerts that "*solving findability and accessibility of data within a discipline by bringing the data together in a virtual research environment can result in a larger silo of data that no longer interoperates with other disciplines*". While virtual labs have a clear, demonstrative value in the short to medium term, building them to **anticipate easy integration into larger environments** that allow their **exploitation** by **larger, interdisciplinary audiences** can be a way of anticipating and mitigating this risk.

FAIR data practices are equally challenged across **public agencies**, the **private sector** and **NGOs**. Showing the benefits of FAIR data management practices is equally important to progress public-private collaboration efforts. Initiatives like EMODnet's **data ingestion service** are an important step in the right direction, as they encourage all types of actors across the public sector, private sector and civil society to contribute and share datasets for further processing as **FAIR data** and in most cases towards publishing as **open data**. Like amongst the research community, making the **benefits** of this **win-win collaboration visible** is critical to achieve buy-in from the private sector and civil society. However, again, FAIR data is a necessary but not a sufficient condition for Open Science. The **drivers** and **motivations** of private companies and NGOs to engage in sharing of data, processes and analysis results are different from those of the academic and the scientific community. Understanding these drivers and motivations and creating the conditions to respond to their needs and expectations will be key to bring these actors into collaborative Open Science schemes geared at delivering solutions to societal challenges. On-going initiatives powered by the private sector (e.g. Kaggle, featured in the box below) can provide inspiration towards developing **collaborative Open Science environments** that provide win-win results to all stakeholders involved, by generating value that aligns with their distinctive motivations.

Kaggle: An example of win-win collaboration in support of data science in the private sector

Recently acquired by **Google**, online data science and machine learning community Kaggle is home to over **one million users** ranging from computer science Ph.D. holders conducting cutting edge research to absolute beginners. Kaggle is best known for its data science competitions that offer (substantial) cash prizes, but it also serves as an educational tool for autodidacts as well as a place to present one's portfolio of related work. By connecting talented data scientists with tough problems, motivating them through lucrative cash prizes, and assisting their professional development through educational and portfolio resources, Kaggle creates substantial value for its users. Private partner organizations can similarly develop custom-built solutions for their business challenges while identifying the best talent to recruit. Finally, Google earns revenue from these partners while at the same time building its credibility in the arena of data science, familiarizing users and building customer loyalty towards their tools and lowering barriers for the uptake of other related commercial services[29].

In summary, in spite of its impressive capabilities, Europe's marine knowledge value chain still faces challenges towards further enhancing its performance and unlocking the full potential of Open Science, for the benefit of all engaged actors and ultimately, European citizens. On the one hand, it requires "lubrication" at different points of its underpinning gear, including **advancing data management practices towards FAIR principles** across blue data infrastructures; making more of the currently available data **open** (and to the extent possible, **free**); and **connecting to other data spaces** (as described in section 2.2) to support interdisciplinary research. But alongside responding to those challenges, Europe's marine knowledge value chain can benefit from developing more robust **frameworks** for **analysing data**. Harnessing the power of **web-based, "cloud" technologies** can boost the value chain's **productivity**, bringing a huge leap forward in the way that **marine data** -as well as the **processes** and **results** linked to **exploiting** and **analysing** such data- is **used**, **shared**, **enriched** and **applied**. Previous experiences have led the way towards testing and advancing technological infrastructures, methods and tools towards this end, introducing opportunities to pilot supporting frameworks for Open Science along the marine knowledge value chain. However, they were limited in their scope to influence changes both upstream ("data access") and downstream ("sharing outputs" across different communities and gearing efforts towards bridging the "science-policy" gap). The **Blue-Cloud project** is a step forward in that direction, **evolving existing computing** and **analytical capabilities** and underpinning them with **easier access to data** delivered across key blue data infrastructures through the use of **common languages** (standardized metadata), seeking to signpost potential solutions towards **federating their data services**, while showcasing how **Open Science** can deliver **impact** on key, **priority policy areas** catering to a **broader range of users**.

[29] Kaggle: Building a Market for Data Science (and Scientists) https://digital.hbs.edu/platform-digit/submission/kaggle-building-a-market-for-data-science-and-scientists/#:~:text=The%20online%20data%20science%20and,edge%20research%20to%20absolute%20beginners.

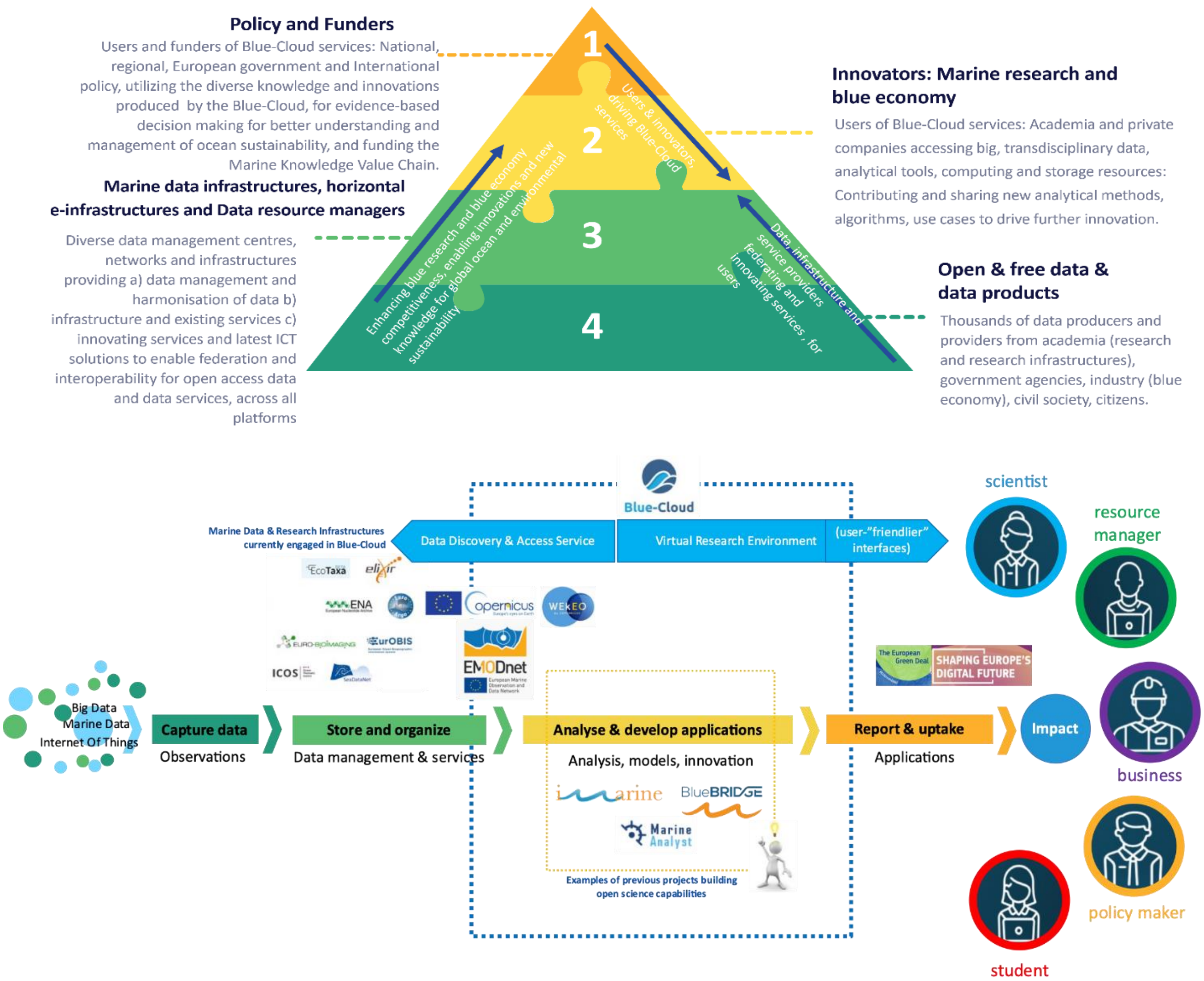


Figure 4: Blue-Cloud's contribution to Europe's marine knowledge value chain. **Credit**: Seascape Belgium

In the next section, we look deeper into the Blue-Cloud's efforts and how they are, and will contribute to, **unlocking the potential of web-based Open Science** along Europe's **marine knowledge value chain**, in support of the **European Green Deal** and the **UN Agenda 2030**.

3. Blue-Cloud's added value: Demonstrating the potential of Open Science in the marine domain

The Blue-Cloud's efforts are geared at enhancing Europe's **marine knowledge value chain**, leveraging on **Open Science** to deliver on the objectives of the **EU Green Deal** and to contribute to the **Sustainable Development Goals** of the **UN Agenda 2030 SDGs**. What specific **added value** will the 3-year Blue-Cloud project contribute by 2022? What challenges and opportunities will require additional time and resources to be properly addressed, towards delivering **higher impact** into the **medium** and **long term**? In this section we respond to these key questions first, before looking deeper into the practical, technological and demonstrative solutions developed to date through the Blue-Cloud project and analysing potential ways of further developing them into the future.

3.1 Defining Blue-Cloud's vision, mission & strategy 2022

The Blue-Cloud project launched in October 2019 aiming to **demonstrate the potential of Open Science in the marine domain**. This is the vision that captures what the Blue-Cloud is seeking to achieve by 2022.

Blue-Cloud Vision 2022
"Demonstrating the potential of Open Science in the marine domain".

Initial stakeholder dialogue and consultations in 2019-2021 have led to the articulation of a short-term **mission** that will guide the project through its implementation, towards delivering on its vision:

Blue-Cloud Mission 2022
"To promote the sharing of data, processes and research findings in the marine domain by delivering a ***collaborative web-based environment*** *that enables* ***Open Science****, underpinned by simplified access to a wealth of easily discoverable and interoperable marine data and products".*

The Blue-Cloud project will deploy a **three-fold strategy** towards delivering on its 2022 mission, building the foundations towards future capitalization and continuation of its efforts:

- Showcasing how using web-based "**cloud**" tools to share **data**, **analysis** and **processes improves the productivity of scientific research** (delivering a strong "case" for **researchers** to engage in collaborative Open Science).
- Showcasing how **federating access to marine data** made available through different infrastructures through use of **common standards** and coupling them with **computing resources, enhances** the **users' experience** and the **perceived value** delivered by marine data and research infrastructures to users, as well as attracting new users (delivering a strong "case" for **infrastructures** to continue collaborating beyond the Blue-Cloud project and to invite new infrastructures to join in this collaboration).
- Showcasing how **enhanced scientific productivity** leads to specific Ocean **knowledge** and **solutions** in support of the **EU Green Deal** and its contribution to the **UN Agenda 2030**

(delivering a strong "case" for funders and **policy makers** to continue building on the Blue-Cloud's efforts).

To implement this three-fold strategy, the Blue-Cloud project is developing a **web-based, cyber platform** (as described in Section 4) that provides users of marine data with enhanced **analytical capabilities** to enable **Open Science**, by providing them with:

- Powerful cloud-computing resources.
- A range of analytical tools.
- Simplified access to data from in-situ and satellite observations, data products and model outputs available across different blue data infrastructures.

In the **short-term**, the project will deliver key **assets** in support of delivering on its vision, including:

- A **Data Discovery & Access Service**, offering quick discovery and access to a range of data catalogues from different blue data infrastructures (as described in **Section 4**).
- A **Virtual Research Environment** (as described in **Section 4**), offering a range of **virtual labs** catering to different **users** in the marine domain, which provide a range of **domain specific services** (as described in **Section 5**) and providing the necessary tools and services for other users to **open new, virtual labs** to support their **research** and **innovation** objectives, on demand.
- A **community** of **Open Science "early practitioners"** revolving around the 5 Blue-Cloud demonstrators that showcase how Open Science can contribute to key policy objectives set out by the EU Green Deal and the UN Agenda 2030 (as described in **Section 6**).

In the **medium-** and **long-term future**, the Blue-Cloud aspires to upscale the results of this effort, aligning with wider developments at European level to unravel new opportunities for innovation.

3.2 Building on existing assets

To deliver on its vision of **demonstrating the potential Open Science in the marine domain**, the Blue-Cloud project has brought together for the first time some of the key **blue data infrastructures**, **Research Infrastructures** and **e-infrastructures** currently servicing the **marine knowledge value** chain in **Europe**, namely (see Figure 5):

- **Blue (marine data & research) infrastructures:**
    - Copernicus & CMEMS (climate and ocean analysis and forecasting)
    - EMODnet (bathymetry, biology, chemistry, geology, physics, seabed habitats and human activities)
    - ELIXIR-ENA (life sciences, biogenomics)
    - Euro-Argo and Argo GDAC (ocean physics and marine biogeochemistry)
    - EuroBioImaging (microscopy)
    - EurOBIS (marine biodiversity)
    - ICOS-Marine (carbon)
    - SeaDataNet (marine environment)
- **e-Infrastructures:**
    - D4Science

▪ EUDAT
▪ WEkEO

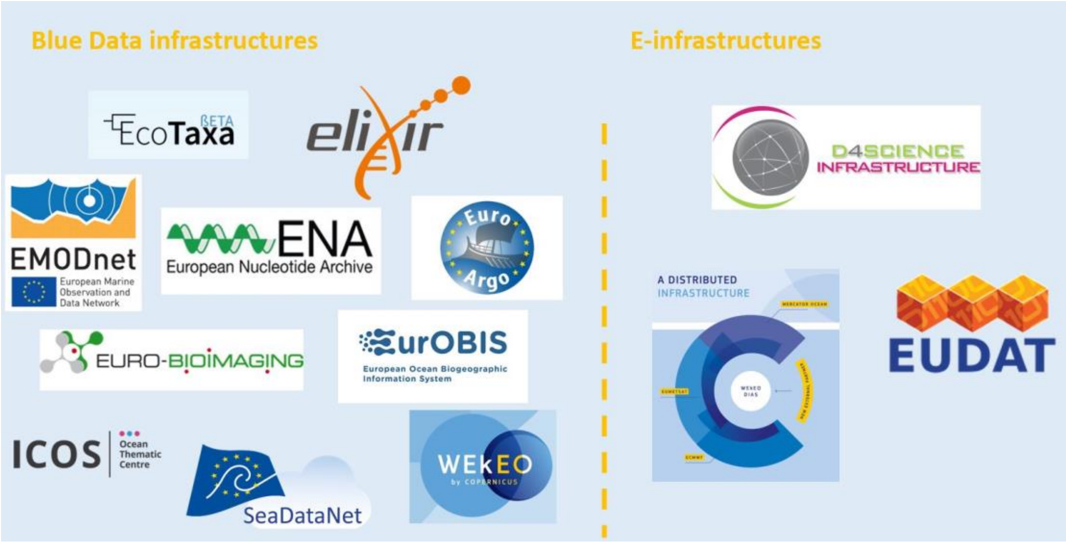


Figure 5: Leading infrastructures bundling their forces for the pilot Blue-Cloud project. **Credit**: MARIS

The Blue-Cloud project has opened a dialogue amongst these key players to explore a common approach towards making the wealth of data available across these different infrastructures more easily **discoverable** and **accessible** to users, but also more easily **combined** and **exploited** by them to support their **analytical research needs** and **objectives**. While the Blue-Cloud project will only be working with selected data catalogues and not with the full offer available across the participating blue data infrastructures, this effort will deliver a guiding reference towards further broadening its scope with additional catalogues from these and potentially other European (and international) blue data infrastructures and towards integration with **EOSC**.

The challenges around data above are closely related to the recent discussions around Data Space concepts[30] aligned with EOSC Core and EOSC Exchange of delivering services. New rules from the European Commission in a proposal for a Regulation on European Data Governance issued last November 2020 will pave the way for data to be harnessed and pave the way for sectoral European data spaces to benefit society, citizens and businesses. The selected data catalogues within Blue-cloud can make up part of the Data Spaces' concepts offering an environment for hosting and for processing research data in support of European science.

3.3 Targeting and delivering added value to key stakeholders

Within the **marine knowledge value chain**, the **Blue-Cloud** project will contribute to bridge and connect different communities, bringing value to **six stakeholder groups** that are central to its implementation, evolution and sustainability, namely:

- The **marine & maritime research community**, which produces and analyses data to create knowledge of the Ocean.

[30] https://ec.europa.eu/digital-single-market/en/news/commission-proposes-measures-boost-data-sharing-and-support-european-data-spaces

- The "**blue economy**" **entrepreneurs**, **SMEs** and **industry** that make use of the Ocean and of available knowledge to deliver products and services that satisfy society's needs.
- The **Policy, decision-making** and **governance institutions** that provide the legal and administrative frameworks to manage and preserve the Ocean.
- The **marine data, RIs** and **e-infrastructures** that are in dialogue with the Blue-Cloud but also other existing ones that are contributing towards the abundance of FAIR and open data in the marine domain and towards providing computing resources and digital services.
- The **European Open Science Cloud**, which the Blue-Cloud seeks to connect to, to enable trans-disciplinary and transformative research and innovation.
- The **ICT sector** that is driving new breakthroughs in the use of artificial intelligence, big data and machine learning.

These key groups are central to the **Blue-Cloud** as they represent: the Blue-Cloud's key "**target users**" (marine & maritime researchers; "blue economy" entrepreneurs, SMEs & industry; Ocean monitoring, management and governance institutions); the Blue-Cloud "**strategic allies**" (partners along the marine knowledge value chain, including blue data infrastructures and e-infrastructures; AI, ML & Big Data and data service providers; EOSC); and the Blue-Cloud "**enablers**" (policy makers and international funding institutions).

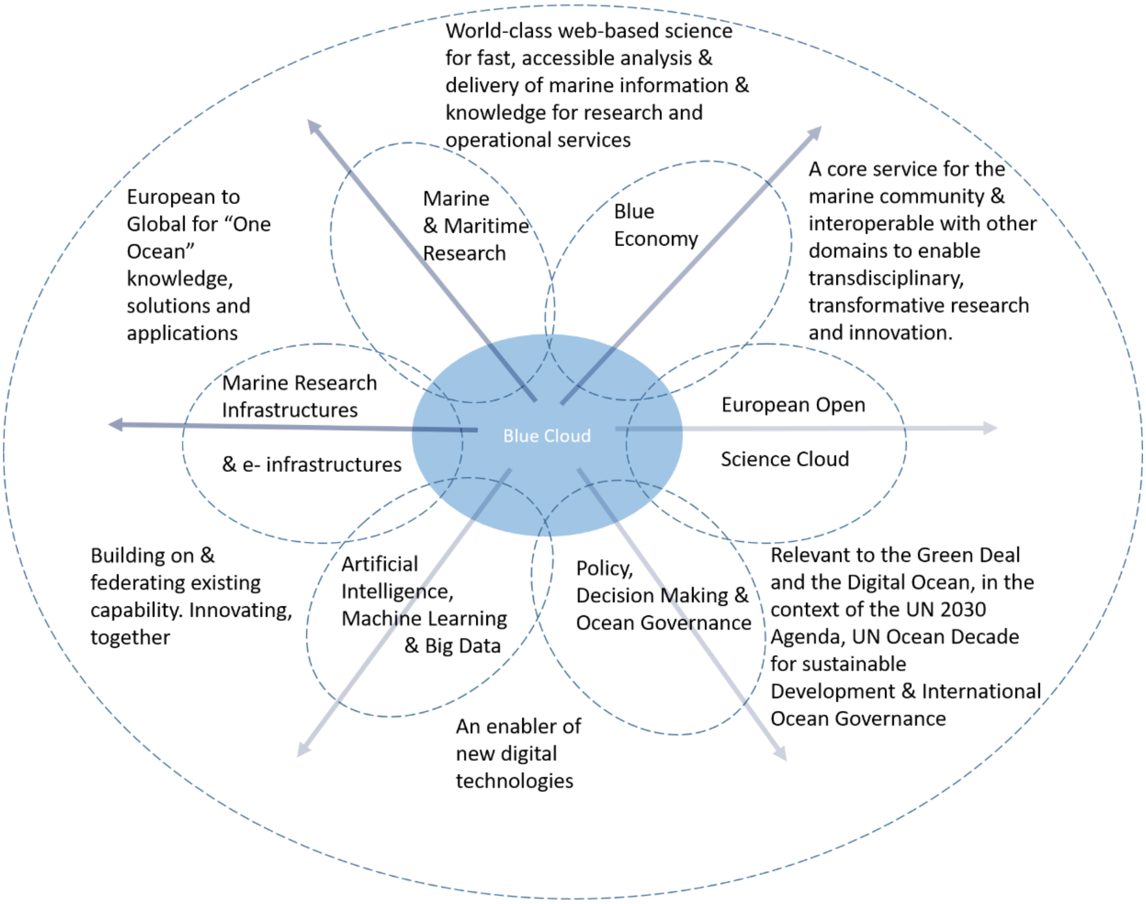


Figure 6: The Blue-Cloud Stakeholder Communities. Credit: Seascape Belgium

The Blue-Cloud seeks to deliver **added-value** to all these communities, aligning a **user-centric approach** towards the creation of such added value:

- **Target users:** Enhancing their capabilities to perform **Open Science** in the marine domain, leading to improved knowledge of the Ocean and of our collective ability to sustain its productivity and good health through science-based management, innovation and policies.
- **Strategic allies:** Providing a clear case and **practical frameworks for closer, long-term collaboration** along the marine knowledge value chain by enhancing their value propositions to shared user communities.
- **Enablers:** Boosting the **productivity** of **European infrastructures** and **assets** in the marine knowledge domain by building synergies around their value to users, in support of the EU Green Deal and of European stewardship in the global Ocean governance arena.

In Figure 6, the arrows show the direction of expansion of the Blue-Cloud community, from a core group of stakeholders in the centre, to engagement of wider stakeholder groups and sectors spanning the full spectrum of the marine and maritime domain, and wider sectors in new digital technologies. Whilst the focus is on Europe, Blue-Cloud is set in a global context and sets out to deliver a world-class capability for web-based Open Science.

3.4 Setting strategic goals & objectives 2022

The Blue-Cloud project is building on existing European assets and engaging with key stakeholders along the European marine knowledge value chain to demonstrate the potential of Open Science in the marine domain. B-C follows a three-fold strategy towards delivering clear added value to key stakeholders along Europe's marine knowledge value chain. To implement its strategy, it will work around accomplishing **five overarching, strategic goals** by 2022:

1. Build the case for Open Science in the marine domain.
2. Showcase approach towards federation of blue data infrastructures and e-infrastructures.
3. Promote wider user uptake of results of Open Science.
4. Develop a thriving environment that attracts engagement and participation of different communities of stakeholders.
5. Demonstrate value to EU Green Deal and UN Agenda 2030.

These overarching, **strategic goals** have been set to inform and prepare the ground for future developments that could benefit from capitalizing on its results, as explained in **Section 7**. The specific **objectives** that will be pursued by the Blue-Cloud project by **2022** are captured in the table below.

Strategic Goal	Objective		Target (2022)
Build the case for Open Science in the marine domain	1	Showcase benefits of Open Science to the marine community	+600 users subscribed to B-C VRE by 2022
Showcase approach towards federation of blue data infrastructures	2	Improve data interoperability through a standard approach towards metadata across participating blue data infrastructures, demonstrating value of pooling data into a common Virtual Research Environment (i.e. test "*bring data to analytical framework"* approach)	1 Shared (Meta) Data Discovery & Access Service for selected catalogues

and e-infrastructures	3	Showcase how computing resources could be brought to blue data infrastructures, driving users to infrastructures, in support of their "business" case (i.e. test "*bring analytical framework to data source*")	Technical recommendations included in Roadmap to 2030
Nurture and promote wider user uptake of results of Open Science	4	Support EOSC strategy towards FAIR by making resulting scientific outputs (i.e. virtual labs, workflows) of demonstrators findable, accessible, interoperable and reusable.	1 B-C Catalogue
	5	Build synergies with initiatives and projects in the marine domain to encourage uptake of the Blue-Cloud services by their communities.	30 synergies established with EU initiatives
	6	Showcase how the B-C Virtual Labs could run in other environments, making them FAIR.	1 User Test running virtual lab on WEkEO
Develop a thriving environment that attracts engagement and participation of key stakeholders	7	Attract engagement and participation of wider "research" audiences beyond early users engaged in the B-C demonstrators.	+600 users subscribed to B-C VRE by 2022
	8	Test the environment with users from other stakeholder communities.	60-100 users through 1 open, virtual hackathon
Demonstrate value to EU Green Deal and UN Agenda 2030	9	Communicate results and how the Blue-Cloud demonstrators contribute to policy objectives.	+10.000 visitors of B-C website
	10	Inform priority areas for development of new demonstrators to guide **Horizon Europe** and other financial frameworks seeking to further support Open Science in the marine domain.	1 B-C Roadmap to 2030

These efforts will leave a long-lasting print laying the foundations for their future evolution, as briefly introduced in the next section and fully considered in **Section 7**.

3.5 Evolving user-driven added value into the future

Early dialogue with the Blue-Cloud community (as described in Section 1.2) has been established to gather the opinion of **key stakeholders** regarding what the Blue-Cloud efforts should aim to achieve in the medium- and long-term future, providing initial insight into their **priorities** and **expectations**. This dialogue has confirmed that the Blue-Cloud's efforts and its results should be aligned towards delivering on a long-term vision of **supporting Open Science in the marine domain**, catering to the **marine research & science community** first, but seeking to **inform better science-based policies** and to **drive innovation across the blue economy** in the medium- to long-term, servicing a much wider community of users, from businesses to trade associations to national, regional and international governmental agencies, amongst others. From a **policy** perspective, when looking into the future, key stakeholders conveyed that the Blue-Cloud's efforts should prioritize delivering value to those policies that are ambitious towards solving **societal challenges**, namely the **EU Green Deal** and the **UN Agenda 2030** (as introduced in Section 2), rather than to horizontal policies with other, different objectives.

While the Blue-Cloud Open Science "use cases" or "demonstrators" will still need additional time to convey the full reach of their efforts, progress achieved to this date (described in Section 5) already anticipates their **demonstrative effect** in terms of **showcasing how Open Science can contribute to**

support specific policy objectives, but also in terms of **igniting interest** for **collaborative Open Science** and generating **other positive, indirect impacts** (like enhancing the visibility of existing marine data infrastructures and creating new opportunities for data ingestion, towards richer data coverage of our seas and Ocean – as evidenced by the image below, which captures an exchange between attendees to one of the several Blue-Cloud webinars which the project has held throughout 2020 and 2021 to introduce its demonstrators).

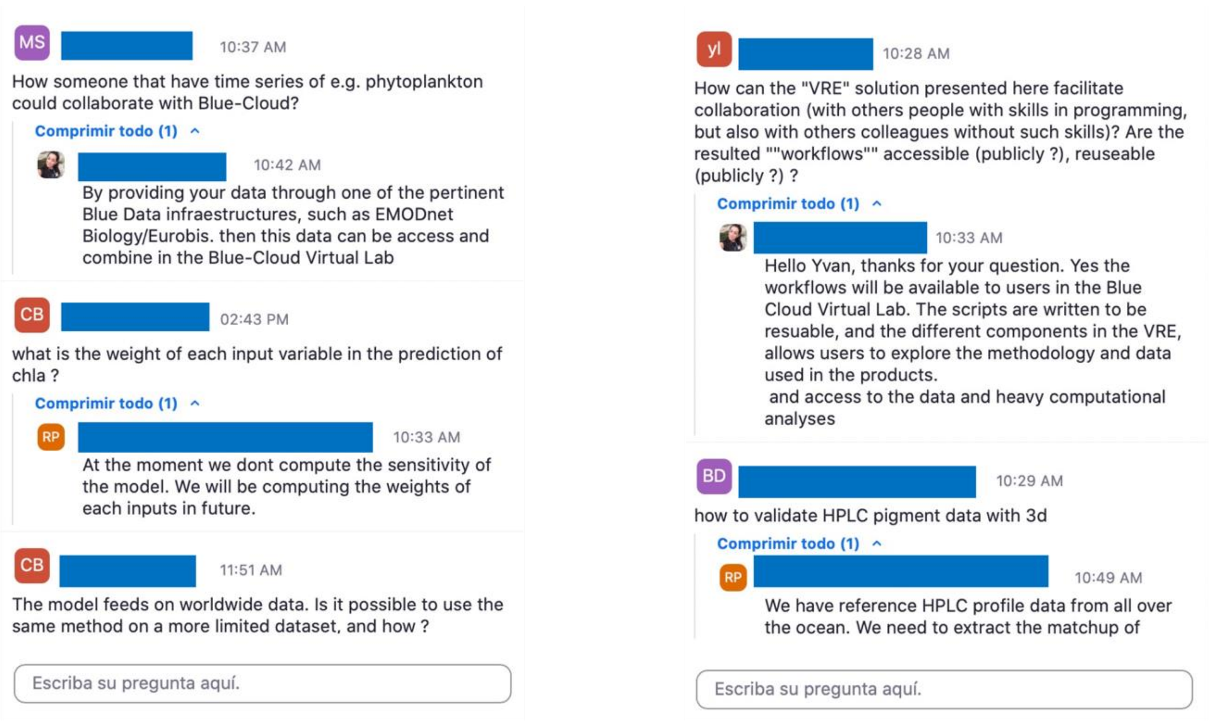


Figure 7: Participants of a B-C webinar showcasing one of its "demonstrators" interact during the webinar, expressing interest to sharing their data with other researchers in the B-C environment and to (re)use their "workflows" to support their research objectives

With this in mind, it seems logical to presume that if the **Blue-Cloud Vision 2022** has been built towards "**demonstrating the potential of Open Science in the marine domain**", the **Blue-Cloud Vision 2030** should be articulated around the idea of "**realizing the full potential of Open Science in the marine domain, in support of the EU Green Deal and the UN Agenda 2030**".

Blue-Cloud (Initial) Vision Statement 2030

"Realizing the full potential of Open Science in the marine domain, in support of the EU Green Deal and the UN Agenda 2030".

As described in **Section 7**, having an initial vision statement is useful and part of our methodological approach towards developing this Roadmap. However, defining a more refined, concrete **vision** and a **mission statement** (how the Blue-Cloud community should evolve its efforts to achieve that vision) and defining the **strategic goals** and **actions** (all of which are key components of a full-fledged "vision") that should guide its realization (the "roadmap") requires aligning action with a wealth of stakeholders who are currently knitting and building the underlying foundations for a thriving Open Science community in the marine domain, interacting with each other at different stages of the marine knowledge value chain.

In **Section 7**, we suggest an initial approach towards co-creating such a full-fledged **Vision 2030** for the capitalization and future development of the Blue-Cloud's efforts, including **policy** and **strategy recommendations** to 2030. But before looking ahead towards the future, let's focus on the present to gain additional insight into the solutions that are shaping **Blue-Cloud**'s **added value** contribution to the **marine knowledge value chain**, today.

4. Blue-Cloud's core services: Technology & innovation

The Blue-Cloud project is building on EU's **marine data landscape**: it is working with, and building on leading initiatives like **EMODnet** and **CMEMS**, well-known **research infrastructures** and **e-infrastructures** and on the drive towards a **European Open Science Cloud** to address the need and realize the potential benefits of bringing **closer integration**, **openness** and **enhanced analytical capabilities** into Europe's **marine knowledge value chain**. In this section we explore the **technology** underpinning the Blue-Cloud services and signpost how their medium term (2025) and long term (2030) evolution could bring wider potential and **innovation**.

Blue-Cloud's **underpinning technology** and **innovative concept** (see Figure 8) has been designed to deliver **core services** which are deployed through smart federation of leading marine data and e-infrastructures, as introduced in Section 3.

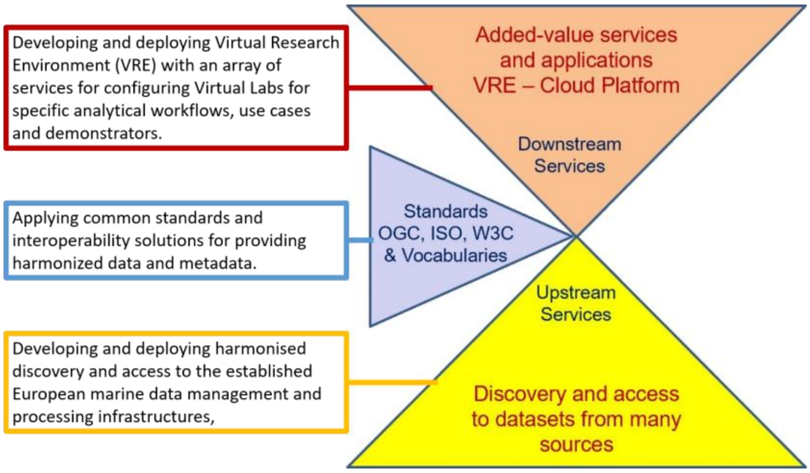


Figure 8: Leading concept and services for the Blue-Cloud development

Building on this concept, the Blue-Cloud project is developing a federated infrastructure that will feature an Open Science environment offering key services, which are considered **B-C's core services**:

- **Blue-Cloud Data Discovery & Access Service (B-C DD&AS)**: Federating and increasing interoperability between key European data management infrastructures (EMODnet, CMEMs, SeaDataNet, EurOBIS, Euro-Argo and Argo GDAC, ELIXIR-ENA, EuroBioImaging, EcoTaxa, and ICOS-Marine) to facilitate users in finding and retrieving multi-disciplinary datasets from multiple repositories.
- **Blue-Cloud Virtual Research Environment (VRE):** Providing a cloud-based, analytical and publishing framework as a federation of computing platforms and analytical services for constructing, hosting and operating **virtual labs** for specific applications.

4.1 Blue-Cloud Data Discovery & Access service

<u>**Initial delivery of B-C DD&AS**</u>

Blue-Cloud's **Data Discovery & Access service** delivers a one-stop shop towards leading European **blue data infrastructures** for exploring and retrieving their joint offer of data sets and data products.

An overarching **discovery interface** and **shopping mechanism** is being structured to allow the interfacing with each of the participating **blue data infrastructures**. This will allow users and machines to find and retrieve data sets from an impressive and highly diversified array of key infrastructures dealing with **physics**, **biology**, **biodiversity**, **chemistry**, and **bio genomics** data and data products. The **Discovery** functionality will be implemented with a two-step approach. The first step is a **user query** on a few **universal criteria** and at a collection level. For this purpose, use is made of the existing **GEODAB metadata brokerage solution**, which supports dynamic harvesting of a set of metadata from each blue data infrastructure. It allows transforming and ingesting these into a harmonised **Blue-Cloud metadata catalogue** at collection level, which makes it easier for users to identify which of the blue data infrastructures might have interesting data or data products and in how many collections. The second step is a **user query** which allows to **formulate additional search criteria** and to interact with each infrastructure, whereby the search profiles are driven by the characteristics of each of the blue data infrastructures and their specific contents and services. In this second case, there is no central caching, but direct use of **machine-to-machine services** to support **dynamic queries** per blue data infrastructure at data granule level, thereby identifying **individual data sets**. The user can include selected data sets from multiple blue data infrastructures in a **shopping basket** which is managed through Blue-Cloud's **data broker service**. This component uses the shopping list and centrally stored business knowledge about each blue data infrastructure to build a list of data requests which are then deployed to each of the relevant infrastructure to **retrieve** the associated data sets. These are temporarily stored in a **central cloud cache**, from which users can **download** them. For this purpose, Blue-Cloud shopping and brokerage mechanism features a **MyBlue-Cloud dashboard** on which each user can follow its shopping requests and their processing until providing links for efficient downloading of the full baskets. This process is coupled with an email communication informing the user of their **shopping requests** and their **processing status**. All data and data products are openly delivered, so the shopping process does not require further negotiation with data providers, but only depends on processing time by the Blue-Cloud broker and each of the blue data infrastructures. To support this process, Blue-Cloud service has to interact with web services and/or API's of each of the infrastructures, which in practice requires dealing with a range of **different service protocols** and with different **metadata** and **data formats**.

The core added value of **B-C Data Discovery & Access Service** to the user is overcoming not only differences in source metadata profiles across different blue data infrastructures, but also in machine-to-machine protocols for queries, aggregation levels of local data resources, delivery mechanisms, and local configurations. It makes it easier for users to **interact with multiple infrastructures** without first having to study and learn to operate with each of them. Moreover, it provides an innovative Blue Cloud shopping ledger for users and providers to keep track of their transactions, their shopping profiles, and allowing for monitoring and analysis of key indicators regarding users' **shopping behaviour**.

Medium-term (2025) potential and evolution of B-C DD&AS

The initial pilot version of **B-C Data Discovery & Access Service** is being deployed by interacting with existing web services and APIs from participating blue data infrastructures. However, this means that both the central services and also its users will still be confronted with differences in metadata formats and in supporting vocabularies. In the medium term, this could be further streamlined by upgrading and increasingly harmonising the underlying metadata catalogues. For example, in practice, each of the blue data infrastructures adopting and supporting more and multiple key elements of **ISO 19115 -19139 metadata model** (which is also in use in **INSPIRE**[31]).

In addition, **semantic interoperability** could be deployed, allowing different approaches. One approach could be encouraging blue data infrastructures to adopt **SeaDataNet**'s **Common Vocabularies,** which are leading in the marine data community in Europe and beyond and which currently consist of more than 110 different lists of topics, including variables, units, regions, platforms, instruments, organisations, instrument manufacturers, processing methods, biota, etc. and which are maintained as a crowd source with a clear governance structure, following requirements of user communities. Another option for semantic interoperability could be a **semantic mapping and brokerage at central level**, taking into account the different vocabularies in use by the blue data infrastructure, if any. This way, the metadata profiles and support by controlled vocabularies at each of the infrastructures might be upgraded. In this process, blue data infrastructures might also consider upgrading and/or changing types of protocols for their web services and API's, as currently there is a wide range of protocols being applied and several deployments are lacking functionalities, proper documentation and reliability, in particular in an operational setting.

In a comparable way, also a **further harmonisation of data formats** might be pushed forward, e.g., by adopting two data container models, such as an **ASCII model** and a **NetCDF model**, including **support for controlled vocabularies**, which would allow to formulate dedicated profiles for each of the types of data sets and products that are currently provided by the blue data infrastructures. This would allow to harmonise data models for possibly 90% of all data and products, while still allowing other data model types for specific data types, like e.g. SEG-Y for seismic, FASTQ for genomic data, and others, which are supported by often extensive and precious software packages in specific scientific domains.

These upgrading activities should fit perfectly in the strive of blue data infrastructures for improved **FAIRness** and for which several infrastructures are already underway, through projects like **ENVRI-FAIR** and other **EOSC** related projects. Such an upgrading of the underlying blue data infrastructures would contribute to optimising the functioning of the overarching query mechanism, possibly introducing a unified interface also at the second query level and facilitating delivery of data sets from several blue data infrastructures in data formats following the common data models and vocabularies, making the life and experience of users much easier.

In the medium term, the range of blue data infrastructures in B-C Data Discovery & Access Service could be **expanded** with **several other relevant marine repositories**. To this effect, a clear distinction should be made between well established and operational blue data infrastructures, which are European aggregators, and project repositories. The recommendation would be for the latter to preferably connect and become a provider node for the former, and in this way to become integrated into B-C Data Discovery & Access Service without extra effort on their side.

[31] https://inspire.ec.europa.eu/

The current B-C Data Discovery & Access Service has a focus on **delayed mode data sets** and **data products**. In the medium term, this might be expanded towards the **Near Real-Time and Real-Time metadata and data brokerage**, however in a close cooperation with **EMODnet Physics**.

In the transition from initial to medium-term, also the **Technology Readiness Level (TRL)** of B-C Data Discovery & Access Service and each of its underlying blue data infrastructures should increase from the **current average TRL7** (system prototype demonstration in operational environment) **towards TRL8** (system complete and qualified) or **better TRL9** (actual system proven in operational environment). This will require further attention for robustness of services and for an overall monitoring system, considering availability and performances of all services and service chains.

Long-term (2030) potential and evolution of B-C DD&AS

In the longer term, considering the earlier medium-term activities for **harmonising metadata** and **data formats** and **associated delivery mechanisms** of blue data infrastructures and the central **Blue-Cloud Data Discovery & Access Service**, a next step will be the **development of a central Data Lake**, which would serve as a **central data cache** for a major subset of all available data and data products. The selection should be made on **user requirements** of much needed data sets for various applications.

The Data Lake could be dynamically maintained through the **Blue-Cloud brokerage system**, interacting with each of the blue data infrastructures, and following their updating and increased population. In addition, several direct APIs with other major systems such as **WEkEO** can be foreseen. The function of the Data Lake would be to build and maintain a service ready **Big Data Collection** which could directly interact with various demanding applications in **B-C Virtual Research Environment**. These applications could be **statistical algorithms**, **data mining**, or **on-demand numerical models**. For that purpose, the Data Lake would not only store big amounts of data, but also transform those into other computer formats which are more suited for fast data throughput and interaction with numerical models and other data intensive applications. Use could be made of **Object Storage** technologies to ensure **high scalability**, **availability** and **durability**. The Data Lake could support cloud-optimised formats (COG/ZARR), including STAC for metadata. As the range of connected blue data infrastructures expands, probably more repositories will get connected to the European infrastructures, thus becoming available in Blue-Cloud Data Discovery & Access Service and Data Lake.

New Data Lake components should also strive to achieve **Technology Readiness Level – TRL9** (actual system proven in operational environment) and as part of this, the overall Blue-Cloud monitoring system should be expanded, considering availability and performances of all services and service chains. Considering the higher data throughput and expected large increases in number and weight of user transactions, considerable upscaling of computing and storage resources and associated bandwidths would be necessary.

4.2 Blue-Cloud Virtual Research Environment

Initial delivery of of B-C VRE

The initial **Blue-Cloud Virtual Research Environment (B-C VRE)** is based upon **D4Science**, an e-infrastructure allowing the Blue-Cloud community to make optimal use of major experience gained in earlier projects undertaken towards developing and operating a generic VRE with many core services, including building and running multiple **Virtual Laboratories** ("Virtual Labs"), each dedicated to specific research targets. From the start, D4Science has provided proven solutions for connecting to external computing platforms and means for orchestrating distributed services, which in practice proves to be instrumental for the further development of the initial Blue-Cloud VRE and its smart federation with the **EUDAT**, **DIAS**, **EGI**, and **EOSC e-infrastructures**.

B-C VRE facilitates collaborative research using a variety of **data sets** and **analytical tools**, complemented by generic services such as **sub-setting**, **pre-processing**, **harmonizing**, **publishing** and **visualization**. In the framework of the Blue-Cloud project, a number of demonstrators are being developed (as previously explained and further described in Section 5). Each demonstrator is enacting a family of **analytical workflows** (or pipelines) that consist of a series of applications and make use of selected datasets as input. Results of analytical workflows can be documented with **provenance information** for **reproducibility**, provided with **DOI**'s for **citation**, **published**, and **visualized**. For each demonstrator, a **Virtual Lab** has been deployed within B-C VRE, with their researchers in group accounts. Multi-disciplinary datasets can be retrieved by means of **B-C Data Discovery & Access Service** and loaded into the **VRE Data Pool**. Also, direct APIs are made possible, such as the one used for DIAS WEkEO service, facilitating regular and automatic provision of new data sets for specific Virtual Labs, following pre-set requirements. Moreover, users can introduce their own data sets, and ingest data sets retrieved and ingested from other major data portals and resources, like **NOAA World Ocean Database**[32], **FAO** databases, and many others.

Analytical services for stocking the overall VRE toolbox and equipping Virtual Labs and their analytical workflows are already available in B-C VRE, are partly added by the developers of the demonstrators, and are running on associate computing platforms, which are federated to become part of B-C VRE. For instance, the **WEkEO DIAS infrastructure** provides access to the **Sentinel** satellite images and also has several tools to analyse and process these, on the same computing platform, close to the images. Results of the processing can be ported into B-C Virtual Labs.

B-C VRE has a **common dashboard** for accessing Virtual Labs for performing collaborative research. The dashboard includes **common facilities** for accessing a **shared workspace**, a **social networking** area, a **data analytics platform** and a **publishing platform**. These facilities are all correlated each other and realize a system where (i) datasets can seamlessly flow across the various components to be easily shared among users and openly discussed by social networking practices; (ii) generated datasets and products are automatically enriched and enhanced with metadata capturing their entire lifecycle, their versions, and the detailed list of authors and tasks performed leading to the development status and shapes. The dashboard allows also to configure new workflows as combinations of services and for future expanding of B-C VRE with more connected computing platforms and more analytical and generic services. One important functionality is providing support for a federation of identities between different AAI systems that the engaged computing platforms might operate. Interoperability is arranged with social media login (OpenID, OAuth 2.0, and SAML), the authoritative eduGAIN network, and others.

---

[32] https://www.ncei.noaa.gov/products/world-ocean-database

OGC standards[33] for interacting with the services and the datasets and resulting data products are used as much as possible.

All the provided facilities enabled through B-C VRE are exploitable as web applications accessible either through the Blue-Cloud VRE **gateway menu** for common users or through the **APIs** for developers. The data analytics platform offers several types of facilities to users of B-C VRE:

- **Jupyter notebooks for expert users:** This allows users to execute code fragments leading to the generation of the results. The notebooks are very suitable as they combine code fragments, text cells, and figures or animations that illustrate different steps in the process. Also, notebooks can be exported to HTML and PDF and are easy to share. Sharing notebooks along with the generated products facilitates improving reproducibility and peer-reviews. Jupyter notebooks are designed for a single user, while a multiple-user instance has been established using Jupyterhub.
- **RStudio**: Offered to deliver a console, syntax-highlighting editor that supports direct code execution, as well as tools for plotting, history, debugging and workspace management;
- **A software importer tool:** For analytical methods implemented in **Java**, **Python**, **R**, **C** and many other languages and a high throughput computing engine are also available for expert users wishing to execute their software code in a distributed infrastructure. Any imported software become accessible via a standard Web Processing Interface (WPS) and any execution generates automatically provenance information making the results sharable, re-usable, and reproducible.

The **initial B-C VRE** provides an **innovative computing platform** capable to hide the complexity of the infrastructure while enabling easy integration and deployment of analytical algorithms and software tools. Moreover, additional services might be deployed as Docker containers which is an easy way for adding new services without configuration efforts. It also delivers an innovative approach for the publication of the analysis results that respects ownership, provenance, and controlled access while supporting collaborative research from the formulation of a new analytical approach to its experimentation, validation and delivery.

In terms of **Technology Readiness Levels (TRLs)**, the initial **B-C VRE common facilities** can be qualified as **TRL8** (system complete and qualified) with the ambition to achieve at least **TRL7** level for all the **Virtual Labs** operated by the B-C VRE within the lifetime of the Blue-Cloud project. A selection of the analytical methods, relevant for the B-C Virtual Labs, are available as prototypes at **TRL6** (technology demonstrated in relevant environment), while other integrated tools are already at **TRL7** (system prototype demonstration in operational environment). However, bringing the integration of all components to TRL7 will still require additional effort.

<u>Medium-term (2025) potential and evolution of B-C VRE</u>

With its generic functionality and its wide connectivity options, **B-C VRE** offers researchers major opportunities to undertake **world-class science**. The Virtual Labs come to demonstrate this and serve as a way to promote and market B-C VRE, reaching out and attracting the interest of many marine scientists and researchers wishing to join Blue-Cloud to access its services. In the **medium term**, this

[33] https://www.ogc.org/docs/is

should result in an expanding number of users and uses. The modular architecture of Blue-Cloud VRE is scalable and sustainable, being fit for **connecting additional infrastructures**, implementing more and advanced **blue analytical services**, configuring more dedicated **Virtual Labs**, and targeting broader (groups of) **users**. This architecture will be expanded to integrate cluster of GPUs to empower the cloud computing platform with native capacities to handle **analytical methods** based on **Artificial Intelligence**. Those resources will become accessible upon request, within the limit assigned to each Virtual Lab. New ways will be explored for adopting additional **cloud storage**, **cloud computing**, **deep learning** and **neural networks** for supporting **big data processes** for **validation**, **extraction**, **interpolation** and **products generation**. Moreover, there will be challenges to serve existing and potential new users with improved and new functionalities for analysing and processing data sets as part of research and for generating data products and knowledge. Data scientists will be empowered to dig into large data sets with new classes of algorithms such as **Artificial Intelligence**, **Machine learning**, and **Data Mining**. This will be implemented and validated through the Blue-Cloud demonstrators. The overall **Technology Readiness Level** of **B-C VRE** should be increased from **TRL7** (system prototype demonstration in operational environment) to **TRL8** (system complete and qualified), which requires all components to be upgraded and made more operationally robust.

Long-term (2030) potential and evolution of B-C VRE

On the longer term, the so far "**data centric approach**" of B-C VRE could be complemented with **mathematical modelling** of **processes**. This could be partly realized in cooperation with **CMEMS**, which is developing and running **large scale ocean mathematical models**. This could be complemented with a set of **on-demand models**, allowing to set up and run **what-if scenarios** for the ocean and marine environment. This way, B-C VRE could develop into a **Digital Twin of the Ocean**, capable of mirroring real-life processes, measures and impacts from within its virtual environment.

Progressing in this direction will require further efforts, including developing and expanding **visualization techniques**, being able to run **what-if scenarios** as interactive games in 4-D (gamification of decision-making processes). This could be very intuitive and useful for decision processes, such as Marine Spatial Planning, developing climate scenarios or for gaining new scientific insights. Techniques from the gaming and animation industry should be studied and brought in for this purpose, also demanding **High Performance Computing** facilities and availability of **GPUs** becoming in same order of magnitude (tens of thousand) of the **classic CPUs**. To lower the threshold for using new models and visualization techniques by researchers, dashboards should be designed and implemented into B-C VRE to **ease interfacing**. Moreover, **visualization techniques** are expected to become part of the publishing of scientific results, in order to bring science closer to the **general public** in a more effective way. At the same time, use of **Artificial intelligence** and **Data Mining** will be amplified, making use of progressing state-of-the-art and increasing computing resources and capabilities.

B-C VRE could potentially attract many thousands of researchers that will develop and deploy hundreds of **Virtual Labs**, making use of and interacting with models of major processes. Also, B-C VRE could become a platform that is used by many EU Research Projects for **specific workflows** and **products**. To this end, the overall **Technology Readiness Level** of the core Blue-Cloud VRE should be increased from **TRL8** (system complete and qualified) to **TRL9** (actual system proven in operational environment), which requires further upgrading of all components and operations.

5. Leveraging on innovation to deliver societal impact

The Blue-Cloud's "**demonstrators**" are case studies showcasing how scientists and researchers are using the technology and services offered by **Blue-Cloud**'s **web based, Open Science environment** described in the previous section. In this section we explore how they are using **B-C core services** to support their research objectives, inviting collaboration and contributions from other users to upscale and capitalize on their results, increasing the overall productivity of science, in support of the **EU Green Deal** and the **UN Agenda 2030**. We look deeper into how their efforts can contribute to **specific policy objectives** and what we have learned so far through their use of the "pilot" Blue-Cloud.

5.1 Blue-Cloud's demonstrators: Showcasing uses in the marine domain

The Blue-Cloud demonstrators are practical examples showcasing how **Open Science** can enhance the **marine knowledge value chain** in support of greater societal objectives. They all have one thing in common: they demonstrate how providing researchers with **web-based**, **analytical tools** and **greater cloud computing power** -allowing them to perform heavy computational analyses otherwise not possible or too timely to perform on local computers-, underpinned by **open access to data** available across different European blue data infrastructures, can lead to a more effective and efficient development of **innovative data products** that contribute to **improve knowledge** and **understanding** of the Ocean, and thus the overall performance of Europe's marine knowledge ecosystem. Below we introduce their research objectives and how they will leverage on the Blue-Cloud to deliver societal impact. They all have publicly shared abundant information on their work through different webinars, which are available at the Blue-Cloud website, for those seeking additional technical insight[34].

<u>**Demonstrator 1: "Zoo- and Phytoplankton Essential Ocean Variables (EOVs)"**</u>

The **Zoo- and Phytoplankton Essential Ocean Variables (EOVs)** demonstrator is working on the development of **innovative data products** that help to estimate the **abundance** and **concentration** of **plankton** (which is an Essential Ocean Variable - EOV). Using the Blue-Cloud Open Science environment and feeding from different data sources, it has developed an **online service** (**Virtual Lab**) that provides users with built-in, **open, transparent methodologies** to estimate **plankton abundance**, **distribution** and **dynamics**, based on **big data analysis** and **machine learning** (e.g. neural networks).

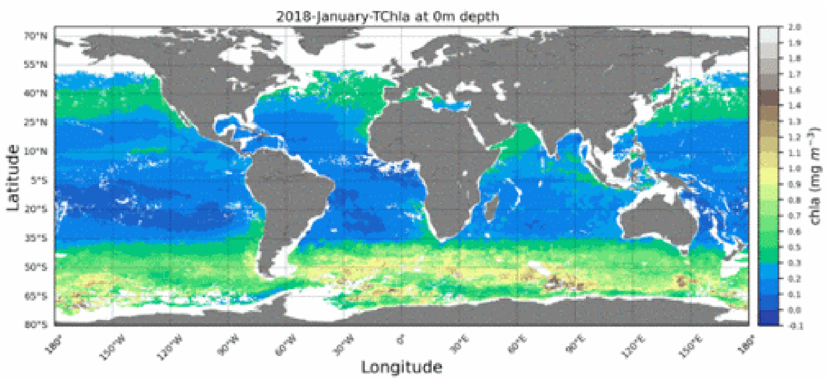


Figure 9: Main output of the model to quantify the relative contributions of **phytoplankton abundances** in one station. The model allows the integration of different EOV variables, not only to display data-driven trends, but also to understand interactions in a mechanistic way.

[34] https://www.blue-cloud.org/blue-cloud-webinar

Shortly, any **user** will be able to join this Virtual Lab and take up these available **EOV methodologies** to easily re-run analysis with different datasets, as well as sharing and comparing results. Users of the **Zoo-phytoplankton Essential Ocean Variables (EOV)** demonstrator will be able to capitalize on the offered reusable methodologies and tools to:

- Generate **maps of zooplankton abundances**.
- Produce **global ocean 3D key phytoplankton products of chlorophyll-a (Chla) concentration**, as a proxy for total **phytoplankton biomass** and phytoplankton functional types, as a proxy for **phytoplankton diversity**.
- Quantify the relative contributions of the top-down and bottom-up drivers in **phytoplankton dynamics**, applying a mechanistic model using near real-time data.

The resulting **EOV products** contribute to **improve knowledge** and vastly **reduce uncertainty** on the present state of **marine plankton ecosystems** and their response to ongoing and future **climate change**, providing valuable information for the **modelling**, **assessment** and **management** of marine ecosystems. **Plankton** is considered the **cornerstone** of marine ecosystems and of the food supply throughout the Ocean, as well as being essential in the global carbon and oxygen cycles producing an estimated 80% of the Planet's oxygen. This resulting knowledge can contribute to inform the **EU Biodiversity Strategy for 2030** and **"Farm To Fork" Strategy** in support of the **EU Green Deal**, as well as the **UN Agenda 2030**'s **Sustainable Development Goals 2** (Zero Hunger) and **14** (Life Under Water).

<table>
<tr><td colspan="2">Virtual Lab: Zoo- and Phytoplankton Essential Ocean Variables (EOVs)
Lead Partner: VLIZ · Scientific Partners: VLIZ, Sorbonne University, CNRS-LOV, University of Liege</td></tr>
<tr><td>Potential users and applications
• Fisheries advisory organisations: To study the availability of food resources for fish stocks and assess the effects on fish stocks.
• Marine policy officers: To address threats such as food insecurity, foreseen under the EU Biodiversity Strategy for 2030.
• Researchers (e.g. from environmental agencies): To contribute to the understanding of environmental conditions and factors at new scales of observations.</td><td>Policy Impact
Knowledge generated through this Open Science ecosystem can help marine policy officers to address threats related to food security, as foreseen under the EU Biodiversity Strategy for 2030 and the "Farm to Fork" Strategy underpinning the EU Green Deal, but also the UN SDGs 2 (Zero Hunger) and 14 (Life Under Water). Moreover, the proposed EOV products are of interest for fundamental research, contributing to modelling of marine ecosystems. These models will be available to feed the future EU pilot Digital Twin of the Ocean, which will unlock new, innovative applications of these Open Science outputs.</td></tr>
<tr><td colspan="2">Users of this Virtual Lab will be exploiting and benefiting from open access to the following data sources:
• Argo GDAC: Salinity, oxygen, chlorophyll data
• CMEMS: Satellite-derived reflectance, sea level anomaly, PAR, physical data
• EurOBIS: Zooplankton abundances
• EMODnet Biology: Zooplankton (from MBA CPR survey) & phytoplankton abundance
• GEBCO: Bathymetry
• GlobColour: Satellite derived Photosynthetically Available Radiation
• GSFC, NASA: Distance to nearest coastline
• LifeWatch: Nutrients, PAR and temperature
• SeaDataCloud: Temperature and salinity
• World Ocean Atlas 2018: Nitrate, Silicate and Phosphate</td></tr>
</table>

Demonstrator 2: "Plankton Genomics"

Using the Blue-Cloud Open Science environment -underpinned by **access** to **biomolecular, imaging** and **environmental data** available across different European blue data infrastructures and cloud computing- the "**Plankton Genomics"** demonstrator has developed an **online service** (**Virtual Lab**) that provides users with **analytical methodologies** and **tools** ("**workflows**") to extract new **"plankton" data products** from existing data, in particular from genomic samples collected by the "Tara Oceans" project[35]. These analytical tools are relevant to study **plankton diversity** at basin and global scales when observations are scarce, which is of particular interest for researchers in the fields of **plankton biogeography**, **marine biogeochemistry**, and **ecosystem health**.

Users of the **Plankton Genomics** demonstrator will be able to capitalize this **analytical framework** to:

- Generate georeferenced tables of **occurrence of known** and **unknown plankton taxa & genes**, together with environmental variables, allowing **species** and **functions discovery** from the **genomic dark matter** of Tara Oceans data.
- Produce **global maps** of potential **distribution** of the selected **plankton taxa** and **genes**, through machine-learning based regression on environmental variables.

Scientific researchers, in particular taxonomists, computational ecologists and bioinformaticians in quest of the **identification** of **unknown sequences** in the oceanic environment will be able to 1) **obtain lists** of unknown taxonomic and functional units, 2) **correlate** these unknowns with **environmental variables**, 3) **project** the **results** of this correlation over the **world's ocean** and 4) **visualize** these bio-geographies as **maps**.

Plankton Genomics **Lead Partner:** EMBL · **Scientific Partners:** Sorbonne Université, CNRS, VLIZ	
Potential users and applications	**Policy Impact**
• **Scientific researchers** (in particular taxonomists, computational ecologists and bioinformaticians): To pursue the identification of **unknown taxa** and **genes** in the ocean, predict other locations where the sequences are likely to occur, and overall, to advance research on **plankton biogeography**, **marine biogeochemistry** and **ecosystem health**.	Knowledge generated through this **Open Science ecosystem** can contribute to advance scientific progress in the fields of **ecosystem health**, informing progress towards the **EU Biodiversity Strategy 2030**. It can also contribute to advance **sequencing** of the **DNA** of our ocean and waters by learning more about those sequences, which is one of the objectives of **Mission Starfish** and offers unprecedented opportunities to discover and use **new molecules** and **biotechnologies**.
Users of this **Virtual Lab** will be exploiting and benefiting from open access to the following data sources: • **ELIXIR-ENA:** Genomics data, in particular that of Tara Oceans and Arctic (global ocean, plankton) • **CMEMS:** Ocean colour, altimetry, temperature and salinity field data, ocean and climate variables • **EcoTaxa:** Quantitative plankton images • **Tentatively:** Data from other (external) sources, such as NOAA's World Ocean Atlas (climatologies)	

The resulting new data products can contribute to generate **new hypotheses** and **guide research** in the fields of **plankton biogeography**, **marine biogeochemistry** and **ecosystem health**, leading to new, deeper assessments of **plankton distributions**, **dynamics** and **fine-grained diversity** (to a molecular

[35] https://oceans.taraexpeditions.org/en/m/about-tara/les-expeditions/tara-oceans/

resolution). The fine scale genomic data used gives unprecedented ability to **detect species** and unknown, but coherent, **genetic entities**, which is particularly relevant in scenarios where collection of new samples is difficult or impossible. This resulting knowledge can contribute to inform the **EU Biodiversity Strategy for 2030** in support of the **EU Green Deal**, as well as the **UN Agenda 2030**'s **Sustainable Development Goal 14** (Life Under Water). More importantly, sequencing and learning about the **DNA** of our ocean and waters, as a **common public good**, is one of the objectives of **Mission Starfish** and will offer immense opportunities to discover and use new molecules and biotechnologies for the **wellbeing** and **health** of **European citizens**.

Demonstrator 3: "Marine Environmental Indicators"

The "**Marine Environmental Indicator"** demonstrator is showcasing how using **analytical tools** and greater computing power underpinned by access to data available across different European blue data infrastructures can ease the way towards developing dynamic **Marine Environmental Indicators** (MEI) that inform the **environmental quality** of the **Ocean**. As a scoping effort focusing on the **Mediterranean Sea**, the demonstrator is using the Blue-Cloud Open Science environment to set up an online service (**Virtual Lab**) allowing users to:

- Calculate and share **marine environmental indicators** feeding from data sets available across different sources.
- Obtain **new, added-value data** applying **big data analysis** and **machine learning methods** on existing data sets.
- Perform online and "on the fly" operations, such as **selecting** a portion of data for a specific area and period of time, **performing analytics** with several methodologies on the selected variables or **displaying** the available indicators on **tables**, **maps** and other **graphic visualizations**.

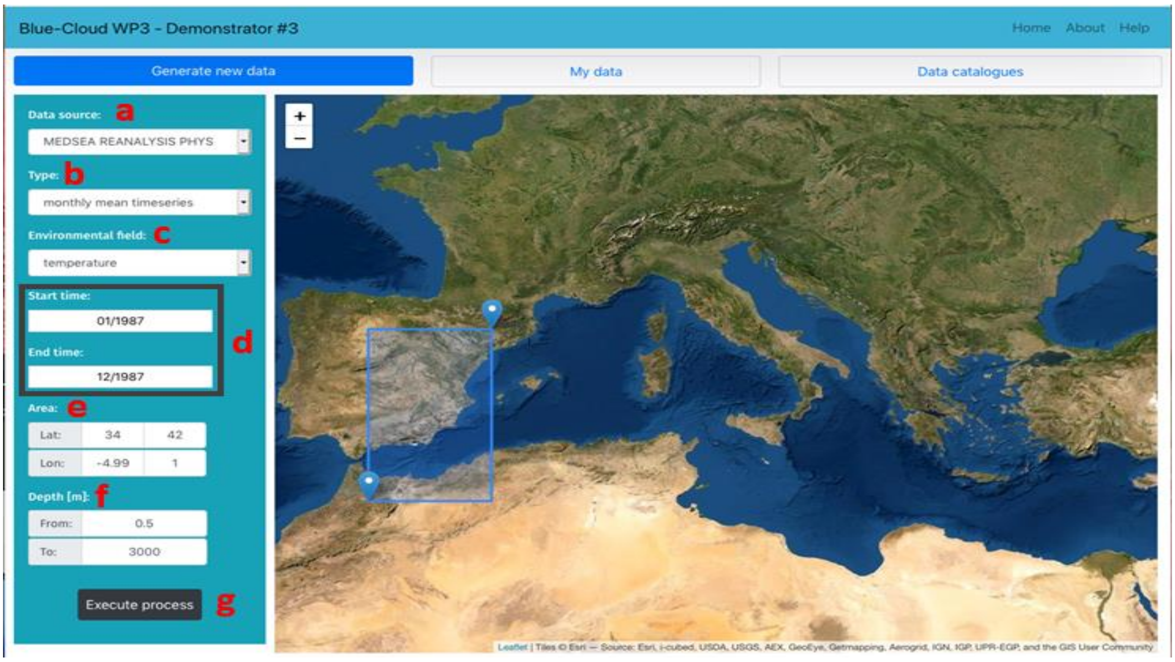


Figure 10: Demonstrator offers a user-friendly interface for execution of WPS methods allowing section of: Output type [b] [c] Lon/lat area [e] & depth layer [f]

To enable these capabilities, the Virtual Lab offers:

- **Prototype Marine Environmental Indicators (MEI) applications**, which allows users to display available data to generate new added- value data, and to apply methods to deliver new **marine environmental indicators**.
- **Prototype of methods**, adopting a big data approach to extracting information from data sources and generating new marine environmental indicators.
- **Data processing workflows** delivered in notebooks for specific research on Ocean patterns and Ocean regimes, delivering **marine physical indicators** based on machine learning. Unsupervised classification algorithms are applied on vertical profiles ("patterns") or time series ("regimes").
- A **Storm Severity Index (SSI)** providing insight into changes of the storm climate of a sea region and into the impact of such changes on sea circulation patterns, delivering relevant **marine meteorological indicators**.
- A prototype for recalculating missing sub-variables of the "Essential Ocean Variable" **Inorganic Carbon** and (potentially) providing calculation of uncertainties, to produce fit-for-purpose data products catering different user segments.

Marine Environmental Indicators **Lead Partner:** CMCC · **Scientific Partners:** IFREMER, MOi, KNMI, University of Bergen	
Potential users and applications	**Policy Impact**
• **Governmental & Environmental Protection Agencies:** To perform statistical analyses of the quality and characteristics of the marine environment (Mediterranean Sea) towards MSFD compliance. Potential to scale up in the next version to the global Ocean. • **IOC UNESCO:** To measure progress towards SDG 14.3 (Ocean acidification). • **Researchers:** To facilitate the discovery of new climatic indicators based on machine learning, and a simplified way to analyse oceanographic data.	Knowledge generated through this **Open Science ecosystem** can contribute to measure progress towards the objectives set forth by the **EU MSFD**, as well as to inform progress towards **SDG target 14.3** on **Ocean acidification**. In the future, it could be instrumental to deliver a more dynamic, timely understanding of the impact of climate change mitigation measures, potentially contributing towards more effective **carbon pricing** throughout the economy. With more adequate interfaces, it could evolve as an effective **educational** and **informational** tool to the **general public**.
Users of this **Virtual Lab** will be exploiting and benefiting from open access to the following data sources: • **CMEMS & C3S (DIAS-WEkEO):** Ocean and climate model data (subset of MEDSEA_REANALYSIS_PHYS_006_004 and GLOBAL_REANALYSIS_PHY_001_030 CMEMS, ERA5 data) and observations • **ICOS-Marine and ENVRI-FAIR:** Inorganic carbon data. Throughout the project, direct access to the following data sources might also be contemplated: • **EMODnet:** Physics, biology and chemistry data. • **Euro-Argo and Argo GDAC:** Salinity, oxygen, chlorophyll data. • **SeaDataNet:** Biogeochemistry, physics, biology, environmental data	

There are different ways in which this demonstrator can contribute to **enhance marine knowledge**, in support of the EU Green Deal and the UN Sustainable Development Goals (SDGs):

- Sustaining **Good Environmental Status** of EU marine waters by progressing towards the targets set out in the **European Marine Strategy Framework Directive** (MSFD) will remain central to EU action on Ocean matters. As anticipated in Section 2, the **MEI Virtual Lab** can provide practical ways for Open Science to ease the way towards **monitoring progress** towards policy objectives, providing a more accurate picture of the status of the Ocean, including the visualization of long-term series to better assess **environmental trends**. Other specific applications might include, for example, better insight into the impact of toxic algae blooms in the environmental quality and productivity of marine ecosystems.

- As also discussed in Section 2, the EC has also set out a clear ambition to achieve **climate neutrality** by **2050**. To deliver on this objective, the EC is intending to review all relevant climate-related policy instruments, comprising the Emissions Trading System; Member State targets to reduce emissions in sectors outside the Emissions Trading System; and the regulation on land use, land use change and forestry. The role of oceans in mitigating and adapting to climate change is increasingly recognised, as lasting solutions to climate change will require greater attention to **nature-based solutions**, including **healthy** and **resilient seas** and **oceans**. Although marine ecosystems are yet hardly, formally considered within the scope of climate-related policy instruments at international nor European level, their role in carbon sequestration is increasingly recognized[36]. Whereas their formal recognition in governance mechanisms might remain a long-term aspiration, MEIs can already start delivering a more dynamic, timely understanding of the **impact** of **climate change mitigation measures**, which currently take years to assess. In the future, the information derived could potentially contribute towards more effective **carbon pricing** throughout the economy, which is one of the key objectives of upcoming climate change related European policy reforms. Although the mechanisms to bring this information into governance mechanisms and pricing systems need to be further matured, the value of MEI's for climate change policy is already starting to be explored and progress can be catalysed through the **MEIs Virtual Lab**, which could guide and provide feedback to EU national agencies and ministries on their national commitments to reduce GHG emissions. The virtual lab could also inform **UNESCO**'s **Intergovernmental Oceanographic Commission (IOC-UNESCO)**, as the custodian agency for the **SDG target 14.3** on **Ocean acidification**.
- As it is the case with demonstrators 4 and 5, in the long term, this **Virtual Lab** has the added value potential of opening a window of opportunity to be used as a **powerful educational** and **informational tool** to the **general public**, for a better-informed civil society, but also so that citizens become more aware of the connections between our use of the Ocean and its impact on marine ecosystems and in turn, on our own health. Although the Virtual Lab would have to develop new interfaces to be fit for this purpose, this fact brings yet another potential, future contribution of **Open Science** to the marine knowledge value chain a step closer to its realization.

<u>Demonstrator 4 "Fish, A Matter of Scales"</u>

The "**Fish, A Matter of Scales**" demonstrator is working on **innovative ways** of **delivering knowledge on fisheries** to **improve fisheries management** and **monitor trends** in **fisheries indicators**. It aims to implement a fully FAIR data management approach across findable and accessible fisheries data and,

[36] Climate action requires new accounting guidance and governance frameworks to manage carbon in shelf seas (T. Luisetti et al, 2020) https://www.nature.com/articles/s41467-020-18242-w

in combination with interoperable (meta)data across different blue data infrastructures, deliver reproducible **data analytics** through a variety of products.

Using the Blue-Cloud Open Science environment, and building on previous efforts, it has set up two different **Virtual Labs**. The first one allows users to search for reliable and fact-checked fisheries content, with features ranging from **global fisheries maps**, **catch statistics** and **overviews**, high resolution bathymetry and **aggregate maps** of **major world fisheries**. The second one provides users with a suite of **analytical services**, including indicators, interactive maps, models and methods (currently under development) based on widely published software to serve a community of **fisheries data analysts**, as well as expanded information from other sources to bring context and perspective to the fisheries maps and for **approved status assessments of fisheries**:

- **Global Tuna Fisheries Atlas Virtual Lab:** Updated for 2019, this **virtual lab** provides a complete overview of **FAO statistics** (global tuna and billfish fisheries), showcasing a **workflow** that facilitates the visualization of these statistics in combination with other maps (e.g. **CMEMS** and **EMODNet** layers) in unprecedented detail and precision. Through a **user interface**, it offers global thematic maps showing the **Earth's fisheries**, their **production** and **trade**, as well as **fish distributions maps** and **ecological zones**; added maps created through the full benefit of CMEMS satellite products based on Copernicus data; and measurements and other accurate analytical techniques to place and analyse fisheries in a wider environmental context.
- **Global Record of Stocks & Fisheries Virtual Lab:** This **virtual lab** provides a reference repository of fish stocks and fisheries combining harmonized data from three global data providers, offering different working environments to **validate** and **harmonize new data sources** (not public) for **approved status assessments of fisheries** and to publish data (making it public). It uses **semantic technology** for data harmonization.

<table>
<tr><td colspan="2">Fish A Matter of Scales
Lead Partner: FAO · Scientific Partners: FAO, FORTH, IRD</td></tr>
<tr><td>Potential users and applications</td><td>Policy Impact</td></tr>
<tr><td>• Regional fisheries data analysts: To help them assess how fisheries develop over time in their area of interest and forecast future evolution under different scenarios, informing fisheries managers for decision making processes. To train data analysts in the use of algorithms in support of SDG 14.4.1 analysis.
• Developers: To access a “boiler plate” solution for the management of fisheries time-series on catch and effort that brings collated statistical data into a data harmonization and quantitative analysis process.
• Other communities: To benefit from a FAIR compliant, data management solution spanning statistical and geospatial data workflows in the fisheries, aquaculture and related aquatic and land-based domains.</td><td>Knowledge generated through this Open Science ecosystem can inform and support decision making across different EU policies, including the Common Fisheries Policy, the EU “Farm-To-Fork” Strategy and the EU Biodiversity Strategy by enhancing policy monitoring capabilities. It can also inform fisheries management decisions at a regional level and support capacity building to monitor progress towards the UN Sustainable Development Goals (SDGs), more specifically SDG 2 (Zero Hunger), SDG13 (Climate Action) and SDG 14 (Life Under Water).
In the long-term, it could evolve as an educational and informative tool targeting the general public, opening an opportunity for citizens to gain insight into fish stocks and fisheries and the impact of their consumption choices, with easy access through web or QR codes.</td></tr>
<tr><td colspan="2">Users of this Virtual Lab will be exploiting and benefiting from open access to the following data sources:</td></tr>
</table>

- **FAO:** Global stocks and fisheries data; FAO global project data and fish tagging data
- Global Effort Maps of Fisheries
- **FNS Cloud:** Data on fish composition
- Any ISO/OGC dataset that is FAIR compliant, including CMEMS and EMODnet

From a technical perspective, this demonstrator shows how **FAIR data management** practices allows to **harmonize** and **combine** data, which can be made available for generic analytical services. From a policy perspective, this demonstrator shows how the availability of **user-friendly analytical frameworks** can support the marine knowledge value chain, servicing **intermediate users** (such as fisheries data analysts) but also informing **policy makers**, with a potential to evolve towards delivering useful resources that bring the issues at stake closer to the **general public**. Again, its contribution to delivering on the objectives of the EU Green Deal is worth analysing not only from a short-term perspective, but also with a view to the potential evolution of the services and added value delivered by this demonstrator:

- With the launch of the **EU Biodiversity Strategy**, all EU policies will be expected to contribute to preserving and restoring Europe's **natural capital**. The **EU "Farm to Fork" Strategy** will continue to work under the common fisheries policy to reduce the **adverse impacts** that fishing can have on **ecosystems**, especially in sensitive areas. **Combining state-of-the-art data** on **fish stocks** and **fisheries** with other **marine environmental indicators** is essential to understand the relations between fisheries and marine ecosystems, informing policy making processes.
- The **"Farm to Fork" Strategy** will also strive to stimulate sustainable food consumption and promote affordable healthy food for all, making European food the global standard for sustainability. The EC will explore new ways to give consumers **better information**, including by digital means, on details such as where the food comes from, its nutritional value, and its **environmental footprint**. **Fish provenance** and **traceability** are relevant information that can contribute towards better assessing the environmental footprint of fisheries, as well as informing more sustainable consumption choices. In the short and medium term, this virtual lab can contribute to support **fisheries monitoring** to inform more sustainable **fisheries management** decisions. In the future, this virtual lab further opens a window of opportunity to evolve to be used as a powerful **educational and informational tool** to the general public, so that consumers become more aware and are able to connect the **impact** of their **consuming behaviour** on available fish stocks.
- In an international context, the "**Fish A Matter of Scales**" demonstrator provides **data analysts** from **Regional Fisheries Management Organizations** (RFMOs) with fisheries analytical models and tools to monitor how fisheries in their area of interest develop over time, also in relation to environmental variables and other ancillary data (for example, correlating fish stocks with environmental data to assess the impact of climate change on fish stocks behaviour). This will contribute to **inform** fisheries management decisions, but also support **capacity building** to monitor progress towards the **UN Sustainable Development Goals** (SDGs), and more specifically **SDG 2** (Zero Hunger), **SDG 13** (Climate Action) and **SDG 14** (Life Under Water).

Last but not least, the **Fish A Matter of Scales** demonstrator is currently exploring synergies with the "**Food Nutrition Security Cloud**"[37], with the potential to broaden the analytical capabilities developed to include nutrition considerations when addressing fisheries considerations.

Demonstrator 5 "Aquaculture Monitor"

The "**Aquaculture Monitor**" demonstrator showcases how **artificial intelligence**, underpinned by **open access** to **multi-source data**, can deliver **remote sensing capabilities** for **aquaculture cage detection**, **monitoring** and **classification** (coastal/pond; land-type), supporting **aquaculture planning** in coastal areas. Using the Blue-Cloud Open Science environment, the demonstrator has set up an **"Aquaculture Atlas Generation" Virtual Lab** delivering a range of products and capabilities to users:

- **Aquaculture Cage Atlas:** The Atlas offers an online overview of satellite data derived **maps of cages** and **cage clusters**, allowing users to visualize geospatial data and to access information at farm level (including feeding systems, farm material, farm design and fish species). Users can overlay farms with satellite images and perform different analysis, like comparing the footprint of coastal ponds and rice paddy fields and check their evolution over a period of 20 years. The products are delivered through a map viewer, allowing registered users to edit features of the detected cages and cage clusters and to automatically map across feature sets to enrich maps. Estimates of **cage activity** over a production season can be produced, provided there is a large enough data sample available. "Super users" can also edit farm attributes and validate them at farm level, publishing the results in a geo-network catalogue.
- **Aquaculture Ponds Atlas:** This test-service uses some of the same data sources and similar analytical processes as the Cage Atlas to deliver a **coastal land-use classification map** of **aquaculture ponds**, fully based on Copernicus data for its remote sensing component.

Aquaculture Monitor **Lead Partner:** FAO · **Scientific Partners:** FAO, CLS	
Potential users and applications	**Policy Impact**
• **Aquaculture data analysts and managers:** To monitor how aquaculture develops over time in their area of interest and inform management decision-making processes. • **Data managers in spatial advisory units for maritime spatial monitoring and planning:** To monitor and inform decision-making processes related to maritime spatial planning. • **System developers:** To access a "template" solution for the management of sentinel and other satellites data access and processing in WeKEO. The demonstrator example workflow on cage monitoring can be adapted to other data and analytical WeKEO processes. • **Other communities** (e.g. aquaculture industry and SMEs)**:** To benefit from a FAIR compliant data management solution spanning statistical and geospatial data workflow in the fisheries,	Knowledge generated through this **Open Science ecosystem** can contribute to support the **EU** "**Farm to Fork Strategy**" in its objective to develop the potential of **sustainable seafood** as a **source of low-carbon food**, but also as the source of new sources of **protein** that can relieve pressure on agricultural land. It can specifically support better **aquaculture monitoring**, better **impact assessments** of fish and plant farming on marine ecosystems and inform more **sustainable aquaculture management** decisions. It can likewise support monitoring progress towards the UN SDGs, more specifically **SDG 2** (Zero Hunger), **SDG 13** (Climate Action) and **SDG 14** (Life Under Water). In the future, it could be scaled up and improved with **more user-friendly interfaces** and with new application interfaces (APIs) to provide information to the **general public** on aquaculture

[37] https://www.fns-cloud.eu/

aquaculture and related aquatic and land-based domains.	locations, production and tracking throughout different web portals.
Users of this **Virtual Lab** will be exploiting and benefiting from open access to the following data sources: ● **WekEO:** Sentinel 1 (S1) and Sentinel 2 (S2) ● **CLS:** VHR images ● **FIRMS**	

To support users in the deployment of these capabilities, the Virtual Lab currently provides use cases of fish farm detection for **Greece** and **Malta**; and coastal ponds and rice fields detection for **Sulawesi** and **Indonesia**. Advanced users will be supported in 2021 with facilities to run **additional analytical services** using Copernicus and other data, while FAO will provide a **mapping service** to connect to **local statistical datasets**. The long-term ambition of this effort is to deliver a full-fledged service providing overviews of **national aquaculture sectors**, using OGC compliant data services and available capabilities across different Virtual Labs available in the Blue-Cloud Open Science environment.

In a similar way to demonstrator 4, the **Aquaculture Monitor** demonstrator shows the potential of servicing **intermediate users** (such as fisheries data analysts) with a view to also informing **policy makers**, potentially evolving in the future towards engaging the **general public**, with wider policy implications:

- In the short- to medium-term, this demonstrator can contribute to support the "**Farm to Fork Strategy**", which is seeking to develop the potential of **sustainable seafood** as a **source of low-carbon food**, but also as the source of new sources of **protein** that can relieve pressure on agricultural land. **Aquaculture** has a crucial role in meeting the challenge of how to feed a growing population while reducing the environmental impacts of food production. It has already overtaken fisheries as the main source of fish for human consumption and accounted for 52% of global production in 2018. FAO forecasts that aquaculture's share of production will increase to 59% (109 million tonnes) by 2030. While fish is less resource-intensive to farm and has a lower carbon footprint than meat, the decline of wild-fish stocks at biologically sustainable levels will continue to drive demand for Ocean aquaculture upwards, requiring **closer monitoring** to keep its **impact** on **marine ecosystems** in **balance**. As the **Aquaculture Atlas Generation Virtual Lab** allows mixing in-situ data (farm inventories) with satellite data, it can support EU and global international monitoring efforts, by for example providing **early warning of diseases** (combining aquaculture data with CMEMS data on currents) or facilitating environmental impact assessments (combining aquaculture data with CMEMS ocean colour and seagrass maps). In short, combining current data on aquaculture with other marine environmental variables, indicators and other ancillary data such as site-inventories can contribute to support better **aquaculture monitoring**, to better **assess the impact** of **fish and plant farming** on **marine ecosystems** and **informing more sustainable aquaculture management decisions** and **relevant policy making processes**.
- As with the "**Fish A Matter of Scale**" demonstrator, in the future, this virtual lab also opens a window of opportunity to be used as a sound **educational tool** to inform the general public, so that consumers become more aware and are able to connect the impact of their **food consuming behaviour** on **marine ecosystems**.

At the moment of writing this document, the Blue-Cloud demonstrators are still working on closed environments. However, their full potential will unleash once their **Virtual Labs** are open to the public, inviting users to use and test the different **analytical frameworks** they provide to **service** and **support Open Science** in the **marine domain**. Over the next year, the Blue-Cloud project will have the opportunity to gather feedback from all early users of these Open Science environments to finetune its infrastructure and its related services.

In the first years of its development, the Blue-Cloud Open Science environment is expected to mainly service **scientists & researchers** of the **marine community**. However, as an Innovation Action, its future ambition is to serve users **all along Europe**'s **marine knowledge value chain**, evolving its capabilities to service policy makers and blue economy SMEs and industries, and ultimately citizens, **closing the gap** between **scientists** and **society** to **advance science-based solutions** and **align collective action**. Its services could be used to develop tailored applications to different societal challenges, including helping to track pollution, climate change or other impacts of human activities in real-time. New services developed through future, upscaled Virtual Labs could be open for use by all economic actors, public authorities and civil society to monitor, observe and enable responses to **climate change**, enhance **maritime spatial planning**, support **safe navigation** or promote **citizen participation** in **ocean** and **water governance** and **policy making**. They could serve as **risk-management tools** for the financial and insurance sectors to assess sustainability credentials of investments. **The possibilities are endless**. However, while it is good to aim high, as resources are not endless but **scarce**, in Section 5.3 we explore some **concrete use cases** that could be prioritized for future development. But before getting ahead of ourselves, in the next section we explore some of the lessons learned through the development of these use cases, which are currently testing the Blue-Cloud functionalities and can signpost some of the challenges that will still need to be addressed in the medium- to long-term, and thus are relevant to inform future developments.

5.2 Lessons learned from the Blue-Cloud demonstrators so far

The Blue-Cloud project kicked off in October 2019 and it has invested its first 15 months in developing the infrastructure and services supporting its demonstrators (as described in **Section 4**), interacting with them and feeding from their internal developments to set up the services that are supporting and enabling their research objectives (such as the **B-C Data Discovery & Access Service**, the **B-C Virtual Research Environment** or the different **Virtual Labs** set-up). These assets will continue to evolve and improve throughout the project. As early adopters of the Blue-Cloud's core services, the demonstrators provide useful insight into some of the emerging, (direct and indirect) positive impacts of the project, but also the challenges still faced to fully cater to their needs. Some of these challenges will be resolved through the project, with the resources available. Others signal greater challenges beyond the scope of the project, which need to be properly identified and understood, in order to tackle them through future efforts.

Patricia Cabrera (VLIZ)

*"Thanks to the Blue-Cloud project, our "Zoo-and Phyto Plankton EOVs" demonstrator has gained interest from **major biodiversity initiatives**. For instance, we are currently exploring how the demonstrator can contribute to the development of the post-2020 global biodiversity framework in the **Convention on Biological Diversity**".*

From a positive perspective, the value of weaving a multi-stakeholder, Open Science community in the marine domain is showing clear benefits. Bringing the **multidisciplinary teams** engaged in **blue data infrastructures**, **e-infrastructures** and **scientific research together** is demonstrating value towards a much-needed dialogue and interaction that contributes to a better understanding of "user needs" and interoperability challenges, creating momentum for closer collaboration. But the value of the community extends beyond this "technical" dialogue and into the realm of signalling potential, specific applications of the data products developed for specific blue economy industries and into advancing science-based policies. For example, the **Zoo- and Phytoplankton EOVs** demonstrator's innovative model can lead to a better understanding of what are the most important **pressures** to **marine ecosystems** at the **base of the food web**, which is relevant to the **fisheries industry** and can further inform the EU "**Farm-To-Fork Strategy**", EU "**Biodiversity Strategy**" and **UN SDG** "**Zero Hunger**". As a prove of value to policy, through the networking opportunities raised by the project the demonstrator has gained interest from major biodiversity initiatives, like the **Convention on Biological Diversity**[38]. The Blue-Cloud demonstrators are now exploring ways of linking and cross-fertilizing their efforts, which could lead to new, significant proof of added value (e.g., showcasing how "data products" can evolve towards "indicators" to inform specific policies).

On the more challenging side, issues connected to **data harmonization** and **interoperability** (due to the lack of common standards across blue data infrastructures, especially for in-situ observations), but also to the **interoperability between blue data infrastructures and computing infrastructures** for the execution of processes still remain, limiting the possibilities of Open Science practitioners. Potential approaches to address these challenges are being explored and will be part of **technical recommendations** delivered by the Blue-Cloud project through the final **Blue-Cloud Roadmap to 2030**. For example, one important technical consideration emerging from the internal dialogue established amongst infrastructures and researchers is the fact that **standardized**, **harmonized** and **scalable data access systems** are a necessary condition for the deployment of scalable algorithms (i.e., the capability to effectively exploit big data). Interoperability and data harmonization standards are the way for cost-efficient and sustainable developments, both for scientific research and applications. Related to this is the need to achieve separation between **computing infrastructures** and **algorithms/analytical methods** (the "workflows" referred to when describing the capabilities and services developed by the Blue-Cloud demonstrators). Ideally, a **FAIR catalogue of algorithms** should make it possible to **deploy any desired, available algorithm on any computing infrastructure**, **on-demand**, **closer to** the required "input" **data**. Making this possible requires the definition of a **protocol for interoperability** of **algorithms** across **computing infrastructures**. This would open new opportunities, enabling the possibility of implementing a more reliable, distributed "system of systems" (underpinned by all federated infrastructures) that could serve a wealth of input data into analytical applications (which in the case of big data is a necessary condition for successful discovery and exploitation).

Feeding from the comprehensive input received through initial stakeholder dialogue throughout 2019-2021, we have summarized the Blue-Cloud "**strengths**" and "**weaknesses**" and brought them together with the landscape of "**opportunities**" and "**threats**" surrounding the development of web-

[38] https://www.cbd.int/conferences/post2020

Blue-Cloud

based Open Science in the marine domain (as described in Section 2), to build a **SWOT analysis** that can be useful towards guiding the effort of developing the Blue-Cloud Roadmap to 2030 (see Table 2). This **SWOT analysis** is included below for consideration in the context of the discussion laid out in **Section 7**, which outlines the approach towards building a **long-term vision** for the development of the Blue-Cloud's efforts into the future.

Table 2: Current outlook of B-C Strengths, Weaknesses, Opportunities and Threats (SWOT)

Strengths	Weaknesses
Internal: Specific to the Blue-Cloud's efforts	
▪ Brings together representative **Blue data and e-infrastructures** for further dialogue and concrete actions to further address interoperability issues. ▪ Improved, **stronger analytical framework** (Virtual Research Environment) developed building on previous efforts (i-Marine, Blue-Bridge project) and tested for different uses across different domains. ▪ **Builds** on **existing European initiatives** and enhances them, not replaces them nor reproduces them. ▪ Enables and supports implementation of **FAIR principles** in the marine domain (through use of **standardized metadata**). ▪ Will provide feedback to improve and consolidate **good practices** towards **FAIRness** of **research outputs** (concerning metadata and publication protocols). ▪ Supports a **culture change** towards **Open Science** by providing an example of best practice. ▪ Is instrumental in familiarizing and **training users** in the deployment of Open Science tools in the marine domain, in anticipation of future developments (EOSC, DTO). ▪ Geared at showcasing the value of Open Science towards **EU Green Deal** and **UN Agenda 2030**.	▪ Engaged blue data and e-infrastructures and teams need financial resources to continue supporting services "free of charge" to users beyond the project, ensuring **sustainability** of efforts. ▪ Data harmonization and interoperability across participating blue data infrastructures not yet solved. ▪ Lack of **interoperability protocols** between blue data infrastructures and e-infrastructures. ▪ **Limitations** of "**data duplication**" **model** (i.e. data sets are "copied" into analytical framework). Hosting = expensive. ▪ APIs developed to connect with blue data infrastructures depend on infrastructures maintaining them **updated** with technical specifications (or else they stop functioning). ▪ **Small user base**, limiting potential towards triggering ideas for innovative applications. ▪ Working in an **"enclosed" analytical environment** could potentially raise barriers to reach larger pools of potential users of marine data.
Opportunities	**Threats**
External: Framing the context in which the Blue-Cloud project is developing	
▪ Current technological developments anticipate an unprecedented **growth** in **abundance** of **marine data**. ▪ The price of computing resources is lowering for users as "**pay-on-demand**" for **cloud-computing** is introduced. ▪ Strong interest in the marine community in **sharing marine data**, **processes** and **results**, creating momentum towards **advancing metadata standards** and to agreeing on basic principles for data sharing. ▪ **GreenData4All** space opens the opportunity for linkages with other domains/sectors (i.e "human health", other economic sectors) and multidisciplinary research. ▪ EOSC will expose existing blue data infrastructures and researchers to much **wider audiences**, providing more "reasons" for effective collaboration. ▪ The global transition towards **data science** will make a wealth of scientific methods, processes, algorithms and systems available to extract knowledge and insights from big data. Emerging initiatives such as Digital Twin of the Ocean will welcome **common languages** in support of **data**	▪ Lack of **FAIR data management** (as not fully appreciated nor mainstreamed in the marine domain is **impeding** progress towards fully seizing the potential of **Open Science**. ▪ Open science currently relies on **individual scientists** with heavy workloads, lack of time/resources/recognition for data management and balanced priorities towards delivering academic publications (favouring data embargoes). ▪ Resources required for **FAIR data management** (estimated at 10% of research cost) to enable Open Science are **not** currently **embedded** in **research budgets**. ▪ **Lack** of agreement on **metadata schemes** in marine domain. ▪ Virtual Labs offering analytical tools to marine researchers are widely set up across different e-infrastructures, with a **fragmented offer** resembling the "old days" of marine data. ▪ Data management efforts are seen as **diluting resources away** from marine observing systems and marine research, which need continuity of long time series to support the system.

interoperability and **interrogation** and the integration of **analytical resources** to support their simulation capabilities. ▪ International developments (UN Decade of Ocean Science for Sustainable Development, G7 FSOI) open opportunities for **global collaboration** beyond EU efforts.	▪ Unproductive **duplication of efforts** and **sustainability** caused by lack of incentives pushing for interoperability amongst existing blue data infrastructures. ▪ Lack of **standard protocols** of **interoperability** between blue data infrastructures and e-infrastructures limits potential. ▪ Data ownership and the road to FAIR data vs open data (differences in data policies and access rights).

Is this SWOT analysis complete?

The roadmap for the future evolution of the Blue-Cloud efforts needs to be built on a **comprehensive understanding** of the strengths and limitations of Blue-Cloud's assets, as well as of the **landscape** of **opportunities & threats** surrounding its development. Is the landscape above complete? What is missing? **We will be inviting feedback on this subject through the Blue-Cloud Roadmap to 2030 consultation.**

5.3 Wider use cases for future development

As described in **Section 4**, the assets developed through the Blue-Cloud project have been set up to continue servicing Open Science beyond project end. While the B-C demonstrators showcase the value that Open Science can deliver towards supporting the realization of the overarching goals of the **EU Green Deal** and **UN Agenda 2030**, Open Science needs to be deployed at a much larger scale to contribute a significant impact. In **Section 7**, we argue in support of providing **clear guidance** and **incentives** for the development of Open Science in the marine domain, prioritizing those areas where **collaborative research** has a higher degree of **feasibility** and can deliver relative **higher impact** towards our target policy objectives (see Figure 11).

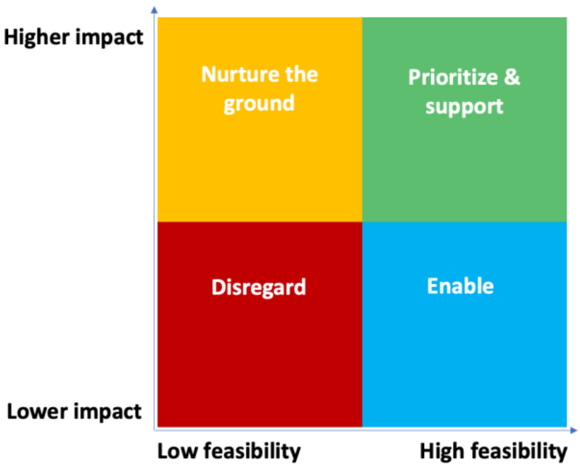


Figure 11: Tentative criteria for the development of future demonstrators. Credit: Adapted by Seascape Belgium

To better understand what "**feasibility**" and "**impact**" could mean in our context, we have outlined some general criteria:

- **Feasibility:** Abundant **data** and **knowledge** already exist around the topic; plenty of **actors** are already working on the field and making progress; and actors could **benefit** from sharing efforts, with potential **win-win scenarios** for **collaboration**.
- **Impact:** Collaboration in a virtual, web-based environment can deliver **significant added value** (i.e., societal benefits) along the priorities of the **EU Green Deal** and **UN Agenda 2030**.

Through the consultation on the **B-C Roadmap to 2030** we will seek to further qualify these criteria, as well as gathering ideas for **potential demonstrators** that could possibly satisfy them, so as to articulate priorities for further support and development.

Initial stakeholder dialogue has already provided insight for tentative "**use-cases**" that could exploit the Blue-Cloud services to deliver societal value. Ideas stemming from the Blue-Cloud community have signposted ideas for **virtual labs** across a broad range of topics, including collaborative efforts towards **Marine Spatial Planning (MSP)**; **monitoring biodiversity**; designating and monitoring **Marine Protected Areas**; artificially intelligent systems to map indicators and predict **coastal flooding** and **erosion**; but also transdisciplinary demonstrators that integrate social science and economic data or health data and wellbeing data to inform future decisions or to influence and support **behavioural change**. Blue-Cloud is in dialogue with several EU projects to identify synergies and opportunities towards using Blue-Cloud infrastructure and services to support on-going marine research and innovation initiatives. The outcomes of this dialogue, together with feedback gathered through the public consultation on the B-C Roadmap to 2030 (see Section 1) will support the identification of future applications of web-based open science in the marine domain around specific user needs, such as the one showcased as example in the box below.

Open Science applications in support of Marine Spatial Planning

Credit & innovators: DG Joint Research Centre · Cogea (https://www.cogea.it) · CETMAR (https://cetmar.org/?lang=en)

MSP -the process of ensuring that human activities at sea take place in an efficient, safe and sustainable way- has been gaining momentum in Europe over the past decade, as it encourages compatible uses of marine space and resources, reducing tensions around conflicting uses and balancing sustainable use and marine conservation. Striking a balance among the ecological, biological, socioeconomic and institutional aspects within ecoregions is one of the pillars of **ecosystem-based management** (EBM). The increasing availability of data on **human activities** in **European coastal regions** and **seas** (readily accessible through **EMODnet**) opens an opportunity for **Open Science applications** deploying **analytical models** to support better and more timely **planning decisions**, contributing to optimizing the sustainable development of **EU blue economy**.

Supporting the implementation of the EU Marine Spatial Planning Directive

The **MSP Directive** will reach an important critical point in 2021, as the deadline for the establishment of **maritime spatial plans** by EU Member States is met. To date, most maritime spatial planning efforts have taken place at local and national level[39]. With the new spatial plans, setting up a **collaborative**, **EU-wide platform** that could support their aggregation, providing the context and analytical tools to ensure spatial and ecological coherence across borders and sectors and suggesting eventual adaptations would not only be **feasible**, but contribute considerable **added value** towards the policy objectives of the directive and creating the conditions for a more **sustainable use** of our seas and Ocean.

Modelling and assessing the socioeconomic impact of Marine Spatial Planning on coastal communities

In the last years, different initiatives have developed **models** to quantify the **socio-economic effects** of **MSP** on coastal communities[40]. Evolving and finetuning these models could further contribute to use existing data on **human activities** in coastal areas and at sea to develop different **use scenarios** of **available marine space**, providing policy makers with tools to test different planning options in real time. Besides gaining insight into uses that contribute **higher social** and **economic value**, with a **lower environmental footprint**, collaborating in an EU platform would allow them to factor in how **neighbouring countries** are planning their ocean space, enabling for example **cost-efficient, collaborative investment decisions** (i.e., on deployment of shared Ocean energy grid infrastructures).

[39] https://www.msp-platform.eu/msp-practice/msp-projects
[40] Surís-Regueiro J.C. *et al* (2021) An applied framework to estimate the direct economic impact of MSP. Marine Policy 127

Potential users:

- **Public authorities:** To ensure and demonstrate a transparent and sustainable use of marine space.
- **Blue economy entrepreneurs, SMEs and industry:** To identify and evolve business opportunities benefitting from a level-playing field and sound economic planning in the marine space.
- **Civil society:** To monitor and influence the use of marine space according to societal needs.
- **Citizens:** To actively monitor and participate in policy decisions and Ocean governance.

Can we bring new ideas for future applications of Blue-Cloud's services?

The roadmap for the future evolution of the Blue-Cloud's efforts into the future will provide recommendations for funding and development of new collaborative, web-based Open Science applications that could benefit from exploiting the Blue-Cloud's assets and services to address specific user needs. **We will be inviting feedback on this subject through the Blue-Cloud Roadmap to 2030 consultation.**

6. Evolving the Blue-Cloud community

The **Blue-Cloud community** can be considered one of the **key assets** stemming from the Blue-Cloud project and one of the cornerstones of its successful evolution into the future. In this section we look deeper into the efforts invested in creating **awareness** and **traction** for Blue-Cloud's services up to date, the current constituents of the community that is evolving around these efforts, the actions planned to **open** and **expand** this community throughout the project and how it is progressing to shape up as a **model thematic service** for **EOSC**.

6.1 Creating awareness and traction for Blue-Cloud

The B-C Project has established different **mechanisms** to invite targeted **stakeholder communities** (as described in Section 3) to engage with its efforts. These mechanisms span from basic **communication activities**, including web-based news items, newsletters, webinars and wide social media outreach, to **motivation mechanisms** articulated around the development of **synergies** with other initiatives.

From a **communication perspective**, the Blue-Cloud project's website[41] plays a central role in communicating with **stakeholders** but also with **users**, being the entry point to the **B-C web-based Open Science environment (Virtual Research Environment**, as described in **Section 4**). It also provides access to key assets, such as the **B-C Catalogue** and different **user-oriented training materials**. It showcases progress and results achieved by the **B-C demonstrators** (as introduced in **Section 5**), including **webinars** and **interviews** with their teams and with representatives of the blue data infrastructures teaming up with their efforts. At the time of writing this roadmap, the website has reached well over **11,000 visitors**.

41 https://blue-cloud.org/

From a **motivational perspective**, as the B-C seeks to evolve towards shaping up as a **thematic service** for **EOSC**, interacting with peer projects that could potentially benefit from its services (e.g., FNS Food Cloud, amongst many others) and teaming up with other efforts (e.g., EOSC-FUTURE project, where science clusters such as ENVRI-FAIR, EOSC-Life, ESCAPE, PANOSC or SSHOC will connect and interact with EOSC core services) is key to better understand how the B-C services should evolve to address their needs and expectations, and build on complementarities, generating **synergies**. An overview of synergies established up to date is available here: https://blue-cloud.org/synergies.

6.2 Expanding the Blue-Cloud's user base

Up until now, the B-C services have been operating on a "testing" mode, restricted to the B-C Project's circle of "early users". Over 300 users from the B-C Consortium and beyond have accessed these services. In the **second semester** of **2021**, the B-C Demonstrators will start opening their **virtual labs** to the "public", inviting other **scientists** and **researchers** to **test** and **experiment** with their many functionalities, "enriching" them with new data, suggesting improvements of their algorithms, developing new data products, models, hypotheses and/or results, including potential new, innovative applications.

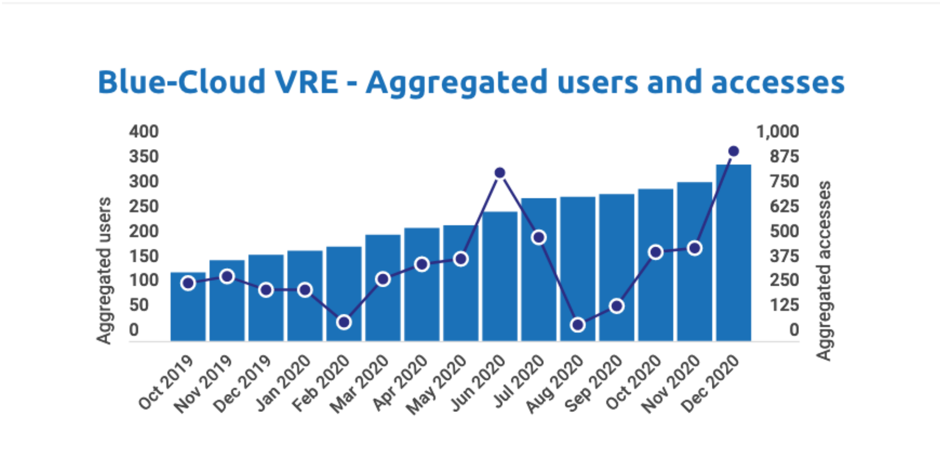


Figure 12: Aggregated users and accesses to B-C Virtual Research Environment. Credit: Trust-IT

In **January 2022**, a **B-C Virtual Hackathon** will be organized and delivered to promote the **Blue-Cloud** to various user communities and serve as a training and capacity building activity, promoting the **testing**, **uptake** and **further use** of the **B-C services**, but also to collect **user feedback** on such services. The hackathon will gather participants from different user communities, tentatively including users of marine data & data service providers, students, ICT experts and hacktivists, innovators, both from the public and private sector (blue economy & industry), government officials and policy makers.

In the future, and as discussed in Section 7, the B-C services will seek to cater to a broad range of user segments. Improving the offer with adequate, fit-for-purpose **user-friendly interfaces** will be key to achieve this objective.

6.3 Aligning with wider developments

The Blue-Cloud services are being deployed with attention to other, wider developments that in the medium term will positively impact Europe's **marine knowledge value chain**, such as the upcoming **EU pilot Digital Twin of the Ocean** referred to in **Section 2**; the **EuroHPC**[42] joint undertaking, which will develop a world class, supercomputing ecosystem in Europe; or the Gaia-X project[43], building the requirements for an EU data infrastructure. Through this roadmap, the B-C project will reflect on ways in which its **key exploitable results** could be **further capitalized** by and **benefit** from these initiatives.

6.4 Blue-Cloud as a model thematic service for EOSC

> *"Europe is going to co-create a framework to allow the use of data. It should consist of a trusted pool of non-personalized data that governments, businesses and other stakeholders can contribute to. And thereby, we open up data as a resource for innovation and bring new solutions to the market. And our scientists are already beginning to do this. We are creating an European Open Science Cloud (EOSC) now. It is a trusted space for researchers to store their data and to access data from researchers from other disciplines. We will create a pool of information leading to a web of research insight."*
>
> *Ursula von der Leyen, President of the European Commission*

On January 21st, 2020, during the annual meeting of the **World Economic Forum**, Ursula von der Leyen, President of the European Commission pointed out the value of the **European Open Science Cloud** and the absolute importance of data as a key resource for innovation. The Blue-Cloud project is based upon the same beliefs and is working to bring the following elements to EOSC:

- The experience of a **thematic** ("marine") **EOSC** that can serve as a role model and 'blueprint' for the development of other thematic clouds within EOSC. The cyber-platform built by the Blue-Cloud project will provide access to multidisciplinary data, analytical tools and computing and storage facilities to support collaborative research in support of the **EU Green Deal** and **UN Agenda 2030**.
- A set of **domain-specific services** deployed through B-C pilot demonstrators which address specific research and innovation needs for Oceans, seas and freshwater bodies. These services will support marine ecosystem research, conservation and forecasting, as well as enabling innovation in the Blue Economy. The **Blue-Cloud Catalogue** -Blue-Cloud's thematic catalogue of services made available through its Virtual Research Environment- is being designed and developed for interoperability with **EOSC**, with the purpose of making B-C's assets and services

[42] https://eurohpc-ju.europa.eu/
[43] https://www.data-infrastructure.eu/GAIAX/Navigation/EN/Home/home.html

findable and accessible via EOSC's **Portal Catalogue & Marketplace**[44], creating opportunities for new applications and workflows. The included service metadata will include APIs to access specific services with blue thematic functionalities. The **B-C Catalogue** is one of the ways in which the B-C project will be contributing to cluster **Open Science marine thematic services** through EOSC by 2022, for easy discovery, access, interoperability and re-use by a much wider community. In the medium to long term, EOSC will be instrumental in deploying B-C's strategy of promoting wider user uptake of the results of Open Science, supporting the transformative science and innovation needed to deliver the **EU Green Deal** and **UN Agenda 2030.**

- A mechanism to **easily access** and **discover blue data**. The blue data infrastructures, federated in the Blue-Cloud Data Discovery & Access service manage important volumes of blue data, which are made easily discoverable and accessible. The metadata on the Blue-Cloud data resources will be shared and included in the EOSC metadata catalogue service with deep links to Blue-Cloud Data Discovery & Access service to facilitate the actual data retrieval by EOSC user communities.
- A proven methodology showing how researchers interacting with e-infrastructure developers can establish a cyber platform with tools and services, which support many scientific challenges and are fit-for-purpose, meanwhile building a thriving community and synergies between IT developers and blue domain researchers, which can provide a learning curve for the further deployment and exploitation of EOSC into a successful initiative.
- Bringing in the Blue-Cloud community towards exploring the added-value and opportunities of EOSC and taking an active role in making EOSC more attractive and dynamic.

[44] https://marketplace.eosc-portal.eu

7. The road ahead: Vision & paths of action to 2030

The Blue-Cloud project is building a clear case for web-based Open Science in the marine domain. As a three-year project, the investment, effort and results delivered will require resources for their future capitalization and further development. What should be the **overarching vision** for the future evolution of these efforts? What **actions** can be articulated to guide progress towards such a vision? **Who** would have to **join forces** in this **mission**? What **resources** will need to be mobilized? What **impact** should be expected from these efforts? In this section we explore the **road ahead**.

7.1 Methodological approach & strategic framework

As stated in **Section 3**, defining a **vision** and identifying the **strategic goals** and **actions** that should guide the future development of the Blue-Cloud's efforts requires **aligning action** with a wealth of stakeholders who are currently knitting and building the underlying foundations for a thriving Open Science community in the marine domain. While in times of Covid-19 convening a large community of stakeholders to invite a long conversation to build such a shared vision has become a difficult task, intensive efforts towards initiating a remote conversation have delivered considerable input and insight. Feeding from initial dialogue with key stakeholders (as introduced in Section 1), we have produced this early draft of the Blue-Cloud Roadmap to 2030 following a "*backcasting*" approach, consisting of working backwards from a very broad "**vision of success**" ("*realizing the full potential of Open Science in the marine domain, in support of the EU Green Deal and the UN Agenda 2030", as introduced in Section 3*) to deliver a comprehensive overview of the challenges ahead and to reflect on the different ways in which the Blue-Cloud efforts could evolve to contribute to the realization of such vision. We suggest using this exercise as the starting point for a conversation that will evolve throughout 2021 and into 2022 to narrow the scope of this vision and underpin it with more specific objectives, feeding from the comments and contributions of the Blue-Cloud Community to co-create a shared, targeted **Vision 2030**, including a set of **policy** and **strategy recommendations** to **2030**.

To start infusing concrete meaning into this **vision of success**, we have sought to respond to the following **questions**:

- What does successful, web-based Open Science in the marine domain look like?
- What does successful, web-based Open Science in the marine domain need to realize its full potential?
- What key actors are currently servicing the needs of the community to enable web-based Open Science?
- What specific added-value will the Blue-Cloud bring to this landscape by 2022?
- What are the benefits to users of the Blue-Cloud assets and services?
- How could the Blue-Cloud's efforts evolve - working with key actors along the marine knowledge value chain - to provide an open-source community, web-based framework that supports a wealth of users in conducting web-based Open Science in the marine domain, with key assets including open access to **interoperable data**, **analytical tools** and **shared computing power** by 2025? By 2030?

In answering these questions, we seek to build a **strategic framework** to guide our backcasting exercise towards developing **strategic goals** and **actions** that can help us connect our desired future with the present. As a next step, departing from the **key exploitable results** that will emerge from the Blue-Cloud project, we will then establish tentative, specific paths of action for their future evolution and capitalization. Through this document, we will consult with the Blue-Cloud community on the validity of these paths of action (*"Are they appropriate? Are they enough? What specific actions should we undertake and prioritize?")*, as well seeking to gather additional input to align a set of policy and strategic recommendations towards their realization.

It is important to note that the task that we have ahead of us is to define a vision for the **future evolution of Blue-Cloud's efforts**, and **not** for the future evolution of Open Science in the marine domain. While defining "*successful Open Science in the marine domain*" provides a framing reference and a sense of direction for our purpose, delivering such a vision would require a much wider analysis and reflection, deeper attention to other drivers and motivations and activating broader communities of stakeholders, all of which falls outside the scope of the Blue-Cloud's efforts. This boundary should be kept in mind when reading through this section.

<u>What does successful web-based Open Science in the marine domain look like?</u>

From a policy perspective, and for the purpose of our efforts, we depart from the assumption that Open Science in the marine domain will have cusped its momentum when it is recognized as driving the **science-based knowledge**, **transformative solutions** and **innovation** required to restore our Ocean and waters by 2030 and to deliver on the objectives of the **EU Green Deal**, contributing with European leadership, concrete actions and results towards meeting the **UN Agenda 2030 SDGs**.

<u>What does successful web-based Open Science in the marine domain need to realize its full potential?</u>

Feeding from different initiatives and expert insight[45], it can be argued that successful web-based Open Science in the marine domain will require:

- **Open and FAIR data:** A wealth of marine observations and data that fully capture the state of the Ocean and seas around the world that is openly available to anyone with an interest in exploiting it. This wealth of data should be FAIR (findable, accessible, interoperable and reusable) and to the extent possible open and free, but not necessarily so (in full alignment with EOSC recommendation regarding the protection of security and property rights[46]).
- **Analytical resources:** A wealth of cloud computing resources and supporting web-based analytical tools allowing anyone interested in transforming data into knowledge to be able to attempt to do so, at a reasonable cost.
- **Infrastructures:** Basic physical and organizational structures and facilities needed to host, support and operate web-based, open data and analytical resources.
- **People:** A critical mass of participants across the marine and maritime community and other interdisciplinary communities, requiring incentives; rewards and metrics to recognize FAIR

[45] Gagliardi, Dimitri & Cox, Deborah & Li, Yanchao. (2014). *What are the factors driving and hindering the adoption of Open Science? An exploratory study.* Manchester Institute of Innovation Research working paper series. 10.13140/2.1.3016.9602
[46] EOSC Strategic Research and Innovation Agenda Version 1.0 https://www.eosc.eu/sites/default/files/EOSC-SRIA-V1.0_15Feb2021.pdf

and OPEN practice; training; and fit-for-purpose interfaces that facilitate, encourage and support wide participation.

- **Policies:** Principles of action for web-based Open Science, widely accepted by the broad marine and maritime community, within the overarching principles provided by EOSC[47].

What key actors are currently servicing the needs of the community to enable Open Science?

As introduced in Section 2, the **marine knowledge value chain** has seen extraordinary progress in the last 30 years. European and international efforts to enhance this knowledge ecosystem are converging with European efforts towards **EOSC**, which means that the "**needs**" of Open Science in the marine domain -as described above- are being served and addressed by a wealth of **actors** and **initiatives**, with overarching **goals** (i.e. building marine observation capabilities in Europe; promoting stewardship in data management and facilitating data access; guiding and nurturing efforts towards FAIR and OPEN data; developing e-infrastructures to computing and digital needs; setting up systems of incentives and rewards for researchers; defining policies in support of Open Science, etc) that go well beyond the scope of the Blue-Cloud project. Servicing the needs of Open Science in the marine domain is a **colossal effort** and the road to success one facing significant hurdles across different topics, calling for comprehensive, aligned, multi-stakeholder action. When designing the Blue-Cloud Roadmap to 2030, we should be asking ourselves: *Who are the key actors currently servicing the needs of the marine community towards enabling Open Science? What overarching goals are they seeking to achieve? How can the Blue-Cloud efforts evolve to support these actors in delivering on the "needs" of Open Science in the marine domain, towards the goal of achieving successful Open Science?* Responding to these questions will allow us to define the boundaries of the Blue-Cloud **added value into the future**, providing a clear and concrete "**mission**" to guide its future evolution.

What specific added-value will the Blue-Cloud bring to this landscape by 2022?

The answer to this question lies in **Sections 3, 4, 5** and **6**, although a clearer outlook will emerge in the coming months, as the Blue-Cloud services are tested and exploited by its early users, confirming their value but also signposting outstanding challenges.

How could the Blue-Cloud's efforts evolve to service a wealth of users and support key actors towards achieving progress in the delivery of a collective vision of success for web-based marine Open Science by 2025? By 2030?

This is the key question that we are seeking to respond through this Roadmap. To facilitate the discussion on the priorities that the evolution of the Blue-Cloud's efforts should seek to attend, we have captured answers to some of the questions posed above. We also have articulated a set of overarching goals that the Blue-Cloud efforts could evolve to serve, towards delivering on the vision of unlocking (2025) and realizing (2030) the full potential of web-based Open Science, in support of the **EU Green Deal** and **UN Agenda 2030**. Table 3 below brings together our projected "vision of success" for web-based Open Science in the marine domain, the different components of such vision, the results that B-C will contribute by 2022 through its **key exploitable results** (KER) and **potential paths of action** that could be laid out for the **future** evolution of these assets **(2025-2030)**.

[47] EOSC Strategic Research and Innovation Agenda Version 1.0 https://www.eosc.eu/sites/default/files/EOSC-SRIA-V1.0_15Feb2021.pdf

Table 3: Strategic Framework Towards Developing a Vision 2030

The vision of success for each "Open Science supporting component" stems from the relevant policies motivating the B-C efforts (EU Green Deal & Mission Starfish · UN Agenda 2030 & Decade of Ocean Science for Sustainable Development)

Vision of Success 2030 What does successful Open Science in the marine domain look like in Europe by 2030?		Blue-Cloud 2022	Into the future 2025 - 2030
Open science drives **science-based knowledge, transformative solutions** and **innovation** to deliver "**Mission Starfish**" objective of restoring our Ocean and waters by 2030 and to support the objectives of the **EU Green Deal**, contributing with European leadership and concrete actions towards meeting the **UN Agenda 2030** SDGs.		**Demonstrate value of Open Science to EU Green Deal and UN Agenda 2030**	**Unlock and realize potential of Open Science in support of the EU Green Deal and UN Agenda 2030**
What does Open Science need to support this vision?		**Possible Paths of Action**	
Open & FAIR data		**What will the B-C deliver? How should the B-C efforts evolve into the future to support these objectives?**	
Ocean observations	• European seabed fully & coherently mapped in high-resolution by 2030.	• Support FAIR through standard metadata, a Data Discovery & Access Service (DD&AS) and an EOSC compliant, scalable catalogue of "blue" research outputs. **KER:** DD&AS, underpinned by metadata standard for marine data sets. **KER:** B-C Catalogue, underpinned by metadata scheme for research outputs.	• Achieve wide adoption of common metadata standards across blue data infrastructures and marine community. • Support international efforts towards global Ocean data coverage and EU quest for FAIR data in the marine domain by gaining buy-in from the public and private sectors towards FAIR data sharing making a strong case for Open Science and innovation.
Open access to data	• All public marine data collected by Member States & EU is pooled centrally and freely available and all private marine data is FAIR.		
Standards for interoperability	• EU standards for marine metadata are available and aligned with international standards, enabling full machine-to-machine discovery and interoperability.		
FAIR data & research objects	• All EU marine data & research infrastructures are federated and have their full catalogues available in an integrated, open access EU Data Discovery & Access Service that supports a wealth of Open Science services and disciplines through EOSC.		
Federated infrastructures & analytical resources			
Marine data infrastructures, RIs, e-infrastructures	• Marine data, research infrastructures and e-infrastructures are federated into a single data space allowing users to access and exploit **data**, **computing** and **analytical resources seamlessly**, pooling their hosting, HPC and cloud computing power together to service the needs of 1.7 million researchers in Europe.	• Showcase approach to interoperability of infrastructures and signpost outstanding challenges. **KER:** DD&AS • B-C VRE offers users a virtual environment with a broad range of analytical services, underpinned by access to data across blue infrastructures and metadata standard for research outputs. **KER:** Virtual Research Environment	• Achieve wide adoption of DD&AS in support of federation of European blue data infrastructures. • Align with technical interoperability efforts towards standards and protocols in support of federation of data AND e-infrastructures. • Support simulation capabilities of European Digital Twin of the Ocean with analytical models and tools developed in virtual labs.
Virtual Research Environments			
Cloud computing resources			
Standards for interoperability	• Access and interoperability standards and protocols between computing infrastructures and workflows (i.e methods & algorithms) are available, allowing them to be deployed on demand, closer to data.		
Digital Twin Ocean	• EU global digital twin of all oceans and waters is operational.		
A diverse, representative and thriving community of practitioners, users and advocates			
Critical mass of participants	• 1.7 million researchers and 1 million professionals engage in Open Science, boosting research productivity and innovation along the marine knowledge value chain. • 1 million students use available Open Science capabilities to support their academic objectives. • Each European is a citizen of our ocean and waters.	• Build the case for Open Science in the marine domain. • Nurture and promote wider user uptake of B-C services. **KER:** Virtual Labs **Asset:** Community	• Mainstream and service Open Science in the marine domain, servicing a wealth of users through multidisciplinary, multipurpose virtual labs. • Achieve a diverse and representative community with incentives and fit-for-purpose user interfaces to bring all "blue" stakeholders onboard.
Incentives			
Rewards			
Training			
User-interfaces			
Policies			
Data policies	• Transparent and consistent terms for participation in Open Science are established and widely understood and accepted, achieving the trust and confidence required for broad engagement in Open Science.	• Alignment with policies established by EOSC	• Support EOSC efforts for wide adoption of policies, processes and procedures for Open Science in the marine domain that provide assurance of quality and trust.

7.2 Setting clear paths of action towards 2030

Extracting from the **strategic framework** in the previous section, the different **paths of action** identified and proposed have been structured to develop a **strategy** for the future evolution of the Blue-Cloud's efforts that revolves around **four strategic pillars**, namely:

- **Strategic pillar 1:** Grow a wealth of **applications of web-based open science in the marine domain**, delivering on **user needs** by supporting **collaborative research** around topics with a relatively higher degree of **feasibility** and **impact** towards the **EU Green Deal** and the **UN Agenda 2030**.
- **Strategic pillar 2:** Evolve a thriving, diverse and representative community of Open Science **practitioners**, **users** and **advocates** in the marine and maritime domain, including scientists & researchers, blue economy SMEs, industry, monitoring agencies, policy makers, civil society & citizens.
- **Strategic pillar 3:** Support efforts towards **FAIR & OPEN** Ocean data to make marine data more readily available and **accessible to all**.
- **Strategic pillar 4:** Align with efforts towards full **interoperability** of marine data infrastructures, Research Infrastructures and e-infrastructures to make **digital objects** and **computing resources** available and deployable on any infrastructure, **on-demand**, - supporting open and standardized approaches, where possible.

Buy building upon these **strategic pillars** (**Open Science applications** in the marine domain; **people**; **FAIR data**; and **federated infrastructures**), we seek to generate and exploit **positive feedbacks** that will reinforce achievements around each **pillar** (see Figure 13):

- Fostering and incentivizing applications of **web-based Open Science** in the **marine domain** that have a higher chance of materializing will help to catalyse the development of a **growing community of practitioners**, who in return can contribute to make collaboration for other research and innovation topics **more feasible**, **broadening the range of topics** covered.
- An expanding community of Open Science practitioners will contribute to win **buy-in for Open Science** and collaborative research, building a stronger case for **data sharing** and opening new opportunities for **data ingestion** for participating blue data infrastructures, which in turn will contribute towards FAIRification of available data and to **bring new users** and **advocates** to the community.
- Advancing **FAIR data** will make analytical environments more attractive, **driving more users to e-infrastructures** to seize their computing resources to combine, process and run analysis on growing catalogues of interoperable data sets. Federating e-infrastructures will in turn contribute to driving more users to those catalogues, creating a **win-win scenario for collaboration** between blue data infrastructures and e-infrastructures.

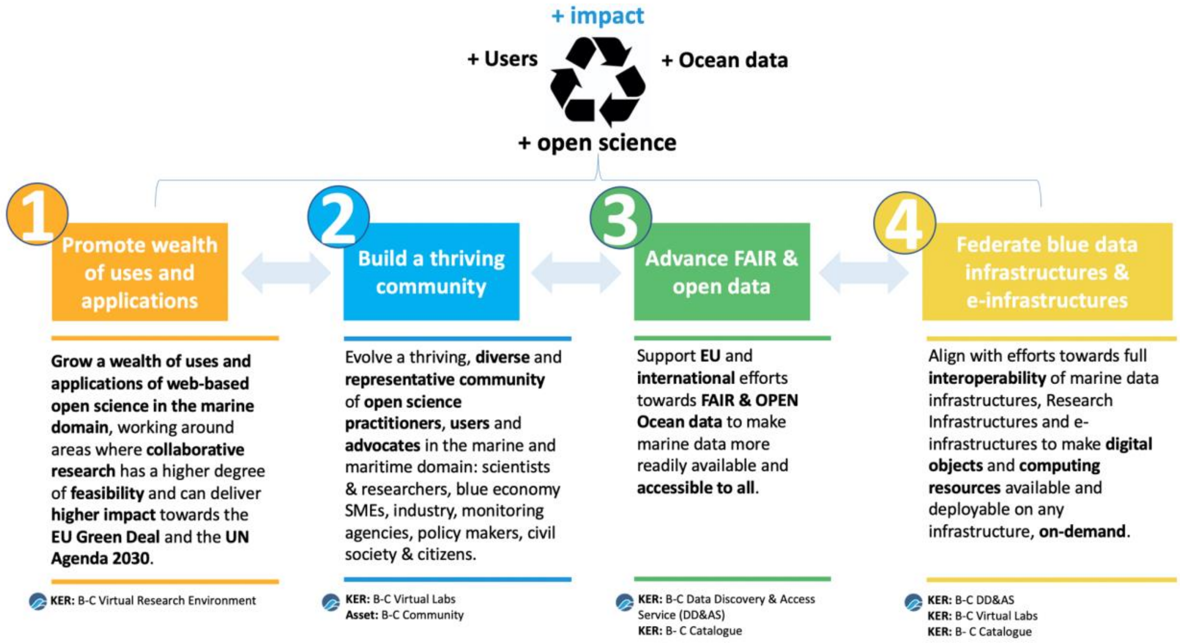


Figure 13: Positive feedback between strategic pillars and impact resulting from these interactions

Feedback from the **B-C community** on these strategic pillars will contribute to confirm or review their formulation, as well as guiding the **identification** and **prioritization** of **policy recommendations**, including **key actions** underpinning their implementation. The public consultation on the B-C Roadmap to 2030 that will launch in June 2021 (see Section 1) will seek to respond to questions such as:

- How should **Blue-Cloud** evolve to support a **thriving environment** for web-based **Open Science** and **Open data** in the **marine domain**?
- What **governance mechanisms** are most appropriate and desirable to ensure broad **engagement** of the **marine community** and to ensure widely accepted **rules of participation**?
- What **applications** of **Open Science** in **the marine domain** could have a higher probability of success, given current availability of data, models and actors willing to engage in collaborative science, across a broad range of topics? Which of such applications should be prioritized towards addressing current **user needs** and delivering **highest societal impact**?
- How should B-C's Open Science environment and services evolve to be fit-for-use not only for scientists, but also for other Open Science users such as **policy makers** and **blue economy SMEs** and **industry**? What needs do these users have that B-C could evolve to address? What **incentives** can contribute to bring Open Science practitioners on board?
- How can B-C evolve to **further connect** with **marine data infrastructures** and **research infrastructures** to deliver full interoperability of marine data through the **B-C Data Discovery & Access Service**, aligning and in collaboration with other international efforts?
- What actions would be required to enable **B-C's Catalogue** of **analytical methods**, **algorithms** and **applications** to **be deployed in EOSC,** but also in other infrastructures **-closer to data-** or across **supercomputing platforms** in Europe?
- How should **B-C's assets** evolve to align with future **EU pilot DTO** and **DestinE** developments?

7.3 Aligning a shared vision of added value to 2030

Moving forward with this backcasting exercise, the agreed **paths of action** shaped through wide stakeholder consultation will help us to better define the "**added value**" that the Blue-Cloud's efforts should seek to deliver into the future, on its journey to contribute to the **EU Green Deal** and **Mission Starfish**. With this in mind, we have attempted to articulate a **vision** of what we would like to achieve through our journey and a travel plan or **mission** indicating how we plan to get there, as a starting point guiding the conversation. Through the early draft **Roadmap** and the consultation, the community will be invited to convey their thoughts on this shared journey, welcoming all contributions. Our efforts will be carefully designed to invite as many travel companions on board as possible, to go as far as possible, together.

Vision 2030

"Big picture" of what we want to achieve

"Unlocking the potential of Open Science in the marine domain in support of the EU Green Deal and UN Agenda 2030 (including Mission Starfish & Decade of OSSD) by 2025 and realizing the full potential of Open Science by 2030 to guide the next decade of challenges".

Mission 2030

General statement on how we will achieve our Vision 2030

*"To contribute towards a **European cloud-based data space** that provides access to a thriving portfolio of **analytical**, **simulation** and **visualization capabilities** underpinned by **seamless access** to a **wealth of FAIR**, transdisciplinary **ocean** and **freshwater data**, enabling **Open Science** to deliver better science-based policies, innovation, public awareness and citizenship for a healthy Ocean, in support of the **EU Green Deal** and the **UN Agenda 2030**".*

Is this overall strategic framework appropriate?

Are there any important considerations missing in the **strategic framework** developed towards co-creating the **Blue-Cloud Vision 2030**? Will action around the **four strategic pillars** identified provide leverage for the future evolution of the Blue-Cloud's effort? Does this **Vision & Mission 2030** reflect the overall goal and path of action that the evolution of the Blue-Cloud's efforts should follow? Is there anything **missing**? **We will be inviting feedback on this subject through the Blue-Cloud Roadmap to 2030 consultation.**

7.4 Service exploitation and sustainability

The Blue-Cloud Project Partnership is made up of 20 beneficiaries[48] who are involved in different aspects of the development of the Blue-Cloud project and in delivery of its planned **Key Exploitable Results**, either directly or in other supporting activities, such as management, promotion, marketing, business planning or long-term strategy. Throughout the next months, Project Partners will have to convene who will continue as active partners in medium-term exploitation of the KERs and/or who will team up in the search for long-term sustainability. Moreover, the future roles of individual partners will have to be determined with regard to each KER and the different activities connected to their future exploitation. A first draft **Service Exploitation & Sustainability Plan**[49] has been produced to guide these discussions and the outcome will be available to inform the final version of this Roadmap, into 2022.

7.5 Expected Impact

The **Recommendations to 2030** and the **Action Plan** underpinning them will include a proposal of **Key Performance Indicators** to set targets, monitor progress and quantify impact of the evolution of Blue-Cloud's efforts into the future, towards measuring success.

[48] https://www.blue-cloud.org/partners
[49] D6.3 Blue-Cloud Services Exploitation and Sustainability Plan (Release 1)

Conclusions

This early draft of the **Blue-Cloud Roadmap to 2030** introduces the policy context, opportunities and challenges motivating **Blue-Cloud**'s efforts, exploring the emergence of **Open Science** in the marine domain in the context of the digital age and how it can contribute to **support** the **European Green Deal** and the **United Nations Agenda 2030 for Sustainable Development**. It analyses relevant developments shaping Europe's **marine knowledge value chain** and describes the **added value** that **Blue-Cloud** will bring to this landscape by **2022**, in the form of **Key Exploitable Results**.

Progress achieved so far in the development of Blue-Cloud's assets confirms the key role of **Open Science** in the marine domain in **delivering science-based knowledge**, **solutions** and **innovation** to address societal challenges. B-C's pilot demonstrators are showcasing value to **specific policy objectives**, identified in **Mission Starfish** and **SDGs 2, 13** and **14**. But also, through its services, Blue-Cloud is building on the solid foundations of Europe's marine knowledge value chain to continue making progress towards delivering on the needs of **successful open science in the marine domain**, fostering the development of science that is shared openly and available for re-use. Through its **B-C Data Discovery & Access Service**, it is simplifying access to data available across different blue infrastructures from within its **Virtual Research Environment**, which provides a strong analytical framework for different open science practitioners to meet and work together. **B-C**'s **Catalogue** will further make a wealth of analytical resources available for re-use. Last but not least, an **emerging community** of **early open science practitioners** is gathering around its existing **Virtual Labs**, which will inspire other users and practitioners to join and build a larger, thriving Open Science community in the marine domain, into the future.

Capitalizing and evolving the Blue-Cloud's efforts into the decade will be key to **unlock the full potential of Open Science in the marine domain**, enabling the development of a wealth of Open Science applications that can deliver science-based solutions and innovation to address policy objectives at a proper scale. This will require aligning action with a wealth of stakeholders who are currently knitting and building the underlying foundations for a thriving Ocean Open Data and Open Science community in Europe. This early draft of the B-C Roadmap to 2030 introduces an overarching **vision** to guide the future capitalization and further development of Blue-Cloud's **KERs**, proposing strategic, high-level **paths of action** that could be activated to guide progress towards such vision.

As the **first early draft "roadmap"**, this document **does not yet include policy recommendations**, but sets the scene for **wide stakeholder consultation** towards co-designing a **community vision** and identifying **concrete actions** for the future evolution of Blue-Cloud's **added value** into the decade, delivering insight into Blue-Cloud's current **strengths** and **weaknesses** and into the **opportunities** and **challenges** ahead. Its contents will be further discussed amongst **Project Partners** and the **B-C ESEB** throughout April and May 2021, leading to a shorter, more focused draft of the B-C Roadmap to 2030 that will be released in June 2021 for consultation, welcoming contributions throughout the Summer. The **feedback** and **contributions** gathered through this public consultation, together with additional dialogue with key stakeholder communities throughout 2021 and 2022 will shape the final, community-driven **Blue-Cloud Roadmap to 2030** which will be delivered to the **European Commission** by **July 2022** (D6.4).