### Project Details

<table>
<thead>
<tr>
<th><strong>Project Acronym</strong></th>
<th><strong>BlueBRIDGE</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Project Title</strong></td>
<td>Building Research environments for fostering Innovation, Decision making, Governance and Education to support Blue growth</td>
</tr>
<tr>
<td><strong>Project Number</strong></td>
<td>675680</td>
</tr>
<tr>
<td><strong>Deliverable Title</strong></td>
<td><strong>Best Practices &amp; Socio-economic Impact: Final Report</strong></td>
</tr>
<tr>
<td><strong>Deliverable No.</strong></td>
<td>D3.5</td>
</tr>
<tr>
<td><strong>Delivery Date</strong></td>
<td>November 2017</td>
</tr>
<tr>
<td><strong>Authors</strong></td>
<td>Sara Garavelli, Holly Ellis (Trust-IT Services)</td>
</tr>
</tbody>
</table>
## DOCUMENT INFORMATION

### PROJECT

<table>
<thead>
<tr>
<th>Project Acronym</th>
<th>BlueBRIDGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Title</td>
<td>Building Research environments for fostering Innovation, Decision making, Governance and Education to support Blue growth</td>
</tr>
<tr>
<td>Project Start</td>
<td>1st September 2015</td>
</tr>
<tr>
<td>Project Duration</td>
<td>30 months</td>
</tr>
<tr>
<td>Grant Agreement No.</td>
<td>675680</td>
</tr>
</tbody>
</table>

### DELIVERABLE

<table>
<thead>
<tr>
<th>Deliverable No.</th>
<th>D3.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deliverable Title</td>
<td>Best Practices &amp; Socio-economic Impact: Final Report</td>
</tr>
<tr>
<td>Contractual Delivery Date</td>
<td>November 2017</td>
</tr>
<tr>
<td>Actual Delivery Date</td>
<td>March 2018</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Sara Garavelli (TRUST-IT), Holly Ellis (TRUST-IT)</td>
</tr>
<tr>
<td>Editor(s)</td>
<td>Sara Garavelli</td>
</tr>
<tr>
<td>Reviewer(s)</td>
<td>Donatella Castelli</td>
</tr>
<tr>
<td>Contributor(s)</td>
<td>Gerasimos Antzoulatos (I2S), Debhasish Bhakta (FAO), Leonardo Candela (CNR), Gianpaolo Coro (CNR), Anna Davies (ICES), Anton Ellenbroek (FAO), Aureliano Gentile (FAO), Gabriele Giammatteo (ENG), Georgios Kakaletris (UOA), Yann Laurent (FAO), Miles Macmillan Lawler (GRID Arendal), Jean-Yves Le Bras (CLS), Nicolas Longépé (CLS), Yannis Marketakis (FORTH), Nikos Minadakis (FORTH), Eleni Petra (UOA), Paolo Scarponi (CNR), Anais Turpault (PMBret), Levi Westerveld (GRID Arendal)</td>
</tr>
<tr>
<td>Work Package No.</td>
<td>WP3</td>
</tr>
<tr>
<td>Work Package Title</td>
<td>Communication, Stakeholder Engagement and Knowledge transfer</td>
</tr>
<tr>
<td>Work Package Leader</td>
<td>TRUST-IT</td>
</tr>
<tr>
<td>Work Package Participants</td>
<td>CITE, CNR, FAO, I2S, PMBret, UOA</td>
</tr>
<tr>
<td>Nature</td>
<td>Public</td>
</tr>
<tr>
<td>Version / Revision</td>
<td>1.0</td>
</tr>
<tr>
<td>Draft / Final</td>
<td>Final</td>
</tr>
<tr>
<td>Total No. Pages (including cover)</td>
<td>48</td>
</tr>
<tr>
<td>Keywords</td>
<td>Best Practices; Innovative Solutions; Open Science, education and skills; Stakeholders Engagement.</td>
</tr>
</tbody>
</table>
DISCLAIMER

BlueBRIDGE (675680) is a Research and Innovation Action (RIA) co-funded by the European Commission under the Horizon 2020 research and innovation programme.

The goal of BlueBRIDGE, Building Research environments for fostering Innovation, Decision making, Governance and Education to support Blue growth, is to support capacity building in interdisciplinary research communities actively involved in increasing the scientific knowledge of the marine environment, its living resources, and its economy with the aim of providing a better ground for informed advice to competent authorities and to enlarge the spectrum of growth opportunities as addressed by the Blue Growth societal challenge.

This document contains information on BlueBRIDGE core activities, findings and outcomes and it may also contain contributions from distinguished experts who contribute as BlueBRIDGE Board members. 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 BlueBRIDGE Consortium and its experts, 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 the creation and publication of this document hold any sort of responsibility that might occur as a result of using its content.

The European Union (EU) was established in accordance with the Treaty on the European Union (Maastricht). There are currently 27 member states of the European Union. It is based on the European Communities and the member states’ cooperation in the fields of Common Foreign and Security Policy and Justice and Home Affairs. The five main institutions of the European Union are the European Parliament, the Council of Ministers, the European Commission, the Court of Justice, and the Court of Auditors (http://europa.eu.int/).

Copyright © The BlueBRIDGE Consortium 2015. See http://www.bluebridge-vres.eu for details on the copyright holders.

For more information on the project, its partners and contributors please see http://www.i-marine.eu/. You are permitted to copy and distribute verbatim copies of this document containing this copyright notice, but modifying this document is not allowed. You are permitted to copy this document in whole or in part into other documents if you attach the following reference to the copied elements: “Copyright © The BlueBRIDGE Consortium 2015.”

The information contained in this document represents the views of the BlueBRIDGE Consortium as of the date they are published. The BlueBRIDGE Consortium does not guarantee that any information contained herein is error-free, or up to date. THE BlueBRIDGE CONSORTIUM MAKES NO WARRANTIES, EXPRESS, IMPLIED, OR STATUTORY, BY PUBLISHING THIS DOCUMENT.
## GLOSSARY

<table>
<thead>
<tr>
<th>ABBREVIATION</th>
<th>DEFINITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>BlueBRIDGE</td>
<td>Building Research environments for fostering Innovation, Decision making, Governance and Education to support Blue growth</td>
</tr>
<tr>
<td>CNR</td>
<td>Consiglio Nazionale delle Ricerche</td>
</tr>
<tr>
<td>FAO</td>
<td>The Food and Agriculture Organisation of the United Nations</td>
</tr>
<tr>
<td>ICES</td>
<td>The International Council for the Exploration of the Sea</td>
</tr>
<tr>
<td>IRD</td>
<td>Institut de Recherche pour le Développement</td>
</tr>
<tr>
<td>UOA</td>
<td>University of Athens</td>
</tr>
<tr>
<td>FORTH</td>
<td>Foundation for Research and Technology of Hellas</td>
</tr>
<tr>
<td>VRE</td>
<td>Virtual Research Environment</td>
</tr>
</tbody>
</table>
## TABLE OF CONTENT

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOCUMENT INFORMATION</td>
<td>2</td>
</tr>
<tr>
<td>DISCLAIMER</td>
<td>3</td>
</tr>
<tr>
<td>GLOSSARY</td>
<td>4</td>
</tr>
<tr>
<td>TABLE OF CONTENT</td>
<td>5</td>
</tr>
<tr>
<td>DELIVERABLE SUMMARY</td>
<td>6</td>
</tr>
<tr>
<td>EXECUTIVE SUMMARY</td>
<td>7</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>8</td>
</tr>
<tr>
<td>2 Providing the Blue Growth sector with innovative solutions</td>
<td>10</td>
</tr>
<tr>
<td>2.1 Optimizing the use of open and sub-meter (VHR) resolution satellite data to generate an aquaculture atlas</td>
<td>10</td>
</tr>
<tr>
<td>2.2 Reporting on ecological seafloor features in marine protected area networks</td>
<td>12</td>
</tr>
<tr>
<td>2.3 Scientific Data modelling and aggregation of marine data</td>
<td>14</td>
</tr>
<tr>
<td>2.4 Publishing AquaMaps Native Habitat Data and Metadata as Exportable NetCDF Files</td>
<td>17</td>
</tr>
<tr>
<td>2.5 Assignment of unique identifiers for harmonised stock and fishery data</td>
<td>20</td>
</tr>
<tr>
<td>2.6 Data collation for the implementation of a Regional Database</td>
<td>23</td>
</tr>
<tr>
<td>2.7 How to stimulate private companies to share data by safeguarding their competitive advantage</td>
<td>26</td>
</tr>
<tr>
<td>3 Boosting Open Science, education and skills</td>
<td>29</td>
</tr>
<tr>
<td>3.1 VRE as the instrument to efficiently manage specific and tailored communities needs</td>
<td>29</td>
</tr>
<tr>
<td>3.2 Publishing software in Zenodo</td>
<td>31</td>
</tr>
<tr>
<td>3.3 Making environmental science reproducible</td>
<td>32</td>
</tr>
<tr>
<td>3.4 Virtual Research environments to support scientific trainings</td>
<td>34</td>
</tr>
<tr>
<td>3.5 Boosting Knowledge transfer</td>
<td>36</td>
</tr>
<tr>
<td>3.6 Using public eInfrastructures for commercial &amp; Operational purposes</td>
<td>38</td>
</tr>
<tr>
<td>4 Engaging key stakeholders in innovative ways</td>
<td>40</td>
</tr>
<tr>
<td>4.1 Organising Datathons to fuel new thinking and develop new skills</td>
<td>40</td>
</tr>
<tr>
<td>4.2 Boosting impact through multipliers</td>
<td>42</td>
</tr>
<tr>
<td>4.3 Leveraging call for SMEs and webinars to engage potential users</td>
<td>43</td>
</tr>
<tr>
<td>4.4 Managing a large consortium through VREs</td>
<td>45</td>
</tr>
<tr>
<td>5 Conclusions</td>
<td>47</td>
</tr>
<tr>
<td>References</td>
<td>48</td>
</tr>
</tbody>
</table>
DELIVERABLE SUMMARY

This deliverable documents the set of best practices identified within the four pillars (WP5-WP8) of BlueBRIDGE and beyond and the resulting benefits and socio economic impact. The resulting best practices are clustered in three categories: Providing the Blue Growth sector with innovative solutions; Boosting Open Science, education and skills and Engaging key stakeholders in innovative ways. The methodology for identifying best practices was provided in D3.2 Best Practices & Socio-economic Impact: Interim Report.
EXECUTIVE SUMMARY

A best practice is a practice that has proven to work well and produce good results, and is therefore recommended as a model. It is a successful experience, which has been tested and validated, in the broad sense, which has been repeated and deserves to be shared so that a greater number of people can adopt it and benefit from it. The use of the word “best” should not be considered in the superlative sense. In other words, the term “Best Practice” is not about “perfection”. Rather, it could be related to only one or more successful components of the practice under consideration.

The BlueBRIDGE best practices can be clustered in three main categories.

**Providing the Blue Growth sector with innovative solutions.** These best practices can support Blue Growth practitioners in addressing issues related (i) to monitoring, identification and location of aquaculture sites, (ii) to identify which features are important for biodiversity and ecosystem services within marine protected areas, (iii) to integrate and aggregate marine data coming from heterogeneous data sources, (iv) to stimulate private companies to share data by safeguarding their competitive advantage. In particular, they can support:

- Optimizing the use of open and sub-meter (VHR) resolution satellite data to generate an aquaculture atlas;
- Reporting on ecological seafloor features in marine protected area networks;
- Scientific Data modelling and aggregation of marine data;
- Publishing AquaMaps Native Habitat Data and Metadata as Exportable NetCDF Files;
- Assignment of unique identifiers for harmonised stock and fishery data;
- Data collation for the implementation of a Regional Database;
- How to stimulate private companies to share data by safeguarding their competitive advantage.

**Boosting Open Science, education and skills.** These best practices can support data scientists to manage tailored communities needs and to make science open and reproducible. The practices can be inspirational for educators dealing with scientific trainings. The best practices part of this category are:

- VRE as the instrument to efficiently manage specific and tailored communities needs;
- Publishing software in Zenodo;
- Making environmental science reproducible;
- Virtual Research environments to support scientific trainings;
- Boosting Knowledge transfer;
- Using public eInfrastructures for commercial & operational purposes.

**Engaging key stakeholders in innovative ways.** These best practices can support communication and training managers and coordinators working on European Projects dealing with diverse and multiple stakeholders and partners. The best practices part of this cluster are:

- Organising Datathons to fuel new thinking and develop new skills;
- Boosting impact through multipliers;
- Leveraging call for SMEs and webinars to engage potential users;
- Managing a large consortium through VREs.
INTRODUCTION

A best practice is a practice that has proven to work well and produce good results, and is therefore recommended as a model. It is a successful experience, which has been tested and validated, in the broad sense, which has been repeated and deserves to be shared so that a greater number of people can adopt it and benefit from it.

The use of the word “best” should not be considered in the superlative sense. In other words, the term “Best Practice” is not about “perfection”, rather, it could be related to only one or more successful components of the practice under consideration. Indeed, documenting and applying lessons learned on what does not work and why it does not work is also an integral part of “Best Practice”, so that the same types of mistakes can be avoided in the future.

The main rationale behind documenting and sharing “Best Practices” is to enable individuals and organizations to avoid “re-inventing the wheel”; to “learn in order to improve performance” and; to “avoid the mistakes of others”. Documenting and sharing “Best Practices” gives the opportunity to acquire knowledge about lessons learned and to continue learning about how to improve and adapt strategies and activities through feedback, reflection and analysis in order to implement larger-scale, sustained, and more effective interventions.

The aim of D3.5 Best Practices & Socio-economic Impact is to provide a summary of the best practices resulting from the BlueBRIDGE project that can be used as a reference point for Blue Growth practitioners, data scientists, and researchers dealing with the production of scientific knowledge and highlight the benefits coming from the implementation of such best practices (the aspects more related to the economic impact are documented in D2.5 Exploitation and Sustainability Plan).

In particular, the best practices that have been selected are those which:

- Have a tangible impact on end users and provide a more efficient and innovative approach with respect to common practices: Best practices have considerable and measurable, positive effects on strategic business and policy targets. Best practices provide value across actor groups and address current challenges and problems. Best practices document solutions including products, processes, services, technologies, or ideas that are more effective than previous ones and are accepted by markets, governments, and society.
- Can be adapted for replication in similar contexts: Best practice should be transferable to other companies, initiatives, contexts or domains.
- Have demonstrated sustainability: BlueBRIDGE adopters have expressed their willingness to adopt the best practices in the future.
- Bring innovative approaches to the target community/domain.
- Contribute to address the challenges of the long-term strategy of Blue Growth and Open Science.

The process on how to document best practices was described in detail in D3.2 Best Practices: Interim Report and for this reason is not reported in this deliverable.

Originally the deliverable was aimed to document the set of best practices identified within the four pillars (WP5-WP8) of BlueBRIDGE but the analysis of the best practices widened a lot the scope. The identified best practices do not only cover aspects related to Blue Growth (WP5-WP7) and education (WP8) but also, among the others, to open science, data sharing and stakeholder engagement. For this reason, the best practices have been clustered in three main categories:
1. Providing the Blue Growth sector with innovative solutions
2. Boosting Open Science, education and skills
3. Engaging key stakeholders in innovative ways

For each category, a set of best practices is described explaining why they are considered as such and the impact and applicability that they have on other sectors/domains.
2 PROVIDING THE BLUE GROWTH SECTOR WITH INNOVATIVE SOLUTIONS

The best practices belonging to this area are:

- Optimizing the use of open and sub-meter (VHR) resolution satellite data to generate an aquaculture atlas (cf. Sec. 2.1);
- Reporting on ecological seafloor features in marine protected area networks (cf. Sec. 2.2);
- Scientific Data modelling and aggregation of marine data (cf. Sec. 2.3);
- Publishing AquaMaps Native Habitat Data and Metadata as Exportable NetCDF Files (cf. Sec. 2.4);
- Assignment of unique identifiers for harmonised stock and fishery data (cf. Sec. 2.5);
- Data collation for the implementation of a Regional Database (cf. Sec. 2.6);
- How to stimulate private companies to share data by safeguarding their competitive advantage (cf. Sec. 2.7).

2.1 OPTIMIZING THE USE OF OPEN AND SUB-METER (VHR) RESOLUTION SATELLITE DATA TO GENERATE AN AQUACULTURE ATLAS

Aquaculture, and in particular marine aquaculture, is expected to grow significantly over the following decades (source FAO, OECD). This growth should be sustainable and the impact of aquaculture on the environment should then be carefully monitored. A prerequisite for this monitoring is the identification and location of aquaculture sites, particularly in coastal areas where conflict of spatial usage is frequent. BlueBRIDGE has implemented a tool to generate an Aquaculture Atlas and has tested it in three countries: Greece, Malta and Indonesia (See [https://bluebridge.d4science.org/web/aquacultureatlasgeneration](https://bluebridge.d4science.org/web/aquacultureatlasgeneration)).

For Greece and Malta use cases, the aim was to create an inventory of all fish farms using in input existing information available at FAO (NASO database), and very high resolution Earth Observation data required to detect relatively small aquaculture structures (e.g. cages with a diameter of 10 metres).

For the Indonesian use case, the aim was to provide a preliminary regional mapping of the aquaculture areas (shrimp ponds) in South Sulawesi and to discriminate them from other wet coastal areas (rice paddies).

Remote sensing was the key tool for production of an aquaculture atlas, and the best practice implemented in BlueBRIDGE is related to the optimization of the use of open and VHR satellite data.

2.1.1 THE BLUEBRIDGE BEST PRACTICE

To optimize the use of open and VHR satellite data the following approaches were implemented:

- Exploiting Copernicus and other relevant Earth Observation (EO) data: Copernicus data can be used for identification and discrimination of large aquaculture features in remote coastal areas (for example in BlueBRIDGE, Copernicus data have been used to discriminate rice paddies and shrimp ponds in Indonesia). Copernicus data can be displayed as ancillary information when the resolution is not sufficient to detect systematically small aquaculture features (this data has been used to generate the Greek Atlas).
- Avoiding duplication of facilities and data storage: The separation of the process between two infrastructures (1 - at CLS for image processing and data store and 2- BlueBRIDGE infrastructure for aquaculture product storage and further exploitation by users or other VREs) avoids duplication of image storage and processing facilities;
- Using existing, proven, validated and standard features: Both CLS and BlueBRIDGE VRE rely on existing infrastructure services that serve multiple communities. All elements aim to provide re-usable and adaptable services that can also serve other purposes. Examples include a single
webviewer for very different products, the use of infrastructure features for users and GUI management, and the use of interoperable services between CLS and BlueBRIDGE infrastructures.

The above mentioned best practice is at the heart of the Aquaculture Atlas VRE. This VRE can support both institutional stakeholders with a mission of aquaculture status assessment (at global, regional or national level) and scientific users.

### 2.1.2 WHY THIS IS CONSIDERED A BEST PRACTICE

<table>
<thead>
<tr>
<th><strong>Best Practice Analysis</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Validation</strong></td>
</tr>
<tr>
<td>The best practice was validated by analysing the results of the Atlas generated by the BlueBRIDGE VRE. The analysis of the results relied on internal validation procedures (at CLS): once approved by the CLS analysts, the results were accessed in superuser by experts designated by FAO. Once validated by the superuser, the features were made accessible to the standard user.</td>
</tr>
<tr>
<td><strong>Impact</strong></td>
</tr>
<tr>
<td>It is expected that the resulting Aquaculture Atlas VRE will contribute to updating existing databases in use (NASO at FAO), and support the development of this capacity in emerging economies who can operate the VRE.</td>
</tr>
<tr>
<td><strong>Innovation</strong></td>
</tr>
<tr>
<td>The practice contributed to an innovative way to use VHR Data (Bing or Google Earth, subject to license conditions) or Sentinel data (time series of Sentinel 1 and 2). It also equipped the service with a data management and review service, which allows the community to collaborate across the EA and GIS boundary. The images produced can be reviewed interactively.</td>
</tr>
<tr>
<td><strong>Success Factors</strong></td>
</tr>
<tr>
<td>Agreement to exploit VHR data in this context at an affordable cost, where required (e.g. Greek use case), support for capacity building actions (e.g. for Indonesian use case) to extend the services portfolio of the VRE to align with local needs (such as advanced GIS features with post-process the results).</td>
</tr>
<tr>
<td><strong>Sustainability</strong></td>
</tr>
<tr>
<td>For future operations some consideration is needed to guide the process from the research context and maturity level (proof of concept through three use cases) achieved in BlueBRIDGE, to operational provision of HR EO data to real communities. This is being refined (See D2.5) with FAO leading the development of a sustainable business model. This model will cover the MoU/agreements with commercial data providers/national space agencies or through FAO.</td>
</tr>
<tr>
<td><strong>Replicability and/or up-scaling</strong></td>
</tr>
<tr>
<td>The Greek (or Malta) use case is easy replicable in any other country of a similar size. It only requires around 8 man-days + VHR license cost.</td>
</tr>
<tr>
<td>The Indonesian use case instead requires around 30 man-days as there are less input data available, and it needs more time for algorithm tuning and validation (Sentinel use). Estimation and production model (need for other partners to be considered) to be refined for fine resolution.</td>
</tr>
</tbody>
</table>

### 2.1.3 LESSONS LEARNT

The Atlas Generation VRE, resulting from the implementation of the above described practice, meets all the community requirements, and can be used. However, for a complete analysis of the imagery, more contextual information is required such as in situ measurements, hydrographic data, and the recommendation is therefore to embed the Atlas Generation VRE in a wider spatial planning or country-atlas approach.
Maritime spatial planning is a rapidly developing area of marine management. The desire to develop the blue economy sector is one of the main drivers of this trend. Coupled with this development of this sector is the need for sustainability - with the protection of the marine environment, on which many of the blue economy sectors rely, of critical importance. Countries have committed to the protection of the marine environment with respect to number of issues. For example, through the Convention on Biological Diversity Aichi Target 11, countries have committed to the following

By 2020 [...] 10 per cent of coastal and marine areas, especially areas of particular importance for biodiversity and ecosystem services, are protected [...] 

In order for countries to understand how well they are achieving this target there is a need to identify which features important for biodiversity and ecosystem services are within their jurisdiction and what proportion of these are represented within the protected areas. BlueBRIDGE has addressed this need through implementation of the Protected Areas Impact Map (PAIM) VRE (See https://bluebridge.d4science.org/web/protectedareaimpactmaps).

2.2.1 THE BLUEBRIDGE BEST PRACTICE

The protected area impact maps VRE provides cloud-based visualization, analysis and reporting on important seafloor features in protected area networks at the country jurisdiction and ecoregion level. It has been designed to provide the following:

- Data and metadata access - data management and access is a key feature of the PAIM VRE. The VRE draws data from a number of sources to provide an integrated analysis environment. Each data set includes full metadata so that users can understand where the data comes from and its use constraints. Data is stored in Geoserver and retrieved through web services for visualization and analysis. The use of open geospatial consortium (OGS) compliant web services means that the data is accessible from both within the VRE as well as from external clients. For example, a user can download and run the analysis algorithm on their desktop computer.

- Data visualisation - the PAIM VRE uses web mapping technologies to display the spatial data in a number of contextual views. The data visualisations use the OGS web service standards to retrieve data and metadata. The MPA reporting application within the PAIM VRE provides a visual interface to drive the analysis. The Use is able to select a region for analysis from the map interface and then visualise the results both graphically and spatially.

- Analysis - analysis within the PAIM VRE utilises custom scripts developed in R-Studio. R is a powerful statistical computing language that allows both statistical and spatial analysis. The algorithm developed for the PAIM VRE is open access and hosted in Git Hub (https://github.com/grid-arendal). The use of the R scripting language makes the algorithm accessible to a wide range of users.

- Standardised reporting and visualisation of results - The final output of the PAIM VRE provides standardised reporting and visualisation of the results. All results are presented as interactive tables, charts and maps. These results can be downloaded as a pdf report, or as raw data tables for further analysis.

The Protected area impact maps VRE is targeted at marine managers, spatial planners and the scientific community. The VRE has been designed to make complex spatial analysis accessible to non-specialists thanks to the power of cloud computing and web mapping in delivering a user-friendly solution.
The PAIM VRE allows the end user to rapidly analyse the seafloor features represented within their protected area networks. This process, which would usually take a GIS specialist hours or days between compiling the data, running the analysis and collating the results can now be completed in seconds or minutes by a non-GIS specialist using the PAIM VRE. The reporting is standardised so that it is comparable at a global scale and directly applicable to country commitments on target 11 of the Aichi Convention on Biological Diversity.

### 2.2.2 WHY THIS IS CONSIDERED A BEST PRACTICE

<table>
<thead>
<tr>
<th>Validation</th>
<th>The PAIM VRE has been reviewed by the EC Joint Research Council. They have provided input as to the usability and functionality. They have also advised on the input data and processing to ensure that the PAIM VRE produces both accurate and relevant results to the end user.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Innovation</td>
<td>The PAIM VRE is being leveraged by GRID-Arendal to support capacity-building projects around maritime spatial planning in Africa and the Pacific. The VRE makes both data and analysis capacity available to a wide audience. One of the main benefits of this approach is that users do not need to download gigabytes of data to set up complex data models in specialist GIS software. This is especially an advantage in countries where the internet and computing infrastructure are limited. The PAIM VRE also reduces the need for specialist GIS software and expertise in order to understand which seafloor features are being represented in national MPA networks. This is an advantage for both marine managers, spatial planners and scientists in both developed and developing countries.</td>
</tr>
<tr>
<td>Success Factors</td>
<td>In order to achieve a high level of success for the PAIM VRE there are several requirements:</td>
</tr>
<tr>
<td></td>
<td>● Functionality for the user to upload their own MPA networks and country/region specific data layers (for example the EMODnet seafloor data) for analysis</td>
</tr>
<tr>
<td></td>
<td>● Continuous promotion amongst the community of users and feedback to improve usage of the application</td>
</tr>
<tr>
<td></td>
<td>● Continued access to the application beyond the BlueBRIDGE project lifespan</td>
</tr>
<tr>
<td>Sustainability</td>
<td>In order for the PAIM VRE to be sustainable beyond the lifetime of the BlueBRIDGE project, several steps are needed.</td>
</tr>
<tr>
<td></td>
<td>● Access to sustainable computing infrastructure (including a funding model to support that infrastructure)</td>
</tr>
<tr>
<td></td>
<td>● Updates to key data (e.g. MPA data) to keep the analysis relevant</td>
</tr>
<tr>
<td></td>
<td>● Updates to functionality, including more flexibility in the analysis and more features included in the analysis (for example inclusion of the EMODnet seafloor data)</td>
</tr>
<tr>
<td>Replicability and/or up-scaling</td>
<td>The aim is to go further than the section &quot;Innovations / critical success factors&quot; in specifying the requirements for replication of the practice on a larger scale (national, regional, international).</td>
</tr>
</tbody>
</table>

### 2.2.3 LESSONS LEARNT

The key lessons learned through the development of the PAIM VRE are related to data management and data access. The importance putting data owners in a position to provide data using OGS compliant web services cannot be understated. The use of these standards allows easy access and analysis of data from a variety of sources and circumvents the need for data to be centrally stored.
In addition, one of the major challenges in developing the PAIM VRE was related to data access: a considerable proportion of the data is not easily available or requires significant pre-analysis prior to ingestion in the VRE. The uptake of OGS web services by data custodians would greatly facilitate the ingestion of data in the VRE.

Several of the key data sources will be routinely updated, for example the global Marine Protected Area layer is regularly updated to account for newly declared protected areas. This layer also requires significant pre-processing prior to the analysis. There is a need to develop better pathways to ingest and pre-process this data. Currently the PAIM VRE relies on the efforts of EC JRC to provide a cleaned version of this data for ingestion.

### 2.3 Scientific Data Modelling and Aggregation of Marine Data

Scientific data integration and aggregation is a challenging, time consuming and error prone process. Marine Data (e.g. biodiversity data, data about Fisheries and Stocks etc.) is scattered across different and heterogeneous databases, with no standard structure, not intended to interoperate with others (database silos) and the guidelines for populating these databases are also heterogeneous.

The main challenge in BlueBRIDGE was to integrate and aggregate marine data coming from heterogeneous data sources in a semantic-rich way to build the Global Record of Stocks and Fisheries VRE.

Although the above processes were carried out in the marine domain in the context of the BlueBRIDGE project, these practices can be applied to any other domain.

#### 2.3.1 The BlueBRIDGE Best Practice

To map, integrate and aggregate marine scientific data BlueBRIDGE has adopted a six steps process:

1. **Selection of Competency Queries**: The Competency Queries express the scientific questions that the final system (Knowledge Base, Application etc.) should be able to answer. Competency queries are usually referred as query requirements. They are formulated by the potential users that have adequate knowledge of the domain, assisted by the semantic model developers and are critical for the success of the practice. The competency queries affect the selection of the semantic model and its extension, the mappings, the design of the services and even the final GUI. An example competency query can be “Give me all the stocks for the species Thunnus Albacares that are harvested by fisheries in the Atlantic Ocean”.

2. **Selection / extension of the Semantic Model**: The second step of the practice is to select or extend a semantic model that is adequate for the domain and has the expressive power to support the competency queries. This model will act as the semantic backbone and all the data will eventually be transformed according to its constraints. In the context of BlueBRIDGE a MarineTLO ontology [2] was adopted that a) provides the consistent abstractions or specifications of concepts included in all data models or ontologies of marine data sources and b) contains the necessary properties to make GRSF a coherent source of facts relating observational data to the respective spatiotemporal context.

3. **Exploit mapping technologies**: The step that follows the selection of the semantic model is the creation and application of the mappings between the model constraints and the source schema. In the context of BlueBRIDGE X3ML language was used to express the mappings. The X3ML framework [3] offers a plethora of applications. An indicative one is a 3M Editor that offers an interactive way of defining mappings and makes it relatively simple (even for users without an IT background).
4. **Transform Data into a common (semantic) format**: One basic step for integrating data from heterogeneous data sources, is to transform them into a common format. This requires the mappings (defined during the previous step). More specifically, we an X3ML engine [4] was used to receive input the sources data and the formulated mappings as input and transform the data RDF in accordance with the semantic model.

5. **Ingest data into a semantic warehouse**: The final step of the process is the actual creation of the aggregator warehouse. In the context of BlueBRIDGE, MatWare [5] was exploited. This is a tool to automatically create semantic warehouses by importing RDF files into a semantic triplestore.

6. **Assess the connectivity of the semantic warehouse**: After ingesting all the data into the semantic warehouse, it is important to inspect how connected the integrated dataset is. Connectivity is assessed in terms of connectivity metrics [6]. This measures how connected the resources of the semantic warehouse are based on metrics (such as common URIs and literals, average degree, etc.). These metrics can be computed automatically after the construction (or refresh) of the semantic warehouse, and also enable errors and redundant data sources to be spotted.

The “key” features of the aforementioned steps that guarantee the efficient and aggregation of heterogeneous marine data are the following:

- The exploitation of mappings that separates the work of domain experts (that actually define the mappings), and the IT people (that define how the information will be transformed). This decoupling of roles eliminates the bottleneck that usually appear when integrating data.
- The assessment of the connectivity of the integrated data which is of paramount importance and guarantees that they the data from the original sources are properly connected and also spot errors (i.e. if there are or irrelevant data sources).
- The automation of the entire process that allows reconstructing from scratch or refreshing particular parts of the integrated data set, by relying on innovative technical components (i.e. MatWare).

### 2.3.2 WHY THIS IS CONSIDERED A BEST PRACTICE

**Best Practice Analysis**

| Validation | The output of the above-described marine data modelling and aggregation best practice is the Global Record of Stocks and Fisheries VRE (GRSFVRE). The GRSF VFRE and therefore the full modelling and aggregation process has been validated by both FAO (FIRMS Database), University of Washington (RAM Database) and SFP (FishSource Database). The feedback was positive as the users (experts) were able to report that:
| 1) They are able to answer queries that could not answer before
| 2) They are able to browse through and discover complete sets of data about stocks and fisheries that could not be performed before
| 3) The data is of better quality, providing them with “hints” as to how to improve their own database. |
| Innovation | The best practice contributed to delivering the (first) Global Record of Stocks and Fisheries. It enabled scientists to adopt it as a methodology to improve the final result with the overall aim of delivering a high-quality product. On this aspect, when a draft version of the registry was available, scientists adapted the competency queries to alleviate issues and capture more concepts, which in turn triggered all the subsequent steps towards re-constructing the registry. |
As a result, the final version of the registry can be considered as a knowledge base containing a coherent set of facts on stocks and fisheries that can be used to carry out advanced stock and fishery assessment activities.

**Success Factors**

The main factors to guarantee that the practice was successful are the following:

1. The practice has been adopted for constructing several versions of the Global Registry of Stocks and Fisheries.
2. It has been presented and approved in three different technical working groups organized by FIRMS (FAO division) with the participation of different marine stakeholders.
3. It has been published and presented in a scientific conference [7].

**Sustainability**

The activities for curating and maintaining the integrated dataset is an activity that without cease. The semantic warehouse relies on data coming from different data sources, and, as such, it might require adaptations as data from the source changes. These adaptations regard changes in the mappings, as well as changes in the algorithms that integrate the data.

In addition, competency queries might change. As new user requirements emerge, mappings and algorithms that integrate data need to be updated to capture them.

The aforementioned activities should be carried out by a person familiar with mapping technologies and software. Although we cannot estimate the potential effort and cost to carry this out (clearly this depends on the complexity and the number of new requests), there is confidence that the benefits to be obtained will be great.

**Replicability and/or up-scaling**

The proposed practice can be used by:

a) Institutions that perform data mapping or data aggregation activities;

b) Companies that exploit integrated data for a profit (e.g. companies exploiting marine resources)

c) Non-profit organizations that use scientific data for decision making and prediction of potential disastrous scenarios (i.e. FAO for predicting the depletion of a marine resource)

The organisations involved in the context of the BlueBRIDGE project were:

- FORTH who lead the technical activities for the designing and implementation of services integrating marine data.
- FAO who provided data, validated the results and assisted in delineating the policies.
- Sustainable Fisheries Partnership who provided FishSource data and validated the results.
- CNR who assisted the front-end development activities.
- RAM Legacy Stocks Assessment Database who provided data.

Although, the best practice was tailored to integrate data for the marine domain, it should be clarified that it can be applied to any other domain, with little or no modifications (i.e. only the semantic model could be updated).

### 2.3.3 LESSONS LEARNT

The collation of information is both difficult and time-consuming, as the information is scattered across different databases and is modelled using different formats and standards. The proposed best practice
defines the steps that are required to integrate data from different sources, allowing it to be exploited for to respond to queries that could not be answered before (as they required combining knowledge from heterogeneous data sources), and assess the connectivity of the resulting semantic warehouse. Thanks to this practice, scientists can easily discover big and complete datasets to assist them on their research, investors have central access to rich information to assist them in their decisions to invest in a specific field (e.g. invest on harvesting tuna in the North Atlantic), industry can acquire and analyse data to predict the sustainability of their businesses, and improve their profits and simple users can gain valid information in their domain of interest.

The main lessons learned during the application of the best practice in the context of BlueBRIDGE were:

1) Collaboration with the providers of the data sources is critical. Without their support in understanding the semantics, the clarifications on their data and the validation of the mappings, it is extremely difficult attain successful aggregation.

2) The use of a central semantic model is mandatory to overcome heterogeneity that exist among the schemata of the original data sources. Furthermore, it removes the complexity of updating technical components when the schemata of the original sources change (in this case only the mappings should be updated and all the technical components remain unchanged).

3) Collection of the query requirements (the competency queries) should always be the first step. A wrong estimation of the querying capabilities of the resulting system can lead to a completely wrong design incapable of serving the query needs of the marine community.

2.4 PUBLISHING AQUAMAPS NATIVE HABITAT DATA AND METADATA AS EXPORTABLE NETCDF FILES

Communities interested in Niche Modeling require structured data on species distributions. Distributions are often published as text, image files, or vector data not suitable for the kind of processing scientists usually perform, while a portable self-describing format such as the NetCDF enables faster access to data and metadata, as well the creation of faster ecological models. For this reason, BlueBRIDGE decided to reduce the inertia communities encountered when dealing with geospatial data for ecological modeling purposes, significantly lowering the data preparation time.

In particular, the goal of this activity was to create a collection of AquaMaps Native Range layers in a NetCDF format, while defining a standard general procedure to create such repositories of raster maps starting from other representation formats. It is worth pointing out that it is not a mere conversion, but implies data value enhancement.

2.4.1 THE BLUEBRIDGE BEST PRACTICE

The workflow that has been implemented to achieve the above mentioned objectives can be summarized in 4 general steps:

1) Programatically download all the AquaMaps Native layers from GeoServer in CSV format via WFS requests

2) Convert every polygon map into a latitude-longitude format, using the centre of mass for every square
3) Convert the CSV into a NetCDF using an automatic procedure, in our case a DataMiner process referenced in the “URL of the practice” section of this table.

4) Programmatically upload the new NetCDF files onto Threads and enrich them with metadata registered on GeoNetwork. This step was also carried out using a DataMiner process referenced in the “URL of the practice” section of this table.

Before BlueBRIDGE, the species distribution information was mainly published in a text and image format, forcing the communities to pre-process and prepare the data before being able to actually use them for their purposes. Moreover, proprietary and textual formats are not easily portable and do not allow for metadata embedding. After BlueBRIDGE, much of this data will be available in a more portable format and the workflow to produce it can be reused for other sets of data. A result of this activity has also been a process to convert a generic CSV file into a NetCDF file, made available as a service accessible from DataMiner. While a manual conversion of this kind would require deep knowledge of the NetCDF format itself, this algorithm is easy to use and masks all the complexity of such translation.

Many stakeholders can benefit from the outputs of this workflow as well as in the workflow itself, which can be easily generalized. Specifically, all the communities involved in BlueBRIDGE, the AquaMaps and the Fishbase communities, and the research groups operating in Ecological Niche Modeling and Fisheries in general. In fact, millions of people in these communities access and download species distribution data every month, for taxonomic, ecological, and fisheries research. Having the information they are searching for in a standard portable format such as the NetCDF allows them to skip the data pre-processing phase and obviously enhance their productivity.

### 2.4.2 WHY THIS IS CONSIDERED A BEST PRACTICE

<table>
<thead>
<tr>
<th><strong>Best Practice Analysis</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Validation</strong></td>
</tr>
<tr>
<td><strong>Innovation</strong></td>
</tr>
<tr>
<td><strong>Success Factors</strong></td>
</tr>
<tr>
<td><strong>Sustainability</strong></td>
</tr>
</tbody>
</table>
| **Replicability and/or up-scaling** | The whole procedure can be easily reproduced and the scalability depends on many factors:  
  1. The optimization of the conversion algorithm  
  2. The computation platform used to perform the conversion. In this case the DataMiner guaranteed the required scalability  
  3. The ability of GeoServer and Thredds to process multiple download/upload requests  
  4. The ability of GeoNetwork to process multiple metadata creation requests |

### 2.4.3 LESSONS LEARNT
The conversion of a certain probability distribution into a raster dataset, as well as being pretty challenging and requiring a profound knowledge of the data formats, is also demanding in terms of computational power, thus making a scalable architecture an almost mandatory choice. The validation process is also fundamental and should be carried out using an automatic procedure. As well as defining a very straightforward portable services workflow to accomplish a polygon-to-raster layer conversion, provides the communities with a wide collection of layers. The NetCDF is much more reusable and portable than the other formats already supported by WFS and is naturally suited to represent raster information (like uniformly spaced species distributions). In fact, on top of being self-describing and designed to represent n-dimensional data with n >= 2, it is also widely used as a standard and there are plenty of tools to visualize and manipulate this format. Moreover, additional information on the data can be included in the file itself as attributes, creating more complex objects without the need for any external references to be fully understandable, and thus reusable and portable.

This activity facilitates the usability of species distribution maps and lays the foundations for collaboration with the AquaMaps and Fishbase communities, as well as any other group interested in global habitat distributions. It also simplifies use of the information and the metadata recovery process, as the metadata is directly attached to the data.

A scientific publication has been produced and published [8]. Moreover, several converters have been released (only the first one has been used for this activity, but the others may be used for similar tasks):

- **CSV_TO_NETCDF_CONVERTER_XY**: A process to convert a generic CSV file into a basic NetCDF one with 2 dimensions (latitude, longitude)
  

- **CSV_TO_NETCDF_CONVERTER_XYT**: A process to convert a generic CSV file into a basic NetCDF one having 3 dimensions (latitude, longitude, time)
  

- **CSV_TO_NETCDF_CONVERTER_XYZ**: A process to convert a generic CSV file into a basic NetCDF one having 3 dimensions (latitude, longitude, altitude/depth)
  

- **CSV_TO_NETCDF_CONVERTER_XYZT**: A process to convert a generic CSV file into a basic NetCDF one having 4 dimensions (latitude, longitude, altitude/depth, time)
  

- **CSV_TO_NETCDF_CONVERTER_DIMCHAR**: A process to convert a generic CSV file into a basic NetCDF one having a single string-type dimension
  

Link to the metadata publication process:
The global fisheries community needs reliable information that identifies where stocks live, and which fisheries target them. This data is essential for management and trade, e.g. to guarantee the provenance of sustainable marine products. Until recently, data was collected by different teams with different objectives, using different formats and references, and a global integration in this sense was long overdue.

2.5.1 THE BLUEBRIDGE BEST PRACTICE

BlueBRIDGE has developed the Global Record of Stock and Fisheries (GRSF) VRE to address the need to have a global integrated knowledge base on stocks and fisheries; an essential asset to uniquely identify the provenance of fish and associated data. This is achieved by: i) collating records from either national or regional sources, ii) organising and storing information according to specific data standards and protocols which allow comparability between records and consistency of the database, and iii) assigning and publishing Universal Unique Identifiers for single stock and fishery standard identifications.

The GRSF VRE specialises in the harmonization, visualization and analysis of the existing data coming from three existing databases (Fisheries and Resources Monitoring System (FIRMS)\(^1\); RAM Legacy Stock Assessment Database\(^2\) and FishSource\(^3\)), a process that is guided by the CWP (the FAO Coordinating Working Party on Fishery Statistics) as concerns the standardization of reference data on species, areas, countries, gears and other classifications. Unifying information requires careful attention to data policies and all the three contributing parties can now endorse all the information they share in the knowledge base. For the integration work, technological challenges in mapping and managing the content, were solved with semantic technologies feeding a CKAN registry.

Overall, the assignment of unique identifiers for harmonized stock data is the best practice sought. To achieve this, the VRE applies the FAIR principles:

- (meta)data are assigned with a globally unique and eternally persistent identifier. These UID are also available as ‘human readable’ identifiers, to facilitate the use and acceptance of the GRSF records. In addition, the harmonization process relies on public codelists and reference data (e.g. from FAO of the UN, Area and country codes, ASFIS and WoRMS species codes) all under their own UUIDs, thus implementing this FAIR principle.
- data are described with rich metadata; the GRSF VRE is designed to merge data from 3 systems, and especially focused on data workflows; at each step in this workflow metadata is collected; from the registration of new data (Provenance), its Harmonization (Process), the terms of use (License and validity), and publication (Citation and access points).
- (meta)data are registered or indexed in a searchable resource; all data are stored in a semantic knowledge base that is harvested automatically by a CKAN registry (Accessible through the D4Science platform), and where applicable.
- metadata specify the data identifier.
- (meta)data are retrievable by their identifier using a standardized communications protocol. The GRSF records are published in a CKAN registry, and the standardized protocols of the registry govern the access not only to GRSF records, but to other information sources as well.

\(^1\) http://firms.fao.org
\(^2\) http://ramlegacy.org
\(^3\) http://www.fishsource.com
• (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation. The GRSF enables interoperability between currently separate systems by ingesting them, and then storing representations of the information in a semantic knowledge base.
• (meta)data use vocabularies that follow FAIR principles. Under the aegis of the CWP, the fisheries domain is developing Master Data Management Services (MDM) for all domain relevant codes and reference data. The contributing parties have committed to expose their data under these guidelines and principles.
• (meta)data include qualified references to other (meta)data. These GRSF records use reference links to identify both content (reference data) and context (the business metadata).
• (meta)data have a plurality of accurate and relevant attributes. All data passing through this VRE are enriched with relevant metadata, and are even dependent on these for their life-cycle.
• (meta)data are released with a clear and accessible data usage license. The terms of use of BlueBRIDGE are clear on the license, and this is added to most datasets.
• (meta)data are associated with their provenance. BlueBRIDGE ensures that the GRSF record source is well described and contains links and reference to data contributors source page and their citation.
• (meta)data meet domain-relevant community standards. The development of the VRE was driven by community (Including the FAO CWP on statistical reference data) to ensure the pervasive and correct use of community standards.

The main beneficiaries of this practice are data managers in the fisheries domain and fisheries policy makers. In addition, the need for unequivocal stock exists in trade and industry, especially for traceability and sustainable market initiatives.

<table>
<thead>
<tr>
<th>2.5.2 WHY THIS IS CONSIDERED A BEST PRACTICE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Best Practice Analysis</strong></td>
</tr>
<tr>
<td><strong>Validation</strong></td>
</tr>
<tr>
<td>The services provided via GRSF were tested and validated by VRE Data managers using a FURPS approach (Functionality – Usability – Reliability – Performance – Security). After the FURPS validation of the technical readiness of the services, the data were further validated by experts in the sector using a dedicated VRE (GRSF Admin) as a data preparation and staging area. Only qualified records are released to the public.</td>
</tr>
<tr>
<td><strong>Innovation</strong></td>
</tr>
<tr>
<td>The GRSF unique identifiers enable the inclusion of stock and fishery data across organizations in workflows related to fisheries management; this can drive innovation in the traceability of marine products. At technology level, the availability of GRSF services will enable the development of similar services for data associated with the GRSF, such as unique identifiers to capture data and model output identifiers. This discussion has already been initiated.</td>
</tr>
<tr>
<td><strong>Success Factors</strong></td>
</tr>
<tr>
<td>The GRSF has to prove at the level of stock records that it can offer a cost-effective way to merge datasets and enrich rather than replace the data sources with additional information. The GRSF provides unique identifiers across systems, and thus requires that a community sees the value of these over their own systems. There is therefore a need for an institutional awareness regarding data sharing policies which must go beyond the single organization’s policy, and this requires long-term vision. The GRSF overcomes this with a community-driven approach and by putting a well-connected governance team in place with weekly or monthly e-meetings, with outcomes documented in the GRSF Wiki. The e-meetings were complemented by three Technical Working Group meetings were stakeholders had the opportunity to discuss face-to-face the key topics of the GRSF development and its sustainability. Similar initiatives can learn from the GRSF how to establish the community, identify the shared need, and how to work towards a sustainable business model.</td>
</tr>
</tbody>
</table>
10 reasons make the Global Record of Stocks and Fisheries (GRSF), a key instrument of global fish stocks status monitoring and traceability

- A comprehensive global information repository on fish stocks and fisheries
- Collation of independent repositories with complementary data coverages (3 are in already)
- Information harmonized through international and dedicated standards
- Assignment and publishing of unique identifiers to stock records for easy information management, retrieving, and interoperability
- Information managed and displayed in a powerful and effective web-based collaborative environment allowing multiple contributions in a data validation workflow
- Updates and incremental contributions facilitated
- Transparency valued through source and ownership metadata, proper citations and credits, and respect of ownerships
- Contribution to international initiatives such as the U.N. Sustainable Development Goals (14, 17, ...), the FAO State of the World Fisheries and Aquaculture (SOFIA), ...
- Users accessing a huge amount of quality data collected in a cost-effective way with distributed effort among authoritative sources
- Authoritative source of information for traceability and certification in the seafood industry

| Sustainability | Sustainability will have to start with a good value proposition that can be developed into a business plan. This is easiest done when there are clear layers of responsibility between infra providers, infra services developers, and communities. A governance model that can develop matches between communities and services providers is essential. The sustainability of the GRSF was discussed at the last Steering Committee (June 2017, FSC10) of the Fisheries Resources Monitoring System (FIRMS) Partnership between Regional Fisheries Bodies and FAO. It was decided that the Partnership should work towards taking the ownership of GRSF which can help achieve its goal: “facilitate the monitoring of the status and trends of all fishery resources.” GRSF can be the tool, together with related collaborations, to disseminate the information on national stocks monitored by countries under SDG14.4.1, to contribute to SOFIA indicator on Stock status, to serve traceability needs, etc., and promote their efforts to render their fisheries sustainable to the general public and their own benefit. This would proceed through expanding and strengthening the current FIRMS Partnership to bodies, countries and resource partners willing to contribute to GRSF, and the FIRMS goal in general. The third GRSF technical working group meeting (Feb. 2018) jointly held with FIRMS partners, and FAO, Univ. of Washington and SFP representatives discussed the value proposition, a governance model and a business plan to be submitted to FIRMS Steering Committee for its possible final adoption. |
| Replicability and/or up-scaling | The partnership approach and the focus on data policy offer a possibility to extend the practice beyond the current scope. The data needs are clearly stated, and clear objectives have to be agreed upon for data sharing, focussing on what can be obtained and not only what needs to be given. Since the GRSF is already on a global scale, it applies global standards. These must be flexible enough to allow (local) extensions and... |
filters. In addition, different scales (in spatial resolutions and time) need to be manageable, calling for rather open data structures over well-defined databases. The GRSF implements a vision that can either grow in details (by adding content) or can be replicated to other domains (agriculture, shipping, etc) where global repositories need to be mapped.

2.5.3 LESSONS LEARNT

Collaboration on key objectives is fundamental. A community-driven approach was needed for a proper formulation of requirements for offering key services to: i) Stakeholders involved in global/regional/national state of stocks indicators, and ii) Public and private actors involved in eco-labelling, traceability and sustainable fisheries.

The data in this VRE is better described than the original fragmented and dispersed datasets; some contributors adapted new formats and reference data in their own system. The GRSF data now has more access methods, and re-creating a dataset for sharing has become less time consuming. Immediate impacts have been that stock and fishery records are now uniquely identified and available under one umbrella, represented on maps and a global search is enabled across all available sources. Additional services can be built on top of the GRSF knowledge base to further support stock and fishery status dissemination and monitoring.

2.6 DATA COLLATION FOR THE IMPLEMENTATION OF A REGIONAL DATABASE

Regional data, statistics and information are key assets to support evidence based policies making, especially to develop and monitor regional fisheries management plans (FMP) such as the regional WECAFC FMP for the flying fish, the queen conch and the Caribbean spiny lobster. The challenges in regional databases are the national sources of fisheries data and statistics. Collection of data, processing of statistics and information are carried out by national institutions with national focus on countries’ communities like small scale fishers to develop food security policies or larger scale, industrial fisheries to develop the economy while exploiting the fish resource in a sustainable way. It is difficult to build a system that informs where species live, what fisheries target them, and how much effort is deployed to catch the fish. These data are essential for management, e.g. to guarantee the sustainability of marine products. Data can be collected by different national institutions with different objectives, using different formats and references, and a regional integration effort that can cope with different data flows and formats is needed.

2.6.1 THE BLUEBRIDGE BEST PRACTICE

BlueBRIDGE has developed the RDB features in the WECAFC-FIRMS VRE to address the need for support for a regional (i.e. across selected countries) database for selected fisheries data and models (most fish stocks of high values are shared between countries across a region, especially in the Western Central Atlantic Fisheries Commission (WECAFC) area).

The RDB approach implements data harmonization, storage, visualization and analysis of the regional fisheries data, a process that is guided by data standards for fisheries (CWP⁴), storage and harmonization (SDMX⁵ and OGC⁶) where it concerns the standardization of reference data on species, areas, countries, gears and other classification, and for visualization and mapping (OGC).

Given the often sensitive nature of the data, careful attention to data policies and data confidentiality is needed, and contributing parties need a stage to endorse any information they share.

---

⁴ Coordinating Working Party on Fishery Statistics
⁵ Statistical Data and Metadata eXchange
⁶ Open Geospatial Consortium
These are the approaches adopted in the implementation of the VRE:

- (meta)data are assigned a globally unique and eternally persistent identifier. All data are assigned a UID on entry to the system. However, over the lifetime of the data a UUID is not always relevant, and so a persistent identifier is only applied for published data. In addition, the harmonization process relies on public codelists and reference data (e.g. from FAO of the UN, Area and country codes, ASFIS and WoRMS species codes) all under their own UUIDs, thus implementing this FAIR principle.

- data are described with rich metadata; the RDB VRE is designed to merge data from countries, and to publish these as SDMX, a format focused on data descriptions in the statistical domain. At each step in this, WF metadata is collected; from the registration of new data (Provenance), its Harmonization (Process), the terms of use (License and validity), and publication (Citation and access points).

- (meta)data are registered or indexed in a searchable resource; all data are stored in a specialized instance of the Tabular Manager service and published in a SDMX registry (Fusion).

- metadata specify the data identifier.

- (meta)data are retrievable by their identifier using a standardized communications protocol. The reference data are published in a SDMX registry, and potentially in a CKAN registry. This also contains a dataset identifier.

- (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation. The RDB data are published in a SDMX fusion registry that enables interoperability between SDMX capable systems. Derived products such as model output cannot be shared, and no language for knowledge representation can be applied.

- (meta)data include qualified references to other (meta)data. The RDB datasets contain reference links to identify both content (reference data) and context (the business metadata).

- meta(data) have a plurality of accurate and relevant attributes. All data passing through this VRE are enriched with relevant metadata (automatically), and the RDB aims at replacing local classifications with global ones to increase global relevance.

- (meta)data are released with a clear and accessible data usage license. The terms of use of BlueBRIDGE are clear on the license, and this is added to most datasets. Users in the VRE are not always authorized to share confidential country data, and a private license can be applied.

- (meta)data are associated with their provenance. BlueBRIDGE ensures that the RDB Data are well described, and contain links and reference to data contributors name.

- (meta)data meet domain-relevant community standards. The development of the VRE was driven by community (Including the FAO CWP on statistical reference data) to ensure the pervasive and correct use of community standards. The mapping between local and regional classification is key in this VRE.

The primary group of stakeholders for this VRE are data managers in the regional fisheries domain and fisheries policy makers. They need a toolset to collate and analyse their data, and if possible report and visualize on the data. The secondary group of stakeholders are national data managers that will benefit from the standardization and harmonization process to identify gaps/improve quality of their collected data to match regional needs; fisheries management and stock assessments officers/experts will benefit from the regional data, statistics and information.

- (meta)data meet domain-relevant community standards. The development of the VRE was driven by community (Including the FAO CWP on statistical reference data) to ensure the pervasive and correct use of community standards. The mapping between local and regional classification is key in this VRE.

The primary group of stakeholders for this VRE are data managers in the regional fisheries domain and fisheries policy makers. They need a toolset to collate and analyse their data, and if possible report and visualize on the data. The secondary group of stakeholders are national data managers that will benefit from the standardization and harmonization process to identify gaps/improve quality of their collected data to match regional needs; fisheries management and stock assessments officers/experts will benefit from the regional data, statistics and information.
### 2.6.2 WHY THIS IS CONSIDERED A BEST PRACTICE

**Best Practice Analysis**

<table>
<thead>
<tr>
<th>Validation</th>
<th>The need for a RDB VRE has been identified in different functional domains such as trade, tourism, agriculture, tourism to support regional policy making. The need in the WECAFC region is also well identified for fisheries as certain species require already similar collation processes (tuna). The need to extend to all species has been validated through different workshops in the WECAFC region. Once the actual services are delivered by development teams, they are first tested and validated by VRE Data managers using a FURPS approach (Functionality – Usability – Reliability – Performance – Security) with data requested from national data managers through different channels. The VRE level tests are done after the software test, and thus only capture the VRE relevant comments. After passing this test, validation of this VRE is mostly done in workshops with real data users (such as the coming first meeting of the WECAFC Fisheries Data and Statistics Working Group planned for early 2018) and training events (E.g. WECAFC training events).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Innovation</td>
<td>At this stage, the collation of regional datasets can be linked to new methods for stock data analysis across organizations; this can drive innovation in the assessment of marine resources. At technology level, the availability of RDB services that are quite generic for data collation will enable the development a stable data platform for fisheries data collation and analysis.</td>
</tr>
<tr>
<td>Success Factors</td>
<td>The RDB has to prove at the level of data collation that it can offer a flexible and cost-effective way to merge datasets, and publish these in a managed repository. The RDB meets the conditions for the institutional quality requirements to share data, but requires that a regional organization accepts the responsibilities. The RDB VRE facilitates this by offering a well-connected governance team and model. Similar initiatives can learn from RDB how to establish the community, identify the shared need, and how to work towards a sustainable business model that meets the needs of statistical data reporting.</td>
</tr>
<tr>
<td>Sustainability</td>
<td>The RDB VRE is in essence a tool to provide sustainable access to national and regional datasets. If data losses occur at national level, data are still available at regional level. The key is to provide sustainable access to the VRE. This is easiest done when there are clear layers of responsibility between infra providers, infra services developers, and communities. A governance model that can develop matches between communities and services providers is essential. SLA or MoU should be discussed at regional level with the WECAFC secretariat and the new Fisheries Data and Statistics Working Group, most likely to act as a steering committee to the RDB. Once identified, maintenance cost could be shared across the regional and sub-regional organizations that will benefit from the RDB such as the Caribbean Regional Fisheries Mechanism (CRFM), OSPESCA and WECAFC.</td>
</tr>
<tr>
<td>Replicability and/or up-scaling</td>
<td>The RDB implements a complete data-driven workflow for the collation and publication of statistical datasets. Any organization in need of merging localized systems data output can benefit from the approach. The current RDB is agnostic of the data structure, relies on complete description of data structure and code lists (metadata) and thus can be applied to other domains with similar needs to aggregate and harmonize national datasets into regional databases. Regional Economic Commissions such as SADC (South African Development Community), EAC (East-African Commission) could benefit from this approach for agriculture, forestry, health, tourism etc., RDB focused on manual entry of larger datasets, and is thus best applied at data collation level.</td>
</tr>
</tbody>
</table>

### 2.6.3 LESSONS LEARNT

Collaboration on key objectives is fundamental. Raising awareness on the key principles underlying the implementation of a RDB is underpins its success: definition of minimum data requirements to define
standard information to be collated from countries, definition of regional classification and mapping with national ones to ensure harmonization of data.

Two profiles of VRE end-users that can benefit from the VRE can be identified: the national data manager that contributes to the regional data collation process and the fisheries expert that uses the collated data. For the national data manager, the need to submit harmonized data has a direct impact on the quality of data (some data might be missing, might too fragmented; national data collection can be improved to match the regional data requirement – national classification can be reviewed to propose more stable lists especially on species list) and on the quantity of data (Effort is deployed by the national data manager to use/exploit/analyse data piled up in the country’s drawers). For the fisheries expert, access to the original fragmented and dispersed datasets is easier: lots of data can be available at a national level but on paper or in excel files stored in local computers. The RDB allows an overview of stocks across country boundaries, and to apply models on stocks shared across countries, something that is very time-consuming to prepare. These users can now apply more methods, and discuss the results in a regional setting. This significantly helps regional discussions.

However, there is a clear difficulty in establishing stable communities in developing countries that can provide useful data. Turn-over of trained staff is high in the WECAFC region, either because of competition with the private sector or because of national budget cuts (impact of oil price drop on Trinidad and Tobago, Venezuela incomes). There is also resistance to replacement of standing practices that are difficult to change, such as data formats, as this may require changes in national legislation, which takes time. Since RDB maps from existing formats to a global / regional one, it should be easier for organizations to use this system.

### 2.7 HOW TO STIMULATE PRIVATE COMPANIES TO SHARE DATA BY SAFEGUARDING THEIR COMPETITIVE ADVANTAGE

One of the major objectives of BlueBRIDGE was to improve the efficiency and support the growth of aquafarms by providing them with information on benchmarking, currently unavailable for the sector. This is a completely new perspective in a domain where stakeholders operate individually, without the ability to comprehend the details of their business performance w.r.t. their competition.

To achieve this, BlueBRIDGE has built tools to collect data from the companies and to transform it into valuable benchmarking information. One of the major issues in this process, was to convince aquafarms to share their own production/operation data in order to create a substantial volume of high quality data to be transformed into benchmarking information. A factor that needs to be taken into account is that those production/operation data is sensitive information and is one of the assets of the company. However, for the approach set by BlueBRIDGE, the more, in quantity and accuracy, data collected, the better for the coverage, accuracy and privacy of the analysis performed by the tools of the project. In essence SMEs are invited to share their data in exchange of better quality of analysis of their performance, comprehending of the competition trends and more advanced tools for their executive skill building.

#### 2.7.1 THE BLUEBRIDGE BEST PRACTICE

To overcome the confidentiality issue and encourage companies to share data, BlueBRIDGE has defined the following set of policies and tools:

- **(a)** data anonymization to prevent exposure of installation-specific production and performance information to competitors and
- **(b)** Removal of location information in favour of location characteristics.
- **(c)** Indirect exposure of outputs through simulation models, rather than direct access to production/operation data.
- **(d)** Delivery of comparative results instead of absolute performance estimation accuracy.
(e) Establishment of thresholds: Unless a threshold of data availability is reached, particular services are not available for a given set of system parameters (e.g. location characteristics), in order to preserve anonymity;
(f) Return data with improved quality.

The adoption of this practice allows private stakeholders to gain information on benchmarking without the fear of revealing their competitive advantage when sharing confidential data. This clearly brings a wide range of benefits to the different stakeholders:

- Aquafarming production SMEs get additional data to understand their own performance, leading to better informed decisions that will boost efficiency and financial results.
- Aquafarming consultants get the means to perform better analyses for their target clients.
- Regulators and policy makers take more informed decisions on policies and rules to be applied to the sectors.
- Strategic investors have better chances of comprehending the competitive advantages of one way/location of operating a business vs another one.

2.7.2 WHY THIS IS CONSIDERED A BEST PRACTICE

<table>
<thead>
<tr>
<th>Best Practice Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validation</td>
</tr>
<tr>
<td>Anonymization techniques are the primary means for protecting sensitive data in the literature and in all cloud applications where such sharing of information applies. Furthermore, the adoption of such techniques allowed BlueBRIDGE to overcome the barriers preventing SMEs from using the platform. So far 10 SMEs are providing data to feed the BlueBRIDGE benchmarking tools.</td>
</tr>
<tr>
<td>Innovation</td>
</tr>
<tr>
<td>Innovation of the practice resides in the new data sharing business model, as well as on the tools that emerged previously and were not implementable due to the lack of data. Engaged SMEs also enter a new mentality w.r.t to cloud infrastructures and information sharing.</td>
</tr>
<tr>
<td>Success Factors</td>
</tr>
<tr>
<td>The success of the practice depends on the following factors.</td>
</tr>
<tr>
<td>- Flexible SMEs that are open to exchange and share data needs to be involved in the process.</td>
</tr>
<tr>
<td>- The infrastructure needs to ensure data confidentiality.</td>
</tr>
<tr>
<td>- A significant amount of data sets need to be continuously fed into the infrastructure.</td>
</tr>
<tr>
<td>- The companies need to be provided with added value information.</td>
</tr>
<tr>
<td>Sustainability</td>
</tr>
<tr>
<td>The practice requires that the infrastructure provider applies measures to evaluate the performance of the infrastructure w.r.t. effectiveness of the confidentiality measures taken. This is not a once-off action (i.e. at the time of definition of the techniques and processes around data confidentiality preservation) but needs to be a continuous activity enumerating new threats and measures as technology and comprehension of data involves.</td>
</tr>
<tr>
<td>Replicability and/or up-scaling</td>
</tr>
<tr>
<td>The practice applies to all domains where primary data would be of value to analysts and where performance benchmarking would be valid. Agriculture production, production and supply lines, sales and services lines, fisheries etc. could be domains to replicate the strategy.</td>
</tr>
<tr>
<td>Upscaling increases the value of the data gathered and reduces the barriers for adoption.</td>
</tr>
</tbody>
</table>
2.7.3 LESSONS LEARNT

Operational information sharing can enhance informed decision making at all levels of a sector (operational, strategic, regulatory).

Stimulating data provisioning in return for advanced or higher quality services seems to be a valid assumption however no rule-of-thumb can be applied regarding achievement and sustainability, as the success of the endeavour would heavily depend on the landscape and maturity of tools in the domain, the mentality of the stakeholders and the innovation of ideas and services delivered in return.

Addressing confidentiality on data that could potentially reveal competitive advantages is most certainly a key element and it is a domain-specific task that cannot be applied without good knowledge of the business domain, its reservations and the data processes engaged. Data anonymization is the key to unlocking confidential data sharing among aquafarming stakeholders and a good communication of confidentiality measures can assuage data owners of their concern in this respect.
3.1 VRE AS THE INSTRUMENT TO EFFICIENTLY MANAGE SPECIFIC AND TAILORED COMMUNITIES NEEDS

“Spontaneous” communities are dynamically aggregating practitioners sharing an interest for a research activity or, more generally, a task to be performed in a collaborative way. Providing these communities with working environments enabling effective work to be carried out in a shared interest is challenging due to several factors, with cost the most critical. Setting up a working environment from scratch is time and effort consuming; it requires collation of heterogeneous resources (data, computing, services) usually spread across several providers, and complementing this interest-specific set of resources with collaboration-oriented facilities (e.g. a shared area for exchanging material, a communication area for discussions) and management-oriented facilities (e.g. managing users, monitoring service availability).

3.1.1 THE BLUEBRIDGE BEST PRACTICE

To enact the creation and management of web-based working environments for “spontaneous” communities BlueBRIDGE relies on an underlying infrastructure (D4Science.org) that put a wizard-based mechanism for creating dedicated Virtual Research Environments in place. Such environments are web-based, built by aggregating the resources required to serve the needs arising for a specific context of application. The aggregated resources cover the entire spectrum of components requested to implement the envisaged working environment. These range from the machines needed to operate the services, to the data to be made them easily available, the services themselves, and the GUI components to enable easy access and use of aggregated resources. The process is started by a request for a VRE: a user on behalf of a spontaneous community can create a ticket with the request for a VRE (including a brief description). The managers of the underlying infrastructure analyse the request and, if the proposal is suitable, they provide the requesting party with a VRE Manager role. By using a wizard, the VRE manager (i.e. the actor requesting a VRE) is allowed to characterise the environment s/he needs in terms of data and facilities by selecting potential assets from a catalogue. Once the specification of the VRE is produced, it is automatically transformed into a deployment plan consisting of a set of services to be deployed and configured to operate according to the expectations of the VRE Manager. Once this deployment and configuration phase is complete, the VRE Manager is provided with a ready-to-use web-based application realizing the envisaged VRE. In some cases, the deployment of the VRE requires the development of new components. This development is performed in the context of the VRE itself so that the component can be almost immediately plugged in to the expected working environment and interact with the rest of

7 [www.d4science.org](http://www.d4science.org)
components. Once the VRE is complete, the VRE Manager can start promoting it and managing the requests for access.

3.1.2 WHY THIS IS CONSIDERED A BEST PRACTICE

<table>
<thead>
<tr>
<th>Best Practice Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Validation</strong></td>
</tr>
<tr>
<td><strong>Innovation</strong></td>
</tr>
<tr>
<td><strong>Success Factors</strong></td>
</tr>
<tr>
<td><strong>Sustainability</strong></td>
</tr>
<tr>
<td><strong>Replicability and/or up-scaling</strong></td>
</tr>
</tbody>
</table>

3.1.3 LESSONS LEARNT

The approach promoted here makes it possible to create VREs easily and efficiently. The immediate impact of this approach on end users is the time reduction to create a working environment and a limited “yet another tool to learn” effect. In fact, VRE deployment time is certainly reduced with respect to the “development from scratch” case thanks to the re-use of existing components. Moreover, VREs share common services for specific tasks users are familiar with, e.g. a shared workspace resembling a typical file manager, a post-based mechanism for discussions, a catalogue for searching and browsing the resource space.

However, developing VREs is a challenging process: it is largely affected by the almost open-ended set of facilities and resources a community can be interested in. BlueBRIDGE has limited this issue by providing each community with a core set of common facilities (file repository, communication tools, wiki, VRE members management, etc.) that largely simplifies the task of developing a new VRE.
Among the main challenges underlying this approach is the fear of communities losing the control of their own resources, lock-in disputes, and sustainability discussions. Such challenges have been addressed by having well defined terms of use, making the system and its services as open as possible and standard-based to simplify the access to what is (produced) in a VRE, having specific Service Level Agreements regulating the provisioning of the VRE.

3.2 PUBLISHING SOFTWARE IN ZENODO

A key aspect of an open source software project is to make the software discoverable, searchable and referenceable to the largest possible number of communities. To achieve this objective is essential to publish the software on multiple platforms/channels and provide a rich set of metadata. An additional challenge for the gCube system (the software powering the BlueBRIDGE VREs) is the high number of components (over 500) and the high number of releases (about one per month) that need to be published.

Before BlueBRIDGE the source code of gCube was only hosted in a Subversion repository internal to CNR and the binaries packages were only published on the gCube website. This resulted in a software that was difficult to discover and access.

3.2.1 THE BLUEBRIDGE BEST PRACTICE

In order to improve the gCube distribution process, it was chosen to publish the releases of gCube components in the Zenodo (https://zenodo.org/communities/gcube-system/) portal. Zenodo offers a rich set of metadata (e.g. authors, description, funding, license, keywords, relationships with other objects) that can be associated to each software component uploaded. This makes the software uploaded in Zenodo easily discoverable and searchable (also programmatically through REST and OAI-PMH interfaces). Zenodo also assigns a unique identifier (i.e. DOI) to each object uploaded. This solves the problem of identifying and referencing a particular software component (also support versioning). Finally, Zenodo offers storage capabilities to host the binary and source packages of gCube system.

In addition to Zenodo, the source code of gCube components is also hosted in GitHub (https://github.com/gcube-team/gcube-releases). GitHub is the biggest community for open source components and this allows the visibility of gCube software to be increased considerably. Furthermore, this integrates nicely with Zenodo, allowing each object uploaded in Zenodo to be linked to the corresponding source code in GitHub.

Given the size of gCube software and the release frequency, all the publication steps (both in Zenodo and GitHub) have been automated.

This practice can be very beneficial for the following stakeholders:

- Developers who need access to software to inspect the source code, fix and/or improve it, build new functionalities on top of it;
- Researchers using the software to do their experiments, as in order to publish their works and assure the repeatability of the experiments, they need a way to reference and cite the software used during the experiments;
- Infrastructure Managers who need to access software packages and the documentation to install and maintain a gCube infrastructure.

3.2.2 WHY THIS IS CONSIDERED A BEST PRACTICE

<table>
<thead>
<tr>
<th>Best Practice Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validation</td>
</tr>
</tbody>
</table>
Innovation | The innovation relies in the application of an open science best practice.

Success Factors | The software development process in the organization must be mature enough to have a meaningful and coherent versioning system, include and maintain meaningful metadata in the software, set-up specific procedures for software integration, release and distribution. Technologically, this practice relies heavily on the automated extraction of metadata information (e.g. authors, description, license) from the software source code. To make it work, the software must include this information in a semi-structured format coherent with all components of the software.

Sustainability | N.A.

Replicability and/or up-scaling | This practice has been used for gCube software, but can be easily applied to any other software. In fact, all the procedures, tools and data used do not make any assumption as to the type of software. This practice has already been proven to work well for big software projects, since gCube is composed of over 500 components and about one release per month. However, thanks to the automation put in place, it is easy even to apply this practice to bigger projects without considerable extra-effort.

### 3.2.3 LESSONS LEARNT

The main challenge of applying this practice is to make sure that the data (source code, binaries and metadata) published is correct - also considering that once published on Zenodo, the information cannot be removed.

To overcome this challenge in gCube, a set of compliance rules have been defined and communicated to the developers. Components compliance to these rules is automatically checked during the integration process to assure the quality of components released.

### 3.3 MAKING ENVIRONMENTAL SCIENCE REPRODUCIBLE

Most users of environmental datasets are trying to do reproducible and accountable science, but different post-processing workarounds and tools can lead to published results which are not repeatable or comparable.

#### 3.3.1 THE BLUEBRIDGE BEST PRACTICE

The European Commission Joint Research Centre (JRC) delivers and maintains the Digital Observatory for Protected Areas (DOPA\(^8\)). The DOPA is a set of web services and applications primarily used to assess, monitor, report and possibly forecast the state of and the pressure on protected areas at multiple scales. The data, indicators, maps and tools provided by DOPA are relevant, for example, to support spatial planning, resource allocation, protected area development and management, and national and international reporting. Indeed, DOPA was acknowledged by the Convention on Biological Diversity (CBD) Secretariat as a reference tool for Country reporting.

Maintaining DOPA requires management of large datasets with highly complex geometries, topological inconsistencies, multiple representations of the same geographical entities, for example coastlines, and

---

\(^8\) dope.jrc.ec.europa.eu
licensing requirements in order to continuously update indicators in response to monthly changes in authoritative data. In order to compute and publish these arrays of indicators, JRC is using a range of open source tools (including GRASS, R, python, GDAL, PostGIS, geometry libraries for Hadoop, Geoserver, Geonode, Mapserver) coupled with some commercial software (such as ArcGIS Pro and the Google Earth Engine platform).

To make all of this reproducible, JRC is trying to move the entire processing chain to open source tools and share it as a versioned resource. The latter is done with the help of BlueBRIDGE with whom JRC is collaborating.

JRC and BlueBRIDGE have developed the PAIM Virtual Research Environment (https://bluebridge.d4science.org/web/protectedareaimpactmaps/) aimed at reporting which features are represented in protected area networks and other managed areas. In particular, the ongoing use case has been developed in the context of the Biodiversity and Protected Areas Management Programme (BIOPAMA Reference Information System9), which aims to address threats to biodiversity in African, Caribbean and Pacific (ACP) countries, while improving socio-economic conditions of the local communities in and around protected areas. The DOPA can directly consume the outputs of the PAIM Virtual Research Environment to update statistics on ecologically important seafloor features represented in protected areas.

### 3.3.2 WHY THIS IS CONSIDERED A BEST PRACTICE

<table>
<thead>
<tr>
<th><strong>Best Practice Analysis</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Validation</strong></td>
</tr>
<tr>
<td><strong>Innovation</strong></td>
</tr>
<tr>
<td><strong>Sustainability</strong></td>
</tr>
<tr>
<td><strong>Replicability and/or up-scaling</strong></td>
</tr>
</tbody>
</table>

---

9 http://rris.biopama.org
3.3.3 LESSONS LEARNT

Different post-processing workarounds and tools can lead to published results which are not repeatable or comparable. To work more effectively, the ideal process would be to share value-added data processed to an agreed standard and format. Since legal restrictions currently forbid this type of redistribution, the next best solution is to share the processing workflow, including code and environmental settings or parameters. The PAIM Virtual Research Environment provides access to a cloud based processing workflow that can be reproducible executed, using the same source data, to produce repeatable results.

3.4 VIRTUAL RESEARCH ENVIRONMENTS TO SUPPORT SCIENTIFIC TRAININGS

Education in an interdisciplinary scenario is usually delivered through university courses, focussed training events and workshops provided by specialised scientific institutes. As a result, there can often be a gap in making the necessary technological support available to scientists, trainees and students.

All the steps in preparing courses and workshops require:

- manual work, which is repeated each time a new course is held.
- installation of complex software on the users’ computers to enable use of data processing services and models that usually come under heterogeneous programming languages.
- an extensive phase of data preparation and powerful hardware installation to allow the execution of the data-intensive models.

In addition, the interaction between teachers and students is usually limited to the duration of the face-to-face part of the course.

A VRE can ease the process of communication, set-up and delivery of training, reducing the need for specific software, and facilitating learning.

3.4.1 THE BLUEBRIDGE BEST PRACTICE

BlueBRIDGE offers Virtual Research Environments (VREs) to set up and deliver training courses in a cost-effective way. The BlueBRIDGE VREs are collaborative, web-based applications which enable collaboration and integrated access to potentially unlimited digital research resources, as well as cross-disciplinary and cross-community tools and services.

The instructor with the support of the BlueBRIDGE consortium sets up the training environment for the course equipping it with the datasets and the data analytics tools that s/he needs. In addition, the file repository facilities available in the VRE, allows the trainer to upload all of the training material needed for the course (powerpoints, word and pdf files, etc.).

Once the environment is ready, the trainer invites all the trainees to register to the environment. From that moment, all of them have access to all the resources available in the VRE. Finally, by exploiting the social media features, the trainer can exchange communications with the course participants without spending time in setting up mailing lists and other dedicated tools.

3.4.2 WHY THIS IS CONSIDERED A BEST PRACTICE

Best Practice Analysis
### Validation
The BlueBRIDGE VREs have been used to deliver a set of scientific trainings organised by ICES, UOA and FAO (See [http://www.bluebridge-vres.eu/supporting-education](http://www.bluebridge-vres.eu/supporting-education)). All the training courses have proceeded without hindrance, and with smooth use of the VRE. This can is further confirmed by the feedback collected on the course through the survey facility and from the course instructors.

### Innovation
The usage of the VRE has allowed the trainers to focus on the preparation of the content of the training course instead of spending time and energy in setting up the tools to run the course. In addition, the usage of the training VRE, with the basic functionality of a cloud infrastructure encompassing data and services for a specific thematic sector, e.g. aquaculture, together with common services and data for the concurrent usage of all participants – trainees, conveys an innovative educational approach for both trainers and trainees in the academic sector.

### Success Factors
A sufficient preparation period is required for the VRE education providers to advertise the efficiency and the innovative approach conveyed of the particular method, in addition to the usefulness and effectiveness of it, especially for the thematic educational domains that are unrelated to the information society world. Access to the platform hosting the VREs is required for every participant, and this may not be the case for organizations and universities not related to the ICT domain. For the use of a VRE in this context to be useful, the courses should foresee the usage of datasets and data analytics.

### Sustainability
The conditions that are required, for this particular best practice to be sustainable, are:
- Sufficient number of trainers to be continuously trained based on the VRE technological progress and results, in order to be easily transferred to the potential target groups.
- Strategic roadmap of the VRE provider for a clear timetable of regular training activities.
- Relevant manuals and videos have to be produced for the complete coverage of the training process.
- 2-person weeks are required for the course preparation; 2-5 person-days are required for the trainers (1 per trainer) for the day of the training plus the travel and accommodation expenses, if the tutorial location is different to the VRE production place.

### Replicability and/or up-scaling
The usage of the VREs as training environments is domain-agnostic therefore applicable to different sectors and domains.

### 3.4.3 LESSONS LEARNT
The usage of VREs to deliver scientific trainings has demonstrated that course instructors, course administrators and course participants have saved a huge amount of time as they didn’t have to deal with the set-up of different tools and environments for the course delivery. The course material documents and information were easily distributed and accessible. In addition, thanks to the computational resources made available by the underlying BlueBRIDGE infrastructure, the course participants were able to run experiments in parallel during the course.

Challenges in the application of the best practice include some hesitation from instructors to make full use of, and take advantage of the new tools/facilities which the VREs offer. With the time constraints that we
face, instructors and participants may be unwilling to spend the time to become familiar with the new interface, and tools, despite assurances that it will save time and effort in the long run. Lessons learned include the importance of distribution of the VRE in good time prior to the course, to allow participants and instructors the possibility to become familiar with the interface. A good explanation of the options / facilities offered by the VRE is also very useful. That’s why a brochure dedicated to explain the different facilities offered by the VREs has been created and disseminated. See http://www.bluebridge-vres.eu/sites/default/files/flyer_bblueskills_17102017.pdf_brouchure_3.pdf.

3.5 BOOSTING KNOWLEDGE TRANSFER

Boosting education means also creating new skills related to diverse topics. Some of the topics related to BlueBRIDGE and the tools delivered by the project became part of university programmes at the University of Athens, at the Sorbonne University and at the University of Pisa.

3.5.1 THE BLUEBRIDGE BEST PRACTICE

Knowledge transfer has been since the beginning one of the major objectives of the BlueBRIDGE project. Given its diversity in scope and target audience, BlueBRIDGE was able to address different topics that were able to gather the interest of some universities. The interest was mainly due to the fact that BlueBRIDGE was able to couple the active usage of VREs with the thematic knowledge of the specific domain, i.e. aquaculture, as well as the specific knowledge of the spatial data, indicators, models, e-infrastructures. The most successful examples are the courses organised by BlueBRIDGE for four Greek university departments:

- The Department of Informatics and Telecommunications – National Kapodistrian University of Athens,
- the Department of Economics - National Kapodistrian University of Athens,
- the Department of Ichthyology and Aquatic Environment – University of Thessaly and
- the Department of Fisheries – Aquaculture Technology – T.E.I of Western Greece

For all these institutes BlueBRIDGE delivered training courses related to the general objectives of e-infrastructures, the particular BlueBRIDGE approach, the VREs context and the particular VRE for the integrated aquaculture production under the generic title “BlueBRIDGE: New generation tools for aquaculture” covering performance evaluation, benchmarking and decision making in aquaculture VRE, strategic Investment analysis and Scientific Planning/Alerting VRE, social and environmental monetization models for Blue Economy and Geoanalytic services. Each course was adapted to the specific scientific directions of each department. Special focus on the architecture and functionality of the VREs has been given in the ICT departments, special analysis and detailed presentation of the technoeconomic and socioeconomic approach and models at the Economic Department and finally, special attention on the aquatic concepts and needs covered when the courses were addressed to the students of the Ichthyology departments. The interactivity of each course given by the practical usage of the dedicated training VRE “AquacultureTrainingLab” was a necessity for the successful completion of every course.

Another best practice related to knowledge transfer is the one related to the PhD Course - “Signal Processing and Mining of Big Data: Biological Data as a Case Study” delivered by Gianpaolo Coro, CNR. Gianpaolo presented a course on Signal Processing and Mining of Big Data: Biological Data as a Case Study at the University of Pisa, 2-6 May 2016. 25 doctoral students from various universities attended with interested parties from such background as Computer Engineering and Computer Science.
The course focused on applications of Big Data analytics in the biology domain to predict climate change impact on species' distribution, to monitor the effect of overfishing on economy and marine biodiversity and to prevent ecosystems collapse. The course, has showcased practical applications of Big Data analytics, with focus on several signal processing and machine learning-based techniques. The course clarified the general concepts behind these techniques, with an educational approach making these concepts accessible also to students with intermediate mathematical skills. The examples will regard real cases involving data that would have been unpractically to be human-analysed and corrected, especially in the biology domain: time series forecasting, periodicities detection, comparison of geographical distribution maps, assessment of environmental similarities between different areas, global scale species distributions.

### 3.5.2 WHY THIS IS CONSIDERED A BEST PRACTICE

<table>
<thead>
<tr>
<th>Best Practice Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Validation</strong></td>
</tr>
<tr>
<td>The students and the Universities appreciated BlueBRIDGE to deliver these courses demonstrating their interest in the topics addressed by the project. Universities have decided to continue using the BlueBRIDGE e-Infrastructure in courses that will be held after the end of the project. For example, University of Pisa will host one course per year and the Sorbonne University is going to host at least one course in 2018. These courses will be funded by the same Universities. Adopting the same approach, the two Greek ichthyology departments will adopt a generic course within the undergraduate and postgraduate program related to the EU projects’ outcomes and their practical usage.</td>
</tr>
<tr>
<td><strong>Innovation</strong></td>
</tr>
<tr>
<td>It is quite unusual that the topics of an EC-funded project become part of a University programme. This is the innovation brought by BlueBRIDGE to the education sector. The e-Infrastructure demonstrated to be able to support sharing of course material, discussions between students, communication between the students and the teacher, direct sharing of exercises results with the teacher, support to examinations via online questionnaires. Finally, the special skill of adopting new generation tools of ICT in the aquaculture provides an extra capacity to the today students and future aquaculture managers.</td>
</tr>
<tr>
<td><strong>Success Factors</strong></td>
</tr>
<tr>
<td>The technology developed by BlueBRIDGE for capacity building and to support teachers demonstrated to be very powerful. It made attendees use multi-disciplinary approaches and perform complex experiments in the time frame of a course, which would have been nearly impossible without a network of services able to interoperate and to connect data and models all together. The computational capabilities offered to the classes allowed executing data-intensive experiments and giving an overview of the potential of data-intensive science. Online experimentation spaces allowed people to work together in this context and made attendees share their expertise. Further, as a side effect, in one of the courses the infrastructure facilitated a student impaired at 90% and participating to the course remotely, to execute the exercises together with his colleagues during the course and to sustain the exam without distinction because of his impairment. This was a surprising side effect that made the University (and the student) appreciate very much the e-Infrastructure.</td>
</tr>
<tr>
<td><strong>Sustainability</strong></td>
</tr>
<tr>
<td>The conditions that are required for this particular best practice to be sustainable, are:</td>
</tr>
<tr>
<td>- Availability of trainers aware of the VRE technology and features;</td>
</tr>
<tr>
<td>- Availability of the e-Infrastructure and possibility to create new VREs for courses;</td>
</tr>
<tr>
<td>- Availability of funds by Universities to host a teacher and to possibly sustain the VRE expenses.</td>
</tr>
</tbody>
</table>
3.5.3 LESSONS LEARNT

BlueBRIDGE was able to efficiently support capacity building in multi-disciplinary contexts. The technology is mature enough to support hands-on training courses for students and scientists as well as live challenges like datathons etc. In some of the courses the teachers observed a practical application of the Open Science principles with attendees creating new ways to exploit the e-Infrastructure through combinations of data and models that had not been foreseen. Further, the support of an impaired person and the possibility to make him overcome his difficulties during a hands-on course was a very positive discovery, which opened the way to other possible applications in this direction.

3.6 USING PUBLIC EINFRASTRUCTURES FOR COMMERCIAL & OPERATIONAL PURPOSES

BlueBRIDGE strongly believes that Small, Medium and Large enterprises can benefit from the data, resources and tools made available by the existing European infrastructures and leverage them to build new services, or enhance existing ones. As part of its mandate, BlueBRIDGE supports IT companies operating in the aquaculture sector by providing them with data, computational resources, data analytics and expertise. But that is not all. During the project, the BlueBRIDGE consortium realised that not only companies operating in the aquaculture sector could benefit from its offering and therefore launched a Call for SMEs (See http://www.bluebridge-vres.eu/bluebridge-competitive-call-data-management-services-smes) through which 5 SMEs operating in the Blue Growth sector were engaged and started using the BlueBRIDGE services.

3.6.1 THE BLUEBRIDGE BEST PRACTICE

The collaboration with SMEs raised a couple of criticalities to which BlueBRIDGE had to respond promptly:

a) SLA level and orientation of e-infrastructures does not present the terms met by commercial infrastructures and
b) operational policies of e-infrastructures formally oppose utilization for indirect or direct commercial activities.

To overcome the above issues BlueBRIDGE adopts the practice of including private companies as resource providers in the context of BlueBRIDGE. This allows a mixture of services coming from public and private sector that can cover the aforementioned challenge. As a best practice too, the project suggests that no revenue generating services are executed on infrastructure resources provided by public bodies. If revenue streams emerge, commercially powered resources may be utilized for hosting the relevant activities. Those can be provided by private companies engaged in the project as infrastructure providers. In that case the only resources that will remain provided via the eInfrastructure will be the services that coordinate the resource tracking and monitoring.

Another part of the practice is related to the licensing of the software used to power the infrastructure. The fact that the BlueBRIDGE software is provided under EUPL licensing 1.1 that allows commercial partners to replicate the infrastructure outside the e-infrastructure landscape.
Additionally, BlueBRIDGE, although it cannot go beyond the typical Research Infrastructures SLAs offered by partners has adopted the practice of thorough monitoring and accounting of infrastructure status and activity, which is an evidence of the quality of the services. Although not an SLA type of commitment, proves that the statistical level of services is of premium quality and adequate for the operations hosted on it.

Note: EOSC is expected to deliver more open and concrete policies on e-infrastructure (at least indirect) commercial utilization in the near future, and better address the SLA requirements.

The implementation of this practice has allowed private companies to proceed with the integration of their services on the BlueBRIDGE e-infrastructure without spending time and resources in addressing all the legal issues related to the commercialisation of a product via the e-infrastructure.

### 3.6.2 WHY THIS IS CONSIDERED A BEST PRACTICE

#### Best Practice Analysis

<table>
<thead>
<tr>
<th>Validation</th>
<th>The infrastructure SLAs have been analysed by commercial partners (project partners and platform boarding aquafarming SMEs) and, as there are still grey areas related to how e-infrastructures can be utilised to support commercialisation of products/services, the solution described in the best practice is the one that allowed companies to get closer to the e-infrastructure world and to further explore the possibilities offered by this area.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Innovation</td>
<td>Engaged stakeholders managed to obtain and exploit services leveraging on e-infrastructures, respecting the fundamental rules of non-commercial exploitation of the latter.</td>
</tr>
<tr>
<td>Success Factors</td>
<td>The practice can be replicated in a flexible service provider environment that can shift the business model to comply with e-infrastructure approaches.</td>
</tr>
<tr>
<td>Sustainability</td>
<td>The practice may require commercial resources to be provided, along with the e-infrastructure ones, while it is necessary that this has to be supported or guided by higher (EC) level strategic decisions on e-infrastructure utilization.</td>
</tr>
<tr>
<td>Replicability and/or up-scaling</td>
<td>The practice is replicable in all cases where commercial partners offer services on top of e-infrastructures. As such utilization remains limited, and no direct commercial exploitation is assumed although operating in a grey area, this may be acceptable. The same holds for the acceptance of e-infrastructure resource providers’ SLAs. However, this practice is not appropriate for scaling until strategic decisions are taken.</td>
</tr>
</tbody>
</table>

### 3.6.3 LESSONS LEARNT

The private sector is highly interested in understanding the benefits of using e-infrastructures. Utilization of e-infrastructures by commercial actors is still a grey area both in the SLA and Terms of Use, even when no direct commercial exploitation of the services is performed. Complementing an e-infrastructure with additional commercial resources, as well as acceptance of lower level SLAs may overcome this issue, however this does not provide a complete solution to the challenge, which needs strategic decisions to be taken in order to be rectified.
4 ENGAGING KEY STAKEHOLDERS IN INNOVATIVE WAYS

The best practices belonging to this area are:

- Organising Datathons to fuel new thinking and develop new skills (cf. Sec. 4.1);
- Boosting impact through multipliers (cf. Sec. 4.2);
- Leveraging call for SMEs and webinars to engage potential users (cf. Sec. 4.3);
- Managing a large consortium through VREs (cf. Sec. 4.4).

4.1 ORGANISING DATATHONS TO FUEL NEW THINKING AND DEVELOP NEW SKILLS

One of the objectives of BlueBRIDGE was to stimulate adoption of its resources (datasets, data analytics and technologies) and transfer knowledge to key stakeholders in different domains. Traditionally these objectives are achieved through promotional activities and ad hoc face-to-face meetings with potential users and through training courses. Very often these are separate activities which requires quite a deal of effort for their preparation and financial resources. That’s why BlueBRIDGE implemented a different approach that can be considered a best practice: the organisation of datathons.

4.1.1 THE BLUEBRIDGE BEST PRACTICE

In June 2017, BlueBRIDGE decided to organise a datathon. The generic challenge of the BlueBRIDGE datathon was to fuel new thinking in the holistic analysis of satellite, environmental, biological and socioeconomic data and the interactions between fisheries, aquaculture and marine spatial planning.

The datathon-specific challenges were:

- Developing a holistic approach to spatial planning for marine protected areas based on knowledge transfer and integration
- Assessing new fisheries and aquaculture activities and mainstreaming environmental and socioeconomic indicators
- Incorporating socio-economic and environmental data in aquaculture assessment, strategic planning and performance analysis

A datathon of this type usually takes place at a research centre or environmental organization, with a usual duration of 2 days, that’s why BlueBRIDGE selected the HCMR venue.

The datathon mainly focused on developing innovative ideas and applications using the data formats, services and products offered by BlueBRIDGE and by other related initiatives (Copernicus, EMODnet, etc.).

The agenda of the datathon was very interactive including presentations of the challenges and of the different tools and datasets provided by the different supporting initiatives, along with related use cases for the specific thematic area of the datathon (such as aquaculture, fisheries, etc). One day and a half then were entirely dedicated to the working sessions: participants teamed up in working groups which, supported by a set of coaches with diverse background and skills, transformed their ideas in working applications.

In this way participants were able to start developing applications while enhancing their skills.

4.1.2 WHY THIS IS CONSIDERED A BEST PRACTICE

Best Practice Analysis
### Validation

A special evaluation session was organized at the end of the datathon engaging the audience, invited speakers and participants. This allowed the organizers to collect feedback from the attendees which resulted to be very positive.

### Innovation

The datathon can be considered as an innovative approach as it managed: — based on the main idea of data exploitation - 1) to attract many participants, having diverse scientific and professional backgrounds, 2) to highlight the challenges of different but complementary domains (such as fisheries, aquaculture, business and ICT) and 3) to stimulate the participants to work on innovative and efficient ideas.

### Success Factors

Based on the experience gained, the next successful datathon requires:

1. Nice environment supporting the cooperative atmosphere (location, weather, food, recreational activities) in order to provide pleasant working conditions
2. A concrete scope of the datathon
3. Complementary areas of interest, such as aquaculture and information technology
4. Excellent and convincing speakers
5. Sufficient amount of open datasets, services and tools
6. Sufficient mixture of differently experienced participants, such as engineers and aquaculture scientists and researchers
7. Prizes for attracting participation

### Sustainability

- Methodology steps to be kept as a practice
- Datathon events to be included as part of training strategy of EC funded projects
- Frequent dissemination activities regarding the usefulness and the effectiveness of the “datathons” to potential organizations

### Replicability and/or up-scaling

Datathons can be replicated in each data-intensive domain. In particular the BlueBRIDGE datathon inspired the organisation of the EMODnet Open Sea Lab Hackathon which was as well a successful datathon in the marine domain

### 4.1.3 LESSONS LEARNT

A Datathon is an excellent opportunity to meet people, projects, results (data and services) from different domains and provide promising outcomes for further development. It additionally provides the framework for combining the needs and answers in a very cooperative environment. New technologies become familiar to a wider audience of varying expertise and backgrounds and new synergies evolve.

The Datathon proved to be an innovative, interactive and user-friendly approach to convince students, researchers and managers from thematic areas unfamiliar with ICT, to use and exploit the up-to-date data models and services made available by BlueBRIDGE for their needs.

The datathon is a complex concept that is difficult to communicate to the candidate participants. Multiple methods of approach are suggested such as: f2f meetings, social media campaigns, traditional announcements tailored to the thematic areas experts that are not familiar with the “thon” events, such as hackathons, datathons, etc. An evaluation session should also be part of the datathon agenda, in order to ensure the participation and response of the participants.

---

10 http://www.opensealab.eu/
4.2 BOOSTING IMPACT THROUGH MULTIPLIERS

One of the major objectives of BlueBRIDGE was to promote the uptake of its services to the different stakeholders. But how could all the different stakeholders (industry, policy makers, academia, etc.) be reached in a cost effective way?

Multipliers are organisations that have built a strong network of stakeholders over the years. Through these organisations the messages to the stakeholders can be conveyed easier and quicker with less resources. Examples of multipliers are the Business and Innovation Clusters.

4.2.1 THE BLUEBRIDGE BEST PRACTICE

It was already clear in the proposal preparation phase that to reach the BlueBRIDGE key stakeholders it would have been important to use the right channels, especially the multipliers. In particular, as BlueBRIDGE was aiming to promote the uptake of its services to SMEs, since the beginning of the project a Business and Innovation Cluster, the Pôle Mer Bretagne Atlantique, was involved in the consortium.

The Pôle Mer Bretagne Atlantique is a French economic development cluster that facilitates the emergence of collaborative research and development projects and support the growth of their member companies. Its role in its region ecosystem is particularly to place new products, services and processes resulting from research on the market. Their input ensures that the companies involved aspire to a privileged position on both domestic and international markets, stimulating growth and consolidating jobs in different fields including those targeted by the BlueBRIDGE project, especially marine biological resources and environment and marine coastal planning.

With its network of 350 members including 50% of SMEs directly linked to the marine and maritime sector, the Pôle Mer Bretagne Atlantique was able to give advice on the markets to the consortium, to engage SMEs within the Call for SMEs and to raise the expectations of the private sector and to advocate appropriate methods for instance.

4.2.2 WHY THIS IS CONSIDERED A BEST PRACTICE

<table>
<thead>
<tr>
<th>Best Practice Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Validation</strong></td>
</tr>
<tr>
<td><strong>Innovation</strong></td>
</tr>
<tr>
<td><strong>Success Factors</strong></td>
</tr>
</tbody>
</table>
**Sustainability**
The sustainability of the practice is linked to the possibility to reserve some funding for these kind of actors in an EC-funded project or to find diverse ways to engage them.

**Replicability and/or up-scaling**
This best practice is replicable in any domain especially in those where the industry component is relevant.

### 4.2.3 LESSONS LEARNT

The communication of the outcomes of the BlueBRIDGE project is extremely complicated, as it involves so many projects in one. Different services developed for different stakeholders (public and private sector) in different domains and sectors (aquaculture, biodiversity, support to stock assessment, education, etc.). The involvement of a Business and Innovation Cluster was a good means to reach the key stakeholders with a reasonable amount of effort and a short timeline (the BlueBRIDGE project is a 30-month project). The involvement of one cluster was also useful to engage other clusters in Europe which helped amplify the message. This was particularly important as the regional specificities of the area covered by a cluster may not be representative enough of each country or regions’ situation (this was clear when the results of the needs and expectations survey that was carried out through PMBA’s network was presented to the consortium and the reviewers, who despite the high number of respondents did not consider it sufficient as a response, as representative of only one region).

The main interest of engaging a business cluster in a project remains that it can represent an economic sector in its entirety instead of individual interests, as can be the case when a private actor gets involved in a consortium.

### 4.3 LEVERAGING CALL FOR SMES AND WEBINARS TO ENGAGE POTENTIAL USERS

The success of BlueBRIDGE is linked to the adoption of its services and products which target a specific niche of stakeholders in different domains. In order to promote adoption, stakeholders need to understand what are the benefits in using them.

#### 4.3.1 THE BLUEBRIDGE BEST PRACTICE

BlueBRIDGE to maximize the promotion of its services and tools to the widest audience as possible adopted two main approaches:

- It launched a call for SMEs
- It organised a set of webinars

The Call for SMEs\(^\text{11}\) was aimed to attract potential users from the private sector. A huge dissemination campaign, leveraging also on the network of many EU Business and Innovation Clusters, was performed to engage SMEs operating in the Blue Growth sector willing to test the BlueBRIDGE services for their scopes. The aim of the call for SMEs was to raise awareness of the BlueBRIDGE services and to concretely support a certain number of SMEs in addressing their issues. 5 SMEs were selected:

- AquaBioTech Group, an international research and development, engineering, technology provider and consulting company located in Malta. The AquaBioTech Group undertakes a variety

of aquaculture, fisheries, marine surveying, aquarium and aquatic environmental projects through its regional offices and partners throughout the world. The vast majority of the company’s work is related to the marine or aquatic environment, encompassing aquaculture developments, market research / intelligence, through to project feasibility assessments, finance acquisition, project management, technology sourcing and technical support and training. With the data and the models made available by BlueBRIDGE, AquaBioTech wants to study the nutrition dispersion and retention in the environment surrounding a local fish farm. The results of the collaboration with BlueBRIDGE will be also analysed in the context of the TAPAS - Tools for Assessment and Planning of Aquaculture Sustainability - H2020 project, in which the SME is involved, to widening the scope of the collaboration.

- Flyby, an independent Italian company, specialized in the development of Decision Support Systems exploiting edge technologies in the field of Remote Sensing, Signal Processing and Big Data Analytics. The company has been founded in 2001 with the aim to develop innovative solutions exploiting data analytics. Flyby operates in five different business sectors (Defence, Space, Health, Maritime & Fishing, Energy). Currently, Flyby is developing a new commercial solution to monitor and provide advice on the sustainability of fish resources: the work going on in BlueBRIDGE related to fisheries stock assessment could enhance the solution under development.

- Kamahu, a French start-up created at the beginning of 2017 in order to conceive and market farm control SAAS solutions for the aquaculture industry wanting to offer benchmarking services to its clients.

- Sinay, a French Environmental Consulting company specialised in underwater acoustics, maritime surveillance and fishery science. Founded in 2008, the company provides a wide range of services such as environmental impact and risk studies related to human activities in the oceans, marine wildlife population surveys and fish stock assessments. Sinay aims to exploit the BlueBRIDGE resources to study the impact of man-made underwater noise on different key marine species in the Mediterranean area. Sinay is already working on these kind of assessments, therefore BlueBRIDGE will be useful to compare the results of the different models used and to assess an alternative solution to explore eventual cost-efficiency benefits.

- StatnMap, a French micro-enterprise providing tutorials and expertise in statistics, modelling, spatial data analyses and mapping for beginners or advanced users from universities, research laboratories and companies. The main field of expertise is marine and aquatic ecology. By exploiting the data, the models and the computational resources made available by BlueBRIDGE, StatnMap aims to build species distribution models in a more efficient way and enhance its offering.

The SMEs are currently exploiting the BlueBRIDGE resources.

The second engagement method that BlueBRIDGE adopted was the organisation of webinars¹²: one hour live presentation/demo delivered online via the BlueBRIDGE website. Overall the project organised 7 webinars attracting around 300 relevant stakeholders.

1. Innovative services to monitor the spatial distribution of human activities
2. The Global Record of Stocks and Fisheries
3. Managing tuna fisheries data at a global scale: the Tuna Atlas VRE
4. New generation tools for Aquaculture

¹² http://www.bluebridge-vres.eu/webinars
5. A web-application to understand ecologically important seafloor features in Marine Protected Areas
6. Semantic Integration of Marine Data
7. Using e-Infrastructures for Biodiversity Conservation

The webinars allowed to save costs to the project but also to the participants and to outreach a very wide geographical distribution of participants. In the BlueBRIDGE case, as the project is dealing with very different topics and stakeholders, the webinars were a perfect mean to attract potential users. Many of the webinars attendees registered to the VREs right after the webinar and becoming users of the BlueBRIDGE infrastructure.

### 4.3.2 WHY THIS IS CONSIDERED A BEST PRACTICE

<table>
<thead>
<tr>
<th>Best Practice Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Validation</strong></td>
</tr>
<tr>
<td><strong>Innovation</strong></td>
</tr>
<tr>
<td><strong>Success Factors</strong></td>
</tr>
<tr>
<td><strong>Sustainability</strong></td>
</tr>
<tr>
<td><strong>Replicability and/or up-scaling</strong></td>
</tr>
</tbody>
</table>

### 4.3.3 LESSONS LEARNT

The call for SMEs required a huge dissemination and management (after the selection of the SMEs) effort. In the future it would be good to include such process as a part of the workplan and allocate some financial resources to the winning SMEs. This will facilitate their participation as one of the main barrier emerged during the numerous phone calls that the consortium had with candidate SMEs during the call promotion period was that SMEs do not have resources to allocate to run these pilots. A financial support might help. The call needs also to be planned from a consortium perspective as in many cases SMEs are not familiar with e-infrastructures and considerable time has been spent by the consortium in understanding the requirements of the companies and in explaining how to address them with the e-infrastructure resources. In many cases ad hoc implementations have been requested.

For what concerns webinars they work really well in BlueBRIDGE because they allowed the consortium with a minimum effort to engage different stakeholders.

### 4.4 MANAGING A LARGE CONSORTIUM THROUGH VRES

One of the major issue in managing an EU-funded project is to ensure a smooth communication among the partners and make sure that each partner has access to all the relevant resources and latest information.
4.4.1 THE BLUEBRIDGE BEST PRACTICE

The Virtual Research Environments developed by BlueBRIDGE are a powerful mean to manage collaboration between different partners. That’s why BlueBRIDGE decided to adopt a VRE as its project management tool. The BlueBRIDGE Project VRE was created by the coordinator and all the partners registered to it. Thanks to the integrated file repository all the partners had access to the project documents (deliverables, communication materials, etc.). The social framework part of the VRE (an area similar to the Facebook wall where the VRE members can post new messages and reply to existing ones) enabled communication among the members of the VRE (no mailing lists were set up) allowing to keep track of all the project discussion on the VRE wall. Finally, the VRE was equipped with a wiki to manage the production of all the project relevant content and with a real-time dashboard to track activities and Key Performance Indicators (KPIs) which allowed the project coordinator to have the status of the project under control in every moment. Beside the BlueBRIDGE project VRE other VREs were created for the management of the project like for example the PSC VRE to track the status of the activities of the Steering Committee or the VREs that supported the development of the tools of the different pillars. Clearly, all the project related VREs are private therefore not accessible to externals.

4.4.2 WHY THIS IS CONSIDERED A BEST PRACTICE

<table>
<thead>
<tr>
<th>Best Practice analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validation</td>
</tr>
<tr>
<td>Innovation</td>
</tr>
<tr>
<td>Success Factors</td>
</tr>
<tr>
<td>Sustainability</td>
</tr>
<tr>
<td>Replicability and/or up-scaling</td>
</tr>
</tbody>
</table>

4.4.3 LESSONS LEARNT

The facilities offered by VREs for supporting communication and collaboration among the members proved to be useful when exploited to support large scale consortium whose members are spread across several institutions and regions. The learning curve is really low, the frustrations resulting from the “yet another system to use” are reduced because the services are implemented to offer familiar interfaces and usage patterns.
5 CONCLUSIONS

Given the diversity of the topics and technologies addressed by BlueBRIDGE, the project was able to develop a high number of best practices that can be re-used in different domains and by different stakeholders.

The best practices include:

- **Innovative solutions to support the Blue Growth sector practitioners to address data challenges:**
  - Optimizing the use of open and sub-meter (VHR) resolution satellite data to generate an aquaculture atlas;
  - Reporting on ecological seafloor features in marine protected area networks;
  - Scientific Data modelling and aggregation of marine data;
  - Publishing AquaMaps Native Habitat Data and Metadata as Exportable NetCDF Files;
  - Assignment of unique identifiers for harmonised stock and fishery data;
  - Data collation for the implementation of a Regional Database;
  - How to stimulate private companies operating in the aquaculture sector to share data by safeguarding their competitive advantage;

- **Practices to boost Open Science, education and skills:**
  - VRE as the instrument to efficiently manage specific and tailored communities needs;
  - Publishing software in Zenodo;
  - Making environmental science reproducible;
  - Virtual Research environments to support scientific trainings;
  - Boosting Knowledge transfer;
  - Using public eInfrastructures for commercial & Operational purposes.

- **Innovative stakeholder engagement methods:**
  - Organising Datathons to fuel new thinking and develop new skills;
  - Boosting impact through multipliers;
  - Leveraging call for SMEs and webinars to engage potential users;
  - Managing a large consortium through VREs.

All these best practices can be leveraged upon by future initiatives and be of inspiration for stakeholders approaching the data management issues in the Blue Growth area.
REFERENCES


