The JERICO-S3 project is funded by the European Commission's H2020 Framework Programme under grant agreement No. 871153. Project coordinator: Ifremer, France.

JERICO-S3 DELIVERABLE

Joint European Research Infrastructure for Coastal Observatories

Science, Services, Sustainability

DELIVERABLE #, WP# and full title	JERICO-S3 D11.2 - WP11 - "VA Intermediate Report from External International Board"
5 Keywords	Access, Virtual Access, Virtual Access Expert Panel, Metrics, Assessment
Lead beneficiary	SOCIB
Lead Author	Damià Rita (SOCIB), Juan Gabriel Fernández (SOCIB), Miguel Charcos (SOCIB), Anca Hienola (FMI), Antonio Novellino (EMODnet, ETT), Thierry Carval (IFREMER), Simon Keeble (Blue Lobster), Shaun Deyzel (SAEON), Sebastien Mancini (AODN-IMOS) and Joaquín Tintoré (SOCIB).
Co-authors	
Contributors	Susana Rubio (PdE), Julien Mader (AZTI), Laura Alonso (PdE), Lauri Laakso (FMI), Emilie Breviere (SMHI), Maria Sotiropoulou (HCMR), Kate Collingridge (CEFAS), Leonidas Perivoliotis (HCMR), Jean-Olivier Irisson (CNRS-LOV), Milla Johansson (FMI), João Vitorino (IH), Emilie Poisson Caillault (IFREMER, ULCO-LISIC), Pierre-Alexandre Hébert (IFREMER, ULCO-LISIC), Andrew King (NIVA), Luis Felipe Artigas (CNRS-LOG, ULCO), Arno Lambert (UNESCO/IOC-IODE), Taavi Liblik (TALTECH), Hele-Riin Pihel (TALTECH), Carlo Mantovani (CNR), Lennert Schepers (VLIZ), Lars Stemmann (CNRS-LOV), Begoña Gómez (PdE), Patrick Gorringe (SMHI), Lorenzo Corgnati (CNR), Pauline Simpson UNESCO/IOC-(IODE), Klas Ove Möller (Hereon), Veronique Creach (CEFAS), Jay Pearlman (IEEE), Peter Thijsse (Maris), Patricia Cabrera (VLIZ), Paulo Nunes (IH), Gregor Rehder (IOW), Jukka Seppälä (SYKE), Gisbert Breitbach (Hereon), Lohitzune Solabarrieta (AZTI), Yoana Voynova (Hereon), Wanda Holzhüter (IOW).
Final version date/ Submission date (dd.mm.yyyy)	12.09.2022

Nature: **R**

(R = Report, P = Prototype, D = Demonstrator, O = Other)

Dissemination level : **PU**

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

GRANT N°: 871153
PROJECT ACRONYME : JERICO-S3
PROJECT NAME : Joint European Research Infrastructure for Coastal Observatories - Science, services, sustainability
COORDINATOR : Laurent DELAUNEY - Ifremer, France - jerico@ifremer.fr

destroy it thereafter.

DOCUMENT TECHNICAL DESCRIPTION

Document ID	JERICO-S3-WP11-D11.2-12.09.2022-V1.0

REVISION HISTORY

Revision	Date	Modification	Author
V1.0	12-Sept-2022	First Submission	Damià Rita Espada

APPROVALS

	Name	Organisation	Date	Visa
Coordinator	Delauney Laurent	Ifremer	12-Sept-2022	DL
WP Leaders	Joaquín Tintoré & Juan Gabriel Fernández	SOCIB	12-Sept-2022	JT JF

Diffusion list

Consortium beneficiaries	Third parties	Associated Partners	other
X	X	X	

TABLE OF CONTENT

1.EXECUTIVE SUMMARY

This document follows the previous deliverable of JERICO-S3 WP11 (D11.1)[1]. In D11.1, the Virtual Access Metrics System (VAMS), a central system to monitor the access to each VA service, was presented. In that document, the access metrics and outreach activities of each VA service were reported for the 19 first months of JERICO-S3 (from February 2020 to August 2021). D11.1 also presented the Virtual Access Expert Panel whose members, in no particular order, are: Anca Hienola (FMI), Antonio Novellino (EMODnet Physics, ETT), Thierry Carval (IFREMER), Simon Keeble (Blue Lobster), Shaun Deyzel (SAEON) and Sebastien Mancini (AODN-IMOS).

In the second phase of the VAMS, the panel members have assessed all the VA services of JERICO-S3 WP11 and provided their feedback to the VA infrastructures[2]. Each VA service has been assessed by two different members of the panel. The assessments are based on the access metrics reported in D11.1, the outreach activities reported in D11.1 and the direct interaction of the panel members with the VA service. As a result, this deliverable D11.2, contains, for each VA service, a succinct description, the assessments of the assigned members of the panel and, optionally, a clarification from the VA infrastructure.

[1] D11.1 First report on VA JERICO Resources access statistics and service provision: (https://www.jerico-ri.eu/download/jerico-s3_deliverables/JERICO-S3-D11.1-First-report-on-VA-JERICO-Resources-access-statistics-and-service-provision-FINAL-V2-Compressed.pdf)

[2]VA Infrastructure is to be understood as the set of VA services managed by an institution and the team that manages them.

2.INTRODUCTION

Virtual Access (VA) means free access to users provided through communication networks; the available services or resources can be simultaneously used by an unlimited number of users and the users are not selected[3]. A VA service can have very different formats: a website, an API, a repository on GitHub… and it can expose very different types of resources: datasets, added-value products, software, documents, training modules, data processing capabilities…

In order to build on previous research results, encourage collaboration and avoid duplication of effort, speed up innovation and involve citizens and society, the European Commission (EC), via the H2020 Framework Programme, encourages the provision of open access to scientific publications and research data[4]. The EC, in its effort to support open access, helps to financially support Virtual Access that is 'free of charge at the point of use'. At the same time, the EC requires that the VA services offered under a project be assessed periodically by a board composed of international experts in the field, at least half of whom must be independent of the beneficiaries.[5][6]

JERICO-S3 is a project funded by the EC's H2020 Framework Programme. The ambition of its members is the creation of an integrated pan-European multidisciplinary and multi-platform **distributed Research Infrastructure** dedicated to a holistic appraisal of coastal marine system changes.

JERICO-S3 VA services have been described and monitored in the "First report on VA JERICO Resources Access Statistics and Service Provision (D11.1)[7]. In that document, the Virtual Access Metrics System (VAMS), a central system to monitor the access to each VA service, was described. Also, the access metrics and outreach activities of all active JERICO-S3 VA services were presented for the 19 first months of JERICO-S3: from February 2020 to August 2021.

D11.1 also described the creation of the Virtual Access Experts Panel that is responsible for the assessment of JERICO-S3's VA services. Unfortunately, despite his best efforts and

[3] ESFRI White Paper chapter on "Access to RIs Including Trans-national Access" (https://www.esfri.eu/esfri-white-paper/23-access-ris-including-trans-national-access)

[4] Article 18 of the Regulation (EU) No 1291/2013 of 11 December 2013: "Establishing Horizon 2020 - the Framework Programme for Research and Innovation (2014- 2020)"

[5] EU Grants: H2020 AGA — Annotated Model Grant Agreement: V5.2

[6] H2020 Work Programme (https://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/main/h2020-wp1415-infrastructures_en.pdf)

[7] D11.1 First report on VA JERICO Resources access statistics and service provision: (https://www.jerico-ri.eu/download/jerico-s3_deliverables/JERICO-S3-D11.1-First-report-on-VA-JERICO-Resources-access-statistics-and-service-provision-FINAL-V2-Compressed.pdf)

due to his busy schedule, Alan Deidun was not able to contribute to the panel. Shaun Deyzel was kind enough to replace him as a member of the Virtual Access Expert Panel. Therefore, the members of the panel are, in no particular order: Anca Hienola (FMI), Antonio Novellino (EMODnet Physics, ETT), Thierry Carval (IFREMER), Simon Keeble (Blue Lobster), Shaun Deyzel (SAEON) and Sebastien Mancini (AODN-IMOS).

Each VA service was assigned to two different experts (one internal to JERICO-S3 and one external). For the next assessment, at the end of JERICO-S3, different experts will be assigned to each VA service. This document contains the assessments provided by the experts which are been based on:

- The metrics reported in D11.1
- The outreach activities reported in D11.1
- Their direct interaction with the VA services

3.MAIN REPORT

The report starts by describing the Virtual Access Expert Panel and its composition. Secondly, for each VA service, the assessments by the members of the panel and the potential responses from the VA service managers are included.

3.1.VA Expert Panel Composition

Article 16.2b of the JERICO-S3 Grant Agreement states that Virtual Access providers "*must have the virtual access services assessed periodically by a board composed of international experts in the field, at least half of whom must be independent from the beneficiaries*".

It has been the responsibility of the WP11 team to select and invite the members of the Virtual Access Expert Panel. It was decided to create a Panel of six members: three of them, independent from JERICO-S3 and three of them, members of JERICO-S3. Since the publication of D11.1, Alan Deidun has been replaced by Shaun Deyzel as a member of the Panel. This was because, due to his busy schedule and despite his best efforts, Alan Deidun was not able to contribute to the Panel. The following is a short profile of each of the current members of the Virtual Access Expert Panel.

- **Anca Hienola (FMI)**. Anca Hienola was awarded a PhD (Physics) in 2008 from the University of Helsinki, Finland. She works at the Finnish Meteorological Institute as a leading specialist and Open Science Taskforce leader. She has extensive experience in data flows and virtual data access metrics, including evaluation and metrics-based decision-making and policies. She is involved in three major EOSC-related projects - ENVRI-FAIR, EOSC Nordic and EOSC-Future.
- **Antonio Novellino (EMODnet Physics)**. Antonio Novellino has a PhD in Biotechnology and Bioengineering. He is the Research Manager at ETT Spa (Gruppo SCAI). He is a member of the EuroGOOS DATAMEQ group and contributes to several EuroGOOS Task Teams for advising on operational oceanography data management procedures and standards. He serves on the EMODnet Steering Committee, and the EMODnet Technical Working Group. He is the EMODnet Physics coordinator and CMEMS DU deputy coordinator.
- **Thierry Carval (IFREMER)**. Thierry Carval is the head of Ifremer "Scientific Information Systems development". He is particularly involved in the management of in situ Marine data. He is the technical coordinator of the European service "Copernicus Marine in situ TAC" and data manager for the French JERICO sites. He is involved in the development of EOSC through projects such as ENVRI-FAIR, EOSC-HUB, EOSC-FUTURE and Blue-Cloud. He ensures that the JERICO VA services are well aligned with the EOSC e-infrastructure.
- **Simon Keeble (Blue Lobster)**. Simon Keeble is the Managing Director at SME Blue Lobster IT Limited based in Wales, UK. He has extensive experience in big-data systems and web / mobile interfaces and user interaction within the marine and environmental sciences sectors. The Blue Lobster team operated one of the

first services to be part of Virtual Access, in the Framework 7 project FixO3, and they continue to develop and operate a variety of web-based platforms presenting a variety of observational, model and complex derived product data.

- **Shaun Deyzel (SAEON)**. Shaun Deyzel is the Science and Data Coordinator of the South African Environmental Observation Network (SAEON) Elwandle Coastal Node and the Shallow Marine and Coastal Research Infrastructure (SMCRI) initiative of South Africa. He has a PhD in Zooplankton Ecology and has been leading the Pelagic Ecosystem Long-term Ecological Research Programme of SAEON since 2010. His data team develops, implements and maintains data management workflows, standard operating procedures and web-based platforms for public serving of a range of ocean, biodiversity and remote sensing data and data products.
- **Sebastien Mancini (AODN-IMOS)**. Since August 2008, Sebastien Mancini has worked as a project officer, data services team leader and, more recently, as the Director of the Australian Ocean Data Network (AODN), which is in charge of the ingestion, curation and preservation of all data and relevant metadata collected by IMOS. During this period, he has contributed to the creation and improvement of the single integrative framework for data and information management that allow discovery and access of the data by scientists, managers and the public.

3.2.VA Expert Panel Task

The VA services were assigned to three different groups. VA services were grouped in a way that each group implied a similar amount of work to assess. Next, each group was randomly assigned to one internal expert and one external expert. This way, each VA service was assigned to two different members of the VA Expert Panel (one internal and one external).

Each panel member was sent the list of the VA services assigned to her/him and instructions on how to perform the assessments. These instructions included the context of the task and some subjects to consider to include in the assessment (although it was clear that these subjects were not compulsory and that other subjects could be added at their discretion). The suggested subjects were:

- Scientific Relevance
- Technical Implementation
- Scientific Usability/FAIRness
- Access Metrics
- Outreach Activities

3.3.Assessment of Mawenzi (ID 1.2)

Next, the description of Mawenzi, its assessment by Anca Hienola and its assessment by Antonio Novellino are included.

3.3.1. Description of the VA Service

ID	1.2
Name	Mawenzi
Institution	IFREMER, ULCO-LISIC
Short Description	Software packages written in R to help scientists, as well as stakeholders, in the modeling and interpretation of data (time series data from gliders/ferrybox or cytometry data), with some tutorials and user guidelines.
VA Services	The VA Infrastructure consists of the following R packages: DTWBI, DTWUMI, sClust and uHMMweb.

3.3.2. Assessment by Anca Hienola

Description

Software packages written in R

VA

DTWBI, DTWUMI, uHMMW

Scientific relevance

Software to fill large missing values in low/uncorrelated multivariate time series. The method solves a problem prevalent in many domains of signal processing and pattern recognition. Similar work does exist but, according to the authors, this method outperforms the state-of-the-art methods in the case of multivariate time series with low or non-correlated data but effective information on each signal.

Technical Implementation

The software is written in R – which I am not familiar with, and like me, many other researchers. As such, although the packages are openly available, they cannot be used by Python lovers, for instance. R and Python are the two most popular and powerful data science languages on the market and maybe it would be a good idea to have a Python version as well.

Although I was not able to check it due to lack of input data, uHMMW seems easy to use, with clear information for each step. Probably feedback from your regular and not so regular users could help you develop it even better.

Scientific Usability

There are clear usage licences for the packages. But only uHMMW contains instructions on how to cite, which the other packages need to add. There are examples of usage instead of a user manual. I am not sure if that is enough. FAIRness can be improved.

The DTWBI front page does not contain too much information about the software package. It should be brought at least to the level of and DTWUMI front page. Actually, both should be improved. The publications page should also contain DOIs for the scientific articles (or at least links to preprints) listed. Similarly, the conference presentations and seminars (abstracts, ppt, videos linked).

Access Metrics

Number of citations could be included, providing you offer a "how to cite" feature. Otherwise, there are not too many points in the graph to be able to judge the performance.

Outreach

In general, I wouldn't consider creating a documentation package as an outreach activity. Maybe more posters and presentations, even a training session could be included.

3.3.3. Assessment by Antonio Novellino to package DTWBI

Description

Method based on Dynamic Time Warping for imputation

Scientific relevance

The DTWBI is an R package to provide the user with advanced methods on Dynamic Time Warping. The tool is particularly interesting for filling timeseries that may have large gaps and it also provides a series of methods to assess and quantify the gap-filling accuracy.

DTWBI fills a scientific gap, and overlaps with other services. A paper presenting results of this tool versus other available tools, with a discussion of limits or better which datasets may best benefit from this gap-filling method, would be an excellent way to see if it can be improved. Moreover, this works (as far as I understood) with 1d timeseries.

A simple search in scholar[8] shows more than 21 000 results, indicating the relevance of the topic.

This package is specifically designed for R users, and it is easy to download and install, although documentation and examples may be improved. Beginners in R may appreciate more complete documentation and step by step guides.

Technical Implementation

[8] https://scholar.google.it/scholar?hl=it&as_sdt=0%2C5&as_vis=1&q=timeseries+gap+filling&btnG

This tool is well discoverable and has a very good SEO[9], especially its listing in CRAN[10].

Scientific Usability

There are potential improvements in the usage licence and citation instructions. A copyright string is presented, but it does not seem to be correct (usually a copyright is backward and this is indicating a forward coverage…). Usually, if a copyright is registered you have a code to mention (in Italy you can register copyrights @ SIAE – authorship register – or trademark or…). Rather than copyrights, here we are speaking about authorship credits. So, maybe better to indicate a paper/DOI so you can also track the impact.

There is a need for more step-by-step examples. That would improve the Accessibility in FAIR. Licence[11] (GNU GPL3 License) and citation[12] are presented.Citation may be improved and could be good to move to CC (CC-BY)[13]. For improvements in FAIRness, see the Metrics section below.

Access Metrics

Access metrics do not make sense for the relevance of the VA service and its history. They also suggest some improvements.

Access to Mawenzi software packages is monitored in two different ways: the CRAN downloads API and the log files of the documentation website.

One very important metric is the number of publications that are using the tool.

The website has a section[14] about it. These papers have some years nowadays and it would be good to complement/add (in a metric) the papers that are using the tool/VA

---

9 https://www.google.com/search?q=DTWBI&rlz=1C5CHFA_enIT967IT967&ei=RYQXYui0B5G4sAfew6noDA&ved=0ahUKEwio656wtZj2AhURHOwKHd5hCs0Q4dUDCA4&uact=5&oq=DTWBI&gs_lcp=Cgdnd3Mtd2l6EAMyDQguELEDEMcBEK8BEA0yBwgAELEDEA0yBAgAEA0yBAgAEA0yBAgAEA0yBAgAEA0yBwguELEDEA0yBAgAEA0yBAgAEA06CgguEMcBEK8BEA06BggAEA0QCjoFCAAQgAQ6CwguELEAEEMcBEK8BOgcIABCABBAKOgYIABANEB46CAgAEA0QChAeSgQIQRgBSgQIRhgAUNwCWJcNYOUOaAFwAHgAgAGYAYgBwwSSAQMxLjSYAQCgAQHAAQE&sclient=gws-wiz

10 https://cran.r-project.org/web/packages/DTWBI/index.html

11 https://mawenzi.univ-littoral.fr/DTWBI/license/

12 https://mawenzi.univ-littoral.fr/DTWBI/participants/

13 https://creativecommons.org/about/cclicenses/

14 https://mawenzi.univ-littoral.fr/DTWBI/publications/

Maybe to facilitate the tracking it would be useful to assign a DOI to the tool itself (e.g. Zenodo allows to assign a DOI to software packages). Anyhow a simple search in scholar[15] gives 23 hits.

Outreach

D11.1 reports on outreach activities: website

Outreach is probably the weakest point of this VA. More than the website it would be important to set up video tutorials and users channel/blog/mailing (e.g. GitHub)

3.3.4. Assessment by Antonio Novellino to package DTWUMI

Description

Method based on Dynamic Time Warping for imputation of Uncorrelated Multivariate Time series.

Scientific relevance

There are potential improvements in the usage licence and citation instructions. A copyright string is presented, but it does not seem to be correct (usually a copyright is backward and this is indicating a forward coverage…). Usually, if a copyright is registered you have a code to mention (in Italy you can register copyrights @ SIAE – authorship register – or trademark or…). Rather than copyrights, here we are speaking about authorship credits. So, maybe better to indicate a paper/DOI so you can also track the impact.

There is a need for more step-by-step examples. That would improve the Accessibility in FAIR.

The DTWUMI is a R package to provide the user with advanced methods on Dynamic Time Warping for imputation of Uncorrelated Multivariate Time series.

The tool is particularly interesting for filling timeseries which may have several gaps.

This package is specifically designed for R users, and it is easy to download and install, although documentation and examples may be improved. Beginners in R may appreciate a more complete documentation and step by step guide.

Technical Implementation

[15] https://scholar.google.com/scholar?hl=it&as_sdt=0%2C5&q=DTWBI&btnG=

This tool is well discoverable and has a very good SEO[16], especially its listing in CRAN[17].

It is a library and the user has to install R to use it.

Scientific Usability

There are potential improvements in the usage licence and citation instructions. A copyright string is presented, but it does not seem to be correct (usually a copyright is backward and this is indicating a forward coverage…). Usually, if a copyright is registered you have a code to mention (in Italy you can register copyrights @ SIAE – authorship register – or trademark or…). Rather than copyrights, here we are speaking about authorship credits. So, maybe better to indicate a paper/DOI so you can also track the impact.

There is a need for more step-by-step examples. That would improve the Accessibility in FAIR.

Licence[18] (GNU GPL3 License) and citation[19] are presented.Citation may be improved and could be good to move to CC (CC-BY)[20]. Improvement of FAIRness see Metrics

Access Metrics

Access metrics do not make sense for the relevance of the VA service and its history. They also suggest some improvements.

Access to Mawenzi software packages is monitored in two different ways: the CRAN downloads API and the log files of the documentation website.

One very important metrics is the number of publications that is using the tool

---

16 https://www.google.com/search?q=DTWUMI&rlz=1C5CHFA_enIT967IT967&ei=OYQXYvflN8rhgAbJ27mQBQ&ved=0ahUKEwi35vKqtZj2AhXKMMAKHcltDIIQ4dUDCA4&uact=5&oq=DTWUMI&gs_lcp=Cgdnd3Mtd2l6EAM6EQguEIAEELEDEIMBEMcBENEDOgsIABCABBCxAxCDAToICC4QsQMQgwE6BAgAEEM6BQgAEIAEOg4ILhCABBCxAxDHARCjAjoICAAQgAQQsQM6BwgAELEDEEM6CAguEIAEELEDOg4ILhCABBCxAxDHARDRAzoLCC4QgAQQxwEQrwE6BwgAEIAEEAo6BAgAEA06CgguEMcBEK8BEA06BggAEA0QCjoHCAAQsQMQDToHCC4QsQMQDUoECEEYAEoECEYYAFAAWLkbYNkgaABwAXgAgAF2iAHiBJIBAzMuM5gBAKABAcABAQ&sclient=gws-wiz

17 https://cran.r-project.org/web/packages/DTWUMI/index.html

18 https://mawenzi.univ-littoral.fr/ DTWUMI/license/

19 https://mawenzi.univ-littoral.fr/ DTWUMI/participants/

20 https://creativecommons.org/about/cclicenses/

The website has a section[21] about it. These papers have some years nowadays and would be good to complement/add (in a metric) the papers that are using the tool/VA

Maybe to facilitate the tracking it would be useful to assign a doi to the tool itself (e.g. Zenodo allows to assign a doi to sw package)

Anyhow a simple search in scholar[22] gives 11 hits.

Outreach

D11.1 reports on outreach activities: website

Outreach is probably the weakest point of this VA. More than the website it would be important to set up video tutorials and users channel/blog/mailing (e.g. github).

3.3.5. Assessment by Antonio Novellino to package uHMM

Description

Unsupervised learning - Hidden Markov Model - Spectral clustering - Harmful algal blooms - Application to phytoplankton.

Scientific relevance

There are potential improvements in the usage licence and citation instructions. A copyright string is presented, but it does not seem to be correct (usually a copyright is backward and this is indicating a forward coverage…). Usually, if a copyright is registered you have a code to mention (in Italy you can register copyrights @ SIAE – authorship register – or trademark or…). Rather than copyrights, here we are speaking about authorship credits. So, maybe better to indicate a paper/DOI so you can also track the impact.

There is a need for more step-by-step examples. That would improve the Accessibility in FAIR.

The UHMM is a R package to provide the user with advanced methods on Unsupervised learning - Hidden Markov Model - Spectral clustering - Harmful algal blooms - Application to phytoplankton.

HAB[23] is an increasing hot topic with many stakeholders.

[21] https://mawenzi.univ-littoral.fr/ DTWUMI/publications/

[22] https://scholar.google.com/scholar?hl=it&as_sdt=0%2C5&q=DTWUMI&btnG=

[23] Harmful algal blooms

This package is specifically designed for R users, and it is easy to download and install, although documentation and examples may be improved. Beginners in R may appreciate a more complete documentation and step by step guide.

The VA also offers a web application with a GUI clear and easy to use.

Technical Implementation

The user interface is good. However, SEO is low: you need to know the link to find it or you need to search for "mawenzi.univ-littoral.fr/uHMM/".

The tool is well discoverable/listed in R – related link. The algorithms are also usable via web interface. Documentation is well done.

Scientific Usability

There are potential improvements in the usage licence and citation instructions. A copyright string is presented, but it does not seem to be correct (usually a copyright is backward and this is indicating a forward coverage…). Usually, if a copyright is registered you have a code to mention (in Italy you can register copyrights @ SIAE – authorship register – or trademark or…). Rather than copyrights, here we are speaking about authorship credits. So, maybe better to indicate a paper/DOI so you can also track the impact.

There is a need for more step-by-step examples. That would improve the Accessibility in FAIR.

Licence[24] and citation are presented (MIT License). Citation may be improved and could be good to move to CC – XX. For improvement of FAIRness, see Metrics.

Access Metrics

Access metrics do not make sense for the software package[25] and they suggest improvements. On the other hand, for the web application[26], they do make sense and they do not suggest any improvements.

Access to Mawenzi software packages is monitored in two different ways: the CRAN downloads API and the log files of the documentation website.

[24] https://mawenzi.univ-littoral.fr/uHMM/license/

[25] https://mawenzi.univ-littoral.fr/uHMM/software/

[26] https://mawenzi.univ-littoral.fr/apps/uHMMweb/

One very important metric is the number of publications that are using the tool. The website has a section[27] on this which links to 2 papers. It would be good to complement/add (in a metric) the papers that are using the tool/VA.

Maybe, to facilitate the tracking, it would be useful to assign a DOI to the tool itself (e.g. Zenodo allows to assign a doi to sw package).

Outreach

D11.1 reports on outreach activities: website

Outreach is probably the weakest point of this VA. More than the website it would be important to set up video tutorials and users channel/blog/mailing (e.g. GitHub)

3.4. Assessment of EU HFR Node/AZTI (ID 4.1)

Next, the description of EU HFR Node/AZTI, its assessment by Simon Keeble and its assessment by Sebastien Mancini are included.

3.4.1. Description of the VA Service

ID	4.1
Name	EU HFR Node/AZTI
Institution	AZTI
Short Description	Open access to different softwares for processing and analysing coastal data. The tools aim to support sustainable management of the ocean and its resources.
VA Services	The VA Infrastructure consists of a Github repository[28].

3.4.2. Assessment by Simon Keeble

The JRadar Toolbox provides access to an open source git repository for file conversions to the NetCDF European HFR standard.

The README.md markdown file explaining the purpose of the repository is extremely minimal and would benefit from a better description and outlining of the tools. It is noted that a Readme.pdf file gives comprehensive instructions on the of the toolbox but this is not

[27] https://mawenzi.univ-littoral.fr/uHMM/publications/

[28] https://github.com/Fundacion-AZTI/JRadar

editable by any future contributors to the Git repository. Perhaps translation to markdown, or provision of the document original would be helpful.

The licensing 'appears' to be fully open source, but the documentation explicitly says that they cannot be used for commercial purposes under 'Creative Commons'. Is this true, that it cannot in any way be used commercially? Perhaps this needs some reflection and clarification. How was it funded and should it be made fully open access commercially as well?

There are no issues showing, only a single branch (master) and limited commits recently, so it is not possible to analyse how under development the tools are or if they are deemed complete. The number of watchers and forks are also low, which I would expect, so that doesn't give an indication of usage level. It would be good to be able to better understand the usage of the tools in general - How vital are they? What is the size of the user community? Is there a user community? Do the tools have competition or are they industry standard?

The metrics show some unique views of the page month by month but these are too low to interpret usage / impact. They refer to the number of times the repository pages have been viewed rather than tool usage, which could be significantly different. One would assume that once the tool has been installed and in use, there would be little need to revisit the repository regularly, no matter how often the toolbox is used.

A scan for inbound links to the repository did reveal that is is referenced by only two websites: JERICO-RI website[29] and EuroGOOS HF Radar Task Team newsletter[30].

It would be recommended to better advertise / market the availability of the tools for greater uptake and potentially more community contribution.

The toolbox appears to be extremely useful and an important one. Most of the comments above relate to the assessment of its usage, but it appears that the tool is of significant importance for the translation of data formats, and would be critical in the data processing production line.

Although not a code review, it is difficult to resist a quick scan of the codebase, which looks well constructed.

[29] https://www.jerico-ri.eu/va-service/jradar/

[30] https://us19.campaign-archive.com/?e=__test_emil__&u=e5fd08d8d94228eecb45183f9&id=1de066b71f

3.4.3. Assessment by Sebastien Mancini

The following Virtual Access (VA) service provides access to a GitHub repository[31] making available code and associated documentation for a Radar Toolbox. The Radar Toolbox is used to convert CODAR TUV and RUV native files to the netCDF European HFR Standard. This toolbox, provided by the European HFR Node, is developed to be run at an operator level.

Feedback about the features available in GitHub

The GitHub repository contains very minimal metadata on its main page. A PDF document named README.pdf is available for download and contains useful information about the toolbox. It would be very useful to the user to have access to more information on the initial page of the GitHub repository to provide more context about the tool such as, for example, "What is it used for", "How it has been developed" and useful links to other relevant resources.

Moreover, the GitHub repository does not take any advantage of the wiki functionality in the GitHub repository to provide additional information about the tools and the services it provides. The data provider of this VA service could look at other well maintained GitHub repositories in the marine community for example on how to use this functionality to provide more information. A useful example is the GitHub repository hosting the USIOOS Climate and Forecast netCDF checker[32].

Another feature of the GitHub repository that seems to not be used yet is the "Issues" tab. Currently, no issues have been reported about the tool to identify bugs or future improvements. The absence of issues might be explained by the fact that the toolbox has been published recently and the activity will increase as the number of data providers using the tool will increase over time.

It would be great if additional information is presented about the existing users of the tool. Are they only data providers from Europe or has it been used by groups in other countries?

The repository could also provide more information on the files produced by the Toolbox. For example, links to existing THREDDS catalogue or ERDDAP servers where netCDF files produced by the toolbox are made available to end-users.

Findability

It would be useful to create a metadata record describing the tool and publish it in a relevant metadata catalogue. It will increase the findability of the useful resource. My

[31] https://github.com/Fundacion-AZTI/JRadar

[32] https://github.com/ioos/compliance-checker

understanding is that the new infrastructure JERICO CORE will provide a mechanism to search and discover the various types of resources and improve the discoverability.

In searching other repository, I became aware of a DOI associated with the Radar toolbox and made available on the ZENODO platform[33]. A link to this resource should be made available from the GitHub repository.

Licensing

Some information on the license is available in the PDF document. It would be very useful to also include this information on the front page of the GitHub repository so that users are clearly made aware of the conditions associated with the resource.

The PDF document contains information on how to credit this resource but it does not provide a suggested citation for the tool. Users need to go to the associated DOI page on the ZENODO platform.

Feedback on the access metrics

The existing two metrics could be enhanced by including additional metrics on the number of users, number of countries and allowing a geographic analysis of where the end-users are coming from.

The metrics are highlighting peaks of downloads, perhaps related to end-user workshops being performed and new users becoming aware of the tool. It would be great to find a mechanism to correlate outreach activity and peak of downloads.

Outreach activity

The main users of the tool are data providers from the European HF radar network. Outreach activities of end-user workshops and news items in relevant newsletters are highly relevant to increase the exposure of the tool.

3.4.4. Response from EU HFR Node/AZTI to both assessments

Only one section has been created to respond to both reviewers as they pointed to similar aspects.

README.md will be updated to include more detailed information, such as: what the software it is used for, how it has been developed, more details about its dependencies, DOI and link to the European HFR Node tools, which also generate the same output using Matlab software. It will appear in the main page, making the information available at first sight.

[33] https://zenodo.org/record/5665528#.Yg8sQPhxVhE

Readme.pdf is not editable but any user can create a "pull request"[34] and we would update the pdf.

The licence of this VA is the CreativeCommons "Attribution-NonCommercial-ShareAlike 4.0 International" licence, because the main goal of the toolbox is to offer the HF Radar operators the possibility to generate their own standard HFR netcdf files, using open-source tools. In fact, the outputs from this toolbox (the .nc files) are affected by the CreativeCommons "Attribution 4.0 International" licence, which does not separate between commercial and non-commercial uses. Anyway, the concern of Simon Keeble is really interesting and the European HFR Node has opened an internal discussion about the suitability of the existing licence or the possibility to change it to another licence to make it fully open source.

There are only 2 releases available in the GitHub repository. The second one was created to update metadata convention attributes of the output files, to make them completely compliant with the Copernicus requirements. No issue has been reported since we recently uploaded the tool but we hope to have contributions while the number of users increase.

The metrics do not show the usage of the toolbox but only the visits and downloads of the toolbox in GitHub. When the toolbox is installed and running operationally in the HFR operator servers, 1 file per HFR radial site and 1 file per each total HFR system is generated every hour. A specific metadata value has been included in the output files to be able to track and count the files uploaded to the Internet from this toolbox. These metrics will be included in JERICO-CORE.

A DOI has been created for the tool[35] to make it more findable. The DOI will be included in the README.md file in github. The metadata of the output files of the tool already contain records that will be searched by JERICO-CORE infrastructure.

The toolbox has been promoted in the European HFR network meetings, workshops and other events related to the users. And we will continue with these outreach activities, to promote its use.

3.5. Assessment of CefMAT (ID 6.1)

Next, the description of CefMAT, its assessment by Simon Keeble and its assessment by Sebastien Mancini are included.

[34] https://github.com/Fundacion-AZTI/JRadar/pulls

[35] https://doi.org/10.5281/zenodo.6325552

3.5.1. Description of the VA Service

ID	6.1
Name	CefMAT
Institution	Cefas
Short Description	Access to marine environmental assessments to address policy questions. Users can efficiently produce high-quality visualisations and summary statistics tailored to their queries. It also gives access to specific products created under CMEMS and visualisation of aggregated DOI biological datasets from Cefas Data Hub.
VA Services	The VA Infrastructure consists of a website[36].

3.5.2. Assessment by Simon Keeble

CefMAT (The Cefas Marine Assessment Tool) provides scientists access to marine environmental assessments to address policy questions. Registered users can efficiently produce high quality visualisations and summary statistics tailored to their queries.

The landing page of the tools offer the chance to login, to already registered users, and also provides a link to the tools. I should note here that even on a 4k monitor, the Login and Register fonts are not very readable. Even blown up, they are not so clear.

Requesting access to the tools wasn't clear. After clicking on the 'Request access here' link, the following form was presented. I assumed that I needed to choose 'CefMAT Ticket Form' and follow that, but this could have said 'Request access' or something more intuitive. Without being able to log in, none of the data in the platform is available. Following completing this form, I did receive an email within an hour with a new link to register for the site. Perhaps this is the link that should be on the site? However, once completed, this was straightforward.

Once in the tool, it was fairly intuitive to be able to access the data. It is not a real-time data tool, but rather requires the user to submit requests and then view them once the outputs have been completed. Notifications are sent by email to save the user waiting.

Some pages did not load quickly, or occasionally not at all, but they worked after refreshing.

I could only load historic data up to 2014 as there isn't any assessment data which is more recent. This could be considered a weakness of the service.

[36] https://www.cefmat.org

Downloading of outputs took a long time but they did eventually download as a zip file.

The front end technology looks to be AngularJS based, with D3 for visualisations, and would appear relatively well built. There are some validation errors on the site. For example, alt tags missing on images making the images not accessible. There aren't a lot though.

Terms and conditions of use are clearly provided when logging in after registering.

Reviewing the analytics, I offer these thoughts:

1) The bounce rate is unusually high. Given that this is a data tool, I would expect users finding the site to be interested in spending time with the data.
2) The number of users seems appropriate, but this statistic presumably includes users who are not public? So, this statistic is not useful in assessing how the platform is used by people external to Cefas perhaps.
3) The device usage plot claims that access by Mobile / Tablet and Desktop users is equally spread daily. This cannot be true.

In general, Google Analytics is not a useful metric for a platform like this. It doesn't give an indication of the level of use, by people outside of Cefas, by product / service. It is a very rough guide of the number of users increasing / decreasing over time only.

The user experience could be better by making the querying activities interactive and not requiring users to wait for results. I couldn't find any way to access the data via an API to save time loading it through the user interface. Perhaps this isn't a requirement for assessment data though.

I found a number of inbound links to the tools on the internet, so it is clearly known about but I would be interested to know what the source of the traffic is generally.

Finally, I think it should be highlighted what the intended audience for the tools is. Given that most of the platform is not available publicly, who are the tools aimed at? This question is asked bearing in mind that the latest data point appears to be in 2014. I appreciate that this is assessment data rather than the latest observational data, so the intention is that it gives a view over a specific period of time. However, once many years have passed, its value diminishes until the latest assessments have been conducted? Its usefulness lies in comparison with the latest data? This is an observation rather than a criticism.

3.5.3. Assessment by Sebastien Mancini

The Cefas marine Assessment Tool provides scientists access to marine environmental assessments to address policy questions. Registered users can efficiently produce high-quality visualisations and summary statistics tailored to their queries.

The website provides a good description of the tools on offer. It states clearly that users must be logged in to the site to use the assessment tool. The registration process is

straightforward with minimal information requested. It only takes a couple of minutes to fill in and a confirmation e-mail is sent to your e-mail address before being able to access the service. During the initial connection to the service, the user needs to accept the "Terms and Conditions". These Terms and conditions are clearly defined so that users have a clear understanding of what they can or cannot do with the data and its derivative from the service.

Once you are logged in, users have access to several additional tabs to choose from.

The user guide provided on the website clearly explained step by step the functionalities of the tool. It is very simple and provides the right level of information to get users started.

As a user, I was able to access the "Atlantic Assessment tool". It was pretty easy to follow and select the different options available. The only minor issue was that the map was not loading properly in the first place. It was difficult to appreciate what was selected. But once the first query ran and I tried different options, the "area selection" displayed properly.

The different options available are for most of them self-explanatory except for the ones under the "area Layout" section. Additional information on the different acronyms used in this section would be useful (e.g. OSPAR, MSFD, ICES, …).

The additional information provided under the "Metadata" header is particularly useful to identify the source data and the processing that was applied to the data.

For most datasets, there is a link to the input dataset. The only exception I could find is for the "neam" product, the estimated fishing pressure is not available as a data source. Is there a particular reason why this product is not yet publicly available? Also, for the "neam" product, no map is made available to the end-user. Even if the spatial extent is fixed, I believe it is worth providing it to the end-user.

The output file is a PDF document that contains plots and tables with data values. It would be useful to provide users with access to the map in PNG format and tables with data in CSV format.

The "Terms and Conditions" could also be included in the ZIP file for each download.

The performance of the website was relatively good with results made available within a few minutes.

Technology

The website does not provide any information or references to the technology used to process the data and to create the data products. Other data providers might be interested in the technology used and the possibility to re-use some of the components and possibly

extend them? Another question: Are some of these scripts used to process the data publicly available in a GitHub repository?

The tool provides a nice user interface for users to select and define their queries.

A future improvement could be to provide offline access to the tool to "Power users', using for example a Web Processing Service, so that they can have access to the web service without having to go through the user interface.

It would also be useful to include a webpage mentioning any future developments and another webpage where users might be able to request new features.

Datasets

Is there a particular reason why most datasets only include data up to 2014? Are there any plans to include more recent data?

Access Metrics

The dashboard provides the typical information gathered by Google Analytics such as the number of users and the number of sessions.

It would be great if further analytics can be included such as, for example, for the Atlantic Assessment Tool, which parameters are accessed the most by users? Google Analytics might not be the most appropriate tool to gather and publish this type of information, but other tools exist to process information stored in logs.

It would be interesting to derive statistics on how many users are using the tool on a regular basis (e.g. weekly, monthly, several times a year…).

Outreach activities

Targeted workshops are ideal to increase the visibility of this type of tool but also to gather feedback from end-users on future developments.

3.5.4. Response from CefMAT to both assessments

Accessibility

We are fully aware that some modifications are needed in terms of visibility (e.g. fonts), functionality (e.g. delay to download the map). We are also investigating new interactive facilities for queries. Hopefully this will be modified soon for better access to information.

Performance

CefMAT is a tool for assessing the marine environment for MSFD and not a data portal. For this reason, only the results from the last assessment which took place in 2014 are available. More recent results should be available soon because of new assessments coming soon. We will modify the description on CefMAT on the website to give more clarity

on the functionality of CefMAT, the targeted audience, and the future development as suggested.

Terms and conditions

The terms and conditions are accepted when you sign in. We think that adding on a ZIP file will be duplication.

Analytical tool

We are aware that Google Analytics is not the most performant for this platform. Changes in CefMAT have happened the last few months but for consistency of the analysis we thought to continue with the way CefMAT has been monitored from the beginning.

Outreach activities

The suggestions are excellent and will try to follow the recommendations.

3.6. Assessment of CNR TirLig e-infrastructure (ID 7.1)

Next, the description of CNR TirLig e-infrastructure, its assessment by Thierry Carval and its assessment by Shaun Deyzel are included.

3.6.1. Description of the VA Service

ID	7.1
Name	CNR TirLig
Institution	CNR
Short Description	Provide access and distribution of sea surface current data derived from High Frequency Radar acquisitions.
VA Services	The VA Infrastructure consists of a THREDDS data server[37] containing HF Radar data of the Ligurian coast near La Spezia and Cinque Terre and a website[38] with visualisations of that data and links to the mentioned THREDDS data server.

[37] http://150.145.136.27:8080/thredds/HF_RADAR/TirLig/TirLig_catalog.html

[38] http://radarhf.ismar.cnr.it/

3.6.2. Assessment by Thierry Carval

A brief summary of usage statistics

Processing: on the last 12 months: 0,7 billion of processes, 467 years of processing time
There are about 1 million files generated per month.

Software: there are about 40 viewers per month

Thredds: the statistics stopped in January 2022
In 2021, there were 16 billion visits, 469 users from 27 countries
The majority of the traffic is HTTPS, followed by catalogue views
Users are well geographically distributed: Italy, Spain, Europe, North America, China

Web: in the last 12 months, there are about 20.000 pages viewed and 1900 sessions.
A majority from Italy and Spain and good geographical distribution from the remaining users (Europe, America, Asia and then Africa).

Scientific relevance

This is the e-infrastructure of EU HF radars. It organises and performs the real-time and delayed mode data distribution from HF radars. The statistics number shows a huge amount of data processing and significant data distribution (500 users and 16 billion visits).

Technical Implementation

The HFR data production is massive (all EU HFR), the data files are rich and FAIR NetCDF files, compliant with Copernicus Marine and SeaDataNet standards and formats.

Scientific Usability

The HF-Radars terms of use are not mentioned on the website. A possible improvement would be to provide a data licence, such as CC-BY[39] that is human and machine-readable.

To improve its FAIRness on machine-to-machine services, the creation of a DOI for HF-Radar datasets may be beneficial. This DOI would ease the citation, the terms of use (licence) and reproducibility.

HF-Radars NetCDF uses Copernicus Marine standards and vocabularies (CF, SeaDataNet): an excellent solution for interoperability.

The HF-Radars NetCDF files have a clear "citation" statement.

[39] https://creativecommons.org/licenses/by/4.0/

Access Metrics

The THREDDS metrics are not updated since December 2021. Most of the thredds traffic is from HTTPS requests.

An additional data downloads and users distribution would be interesting (without the protocols or geographics subtotals)

Outreach

CNR TirLig contributed to an HF radar summer school in May 2021. A yearly summer school in Europe may be organised.

The website may be improved with specific outreach content (such as presentations, posters, summer school, training sessions).

3.6.3. Response from CNR TirLig e-Infrastructure to the Assessment by Thierry Carval

All HFR netCDF data files contain the global attribute "license" reporting the data license as *"HF radar sea surface current velocity dataset by CNR-ISMAR is licensed under a Creative Commons Attribution 4.0 International License. You should have received a copy of the license along with this work. If not, see http://creativecommons.org/licenses/by/4.0/."* The catalogue page reports as well "*rights: The dataset is licensed under a Creative Commons Attribution 4.0"*

3.6.4. Assessment by Shaun Deyzel

Summary of usage statistics accessed via VAMS dashboards (for 12 months ending May 2022)

Processing: 662,451 processes; 145,7,6GB input files; 203GB output files with 506 processing years.

Software: considerable variation in the number of viewers per month; HFR_Node_tools example fluxing Mar-Apri-May (2022) 47 to 9 and 15 respectively.

Thredds: 47,072,099 visits recorded with up to 1TB data downloaded by 485 users representing 32 countries. Peak downloads from HTTPServer occurred in March 2022 (220.1GB). During the 12-month reporting period, the majority of downloads were made via HTTPServer (>600GB), followed by the Catalog (>250GB) and OPENDAP (<50GB). However, the number of Catalog users (405) far exceeded that of the HTTPServer (52). With the exception of Australia (1 user), the majority of users accessed data from the Northern Hemisphere, with the majority of traffic being from Italy (>13 million visits in April 2022).

Web: 18,051 page views from 2,729 sessions. Once more the majority of these originated from the Northern Hemisphere, particularly Italy and Spain. With the exception of

December 2021, the proportion of new users exceeded those returning (overall 72% vs. 27%).

Scientific relevance

High-Frequency Radar data was acquired from an operating network of two radar stations along the coast of Eastern Liguria, Italy. CNR TirLig serves radar data (real-time and delayed) via a Thredds server and map visualisations via a website. I am not aware of overlap in this sector, however, merely judging by the number of new and returning users and downloads, this resource is indeed receiving considerable interest and issuing large amounts of data.

Technical implications

Noted no log-in requirements. The HFRadar is technically quite simple by design with four tabs, three of which contain project information. Direct links to the thredds catalogue and historical data and maps is shown on the landing page for quick access. Unable to comment on performance and SEO.

Scientific Usability:

I noted a user licence (Creative Commons Attribution 4.0) for the dataset, along with citation text, listed among the Catalog documentation[40], but not on the HFRadar website[41]. A short description outlining the terms of use, applicable licences and citation instructions would be useful for viewing on the HFRadar website.

Access Metrics:

The access metrics for the VA is well developed and varied with subset dashboards for processing, software, Thredds and web metrics. Results can be queried by defined time ranges, which is excellent for refining targeted reporting figures. Software page view data only start in June 2021, which compared to the other subsets offer limited insight into historical information. I suggest adding to the metrics tab of the Web subset, the total number of users and countries.

3.6.5. Response from CNR TirLig e-Infrastructure to the Assessment by Shaun Deyzel

The network is made of five radar stations along the whole coast of Liguria, not just the Eastern part. Two more stations will be installed by the end of the year, bringing the number of stations to seven.

[40] http://150.145.136.27:8080/thredds/HF_RADAR/TirLig/Totals/Aggregated/TirLig_totals_version22_aggr_catalog.html?dataset=CROS_HFRADAR_TirLig_Aggr_NRT_v2.2_Totals

[41] http://radarhf.ismar.cnr.it

3.7. Assessment of EU HFR Node/CNR (ID 7.2)

Next, the description of EU HFR Node/CNR, its assessment by Anca Hienola and its assessment by Antonio Novellino are included.

3.7.1. Description of the VA Service

ID	7.2
Name	EU HFR Node/CNR
Institution	CNR
Short Description	The EU (European) HFR (High Frequency Radar) Node/CNR is the contribution of CNR to the EU HFR Node Competence Center for HFradar data management.
VA Services	The VA Infrastructure consists of a software package written in MATLAB [42] and a remote processing service[43] that uses the software to process HF Radar files.

3.7.2. Assessment by Anca Hienola

Description

Open access to different software for processing and analysing coastal data.

VA

GitHub repository containing Matlab scripts for processing historical data within the workflow of the European HFR Node. Tools for HFR data providers.

Scientific relevance

The VA contains Matlab scripts for data providers which, by themselves, represent good coding practices that ensure reproducibility and fluency and help providers simplify the operations that are involved in supplying data. As such, the need for such tools is unquestionable from both practical and scientific points of view.

Technical Implementation

MATLAB requires expensive licences that put it outside the reach of people outside academia. A very large code and support base has evolved around MATLAB, producing a vast well of knowledge, but is effectively sealed off from everyday people, making it a barrier for scientific openness and advancement. As MATLAB can easily be replaced by

[42]https://github.com/LorenzoCorgnati/HFR_Node__Historical_Data_Processing

[43] http://150.145.136.36/

faster, better and free alternatives like Python (although utility-wise MATLAB includes more signal processing and non-linear optimization toolkits than Python), the idea of moving over to or adding another development environment (Python, that is) that is already significantly more performant and widely adopted at large than MATLAB could increase the accessibility and usability of the VA and its contents.

Scientific Usability

There is a clear usage licence, but there are no instructions on how to cite.

I cannot comment on the performance and easiness of use of the scripts themselves as, since moving out of academia, I do not hold a MATLAB licence anymore. However, I have tried to read some of the scripts and I think the README file should be developed further with clear step by step instructions for newcomers.

The service is interoperable, but the R in FAIR can be improved by developing codes in other environments that MATLAB and better "how to use" instructions

Access Metrics

The access metrics for this particular VA include Page Views vs Time by Package and Unique Page Views vs Time by Package. I would suggest including also the number of downloads and number of users (if feasible).

Outreach

Only one activity is listed under this VA, that is *Sicomar-Plus HF Radars Summer School*. Depending on the demand, one or two similar sessions per year might suffice. Also, training material, user guides, videos can be included in the webpage with links to and from the repository.

3.7.3. Assessment by Antonio Novellino to the Software Package

Description

It's a collection of software tools to process HFR data and make radial and total velocity files in netCDF format according to the European standard data and metadata model for near real-time HFR current data.

Scientific relevance

The tool fills a scientific gap and does not overlap other services. It is designed for historical data, but an open tool to facilitate operators to provide nrt would be very useful, it's in Matlab only and this is a strong limitation, open access tools (Python, R…) should be considered.

The tools are very valuable and help HFR data providers to organise data and publish them in netCDF transport files organised according to the European standards: a community's Milestone.

The main limitation is that it's a MATLAB tool (and MATLAB is not always accessible). A Python version (maybe a Docker-Python version) would be a big milestone.

Technical Implementation

The SEO is good.A simple search links directly to the Corgnati Page

Scientific Usability

A clear usage license and instructions on how to cite are missing. The GitHub page does a good job as a user manual. This tool serves the HFR providers to make their data available in common formats (a piece of interoperability) but FAIRness could be improved to some degree: it's in Matlab only and this is a strong limitation, open access tools (Python, R…) should be considered.

Access Metrics

The software package is monitored using the views at the GitHub repository, the processing service is monitored using the event data stored in a SQL

Although the defined metrics make sense, the number of HFR operators is not increasing on a daily basis, we know there are about 30 operational antennas in Europe and the proposed processing is done at HFR node level (one place) – in the best case at each provider place.

It would make more sense to track the number of end-users who may have had access to HFR data thanks to this tool.

Concerning the software, maybe to facilitate the tracking it would be useful to assign a DOI to the tool itself.

Outreach

Reported outreach is very limited, it's one single event (Interreg project). Another limitation is that such a tool cannot be (only) disseminated under a personal page.

3.7.4. Assessment by Antonio Novellino to the Processing Service

Description

It is the HFR mapping questionnaire for making/updating HFR asset mapping.

Scientific relevance

The service perfectly fills a scientific gap without overlapping other existing services.This tool is very important and useful.

Technical Implementation

The SEO is very low – if you do not know the link you won't get there. However, the user interface is simple while collecting all the needed information. Registered users can fill the form, there is a very basic login.

Scientific Usability

A user manual is missing.

Access Metrics

It may make sense to count the registered users and how often they update the information. I do not think it makes sense to use any metrics on this.

Outreach

This tool is very important for mapping new sources and serves many marine data infrastructure communities, outreach should be a main concern for this tool. The web page should go under an easily searchable domain and should be linked by EuroGOOS, JERICO, EMODnet, CMEMS, SeaDataNet … to facilitate new operators to find and fill it.

And a link to this page should be in any HFR scientific and technical communication.

3.7.5. Response from EU HFR Node/CNR to both assessments

To avoid the limitation of having operational scripts only in Matlab, the EU HFR Node team is in the process of porting all the EU HFR NODE tools in Python. This task should be accomplished by mid 2023 (hopefully by the end of 2022).

Concerning Antonio's review, it only relates to the Historical Data Processing package, but there is also a specific package dedicated to providers for NRT operations: https://github.com/LorenzoCorgnati/HFR_Node_tools

For sake of completeness, in the following, all the repositories related to the EU HFR NODE data processing are reported:

- https://github.com/LorenzoCorgnati/HFR_Node__Centralized_Processing
- https://github.com/LorenzoCorgnati/HFR_Node_tools
- https://github.com/LorenzoCorgnati/HFR_Node__Historical_Data_Processing
- https://github.com/LorenzoCorgnati/HFR_Node__REP_Temporal_Aggregation
- https://github.com/LorenzoCorgnati/HFR_Node__SDC_Processing
- https://github.com/LorenzoCorgnati/EU_HFR_Node__WebForm

Furthermore, Antonio points out that there is no usage licence: this is not true, since each of the repositories includes a licence file (GNU General Public License v3.0).

Concerning the "how to cite" issue, it's a real lack. The EU HFR Node team just added the "how to cite" instructions to the README files of all repositories. Related to this point, Antonio states that DOI should be assigned to the code: it's already like this. Every

software package has DOI-assigned releases, accessible from the GitHub repositories. Now the DOIs are mentioned in the "how to cite" instructions. Of course, the mentioned DOIs are the cumulative ones, i.e. they list all the release versions.

As Antonio points out, these are tools that are mainly run at the EU HFR NODE level, thus there's no meaningful traffic in terms of user access and download. But, the VA metrics also takes into account the datasets produced by these tools, so the request of Antonio is fulfilled.

The EU HFR Node team agrees with Antonio's statement that these resources should not only be disseminated via a personal page: the EuroGOOS HFR Task Team is organising an official GitHub profile for the Task Team that will include all the considered repos. This task should be accomplished by the end of 2022.

Concerning the webform (http://150.145.136.36/), the EU HFR Node team assigned an alias to be propagated via DHCP instead of plain IP: http://webform.hfrnode.eu

The webform has a user manual, accessible at https://cnrsc-my.sharepoint.com/:b:/g/personal/lorenzo_corgnati_cnr_it/Ecoa9zTYUmVNiXuDxiKWslQBqwbtuG3r1RVwdI5Dy-E9zg

The EU HFR Node team just completed the linking of the manual to the homepage of the webform.

According to Antonio's statement that "it may make sense to count the registered users and how often they update the information, I do not think it makes sense to use any metrics on this", the EU HFR Node team is collecting this information and will work to include it in VAMS' dashboard.

The EU HFR Node team agrees with Antonio's statement "The web page should go under an easily searchable domain and should be linked by EuroGOOS, JERICO, EMODnet, CMEMS, SeaDataNet … to facilitate new operators to find and fill". The team will work on having the link to the webform on the suggested portals. In recent weeks, the link was added on the JERICO website[44] and JERICO-CORE, when it is operational, will also provide this link.

Concerning Anca's review, the ongoing task of porting the tools in Python and the addition of the "how to cite instructions" were mentioned above.

[44] https://www.jerico-ri.eu/va-service/eu-hfr-node-cnr-einfrastructure/

3.8. Assessment of CytoFluoTool (ID 8.1)

Next, the description of CytoFluoTool, its assessment by Simon Keeble and its assessment by Sebastien Mancini are included.

3.8.1. Description of the VA Service

ID	8.1
Name	CytoFluoTool
Institution	CNRS-LOG, ULCO
Short Description	Automated techniques, and complementary to automated image analysis, automated flow cytometry (FCM) and multispectral fluorometry (MSF).
VA Services	The VA Infrastructure consists of a software package written in R (RclusTool)

3.8.2. Assessment by Simon Keeble

RclusTool is a Clustering and visualisation R toolbox. Visualisation and Processing of data with different formats: profile/time series, features and images, Unsupervised Clustering, Semi-supervised Clustering, Supervised Classification, Labelling Expert Interface, Constraint Expert Labelling (matching without specific name: points must be link or can not link).

The online documentation is limited / virtually non-existent beyond installation instructions in the tool's web page. Instead, providing the contact details of experts to contact with questions. Full documentation, however, was found in the Cran repository.

The tools appear to be well utilised within a community, contributed to by a respected community of developers and also funded by a respected network, including JERICO.

The download statistics are interesting and show a high number of downloads, presumably from the Cran repository statistics. This is positive, and leads to the natural question relating to who are the users, as there are so many downloads? The statistics relating to the website of the tool show a much lower level of interest, but this is perhaps expected as it's not the main resource for the tools.

I found no reference to version control via Git, or similar. Perhaps this exists?

The licensing is clear and definitely open source, for any purpose etc, which is good.

3.8.3. Assessment by Sebastien Mancini

The RclusTool is a clustering and visualisation R toolbox, also available on the CRAN repositories.

The RclusTool R package is available on the following website: https://mawenzi.univ-littoral.fr/RclusTool/.

The website contains general information about the R package, how to access it and how to install it which most R users will be familiar with.

The R package is also on the CRAN repository[45].

The website contains clear documentation on the licence for the tool.

There is little documentation on what the R package is offering and the type of processing that can be performed. Several research articles are listed in the documentation.

It would be useful to provide some examples of how the tool can be used. For example, Jupyter Notebooks are a popular way to demonstrate the use of a tool through one or two real-life examples.

Feedback on access metrics

The ability to display two dashboards, one for the downloads through the CRAN API and a second one for the access of the documentation website is definitely appropriate.

It would be interesting to include the number of users that are actually clicking Downloads from the website.

Outreach activities

Small workshops are definitely appropriate to promote the tool and increase awareness. Inclusion in teaching materials is also a good way of promoting this type of tool.

It could be interesting to include a metadata record in an existing repository to describe the package. This metadata record could be harvested by other data repositories and will increase its discoverability.

[45] https://cran.r-project.org/web/packages/RclusTool/index.html

3.9. Assessment of JERICO-ECOTAXA (ID 8.2)

Next, the description of JERICO-ECOTAXA, its assessment by Thierry Carval and its assessment by Shaun Deyzel are included.

3.9.1. Description of the VA Service

ID	8.2
Name	Ecotaxa
Institution	CNRS-LOV
Short Description	Web application that provides services for users to handle large image datasets and their associated metadata. System suited for imaging sensors such as UVP5, ZOOSCAN, ZOOCAM, FLOWCAM, IFCB, microscopic imaging, and confocal microscopy, that are used in marine sciences.
VA Services	The VA Infrastructure consists of a website[46].
Monitoring Method	EcoTaxa has its own metrics system that synchronizes to VAMS via a custom CSV file
Metrics	EcoTaxa metrics can be found in Annexe 1.7.
Funders	OCEANOMICS[47] Partner University Fund[48] CNRS LEFE program[49] Belmont Forum WWW.PIC project[50] H2020 Blue-Cloud project[51]

[46] https://ecotaxa.obs-vlfr.fr/

[47] http://www.oceanomics.eu/

[48] https://face-foundation.org/higher-education/partner-university-fund/

[49] https://programmes.insu.cnrs.fr/lefe/

[50] https://www.belmontforum.org/projects/world-wide-web-of-plankton-image-curation/

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

	JERICO-S3[52]

3.9.2. Assessment by Thierry Carval

A brief summary of usage statistics

In February 2022 EcoTaxa stored 192 million objects and 76 million validated objects.

More than 150 000 pages views from 20 000 sessions.

The number of distinct users is not reported.

In the last 12 months the number of objects grew by 20 000, the number of validated objects grew by 11 200. Those numbers are impressive: more than a 20% increase a year.

The number of monthly page views is on average 3 000 for Jerico (10 000 total)

The number of monthly sessions is on average 400 for Jerico (1 400 total)

Scientific relevance

EcoTaxa is a major tool to identify marine species. Any interested party can access and use the EcoTaxa "virtual laboratory". As a leader in its domain, it is not overlapping other similar services.

Technical Implementation

During the last 12 months, the number of objects and validated objects increased regularly. There are 1 400 monthly sessions. The system is therefore used in a sustained manner, its user interface is efficient to support that many sessions.

Scientific Usability

Some EcoTaxa data base used for machine learning are published with a DOI such as https://doi.org/10.17882/55741 or https://doi.org/10.17882/73002

A DOI ensures that data are Findable (indexed by Google, google dataset search and other search engines). It is Accessible with a CC-BY-NC licence (machine- and human-readable).

It is Interoperable as it uses well-defined vocabularies, in particular the UniEuk taxonomy framework for eukaryotes.

[52] https://www.jerico-ri.eu/

An improvement for Reproducibility would be to archive (long term) and provide DOIs to the successive versions of the database with data publishers such as Zenodo or Seanoe.

Access Metrics

The metrics provide a good overview of EcoTaxa growth (content and users). Additional statistics on distinct users should be added to the dashboard. The bounce rate session is 50% illustrating the good findability of EcoTaxa (50% of access comes from search engines)

Outreach

From its website, Ecotaxa is clearly open to new users, any interested scientist can create an account and use or contribute to Ecotaxa.

In 2021, Ecotaxa was presented at 2 conferences and in a workshop.

The website may be improved with specific outreach content (such as presentations, posters, training sessions).

3.9.3. Response from JERICO-ECOTAXA to the Assessment by Thierry Carval

There seems to be an error in the growth of the number of objects reported in Thierry Carval's assessment. The number of objects probably grew by ~20M in 12 months. As stated in the assessment, that corresponds to an important percentage of the total number.

It is also important to clarify that the number of users and sessions referred to as "for JERICO" in the assessment above is not "for Jerico" but rather "for countries involved in JERICO". We cannot provide stats only for data related to JERICO since those projects are not identified in EcoTaxa.

Two DOIs are mentioned in Thierry Carval's text. Now, there are also several data sets meant for scientific exploitation published with a DOI[53].

[53]
https://www.emodnet-biology.eu/data-catalog?module=dataset&dasid=6505
https://www.emodnet-biology.eu/data-catalog?module=dataset&dasid=6506
https://www.emodnet-biology.eu/data-catalog?module=dataset&dasid=6507
https://www.emodnet-biology.eu/data-catalog?module=dataset&dasid=6508
https://www.emodnet-biology.eu/data-catalog?module=dataset&dasid=6510
https://www.emodnet-biology.eu/data-catalog?module=dataset&dasid=6511
https://www.emodnet-biology.eu/data-catalog?module=dataset&dasid=6512
https://www.emodnet-biology.eu/data-catalog?module=dataset&dasid=6513
https://www.emodnet-biology.eu/data-catalog?module=dataset&dasid=6514
https://www.emodnet-biology.eu/data-catalog?module=dataset&dasid=6515

Referring to the comment of minting DOIs for the "database", it is not ours to archive. Each piece of data (project) in EcoTaxa is the property of its project manager who can choose, or not, to publish it.

3.9.4. Assessment by Shaun Deyzel

Summary of statistics (last 1 year ending May 2022 to test filter)

In total 209,657,431 objects listed, of which 87,426.044 are validated. Over the last 12 months (ending May 2022), 160,740 page views and 21,551 sessions have been recorded. Growth in terms of objects is steady, with a 32% and 36% increase in verified and unverified objects respectively, from May 2021 to May 2022. Page views by Jerico Countries still outnumber those from outside the Jerico region by 2.5:1.

Scientific relevance

EcoTaxa is a web application for storage and visual exploration of images captured by an array of compatible instruments with the intention of taxonomic identification of plankton species (and/or particles). The centralisation of objects enhances collaboration within and between projects and institutions.

Technical Implementation

Although a log-in/registration facility is shown on the landing page of the EcoTaxa website, parts of the database can be navigated and verified images can be viewed by visitors. This mode of access is free and indeed EcoTaxa issues a no responsibility disclaimer on their website landing page, regarding the safety of images and data archived with the application. SEO is good, a search for EcoTaxa returns the site at the top of the list.

Scientific Usability

A clear guideline for citation is given on the landing page of the EcoTaxa website. A link to the Quantitative Imagery Platform of Villefranche[54] is shown on the landing page of the EcoTaxa website, that inter alia leads to additional documentation, manuals, literature and a link for registration to training sessions.

Access Metrics

Metrics are limited to the number of objects, page views, sessions and bounce rate. Implementing tracking code with Google Analytics will allow traffic to be assessed by geographical location, operating system and more. This may in turn assist with directing outreach efforts. It was noted that a user is able to turn off Google Tracking on the Privacy page, which then limits these additional information bits to be gathered, unless by choice such information is not channelled to the VAMS dashboards.

[54] https://sites.google.com/view/piqv

3.10. Assessment of Utö Atmospheric and Marine Research Station (ID 15.1)

Next, the description of Utö Atmospheric and Marine Research Station, its assessment by Simon Keeble and its assessment by Sebastien Mancini are included.

3.10.1. Description of the VA Service

ID	15.1
Name	Utö Atmospheric and Marine Research Station
Institution	FMI
Short Description	Software for instrument-computer interfaces, automated warning message system, data transmission systems, basic QC processes, data visualisation tools and a www-page open for public use.
VA Services	The VA Infrastructure consists of a website[55].
Monitoring Method	The website is monitored using Google Analytics API.
Metrics	Utö metrics can be found in Annexe 1.8.

3.10.2. Assessment by Simon Keeble

The Utö Atmospheric and Marine Research Station provides real-time data measured at the Utö Atmospheric and Marine Research Station located on Utö island at the outer edge of the Finnish archipelago.

The site provides up to date wind, wave, current, chlorophyll, sea level pCO2 and marine data.

Access to the data is intuitive and the plots load extremely fast. It is possible to retrieve some of the latest data through the service (the last two weeks) and download as CSV files. It is not clear where to retrieve any historic data. It would be good to see on the site where users can obtain earlier data.

The plots are static rather than interactive.

[55] http://swell.fmi.fi/Uto/

The site shows the JERICO-NEXT logo, it would be good to update this to include and recognise JERICO-S3 also.

The site is using Google Analytics, but there is no privacy or cookie information on the website - no opportunity for a user to opt out of tracking.

I couldn't find any clear explanation of how the data may be used.

The service is, I have assumed, a very useful one for a number of reasons. The visitor statistics are impressive and this is obviously an important and well referenced resource. It would be interesting to know some of the use cases for people using the platform.

3.10.3. Assessment by Sebastien Mancini

The website[56] contains real-time data measured at the Utö Atmospheric and Marine Research Station located on Utö Island at the outer edge of the Finnish archipelago.

The "About" tab provides useful contextual information and provides links to other websites where additional information is useful. One of these links to a webpage hosted by the Finnish Meteorological Institute[57] provides additional information regarding the parameters measured at the station.

The existing web pages do not provide any details on the appropriate acknowledgement and citation. Information is only available in one of the additional links[58].

The website provides graphs for different variables at different timescales. The data for the past two weeks can be downloaded in CSV format. The CSV files contain information about the parameters and their associated units of measure. A future improvement would be to provide links to existing vocabularies such as, for example, P01 BODC Parameter Usage Vocabulary.

Downloading only the last 2 weeks of data is a current limitation. On the JERICO website, it is mentioned that the data is made available through existing platforms such as, for example, the EMODNET Physics infrastructure. Is that correct? If yes, to increase discoverability and ease of access, it would be great if a link to the EMODNET Portal was provided on the website. Is the existing data made available via web services (e.g. OPeNDAP, OGC Web Map Service, OGC Web Feature Service…)?

[56] http://swell.fmi.fi/Uto/about.html

[57] https://en.ilmatieteenlaitos.fi/uto-observations

[58] https://en.ilmatieteenlaitos.fi/open-data

Moreover, it would be useful if a metadata record was created in a relevant repository to describe the entire dataset collection and act as a reference point for other metadata catalogues.

Information about the Quality Control procedures applied to the data was not easily accessible. It would be useful for the user if links to more information were provided.

Feedback on access metrics

The existing suite of access metrics is appropriate.

It might be a good opportunity to include another metric to identify the most viewed webpage on the website.

The number of downloads per month and information about the parameter that is being primarily accessed by end-user would be a great addition to the existing access metrics.

If the data is available on the EMODNET Physics infrastructure, it would be useful to include download statistics from this platform into the existing dashboard.

3.11. Assessment of POSEIDON Multi Platform Observatory Data Center (ID 16.1)

Next, the description of POSEIDON Multi Platform Observatory Data Center, its assessment by Thierry Carval and its assessment by Shaun Deyzel are included.

3.11.1. Description of the VA Service

ID	16.1
Name	POSEIDON Multi platform observatory Data Center
Institution	HCMR
Short Description	Integrated observatory located in the Eastern Mediterranean, which has adopted a multiplatform-multiparameter approach with the current system's status including open and coastal sea fixed platforms, deep-ocean observatories, a Ferrybox system, glider missions and Argo profiling floats.
VA Services	The VA Infrastructure consists of a website[59] and an API[60].

[59] https://poseidon.hcmr.gr/

[60] https://api.poseidon.hcmr.gr/swagger/

3.11.2. Assessment by Thierry Carval

A brief summary of usage statistics

In February 2022, for the last 12 months, Poseidon website reported 570 million page views and 12 million sessions.

The number of distinct users is not reported.

The number of new users is about 50% per month.

Two thirds of users access from a telephone (1/3 from desktop)

The majority of users are from Greece and Turkey

The Poseidon API shows an impressive number of users: 100 per month from 46 countries. The vast majority of API users come from Greece, followed by good numbers from USA and Turkey. A majority of users arrive on the API from the HCMR website.

Scientific relevance

Poseidon provides long series of observations of the Greek seawater. This national service for ocean observations is essential to detect and understand environmental trends (in the short, medium and long term).

Technical Implementation

During the last 12 months, the number of sessions increased (with spikes of frequentation in October and January). There are a huge 800 000 monthly sessions. The system is therefore well used, its user interface is efficient to support that many sessions.

To download data, registration is needed. There is no GDPR (General Data Protection Regulation) mentioned when logging or downloading. It should probably be added.

Scientific Usability

The Poseidon terms of use are not explicitly displayed on the website. There is a mention "For authorization credentials please contact poseidon@hcmr.gr with a short description of the intended data use!".

A possible improvement would be to provide a data licence, such as CC-BY[61] that is human- and machine-readable.

[61] https://creativecommons.org/licenses/by/4.0/

To improve its FAIRness on machine-to-machine services, the creation of a DOI for Poseidon dataset may be beneficial. This DOI would ease the citation, the terms of use (licence) and reproducibility.

Poseidon NetCDF uses Copernicus Marine standards and vocabularies (CF, SeaDataNet): an excellent solution for interoperability.

The Poseidon NetCDF files have a clear "citation" statement.

Access Metrics

The metrics provide a good overview of Poseidon growth (content and users). Additional statistics on distinct users should be added in the dashboard: a number of 700 000 in February is probably not the number of users but the number of pages.

The bounce rate session is 50% and illustrates good findability of Poseidon data (50% of access comes from queries in search engines).

Outreach

There is a good outreach activity reported on Poseidon website, with two chapters: News and Media.

The number of News per year: 4 in 2021, 11 in 2020, 4 in 2019.

In the media section, there are pictures of the 10 pieces of observation equipment managed by Poseidon.

In 2021, the D11.1 mentions a good activity: 2 website updates, one participation to a conference and two press releases.

3.11.3. Response from POSEIDON Multi Platform Observatory Data Center to the Assessment by Thierry Carval

We have checked the review provided by Thierry Carval. All comments are very constructive and we are going to take them into consideration in our future updates.

The part concerning GDPR should be addressed in the registration form of the download service for the in situ data and the API swagger documentation. Relative information is already available in the Poseidon website overall cookies policy[62], but will be addressed in the separate interfaces as well per registration request.

Regarding the statistics, it was not easy to cross-check the numbers with either the VA or our Google Analytics (GA) reports since the exact time period, that corresponds to the

[62] https://poseidon.hcmr.gr/cookies-policy

values in the comments, was not provided. Approximately, they make sense. On the expert's doubt that "*a number of 700 000 in February is probably not the number of users but the number of pages*", we would like to comment that it is indeed the number of users. We guess that since users statistics are not available in the VA, this number is a rough sum of the reported users per country. According to GA there were: ~600 000 users, ~1 400 000 sessions and ~53.5 million pageviews during February 2022. The POSEIDON web page is active for almost 14 years, with a very reliable forecasting section which has received great popularity. With the update of the web site to a more responsive and modern layout, the publicity is constantly increasing (during March it reaches an average of ~80 000 daily users). Indeed the number of users would be interesting to be tracked in the VA as well, if this is a number that is available in the reported logs.

Finally, on the outreach activities, apart from the explicit section of "Outreach" (News and Media), additional information on participation in projects and available publications, can be accessed in the "About us" menu on the Poseidon website.

3.11.4. Assessment by Shaun Deyzel

Summary statistics (recent 1 year ending May 2022 to test filter)

Web metrics (as per VAMS dashboard) report 668,851,033 page views and 16,394,879 sessions over the last year (ending May 2022). By far the majority of traffic originated from Greece and Turkey with >20 million and >9 million views respectively compared to <500,000 by the many other countries accessing this resource. The number of sessions and users mirrored this trend. The number of new users increased dramatically since February 2022, peaking in May 2022 at 91% of the total number of users accessing this resource. Overall, access by mobile constituted 71.9% of the total number of access attempts, which clearly indicates the popularity of the Poseidon App downloadable for iOS, Android and Huawei devices.

The API metrics report 163,090 visits by 1,332 users from 47 countries with 20.7GB of data downloaded from May 2021 to May 2022. An "Unknown" category appears to be dominating much of the additional metrics on this dashboard.

Scientific relevance

The POSEIDON Multi platform observatory Data Centre, through its website, serves the Poseidon system, an observatory research infrastructure of the Eastern Mediterranean basin. In situ data collection is maintained over the long term for multiple observatories with coverage from the atmosphere to the ocean. Data delivery serves decision-making and forecasting by multiple use cases.

Technical Implementation

The Poseidon System website is impressive with an informative and interactive landing page. The "read more" tab leads to a good body of background, context and historical information[63]. Navigation is responsive with little distraction lent to poor resolution or

[63] https://poseidon.hcmr.gr/about-us/evolution

refreshing times. SEO can improve. The only way the Poseidon System website could be found was when a Google search was performed with the term poseidondata. I did not assess the functionality of the App.

Scientific Usability

I could not find indication of a usage licence on the Poseidon System website. When attempting to download a dataset, the process terminated in a request to register/sign-in page. At this point it is not yet clear if the data are freely available or licensed. I could also not find a citation instruction. It was fairly easy to navigate the in situ data service[64] and I don't think a user manual is necessary.

Access Metrics

Two dashboards are available on VAMS, both of which I thought were well equipped with myriad metrics, graphics and components. I was kept quite curious as to the identity of the "Unknown" entry among the Geographical coverage metrics on the API dashboard, simply because it dominated several of the enquiries.

3.11.5. Response from POSEIDON Multi Platform Observatory Data Center to the Assessment by Shaun Deyzel

The review from Shaun Deyzel has been well received and the remarks for improvement will be considered in the next updates whenever possible!

Following the assessment, we are planning on adding a statement of data usage on the landing page of our data access services to make clear the distribution license! At the moment, the distribution statement and the citation statement are provided only in the netCDF files generated upon request for data from a user.

Regarding the "Unknown" category appearing among geographical coverage, we have cross-checked with the respective Google Analytics metrics and indeed the respective "Not set" geographical metric appears in the dominating categories. The description of this metric from the Google Analytics service is the following:

> "Google Analytics uses a third-party datasource to determine your visitors' geographical locations. If our third-party vendor does not have an accurate record of the visitor location, Google Analytics displays a (not set) entry"[65].

[64] https://poseidon.hcmr.gr/services/ocean-data/situ-data
[65] source: https://support.google.com/analytics/answer/1009667?hl=en

3.12. Assessment of the Coastal Observing System for Northern and Arctic Seas (COSYNA) (ID 17.1)

Next, the description of COSYNA, its assessment by Thierry Carval and its assessment by Shaun Deyzel are included.

3.12.1. Description of the VA Service

ID	17.1
Name	COSYNA
Institution	Hereon
Short Description	Monitor real-time conditions and provide short-term forecasts, data, and data products to help assess the impact of anthropogenically induced change. It includes data from Hereon HF Radar, FerryBoxes, Underwater Nodes and operational models (Circulation, Waves) as well as models with assimilated data.
VA Services	The VA Infrastructure consists of a website[66]

3.12.2. Assessment by Thierry Carval

A brief summary of usage statistics

In February 2022 Cosyna reported 7GB of data download per month.

The statistics are focused on the volume of downloaded data.

The majority of downloads are for scientific activities.

On a monthly basis, the number of German data users varies from 90% to 10%

The number of distinct users is not reported.

The number of sessions is not reported.

Scientific relevance

[66] https://www.hereon.de/institutes/coastal_ocean_dynamics/cosyna/index.php.en

COSYNA provides long series of observations of the Northern and Arctic seas. These observations are essential to detect and understand environmental trends (in the short, medium and long term).

Technical Implementation

The interface to select and download data is easy to use and responsive.

The SOS data format is unusual but contains rich metadata (linked data).

Each observation has a SeaDataNet quality control flag.

Scientific Usability

To download data, registration and identification is required. On registration, a clear COSYNA privacy policy[67] is available.

As mentioned for Poseidon, a possible improvement for COSYNA would be to provide a data licence, such as CC-BY[68] that is human- and machine-readable.

To improve its FAIRness on machine-to-machine services, the creation of a DOI for COSYNA dataset may be beneficial. This DOI would ease the citation, the terms of use (licence) and reproducibility.

COSYNA uses linked data and well-established vocabularies (OGC, CF, SeaDataNet): an excellent solution for interoperability.

Access Metrics

The number of distinct users should be reported.

Outreach

On the COSYNA website there is a photo gallery on "observation of the seas". There is also a "Coast in focus" section, with clear explanations and an excellent web app to present the observing network and interactive data display.

The outreach section of D11.1 mentions significant activities in 2021 : organisation of a workshop for more than 100 users, participation in Ocean Science meeting (more than 5000 participants) and news on the web site's new features.

---

67 https://www.hereon.de/innovation_transfer/communication_media/imprint/privacy_policy/index.php.en

68 https://creativecommons.org/licenses/by/4.0/

3.12.3. Response from the Coastal Observing System for Northern and Arctic Seas to the Assessment by Thierry Carval

We have modified improved our integration with VAMS in order to report the number of distinct users every month. We have also upgraded the information that is already in VAMS with that metric. Our live dashboard includes it now.

3.12.4. Assessment by Shaun Deyzel

Summary statistics

Statistics end in April 2022, with the most recent month (May 2022) not visible at the moment of writing this. A summary metrics panel indicating key data will be useful (e.g. data downloaded, users, countries etc.). Data downloads vary considerably between months with a peak over last year being evident for February with >60GB downloaded. In general, downloads average less than 20GB since June 2021. There appears to be a data gap from January 2020 until and including May 2021. By far the majority of downloads are made by the Science sector, followed by administrative and public users respectively.

Scientific Relevance

COSYNA offers an integrated observing and modelling system focused on the Northern and Arctic Seas. The aim is to enable a variety of stakeholders (from science to industry and the public) with respect to investigation and decision-making around issues pertaining to the functioning and state of coastal ecosystems. I am not aware of overlap within the region.

Technical Implication

The data portal leads from the COSYNA landing page with a brief indication that data are freely and publicly accessible via the portal. The data portal tab resolves to an easily explained registration process indicating what the personal information is used for. Users are also urged to read the COSYNA data usage guidelines. Once registered, data can be queried with options to plot or downloaded. What was missing for me is the estimated size of the data being downloaded as it would be useful to know whether I'm stressing my personal system with 9MB or 9GB. The actual download was fairly instant (with a line speed = 50MB), with various format options to choose from.

Scientific Usability

Data is served publicly free of charge pending a short registration process. COSYNA Data Access information is comprehensively outlined on the Data Portal landing page[69]. The information contained therein adequately describes what to expect when filtering and downloading data. There is also a section on Data Usage complete with provision to adhere to scientific best-practice standards with regard to citation and co-authorships. I could not find any explicit statement on citation instructions, not on the website, data portal or once the data were downloaded. I'm assuming it is not necessary.

[69] https://www.hereon.de/institutes/carbon_cycles/cosyna/data_management/index.php.en

Access Metrics

The VAMS dashboard for COSYNA was rather limited. Given the information recorded during the registration process. An analysis by country would provide more detailed insight beyond the current report focussed on data downloads from Germany only, especially given their declaration on their website indicating engagement with 51 countries (in 2017). It may also be useful to introduce a summary metric tab on top to give a glance at total downloads, number of users, countries etc.

Outreach

Links to two popular articles are appended on the landing page. The numerous outreach activities listed in D11.1 (Annexe 2.9 COSYNA ID17.1) is well worth promoting on the main website.

3.13. Assessment of HIDROGRAFICO (ID 19.1)

Next, the description of HIDROGRAFICO+, its assessment by Simon Keeble and its assessment by Sebastien Mancini are included.

3.13.1. Description of the VA Service

ID	19.1
Name	HIDROGRAFICO+
Institution	IH
Short Description	Access to data collected by the real-time monitoring infrastructure for the Portuguese waters (MONIZEE system) that is operated by Instituto Hidrografico.
VA Services	The VA Infrastructure consists of a website[70] which presents data provided by a webservice.

3.13.2. Assessment by Simon Keeble

The platform provides access to data collected by the real-time monitoring infrastructure for the Portuguese waters (MONIZEE system).

It's always a pleasant experience working on a GIS platform that performs quickly and efficiently and this didn't disappoint. It was instantly intuitive to use and quick to respond.

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

Moving around the various parameters was easy and it is clear that this is a service that would be of great interest to recreational ocean users.

Data is up to date and easily retrieved by clicking on features in the map.

Additional useful calculation tools are provided, including voyage planning.

I wasn't able to register on the site, which would have been interesting.

I couldn't find any documentation on the site - perhaps this is not necessary for some, but might be useful for others. Some basic help would be recommended. Contact information, "about" page, how the data is collected etc.

There is no reference to the JERICO-S3 project on the site and credit logos are too small to be appropriate as they are unreadable. In addition, none of them link to the credited agency.

The site works well on a desktop browser, but the UI is quite broken on an iPhone. It is still usable but perhaps needs some corrections.This check was carried out as it was expected that this would be used by the public and mobile phone operation is important.

It's interesting to see the web service statistics, as that gives an impression on the load of the application, but the user statistics were more interesting. The number of monthly users is lower than I would have expected. I'm not based around the Portuguese coast, but to have such a service in my location would be used by all…if it was known about. How is this marketed to the communities? I didn't find many inbound links to the service. Perhaps some link building and public awareness is needed to reach its full market potential?

3.13.3. Assessment by Sebastien Mancini

The VA infrastructure provides easy access to a broad range of data collected by the MONIZEE system in the Portuguese coastal ocean water which is complemented with numerical modelling products such as operational forecasts.

The user interface[71] is very similar to other existing user interfaces providing similar features to users. Users are able to search for a wide range of dataset collections and load the layer on the map interface.

Once a layer is selected, several options are made available to the user to interact with the data. One of the most interesting features is the ability to open/close a timeline tab which enables the user to control a range of options to display gridded data on the map. This feature is quite intuitive, and the response time is relatively fast so that the user does not wait too long for the animation to be prepared. In addition, users are able to select a

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

timeseries at a location and a plot is drawn pretty quickly. The data from the timeseries plot is also available for download in CSV format.

Overall, it is relatively quick to add or remove any layers on the map interface. There is a bit of lag when requesting large timeframes for gridded datasets.

Another important feature is a link to the corresponding metadata in a metadata catalogue to access more information about the dataset collection. The metadata record is well described and contains a number of additional fields including links to the web services (e.g. OGC WMS and WFS…) used by the user interface. Some of the metadata records mentioned the use of Creative Commons but this information could be more clearly displayed for the end user. It was also noticed that some metadata records did not include any licensing information. Also, the content of the metadata record is available in Portuguese.

I assumed that the data for most of the gridded products are stored in netCDF format. It would be useful to include a link in the metadata record to the storage location of these netCDF files (e.g. FTP server, THREDDS catalogue, ….) so that users can access these datasets with their preferred tool.

The infrastructure is built with open-source tools commonly used by other groups in the marine community; Geonetwork for the metadata catalogue, Geoserver to publish OGC web services.

I was unable to register to the website despite multiple attempts to a different combination of username and e-mail address. The error message mentioned that the "username/e-mail" is already in use. Unfortunately, I was not able to test the download feature which seems to be available when a user is logged in.

The website would also benefit from the inclusion of a help section to provide user guides to the end-user on the different functionality made available in the user interface. For example, a user guide describing the "Trip Planning" feature would be very useful to users.

Feedback on access metrics

It is great to include dashboards with metrics for both the website usage but also the use of web services.

For the web service, it would be great to clarify which web service is currently monitored. I am assuming it is the Geoserver instance providing OGC web services to the infrastructure. A link to the Geoserver instance would be useful to avoid any confusion. Are the metrics capturing the number of requests made by the Portal to access the OGC web services or the user statistics accessing the OGC web services without using the Portal interface?

All metrics used in the dashboard are standard outputs for the analysis of website usage.

It would be interesting to identify how many users are actually using the login feature and how many users are actually browsing as Guest. An additional metric would be to report on the list of preferred dataset collections displayed and downloaded by users.

A number of metrics are displaying timeseries by Station ID. These Station ID are not really self-describing and difficult to understand for somebody outside the organisation. Would it be possible to include the station name instead of the station identifier in the metrics?

Outreach activities

Publishing regular news items to highlight the release of new dataset collections to the user interface would be beneficial to raise the profile of the website.

3.14. Assessment of VOS Finnmaid GHG - BGC (ID 21.1)

Next, the description of VOS Finnmaid GHG - BGC, its assessment by Simon Keeble and its assessment by Sebastien Mancini are included.

3.14.1. Description of the VA Service

ID	21.1
Name	VOS Finnmaid GHG - BGC
Institution	IOW
Short Description	Measurement of trace gases on the ferry Travemünde/Germany to Helsinki/Finland.
VA Services	The data is distributed via ICOS (SOCAT[72] database)

3.14.2. Assessment by Simon Keeble

The web page provides an overview of the ship and what it collects.

The data collected is available via SOCAT (ICOS), which is linked at the bottom of the page.

The page does not provide any more information, plots, examples, links to papers or outputs etc

[72] https://www.socat.info/

There is no recognition on the page of the contribution of JERICO-S3 / the EU.

It is unclear from the metrics whether the downloaded data was just for this VA or if it was part of a download of a larger data set, so it's not possible to analyse.

There are limited inbound links to the page. How is the service communicated?

3.14.3. Response from VOS Finnmaid GHG - BGC to the Assessment by Simon Keeble

Thank you for the review and the comments. Please note that the website has been updated at the beginning of the year. All involved and contributing projects are linked and acknowledged. There is also a link to every publication, where data from the Ferryline Finnmaid was involved.

The website will be frequently updated. Thank you for the advice to put more effort into promoting the website of the ferry line in general.

3.14.4. Assessment by Sebastien Mancini

Since 2003 the Leibniz Institute for Baltic Sea Research Warnemünde (IOW) operates the measurement of trace gases on the ferry Travemünde/Germany to Helsinki/Finland, which crosses some of the main basins of the Baltic Sea about every 2 days.

A single webpage[73] on the ICOS website provides contextual information about the type of observations being measured. The webpage does not provide information about the Quality Control procedures applied to the data, the appropriate acknowledgement and citation. The webpage does not include any links to different websites or infrastructures where the users will be able to download the data.

On the JERICO website, it is mentioned that the data collected by this Virtual Access infrastructure is made available on the ICOS Data Portal and the SOCAT database. Distributing the data on these two platforms is totally appropriate as they are the recognised platforms to discover and download data in this field. It also ensures that the data is not duplicated across infrastructure. But it would be very helpful for a first-time user coming to the initial webpage to find instructions on how to access this particular from one of the two data infrastructures.

[73] https://www.icos-infrastruktur.de/en/icos-d/komponenten/ozeane/vos/ostsee/

I have been directed to a link[74] to find the data on the ICOS Data Portal with appropriate filters. The use of these filters was not obvious as a first time user interesting in this dataset.

Moreover, the webpage could also point to other useful information such as, for example, the diagram which provides a good explanation of the data flow[75].

Access Metrics

In a previous report, it is mentioned that the data underlying the plots are obtained via SOCAT's Access Metrics API.

It would be great to clarify which parts of the SOCAT website is used to provide these metrics. There are multiple ways to download the data either directly from the SOCAT website or from the NOAA Liver Server.

Also, are these metrics only capturing the download of this specific dataset or are they also including downloads performed by users on a larger dataset, for example, a user might have requested all data in the Northern Hemisphere.

The ICOS Data Portal is also another user interface to access and download this data. Will it be possible to include metrics on the number of downloads from the ICOS Portal?

Outreach Activities

It would be useful if links to the relevant website or webpages (e.g. Twitter…) are included in the summary table for each outreach activity.

For press releases and newspapers, would it be possible to add a link to the publication?

3.14.5. Response from VOS Finnmaid GHG - BGC to the Assessment by Sebastien Mancini

Thank you for the multiple suggestions on how to improve the data access and website of the ferryline Finnmaid. We are currently working on generating a live access to the data stream from the Finnmaid and an option to visualize and access the data via the IOW Finnmaid homepage until the end of the year 2023. A description of quality control measures will also be included on the website.

74 https://data.icos-cp.eu/portal/#%7B%22filterCategories%22%3A%7B%22level%22%3A%5B1%2C2%5D%2C%22station%22%3A%5B%22iOS_BALTIC-VOS%22%5D%7D%7D

75 https://www.icos-cp.eu/index.php/data-services/data-collection/data-flow

The next update of the website will include citation text and graphics with recent data from the ferry line.

A second website https://www.io-warnemuende.de/Finnmaid.html hosted by the IOW is currently under construction and provides also information on contextual information about Finnmaid.

At https://www.io-warnemuende.de/Finnmaid.html are links included to data library's SOCAT and Pangea.

3.15. Assessment of NorFerry/NorSOOP (ID 25.1)

Next, the description of NorFerry/NorSOOP, its assessment by Anca Hienola and its assessment by Antonio Novellino are included.

3.15.1. Description of the VA Service

ID	25.1
Name	NorFerry/NorSOOP
Institution	NIVA
Short Description	Coastal observing data from FerryBoxes, descriptions of data types, and ocean literacy educational stories are provided for public interaction in the form of 24-inch touchscreen display consoles on passenger vessels that are updated pseudo-real-time.
VA Services	The VA Infrastructure consists of the application available on the touchscreens.

3.15.2. Assessment by Anca Hienola

Description

Coastal observing data from FerryBoxes.

VA

Application available on touchscreen.

Assessment

VA cannot be evaluated properly as the ships were not able to receive passengers due to COVID pandemic.

3.15.3. Assessment by Antonio Novellino

Description

NorSOOP[76] (Norwegian Ships of Opportunity program) is a national research infrastructure that uses ships of opportunity, such as container ships, ferries, and cruise ships, to support oceanic and atmospheric research and observations.

The VA Infrastructure consists of the application available on the touchscreens on passenger vessels that are updated pseudo-real-time. Data are collected and presented to general cruise passengers.

Assessment

It is not possible to do any assessment the system is on the ships so either you get on board or you cannot do it

3.16. Assessment of PORTUS Observing and Forecasting System (ID 28.1)

Next, the description of PORTUS Observing and Forecasting System, its assessment by Thierry Carval and its assessment by Shaun Deyzel are included.

3.16.1. Description of the VA Service

ID	28.1
Name	PORTUS observing and forecasting system
Institution	PdE
Short Description	Oceano-meteorological network and forecasting system that integrates different observing platforms (buoys, tide gauges and HF-radars) with numerical models along the Spanish coast. All data and derived products are distributed through PORTUS early warning system and visualisation tool. Additionally, data from the numerical models, the HF radars and the tide gauges are presently available in the PORTUS OPeNDAPsystem.
VA Services	The VA Infrastructure consists of a website[77] and a THREDDS data server[78].

[76] https://www.norsoop.com/

[77] http://portus.puertos.es

[78] http://opendap.puertos.es

3.16.2. Assessment by Thierry Carval

A brief summary of usage statistics

For the last 12 months, the Puertos Thredds server reported 22 million visits, 10 Tb downloaded data, 2 700 users from 112 countries.

The majority of users are from Spain and the USA, plus a significant number of users from Europe, America, Asia and Africa.

Scientific relevance

It is an OPeNDAP-THREDDS server for accessing all the Spanish data from tide gauges, radars, satellites, models.

The number of datasets and their diversity is impressive.

This comprehensive national service is a major European source of ocean data.

The massive and geographically global number of users clearly shows the importance of this Virtual Access service.

Technical Implementation

The OPeNDAP-THREDDS server is a simple, powerful and efficient data server. It is implemented on top of Copernicus Marine NetCDF4-CF files.

Scientific Usability

The Puertos data have rich, well-documented metadata in the Copernicus Marine NetCDF files. They comply with the CF and SeaDataNet standards, so they are well interoperable.

For findability, a DOI would improve the exposure of datasets.

A possible improvement would be to provide a data licence, such as CC-BY[79] that is human- and machine-readable.

To improve its FAIRness on machine-to-machine services, a DOI for Puertos dataset may be beneficial. This DOI would ease the citation, the terms of use (licence) and reproducibility.

The Puertos NetCDF files have a clear "citation" statement.

[79] https://creativecommons.org/licenses/by/4.0/

Access Metrics

Two additional specific monthly user histograms and data download (without country distribution) would illustrate the important use of Puertos data.

Outreach

There is a rich and diverse outreach activity in 2021: a scientific publication, additional online resources, a best practice publication, two keynote presentations and 5 Twitter posts. Some of these activities should be advertised on Puertos' websites.

3.16.3. Response from PORTUS Observing and Forecasting System to the Assessment by Thierry Carval

It should be said that for a significant part of Spanish data, some institutions don't share their data with us. The lack of DOIs is a serious limitation, we are aware of. We have plans for progressing this year on this subject in collaboration with SOCIB. We will consider adding two additional specific monthly histograms (user and data download) in the future.

3.16.4. Assessment by Shaun Deyzel

Summary of statistics

Data records started in March 2020. Overall, the thredds repository received 76,410,369 visits from 8,115 users representing 129 countries (ending May 2022). During this time 26,4TB of data have been downloaded. The majority of data were accessed via the HTTPServer (19.9TB) followed by OPENDAP (5.2TB) and Catalog (1.2TB). Spain dominates in terms of visitors with 2,444,575 visits in May 2022 compared to <250,000 by the next ranking country, viz. Germany. Chrome is the preferred client by users (3,148), followed by Chrome Mobile WebView (1,413) and Firefox (1,308) respectively.

Scientific Relevance

The PORTUS system is a resource of vast amounts of ocean and atmospheric data served through a THREDDS server with the addition of a web-based mapping and forecasting tool to aid various scientific and maritime users.

Technical Implementation

The OPeNDAP THREDDS server is a standard implementation providing standards-based access to standardised NetCDF files. There is a large number of files available, which can be a bit overwhelming.

Scientific Usability

The OPeNDAP THREDDS server contains a vast collection of data that is easy to navigate in the folder configuration. DOIs for dataset(s) would help make them findable. The FAIRness of the data needs to be improved.

Access Metrics

The VAMS dashboard presents Thredds traffic. Overall comprehensive with an overview panel indicating the number of visits, users and countries of origin, as well as the amount of

data downloaded. Temporal trends are shown for data downloads, users, visits by country, visits by client, users by domain, users by operating system.

Outreach

I could only find outreach activities listed in the D11.1 report (Annexe 2.13). An online search led me to the Puertos del Estado website[80], which contained loads of general information along with the interactive forecasting map dashboards etc. I'm assuming some of the outreach activities will be listed here, although I could not find them.

3.17. Assessment of Swedish Oceanographic Data Centre (ID 32.1)

Next, the description of Swedish Oceanographic Data Centre, its assessment by Anca Hienola and its assessment by Antonio Novellino are included.

3.17.1. Description of the VA Service

ID	32.1
Name	Swedish Oceanographic Data Centre - Toolboxes in marine data management
Institution	SMHI
Short Description	Access to Swedish oceanographic data and with the use of toolboxes; handling, quality control and analyses of such data.
VA Services	The VA Infrastructure consists of 3 websites (SMHI general website[81], OpenData View[82] and Shark Web[83]) and 2 APIs (OpenData[84] and SharkData[85]).

[80] https://www.puertos.es/en-us/oceanografia/Pages/portus.aspx

[81] https://www.smhi.se

[82] https://opendata-view.smhi.se

[83] https://sharkweb.smhi.se

[84] https://opendata.smhi.se

[85] https://sharkdata.smhi.se

3.17.2. Assessment by Anca Hienola

Description

Access to Swedish oceanographic data

VA

2 websites

Scientific relevance

Although promising at the beginning, once one clicks the link *Explorer SMHI´s data*[86] under the Search SMHI's open data, the viewer faces an unexpected challenge (especially in the field of science!): all the information is in Swedish, from variables, explanations, metadata to open data policy. I think it would have a much greater scientific relevance if the data would be easily available also outside Sweden (and Finland).

Technical Implementation

It is laborious and frustrating for a non-Swedish speaking person to access data from this site. Most of the files are XML (not really human friendly) which means they can be used only by those selected few who are able to write a Python (or something else) script to read them, but completely unavailable to the rest of the world. By mistake, I discovered that one can actually download the data in a different format in the opendata-download.smhi.se. Maybe if everything would have been in English this would not have been an issue. As such, I suggest, at the minimum, translating all the pages, documentation, policies and how-to's in English, to ensure wider distribution.

Scientific Usability

Under *Conditions of Use* there is a clear usage license. But no instructions to cite could be found. I am not sure if a user manual exists. But, if it does, a translation is needed. In any case, some Python scripts to read the files would help the potential users.

I don't consider these data FAIR by any stretch of the imagination (that includes the lack of PIDs). As such, there is a lot of room for improvement.

Access Metrics

It seems that the service indeed is targeting only domestic users, as is evident from the metrics. I suggest adding also the number of international users (probably very small at the moment due to the Swedish language only) and the number of downloads.

3.17.3. Assessment by Antonio Novellino

Description

[86] http://opendata-catalog.smhi.se/explore/

The VA Infrastructure consists of 3 websites (SMHI general website[87], OpenData View[88] and Shark Web[89]) and 2 APIs (OpenData[90] and SharkData[91]). The VA is offering different tools and different means to access data.

Scientific relevance

This VA service fills a scientific gap without overlapping other existing services. It might be improved by including potential missing SMHI data.

Technical Implementation

Sharks interfaces are very intuitive. Maybe an English version of all pages would make the interaction even easier. SEO is good: "shark smhi" and "opendata smhi" are well indexed.

Scientific Usability

The service contains instructions on how to cite and usage licence[92]. Although the licence is not always evident in the pages. The service is interoperable but there is degree of FAIRness improvements: user manual, and examples may help accessibility.

This VA is really providing access to data, the linked interfaces are working well, anyhow if the user is interested in bulk downloads it may be not straightforward. Setting up an ERDDAP on top/in parallel would make this VA very performing.

Access Metrics

It consists of google analytics and matomo analytics. They make sense. One more could be the number of downloaded datasets.

Outreach

Four events are listed, they looks like more institute outreaching activities rather than VA outreaching activities

[87] https://www.smhi.se

[88] https://opendata-view.smhi.se/

[89] https://sharkweb.smhi.se

[90] https://opendata.smhi.se

[91] https://sharkdata.smhi.se

[92] https://www.smhi.se/oceanografi/oce_info_data/SODC/datapolicy_en.pdf

3.18. Assessment of SOCIB Data Centre Multi-Platform Observatory (ID 33.1)

Next, the description of SOCIB Data Centre Multi-Platform Observatory, its assessment by Anca Hienola and its assessment by Antonio Novellino are included.

3.18.1. Description of the VA Service

ID	33.1
Name	SOCIB Data Centre Multi-Platform Observatory
Institution	SOCIB
Short Description	Access to the multi-platform observing system of the Balearic Islands in the western Mediterranean.
VA Services	The VA Infrastructure consists of a THREDDS data server[93], a data API[94] and a data catalogue[95].

3.18.2. Assessment by Anca Hienola

Description

Access to the multi-platform observing system, of the Balearic Islands in western Mediterranean

VA

THREDDS data server, data API and data catalogue

Scientific relevance

According to the SOCIB Data repository website, the repository contains a large variety of Oceanographic data, including physics (seawater temperature, seawater salinity, currents, sea surface height, etc.) biochemistry (nutrients, oxygen, turbidity, etc.) and biology (chlorophyll mass concentration, etc.), but also related meteorological variables (atmospheric pressure, wind velocity, and direction, air temperature, and humidity, etc.). This data facilitates the monitoring and forecast of the ocean state, with a particular focus on the Western Mediterranean Sea, and therefore of great scientific value.

Technical Implementation

[93] https://thredds.socib.es

[94] http://api.socib.es

[95] http://apps.socib.es

This is probably one of the best data repositories in use, providing a data catalogue in the form of a browser-like application dedicated to non-operational users, data API for machine to machine access and a data server (Thredds) directed towards operational users. All three "branches" are easy to use (even Thredds, which might be intimidating at the beginning), with clear instructions and intuitive filtering. I was also very pleased to see that the visualization of data is included as a feature, however not always working (this is something to develop more). Although probably not necessary due to its clarity, a service like a Helpdesk might help "lost" users.

Scientific Usability

There is a very clear usage license and there are instructions on how to cite. There is also a very clear user manual. However, maybe, it would be a good idea to have it- upon clicking - open in a new tab instead of having it only downloadable. The service is interoperable.

Regarding FAIRness, It is not very clear to me why some of the datasets are not provided with DOIs. For the benefit of FAIRness and users, I think all the datasets included in the repository should be assigned PIDs.

Access Metrics

SOCIB's dashboard presents an extensive set of metrics obtained from the log files (access, number of views and users vs time by protocol, country, web browser, domain, operating system, observing system, platform type and name, instrument, data processing level, data aggregation level), Google Analytics (access, number of views, sessions and users by time by country, new and returning users, bounce rate, traffic source), Google Analytics API and log files of API server. The metrics are focused on user behaviour characteristics, but none on the repository operations and scientific impact (this comment is also valid for the other VAs, where applicable). While the costs associated with repository operation serve mostly for internal administration, the scientific contribution impact (eg. Number of publications citing SOCIB, number of projects using SOCIB data etc.) would reveal the true value of the data provided by SOCIB.

Outreach

The outreach activities for SOCIB, according to deliverable D11.1, comprise two Twitter messages and two website mentions. This is a modest effort on outreach compared to the complexity and maturity of the repository. Except for social media, other actions could be considered, such as presentations and booths in large European and international field-specific conferences, articles, newsletters etc.

3.18.3. Assessment by Antonio Novellino

Description

The VA Infrastructure consists of 3 webtools to access and consume SOCIB's data. It matches very well the VA concept.

Scientific relevance

The service fills a scientific gap without overlapping other existing services. The tools provide different methods to access SOCIB area data.

Technical Implementation

Both performance and SEO are very good. User interface is super.

Scientific Usability

The service provides a clear usage licence and instructions on how to cite and is very interoperable. Documentation comes with the tools. SOCIB will be the FAIRness champion in Europe if it adds the ERDDAP interface.

Access Metrics

The system is monitored in terms of logs that makes sense. If they would be able to track the number of publications and services developed with the data/datatools they expose it would be a plus.

Outreach

SOCIB has a very well defined outreach strategy. It integrates tools and multimedia channels that let them reach from scientists to citizens.

3.19. Assessment of SYKE-ALG@LINE (ID 34.1)

Next, the description of SYKE-ALG@LINE, its assessment by Simon Keeble and its assessment by Sebastien Mancini are included.

3.19.1. Description of the VA Service

ID	34.1
Name	SYKE-ALG@LINE
Institution	SYKE
Short Description	Monitors the state of the Baltic Sea using ferrybox systems. Collected data includes simultaneous measurements of physics, biogeochemistry and biology carried out in two ferrylines.
VA Services	The VA Infrastructure consists of a set of web pages [96] with information about Alg@line and its data

[96] https://www.marinefinland.fi/en-US/The_Baltic_Sea_now/Automatic_observations_from_ships, https://www.marinefinland.fi/en-US/The_Baltic_Sea_now/Algal_bloom_observations, http://swell.fmi.fi/Algaline/ and https://www.finmari-infrastructure.fi/ferrybox/

3.19.2. Assessment by Simon Keeble

Four websites relating to the monitoring of the Baltic Sea using ferrybox systems.

Page about Ferrybox in Marine Finland[97]

A page explaining the Alg@line-network.There is a plot showing the routes.

There is no reference to JERICO-S3 or funding by the EU.

A link to the current observations (is this part of the VA?) with access to the latest information (5 parameters). This is in the form of plots. There is no data access provided here.

Another link to time-series data. This is presented in the form of static plots.

Page about Algal Blooms in Marine Finland[98]

A web page explaining aglae observations.There is limited information on this page. Some basic information about blue-green algae and links to external resources.

There is no reference to JERICO-S3 or funding by the EU.

Page about Alg@line in SWELL[99]

Alg@line Data from the Baltic Sea. There are some interesting plots showing the current profiles. Unfortunately, the rest of the data for last month and this year don't have enough data to show anything, so it's not possible to view.

It's not possible to download data from here and there is no information about where to get it but on the About page it does link to the next link below.

There is no reference to JERICO-S3 or funding by the EU. There is a reference to JERICO-NEXT.

Page about Ferrybox in Finmari[100]

This is a web page giving an overview of the Alg@line ferrybox and it provides links to the data via European FerryBox Database and EMODnet, as well as others.

[97] https://www.marinefinland.fi/en-US/The_Baltic_Sea_now/Automatic_observations_from_ships

[98] https://www.marinefinland.fi/en-US/The_Baltic_Sea_now/Algal_bloom_observations

[99] http://swell.fmi.fi/Algaline/

[100] https://www.finmari-infrastructure.fi/ferrybox/

The JERICO logo is prominent. Beneath it says Jerico-S3, which should be JERICO-S3.

There are no cookie policies - opt in/out on any of the pages, a requirement to operate in the EU.

The Google Analytics metrics show quite high visitor statistics but it's unclear whether this is for a single page or all combined, or for the SYKE website in general. It's important that this is clarified to enable a review of statistics.

Some statistics relating to the data usage, albeit from EMODnet etc, would be useful.

3.19.3. Assessment by Sebastien Mancini

The infrastructure monitors the state of the Baltic Sea using ferrybox systems.

Several links were provided to the reviewer in order to assess the existing VA infrastructure.

The first link[101] provides contextual information about the project and the observations collected. The website provides links to other relevant web pages such as the "Current observations" where maps and timeseries plots of the different variables are made available for the most recent transect. The data presented on the plots are not available for downloads and no links are provided to the end-user for accessing and downloading this data. The "Timeseries" webpage also provides a series of nice plots with explanations. But again, the user is not able to download the plots in PNG format for example or the data in a tabular form. Moreover, the web pages do not contain any information on how the data has been processed, the Quality Control procedures applied to the different parameters.

The second link[102] provides useful information with links to additional websites. Some of these additional websites are only available in Finnish. As for the first webpage, the data is not available for download for the end-user.

The third webpage[103] is a similar website to the real-time data collected at the Utö Atmospheric and Marine Research Station. Contextual information is available in the "About" tab of the website. Links to other websites are made available but it is difficult to get access to the data and be able to download it. The website contains static plots of the data, several plots are not refreshed properly and are just blanks with no additional information.

[101] https://www.marinefinland.fi/en-US/The_Baltic_Sea_now/Automatic_observations_from_ships

[102] https://www.marinefinland.fi/en-US/The_Baltic_Sea_now/Algal_bloom_observations

[103] http://swell.fmi.fi/Algaline/

There is no capability to download the data even the last two weeks of the data as for the Utö Atmospheric and Marine Research Station.

The fourth website[104] is the most relevant one of all the different links made available. It provides a good description of the project and a link to other related data infrastructure (e.g., EMODNET Physics, EMODNET chemistry, CMEMS…) where users can download the data. The links to the EMODNET Physics website brings the user directly to the dashboard where the user can download the data. For the link to the EMODNET chemistry website, it was more difficult to get access to the data and so perhaps a tutorial or a help page would be useful to guide the user on the steps to download data from this infrastructure. There is also a link directing the user to the European Ferrybox database but users need to log in before downloading and it is not obvious how to register to the website.

The data collected by the different platforms is really valuable and I am assuming it is being made available on a number of different research infrastructures across Europe. It is difficult to understand who the primary custodian of this dataset is and where it should be primarily discovered and accessed.

In all the different websites mentioned above, there is no clear mention of appropriate acknowledgement and citation to this dataset.

Feedback on access metrics

It is unclear what is the actual source of the access metrics presented in this dashboard. A number of websites and web pages have been provided to the review and so the access metrics could refer to several of these websites[105].

In addition to the generic metrics describing user access, it would be useful to include plots describing the number of data downloads from the corresponding data infrastructure such as, for example, EMODNET Physics.

Outreach activities

The list of outreach activities is relevant with links to the corresponding online resource.

It would be interesting to highlights any possible workshops where the data might have been used by students in different countries.

[104] https://www.finmari-infrastructure.fi/ferrybox/

[105] Note from VAMS management: That is a great point. We never added a plot where different sites are shown separately. We have fixed that now and added one row with two plots that separate the traffic by site.

3.20. Assessment of Keri Island Research Station (ID 35.1)

Next, the description of Keri Island Research Station, its assessment by Thierry Carval and its assessment by Shaun Deyzel are included.

3.20.1. Description of the VA Service

ID	35.1
Name	Keri Island research station
Institution	TALTECH
Short Description	Part of the Network of experimental research stations of the Estonian Observatory that measures high-frequency full-depth profiles of temperature, salinity, turbidity, oxygen content, chlorophyll-a and phycocyanin fluorescence in the Central Gulf of Finland.
VA Services	The VA Infrastructure consists of a webpage[106]

3.20.2. Assessment by Thierry Carval

A brief summary of usage statistics

In February 2022, for the last 12 months Keri island station reported 231 sessions and 2000 page views.

There were about 40 sessions per month, most of them from the institute ("domestic").

The number of distinct users is not reported.

There is no information on data downloads.

Scientific relevance

EMIS is the primary distributor of Talltech observations systems; it is therefore plainly relevant.

Technical Implementation

The data and metadata are available as NetCDF files with data and metadata. The user interface is an efficient interface for ocean data discovery. The addition of a more machine-to-machine oriented ERDDAP server may be suggested.

Scientific Usability

[106] https://taltech.ee/meresusteemide-instituut/mereinfo

The data from EMIS (Estonian Marine areas Information System) are available in NetCDF4. The NetCDF files have discovery attributes. The NetCDF data parameters follow the CF-1.0 convention.

A richer implementation of NetCDF files following SeaDataNet or Copernicus Marine NetCDF implementation would improve the FAIRNESS of EMIS data.

There is no licence or citation statement mentioned. An improvement would be to provide a data licence, such as CC-BY[107] that is both human- and machine-readable.

Outreach

In 2021 there were two posts on Facebook. I may suggest adding an outreach section on the Keri Island website, that would link to these posts and outreach material.

3.20.3. Assessment by Shaun Deyzel

Summary of statistics

Records start in February 2020, ending in August 2021. The number of sessions total 945 with 5,525 page views. The number of users gradually decreased over the first few months of data reporting but remained steady in 2021. The number of sessions tracks this trend with a reduction from 120 in February 2020 to 25 in August 2021.

Scientific relevance

Meteorological and ocean observational data are served through an easy-to-use information system in EMIS, along with various mapping and forecasting platforms (e.g. MET-OCEAN portal, TAllinn-Helsinki Ferrybox etc.). The data is easily downloadable with use cases spanning several sectors and thus has scientific relevance.

Technical Implementation

Data are served through various platforms accessible from the website. NetCDF files are downloadable via EMIS through an easy-to-use interface, which includes a forecasting tool. The User Guide link on the "See details" page is not resolving.

Scientific usability

No clear usage licence or citation instruction was listed on the website. A THREDDS server will improve the FAIRness of the data.

Access Metrics

The dashboard for this VA gives basic metrics related to sessions and views, users and new users, bounce rate and session duration. Information on domestic use is also shown and in one instance divided by device. Geographical user information would be useful as well as the amount of data downloaded.

Outreach

[107] https://creativecommons.org/licenses/by/4.0/

Only two outputs to social media reported. I'm sure there were more, which should ideally be showcased on the website.

3.21. Assessment of OBPS-OTGA (ID 36.1)

Next, the description of OBPS-OTGA, its assessment by Simon Keeble and its assessment by Sebastien Mancini are included.

3.21.1. Description of the VA Service

ID	36.1
Name	OBPS-OTGA
Institution	IODE of UNESCO-IOC
Short Description	**Ocean Best Practices System** (OBPS): Open access, permanent, digital repository of community best practices in ocean-related sciences maintained by the IODE of the UNESCO-IOC as an IOC coordinated activity **The OceanTeacher Global Academy** (OTGA) Project aims at building equitable capacity related to ocean research, observations and services in all IOC Member States. **AquaDocs**: Open access thematic document repository covering the natural marine, coastal, estuarine/brackish and freshwater environments maintained by the UNESCO/IOC International Oceanographic Data and Information Exchange (IODE) and the International Association of Aquatic and Marine Science Libraries and Information Centers (IAMSLIC) with support from the FAO Aquatic Sciences and Fisheries Abstracts.
VA Services	The VA Infrastructure consists of three different websites: AquaDocs[108] (an oceanographic documentation open-access repository), OceanTeacher[109] (an oceanographic web-based training platform) and OceanBestPractices[110] (a best practices repository).

3.21.2. Assessment by Simon Keeble

The service covers three websites.

[108] https://aquadocs.org

[109] https://classroom.oceanteacher.org

[110] https://repository.oceanbestpractices.org

AquaDocs[111]

AquaDocs is the joint open access repository of the UNESCO/IOC International Oceanographic Data and Information Exchange (IODE) and the International Association of Aquatic and Marine Science Libraries and Information Centers (IAMSLIC).

It contains an extensive repository of searchable documentation across the natural marine, estuarine/brackish and freshwater environments.

The search facility is fast and the results provide a good overall view of the metadata associated with documents and enables the user to download.

There is no recognition of the JERICO-S3 or EU funding. Although using cookies, there is no opt in / out or privacy policy, required to operate in the EU.

The metrics show a very high level of usage, and this would be expected with such a huge volume of content. The highest level of views comes from the country 'Others'. Perhaps this could be explained?

Ocean Teacher Global Academy[112]

OceanTeacher Global Academy provides a comprehensive web-based training platform that supports classroom training (face-to-face), blended training (combining classroom and distance learning), and online (distance) learning.

The site is well organised and intuitive in terms of the courses it offers. It is not possible to evaluate a course as part of this review.

A calendar is provided but it doesn't show any events.

There is no recognition of the JERICO-S3 or EU funding.

The usage statistics shows that this has been high in the past but there are no usage statistics since summer 2020.

OceanBestPractices[113]

The Ocean Best Practices service is a well known resource. It is well organised into specific communities and collections. It is also searchable.

[111] https://aquadocs.org

[112] https://classroom.oceanteacher.org

[113] https://repository.oceanbestpractices.org

Search results produce well laid out metadata and the ability to download documents.

Although in the footer of the page there are links to feedback / helpdesk, these are hidden out of the view of the user. It would be appropriate to add a contact page to the site.

There is no recognition of the JERICO-S3 or EU funding.

Although using cookies, there is no opt in / out. The privacy policy is inadequate as it is not explicit enough to meet GDPR regulations, and even claims that loss of personal data is not the responsibility of UNESCO.

The metrics show a very high level of usage internationally. Particularly across North America and western Europe.

3.21.3. Response from OBPS-OTGA to the Assessment by Simon Keeble

AquaDocs

Is a hosted product and we will start a dialogue concerning cookies.

I cannot see the category ' Others' on stats provided ? https://aquadocs.org/displaygastats

Ocean Best Practices System (OBPS)

We acknowledge JERICO S3 support on our webpage[114].

The OBPS Privacy Policy[115] has been revised and a planned enhancement update is to include that and the Disclaimer plus other info in an integrated 'Terms of Use' link.

Do not understand the Helpdesk/Feedback comment 'these are hidden out of the view of the user.' - these are displayed at the bottom of every screen and are live links to a Contact form[116].

3.21.4. Assessment by Sebastien Mancini

This Virtual Access Infrastructure is composed of the following three sub-components:

- Aquadocs
- Ocean Teacher Global Academy

[114] https://www.oceanbestpractices.org/home/projects/

[115] https://repository.oceanbestpractices.org/page/policy

[116] https://repository.oceanbestpractices.org/feedback

- OceanBestPractices

AquaDocs is the joint open access repository of the UNESCO/IOC International Oceanographic Data and Information Exchange (IODE) and the International Association of Aquatic and Marine Science Libraries and Information Centers (IAMSLIC) with support from the FAO Aquatic Sciences and Fisheries Abstracts. It is a thematic repository covering the natural marine, coastal, estuarine /brackish and freshwater environments and includes all aspects of the science, technology, management and conservation of these environments, their organisms and resources, and the economic, sociological and legal aspects.

The website clearly highlights its objective, what it is trying to support and also define what is actually not in scope. There are a number of user guides available for people to access for a number of activities.

A registering feature is also available and it allows users to subscribe to collections for e-mail updates and also submit new items to the repository.

The website provides the ability to search and download documents. Different facets are available to the user to filter the number of results from an initial search query.

The metadata associated with a particular dataset is clear and well curated. It contains, where applicable, a link to a Digital Object Identifier (DOI) and appropriate licensing information. The use of the Creative Commons license is clearly displayed.

The Classroom Ocean Teacher[117] provides a comprehensive web-based training platform that supports classroom training (face-to-face), blended training (combining classroom and distance learning), and online (distance) learning.

It is required to register to access the content of the courses. Registering is a two-step process. You must be cleared by an Admin to have access to the course's material. I registered to the website but never received the confirmation e-mail. A couple of days later I was able to perform a password reset and was able to log in to the website.

From the dashboard, you can select the different courses of interest. From the list of courses, it is unclear which one is available for enrollment. Many of the courses mentioned that you cannot enrol since enrolment is already finalised but on the main screen it does not mention it.

I was able to enrol in one of the courses regarding "OceanBestPractices". Once enrolled it was relatively straightforward to have access to the Table of contents of the course and start the different modules.

[117] https://classroom.oceanteacher.org/

OceanBestPractices (OBP)[118] is a secure, permanent document (and other objects) repository. It aims to provide a discovery point for research groups to search and find community accepted existing ocean best practices. This service also invites the ocean research, observation and data/information management communities to submit their own best practice documents to share globally with their colleagues.

The website has a clear user interface and provides users with the ability to search documents using different facets. It has the ability to access more metadata about each document. The metadata is displayed in a consistent way, providing clear licensing information such as Creative Commons and Digital Object Identifier that can be reused by the user. Multiple versions of the same document are available for download by the users. A registration facility is available for members of the community who want to contribute future best practices.

Feedback on access metrics

The dashboards provide the typical information gathered by Google Analytics such as the number of users and the number of sessions.

The maps and the number of users are the most interesting metrics to help understand where the users come from and help define the location for future workshops so that the tools are being used by a broader audience

For each of the three resources, it would be interesting to identify how many users are currently browsing as guests or using login. Moreover, it could be useful to include metrics on the use of facets to understand which ones are preferred by users and could be improved in the future.

Outreach activities

Conferences on data management are a very useful platform to promote these tools. The IODE meetings have been used in the past to present the tools and their future developments.

3.21.5. Response from OBPS-OTGA to the Assessment by Sebastien Mancini

The metrics suggested by Sebastien Mancini are available and would need to be included in the metrics dashboard.

[118] https://repository.oceanbestpractices.org/

3.22. Assessment of VLIZ Marine Data Archive (ID 38.1)

Next, the description of VLIZ Marine Data Archive, its assessment by Anca Hienola and its assessment by Antonio Novellino are included.

3.22.1. Description of the VA Service

ID	38.1
Name	VLIZ Marine Data Archive
Institution	VLIZ
Short Description	The infrastructure is an archival platform that provides a java based interface with a connected MS SQL Server database for upload and documentation of data files.
VA Services	The VA Infrastructure consists of a website[119]

3.22.2. Assessment by Anca Hienola

Description

Online repository/archival platform

VA

Website

Scientific relevance

This is another marine data archive among many others, be they institutional, national, regional or subfield specific. It is not very clear from the description[120] of the repository if it's a national repository and what is the added value of the MDA in comparison to the other 157 marine repositories registered with re3data[121]. Probably a better description in the introductory page would help a newcomer understand the purpose and the target audience (personal note: with larger fonts and better contrast between fonts and background).

Scientific Usability

First of all, the service requires registration in order to upload or search, find and download datasets. Upon registration, the user is taken to the archive page which is presented in a tree-like (but not very intuitive) manner. When searching for specific types of data, the only possibility is the search (free text) box without any other filtering possibilities, nor

[119] https://marinedataarchive.org

[120] https://marinedataarchive.org/introduction.php

[121] https://www.re3data.org/

visualization of the location. It is not either clear how to cite or where to find the DOI or any other PID. Although the repository provides a rich metadata form, most of the datasets are accompanied by scarce metadata, which shows that a minimal curation is needed.

There is no usage license or instructions on how to cite. However, there is a user manual. Finally, FAIRness can be improved.

Access Metrics

The metrics show a constant increase in the number of users, although those who upload or download data into or from the archive are an order of magnitude less. As the repository/archive deals with scientific data, metrics regarding the scientific impact are needed.

Outreach

The list of outreach activities for this VA is extensive and diverse, comprising workshops, conference presentations, training sessions, Twitter messages etc., the fact that is translated into a constantly increasing number of views/users. I would like to suggest a series of sessions where users will voice their opinion and vote for the functionalities they most desperately want to see introduced in order to make the service even more attractive.

3.22.3. Assessment by Antonio Novellino

Description

The Marine Data Archive (MDA) is an online repository specifically developed to independently archive data files in a fully documented manner. The MDA can serve individuals, consortia, working groups and institutes.

Assessment

After the registration to the service "No access to archive" is provided. It's impossible to assess the VA, not really FAIR.

On the other hand, SEO is good: searching for "VLIZ marine data Archive" points to the right place.

3.22.4. Response from VLIZ to the Assessment by Antonio Novellino

New MDA users get auto-approved - they automatically have access to a personal space and the public space, so this must be an exception. Antonio already got an account to our systems in 2015, we're looking further into the issue.

Data in the MDA archive are made public and FAIR by linking them to the Integrated Marine Information System (IMIS) - a FAIR metadata catalogue with DOIs etc. (i.e: https://doi.org/10.14284/549).

4.OUTREACH, DISSEMINATION AND COMMUNICATION ACTIVITIES

The results of these documents have been communicated, mainly, to the members of JERICO-S3, especially to the partners of WP11. These communication activities consist of:

- Online meeting with all JERICO-S3 WP11 partners on May 25th, 2021
- Online workshop in the Nov 2021 JERICO General Assembly
- Online meeting with all JERICO-S3 WP11 partners on Feb 2nd, 2022
- Online presentation in the March 2022 JERICO General Assembly
- Communication of the assessments by the VA Expert Panel of each VA service to the concerned organisation.

These communication activities are helping the partners of WP11 improve their VA services. They are also fostering collaboration between WPs.

Annexe 2 of D11.1[122] includes a report of the outreach activities carried out by the JERICO-S3 WP11 partners to promote their VA services. The report corresponds to the period from February 2020 to August 2021.

[122] D11.1 First report on VA JERICO Resources access statistics and service provision: (https://www.jerico-ri.eu/download/jerico-s3_deliverables/JERICO-S3-D11.1-First-report-on-VA-JERICO-Resources-access-statistics-and-service-provision-FINAL-V2-Compressed.pdf)

5.CONCLUSIONS

Each VA service has been assigned to two of the six members of the VA Expert Panel of JERICO-S3. This document contains these assessments together with a succinct description of each service and, optionally, the response of the VA infrastructure. The assessments are based on the access metrics reported in D11.1, the outreach activities reported in D11.1 and the direct interaction of the panel members with the VA service.

Although each VA service has had specific feedback provided by the assigned VA Expert Panel members, the following are the general conclusions we could draw from the collection of assessments.

As a general rule, the panel members have found the VA service very scientifically relevant and useful for their respective user communities.

On the other hand, the two main points for VA services to improve have been insufficient licensing information and citation instructions and insufficient outreach activities.

Additionally, some tools have been found intuitive and easy to use while others need better instructions to help users.

Finally, VA services that were not open source (or were written in proprietary programming languages) received the recommendations to share the code to make it open to contributions by the community (or to change to an open source language to improve accessibility).

Specific workshops and meetings will be organized to support partners improving such issues.