WP10 – Virtual Research Environment
Progress – SDC plenary

Peter Thijssse, Thomas Loubrieu, Merret Buurman, et al.

Athens, 18/19 October 2017
Outline

• Background: VRE in SeaDataCloud
• Approach and current status
• Workplan how to proceed
SeaDataClouds VRE needs to:

- facilitate collaborative and individual research: Using, handling, analysing and processing ocean and marine data into value-added data products which can be integrated, visualised and published using high level visualisation services.

- combine data with subsets from other data resources, such as the ingested collections

- Have a high capacity and performance for big data processing and state-of-the-art web visualisation services

– <..>
– Respect privacy of users and differences in data policies. Differentiated users, different access to data and data products.

– Be possible to configure virtual work spaces for individuals or groups to work on specific projects, including setting up of dedicated pools of data

– Allow producers to decide whether their outcomes will be shared in the public domain or stay private

– Be based and hosted on EUDAT’s infrastructure based on it B2-… service platforms
This is an ambitious challenge

- High ambitions, many expectations now and near future
- It needs a solid architecture, ready to be expanded over time
- But: The EUDAT platform is strong and already partly operational
- And: We can learn from existing architectures from other projects
2. Approach for VRE and architecture specification
As in latest version D10.1 specs document

• Background, expectation from DOW
• Analysis of existing VRE’s
• SeaDataCloud use cases
• Requirements from use cases and beyond
• Architecture
• Development plan
Conclusions of analysis other VRE’s

- Mostly the same expectations with respect to community building, data sharing, processing and analysis tools
- Authorisation/Authentication layer both in portal layer as well as on top of service layer
- API’s for each (processing) service
- Communication and (meta)data standards are key to success
- Front end applications are various: From self created workflows, to VRE virtual labs, to dedicated user interfaces. But all run on same set of services and data.
- We need to distinguish well between typical VRE modules (communication, GUI, etc) and the fundamental architecture
A flexible framework for versatile use cases

- **SeaDataNet**, T/S qualification and optimal interpolation, biology statistical control
- **EMODNET-Chemistry**, same for bio-geo-chemistry
- **EMODNET-Bathymetry**, DTM processing
- **EMODNET-Ingestion**, convert files
- **Marinet2**, Marine Renewable Energies prototype test analysis
- And much more, Research is innovation…
First use case: SeaDataCloud T/S products

CDI + ODV cache

Qualified buffer

Geospatial interpolation

Publication (catalogue, doi, visualization)
First functions targeted
(abstract of total use case!)

<table>
<thead>
<tr>
<th>Task</th>
<th>Tool/Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>log in with single sign on</td>
<td>B2ACCESS + Marine-ID</td>
</tr>
<tr>
<td>integration GUI development</td>
<td>Javascript library</td>
</tr>
<tr>
<td>apply water column obs quality control with friendly data editor and save result,</td>
<td>webODV</td>
</tr>
<tr>
<td>advise data centre of the regional quality control</td>
<td></td>
</tr>
<tr>
<td>be advised of quality control result (email of log of changes/anomalies sorted per DC)</td>
<td>email</td>
</tr>
<tr>
<td>configure DIVA interpolation</td>
<td></td>
</tr>
<tr>
<td>apply DIVA interpolation, send notification (email) when processing is completed</td>
<td>Jupyter + DIVA library</td>
</tr>
<tr>
<td>visualize interpolation result together with original observations of other observations</td>
<td></td>
</tr>
<tr>
<td>extract and view profiles, time series, hovmuller out of the interpolation result</td>
<td>oceanBrowser+oceanotrop+sextant-dataCite</td>
</tr>
<tr>
<td>publish dataset results (metadata and data), get a DOI</td>
<td></td>
</tr>
</tbody>
</table>
Requirements

• Functional requirements: To be extracted from the use cases
• Non functional requirements: To be extracted from users (performance, layout etc), and other stakeholders

=> To be done soon! (and checked against first ideas of Architecture)
SDC VRE Architecture on top of EUDAT's infrastructure (latest draft)
Ideas for technical solutions

• Authentication: Oauth2 protocol (B2ACCESS + Marine-ID)
• Integrated menu: Application in php + Javascript library
• Private file system, sharing: B2DROP (Owncloud/NextCloud)
• Write, execute code: Jupyter notebook
• Predefined processing: OGC/WPS
• Workbenches, ie applications fit for a specific purpose: web applications deployed with VM or docker, e.g. webODV
• Reference datasets
• Communication (chat, forum…)

=> Separate fundamental components, from dedicated VRE applications
Workplan how to proceed

• D10.1 release
  – Deadline M12 => M14
  – Finalise requirements, architecture, and

• Work on development plan:
  – Start small but having the end goal in sight
  – Work in several development runs ending in a code sprint

• Start as soon as possible after M14 with first development cycle.
End..