



SeaDataCloud

Studying the seas from the cloud

Progress of the SeaDataCloud Virtual Research Environment

Merret Buurman, Peter Thijsse and the complete WP10 team
@IMDIS 2018

sdn-userdesk@seadatanet.org – www.seadatanet.org



Outline

1. The SeaDataCloud VRE
2. Progress of individual components
3. Architecture and EUDAT backend
4. Outlook



1. The SeaDataCloud VRE



SeaDataClouds VRE main requirements

- Facilitate collaborative and individual research
- Combine data with subsets from other data resources
- Have a **high capacity and performance** for big data processing and web visualisation services
- Differentiated users, different access to data and data products
- Configure **virtual work spaces** for individuals or groups, including setting up of dedicated pools of data
- **Sharing options**, if user decides
- Based and **hosted on EUDAT's infrastructure** and its B2-... service platforms

Support 5 versatile use cases:

1. **SeaDataNet Temperature and Salinity** water column analysis
2. **EMODNET - Chemistry**, same for bio-geo-chemistry
3. **SeaDataNet Biology Quality Assessment**
4. **EMODNET - HRSM**, DTM processing
5. **Processing and visualising data sets**

Abstract of T/S use case

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

2. Progress per component

Some highlights

Marine-ID mapped to B2Acces



The screenshot shows a login window for Marine-ID. At the top right is a close button (X). The title "MARINE ID" is centered at the top. Below it is a text input field containing "scott". Underneath is a password input field with ten black dots. A checkbox is followed by the text "Prévenez-moi avant d'accéder à d'autres services." Below this is a large blue button labeled "SE CONNECTER". A security notice reads: "Pour des raisons de sécurité, veuillez vous déconnecter et fermer votre navigateur lorsque vous avez fini d'accéder aux services authentifiés." Below the notice are two columns of links: "Register" and "Forgotten password" on the left; "Forgotten username" and "Change password" on the right. At the bottom, it states "Marine-id supports user authentication for EMODnet, SeaDataNet, Coriolis data services" and features logos for EMODnet, SeaDataNet, and Coriolis.



T & S Lab

Bio Lab

WEBODV



Lorem ipsum dolor sit amet, consectetur adipiscing elit do eiusmod.

JUPYTER



Ut enim ad minim veniam, quis nostrud exercitation.

OCEANBROWSER



Duis aute irure dolor in reprehenderit in voluptate velit esse cillum.

SEXTANT



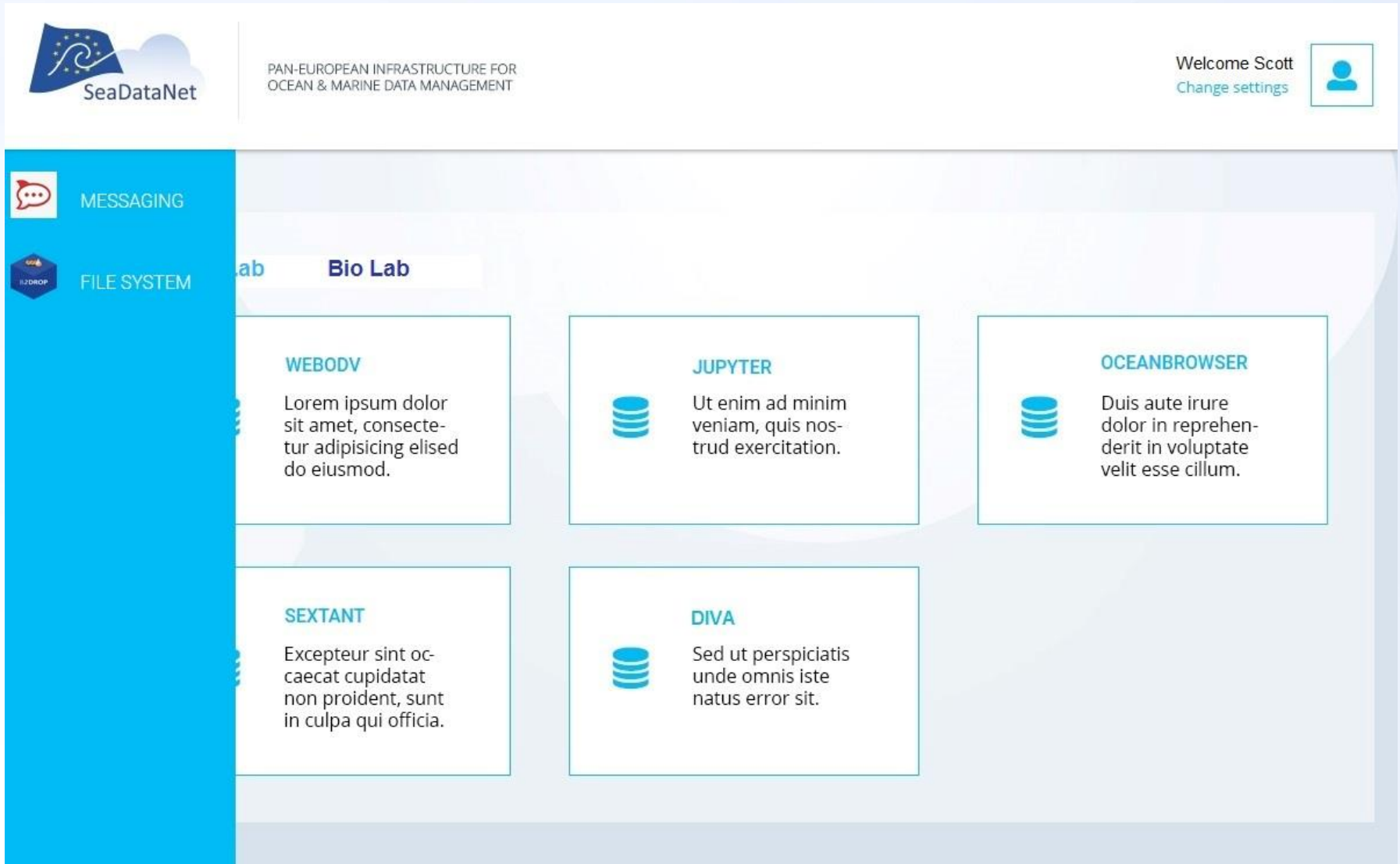
Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia.

DIVA



Sed ut perspiciatis unde omnis iste natus error sit.

Access to messaging a file system



The screenshot displays the SeaDataNet user interface. At the top left is the SeaDataNet logo. The main header contains the text "PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT" and a user profile section for "Scott" with a "Change settings" link and a user icon. A blue sidebar on the left contains "MESSAGING" and "FILE SYSTEM" options. The main content area shows a "Bio Lab" tab and a grid of application tiles: WEBODYV, JUPYTER, OCEANBROWSER, SEXTANT, and DIVA. Each tile includes an icon and a short description.

SeaDataNet

PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT

Welcome Scott
[Change settings](#)

MESSAGING

FILE SYSTEM

lab **Bio Lab**

WEBODYV
Lorem ipsum dolor sit amet, consectetur adipiscing elit do eiusmod.

JUPYTER
Ut enim ad minim veniam, quis nostrud exercitation.

OCEANBROWSER
Duis aute irure dolor in reprehenderit in voluptate velit esse cillum.

SEXTANT
Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia.

DIVA
Sed ut perspiciatis unde omnis iste natus error sit.



Private workspace via B2Drop

The screenshot shows a file management interface with a blue header bar containing navigation icons (home, plus, search, settings). A context menu is open over the 'SDN_Germany.Data' folder, listing options: Upload file, New folder, and New text file. Another context menu is open over the 'SDN_Germany.Data' folder, listing options: Details, Rename, Move, Download, and Delete. The main content area displays a list of folders and files with columns for Name, Size, and Modified.

	Size	Modified
Documents	112 KB	14 hours ago
hallo123	0 KB	a month ago
hallo_123	13 KB	9 days ago
Photos	2.9 MB	14 hours ago
SDN_Germany.Data	663 MB	9 days ago
SDN_test_Ligurien.Data	179.5 MB	17 days ago
thumbnails	243 KB	a month ago
VRE	667 MB	3 days ago
CDI-list-export.zip	9.4 MB	6 months ago
DIVAnd-WebODV3.ipynb	250 KB	6 months ago
gebco_30sec_4.nc	125.1 MB	6 months ago

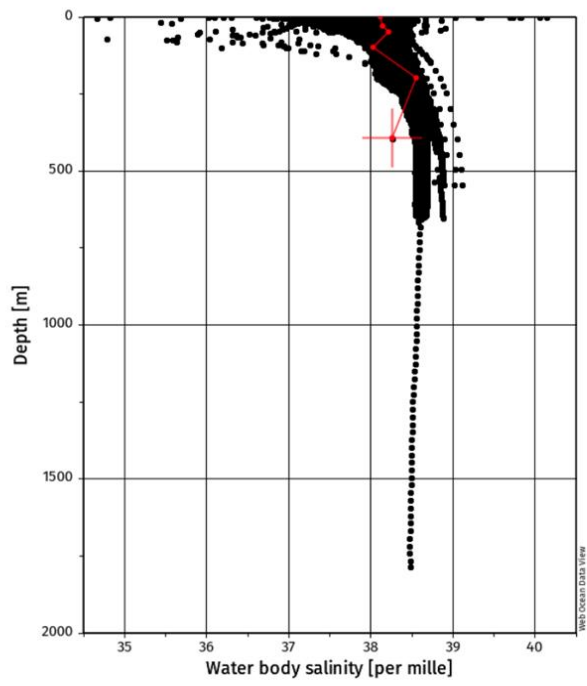
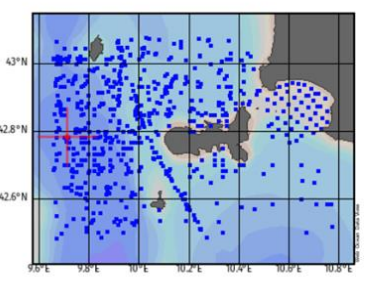
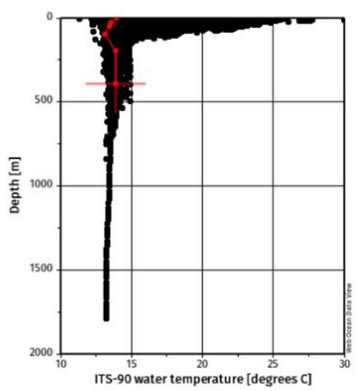
Quality control via WebODV

[video](#)

Quality Control Services

- ← previous
- 1. Select data set
- 2. QC Editor**
- 3. Save/Export/Exit
- next →

Right click on the data windows or map window to open context menus and choose an options. Apply zoom by double left mouse click or *enter* key if *zoom mode* is active. Cancel *zoom mode* by *ESC* key. To assign a quality flag, right click on the value or flag of the respective variable in the "Sample" table at the bottom right of the page. Use the *arrow buttons* below to navigate from sample to sample within one station. Use double-clicking for larger steps.



Status: ▲

Data set: **VRE/SDN_test_EL..**

Mouse:

Station ID: 369 ▲

Accession Num...	369
Cruise	EAA2
Station	36950 (B)
Position	9.717°E / 42.783...
Date	25 March 1976
Time	10:05:59
Depth Range [m]	[0.00 - 396.45]
LOCAL_CDI_ID	158829
EDMO_CODE	120

Sample: 6/6 ▲

variable	value	flag
1: Depth [...]	396.45	1
2: ITS-90 ...	13.88	4
3: Water ...	38.26	4

Data Extractor

- ← previous
- 1. Select stations**
- 2. Select variables
- 3. Download
- 4. Exit
- next →

Select cruises from the *Cruises* menu. Click *Zoom in* to define a sub-region, *Apply* to select the sub-region, or *Zoom out* to return to global domain. Use the *Required variables* as a station filter.

Selection status

Stations:
Output variables:

Cruises

cruises

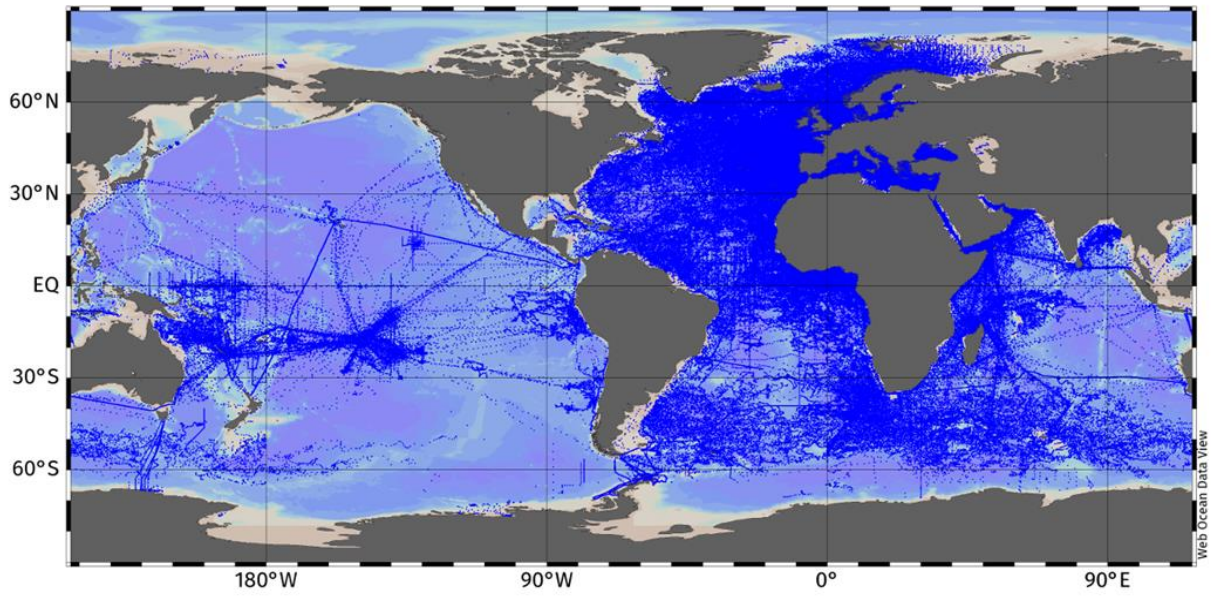
Map domain

Zoom in

Required variables ?

Nothing selected

Reset



ODV-online

SDN_2015-09_TS_MedSea_QC_done_v2.odv

Window 1
STATION

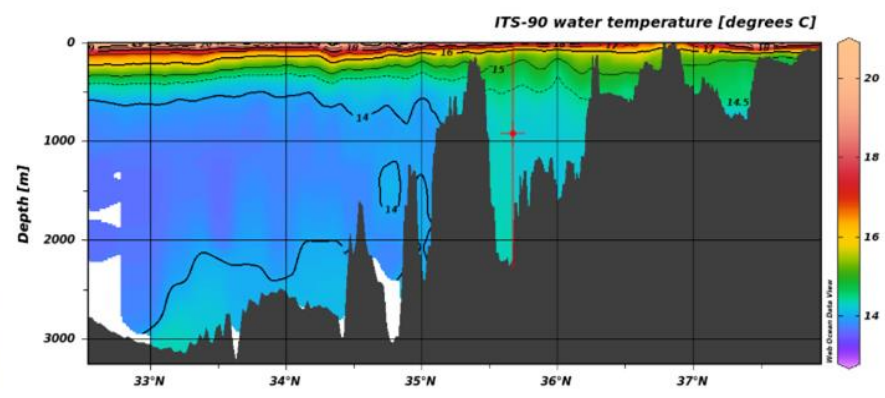
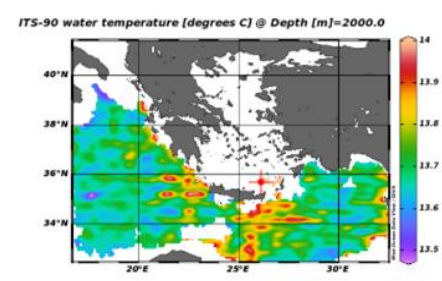
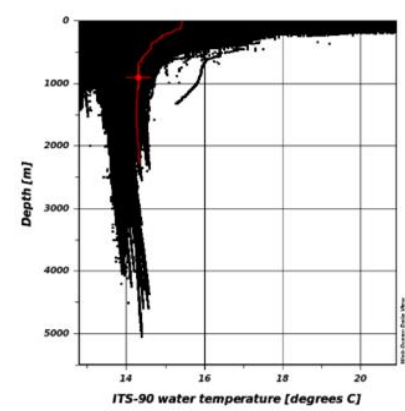
Press ENTER to add the data of the current station to the plot.

ITS-90 water temperature [degrees C] ==>

Window 2
STATION

Press ENTER to add the data of the current station to the plot.

Water body salinity [per mille] ==>



Station ID: 134757

Depth Range [m]	[2.0 - 2285.5]
Bot. Depth [m]	2316
LOCAL_CDI_ID	GN36200503701000_FS109_269
EDMO_code	269

Instrument Info

Reference	
Data set name	GN36200503701c.med(FS109)
Discipline	Administration and dimensions;C

Sample: 923 / 2314

1: Depth [m]	915.2	1
2: ITS-90 water temperature [degrees]	14.33	1
3: Water body salinity [per mille]	36.961	1
drvd: Section Latitude	35.68	1

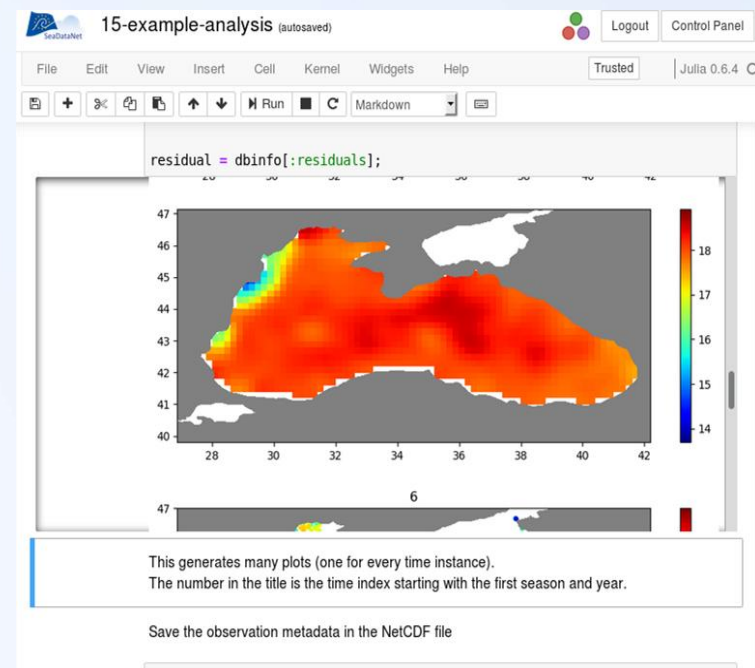
Isosurface Values

Longitude	26.119
Latitude	35.677
Time [yr]	2005.185
Day of Year	68
Depth [m] @ Depth [m]=first	2.0
ITS-90 water temperature [degrees C] @ Depth [m]	15.43
Water body salinity [per mille] @ Depth [m]	36.960

Window 5

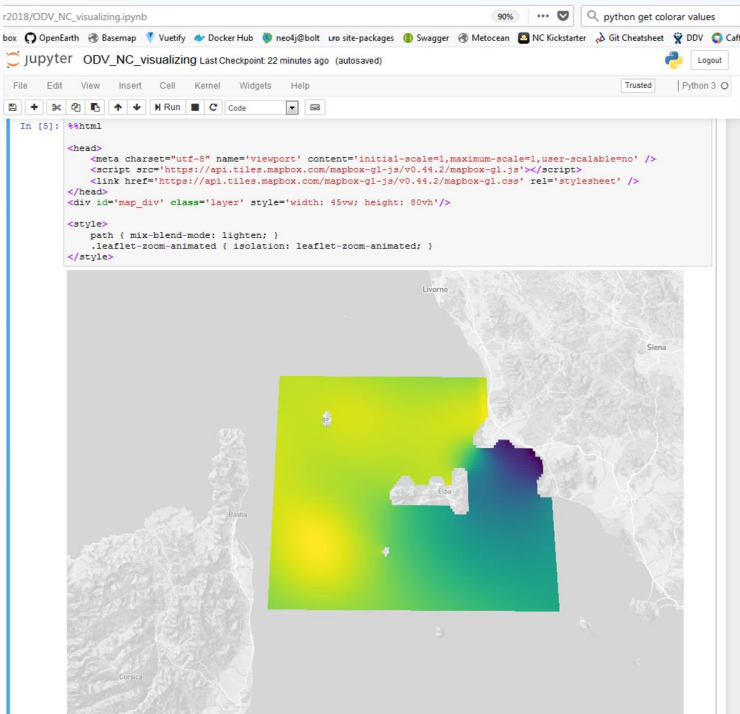
DIVAnd in Jupyter notebooks

- Jupyter notebooks are integrated web environment
 - Computing
 - Visualization
 - Documentation of code
- DIVAnd extension of DIVA in more than 2 dimensions
- DIVAnd distributed as a Julia package installed into Docker containers



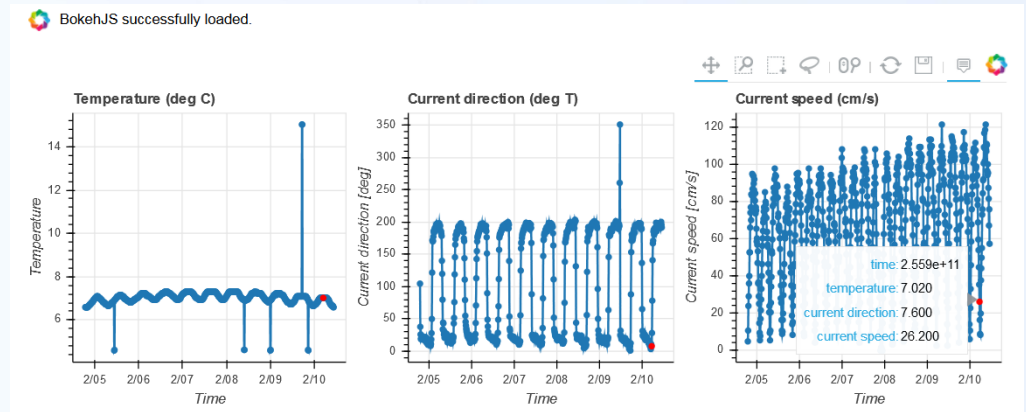
Additional viewing services

- Containers, Notebooks, Package, Visualization
- Dockerfile <https://github.com/openearth/sdc-visualization>
- Data type. NetCDF / tarfile from B2DROP
- Logging, **import** sdc_visualization

```

In [5]: %html
<head>
<meta charset="utf-8" name="viewport" content="initial-scale=1,maximum-scale=1,user-scalable=no" />
<script src="https://api.tiles.mapbox.com/mapbox-gl-js/v0.44.2/mapbox-gl.js"></script>
<link href="https://api.tiles.mapbox.com/mapbox-gl-js/v0.44.2/mapbox-gl.css" rel="stylesheet" />
</head>
<div id="map_div" class="layer" style="width: 45vw; height: 80vh;">
<style>
path { mix-blend-mode: lighten; }
.leaflet-zoom-animatd { isolation: leaflet-zoom-animatd; }
</style>
  
```



3. Architecture and EUDAT backend



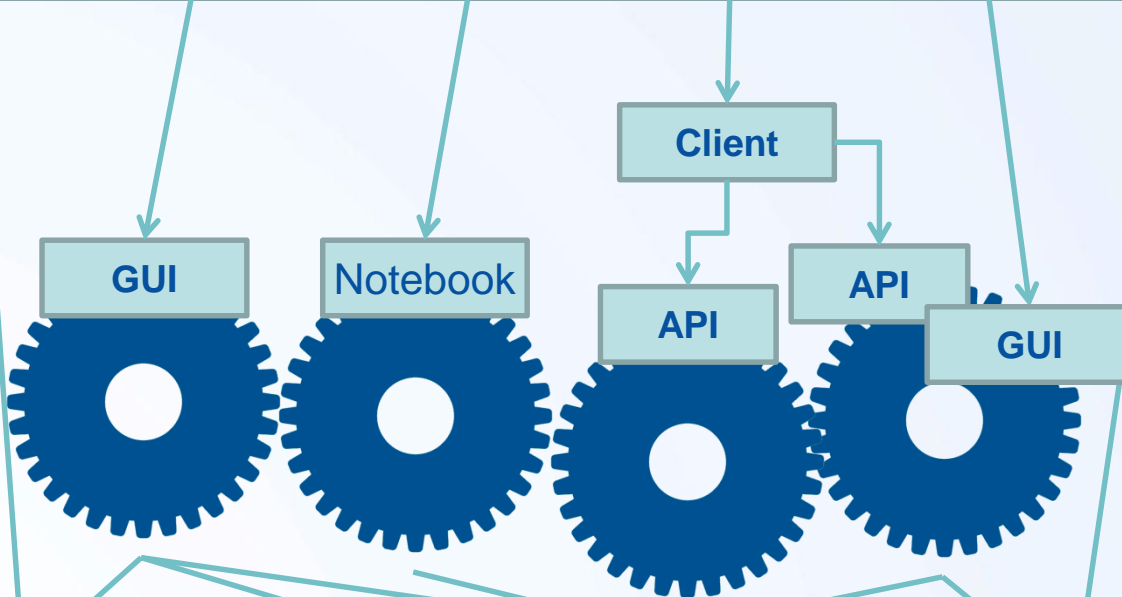
VL1



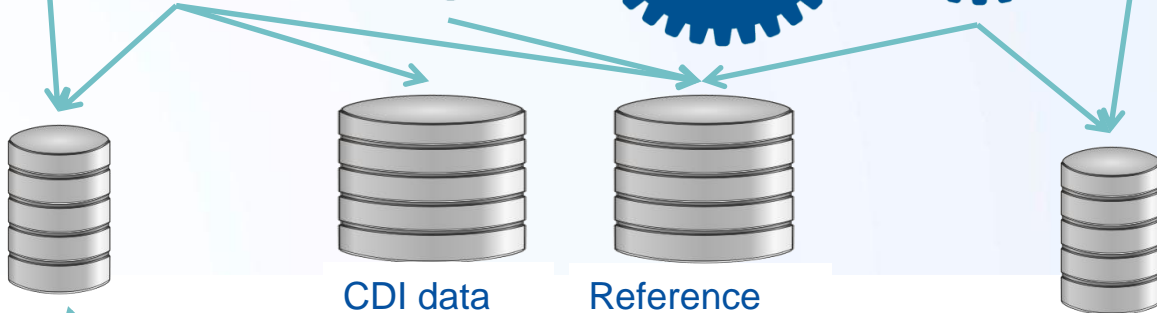
VL2



Frontend



Processing layer



Data access

Data layer

Custom data Alice

CDI data

Reference datasets

Custom data Bob

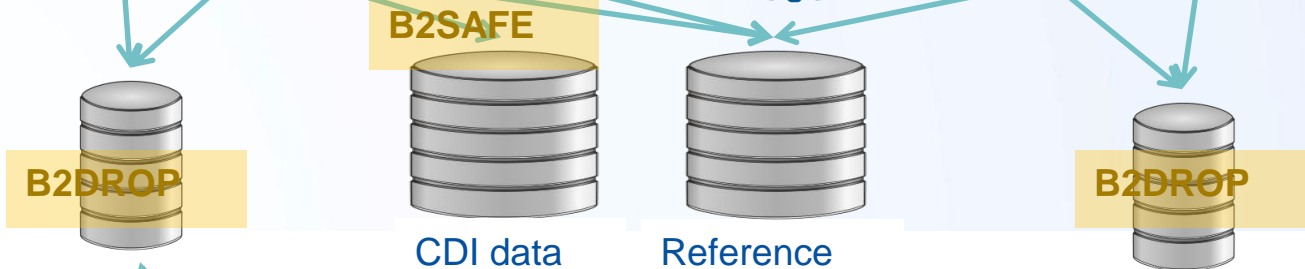
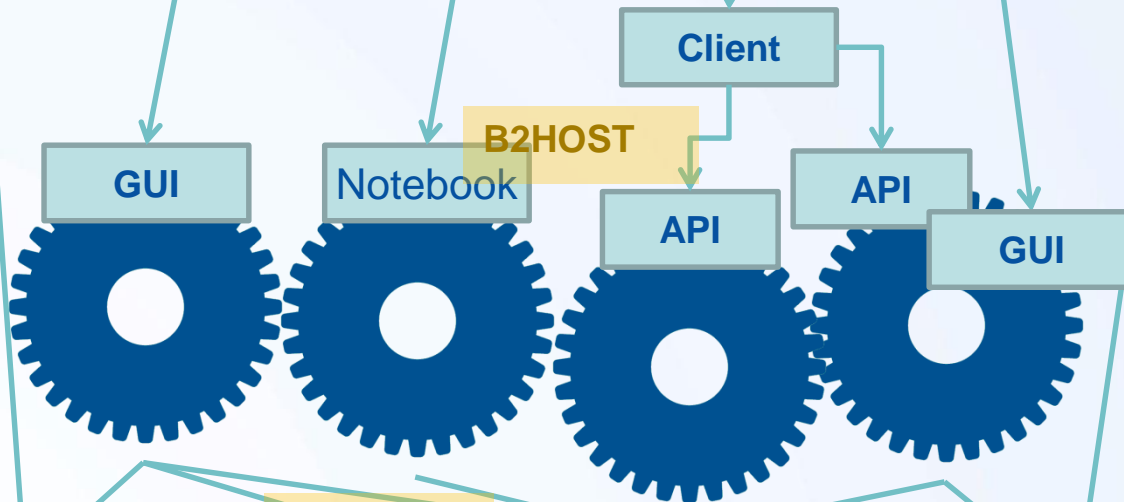
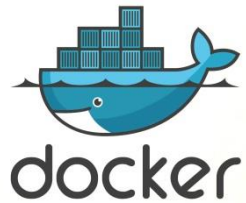


EUDAT services



VL1

VL2



Custom data Alice

CDI data

Reference datasets

Custom data Bob

Processing layer

Data access

Data layer

Technical components

(slide only for later reference)

- B2ACCESS: User Authentication and group management
- B2SAFE and B2STAGE: SeaDataNet CDI data and other external datasets
- B2DROP: User data storage and collaboration
- B2HOST: VRE Execution environment
- Jupyter Notebooks: User code execution
- ... JavaScript Frontend (GUI) to tie all together!

Execution environment (B2HOST) (slide only for later reference)

Applications are provided as docker containers that run on EUDAT computing centres.

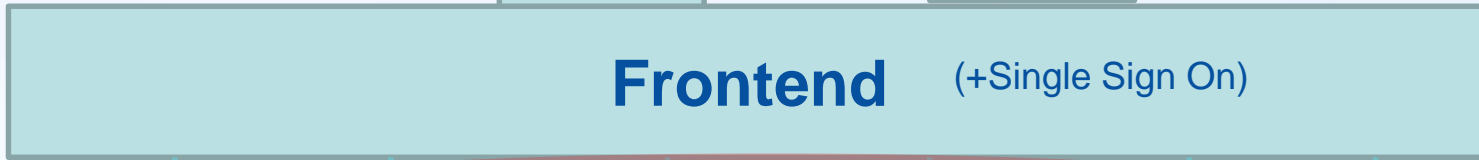
- Processing power
- Efficient access to data
- High availability / redundancy (5 centres)
- Resource accounting
- Service logging (for services run on EUDAT infrastructures)
- Docker registry (SDN-private, incl. security scan)



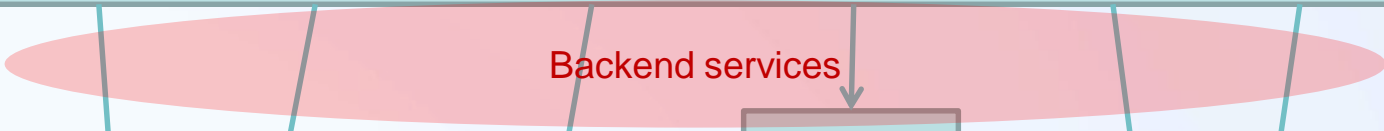
VL1



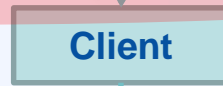
VL2



Frontend



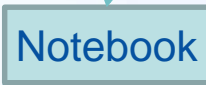
Backend services



Client



GUI



Notebook



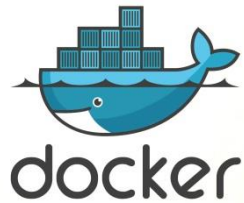
API



API

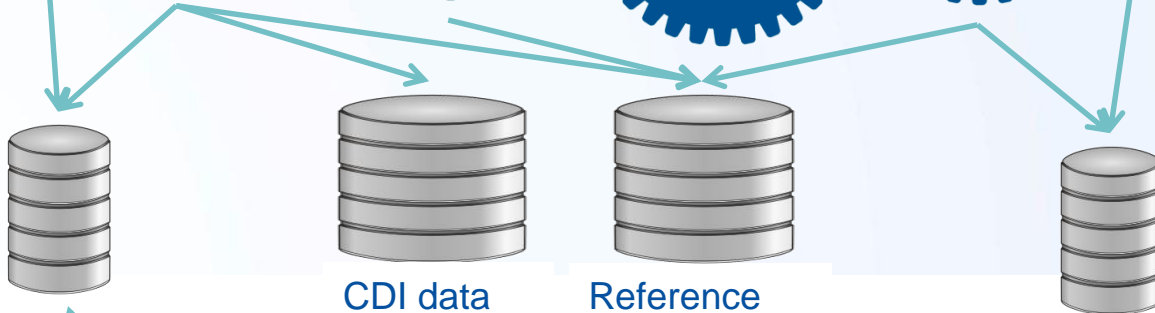


GUI

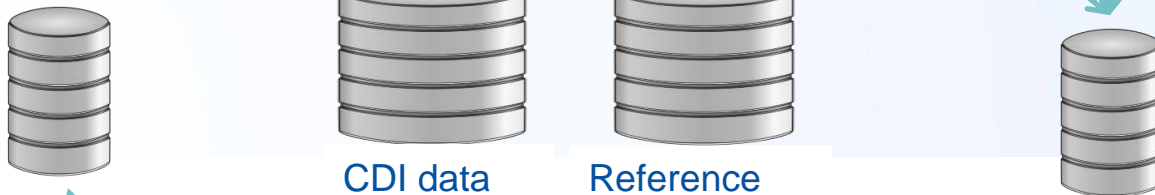


docker

Processing layer



Data access



Data layer

Custom data Alice

CDI data

Reference datasets

Custom data Bob

Backend services

The services that the infrastructure provides to the frontend as HTTP requests.

- Starting/stopping/hibernating services (containers)
- Requesting external datasets
to be moved close to processing
- Requesting B2DROP mount
(unless directly done by services)



VL1

VL2



Frontend

Backend services

Client

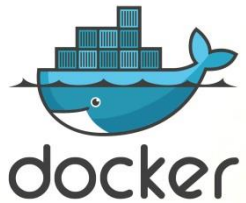
GUI

Notebook

API

API

GUI

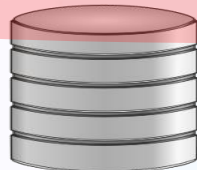


docker

Processing layer

Data access

Data access



Data layer

Custom data Alice

CDI data

Reference datasets

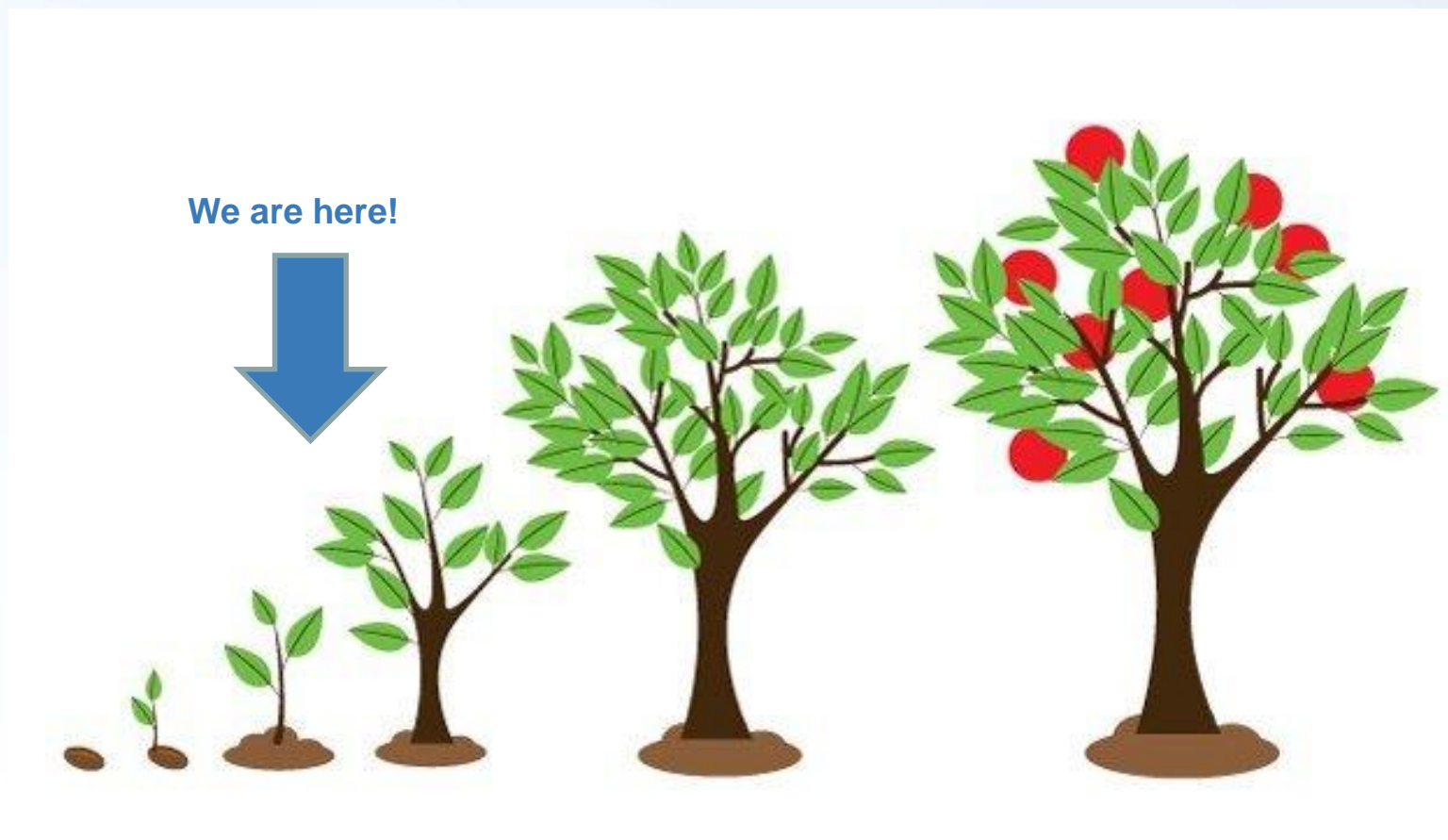
Custom data Bob

Data access

- Challenge: Efficient access to data, but no local copies (for accounting reasons).
- Custom data + CDI data sits on B2DROP (nextcloud) and is mounted via WebDAV.
- SeaDataNet and EMODnet products sit on Sextant, (to be made) accessible via OpenDAP.
- More external datasets will be included.

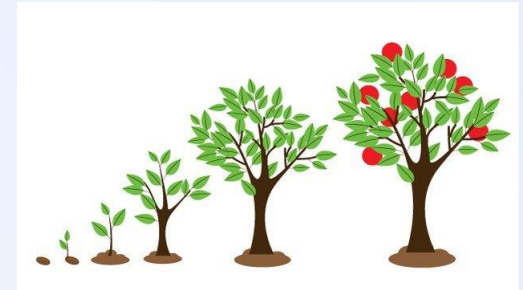
4. Outlook

Outlook



Next steps until June 2019:

- Beautiful design to be integrated
- Customization for different groups
- Push CDI data from portal to VRE
- Enable usage of SDN products
- etc.



After June 2019: Advanced features

(workflows, group management, communications, data publications, etc.)

Questions, or suggestions?