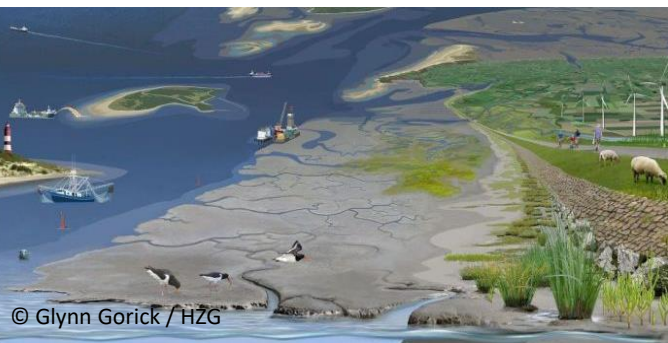


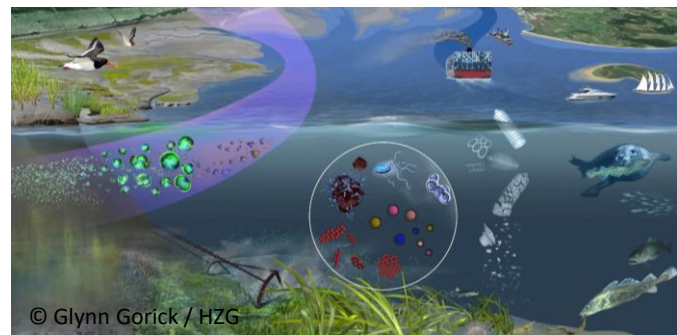
UNIFYING SPECIALIZED DATABASES FOR A CENTRAL SEARCH PORTAL – THE HCDC APPROACH



© Glynn Gorick / HZG

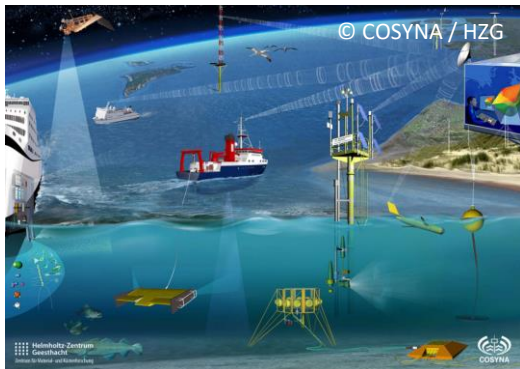


© Christian Schmid / HZG



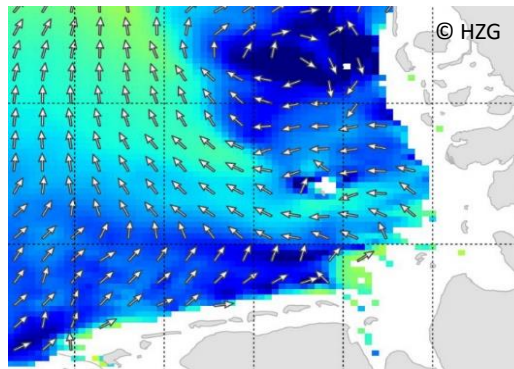
© Glynn Gorick / HZG

Linda Baldewein, Ulrike Kleeberg, Dietmar Sauer,
Robin Luckey, Philipp S. Sommer, Housam Dibeh
Helmholtz Coastal Data Center
Geesthacht, April 14th, 2021



Sensors

- Near real-time data
- Continuous data transfer



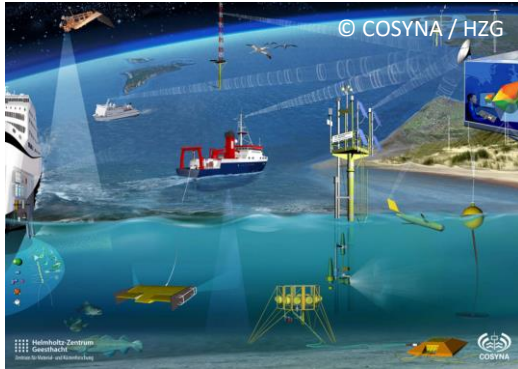
Models

- Large amount of data
- Stored as netCDF files



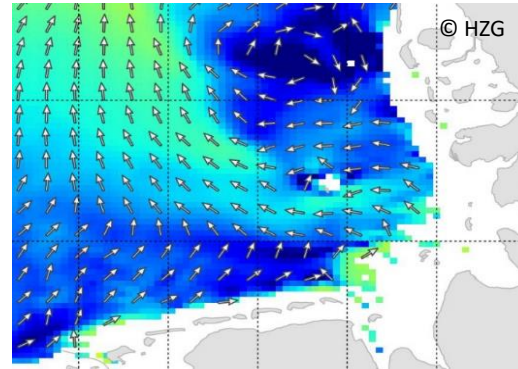
Samples

- Physical samples
- Highly varied metadata



Sensors

- Time series database



Models

- Stored in German Climate Computing Center



Samples

- Relational database

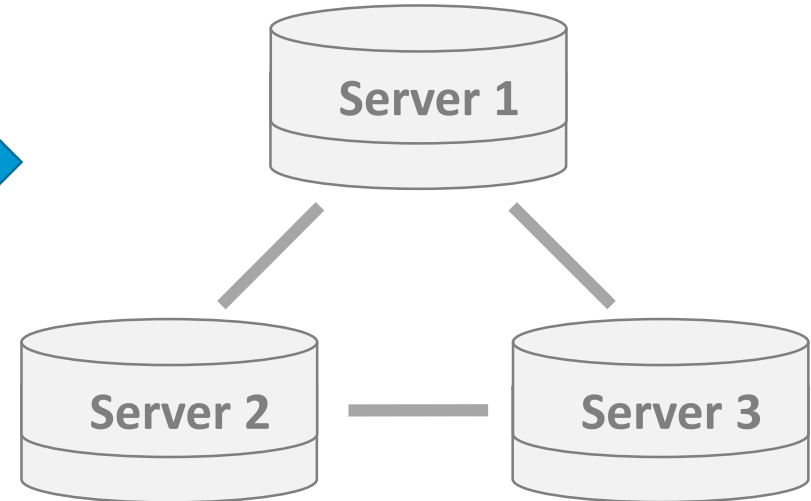
Harmonization of metadata

- Definition of common search options
 - Parameter
 - Author
 - Platform / Campaign / Model
 - Project
 - Date / Time
 - Bounding box
- Standardization of metadata
 - Usage of controlled vocabularies



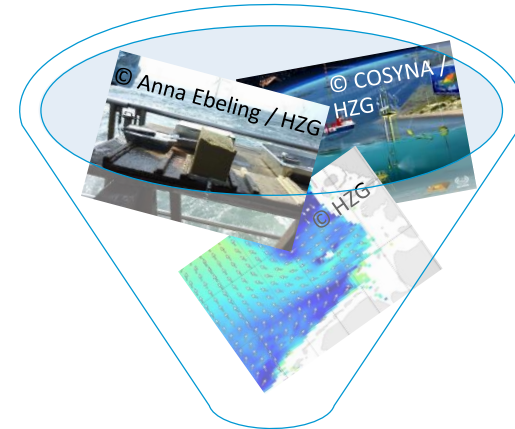
Portal backend

- Elasticsearch cluster



Defining features

- Single point of access to
 - Sensor data
 - Model data
 - Sample data
- User does not see underlying databases
- Download functionalities

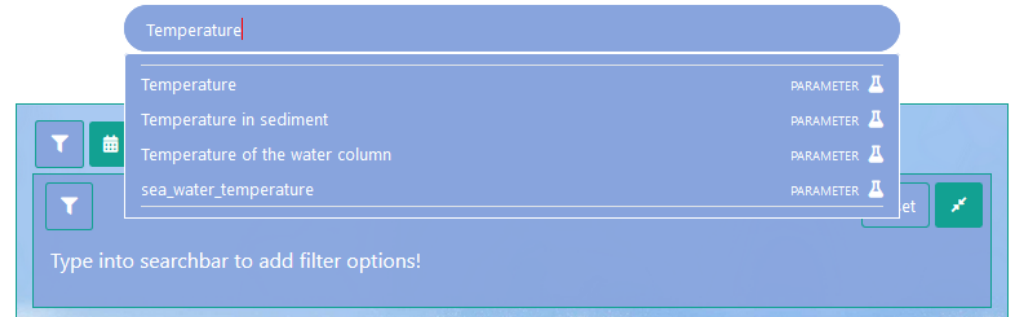


Type here to add filterations ...



Features desired by stakeholders

- “Google“-like search feeling



Features desired by stakeholders

- “Google“-like search feeling
- Filter options
 - Time



Type here to add filteroptions ..

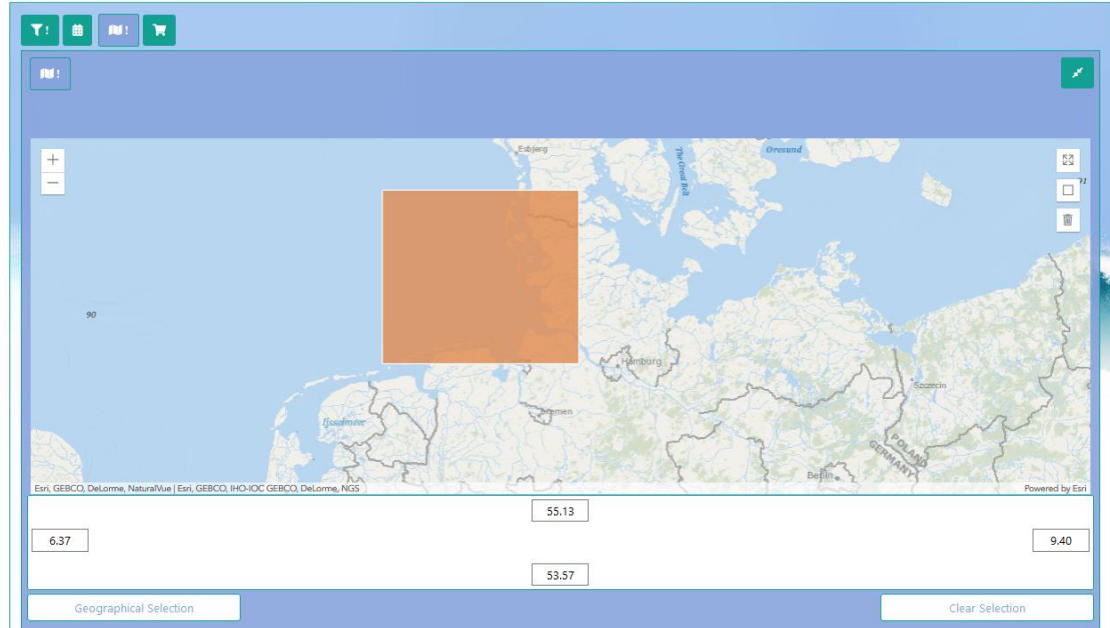
Start date: 2010-01-01

End date: 2021-03-11

Time range in use << January 1, 2010 - March 11, 2021 >>

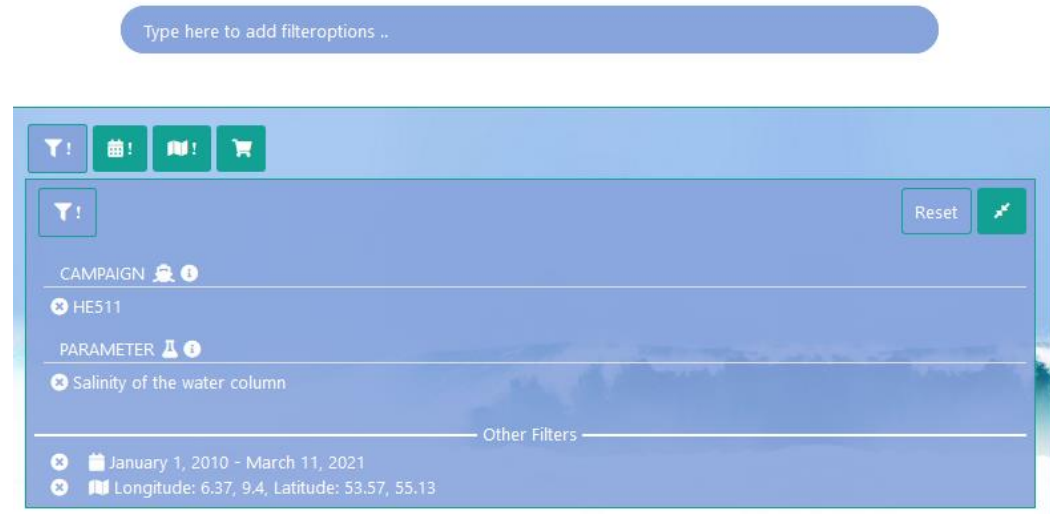
Features desired by stakeholders

- “Google“-like search feeling
- Filter options
 - Time
 - Space



Features desired by stakeholders

- “Google“-like search feeling
- Filter options
 - Time
 - Space
 - Metadata fields, e.g. parameters



Features desired by stakeholders

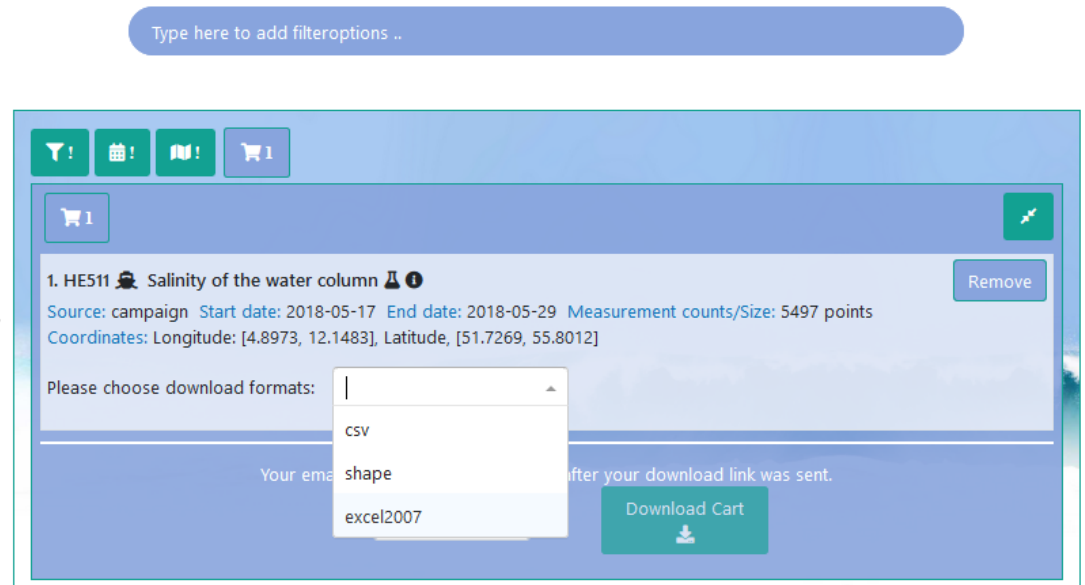
- “Google“-like search feeling
- Filter options
 - Time
 - Space
 - Metadata fields, e.g. parameters
- Simple user interface

Type here to add filteroptions ..

The screenshot displays a search portal interface. At the top, there is a search bar with the placeholder text "Type here to add filteroptions ..". Below the search bar, there are four filter icons: a funnel, a calendar, a location pin, and a shopping cart. The main filter area is divided into sections: "CAMPAIGN" with a dropdown menu showing "HE511", "PARAMETER" with a dropdown menu showing "Salinity of the water column", and "Other Filters" with a date range "January 1, 2010 - March 11, 2021" and a location "Longitude: 6.37, 9.4, Latitude: 53.57, 55.13". A "Reset" button is located in the top right corner of the filter area. Below the filter area, there is a list of search results. The first result is "1. HE511 Salinity of the water column" with a shopping cart icon. Below the result, there is a summary line: "Source: campaign Start date: 2018-05-17 End date: 2018-05-29 Measurement counts/Size: 5497 points". Below the summary line, there are two sections: "NAMES (2)" and "PROJECTS (1)".

Features desired by stakeholders

- “Google“-like search feeling
- Filter options
 - Time
 - Space
 - Metadata fields, e.g. parameters
- Simple user interface
- Download in different formats



Challenges

- Limit amount of data downloaded
- Optimize search results
- Implement tags

Outlook

- Release of portal in Fall 2021
- Implementation of REST API
 - Integration into collaborations such as the German Marine Research Alliance
- Visualization options



Type here to add filteroptions ..

The screenshot shows a search results page with a blue header and a green bar indicating '2 Result(s)'. The first result is for 'sea_water_salinity' from the 'opendap' source, with a warning that the data is not downloadable due to exceeding the 300 megabyte limit. The second result is also for 'sea_water_salinity' from the 'coastdat' source. The interface includes navigation icons at the top and a list of metadata fields for each result, such as PLATFORMS, DATASET, NAMES, PROJECTS, TEMPORALRESOLUTION, METHOD, PARAMETERGROUP_NAME, and UNITNAME.



THANK YOU FOR YOUR ATTENTION!

Linda Baldewein

Helmholtz Coastal Data Center

Linda.baldewein@hzg.de

+49 (0) 1452 87 1801

Max-Planck-Str. 1

21502 Geesthacht, Germany

www.hereon.de

 **Helmholtz-Zentrum
Geesthacht**

Centre for Materials and Coastal Research