O2A – Observations to Archive
Data Flow Framework

Alfred Wegener Institute - Computing Center
presented by Angela Schäfer
D4IR 2017 – 30.11.2017
Use Case: Arctic long-term observatory FRAM

Ice tethered platform networks

... radiation, snow height, ice thickness, temperature, salinity, oxygen, chlorophyll a ...
Use Case: FRAM

Water column

... fluorescence, nutrients, salinity, temperature, conductivity, acoustic Doppler current profiler, water and phytoplankton samples, ...
Use Case: FRAM

Ocean floor

... photo, video, benthic flux, physico-chemical, ...

Medieningenieure Bremen / Sabine Lüdeling
Data Flow Framework

**DASHBOARD**
Monitoring of near real-time data

**ANALYSIS**
Data viewing and analysis solutions; Map-based visualization services

**PORTAL**
One-stop-shop framework Interoperability services

**WORKSPACE**
Solutions for data storage, processing and long-term preservation

**STORAGE ARCHIVE**

**REPOSITORIES**
Data and data products Publications, presentations, field reports

**IMPROVE DATA**
Objectives

- **Generic** infrastructure for data flows
- **Sustainability** and up-to-date services
- **Interoperability** and standards
  - e.g. Open Geospatial Consortium
- Seamless **integration** with our infrastructure
  - Web GIS
  - Workspace/Sandbox
  - Web Portals
  - Data Archive / Publishing
Challenges

- **Heterogeneity** of scientific needs and workflows
- Vast **Number** of different instruments, data sources and formats
- Multitude of **Standards**
- **Integration** with existing solutions, e.g. for the data flow, but also administrative information
- Limited additional **Effort** acceptable by science
- Limited **Bandwidth** ship to shore
Use Case: MOSAiC

Multidisciplinary drifting Observatory for the Study of Arctic Climate, the first year-round expedition into the central Arctic to explore the Arctic climate system during 2019 to 2020.

Based on year round operation of RV Polarstern, drifting with the sea ice across Arctic...
O2A: MOSAiC

MOSAiC Raw Data

Onboard Data Transfer

Polarstern Satellite Link for Data Monitoring and Remote Service

Ship-to-shore Data Transfer

only 2 x 100MB/day ?

“direct” satellite links to partner sites
O2A components

- **SENSOR.awi.de** ready
- **DASHBOARD.awi.de** ready
- **MAPS.awi.de** ready
- **PANGAEA.de** always
- **DATA.awi.de** prototype internal only
Platform and device descriptions for provenance information and reduced data integration effort

- **Versioning and citability**
- Interoperability and standards
- ~1200 descriptions available and counting
User-customizable, flexible dashboards for data monitoring

Automatic data streaming of near-real time and delayed-mode data

Based on sensor descriptions and configurations
Near Real Time Data

Near real time data from research vessel Polarstern is presented in hourly averages and no quality control is applied.

Masterpos on Polarstern

<table>
<thead>
<tr>
<th>Latitude</th>
<th>Longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>71.2</td>
<td>36.3</td>
</tr>
</tbody>
</table>

FerryBox on Polarstern

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDOM</td>
<td>0.1 ppb</td>
</tr>
<tr>
<td>Conductivity</td>
<td>15.2 mS/cm</td>
</tr>
<tr>
<td>Fluorescence</td>
<td>0.8 microg/l</td>
</tr>
<tr>
<td>Flow Rate</td>
<td>1.4 m/s</td>
</tr>
<tr>
<td>Nitrates</td>
<td>0.0 mmol/l</td>
</tr>
<tr>
<td>Phosphate</td>
<td>0.3 mmol/l</td>
</tr>
<tr>
<td>Silica</td>
<td>2.7 mmol/l</td>
</tr>
<tr>
<td>Oxygen</td>
<td>325.1 µmol/l</td>
</tr>
<tr>
<td>pH</td>
<td>8.8</td>
</tr>
<tr>
<td>Phycocyanin</td>
<td>296.0 cells</td>
</tr>
<tr>
<td>Turbidity</td>
<td>10.0 °C</td>
</tr>
<tr>
<td>Salinity</td>
<td>-29.0 mbar</td>
</tr>
</tbody>
</table>

pCO2 on Polarstern

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total CO2 (atm)</td>
<td>386.7 ppm</td>
</tr>
<tr>
<td>CO2 corrected</td>
<td>356.8 ppm</td>
</tr>
<tr>
<td>CO2 density</td>
<td>15.2 mmol/l</td>
</tr>
<tr>
<td>H2O (atm)</td>
<td>32.3 mmol/l</td>
</tr>
</tbody>
</table>

Weather on Polarstern

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air pressure</td>
<td>992.1 hPa</td>
</tr>
<tr>
<td>Temperature</td>
<td>3.1 °C</td>
</tr>
<tr>
<td>Air pressure</td>
<td>50000.0 mbar</td>
</tr>
<tr>
<td>Direct radiation</td>
<td>555.6 W/m²</td>
</tr>
<tr>
<td>Global radiation</td>
<td>34.0 W/m²</td>
</tr>
<tr>
<td>Wind speed</td>
<td>12.0 m/s</td>
</tr>
<tr>
<td>Visibility</td>
<td>72.0 km</td>
</tr>
<tr>
<td>Visibility</td>
<td>355.0 km</td>
</tr>
</tbody>
</table>
DATA.awi.de (release summer ‘18)
COLLECTIONS
Explore data and products thematically grouped

FRAM - FRontiers in Arctic Marine Monitoring
Large-scale Ocean Observation Infrastructure designed to support various types of long-term time series observations in the Arctic Ocean

NEAR REAL TIME DATA
Near real time data is presented in hourly averages and no quality control is applied.

Air temperature
- 18.10 °C Polarstern
- 14.20 °C Heincke
- -5.35 °C Neumayer III

Water temperature
- 9.52 °C Polarstern
- 14.26 °C Heincke

LATEST PUBLICATIONS
Datasets, publications and reports from AWI repositories

Publish your data

Dataset
Jessen, Gerdhard L et al.
Biogeochemical measurements of sediments collected in the Black Sea during the MSM15/1 cruise in 2010 (2017)

Article
Fortelius, Carl et al.
New methodologies to observe wind gusts: research aircraft and Doppler lidar measurements (3201)

Report
Boebel, Olaf
The Expedition PS103 of the Research Vessel POLARSTERN to the Weddell Sea in 2016/2017 (2017)
Data Flow Framework

SENSOR
Manage platform, sensor metadata

STREAM
Near real-time streaming of large data volume

DASHBOARD
Monitoring of near real-time data

ANALYSIS
Data viewing and analysis solutions; Map-based visualization services

PORTAL
One-stop-shop framework Interoperability services

DATA ACQUISITION

NEAR REAL TIME DATA

WORKSPACE
Solutions for data storage, processing and long-term preservation

STORAGE
ARCHIVE

REPOSITORIES
Data and data products
Publications, presentations, field reports
Current work

- Developing a science community workspace for data sharing and data analyses within the **Helmholtz Data Federation (HDF)**
- State-of-the-art storage, AWI-part distributed between Bremerhaven and Potsdam
- User-friendly, high bandwidth compute/analysis solutions with virtual machines and containers
- Hadoop big data analysis based on Hortonworks data flow and data platform
- Raster data management and analysis with rasdaman (**hypercube data analytics**)
Thank you very much for your attention!
We are looking for IT-staff!

Team: Roland Koppe, Peter Gerchow, Ana Macario, Antonie Haas, Christian Schäfer-Neth, Hans Pfeiffenberger, presented by Angela.Schaefer@awi.de
Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research
Bremerhaven, Germany