Persistent Identification of Instruments WG (PIDINST WG)

Markus Stocker (@markusstocker) and The #PIDINST Team
Fun Fact: The #PIDINST WG was conceived in discussions with co-chairs Lou and Ale at the 2017 #THOR#: #ENVRIplus Bootcamp on @ORCID_Org integration in environmental Research Infrastructures held at @AaltoUniversity doi.org/10.5281/zenodo... @freya_eu @ENVRIcomm @resdatall
PIDINST

Proposes a community-driven solution for globally unique and unambiguous identification of instrument instances that are operational in the sciences
Why

- Link data to the instruments that generated them (provenance)
- Aid equipment logistics and mission planning
- Facilitate interoperability and open data sharing
- Improve the discoverability and visibility of instruments and their data
- Metrics that quantify the use of instruments
- ...


Updates since P13 (Philadelphia)

- Completed use case collection (15 to date plus one to come)
- Consolidated PIDINST Metadata Schema
- Adoption with
  - PID Infrastructure Providers DataCite and ePIC
  - Instrument database providers, including HZB, BODC, AWI, ICOS, EISCAT_3D
- Deliverables, in addition to schema
  - Journal article
  - White paper for instrument database providers
Outlook

- PIDINST comes to an end (wrap-up session)
- Transition into “maintenance mode” to
  - Catalyse adoptions, especially among instrument database providers
  - Support the implementation of PIDINST Metadata Schema with DataCite, ePIC, others TBD
  - Maintain schema, resolve issues, etc.
  - Publish results widely, including among other RDA groups
- If new needs emerge, start a new WG
- Short-term: Complete deliverables (article and white paper)
Thanks!

- Co-chairs for the early encouragement and support, throughout the WG
- All those who contributed use cases: Thank you!
- The community for their participation during plenaries
- The regulars during conference calls
- And special thanks to Rolf Krahl
PIDINST Metadata Schema

Rolf Krahl

Persistent Identification of Instruments WG © RDA P14, Helsinki, 23 October 2019
A Metadata Schema for the Persistent Identification of Instruments.

- First step: formulate a generic schema for instruments from scratch (heavily inspired by DataCite Metadata Schema), based on use cases.
- Second step: map generic schema onto DataCite.
- Current status: both versions are considered valid.
- Mostly finalized, a few issues are still open.

Development on GitHub:

https://github.com/rdawg-pidinst/schema
### Generic Schema: Mandatory Properties

<table>
<thead>
<tr>
<th>ID</th>
<th>Property</th>
<th>Occ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Identifier (with identifierType)</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>LandingPage</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>Name</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>Owner (with ownerName, ownerContact, and owner-Identifier)</td>
<td>1-n</td>
</tr>
<tr>
<td>5</td>
<td>Manufacturer (with manufacturerName and manufacturerIdentifier)</td>
<td>1-n</td>
</tr>
</tbody>
</table>

**Identifier**: the persistent identifier.

**LandingPage**: the URL that the PID resolves to.

**Name**: the name by which the instrument is known.

**Owner**: institution(s) responsible for the instrument.

**Manufacturer**: the manufacturer(s). May be the same as the owner.
## Generic Schema: Recommended Properties

<table>
<thead>
<tr>
<th>ID</th>
<th>Property</th>
<th>Occ</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3</td>
<td>modelName (subproperty of Manufacturer)</td>
<td>0-1</td>
</tr>
<tr>
<td>6</td>
<td>Description</td>
<td>0-1</td>
</tr>
<tr>
<td>7</td>
<td>InstrumentType</td>
<td>0-1</td>
</tr>
<tr>
<td>8</td>
<td>VariableMeasured</td>
<td>0-n</td>
</tr>
<tr>
<td>9</td>
<td>Date (with dateType)</td>
<td>0-1</td>
</tr>
<tr>
<td>10</td>
<td>RelatedIdentifier (with relatedIdentifierType and relationType)</td>
<td>0-n</td>
</tr>
<tr>
<td>11</td>
<td>AlternateIdentifier (with alternateIdentifierType)</td>
<td>0-n</td>
</tr>
</tbody>
</table>

**Description:** technical description.

**InstrumentType:** classification of the type of the instrument.

**VariableMeasured:** variable(s) that this instrument measures.

**RelatedIdentifier:** identifiers of related resources.

**AlternateIdentifier:** serial number.
Related Identifier

- RelatedIdentifier value: a PID.
- relatedIdentifierType: controlled list of values: DOI, Handle, URL, URN, ...
- relationType: description of the relationship, controlled list of values: IsDescribedBy, IsNewVersionOf, IsPreviousVersionOf, HasComponent, IsComponentOf, References, HasMetadata, ...

Possible applications:
- link articles describing the instrument (IsDescribedBy).
- large instruments are made off components that are instruments on their own. Mint a separate PID for the component and link using HasComponent / IsComponentOf.
- link other resources providing supplemental information, e.g. the support page of the manufacturer providing the specs.
- describe the instrument instance using a different metadata schema and link this (HasMetadata).
Mapping onto DataCite Schema

Most properties from the generic schema may be mapped onto DataCite, some definitions need to be somewhat stretched though:

- **Identifier** → Identifier
- **Name** → Title, titleType=Other
- **Owner** → Contributor, contributorType=HostingInstitution
- **Manufacturer** → Creator
- **Description** → Description
- **Date** → Date
- **RelatedIdentifier** → RelatedIdentifier
- **AlternateIdentifier** → AlternateIdentifier
Mapping onto DataCite Schema

New properties in the DataCite Schema:

- **Publisher**: entity that created and manages this PID
- **PublicationYear**: the year when this PID has been created
- **Subject**: keyword, classification code
- **ResourceType, resourceTypeGeneral**: Instrument, Other

Missing properties:

- **modelName**
- **InstrumentType**
- **VariableMeasured**
Issues with DataCite Schema

Issues identified in the DataCite Schema:

#70 Add “Instrument” to controlled list of values for “resourceTypeGeneral”

#71 Add a value indicating “was used in” to “relationType”

#72 Add a “Series” property (to accommodate modelName)

#73 Add “Name” to controlled list of values for “titleType”

#74 Where to put a serial number of an instrument?

See:

https://github.com/datacite/schema/issues
PIDINST Journal Article

- Summarize the work done, results achieved, discussion
- Work in progress
- First draft for Introduction and Methodology sections
- Results to be completed
- Discussion, Conclusion TODO
- Aiming at submission before Xmas
- Data Science Journal Collection RDA Results
Institutional instrument providers

White paper

RDA's 14th Plenary - Helsinki, Finland

Louise Darroch (British Oceanographic Data Centre, National Oceanography Centre, UK)
Status

- Dedicated group discussion (2019)
- White paper work in progress....
Who’s it for?

Institutional instrument providers

e.g.

- Data repositories
- Metadata registries
- Hardware repositories
- Research infrastructures etc.
Landing pages

- Landing page for each PID to resolve to
- Enough information (metadata) to identify instrument
# Metadata

## Recommendations for additional, more descriptive information

<table>
<thead>
<tr>
<th>Category</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model version</td>
<td></td>
</tr>
<tr>
<td>Documents</td>
<td></td>
</tr>
<tr>
<td>Classifications</td>
<td></td>
</tr>
<tr>
<td>Identifications</td>
<td></td>
</tr>
<tr>
<td>Calibrations</td>
<td></td>
</tr>
<tr>
<td>Capabilities</td>
<td></td>
</tr>
<tr>
<td>Characteristics</td>
<td></td>
</tr>
<tr>
<td>Servicing</td>
<td></td>
</tr>
<tr>
<td>Funding references</td>
<td></td>
</tr>
<tr>
<td>Geolocation</td>
<td></td>
</tr>
<tr>
<td>Location name</td>
<td></td>
</tr>
<tr>
<td>Mounting platform</td>
<td></td>
</tr>
<tr>
<td>Position</td>
<td></td>
</tr>
<tr>
<td>Instrument orientation</td>
<td></td>
</tr>
<tr>
<td>Ownership dates</td>
<td></td>
</tr>
</tbody>
</table>
## Terminology

Enables unambiguous markup, sharing and machine-readability

Extensive set of instrument-related controlled vocabularies in marine domain

<table>
<thead>
<tr>
<th>Values</th>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observable property: Collections P01</td>
<td>SensorML Events: Collection W03</td>
</tr>
<tr>
<td>Instrument Type: Collection L05</td>
<td>SensorML Capabilities: Collection W04</td>
</tr>
<tr>
<td>Platform Type: Collection L06</td>
<td>SensorML Characteristics: Collection W05</td>
</tr>
<tr>
<td>Platform/device models: Collections L22, B76</td>
<td>SensorML Classifications: Collection W06</td>
</tr>
<tr>
<td>Roles: Collections G04, C86</td>
<td>SensorML Identifications: Collection W07</td>
</tr>
<tr>
<td>Feature of Interest: Collection C19</td>
<td>SensorML Contacts: Collection W09</td>
</tr>
<tr>
<td>Manufacturer: Collections L35, C75</td>
<td>SWE O&amp;M: Collection W09</td>
</tr>
<tr>
<td></td>
<td>SWE Data interface types: W10</td>
</tr>
</tbody>
</table>

[https://www.bodc.ac.uk/resources/vocabularies/](https://www.bodc.ac.uk/resources/vocabularies/) (available on NERC Vocabulary Server (NVS) 2.0)
Sensor web standards

- Standards and ontologies for making sensors discoverable, accessible and usable via the Web
- Uses formal descriptions of sensor instances (physical sensor)

Semantic Sensor Network (SSN) Ontology
Sensor Web Enablement (SWE)
Sea-Bird CT Sail CTD

http://linkedsystems.uk/system/instance/TOOL1188_0226/current/
Sensor Web Standards

- Different users use different ways of identifying sensors (e.g. SWE)
- Requires some community consensus
Instructions on creating PIDs

### ePIC

https://github.com/rdawg-pidinst/schema/blob/master/schema.rst

### DataCITE

https://github.com/rdawg-pidinst/schema/blob/map_to_datacite/schema.rst

---

**Metadata Schema for the Persistent Identification of Scientific Measuring Instruments**

<table>
<thead>
<tr>
<th>ID</th>
<th>Property</th>
<th>Obligation</th>
<th>Occ</th>
<th>Definition</th>
<th>Allowed values, constraints, remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Identifier</td>
<td>M</td>
<td>1</td>
<td>Unique string that identifies the instrument instance</td>
<td>POINT</td>
</tr>
<tr>
<td>1.1</td>
<td>identifierType</td>
<td>M</td>
<td>1</td>
<td>Type of the identifier</td>
<td>Controlled list of values</td>
</tr>
<tr>
<td>2</td>
<td>LandingPage</td>
<td>M</td>
<td>1</td>
<td>A landing page that the identifier resolves to</td>
<td>URL</td>
</tr>
<tr>
<td>3</td>
<td>Name</td>
<td>M</td>
<td>1</td>
<td>Name by which the instrument instance is known</td>
<td>Free text</td>
</tr>
<tr>
<td>4</td>
<td>Owner</td>
<td>M</td>
<td>1-n</td>
<td>Institutions responsible for the management of the instrument. This may include the legal owner, the operator, or an institution providing access to the instrument</td>
<td>Free text</td>
</tr>
<tr>
<td>4.1</td>
<td>ownerName</td>
<td>M</td>
<td>1</td>
<td>Full name of the owner</td>
<td>Free text</td>
</tr>
<tr>
<td>4.2</td>
<td>ownerContact</td>
<td>O</td>
<td>0-1</td>
<td>Contact address of the owner</td>
<td>Email address</td>
</tr>
<tr>
<td>4.3</td>
<td>ownerIdentifier</td>
<td>O</td>
<td>0-1</td>
<td>Persistent identifier of the owner</td>
<td>Free text</td>
</tr>
<tr>
<td>4.3.1</td>
<td>ownerIdentifierType</td>
<td>O</td>
<td>1</td>
<td>Type of the identifier</td>
<td>Free text</td>
</tr>
<tr>
<td>5</td>
<td>Manufacturer</td>
<td>M</td>
<td>1-n</td>
<td>The instrument’s manufacturer</td>
<td>Free text</td>
</tr>
</tbody>
</table>

**Metadata Schema for the Persistent Identification of Scientific Measuring Instruments**

The following table presents the metadata schema for the persistent identification of instruments mapped onto the DataCITE Metadata Schema 4.2. Note that the current version of the DataCITE schema has not been designed to describe instruments. As a consequence, some definitions in the DataCITE schema need to be stretched. For a relevant instrument property, there is even no suitable place in the DataCITE schema at all.

In this presentation, the DataCITE schema is mostly taken as is, assuming that no adaptations are made to accommodate instruments. Nonetheless, there are some shortcomings of this approach, so some amendments of the schema would facilitate its use for instruments and should be negotiated with DataCITE.

<table>
<thead>
<tr>
<th>ID</th>
<th>Property</th>
<th>Obligation</th>
<th>Occ</th>
<th>Definition</th>
<th>Allowed values, constraints, remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Identifier</td>
<td>M</td>
<td>1</td>
<td>Unique string that identifies the instrument instance</td>
<td>DOI</td>
</tr>
<tr>
<td>1.1</td>
<td>identifierType</td>
<td>M</td>
<td>1</td>
<td>Type of the identifier</td>
<td>Controlled list of values</td>
</tr>
<tr>
<td>2</td>
<td>Creator</td>
<td>M</td>
<td>1-n</td>
<td>The instrument’s manufacturer or developer. This may also be the owner for custom built instruments</td>
<td>Free text</td>
</tr>
<tr>
<td>2.1</td>
<td>creatorName</td>
<td>M</td>
<td>1</td>
<td>Full name of the manufacturer</td>
<td>Free text</td>
</tr>
<tr>
<td>2.1.1</td>
<td>creatorNameType</td>
<td>R</td>
<td>0-1</td>
<td>The type of the name</td>
<td>Free text</td>
</tr>
<tr>
<td>2.2</td>
<td>ownerName</td>
<td>M</td>
<td>1</td>
<td>Full name of the manufacturer</td>
<td>Free text</td>
</tr>
<tr>
<td>2.3</td>
<td>familyName</td>
<td>R</td>
<td>0-1</td>
<td>Last name of the manufacturer</td>
<td>Free text</td>
</tr>
</tbody>
</table>
Dos and don’ts (including)

- Duplication
- When to create a new PID (instrument upgrades)
ePIC and PIDINST
Wrapping-up PIDINST WG

Ulrich Schwardmann (GWDG)

Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG)
Am Fassberg, 37077 Göttingen
ulrich.schwardmann [at] gwdg.de

23 October 2019, Helsinki
Interoperability by Registration of Types

RDA working group on
Data Type Registries

- approach to provide *type definitions*
- a PID for each definition
- defines the type structure, its use and semantics
- CORDRA as DTR service
- typical use cases:
  - with given PID find a type and ask for its use at DTR (see left)
  - ask at DTR for types with given semantics and find via PIDs according data
The ePIC Data Type Registry

- **Features**
  - Definition of PID Information Types
  - hierarchical types and automated schema extraction
  - Access via REST API, Browser

- based on CORDRA software
- GWDG is provider on behalf of ePIC
- Who can use the service?
  - public, authorization needed only for type definition

Overview: [http://dtr.pidconsortium.eu/](http://dtr.pidconsortium.eu/)

PID InfoType states are:

- *in preparation* (21.T11148),
- *candidate, approved, deprecated* (21.11104)
  - [http://dtr-pit.pidconsortium.eu/](http://dtr-pit.pidconsortium.eu/)
Hierarchical Type Definitions at ePIC DTR

- types are often dependent from each other, how exactly?
- to exactly describe JSON objects by data types one needs:
  - a distinction between derived objects and basic objects
    - concept of *basic PID info types* and *PID info types*
  - a more exact description of the type dependencies
  - additionally a JSON schema inspired dependency model
- in consequence:
  - possibility to derive JSON schemas for the type values
    - automated server side schema derivation at ePIC DTR
  - one type defines in an exact way its whole dependencies
    - in objects of a certain type one can use the names of its parts (instead of type identifiers)
- see also Schwardmann, U.: Automated schema extraction for PID information types
  - PID: http://hdl.handle.net/21.11101/0000-0002-A987-7
### Defining the PID4Instruments InfoType

```json
{  
  allowOmitSubsidiaries: "Yes",
  name: "Owners",

  identifier: "21.T11148/1f3e82dd0697a497432",

  representationsAndSemantics: {
    0: {
      obligation: "Mandatory",
      repeatable: "No",
      expression: "",
      allowOmitSubsidiaries: "Yes",
      value: "",
      name: "Manufacturers",
    }

    5: {
      identifier: "21.T11148/55f8ebc805e65b5b71dd",
      representationsAndSemantics: {
        0: {
          obligation: "Optional",
          repeatable: "No",
          expression: "",
          allowOmitSubsidiaries: "Yes",
          value: "",
          name: "Description",
        }

        6: {
          identifier: "21.T11148/f76ad9d8324302fc47dd",
        }
      }
    }
  }
}
```
Where to find the definitions

- find the complete JSON of the type definition for type Properties-PID-instruments. with:
  http://hdl.handle.net/
  21.T11148/17ce618137e697852ea6

- for a more user friendly layout use
  http://dtr-test.pidconsortium.eu/#objects/
  21.T11148/17ce618137e697852ea6
  • Here you also has the pointers to all the defined subtypes for PIDINST.
Representing the Nanocluster Example

mapping the subtypes as single entries in the Handle record

* 21.T11998/0000-001A-1379-7

http://hdl.handle.net/21.T11998/0000-001A-1379-7

Create new value Save handle Delete handle Refresh Sort by none Collapse all QR Code Sign Verify

Delete Type: URL

https://www.helmholtz-berlin.de/pubbin/fgama_output?modus=einzel&amp;sprache=en&amp;gid=1848&a...

Delete Type: 21.T11148/8ebd55ee0b

("IdentifierValue": "21.T11998/0000-001A-1379-7", "IdentifierType": "PIDINST")

Delete Type: 21.T11148/3a135a4733d

https://www.helmholtz-berlin.de/pubbin/fgama_output?modus=einzel&amp;sprache=en&amp;gid=1848&a...

Delete Type: 21.T11148/799a23220c

NanoclusterTrap

Delete Type: 21.T11148/4eacc4b0f81

["Owner": "Helmholtz-Zentrum Berlin fu00for Materialien und Energie"]

Delete Type: 21.T11148/1056d2e0f11

["Manufacturer": "Helmholtz-Zentrum Berlin fu00for Materialien und Energie"]

Delete Type: 21.T11148/35ebe00c0c

The Nanocluster Trap endstation at BESSY II combines a cryogenic linear radio-frequency ion trap with an ...

Delete Type: 21.T11148/766d9d032

Experimental station for x-ray magnetic circular dichroism (XMCD)
Many Thanks

Questions ???

Contact at ePIC:
- support [at] pidconsortium.eu

Contact at GWDG:
- **Ulrich Schwardmann**
  - T: +49 551 201-1542, E: ulrich.schwardmann [at] gwdg.de
DataCite DOIs for Instruments at HZB

Rolf Krahl

Persistent Identification of Instruments WG © RDA P14, Helsinki, 23 October 2019
Helmholtz-Zentrum Berlin für Materialien und Energie operates a neutron source BER II (closing down by end 2019) and a synchrotron light source BESSY II:

- 10 beamlines at BER II (in full operation).
- 35 beamlines at BESSY II (in full operation).
- 43 experimental stations at BESSY II (in full operation).

User facilities: external users may submit a proposal and will be granted experiment time at the instruments (based on the scientific merit of the proposal).

We plan to mint DataCite DOIs for these instruments and just started to mint the first four of them.
The NanoclusterTrap experimental station:
https://doi.org/10.5442/ni000004

The corresponding PID kernel record:
https://doi.org/10.5442/ni000004?noredirect

The corresponding search result at DataCite:
https://search.datacite.org/works?query=10.5442%2Fni000004
British Oceanographic Data Centre (BODC)

Louise Darroch
http://hdl.handle.net/21.T11148/17ce618137e697852ea6

http://dtr-test.pidconsortium.eu/#objects/21.T11148/17ce618137e697852ea6
Get capabilities

Use offering

23/10/2019    rd-alliance.org    @resdatall | @rda_europe
Integrated Carbon Observation System (ICOS)

• ICOS operates ca 150 measurement stations in 12 countries
• GHG concentrations & fluxes, meteo, environmental monitoring
• Three domains: Atmosphere, Ecosystem, Marine
• Currently ICOS metadata about instruments include type, serial number, settings, location, calibration coefficients, ...
• These are collected by stations and submitted to ICOS Thematic Centers that process the data
• Workflow to pass on all relevant metadata to the ICOS data repository (Carbon Portal) is being implemented now
Mapping of ICOS to PIDINST data model

• We need to add a few properties to the ICOS data model to be fully compliant

• The PIDINST fields are assembled from different “Classes” or “Categories” within the ICOS metadata specification (see following slide)

• Lesson learned from the mapping:
  • Metadata about metadata entries are not formally described
  • Is it necessary to provide two pieces of information about the URL? E.g. both the URL string itself and then also the specification of what “type” the URL is (“http url link”)...
PIDINST entries are combined from ICOS view of Data Object

Digital Data Object (TimeSeries)

Contribution

Measurement
Pre Processing
Raw Data
Sensor & Instrument Information

- People
- Station
- Instruments

Sensor & Instrument
- Calibration, SerialNumber ...

Immutable Hardware
- Manufacturer, Accuracy, Resolution ...

Installation & Configuration
- Location, Orientation, Calibration ...

Conversion / Transformation
- Timezone, Engineering Unit ...

M. Hellström & C. D’Onofrio, RDA P14 @Helsinki 2019-10-23
Two related Use Cases:

- AWI sensor.awi use case 1 –
  - Observation 2 Archive framework (O2A) includes SENSOR.awi
    - curated metadata repository for sensor information
    - *Handles* as PIDs for sensors
    - Automated data flow from ship to archive
  - Sensor data commonly published in PANGAEA ([https://pangaea.de/](https://pangaea.de/))
  - Use Case focus:
    - Follow standard metadata schema for sensors
    - Moving from *handles* to externally minted PIDs (e.g. DOIs)
    - Solutions for versioning of sensors (provenance info)
Two related Use Cases:

- **AWI sensor.awi Use Case 1** –
  - Focus on FAIR data for users
  - Datasets are archived with their metadata in a relational database
  - Standard terminologies (e.g. parameters, methods, *device types* (1364))
  - Data published and registered with Datacite DOIs
  - for advanced interactions, related web services and APIs available (e.g. OAI-PMH metadata provider, Elasticsearch API, data warehouse)

- **PANGAEAE Data Publisher Use Case 2** –
Two related Use Cases:

- AWI sensor.awi Use Case 1 –
- PANGAEA Data Publisher User Case 2 –

  Use Case focus:

  - Device PIDs as part of event data - enrich metadata (FAIR)
  - Remove ambiguity between method and device use
  - Linking PIDs with graph architecture (PID Graph)

https://project-freya.eu
Wulff, Thorben; Bauerfeind, Eduard; von Appen, Wilken-Jon; Wulff, Uwe; Hagemann, Jonas; Lehmenhecker, Sascha (2018): Vertical profiles of physical and biogeochemical parameters obtained by AWI's AUV "PAUL" during a dive in the vicinity of an ice tongue in the Fram Strait in 2013. PANGAEA, https://doi.org/10.1594/PANGAEA.887579

AWI's autonomous underwater vehicle "PAUL" covered two 10 km long transects in the Fram Strait on July 2nd / 3rd 2013 to investigate the physical-ecological coupling at an ice edge. The dive was orientated perpendicular to a meltwater front. The meltwater front was associated to a large ice tongue extending from the main ice edge. Every 800 - 1000 m, the vehicle ascended vertically from 50 m water depth to a minimal depth of 3 m to gather a high resolution profile of the following parameters: Temperature, Conductivity, Pressure, Chlorophyll a, CDOM, Dissolved Oxygen, Photosynthetically Active Radiation, and Nitrate.

Project(s):
Physical Oceanography @ AWI (AWI_Phyoce)

Coverage:
Median Latitude: 78.753080 * Median Longitude: 5.144880 * South-bound Latitude: 78.714727 * West-bound Longitude: 5.100582 * North-bound Latitude: 78.794343 * East-bound Longitude: 5.185734
Date/Time Start: 2013-07-02T20:45:38 * Date/Time End: 2013-07-03T01:35:26
Minimum DEPTH, water: 1.22 m * Maximum DEPTH, water: 52.62 m

Event(s):
MSM29_440-5 * Latitude Start: 78.714170 * Longitude Start: 5.160830 * Latitude End: 78.715330 * Longitude End: 5.158000 * Date/Time Start: 2013-07-02T19:58:00 * Date/Time End: 2013-07-03T02:58:00 * Elevation Start: -2323.2 m * Elevation End: -2323.0 m * SENSOR AWI: hld:10013/sensor.664525cf.45b9.4f99.hh88.91a5c5e97a5b * Location: North Greenland Sea * Campaign: MSM29 (HALISGARTEN 2013) * Basis: Maria S. Merian * Device: Autonomous underwater vehicle (AUV)

Parameter(s):

<table>
<thead>
<tr>
<th>#</th>
<th>Name</th>
<th>Start Name</th>
<th>Start Name</th>
<th>Name</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DATE/TIME</td>
<td>Date/Time</td>
<td>Wulff, Thorben</td>
<td>Gecode</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>LATITUDE</td>
<td>Latitude</td>
<td>Wulff, Thorben</td>
<td>Gecode</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>LONGITUDE</td>
<td>Longitude</td>
<td>Wulff, Thorben</td>
<td>Gecode</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>DEPTH, water</td>
<td>Depth water</td>
<td>Wulff, Thorben</td>
<td>Gecode - PAR</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Radiation, photosynthetically active</td>
<td>PAR</td>
<td>Wulff, Thorben</td>
<td>corrected</td>
<td></td>
</tr>
</tbody>
</table>
**Related to:**

**Project(s):**
Physical Oceanography @ AWI (AWI_PhyOce)

**Coverage:**
- Median Latitude: 78.753080
- Median Longitude: 5.144880
- South-bound Latitude: 78.714727
- West-bound Longitude: 5.100582
- North-bound Latitude: 78.794343
- East-bound Longitude: 5.185734

**Date/Time Start:** 2013-07-02T20:45:38
**Date/Time End:** 2013-07-03T01:35:26

**Minimum DEPTH, water:** 1.22 m
**Maximum DEPTH, water:** 52.62 m

**Event(s):**
MSM29 440-5
- Latitude Start: 78.714170
- Longitude Start: 5.160830
- Latitude End: 78.715330
- Longitude End: 5.158000
- Date/Time Start: 2013-07-02T19:58:00
- Date/Time End: 2013-07-03T02:58:00
- Elevation Start: -2332.3 m
- Elevation End: -2332.0 m
- SENSOR AWI: hdl:10013/sensor.664525cf-45b9-4969-bb88-91a1c5e97a5b
- Location: North Greenland Sea
- Campaign: MSM29 (HAUSGARTEN 2013)

**Parameter(s):**

<table>
<thead>
<tr>
<th>#</th>
<th>Name</th>
<th>Short Name Unit</th>
<th>Principal Investigator</th>
<th>Method Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DATE/TIME</td>
<td>Date/Time</td>
<td>Wulff, Thorben</td>
<td>Geocode</td>
</tr>
<tr>
<td>2</td>
<td>LATITUDE</td>
<td>Latitude</td>
<td>Wulff, Thorben</td>
<td>Geocode</td>
</tr>
<tr>
<td>3</td>
<td>LONGITUDE</td>
<td>Longitude</td>
<td>Wulff, Thorben</td>
<td>Geocode</td>
</tr>
<tr>
<td>4</td>
<td>DEPTH, water</td>
<td>Depth water m</td>
<td>Wulff, Thorben</td>
<td>Geocode – PAR</td>
</tr>
<tr>
<td>5</td>
<td>Radiation, photosynthetically active</td>
<td>PAR</td>
<td>Wulff, Thorben</td>
<td>corrected</td>
</tr>
</tbody>
</table>
Current Approach: SENSOR PIDs

- Cite-enabled PIDs (type: handle) - minted upon demand
  
  The former automated generation of versioning with individual PIDs for events of type deployment, mission, calibration, etc. (as described in Use Case document) has been adapted to fit recent on-board integration requirements (e.g. synchronization with acquisition system DSHIP on board of RV Polarstern)

- Manufacturer, model, and affiliation attribute not mandatory, free text

- Item type assignment not supported (resourceType with values platform, device, sensor as proposed in DataCite schema)

- SENSOR.awi.de PIDs are linked to data in PANGAEA under „Events“. Example: https://doi.pangaea.de/10.1594/PANGAEA.887579
Future Approach: SENSOR PIDs

- All parent and children items placed in „live archive“ will have individual cite-enabled PIDs (type handle)

- Metadata cross-walk SensorML 2.0 -> PID RDA, DataCite will be supported
  - Adoption of NERC controlled vocabularies for „Manufacturer“ [http://vocab.nerc.ac.uk/collection/L35/current](http://vocab.nerc.ac.uk/collection/L35/current)
  - Adoption of Research Organization Registry [https://www.ror.community/](https://www.ror.community/)
  - Adoption of Standard values for unknown information from DataCite, [DataCite 4.3 Metadata Schema Documentation](https://www.datacite.org/metadata)

- Parent items placed in „live archive“ will be automatically registered against DataCite using a pre-defined DOI syntax. New landing pages will display DataCite DOI as additional identifier.

- Workflows SENSOR ➔ PANGAEA will be revised, optimized
PID Graph

**Established PIDs**
- Publications (DOI)
- Researchers (ORCID)
- Institute/Organisations (ROR IDs)
- Measurement Protocol/best practices (DOI)
- Data (DOI)

**Emerging PIDs**
- Instruments
- Instrument Manufacturers
- Facility
- Research Infrastructure

**Instrument – User Stories:**
- Researcher - Find and evaluate related data
- Facility - Evaluate impact of research from instrument in facility
- RI - Follow careers of instrument users
- Researcher - Find other data from same instrument
- Researcher - Find researchers and their research which use the same instrument
- Manufacturer - Find research that own instruments were used in
Issues:

- **Granularity** - Shall changes in payload be registered in DataCite under a new PID?
  - Alternative is to keep track of payload changes and possible sensor calibration events with a local UID in chronological order, to be displayed in the new landing pages

- No automated mechanism for assigning resourceType (platform, instrument, sensor,..) to items

- Device terminology harmonization

Questions? Email **ana.macario@awi.de** or **tdohna@marum.de**
Discussion

- Transition to “maintenance mode”: Reasonable or do you see immediate needs for a new WG?
- Any early adopters in the room, especially instrument database providers?
- Interest in contributing to the deliverables?
- How to publish the results widely? What are relevant RDA groups? How about beyond RDA?