Persistent Identification of Instruments WG (PIDINST WG)

Markus Stocker (@markusstocker) and The #PIDINST Team

tinyurl.com/y54ycq8d

Markus Stocker
@markusstocker

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

Notes

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

ID	Property	Occ
1	Identifier (with identifierType)	1
2	LandingPage	1
3	Name	1
4	Owner (with ownerName, ownerContact, and owner-Identifier)	1-n
5	Manufacturer (with manufacturerName and manufac-turerIdentifier)	1-n

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

ID	Property	Occ
5.3	modelName (subproperty of Manufacturer)	0-1
6	Description	0-1
7	InstrumentType	0-1
8	VariableMeasured	0-n
9	Date (with dateType)	0-1
10	RelatedIdentifier (with relatedIdentifierType and relationType)	0-n
11	AlternateIdentifier (with alternateIdentifierType)	0-n

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).
  - versioning: after major modification, mint a new PID and link using IsNewVersionOf / IsPreviousVersionOf.
  - 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	$\rightarrow$ Identifier
Name	$\rightarrow$ Title, titleType=Other
Owner	$\rightarrow$ Contributor, contributorType=HostingInstitution
Manufacturer	$\rightarrow$ Creator
Description	$\rightarrow$ Description
Date	$\rightarrow$ Date
RelatedIdentifier	$\rightarrow$ RelatedIdentifier
AlternateIdentifier	$\rightarrow$ 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
  - https://datascience.codata.org/collections/special/research-data-alliance-results/

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.
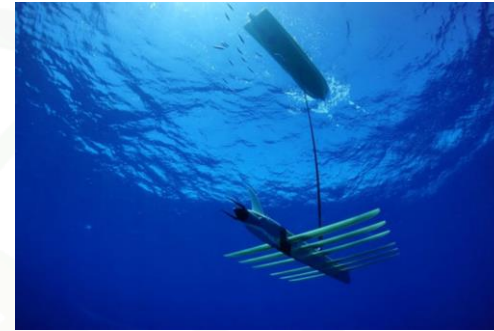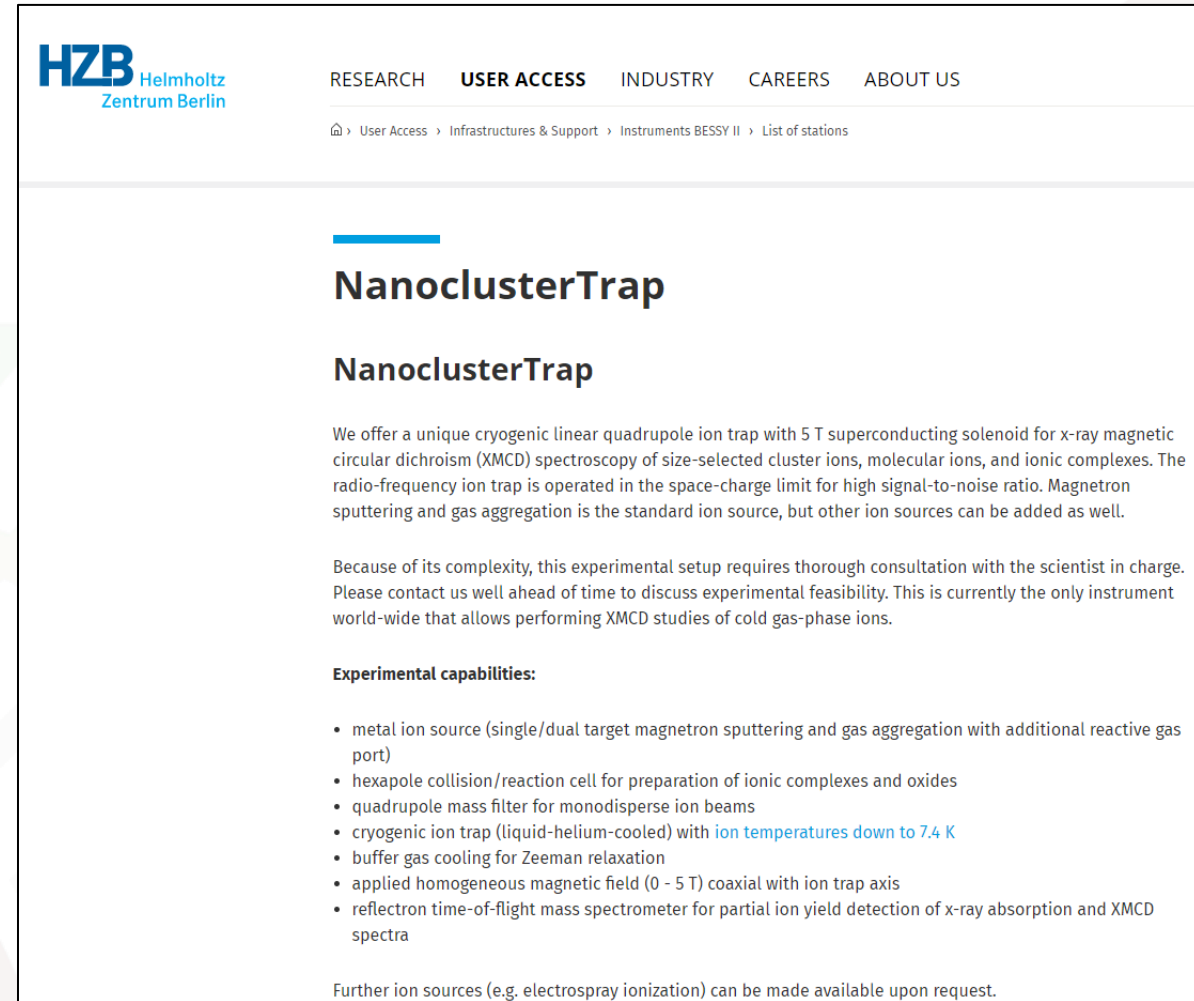
Landing pages

- Landing page for each PID to resolve to

- Enough information (metadata) to identify instrument

HZB Helmholtz Zentrum Berlin

RESEARCH   **USER ACCESS**   INDUSTRY   CAREERS   ABOUT US

⌂ › User Access › Infrastructures & Support › Instruments BESSY II › List of stations

NanoclusterTrap

NanoclusterTrap

We offer a unique cryogenic linear quadrupole ion trap with 5 T superconducting solenoid for x-ray magnetic circular dichroism (XMCD) spectroscopy of size-selected cluster ions, molecular ions, and ionic complexes. The radio-frequency ion trap is operated in the space-charge limit for high signal-to-noise ratio. Magnetron sputtering and gas aggregation is the standard ion source, but other ion sources can be added as well.

Because of its complexity, this experimental setup requires thorough consultation with the scientist in charge. Please contact us well ahead of time to discuss experimental feasibility. This is currently the only instrument world-wide that allows performing XMCD studies of cold gas-phase ions.

Experimental capabilities:

- metal ion source (single/dual target magnetron sputtering and gas aggregation with additional reactive gas port)
- hexapole collision/reaction cell for preparation of ionic complexes and oxides
- quadrupole mass filter for monodisperse ion beams
- cryogenic ion trap (liquid-helium-cooled) with ion temperatures down to 7.4 K
- buffer gas cooling for Zeeman relaxation
- applied homogeneous magnetic field (0 - 5 T) coaxial with ion trap axis
- reflectron time-of-flight mass spectrometer for partial ion yield detection of x-ray absorption and XMCD spectra

Further ion sources (e.g. electrospray ionization) can be made available upon request.

Metadata

Recommendations for additional, more descriptive information

- Model version
- Documents
- Classifications
- Identifications
- Calibrations
- Capabilities
- Characteristics

- Servicing
- Funding references
- Geolocation
- Location name
- Mounting platform
- Position
- Instrument orientation
- Ownership dates

Terminology

Enables unambiguous markup, sharing and machine-readability

Extensive set of instrument-related controlled vocabularies in marine domain

Values	Properties
Observable property: Collections P01	SensorML Events: Collection W03
Instrument Type: Collection L05	SensorML Capabilities: Collection W04
Platform Type: Collection L06	SensorML Characteritics: Collection W05
Platform/**device models**: Collections **L22**, B76	SensorML Classifications: Collection W06
Roles: Collections G04, C86	SensorML Identifications: Collection W07
Feature of Interest: Collection C19	SensorML Contacts: Collection W09
Manufacturer: Collections L35, C75	SWE O&M: Collection W09
	SWE Data interface types: W10

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)
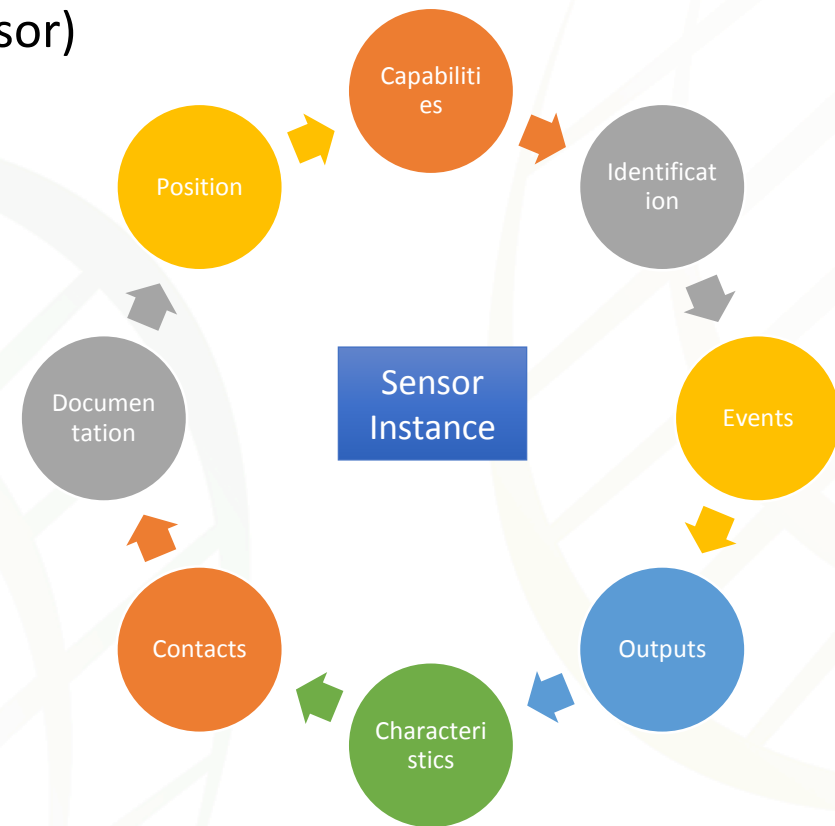
Sensor Web Enablement SensorML

http://linkedsystems.uk/system/instance/TOOL1188_0226/current/

Sea-Bird CT Sail CTD

```xml
<sml:PhysicalSystem xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:gmd="http://www.isotc211.org/2005/gmd" xmlns:gco="http://www.isotc211.org/2005/gco"
xmlns:sml="http://www.opengis.net/sensorml/2.0" xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:gml="http://www.opengis.net/gml/3.2" xmlns:swe="http://www.opengis.net/swe/2.0"
xsi:schemaLocation="http://www.opengis.net/swes/2.0 http://schemas.opengis.net/swes/2.0/swesDescribeSensor.xsd http://www.opengis.net/sensorml/2.0
http://schemas.opengis.net/sensorML/2.0/sensorML.xsd http://www.isotc211.org/2005/gmd http://schemas.opengis.net/iso/19139/20070417/gmd/gmd.xsd http://www.isotc211.org/2005/gco
http://schemas.opengis.net/iso/19139/20070417/gco/gco.xsd http://www.opengis.net/gml/3.2 http://schemas.opengis.net/gml/3.2.1/gml.xsd" gml:id="TOOL1188_0226">
  <gml:description>Seaglider CT sail 0226</gml:description>
  <gml:identifier codeSpace="uniqueID">TOOL1188_0226</gml:identifier>
  <sml:keywords>
    <sml:KeywordList>
      <sml:keyword>Pres_Z</sml:keyword>
      <sml:keyword>CTDCond</sml:keyword>
      <sml:keyword>WC_temp_CTD</sml:keyword>
      <sml:keyword>SigTheta</sml:keyword>
      <sml:keyword>P_sal_CTD</sml:keyword>
    </sml:KeywordList>
  </sml:keywords>
  <sml:identification>
    <sml:IdentifierList>
      <sml:identifier>
        <sml:Term definition="TOOL1188_0226">
          <sml:label>UUID</sml:label>
          <sml:value>TOOL1188_0226</sml:value>
        </sml:Term>
      </sml:identifier>
      <sml:identifier>
        <sml:Term definition="http://vocab.nerc.ac.uk/collection/W07/current/IDEN0006/">
          <sml:label>Short Name</sml:label>
          <sml:value>Seaglider CT sail 0226</sml:value>
        </sml:Term>
      </sml:identifier>
      <sml:identifier>
        <sml:Term definition="http://vocab.nerc.ac.uk/collection/W07/current/IDEN0002/">
          <sml:label>Long Name</sml:label>
          <sml:value>Seaglider CT sail 0226</sml:value>
        </sml:Term>
      </sml:identifier>
    </sml:IdentifierList>
  </sml:identification>
  <sml:validTime>
    <gml:TimePeriod gml:id="validityPeriod">
      <gml:beginPosition>2018-04-30</gml:beginPosition>
      <gml:endPosition>Indeterminate</gml:endPosition>
    </gml:TimePeriod>
  </sml:validTime>
  <sml:history>
    <sml:EventList/>
  </sml:history>
  <sml:typeOf xlink:title="Sea-Bird_Scientific_Unpumped_Seaglider_CT_sail" xlink:href="http://linkedsystems.uk/system/prototype/TOOL1188/"/>
  <sml:outputs>
    <sml:OutputList>
      <sml:output name="P_sal_CTD">
        <swe:Quantity definition="http://vocab.nerc.ac.uk/collection/P01/current/PSALST01/">
          <swe:uom code="UCUMCODE" xlink:href="http://vocab.nerc.ac.uk/collection/P06/current/UUUU"/>
        </swe:Quantity>
      </sml:output>
      <sml:output name="WC_temp_CTD">
        <swe:Quantity definition="http://vocab.nerc.ac.uk/collection/P01/current/TEMPST01/">
          <swe:uom code="UCUMCODE" xlink:href="http://vocab.nerc.ac.uk/collection/P06/current/UPAA"/>
        </swe:Quantity>
      </sml:output>
```

Sensor Web Standards

- Different users use different ways of identifying sensors (e.g. SWE)

- Requires some community consensus

Instructions on creating PIDs

ePIC

ePICC ◦◦◦
Persistent Identifiers for eResearch

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

DataCITE

DataCite

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

---

187 lines (167 sloc) | 15.8 KB   Raw | Blame | History

Metadata Schema for the Persistent Identification of Scientific Measuring Instruments

ID	Property	Obligation	Occ	Definition	Allowed values, constraints, remarks
1	Identifier	M	1	Unique string that identifies the instrument instance	PIDINST
1.1	identifierType	M	1	Type of the identifier	Controlled list of values: PIDINST
2	LandingPage	M	1	A landing page that the identifier resolves to	URL
3	Name	M	1	Name by which the instrument instance is known	Free text
4	Owner	M	1-n	Institution(s) responsible for the management of the instrument. This may include the legal owner, the operator, or an institute providing access to the instrument.	
4.1	ownerName	M	1	Full name of the owner	Free text
4.2	ownerContact	O	0-1	Contact address of the owner	Electronic mail address
4.3	ownerIdentifier	O	0-1	Persistent identifier of the owner	Free text, must be globally unique identifiers.
4.3.1	ownerIdentifierType	O	1	Type of the identifier	Free text
5	Manufacturer	M	1-n	The instrument's manufacturer(s) or	

---

363 lines (340 sloc) | 30.4 KB   Raw | Blame | History

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 few relevant instrument properties 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. Nevertheless, 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.

ID	Property	Obligation	Occ	Definition	Allowed values, constraints, remarks
1	Identifier	M	1	Unique string that identifies the instrument instance	DOI
1.1	identifierType	M	1	Type of the identifier	Controlled list of values:[1] DOI
2	Creator	M	1-n	The instrument's manufacturer(s) or developer. This may also be the owner for custom build instruments	
2.1	creatorName	M	1	Full name of the manufacturer	Free text
2.1.1	nameType	R	0-1	The type of name	Controlled list of values:[2] Organizational Personal
2.2	givenName	R	0-1	First name of the manufacturer, if applicable	Free text
2.3	familyName	R	0-1	Last name of the manufacturer, if applicable	Free text

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

ePIC
Persistent Identifiers for eResearch

Interoperability by Registration of Types
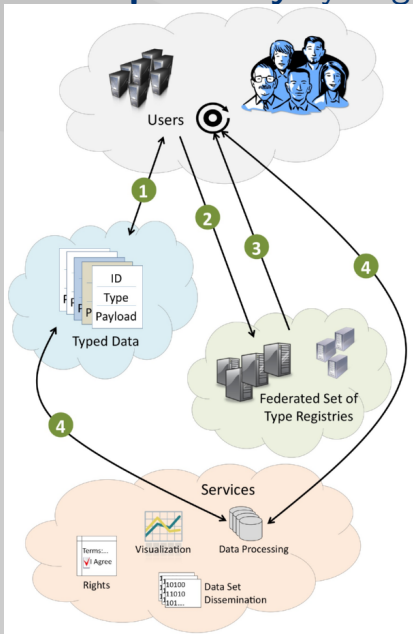
RDA working group on
Data Type Registries

- 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

ePIC and
PIDINST

Ulrich
Schwardmann
(GWDG)

Interoperability
and
Registration of
Types

Data Type
Registries

ePIC DTR
Adaption of
the MD
Schema for
Instruments

Questions

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/`

PID InfoType states are:

- *in preparation* (21.T11148),
  - `http://dtr-test.pidconsortium.eu/`
- *candidate, approved, deprecated* (21.11104)
  - `http://dtr-pit.pidconsortium.eu/`

ePIC and PIDINST

Ulrich Schwardmann (GWDG)

Interoperability and Registration of Types

Data Type Registries

ePIC DTR Adaption of the MD Schema for Instruments

Questions

Hierarchical Type Definitions at ePIC DTR

ePIC and
PIDINST

Ulrich
Schwardmann
(GWDG)

Interoperability
and
Registration of
Types

Data Type
Registries

ePIC DTR
Adaption of
the MD
Schema for
Instruments

Questions

- 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

ePIC
Persistent Identifiers for eResearch

Defining the PID4Instruments InfoType

JSON   Rohdaten   Kopfzeilen

Speichern   Kopieren   JSON durchsuchen

          allowOmitSubsidiaries: "Yes"
          value: ""
      name: "Owners"
  ▼ 4:
      identifier: "21.T11148/1f3e82ddf0697a497432"
      ▼ 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/f76ad9d0324302fc47dd"

Where to find the definitions

ePIC and
PIDINST

Ulrich
Schwardmann
(GWDG)

Interoperability
and
Registration of
Types

Data Type
Registries

ePIC DTR
Adaption of
the MD
Schema for
Instruments

Questions

- 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

ePIC and
PIDINST

Ulrich
Schwardmann
(GWDG)

Interoperability
and
Registration of
Types
Data Type
Registries

ePIC DTR
Adaption of
the MD
Schema for
Instruments

Questions

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

ePIC
Persistent Identifiers for eResearch

DataCite DOIs for Instruments at HZB

Rolf Krahl

Persistent Identification of Instruments WG @ RDA P14,
Helsinki, 23 October 2019

HZB

- 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.

A DOI for an Experimental Station

- 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`

Sensor Web Enablement SensorML

Sea-Bird CT Sail CTD

Conductivity Sensor

Entrance to C Sensor
in Seaglider flowpath

Temperature Sensor

```xml
<sml:PhysicalSystem xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:gmd="http://www.isotc211.org/2005/gmd" xmlns:gco="http://www.isotc211.org/2005/gco"
 xmlns:sml="http://www.opengis.net/sensorml/2.0" xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:gml="http://www.opengis.net/gml/3.2" xmlns:swe="http://www.opengis.net/swe/2.0"
 xsi:schemaLocation="http://www.opengis.net/swes/2.0 http://schemas.opengis.net/swes/2.0/swesDescribeSensor.xsd http://www.opengis.net/sensorML/2.0
http://schemas.opengis.net/sensorML/2.0/sensorML.xsd http://www.isotc211.org/2005/gmd http://schemas.opengis.net/iso/19139/20070417/gmd/gmd.xsd http://www.isotc211.org/2005/gco
http://schemas.opengis.net/iso/19139/20070417/gco/gco.xsd http://www.opengis.net/gml/3.2 http://schemas.opengis.net/gml/3.2.1/gml.xsd" gml:id="TOOL1188_0226">
  <gml:description>Seaglider CT sail 0226</gml:description>
  <gml:identifier codeSpace="uniqueID">TOOL1188_0226</gml:identifier>
  <sml:keywords>
    <sml:KeywordList>
      <sml:keyword>Pres_Z</sml:keyword>
      <sml:keyword>CTDCond</sml:keyword>
      <sml:keyword>WC_temp_CTD</sml:keyword>
      <sml:keyword>SigTheta</sml:keyword>
      <sml:keyword>P_sal_CTD</sml:keyword>
    </sml:KeywordList>
  </sml:keywords>
  <sml:identification>
    <sml:IdentifierList>
      <sml:identifier>
        <sml:Term definition="TOOL1188_0226">
          <sml:label>UUID</sml:label>
          <sml:value>TOOL1188_0226</sml:value>
        </sml:Term>
      </sml:identifier>
      <sml:identifier>
        <sml:Term definition="http://vocab.nerc.ac.uk/collection/W07/current/IDEN0006/">
          <sml:label>Short Name</sml:label>
          <sml:value>Seaglider CT sail 0226</sml:value>
        </sml:Term>
      </sml:identifier>
      <sml:identifier>
        <sml:Term definition="http://vocab.nerc.ac.uk/collection/W07/current/IDEN0002/">
          <sml:label>Long Name</sml:label>
          <sml:value>Seaglider CT sail 0226</sml:value>
        </sml:Term>
      </sml:identifier>
    </sml:IdentifierList>
  </sml:identification>
  <sml:validTime>
    <gml:TimePeriod gml:id="validityPeriod">
      <gml:beginPosition>2018-04-30</gml:beginPosition>
      <gml:endPosition>Indeterminate</gml:endPosition>
    </gml:TimePeriod>
  </sml:validTime>
  <sml:history>
    <sml:EventList/>
  </sml:history>
  <sml:typeOf xlink:title="Sea-Bird_Scientific_Unpumped_Seaglider_CT_sail" xlink:href="http://linkedsystems.uk/system/prototype/TOOL1188/"/>
  <sml:outputs>
    <sml:OutputList>
      <sml:output name="P_sal_CTD">
        <swe:Quantity definition="http://vocab.nerc.ac.uk/collection/P01/current/PSALST01/">
          <swe:uom code="UCUMCODE" xlink:href="http://vocab.nerc.ac.uk/collection/P06/current/UUUU"/>
        </swe:Quantity>
      </sml:output>
      <sml:output name="WC_temp_CTD">
        <swe:Quantity definition="http://vocab.nerc.ac.uk/collection/P01/current/TEMPST01/">
          <swe:uom code="UCUMCODE" xlink:href="http://vocab.nerc.ac.uk/collection/P06/current/UPAA"/>
        </swe:Quantity>
      </sml:output>
```

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

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

Sensor Web Enablement SensorML

http://hdl.handle.net/21.T11998/0000-001A-3905-F?noredirect

Get capabilities

Use offering

```
580.        },
581.        "offering" : {
582.            "allowedValues" : [
583.                "http://linkedsystems.uk/system/offering/1225804/",
584.                "http://linkedsystems.uk/system/offering/1225816/",
585.                "http://linkedsystems.uk/system/offering/1225828/",
586.                "http://linkedsystems.uk/system/offering/1225841/",
587.                "http://linkedsystems.uk/system/offering/1225853/",
588.                "http://linkedsystems.uk/system/offering/1225865/",
589.                "http://linkedsystems.uk/system/offering/1225877/",
590.                "http://linkedsystems.uk/system/offering/1225889/",
591.                "http://linkedsystems.uk/system/offering/1225890/",
592.                "http://linkedsystems.uk/system/offering/1225908/",
593.                "http://linkedsystems.uk/system/offering/1225921/",
594.                "http://linkedsystems.uk/system/offering/1225933/",
595.                "http://linkedsystems.uk/system/offering/1225945/",
596.                "http://linkedsystems.uk/system/offering/1225957/",
597.                "http://linkedsystems.uk/system/offering/1225969/",
598.                "http://linkedsystems.uk/system/offering/1225970/",
599.                "http://linkedsystems.uk/system/offering/1225982/",
600.                "http://linkedsystems.uk/system/offering/1225994/",
601.                "http://linkedsystems.uk/system/offering/1226008/",
602.                "http://linkedsystems.uk/system/offering/1226021/",
603.                "http://linkedsystems.uk/system/offering/1226033/",
604.                "http://linkedsystems.uk/system/offering/1226045/",
605.                "http://linkedsystems.uk/system/offering/1226057/",
606.                "http://linkedsystems.uk/system/offering/1226069/",
607.                "http://linkedsystems.uk/system/offering/1226070/",
608.                "http://linkedsystems.uk/system/offering/1226082/",
609.                "http://linkedsystems.uk/system/offering/1226094/",
610.                "http://linkedsystems.uk/system/offering/1226101/",
611.                "http://linkedsystems.uk/system/offering/1226113/",
612.                "http://linkedsystems.uk/system/offering/1226125/",
613.                "http://linkedsystems.uk/system/offering/1226137/",
614.                "http://linkedsystems.uk/system/offering/1226149/",
615.                "http://linkedsystems.uk/system/offering/1226150/",
616.                "http://linkedsystems.uk/system/offering/1226162/",
617.                "http://linkedsystems.uk/system/offering/1226174/",
618.                "http://linkedsystems.uk/system/offering/1226186/",
619.                "http://linkedsystems.uk/system/offering/1226198/",
620.                "http://linkedsystems.uk/system/offering/1226205/",
621.                "http://linkedsystems.uk/system/offering/1226217/",
622.                "http://linkedsystems.uk/system/offering/1226229/",
623.                "http://linkedsystems.uk/system/offering/1226230/",
624.                "http://linkedsystems.uk/system/offering/1226242/",
625.                "http://linkedsystems.uk/system/offering/1226254/",
626.                "http://linkedsystems.uk/system/offering/1226266/",
627.                "http://linkedsystems.uk/system/offering/1226278/",
628.                "http://linkedsystems.uk/system/offering/1226291/",
629.                "http://linkedsystems.uk/system/offering/1226309/",
630.                "http://linkedsystems.uk/system/offering/1226310/",
631.                "http://linkedsystems.uk/system/offering/1226322/",
632.                "http://linkedsystems.uk/system/offering/1226334/",
```

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

ICOS | Carbon Portal

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")…

ICOS | Carbon Portal

PIDINST entries are combined from ICOS view of Data Object
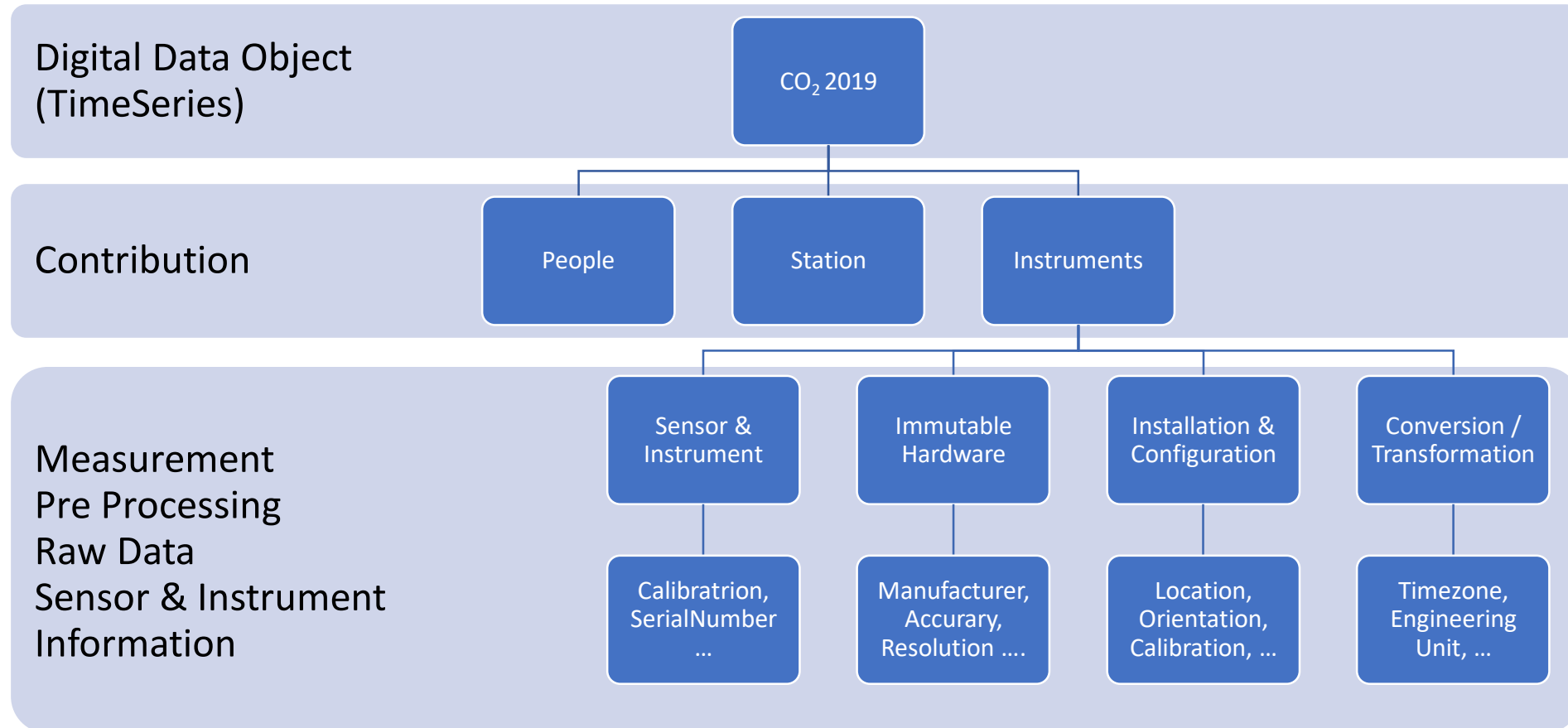
SENSOR.awi.de

Alfred Wegener Institute/PANGAEA

Helmholtz-Zentrum for Polar and Marine Research

14th RDA Plenary, Helsinki 2019

PANGAEA® Data Publisher

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/)
  - 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 –

- PANGAEA Data Publisher Use Case 2 –

PANGAEA® Data Publisher

  - 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)

Two related Use Cases:

- AWI sensor.awi Use Case 1 –

- PANGAEA Data Publisher User Case 2 –

  PANGAEA® Data Publisher

  - 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

https://doi.pangaea.de/10.1594/PANGAEA.887579

– Dropbox   Meine Ablage – Googl...   System Dashboard - P...   PangaWiki   Data Publisher for Eart...   datacite graphQL envi...

PANGAEA.

Data Publisher for Earth & Environmental Science

SEARCH   SUBMIT   ABOUT   CONTACT

Tina Dohna

Citation:

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

📢 **Always quote above citation when using data!** You can download the citation in several formats below.

RIS Citation   BIBTEX Citation   🖺 Copy Citation   Facebook   Twitter   Show Map   Google Earth
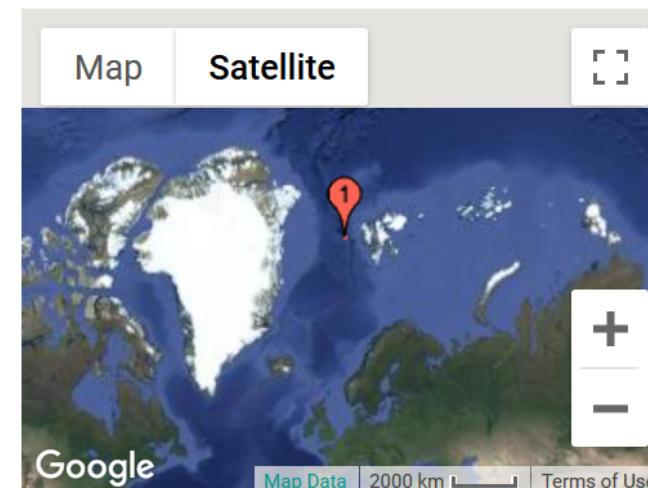
Abstract:

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.

ⓘ 🛡 🔒 https://doi.pangaea.de/10.1594/PANGAEA.887579 📄 ··· ♡ ☆ ↓ ⅢＶ 🔲

· Dropbox 📁 Meine Ablage – Googl... 🔴 System Dashboard - P... 🔴 PangaWiki 🔴 Data Publisher for Eart... 🔲 datacite graphQL envi...

Related to: **Wulff, Thorben; Bauerfeind, Eduard; von Appen, Wilken-Jon (2016):** Physical and ecological processes at a moving ice edge in the Fram Strait as observed with an AUV. *Deep Sea Research Part I: Oceanographic Research Papers*, **115**, 253-264, 🔴 https://doi.org/10.1016/j.dsr.2016.07.001 🔍

*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) 🔍 * *Basis:* Maria S. Merian 🔍 * *Device:* Autonomous underwater vehicle (AUV) 🔍

Search PANGAEA for other datasets related to 'Autonomous underwater vehicle'...

Parameter(s):

# Name	Short Name	Unit	Principal Investigator	Method	Comment
1 DATE/TIME 🔍	Date/Time		Wulff, Thorben 🔍		Geocode
2 LATITUDE 🔍	Latitude		Wulff, Thorben 🔍		Geocode
3 LONGITUDE 🔍	Longitude		Wulff, Thorben 🔍		Geocode
4 DEPTH, water 🔍	Depth water	m	Wulff, Thorben 🔍		Geocode – PAR
5 Radiation, photosynthetically active 🔍	PAR	µmol/m²/s	Wulff, Thorben 🔍		corrected

https://doi.pangaea.de/10.1594/PANGAEA.887579

Dropbox · Meine Ablage – Googl... · System Dashboard - P... · PangaWiki · Data Publisher for Eart... · datacite graphQL envi...

Related to: **Wulff, Thorben; Bauerfeind, Eduard; von Appen, Wilken-Jon (2016):** Physical and ecological processes at a moving ice edge in the Fram Strait as observed with an AUV. *Deep Sea Research Part I: Oceanographic Research Papers*, **115**, 253-264, 🔗 https://doi.org/10.1016/j.dsr.2016.07.001 🔍

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) 🔍 * *Basis:* Maria S. Merian 🔍 * *Device:* Autonomous underwater vehicle (AUV) 🔍

Search PANGAEA for other datasets related to 'Autonomous underwater vehicle'...

Parameter(s):

| #	Name	Short Name	Unit	Principal Investigator	Method	Comment
1	DATE/TIME 🔍	Date/Time		Wulff, Thorben 🔍		Geocode
2	LATITUDE 🔍	Latitude		Wulff, Thorben 🔍		Geocode
3	LONGITUDE 🔍	Longitude		Wulff, Thorben 🔍		Geocode
4	DEPTH, water 🔍	Depth water	m	Wulff, Thorben 🔍		Geocode – PAR
5	Radiation, photosynthetically active 🔍	PAR	$\mu mol/m^2/s$	Wulff, Thorben 🔍		corrected

PANGAEA.
Data Publisher for Earth & Environ

Citation:
Wulff, Thorben; Bauerfeind, Eduard; von Appen,
Sascha (2018): Vertical profiles of physical and
dive in the vicinity of an ice tongue in the Fram
https://doi.org/10.1594/PANGAEA.887579

📢 Always quote above citation when using data! You can download th

RIS Citation · BibTeX Citation · Text Citation · 🔗 Facebook · 🔗 Twitter ·

Abstract: AWI's autonomous underwater vehicle "PAUL" covered two 10
meltwater front. The meltwater front was associated to a large
resolution profile of the following parameters: Temperature, C

The dataset contains the data of the vertical ascends only. Due
parameter has an individual depth stamp.

Related to: Wulff, Thorben; Bauerfeind, Eduard; von Appen, Wilken-Jo
253-264, https://doi.org/10.1016/j.dsr.2016.07.001 Q

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

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

Event(s): MSM29_440-5 Q * 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: https://hdl.handle.net/10013/sensor.664525cf-45b9-4969-bb88-91a1c5e97a5b * Location: North Greenland Sea Q * Campaign: MSM29 (HAUSGARTEN 2013) Q * Basis: Maria S. Merian Q *
Device: Autonomous underwater vehicle (AUV) Q

Parameter(s):

# Name	Short Name	Unit	Principal Investigator	Method	Comment
1 DATE/TIME Q	Date/Time		Wulff, Thorben Q		Geocode
2 LATITUDE Q	Latitude		Wulff, Thorben Q		Geocode
3 LONGITUDE Q	Longitude		Wulff, Thorben Q		Geocode
4 DEPTH, water Q	Depth water	m	Wulff, Thorben Q		Geocode – PAR

https://hdl.handle.net/10013/sensor.664525cf-45b9-4969-bb88-91a1c5e97a5b

SENSOR.awi.de

AWI AUV Polar Autonomous Underwater Laboratory

Overview · Contacts · Actions · Parameters · Resources · Properties · Local Frame · Subdevices · Images

2013-07-02 19:58:00 Deployment MSM29_440-5

SENSOR (2013). Platform metadata for Vehicle AWI AUV Polar Autonomous Underwater Laboratory. Configuration from 2013-07-02 21:58:00. Alfred
Wegener Institute for Polar and Marine Research. https://hdl.handle.net/10013/sensor.664525cf-45b9-4969-bb88-91a1c5e97a5b

| State:	public
ID:	458
Parent:	
Device URN:	vehicle:awi_paul
Short Name:	AWI-PAUL
Long Name:	AWI AUV Polar Autonomous Underwater Laboratory
Collections:	
Description:	The Bluefin-21 is a highly modular autonomous underwater vehicle able to carry multiple sensors and payloads at once. It boasts a high energy capacity that enables extended operations even at the greatest depths. The Bluefin-21 has immense capability but is also flexible enough to operate from various ships of opportunity worldwide.
Serial:	Paul
Manufacturer:	Bluefin Robotics
Model:	Bluefin-21
Type:	Vehicle
Asset Number:	44055

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 valuesf for 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
  - Adoption of Research Organization Registry https://www.ror.community/
  - Adoption of Standard values for unknown information from DataCite, DataCite 4.3 Metadata Schema Documentation)

- 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
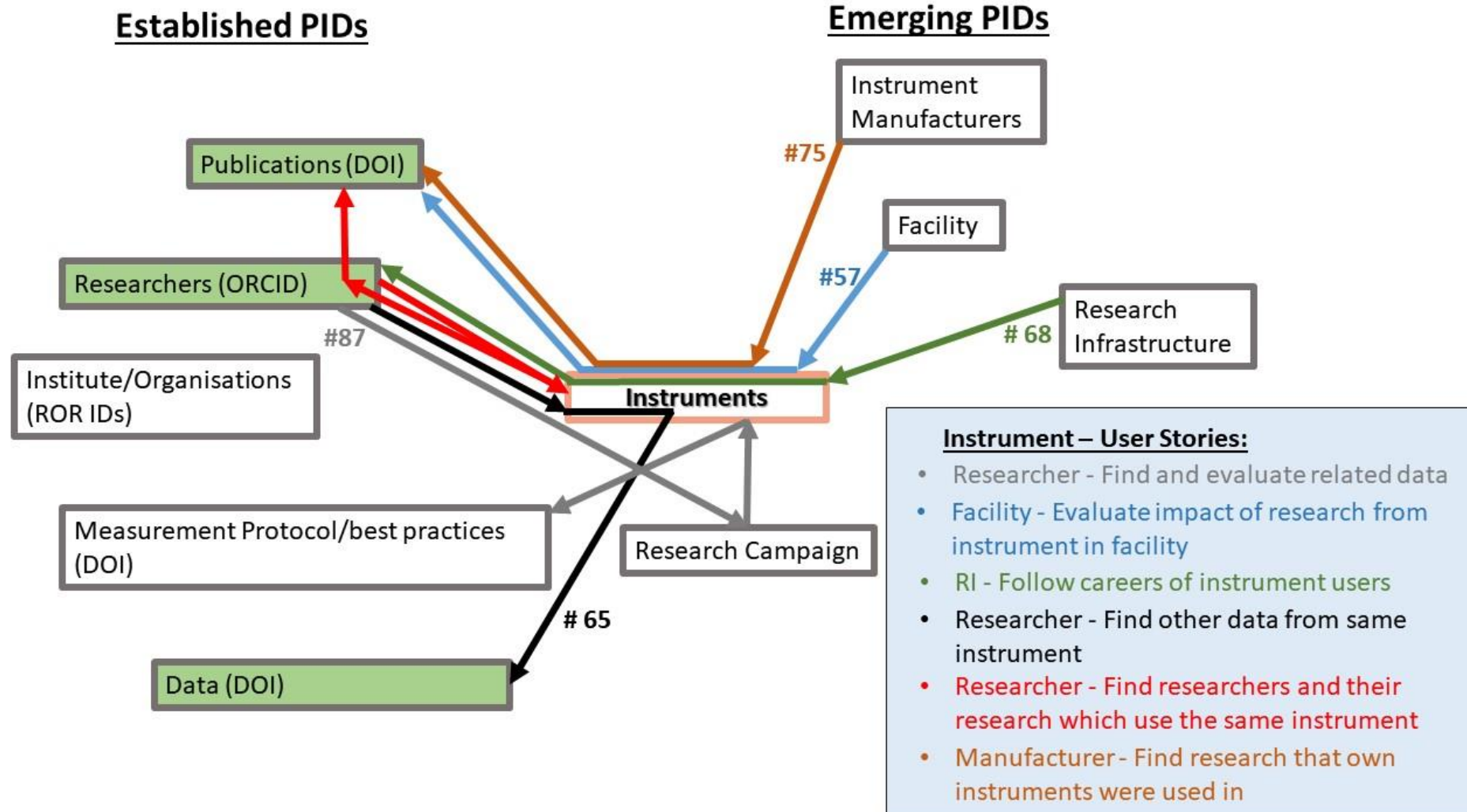
Issues:

- 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?