Persistent Identification of Instruments WG (PIDINST WG)

The PIDINST team

tinyurl.com/yddr3ocv

PIDINST

*Seeks to propose a community-driven solution for globally unique and unambiguous **identification** of instrument **instances** that are operational **in the sciences***

PIDINST

- Measuring instruments
  - A device or combination of devices designed for measurement of quantities
  - A device which quantifies the magnitude of a property, typically by assigning a numerical value or by producing a numerical representation that is subsequently used for evaluating quantities, such as a digital image.
- Information about instruments plays an important role in science
  - “To interpret a digital dataset, much must be known about the hardware used to generate the data, whether sensor networks or laboratory machines.” -- Christine L. Borgman in Big Data, Little Data, No Data, MIT Press, 2015 (p. 46)

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
- ...

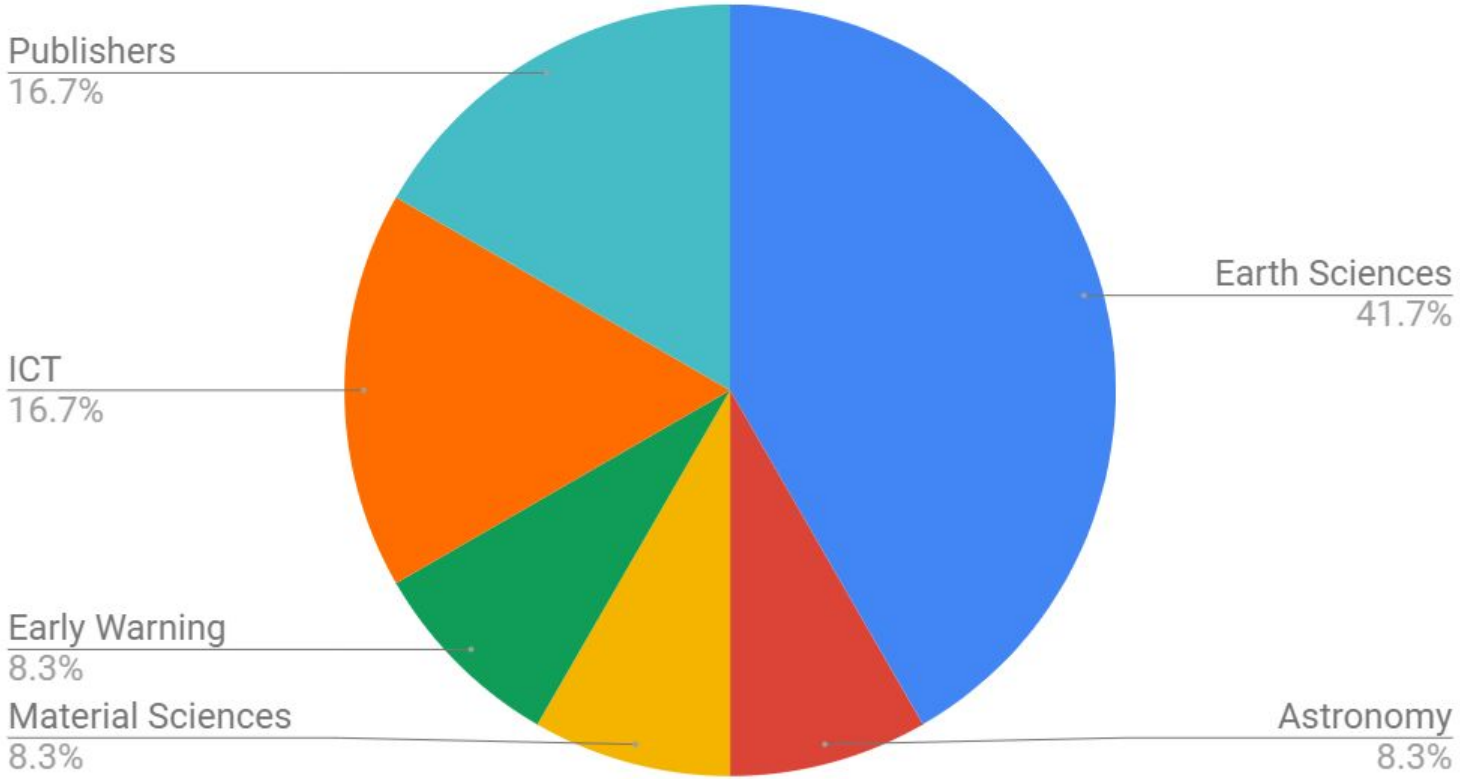
Updates since P11 (Berlin)

- Regular monthly conference calls
  - One call for the whole group
  - More recently, one call for technical subgroup
- Collected 13 use case descriptions
- Collected essential metadata for 10 use cases
- Aligned use case metadata
- Developed first draft schema

Use cases since P11

- SENSOR.awi.de by Ana Macario et al. (April)
- Marine SWE by Robert Huber et al. (May)
- ORCID by Tom Demeranville (May)
- ICOS Carbon Portal by Claudio D’Onofrio et al. (June)
- BODC by Louise Darroch et al. (July)
- ESO by Dominic Bordelon et al. (August)
- FZJ Central Library (JLSRF) by Claudia Frick (September)
- PANGAEA by Anusuriya Devaraju et al. (September)
- EuroGOOS/PSMSL/GLOSS by Louise Darroch (October)

Use Case Distribution



Examples

Use Case: JLSRF

- Journal of large-scale research facilities
  - Publishes articles describing large-scale scientific equipment
  - A means to cite large-scale facilities in publications
- How could JLSRF use an instrument PID
  - State instrument PID in the article published by JLSRF
  - Link article DOI with instrument PID
- For what is such a PID useful
  - Identify publications related to an instrument
  - Instrument PID could provide machine readable information useful for metrics
- Risks
  - Multiple identifiers (article and instrument) could create confusion: Which one should I use?

JLSRF: Metadata

Property	Definition	Datatypes
Instrument identification	Information directly identifying the instrument	Text, URL, ID, ISO 8601, numeric
Ownership	Information about the ownership including history of ownership, i.e. one entry for each past or current owner	Text, ID, ISO 3166, ISO 8601
Contact	Contact information, might be people or generic contact addresses	Text, ID
Related instruments	Links to predecessors to successors, but also to parents and childs for subunits	Text, ID

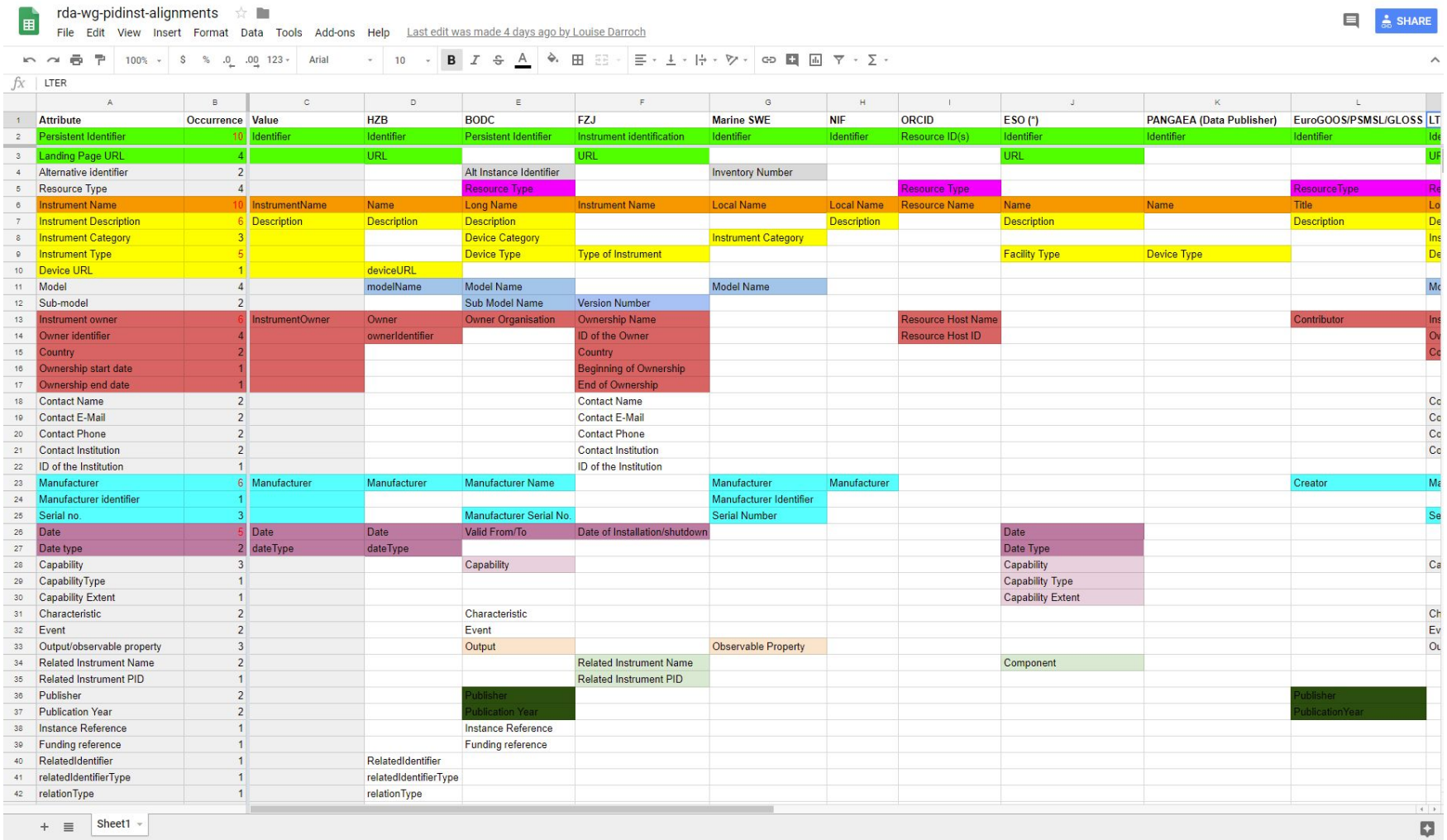
Use case: ESO

- European Southern Observatory
  - Intergovernmental astronomical research organization
  - Design, build, and operate ground-based telescopes, primarily sites in Chile
  - Each telescope may have one or more instruments
  - Each instrument is configurable and thus has an operation setup
  - Sites, telescopes, instruments, and instrument setups are collectively termed facilities
- How could ESO use an instrument PID
  - Tracking data usage in publications
  - Analysis related to the design, maintenance, decommissioning of instruments
- For what is such a PID useful
  - Instruments are an element of the context in which data are generated
  - Capturing this context facilitates reproducibility and verifiability of scientific analyses
  - Advance ESO’s commitment to Open Science
  - Provide practical benefits for the organization

ESO: Metadata

Property	Definition	Datatypes
Identifier	The Identifier is a unique string that identifies the facility	PID
Name	The name identifying the facility (short name, long name,technical name)	Text
URL	Landing Page URL	URL
Description	Description of the facility	Free text
Date	Dates relevant for the facility	ISO 8601
Date Type	Mandatory subproperty for each Date	Controlled list
Facility Type	General facility type, e.g., telescope instrument	Text
Capability	The set of capabilities, i.e., what a facility can do	
Capability Type	Mandatory subproperty of Capability (e.g., wavelength, resolution, etc.?)	
Capability Extent	The quantity/numerical expression/range/extent that accompanies the capability type	
Component	Physical components of the facility	

Alignment



rda-wg-pidinst-alignments

File Edit View Insert Format Data Tools Add-ons Help Last edit was made 4 days ago by Louise Darroch

SHARE

LTER

	A	B	C	D	E	F	G	H	I	J	K	L	
1	Attribute	Occurrence	Value	HZB	BODC	FZJ	Marine SWE	NIF	ORCID	ESO (*)	PANGAEA (Data Publisher)	EuroGOOS/PSMSL/GLOSS	LT
2	Persistent Identifier	10	Identifier	Identifier	Persistent Identifier	Instrument identification	Identifier	Identifier	Resource ID(s)	Identifier	Identifier	Identifier	Ide
3	Landing Page URL	4		URL		URL				URL			UR
4	Alternative identifier	2			Alt Instance Identifier		Inventory Number						
5	Resource Type	4			Resource Type				Resource Type			ResourceType	Re
6	Instrument Name	10	InstrumentName	Name	Long Name	Instrument Name	Local Name	Local Name	Resource Name	Name	Name	Title	Lo
7	Instrument Description	6	Description	Description	Description			Description		Description		Description	De
8	Instrument Category	3			Device Category		Instrument Category						Ins
9	Instrument Type	5			Device Type	Type of Instrument				Facility Type	Device Type		De
10	Device URL	1		deviceURL									
11	Model	4		modelName	Model Name		Model Name						Mo
12	Sub-model	2			Sub Model Name	Version Number							
13	Instrument owner	6	InstrumentOwner	Owner	Owner Organisation	Ownership Name			Resource Host Name			Contributor	Ins
14	Owner identifier	4		ownerIdentifier		ID of the Owner			Resource Host ID				Ow
15	Country	2				Country							Co
16	Ownership start date	1				Beginning of Ownership							
17	Ownership end date	1				End of Ownership							
18	Contact Name	2				Contact Name							Co
19	Contact E-Mail	2				Contact E-Mail							Co
20	Contact Phone	2				Contact Phone							Co
21	Contact Institution	2				Contact Institution							Co
22	ID of the Institution	1				ID of the Institution							
23	Manufacturer	6	Manufacturer	Manufacturer	Manufacturer Name		Manufacturer	Manufacturer				Creator	Ma
24	Manufacturer identifier	1					Manufacturer Identifier						
25	Serial no.	3			Manufacturer Serial No.		Serial Number						Se
26	Date	5	Date	Date	Valid From/To	Date of Installation/shutdown				Date			
27	Date type	2	dateType	dateType						Date Type			
28	Capability	3			Capability					Capability			Ca
29	CapabilityType	1								Capability Type			
30	Capability Extent	1								Capability Extent			
31	Characteristic	2			Characteristic								Ch
32	Event	2			Event								Ev
33	Output/observable property	3			Output		Observable Property						Ou
34	Related Instrument Name	2				Related Instrument Name				Component			
35	Related Instrument PID	1				Related Instrument PID							
36	Publisher	2			Publisher							Publisher	
37	Publication Year	2			Publication Year							PublicationYear	
38	Instance Reference	1			Instance Reference								
39	Funding reference	1			Funding reference								
40	RelatedIdentifier	1		RelatedIdentifier									
41	relatedIdentifierType	1		relatedIdentifierType									
42	relationType	1		relationType									

Sheet1

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.	

A First Draft for a Metadata Schema

Rolf Krahl

Persistent Identification of Instruments WG @ RDA P12,
Gaborone, 5 November 2018

Notes

A Metadata Schema for the Persistent Identification of Instruments.

- Heavily inspired by DataCite Metadata Schema.
- This is the first draft, many things not yet final. Current controlled lists of values may be considered as examples.
- Still need a catchy name for the system, use PIDINST as a place holder here.

Mandatory Properties

ID	Property	Occ
1	Identifier (with mandatory identifierType subproperty)	1
2	LandingPage	1
3	Name	1
4	Owner (with mandatory ownerName subproperty)	1-n
5	Manufacturer (with mandatory manufacturerName name subproperty)	1-n

Identifier: the persistent identifier. identifierType: controlled list of values: PIDINST.

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.

Recommended Properties

ID	Property	Occ
6	Description	0-1
7	InstrumentType	0-1
8	VariableMeasured	0-n
9	Date (with mandatory dateType subproperty)	0-1
10	RelatedIdentifier (with mandatory relatedIdentifierType and relationType subproperties)	0-n

Description: technical description.

InstrumentType: classification of the type of the instrument.

VariableMeasured: variable(s) that this instrument measures.

Date: dates relevant to the instrument. dateType controlled list of values: Commissioned, DeCommissioned.

RelatedIdentifier: identifiers of related resources.

Related Identifier

- RelatedIdentifier value: a PID.
- relatedIdentifierType: controlled list of values: PIDINST, DOI, Handle, URL, URN, ...
- relationType: description of the relationship, controlled list of values: IsDecribedBy, IsNewVersionOf, IsPreviousVersionOf, HasComponent, IsComponentOf, References, HasMetadata, ...
- Possible applications:
  - link articles describing the instrument (IsDecribedBy).
  - versioning: after major modification, mint a new PIDINST and link using IsNewVersionOf / IsPreviousVersionOf.
  - large instruments are made off components that are instruments on their own. Mint a seperate PIDINST for the component and link using HasComponent / IsComponentOf.
  - link other ressources 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).

Optional Property

ID	Property	Occ
11	AlternateIdentifier (with mandatory alternateIdentifier-Type subproperty)	0-n

AlternateIdentifier: other identifier(s) pertaining to the same instrument instance. Possible applications:

- serial number attributed by the manufacturer.
- inventory number for the owning institute.
- entry in some independent instrument registry.

Get Involved

- Follow the development on GitHub: `https://github.com/rdawg-pidinst/schema`.

Discussion

- First draft of the schema: Any comments, suggestions, critique, etc.?
- Any proposed new use cases / communities
- Schema adoption, PIDINST in practice: Who likes to implement the schema?
- Metadata by instrument database providers: It’s complicated

Conclusion

- First period
  - Completed
- Second period
  - Consolidate metadata schema and validate with new use cases
  - Draw feedback from the community
  - Develop additional deliverables
- Third period
  - Outreach to PID infrastructure to catalyze adoption