HZB Helmholtz Zentrum Berlin

Persistent Identification of Instruments

Discussion of a Use Case

Rolf Krahl

Scientific Data Management for Photon and Neutron Facilities,
Berlin, 20 March 2018

RDA WG: Persistent Identification of Instruments

The Persistent Identification of Instruments Group seeks to explore a community-driven solution for globally unique identification of active instruments in the sciences.

- Explore globally unique solution to persistently identify devices.
- Recommend a metadata profile.
- Explore technology to register and resolve the PIDs.
- Operationalize the solution, engaging instrument developers and manufacturers.

Create something like a DOI or an ORCID, but for instrument instances rather then publications or people.

Instruments at Photon and Neutron Facilities

Particularities of instruments at PaN facilities:

- Multiple complex instruments involved in a single measurement: source, insertation device, beamline, experimental station.
  $\Rightarrow$ *May need to reference a combination of instruments at once.*
- Unique instruments. Mostly designed and sometimes even manufactured in-house.
  $\Rightarrow$ *There may be no external manufacturer, no standard type.*
- Built off several components: simple (mirror, slit), complex custom built (monchromator), off-the-shelf products (detectors).
  $\Rightarrow$ *May need to also identify individual components.*
- Setup may change over time.
  $\Rightarrow$ *Need some kind of versioning.*

Instruments at Photon and Neutron Facilities

Particularities of instruments at PaN facilities:

- Multiple complex instruments involved in a single measurement: source, insertation device, beamline, experimental station.
  $\Rightarrow$ *May need to reference a combination of instruments at once.*
- Unique instruments. Mostly designed and sometimes even manufactured in-house.
  $\Rightarrow$ *There may be no external manufacturer, no standard type.*
- Built off several components: simple (mirror, slit), complex custom built (monchromator), off-the-shelf products (detectors).
  $\Rightarrow$ *May need to also identify individual components.*
- Setup may change over time.
  $\Rightarrow$ *Need some kind of versioning.*

Instruments at Photon and Neutron Facilities

Particularities of instruments at PaN facilities:

- Multiple complex instruments involved in a single measurement: source, insertation device, beamline, experimental station.
  $\Rightarrow$ *May need to reference a combination of instruments at once.*
- Unique instruments. Mostly designed and sometimes even manufactured in-house.
  $\Rightarrow$ *There may be no external manufacturer, no standard type.*
- Built off several components: simple (mirror, slit), complex custom built (monchromator), off-the-shelf products (detectors).
  $\Rightarrow$ *May need to also identify individual components.*
- Setup may change over time.
  $\Rightarrow$ *Need some kind of versioning.*

Instruments at Photon and Neutron Facilities

Particularities of instruments at PaN facilities:

- Multiple complex instruments involved in a single measurement: source, insertation device, beamline, experimental station.
  $\Rightarrow$ *May need to reference a combination of instruments at once.*
- Unique instruments. Mostly designed and sometimes even manufactured in-house.
  $\Rightarrow$ *There may be no external manufacturer, no standard type.*
- Built off several components: simple (mirror, slit), complex custom built (monchromator), off-the-shelf products (detectors).
  $\Rightarrow$ *May need to also identify individual components.*
- Setup may change over time.
  $\Rightarrow$ *Need some kind of versioning.*

Use cases and benefits

- Document the provenance of datasets.
- Track the scientific output of a given instrument.
- For a given dataset, search for other datasets created at the same instrument. Search for calibration data.
- Each HZB instrument has a web page providing documentation on the instrument, its design, and capabilities. Link this page from the PID.
- Attribute PIDs also to major components of an instrument, such as the detector. This allows an independent description of the characteristics of these components.
- Provide relevant metadata that can be automatically retrieved for any objects referencing the PID. E.g. the metadata schema for datasets created by the instrument.

Use cases and benefits

- Document the provenance of datasets.
- Track the scientific output of a given instrument.
- For a given dataset, search for other datasets created at the same instrument. Search for calibration data.
- Each HZB instrument has a web page providing documentation on the instrument, its design, and capabilities. Link this page from the PID.
- Attribute PIDs also to major components of an instrument, such as the detector. This allows an independent description of the characteristics of these components.
- Provide relevant metadata that can be automatically retrieved for any objects referencing the PID. E.g. the metadata schema for datasets created by the instrument.

Use cases and benefits

- Document the provenance of datasets.
- Track the scientific output of a given instrument.
- For a given dataset, search for other datasets created at the same instrument. Search for calibration data.
- Each HZB instrument has a web page providing documentation on the instrument, its design, and capabilities. Link this page from the PID.
- Attribute PIDs also to major components of an instrument, such as the detector. This allows an independent description of the characteristics of these components.
- Provide relevant metadata that can be automatically retrieved for any objects referencing the PID. E.g. the metadata schema for datasets created by the instrument.

Use cases and benefits

- Document the provenance of datasets.
- Track the scientific output of a given instrument.
- For a given dataset, search for other datasets created at the same instrument. Search for calibration data.
- Each HZB instrument has a web page providing documentation on the instrument, its design, and capabilities. Link this page from the PID.
- Attribute PIDs also to major components of an instrument, such as the detector. This allows an independent description of the characteristics of these components.
- Provide relevant metadata that can be automatically retrieved for any objects referencing the PID. E.g. the metadata schema for datasets created by the instrument.

Use cases and benefits

- Document the provenance of datasets.
- Track the scientific output of a given instrument.
- For a given dataset, search for other datasets created at the same instrument. Search for calibration data.
- Each HZB instrument has a web page providing documentation on the instrument, its design, and capabilities. Link this page from the PID.
- Attribute PIDs also to major components of an instrument, such as the detector. This allows an independent description of the characteristics of these components.
- Provide relevant metadata that can be automatically retrieved for any objects referencing the PID. E.g. the metadata schema for datasets created by the instrument.

Use cases and benefits

- Document the provenance of datasets.
- Track the scientific output of a given instrument.
- For a given dataset, search for other datasets created at the same instrument. Search for calibration data.
- Each HZB instrument has a web page providing documentation on the instrument, its design, and capabilities. Link this page from the PID.
- Attribute PIDs also to major components of an instrument, such as the detector. This allows an independent description of the characteristics of these components.
- Provide relevant metadata that can be automatically retrieved for any objects referencing the PID. E.g. the metadata schema for datasets created by the instrument.

Properties

- Obvious attributes: name, description, manufacturer, type, owner, landing page, ...
- Reference technical specification.
- Life time: start and end date of the instrument being in operation.
- Documentation: have a "is described by" relation with other resource.
- Versioning: have a "is new version of" and "is previous version of" relation with other instrument.
- Components: have a "has component" and "is component of" relation with other instrument.
- Extensible: Link other related resources.

Properties

- Obvious attributes: name, description, manufacturer, type, owner, landing page, ...
- Reference technical specification.
- Life time: start and end date of the instrument being in operation.
- Documentation: have a "is described by" relation with other resource.
- Versioning: have a "is new version of" and "is previous version of" relation with other instrument.
- Components: have a "has component" and "is component of" relation with other instrument.
- Extensible: Link other related resources.

Properties

- Obvious attributes: name, description, manufacturer, type, owner, landing page, ...
- Reference technical specification.
- Life time: start and end date of the instrument being in operation.
- Documentation: have a "is described by" relation with other resource.
- Versioning: have a "is new version of" and "is previous version of" relation with other instrument.
- Components: have a "has component" and "is component of" relation with other instrument.
- Extensible: Link other related resources.

Properties

- Obvious attributes: name, description, manufacturer, type, owner, landing page, . . .
- Reference technical specification.
- Life time: start and end date of the instrument being in operation.
- Documentation: have a “is described by” relation with other resource.
- Versioning: have a “is new version of” and “is previous version of” relation with other instrument.
- Components: have a “has component” and “is component of” relation with other instrument.
- Extensible: Link other related resources.

Properties

- Obvious attributes: name, description, manufacturer, type, owner, landing page, ...
- Reference technical specification.
- Life time: start and end date of the instrument being in operation.
- Documentation: have a “is described by” relation with other resource.
- Versioning: have a “is new version of” and “is previous version of” relation with other instrument.
- Components: have a “has component” and “is component of” relation with other instrument.
- Extensible: Link other related resources.

Properties

- Obvious attributes: name, description, manufacturer, type, owner, landing page, . . .
- Reference technical specification.
- Life time: start and end date of the instrument being in operation.
- Documentation: have a “is described by” relation with other resource.
- Versioning: have a “is new version of” and “is previous version of” relation with other instrument.
- Components: have a “has component” and “is component of” relation with other instrument.
- Extensible: Link other related resources.

Properties

- Obvious attributes: name, description, manufacturer, type, owner, landing page, . . .
- Reference technical specification.
- Life time: start and end date of the instrument being in operation.
- Documentation: have a “is described by” relation with other resource.
- Versioning: have a “is new version of” and “is previous version of” relation with other instrument.
- Components: have a “has component” and “is component of” relation with other instrument.
- Extensible: Link other related resources.

Journal of Large-Scale Research Facilities (JLSRF)

- Earlier approach to address some of the use cases: JLSRF.
- HZB's instruments have an article in JLSRF describing the instrument.
- Users are asked to cite this article in papers using data created at the instrument.
- The DOI of the JLSRF article is partly used as an substitute for the (not yet existing) instrument PID.
- Nevertheless, both approaches are not redundant: the textual instrument description in JLSRF gives more value to a human reader, while the instrument PID provides much richer options to automatically aggregate information by following the references.

Journal of Large-Scale Research Facilities (JLSRF)

- Earlier approach to address some of the use cases: JLSRF.
- HZB's instruments have an article in JLSRF describing the instrument.
- Users are asked to cite this article in papers using data created at the instrument.
- The DOI of the JLSRF article is partly used as an substitute for the (not yet existing) instrument PID.
- Nevertheless, both approaches are not redundant: the textual instrument description in JLSRF gives more value to a human reader, while the instrument PID provides much richer options to automatically aggregate information by following the references.

Journal of Large-Scale Research Facilities (JLSRF)

- Earlier approach to address some of the use cases: JLSRF.
- HZB's instruments have an article in JLSRF describing the instrument.
- Users are asked to cite this article in papers using data created at the instrument.
- The DOI of the JLSRF article is partly used as an substitute for the (not yet existing) instrument PID.
- Nevertheless, both approaches are not redundant: the textual instrument description in JLSRF gives more value to a human reader, while the instrument PID provides much richer options to automatically aggregate information by following the references.

Journal of Large-Scale Research Facilities (JLSRF)

- Earlier approach to address some of the use cases: JLSRF.
- HZB's instruments have an article in JLSRF describing the instrument.
- Users are asked to cite this article in papers using data created at the instrument.
- The DOI of the JLSRF article is partly used as an substitute for the (not yet existing) instrument PID.
- Nevertheless, both approaches are not redundant: the textual instrument description in JLSRF gives more value to a human reader, while the instrument PID provides much richer options to automatically aggregate information by following the references.

Journal of Large-Scale Research Facilities (JLSRF)

- Earlier approach to address some of the use cases: JLSRF.
- HZB's instruments have an article in JLSRF describing the instrument.
- Users are asked to cite this article in papers using data created at the instrument.
- The DOI of the JLSRF article is partly used as an substitute for the (not yet existing) instrument PID.
- Nevertheless, both approaches are not redundant: the textual instrument description in JLSRF gives more value to a human reader, while the instrument PID provides much richer options to automatically aggregate information by following the references.

Get involved!

- The RDA WG on Persistent Identification of Instruments starts its 18 month lifetime at this RDA plenary.
- Other endeavors to create such an instrument PID system seem to be underway.
- The design of this system is now being laid out.
- Now is the time to get involved, define our requirements, share our ideas and concerns, to make sure that the result will be suitable to our needs!

Discussion