ExPaNDS

European Open Science Cloud Photon and Neutron Data Services

PIDs use in facility proposals and in data publishing

Vasily Bunakov (UKRI STFC)
22 October 2021

A variety of PIDs to consider in facilities context

- Facilities data
  www.datacite.org – yet bear in mind that DataCite PID (DOI) may hide different meanings across facilities

- Facility instruments
  www.jlsrf.org , also RDA recommendation

- Organizations
  https://grid.ac/
  https://ror.org/
  https://www.crossref.org/services/funder-registry/

- Sky is the limit (make your own use of PIDs in facilities)

ExPaNDS
European Open Science Cloud Photon
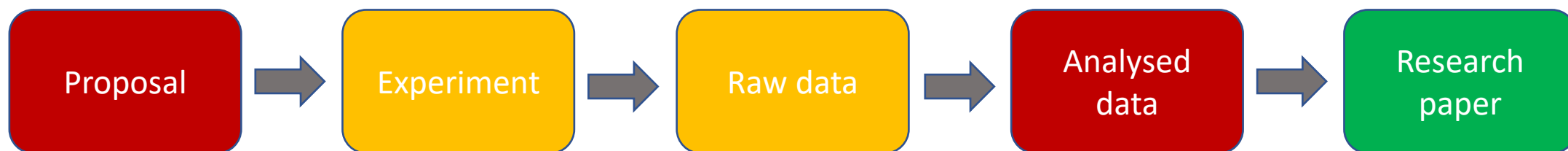and Neutron Data Services

An example of a conceptual challenge: what is a facility?

- It is an organisation, hence we should assign Organisation PID to it

- It is a large-scale instrument, hence we should assign Instrument PID to it

- It is a funder in-kind (and sometimes monetary funder), hence we should assign Funder PID to it

- Or we can assign all these PIDs to a facility and use them depending on research information context

- It will not harm maintaining a common vocabulary to link all PIDs for the same facility

ExPaNDS
European Open Science Cloud Photon and Neutron Data Services

(Simplified) facilities research lifecycle

Proposal → Experiment → Raw data → Analysed data → Research paper

PIDs to consider for:

Proposal	Experiment	Raw data	Analysed data	Research paper
Proposal itself	Experiment itself	Dataset itself	Dataset itself	Paper itself
PI	PI	Creator	Creator	Author
PI affiliation	PI affiliation	Creator affiliation	Creator affiliation	Author affiliation
PI funder	Instrument	Instrument	Instrument	References
Instrument	Instrument scientist	Data publisher	Data publisher	
References			Data format	
			Entities	
			References	

ExPaNDS
European Open Science Cloud Photon and Neutron Data Services
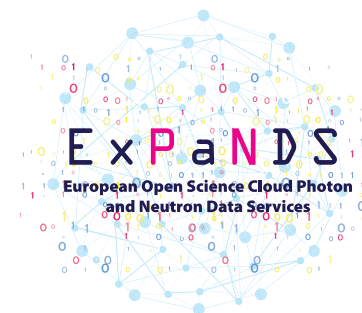
What we could connect in a PID graph for facilities (if we had enough PIDs)
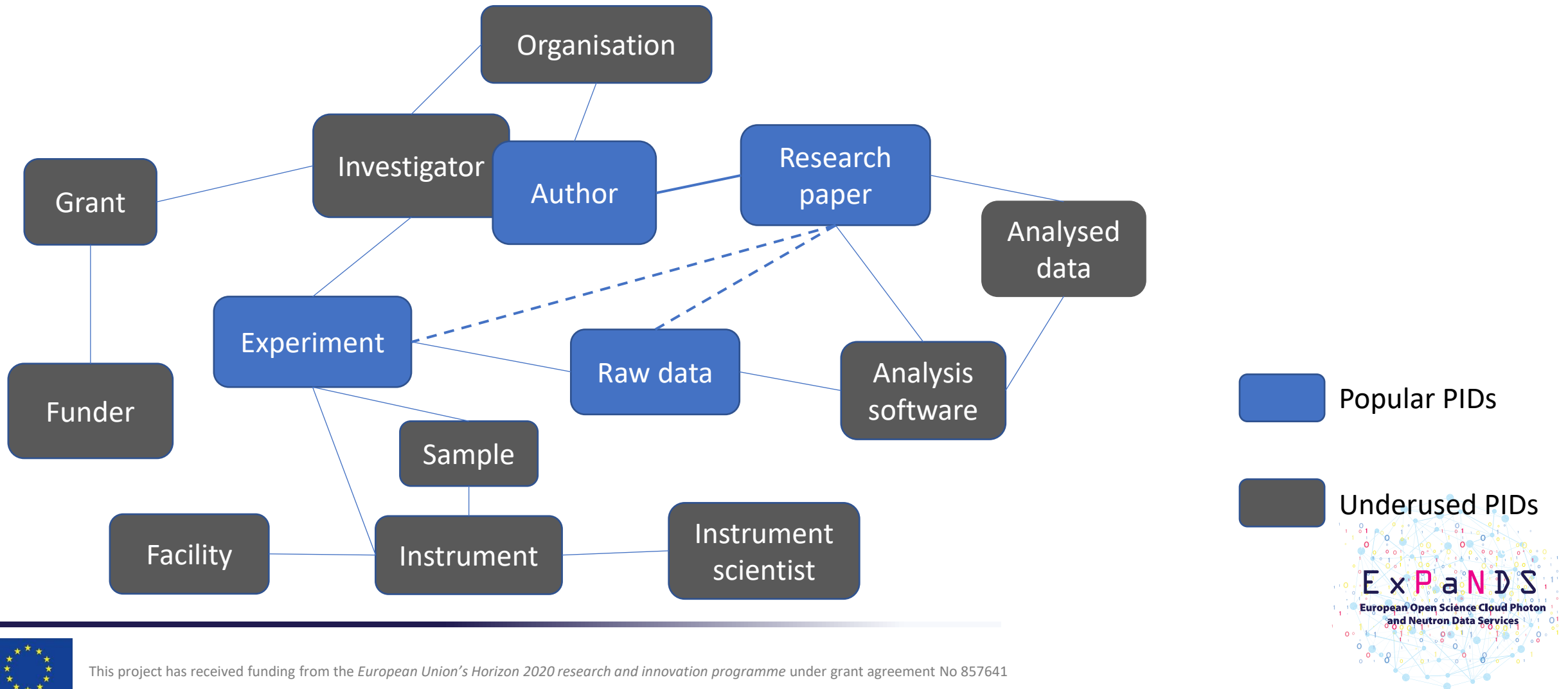
PIDs (or, rather, absence of them) in facility experiment proposals
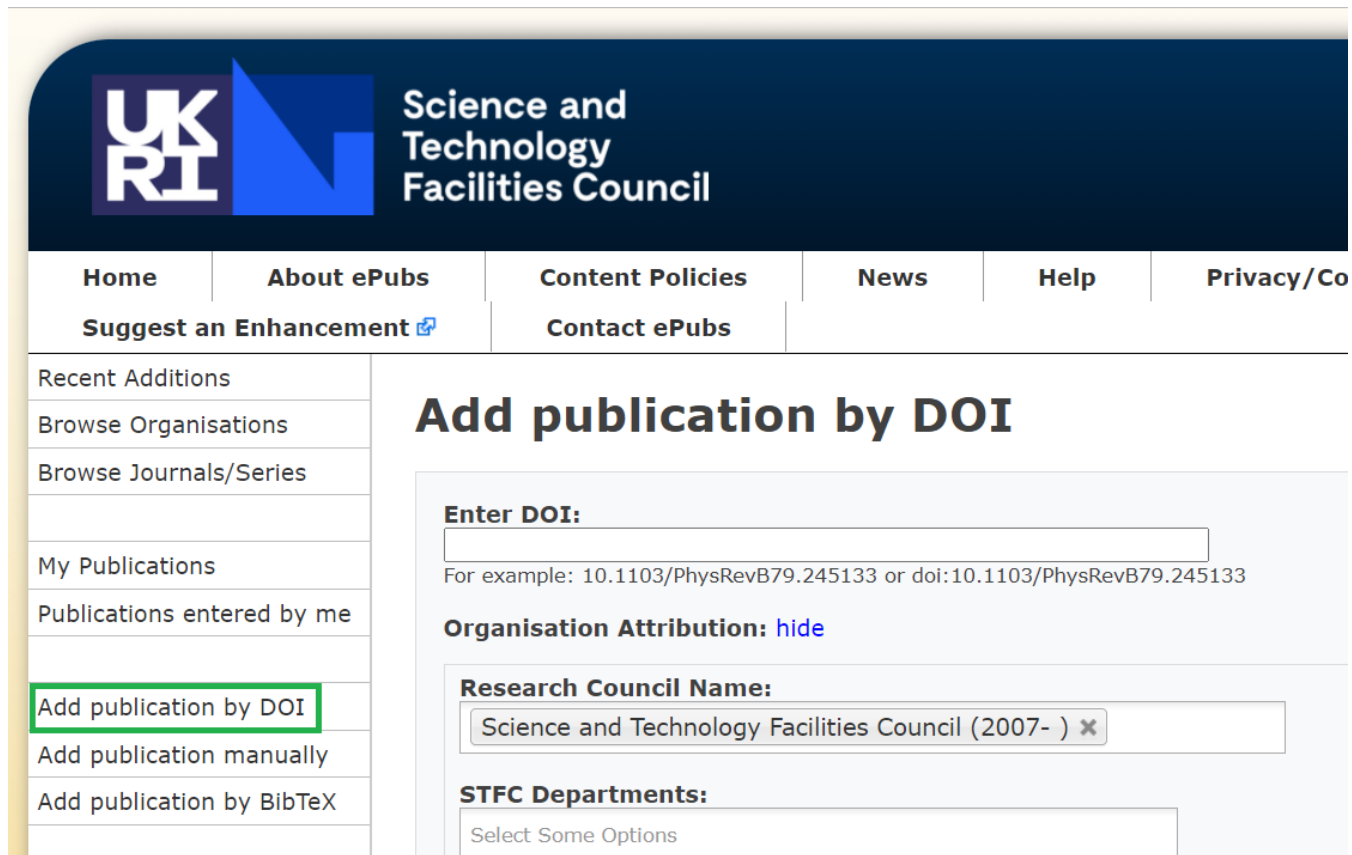

PIDs to consider for:

Proposal itself	Proposal ID is generated but it is not a PID
PI	PI ORCID is not collected
PI affiliation	PI Organisation PID is not collected
PI funder	Funder information is not requested (let alone Funder PID)
Instrument	Instrument name is captured but not a PID
References	Bibliographic references are requested in arbitrary format, without a DOI

User Offices and their proposal systems are first in the facilities research information value chain, and they fail the rest of the chain by not collecting PIDs

ExPaNDS
European Open Science Cloud Photon and Neutron Data Services

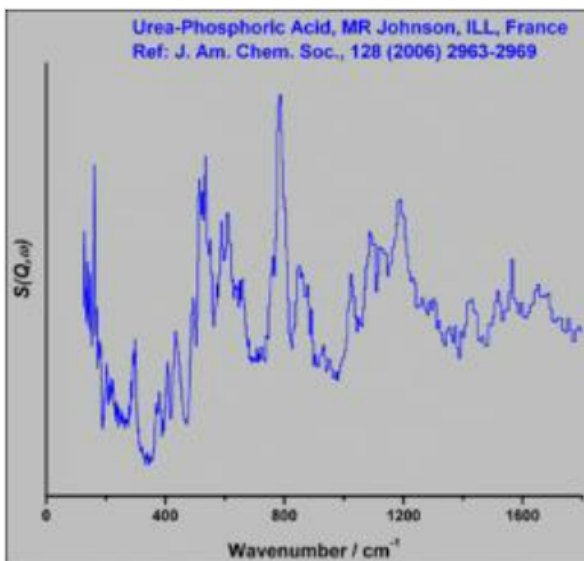# Illustration of why not requesting PIDs in proposal systems is a missed opportunity

The form is auto-populated with quality metadata

Issues with publishing analysed data

Urea/Phosphoric Acid

M.R. Johnson   Institut Laue-Langevin, Grenoble, France

Inelastic Neutron Scattering spectrum of Urea/Phosphoric Acid, NH_2CONH_2/H_3PO_4 measured on the TOSCA instrument

URI
http://purl.org/net/edata/handle/edata/707
http://dx.doi.org/10.5286/edata/687

Collections
INS Database

Related records
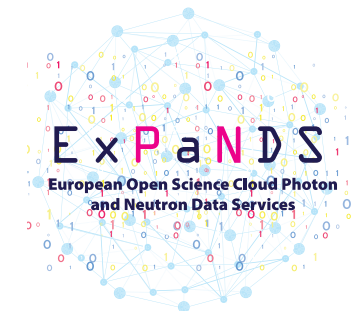J. Am. Chem. Soc., 128 (2006) 2963-2969 **!**

Licence
Creative Commons Attribution 4.0 International

View/Open
- Ureaphosphoricacid.gif (157.7Kb)
- Ureaphosphoricacid.dat (32.03Kb)
- Ureaphosphoricacid.dx (33.43Kb)

ExPaNDS
European Open Science Cloud Photon
and Neutron Data Services

Summary of issues with PIDs

- Even mature PID types, such as DOIs for publications, are underused by facilities User Offices, or when we formally publish analysed data in repositories

- Practices of User Offices are detached from research information management practices further down the facilities research lifecycle: librarians are typically better with PIDs than User Offices but they do not talk enough with each other

- There are win-wins, such as auto-population of proposals with quality metadata fetched via PIDs, but these easy-to-implement features are missing in proposal systems

- Publishers do not care enough about analysed data, so facilities (data repository owners in host organisations) could step in and properly care about their own analysed data originating in facility experiments – but they do not seem interested
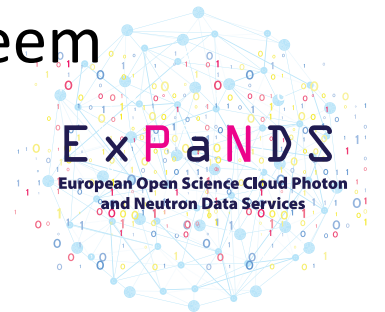
Light at the end of the tunnel…

- Every issue is an opportunity ☺

- Facilities can become hubs for quality research information (if they want to, and if they invest in PIDs and associated metadata)

- User Offices and Libraries in host organisations are key players who can promote PIDs use by co-design and co-implementation of best practices

ExPaNDS
European Open Science Cloud Photon and Neutron Data Services