PID services, what why how?
A short dip in the PID pond

Wednesday 14 November 2018
Belgrade, Serbia

CESSDA Widening Task

Birger Jerlehag
Senior Adviser
Swedish National Data Service, SND

Widening Workshop
What is a Persistent IDentifier

Metadata or article contains URL or PID

DOI HDL

PID Resolver
PID translates to URL

http://

Digital object
What is a PID Service

- Metadata or article contains URL or PID
- PID Resolver
  - PID translates to URL
  - May contain other services

  - 404
    - Object moved to new location

  - Digital object
Most common PID services

<table>
<thead>
<tr>
<th>Service</th>
<th>Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Handle system</td>
<td>The DONA Foundation</td>
</tr>
<tr>
<td>DOI</td>
<td>International DOI Foundation</td>
</tr>
<tr>
<td></td>
<td>DataCite is a member of IDF</td>
</tr>
<tr>
<td>ARK</td>
<td>The California Digital Library</td>
</tr>
<tr>
<td>URN:NBN</td>
<td>National Libraries</td>
</tr>
<tr>
<td></td>
<td>Uniform Resource Name : National Bibliographic Number</td>
</tr>
</tbody>
</table>
Why are PIDs important

Can I trust you with the data I have created or compiled

Can I rely on that data from you is trustworthy

Can data I use and cite be located and accessed in the future

_The use of a trustworthy PID service is an important step to take towards FAIR data and becoming a Trusted Digital Repository_
DataCite is

a global membership-based not for profit organisation

member of IDF (International DOI Foundation)

a Registration Agency for DOI prefixes
DataCite

90+ members/Allocation Agencies serving 1700+ Data Centers/ Clients with PID prefixes

More than 15 million DOIs minted through DataCite

Financed by membership fees and projects funding (Freya ...)

Uses its own metadata schema with 19 properties
DataCite

Runs an array of services publicly available

Runs re3data (Registry of Research Data Repositories)

Works closely with other PID organisations such as CrossRef and ORCID

CESSDA Service Providers GESIS and SND are members
Costs, where to get a DOI prefix,

DataCite members can assign DOI prefixes to clients

There might be a DCmember from your country you can ask

Each DCmember can set their own cost model so costs vary

GESIS da|ra offers DOIs without cost to CESSDA SPs
GESIS da|ra

The use of the da|ra service is open to institutions interested in registering their stored research data with DOI names.

It is free of charge for clients coming from the academic field.

To become a da|ra client the respective institution has to sign a Service Level Agreement with da|ra.

CESSDA SPs GESIS, CSDA, FORS and AUSSDA are all clients at da|ra.
The PID Services Report (D5.1) from 2014 gave a good insight to what demands SPs in all 5 SSH infrastructures had on PID Services.

The PID Task Force was formatted in 2015/2016 and could after a series of interviews, workshops and webinars present a suggestion to a PID Policy which was accepted by the General Assembly in 2017.
CESSDA PID Policy, background and considerations

Distributed infrastructure

Local/National considerations

Different levels of readiness and knowledge

No mention of PIDs in statutes or other guidelines
CESSDA PID Policy

Is scheduled for revision 2019

Shall be fully implemented by all SPs in 2020

Principle 1 – Identifying
Each CESSDA SP shall use globally unique and persistent identifiers to identify their data holdings, which are of interest to CESSDA.

Principle 2 – Locating
All data holdings of each CESSDA SP shall be findable by their global PID via the Internet.

Principle 3 – Resolving
CESSDA SP shall use global PID services that ensure 24/7 resolvability of PID.

Principle 4 – Referencing and Citation
PID need to be used to ensure referencing and citation of the data holdings of each CESSDA SP.

Principle 5 – Visibility
PID must be included in the resource-discovery metadata provided by CESSDA Service Providers for the CESSDA Data Catalogue.

Principle 6 – Flexibility
This Policy shall be reviewed at least every two years and adjusted according to the latest strategic and technological developments within and outside CESSDA.

CESSDA stands for Consortium of European Social Science Data Archives and ERIC stands for European Research Infrastructure Consortium.
A living document where recommendations from the PID community (RDA-groups, Freya, DataCite ...) will be adopted
Implementing the CESSDA Persistent Identifier Policy

Extended Best Practice Guidelines on specific topics related to the FAIR principles was created.

An assessment of the results of the Research Data Alliance (RDA)-EU adoption project on “Persistent Identifier Types for the Social Sciences” was carried out.

A collocated RDA-event ‘The use of Persistent Identifiers in the CESSDA ERIC’ was arranged and a Poster was presented at the RDA Plenary.

GESIS, DANS, SND
CESSDA PID TO DO?

Longitudinal data
Granularity
Relations
Fragments (EQB)
Multinational datasets
PIDs for people and Orgs
Alignment with other ERICs
Other types of data, databases

Tip of the Iceberg — Image by © Ralph A. Clevenger/CORBIS
PID projects and future enhancements

Freya
PID Graphs

PIDappaloza
Research Data Alliance (RDA)

Scholix
Make Data Count

Cite data, an ICPSR Movie
OVER AND OUT.