# CLARIN PID Task Force Meeting

CLARIN Center days 11-5-2022

# Agenda

- Need and place for a CLARIN PID TF?
- Widening scope
- PID Related topics

# Need for a CLARIN PID TF?

- Last actions were
  - a survey on PID use by the centres
  - Investigating PID/DOI brokering by CLARIN ERIC
  - *Gathering input for recommendations for DOI use vs Handle*
- The PID mechanism itself and its use seem no longer a chalenge
  - Interesting problems such as use of PIDs for particular data-types eg. (dynamic) social media, SKOS concepts, is perhaps better discussed in the context of the data-type itself, not for the PID TF perse
  - But some data-management tppics associated with PIDs will be more interesting and needed

# Widening scope

- To be interesting &b relevant we could use a widening of scope
- Widening wrt. RI domain eg. collaboration with DARIAH, CESSDA, CH
  - Quite a few centers have a stake in multiple RI domains
- Widening of subject
  - Data citation
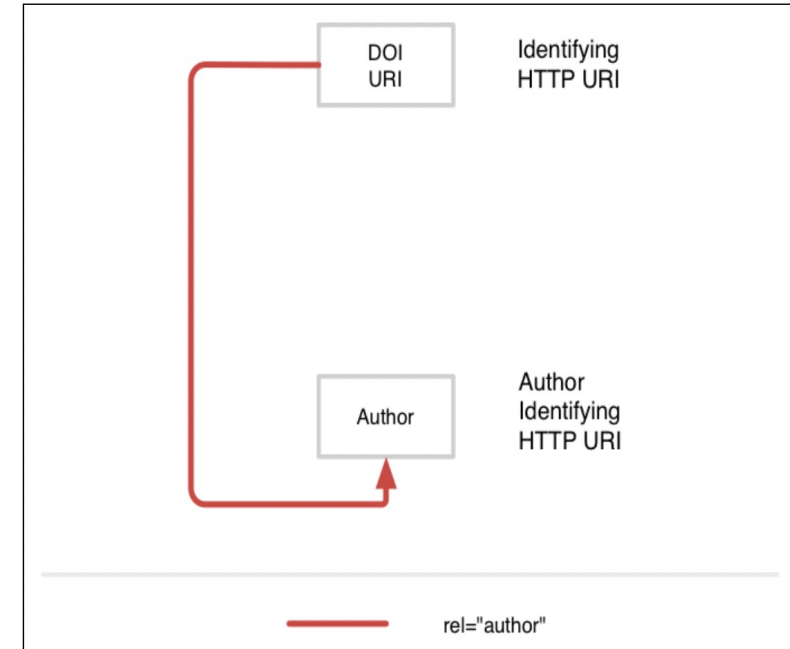  - Machine access to data (FAIR Data PIDs)

# Data Citation

- Recent SSHOC project task led by Nicolas Larrousse investigated the SSH landscape wrt. use of data-citation
- Survey/desktop study on repositories providing data-citation function
- Investigated FAIR data citation concept ie. machines should be able to consume and act on data citations

# Access to data, FAIR data PIDs

- PIDs point to landing pages, require often Human intervention to find and interpret a link or reference to the actual data (bitstream)
- What to do?
- There are a few technologies available who claim to address that:
  - Sign-posting (vd Sompel)
  - DOG – Digital Object Gateway (CLARIN)
  - Fair Digital Objects (FDO Forum)
  - Standardized Repository APIs?
    - [Research Data Repository Interoperabilty WG Recommendations](#)
    - [RDA Collections API, https://zenodo.org/record/2428145](https://zenodo.org/record/2428145)
    - …

# Sign-posting

- It uses [Typed Links](#) as a means to clarify patterns that occur repeatedly in scholarly portals. For resources of any media type, these typed links are provided in [HTTP Link headers](#). For HTML resources, they may additionally be provided in [HTML link elements](#).

- Very HTTP specific, uptake seems limited?

- https://signposting.org

- Van de Sompel, H., Nelson, M.L. Reminiscing About 15 Years of Interoperability Efforts. *D-Lib Magazine*, Volume 21, Number 11/12 https://doi.org/10.1045/november2015-vandesompel.



The HTTP Link Header to link to the authors from the **DOI URI** https://doi.org/10.1045/november2015-vandesompel is:

```
curl -I "https://doi.org/10.1045/november2015-vandesompel"
```

```
Link: <http://orcid.org/0000-0002-0715-6126>
      ; rel="author",
      <http://orcid.org/0000-0003-3749-8116>
      ; rel="author"
```

# DOG

- CLARIN initiative to create a software library that can translate PIDs to data location on the basis of special repository information

- Short paper available

- Partly implemented in Python library

# Fair Digital Objects -FDO

- Long term activities originally from researcher initiative GEDE, active group at RDA,
  - On the basis of set of attributes directly associated with a PID (Handle) machines are able to broker services and process data
  - research interests from: Earth science, Life science, Biomed, support from larger compute/data centres and Handle providers (GWDG), interest from DIN and industry
  - Several discussion groups with strong experts: https://fairdo.org
  - But still few real implementations

# TF member interest

| Action/task | interest | comment |
|---|---|---|
| Continue as before | | |
| Scaling-up RI domain | | Find partners from non-CLARIN RI |
| - | | |
| Expanding TF topic | | |
| - data-citation | | |
| - Access to data | | |
| | | |
| | | |
| | | |