



Use of PIDs within CLARIN: requirements and recommendations

Dieter Van Uytvanck

Max Planck Institute for Psycholinguistics / CLARIN ERIC

dieter@clarin.eu

PID workshop, Utrecht

2014-02-20

- Introduction round
- Context of this meeting:
 - many centre candidates that are about to make a choice
 - EPIC in an organisational transition
 - some new identifier systems: status?
 - need for clarification
- Goals of today's meeting:
 - morning: getting an update about recent developments, question and answer opportunity
 - afternoon: internal discussion & planning
 - later on: document with summary and answers to frequent questions

- **Forenoon - open workshop**
 - 09:00-09:30 Use of PIDs within CLARIN: requirements and recommendations (Dieter Van Uytvanck, CLARIN [ERIC](#))
 - 09:30-10:15 [EPIC](#): current state and future perspective (Tibor Kalman, GWDG)
 - 10:15-11:00 [URN:NBN services](#) (Maarten Hoogerwerf & Marnix van Berchum, [DANS](#))
 - 11:00-11:15 coffee break
 - 11:15-12:00 Running your own handle server: experiences from some CLARIN centres (Rainer Metsvahi, Jozef Misutka)
 - 12:00-12:30 other [PID](#) services: [DataCite](#), [Picturae](#) services, [ISLRN](#)
- 12:30-13:30 lunch break

- **Afternoon - closed meeting of the PID taskforce and interested centres**
 - 13:30-14:00 analysis of the PID questionnaire (Dieter Van Uytvanck, CLARIN ERIC)
 - 14:00-15:30 discussion slot
 - 15:30-16:00 wrap-up and looking forward

Requirements



- Already in the preparatory phase (**2009**), CLARIN put forward some clear recommendations:
 - <http://hdl.handle.net/1839/00-DOCS.CLARIN.EU-30>
- In **2012**, the centre committee adopted the official centre criteria document:
 - <http://hdl.handle.net/1839/00-DOCS.CLARIN.EU-77>
- In **2013**, after several meetings of the PID taskforce, the criteria for the usage of PIDs were made more detailed and explicit (in relation to metadata records):
 - <https://www.clarin.eu/node/3757>

Requirements 2009



- centres should get acquainted with PIDs
- and with repositories that support PIDs
- keep PIDs in mind during software development
- establish CLARIN-wide PID service
- one system which is performant, scalable and robust enough and that offers enough flexibility: **handle**
- talk to CNRI about requirements that are not met yet:
 - Global Handle Registry mirror
 - part identifiers
- make versioning compulsory for digital objects (with PIDs – comply or explain)
- establishing PID service independent of any commercial business model (DOI ok as individual choice but should not be default option)
- investigate various options of sharing a registration and resolution service with other disciplines

- ***Centres need to associate PIDs records according to the CLARIN agreements with their objects and add them to the metadata record.***
- This should be indicated by:
 - An indication of the **Handle** assignment policies and procedures and ways to check that they are systematically applied.
 - An indication where to find **Handle** information in the metadata records and whether clicking on them will bring you to the data.

Requirements 2013



- Centres need to associate (handle) PIDs with their metadata records. These PIDs should be suitable for both human and machine interpretation, taking into account the HTTP- accept header.
- Non-metadata files should receive a PID or a PID in combination with a part identifier, if these files:
 - are accessible via internet
 - are considered to be stable by the data provider
 - are considered to be worth to be accessed directly (not via metadata records) by the data provider

Recommendations



- PID taskforce, november 2013:
 - get your own prefix (costs: 50\$ registration + 50\$/year, see http://www.handle.net/service_agreement.html)
 - it is not obligatory to use EPIC, you can also run your own handle server
 - if you use EPIC, make sure to use API version 2 (in combination with your own prefix, see www.pidconsortium.eu/index.php?page=api_v2)
- Has not been explicitly expressed in a document so far
 - add this to the summary document of this meeting

CLARIN

Common Language Resources and Technology Infrastructure



Thank you for your attention!
