|  |  |  |
| --- | --- | --- |
| **Title** | PID FAQ – update 2016 |  |
| **Version** | 1 |
| **Author(s)** | Dieter Van Uytvanck |
| **Date** | 2016-05-06 |
| **Status** | Draft for feedback |
| **Distribution** | SCCTC |
| **ID** | CE-2016-0781 |

[**What is a persistent identifier (PID)?**](https://www.clarin.eu/faq/what-persistent-identifier-pid)

Persistent identifiers are increasingly often seen as core component for all the many references we are creating at various levels - this can range from references between metadata descriptions and their resources up to references between semantic assertions made by using the RDF (Resource Description Framework). For more information please read the requirements [specification document](https://www.clarin.eu/specification-documents) or [the short guide](https://www.clarin.eu/files/pid-CLARIN-ShortGuide.pdf).

[**Why do I need PIDs?**](https://www.clarin.eu/faq/why-do-i-need-pids)

In the emerging cyberinfrastructure we are creating more and more references between resources, resource fragments and services. The creation of these references is very costly and often is essential for the interpretation of a resource. Therefore we need proper mechanisms to ensure that these references survive despite all the changes that happen in repositories for example. It is known that URLs are not appropriate - they are not persistent even when we believe that they are proper URIs. Therefore special PIDs come into place which identify an object and which are maintained by reliable institutions.

[**How does it function?**](https://www.clarin.eu/faq/how-does-it-function)

Handling PIDs is very simple. First you need to register a [PID](https://www.clarin.eu/glossary#PID) for a resource or service. You can do this very simply by providing the required information to the PID service site, in particular the path to access the resource such as a URL and you will receive back a PID which you can enter into the metadata description for example, so that everyone can use it for referencing. When a user finds such a PID in a resource, he/she can click on this reference and the service will resolve the PID and give access to (one of the copies of) the resource. Normally as user you don't see the intermediate transactions.

[**What if the PID service is down?**](https://www.clarin.eu/faq/what-if-pid-service-down)

If the PIDs cannot be resolved at a certain moment one simply cannot access a resource. Think of a situation where hundreds of users are waiting on a resolution of a [PID](https://www.clarin.eu/glossary#PID) and nothing happens - a nightmare for any cyberinfrastructure scenario! Since this would not be acceptable, we need to make sure that the PID service is based (a) on a very robust and reliable software offering sufficient functionality, (b) on a proper service based on redundant centres with a high availability and persistency guarantee.

[**Is CLARIN offering such service?**](https://www.clarin.eu/faq/clarin-offering-such-service)

CLARIN has an arrangement with the [EPIC consortium](http://www.pidconsortium.eu) that CLARIN members will be able to register PIDs and of course resolve them. This consortium groups a number of reliable European service providers that want to participate in providing a redundant service for the research world, i.e. we are speaking about millions of PIDs and a service at very low costs. The service is based on the Handle System which according to our investigations is the only robust system meeting [all requirements](http://hdl.handle.net/1839/00-DOCS.CLARIN.EU-30). No one is obliged to register Handles, but of course CLARIN centres will need to demonstrate that their PIDs can be resolved in a robust manner and offer the required functionality.

[**Why not PURLs or URNs?**](https://www.clarin.eu/faq/why-not-purls-or-urns)

PIDs are as said unique and persistent identifiers of objects that are made available by proper repositories. For many resources there are additional characteristics such as multiple copies for preservation reasons, a string (e.g. a checksum) that can be used to check authenticity, simple metadata for citation purposes, a reference to the access permission record etc. A proper [PID](https://www.clarin.eu/glossary#PID) system should offer such information immediately when resolving a PID. PURLs can't offer functionality, for URNs we do not know about well-proven and robust resolver, although the big libraries agreed on using URNs for their publications.

[**When should I use a part identifier for a PID?**](https://www.clarin.eu/faq/when-should-i-use-part-identifier-pid)

[**How do I issue a part identifier for an EPIC handle?**](https://www.clarin.eu/faq/how-do-i-issue-part-identifier-epic-handle)

The rewriting behaviour of part identifiers can be configured per handle prefix (actually it can also be done per individual handle but this is not supported for [EPIC](https://www.clarin.eu/glossary#EPIC) at this point). For EPIC (version 1, so with prefix 11858) the choice was made to rewrite [suffix] to ?[suffix]

 So suppose that 11858/1234 resolves to [http://clarin.eu](http://clarin.eu/) then

 11858/[1234@test](mailto:1234@test)=a will be resolved to [http://clarin.eu?test=a](http://clarin.eu/?test=a)

Please note that when you offer PIDs with part identifiers that you are responsible of maintaining the part identification fragment as well. Remember that users will use it to link to your resources and that the resulting end point should always be available.

[**How do I resolve a handle persistent identifier?**](https://www.clarin.eu/faq/how-do-i-resolve-handle-persistent-identifier)

A handle exists of 2 parts:

* a prefix (e.g. *1839)*
* a suffix (e.g. *00-0000-0000-0009-3C7E-F*)

The official way of refering to a handle is:  
  
hdl: + *prefix* + / + *suffix*  
  
e.g.:

hdl:1839/00-0000-0000-0009-3C7E-F

To resolve such a handle (=make it a clickable link that redirects to the resource itself) use the following formula:  
  
<http://hdl.handle.net/>*prefix*/*suffix*  
  
e.g.: <http://hdl.handle.net/1839/00-0000-0000-0009-3C7E-F>

**[How can I get an EPIC account as a CLARIN centre?](https://www.clarin.eu/faq/how-can-i-get-epic-account-clarin-center)**

As [described at the EPIC website](http://www.pidconsortium.eu/index.php?page=process) it is sufficient to send a mail to handle /at/ gwdg.de with a motivated request.

**I want to setup a CLARIN-compliant repository. Should I use the EPIC API or run a handle server myself?**

Although there is no simple answer to this, the flowchart below can offer some guidance. In general, there is some overhead to run a handle server that pays off if:

* a centre is planning to register a lot of handles,
* a centre will use a repository that come with an integrated handle server (LAT/FLAT, DSpace)****