



# Requirements for PID Systems

## nestor standards working group

Jens Ludwig

[ludwig@sub.uni-goettingen.de](mailto:ludwig@sub.uni-goettingen.de)

Garching, 28th of March 2008

## outline



- professional and personal perspective
- basic requirements
- the nestor standards working group
- new developments

# professional and personal perspective

nestor



- The german network of expertise in digital long-term preservation (<http://www.langzeitarchivierung.de>)
- Since 2003
- Funded by Federal Ministry for education and research
- Deals with the non-technical issues of LTP: knowledge, organisational issues, consciousness, training, ...
- 7 working groups, one deals with standards and PIDs

## professional perspective



- LTP basically fails if digital objects can not be accessed because their identifiers are no longer valid
- Classical (libraries, archives, ...) and new (data centers, ...) memory institutions need them for LTP
- They often need/have their own resolving infrastructure.
- Which PID system should be adopted?
- How should the local infrastructure be managed?

## personal perspective



- As a user I am confused by the plurality of PID systems and resolvers. The average user inside of me wants only one system ...
- As an academic user in a field with mainly electronic and internet resources I recently observed two things:

## observation 1



- RLG has merged with OCLC: No URL is any longer valid?
- Seems like a use case for PIDs ...
- I did no research, but I wonder: Has RLG issued PIDs and have they used their domain name in them?
- That would be a case for N2T (<http://n2t.info>).
- Anyway: Does RLG/OCLC have a clear and trustworthy policy concerning their ‘normal’ URLs?

## observation 2



- I tried to locate versions of free academic resources in the Internet Archive.
- Their domains are harvested, but ...
- ...but not the resources linked via PIDs!
- Supposed reason: The links are regarded as outside of the domain. Active resolving is probably difficult for web archiving.
- **Diagnosis: insufficient interoperability!**



# basic requirements

## basic requirements



- Trustworthiness
  - Basic principle of LTP because we can not prove the success of LTP in the present
  - Important aspects: transparency and adequacy

## basic requirements



- Interoperability
  - LTP is interoperability with the future (David Giaretta and others)
  - Independence of infrastructure and individual organisations supports LTP
  - Organisations hesitate to adopt PIDs because of the plurality of systems and the missing interoperability.

# the nestor standards working group

the nestor standards wg



- Had a project in 2007 funded by the German Institute for Standardization (DIN) and the Federal Ministry for Economics and Technology (BMWI).
- Three work packages for PIDs
  1. First sketch of criteria catalogue for trustworthy PIDs and their infrastructure
  2. Requirements from classical memory institutions

the nestor standards wg



### 3. Installation and evaluation of N2T

(<http://134.76.163.149/examples.html>)

– Will continue now in 2008

## criteria catalogue



- The idea is similar to the RLG/OCLC TRAC and nestor criteria catalogue.
- Tries to address different kinds of risks: social, organisational, policy related and technical
- Important: Documentation of the fulfilment of criterias.
- Fulfilment can be accomplished to different degrees.

## criteria catalogue: examples



- Credible commitment to continuing maintenance from PID provider
- Transparency of collaboration between PID provider and content provider: They can not keep PIDs stable on their own.
- Authenticity: Does the resolving of the PID lead us to the correct object? How can we verify this?



## criteria catalogue: examples



- **Fallback procedures and inheritance:** What happens if PID provider goes out of business?
- **Error-handling:** What happens if PIDs break or can temporary be not resolved?
- **Scalability**
- **Maintenance**
- **Standardisation**

## criteria catalogue: examples



- **Extendability**
- **Technology independence**
- **Export functionality**
- **Ease of usage**
- ...

# new developments

new developments to be considered



- Australian PILIN project: Persistent Identifier Linking Infrastructure
- Context: ARROW project, national infrastructure
- Aim: *shared* identifier management infrastructure
  - Persistence of identifiers and identifier services
  - Assist movement of identifiers and resources
- Ended in December 2007

## some of PILIN's developments



- Abstract model/Ontology for identifiers and services:  
Potential for interoperability!  
And danger because it adds another layer of complexity?
- Use cases/scenarios (esp. digital library, e-Research, e-Learning)
- Advices and guidelines on using identifiers
- Service usage and management models and workflows
- Pilot software infrastructure

## future nestor activities



- **As mentioned: small DIN and BMWI project 2008**
  - **New draft of criteria catalogue**
  - **Extension of use cases of memory institutions**
  - **Probably workshop at the end of the year**
- **New DIN committee: LTP with sub committee for PIDs**

Thanks!