

Persistent Identifiers @ DANS

User- and Service Provider stories

Maarten Hoogerwerf

[Workshop Persistent Identifiers within CESSDA](#)

18 February 2016 @ GESIS, Cologne



This work is licensed under [Creative Commons Namensnennung 4.0 International Lizenz](#).

DANS is an institute of KNAW en NWO

Driven by data



Two presentations

- **Thursday: User story** - Persistent Identifiers for DANS' Archive (10m)
- **Friday: Service Provider story** - URN Service by DANS and Royal Library (15m)



User Story

PERSISTENT IDENTIFIERS FOR DANS' ARCHIVE



Outline

- Introduction of DANS & Archive
- Short History on persistent identifiers
- Challenges
- Recommendations



Data Archiving and Networked Services (DANS)

Established in 2005 by KNAW and NWO, to support social sciences and humanities with data archiving/management services.

- Successor of e.g. Steinmetz Archive (1964)
- Online self-deposit archive EASY
- Persistent Identifiers since 2006

- Other services: NARCIS portal, DataverseNL, NESSTAR, DCCD and Mendeley Data (LTP)



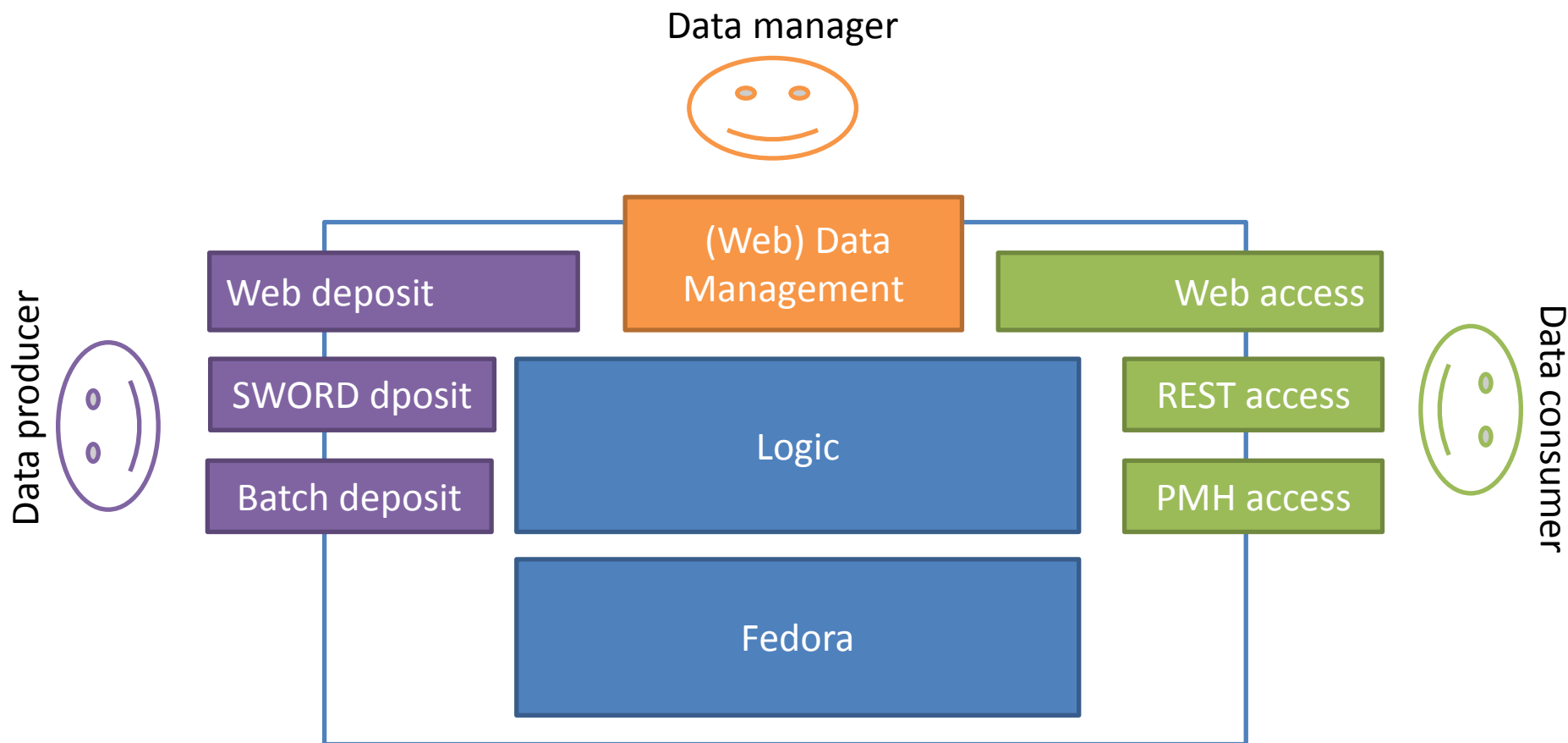


Self deposit Archive EASY

- **Register** - agree to terms & conditions
- **Start Deposit** – Specify community profile
- **Upload Data** – data and metadata files
- **Specify Metadata** – and access conditions
- **Submit** – Confirm and submit to data manager
 - *PID is minted*
- **Validate** - Perform checks (and corrections)
- **Accept** – Publish in archive
 - *PID is registered / resolvable*



EASY – Simplified Overview





A short history on PID @ DANS

Take one (2006) – Archival Perspective

- Identifiers for a Long Term Preservation archive

Take two (2013) – Researcher Perspective

- Identifiers to link data with e.g. publications

Take three (tbd) – Infrastructure Perspective

- Address future challenges



Take one – URN:NBN

Context of LTP archives and institutional repositories:

- User friendly – no uuid
- **Business model** - no per dataset annual fee
- **Technology independent** – no HTTP...
- **Collaborative** – repositories and archives
- **Governance** – to take responsibility

→ **URN:NBN** – relies on RFC, domain of LTP



Take two - Datacite DOI

Context of datasets as part of scholarly communication

- **User friendly** – recognizable
- **Data citation** – part of scholarly communication
- **Business model** – annual fee
- **Collaborative** – too big to fail
- *Technology independent*
- *Governance*

→ **Datacite DOI** – usage increases sustainability



Take three – What next?

- **Support URN and DOI!**
- **Machine access** – To a landing page?
- **Granular access** – Identification of individual files or variables (linked open data!)
- **Global collaboration** – Identifying the same dataset/study
- **Durable context** – Identifying related objects
- **Various** – Integration Mendeley, Google Scholar, unclear policy by Datacite (like/dislike)



Lessons Learned / Recommendations

- We do this for end-users
- Policy and communication are important
- Persistent identifiers are a loaded term
- We serve more than one community



You can't
play with us
if you are
different

We don't do
this for users,
they don't
understand what
we do

Thank you for your attention