



Data Archiving and Networked Services  
**DANS**

# DICE Digital Preservation Service (DDPS) @ IDCC22

Tuesday, June 14th 2022

Wilko Steinhoff (DANS-KNAW), Speaker

Herbert van de Sompel (KNAW-DANS & UGent)

Eko Indarto (DANS-KNAW)



The DICE project has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 under Grant Agreement no. 101017207



# DICE Digital Preservation Service (DDPS)



DICE Project: Data Infrastructure Capacity For EOSC

- EU Horizon 2020 project, Duration 30 months, ending June 2023

WP4: Task 4.3 Long Term Preservation of data.

- DANS-KNAW

- SURF

“Implementation of a Long Term Preservation (LTP) service for B2SHARE in a certified archive (DANS)”

Objectives:

- Technical specification document

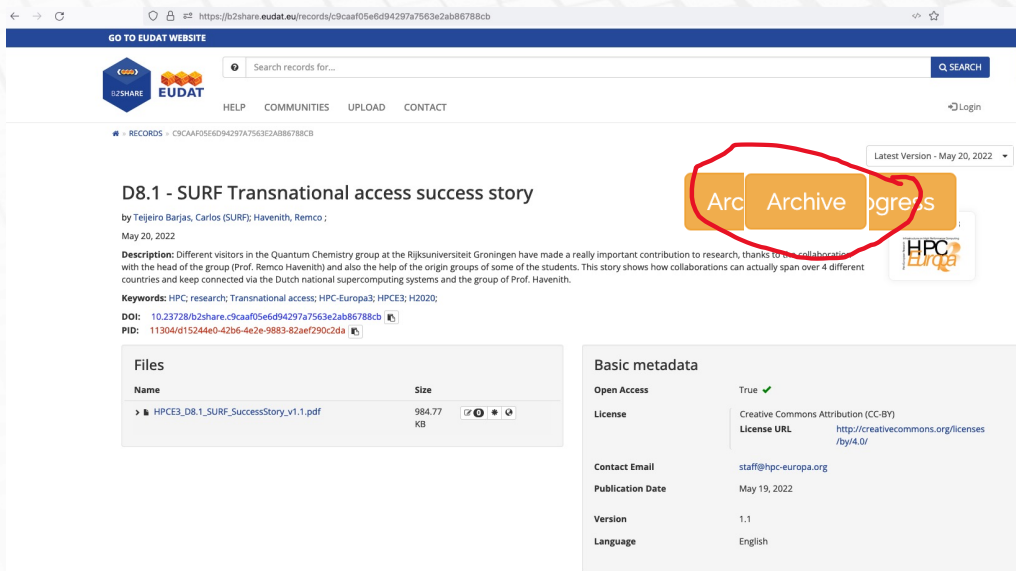
- Reference Long Term Preservation (LTP) implementation: DICE Digital Preservation Service (DDPS)

Implementation:

- “Notification Driven Archiving” using Open Web Standards

## B2SHARE dataset archiving

A single dataset will be archived on demand, according to the repository LTP Policy, by an authenticated repository user.



The screenshot shows a B2SHARE record page for a dataset titled "D8.1 - SURF Transnational access success story". The page includes a search bar, navigation links (HELP, COMMUNITIES, UPLOAD, CONTACT), and a "Login" button. The record details are as follows:

- Title:** D8.1 - SURF Transnational access success story
- Author:** by Teijeiro Barjas, Carlos (SURF); Havenith, Remco
- Date:** May 20, 2022
- Description:** Different visitors in the Quantum Chemistry group at the Rijksuniversiteit Groningen have made a really important contribution to research, thanks to the collaboration with the head of the group (Prof. Remco Havenith) and also the help of the origin groups of some of the students. This story shows how collaborations can actually span over 4 different countries and keep connected via the Dutch national supercomputing systems and the group of Prof. Havenith.
- Keywords:** HPC; research; Transnational access; HPC-Europa3; HPCE3; H2020;
- DOI:** 10.23728/b2share.c9caaf05e6d94297a7563e2ab86788cb
- PID:** 11304/d15244e0-42b6-4e2e-9883-82aef290c2da

The "Files" section contains one file:

Name	Size
HPC_E3_D8.1_SURF_SuccessStory_v1.1.pdf	984.77 KB

The "Basic metadata" section shows:

- Open Access:** True
- License:** Creative Commons Attribution (CC-BY)
- License URL:** <http://creativecommons.org/licenses/by/4.0/>
- Contact Email:** [staff@hpc-europa.org](mailto:staff@hpc-europa.org)
- Publication Date:** May 19, 2022
- Version:** 1.1
- Language:** English

A red circle highlights the "Archive" button in the top right corner of the record page.



## Use Open (Web) Standards:

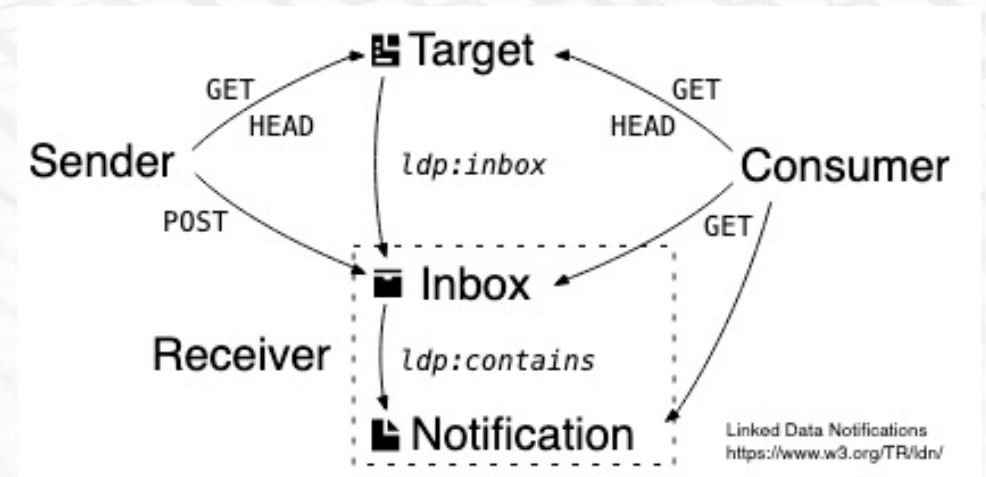
- 📄 Linked Data Notifications (*W3C: Notification system*)
- 📄 ActivityStreams 2.0 (*W3C: Vocabulary*)
- 📄 Signposting (*IETF RFCs based: Navigational*)

- Linked Data Notifications (LDN)

## Notification

📄 W3C Recommendation (Web standard)

📄 <https://www.w3.org/TR/ldn/>



## Vocabulary

A model for representing potential and completed activities using the JSON-LD format.

These make up the messaging/notification payloads.

 W3C Recommendation

 <https://www.w3.org/TR/activitystreams-vocabulary/>

Activity Type	Purpose	Interoperability-Scope
as:Offer	Indicates that the actor is offering the object. If specified, the target indicates the entity to which the object is being offered.	Web-wide
coar-notify:ReviewAction	An action related to a review, or request for a review, of a resource	Scholarly-communication
coar-notify:ArchiveAction	An action related to LTP archiving, or request for LTP archiving, of a resource	Scholarly-communication

## Namespaces:

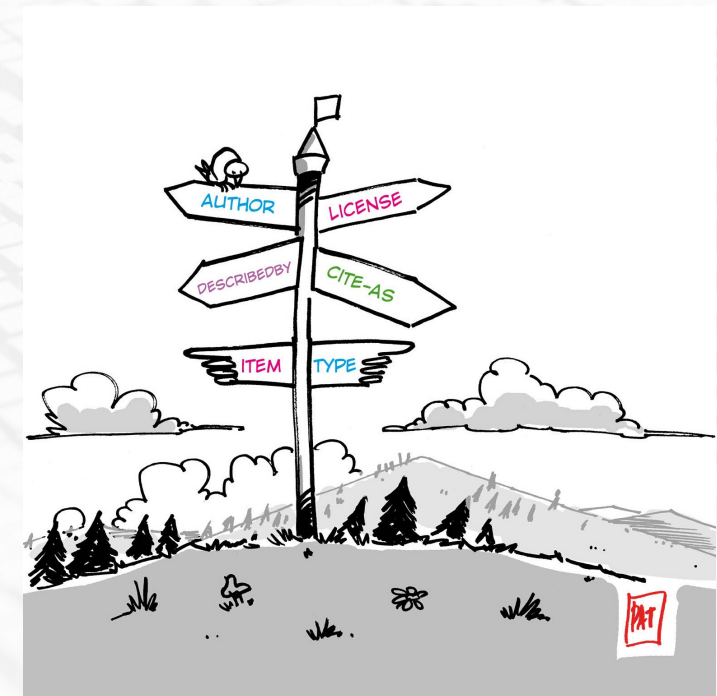
**as** : <https://www.w3.org/ns/activitystreams#>

**coar-notify** : [http://purl.org/coar/notify\\_vocabulary/](http://purl.org/coar/notify_vocabulary/)



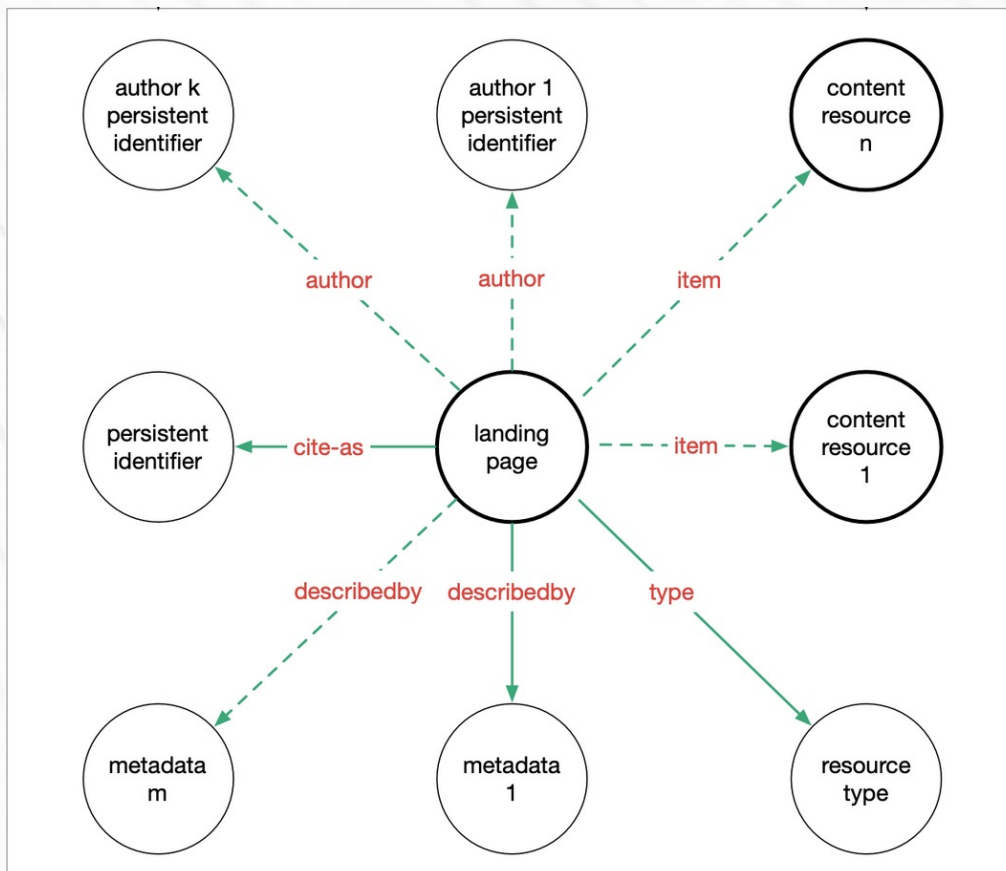
## Navigation

- REST/HATEOAS technology
- Based on IETF RFCs and IANA registered link relation types
- Link Set will very soon be an RFC
- FAIR Signposting Profile:  
<https://signposting.org/FAIR/>

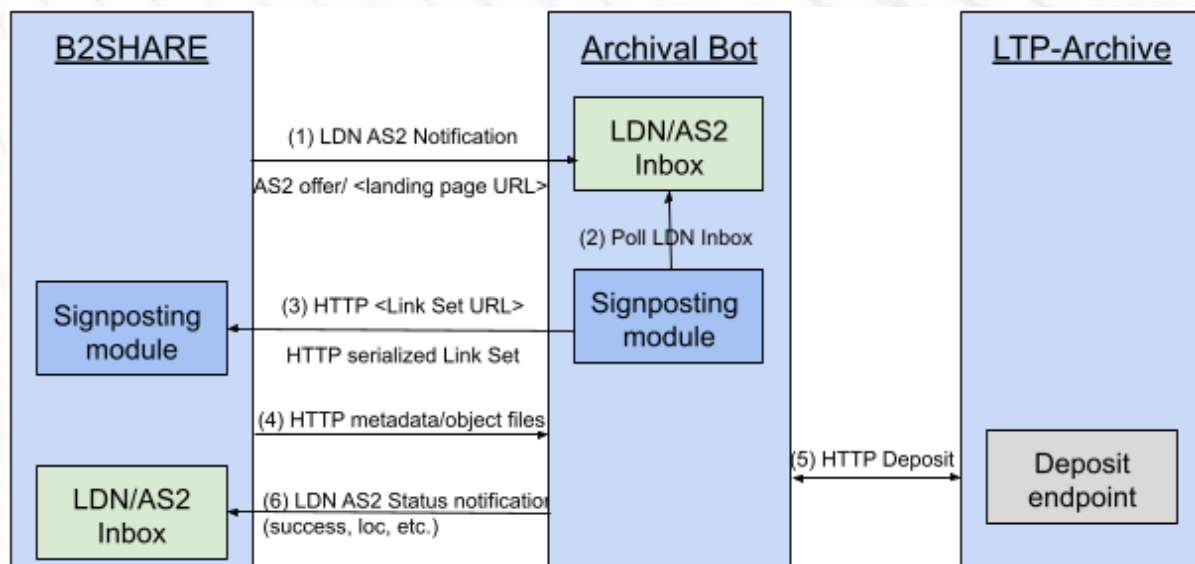


Cartoon by Patrick Hochstenbach





# Putting it all together: DDPS Components diagram



1) An LTP request will be sent from a B2SHARE landing page to the Archival Bot's inbox, conveying the URL of the landing page of the dataset.

2) The Archival Bot will poll the LDN Inbox for such requests.

3) The Archival bot visits the landing page URL, discovers a Link Set provided via Signposting, and retrieves it.

4) The Archival Bot parses the Link Set to obtain URLs for object files and metadata, and then retrieves those.

5) The Archival Bot deposits an archival package containing metadata and object files to the LTP-archive.

6) The Archival Bot reports back on the status of the deposit.




Aligned with



### **COAR** Notify Project

-  Repository and stand-alone review services interoperability Project (Peer review system)

-  <https://notify.coar-repositories.org>

### Researcher Pod: Mellon-funded Project (UGent)

-  Personal research data-pods on the decentralized web

-  <https://mellonscholarlycommunication.github.io/spec-notifications>



Data Archiving and Networked Services

DANS

# Thanks for your attention!

## JOIN OUR COMMUNITY



DICEosc



/company/diceosc

- 📄 Wilko Steinhoff (DANS-KNAW), speaker
- 📄 Eko Indarto (DANS-KNAW)
- 📄 Herbert van de Sompel (KNAW-DANS & UGent)

Acknowledgements:

Hans van Piggelen; Mark van de Sanden; Patrick Hochstenbach; Martin Klein; Paul Walk