DICE
Data Infrastructure Capacity for EOSC

B2INST

ePIC
Persistent Identifiers for eResearch

DataCite
FIND, ACCESS, AND REUSE DATA

EUDAT

SCAN ME

PIDs for Instruments - the B2INST Service

Tibor Kálmán

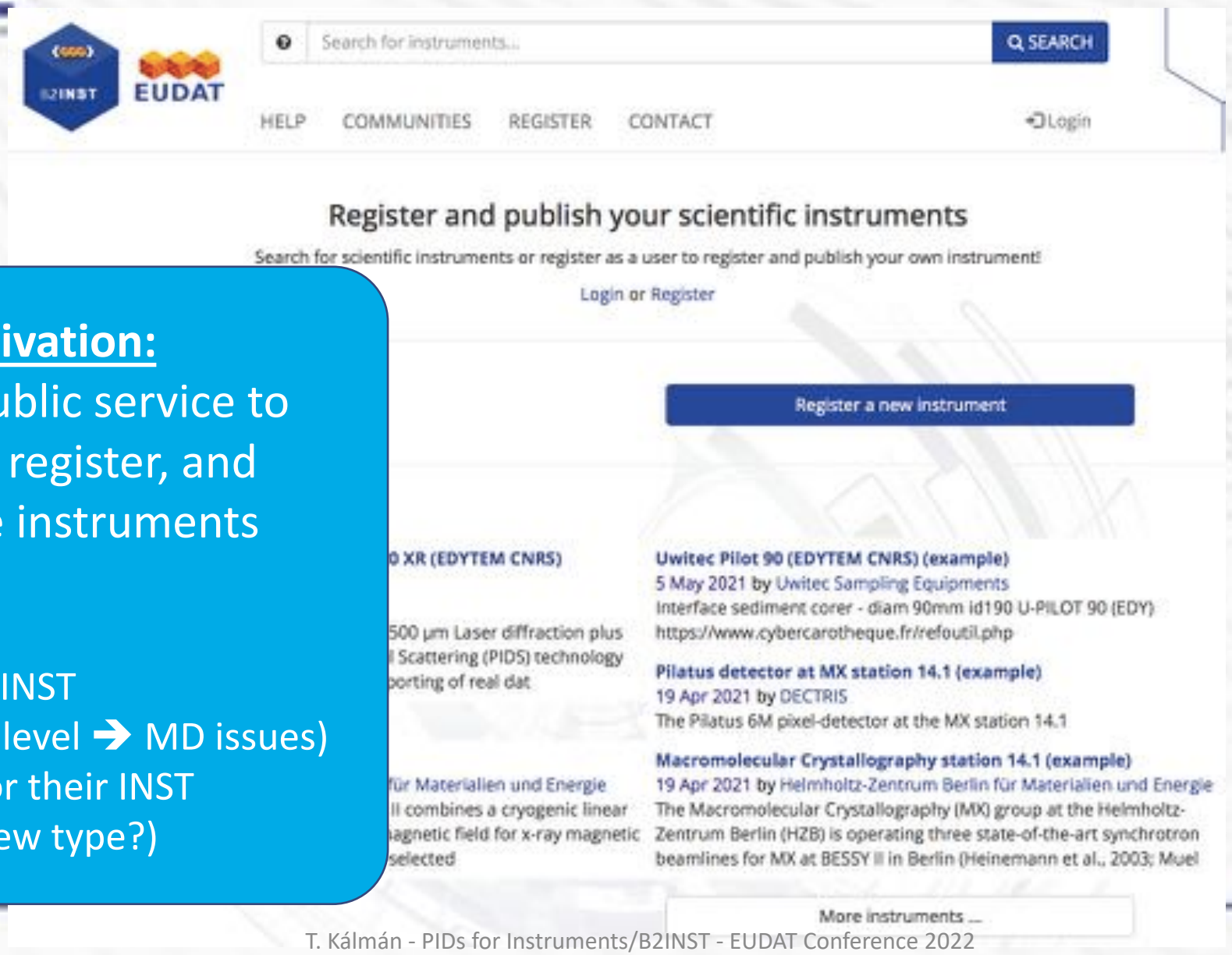
tibor.kalman@gwdg.de

Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen

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

Agenda

- Setting the stage (B2INST)
- Information modelling (PID4INST)
- Management & Strategy
- Metadata indexing & Discovery
- Technology & Development



Motivation:

Create a public service to describe, register, and reference instruments

Trends:

- Registries for INST (community level ➔ MD issues)
- Assign PIDs for their INST (emerging new type?)

B2INST service – landing page

Instrument Name

Instrument Description

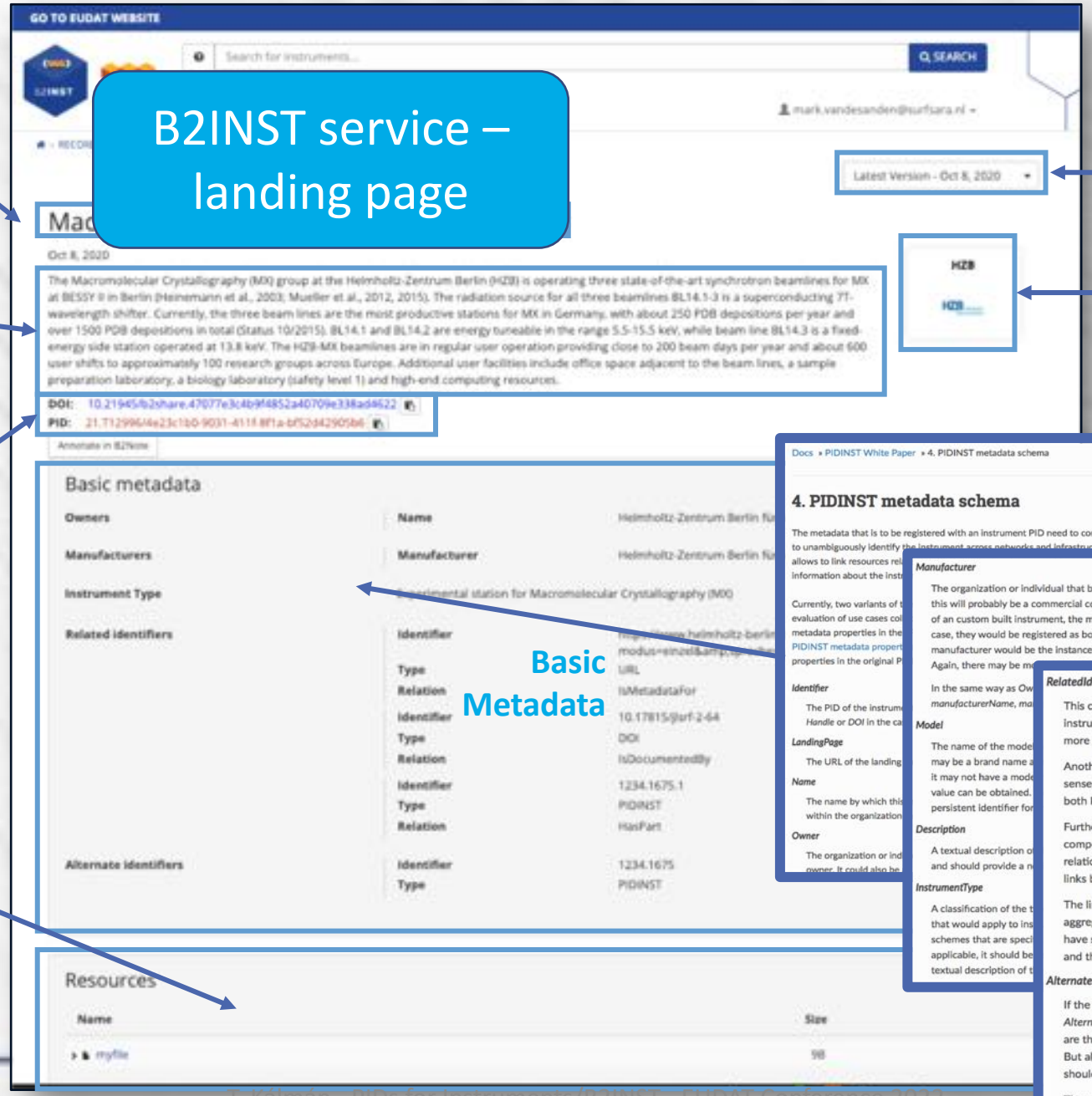
Persistent Identifier (ePIC / DOI)

Optional Files (manuals, pictures of the instrument, etc)

Versioning

Scientific Domain (Community)

Basic Metadata

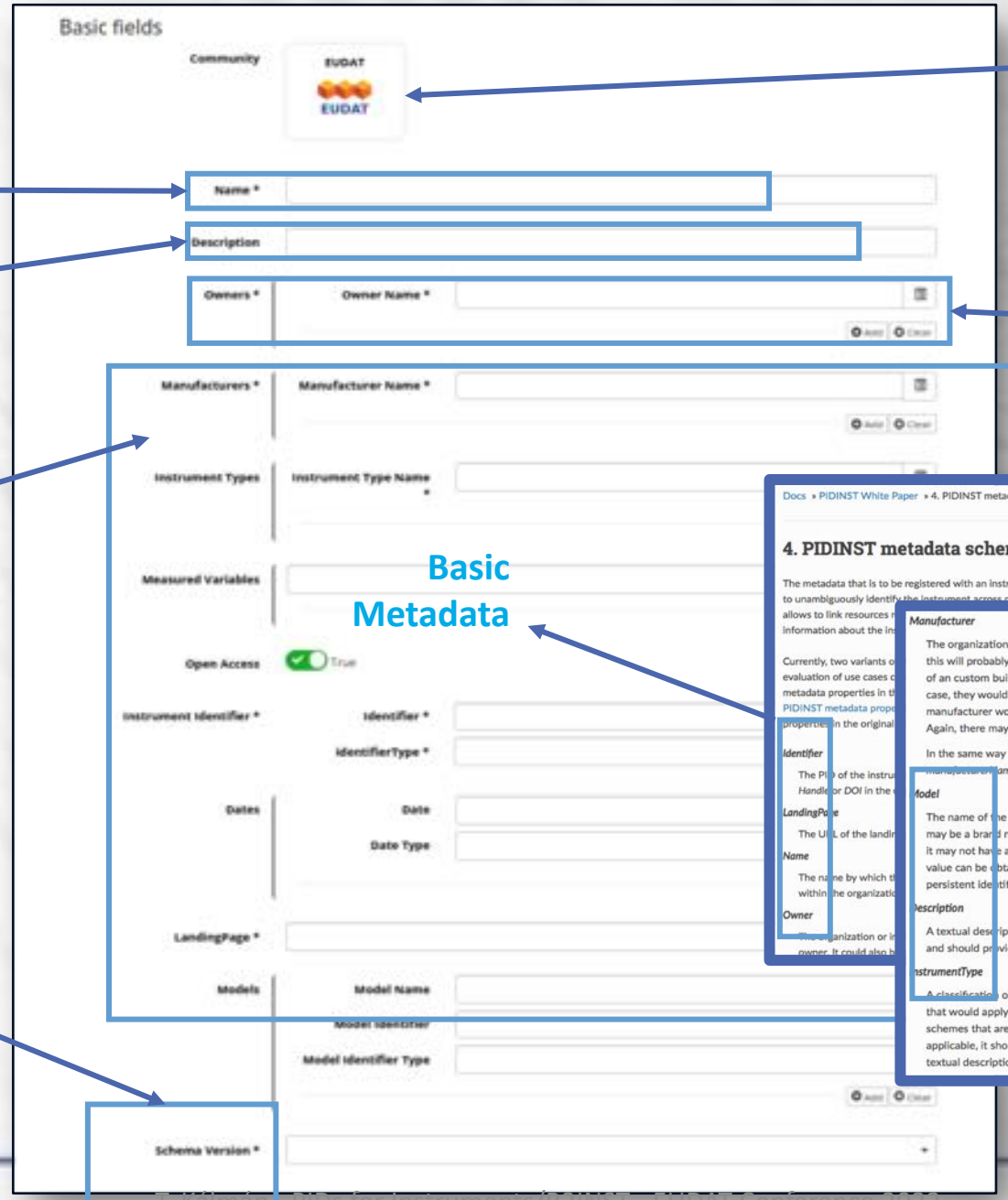


Instrument Name

Instrument Description

INST specific fields

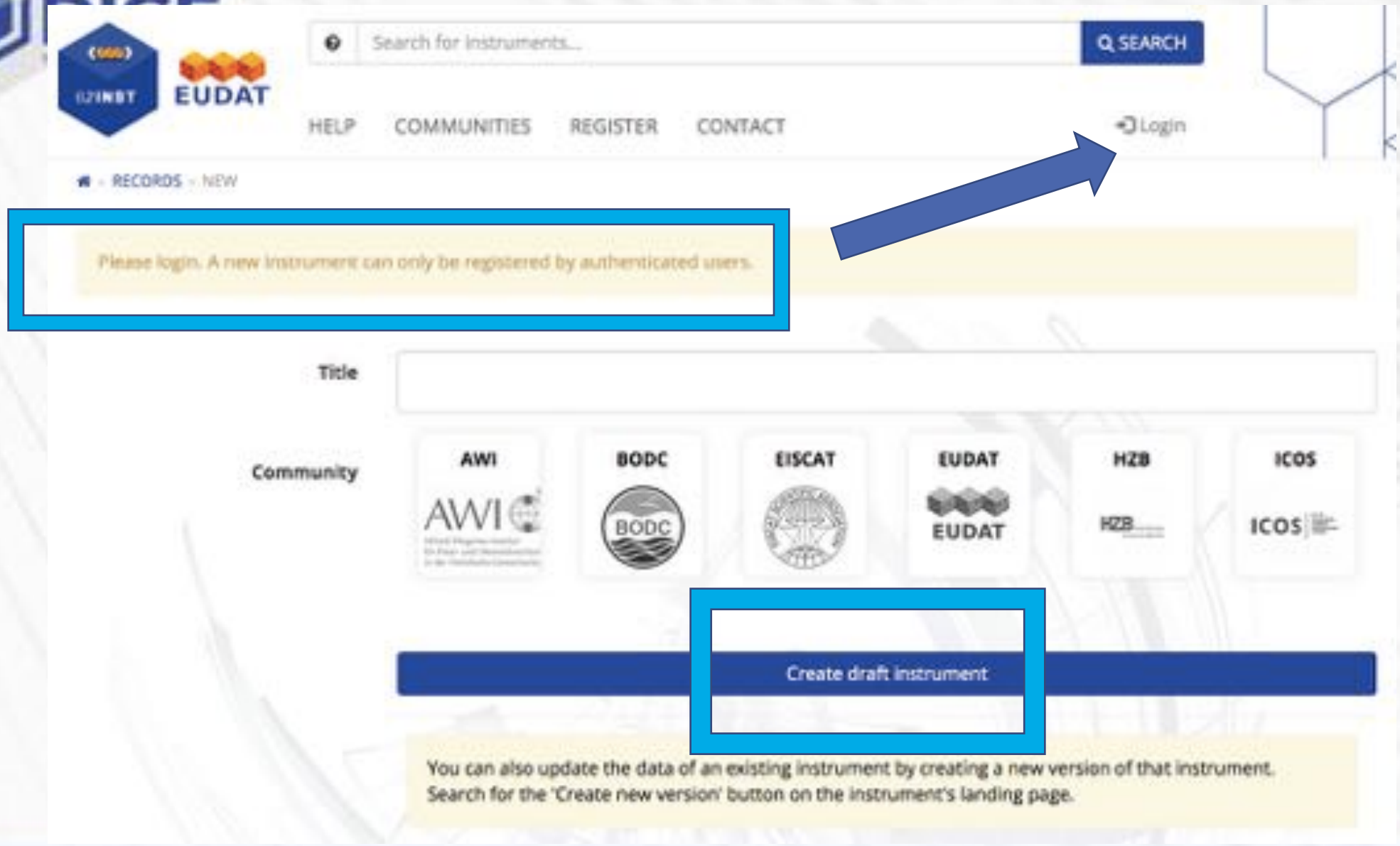
Schema version & Optional Files (manuals, pictures of the instrument)



Community

Owner

Basic Metadata


Search for instruments...
SEARCH
HELP
COMMUNITIES
REGISTER
CONTACT
Login
RECORDS
NEW
Please login. A new instrument can only be registered by authenticated users.
Title
Community
AWI
BODC
EISCAT
EUDAT
HZB
ICOS
Create draft instrument
You can also update the data of an existing instrument by creating a new version of that instrument.
Search for the 'Create new version' button on the instrument's landing page.

Information modeling

Outputs:
1. *Persistent Identification of Instruments*, doi:10.5334/dsj-2020-018
2. *PIDINST White Paper*

g & Data modeling

1. Identify use cases
2. Gather requirements
3. Design an information mode
4. Define a data model
5. Implement a solution

RDA (PID4INST)

Collected real*[*]* use cases (PIDs for instruments)

[*] we deal with researchers!

hat information is needed to describe instruments?

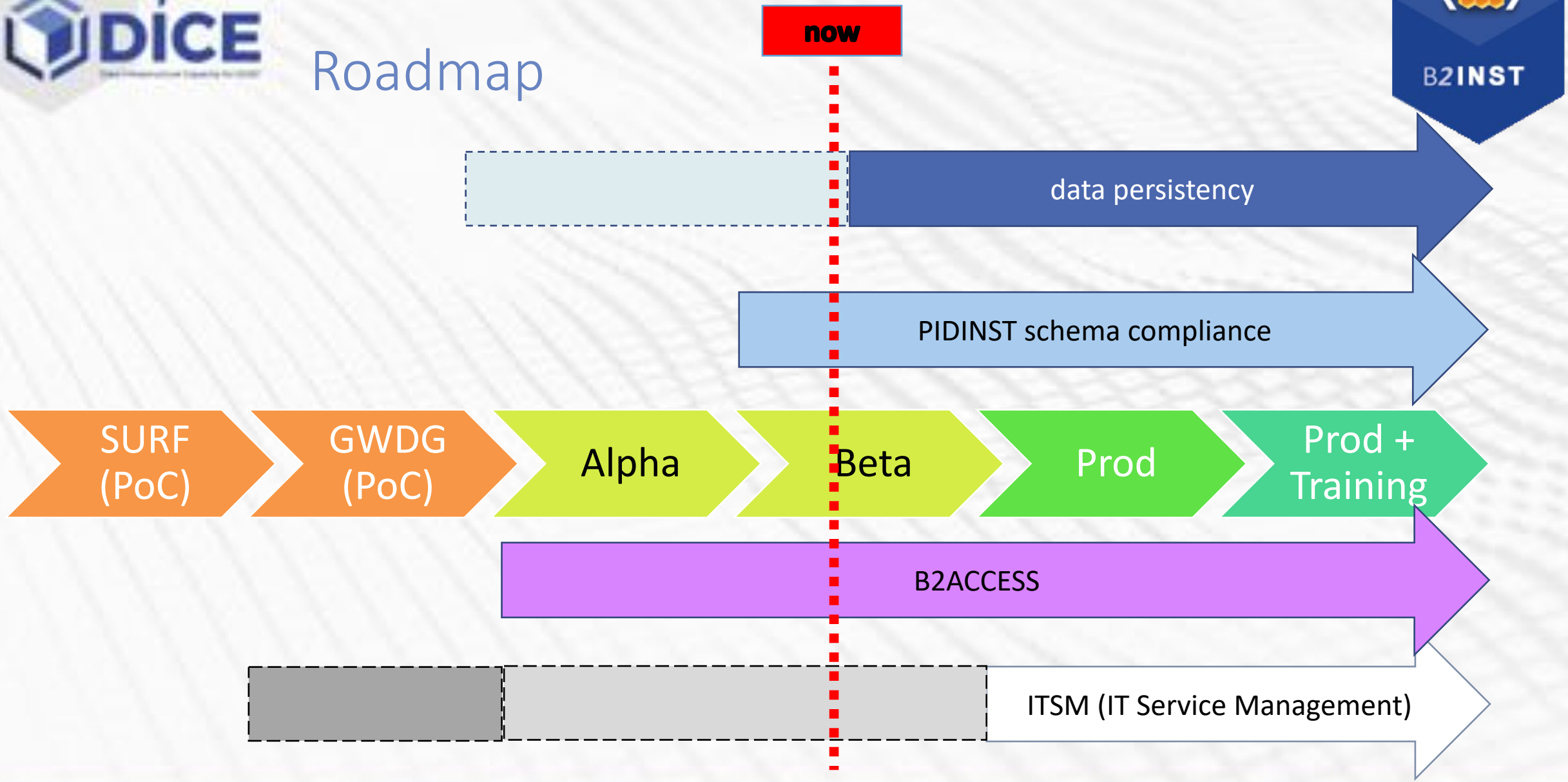
efine a schema and the necessary attributes to store the nformation about the instruments

xplored potential PID service providers & their PIDs. Identified two candidates: ePIC (Handles) / DataCite (DOIs)

???

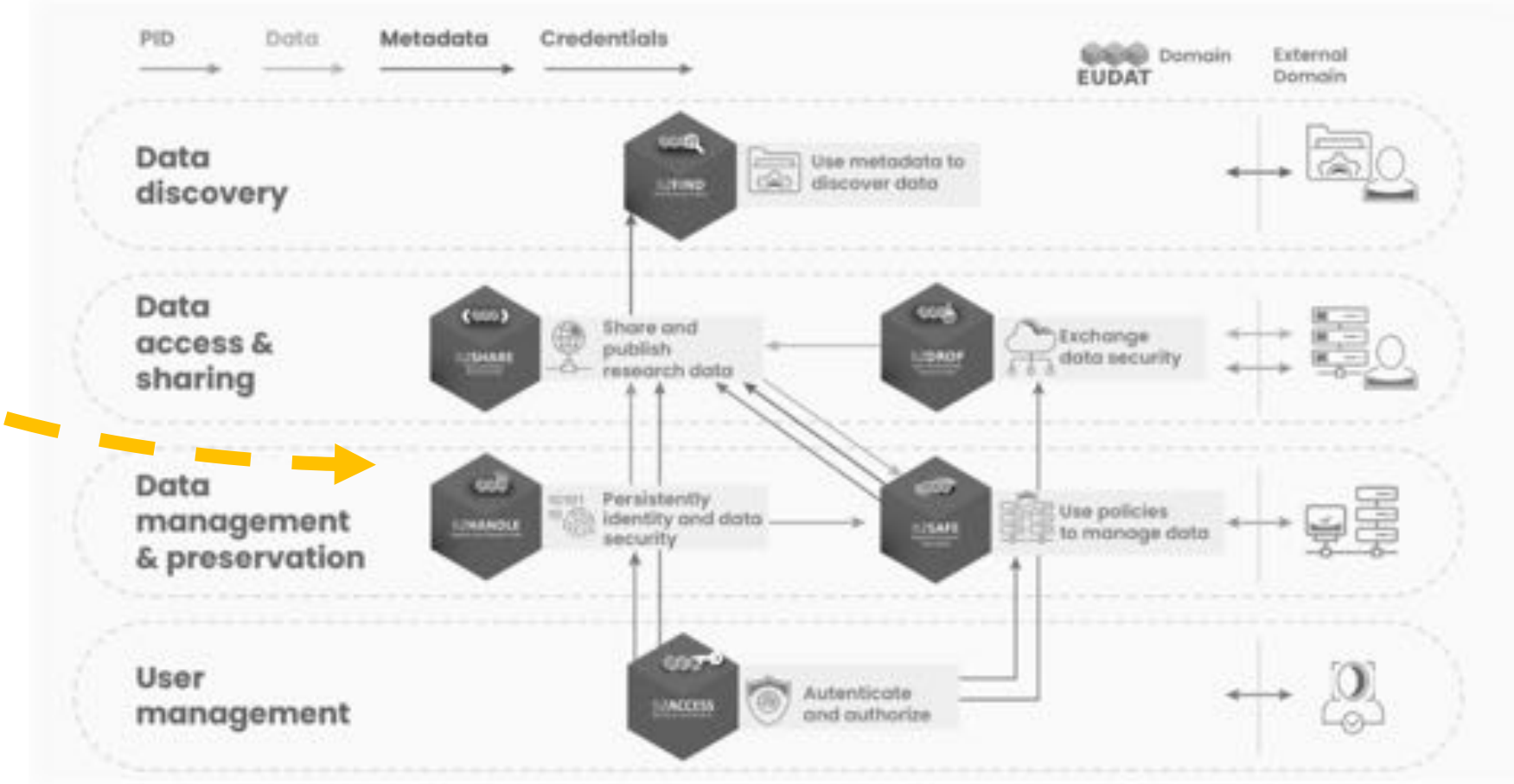
Roadmap, ongoing activities

Roadmap

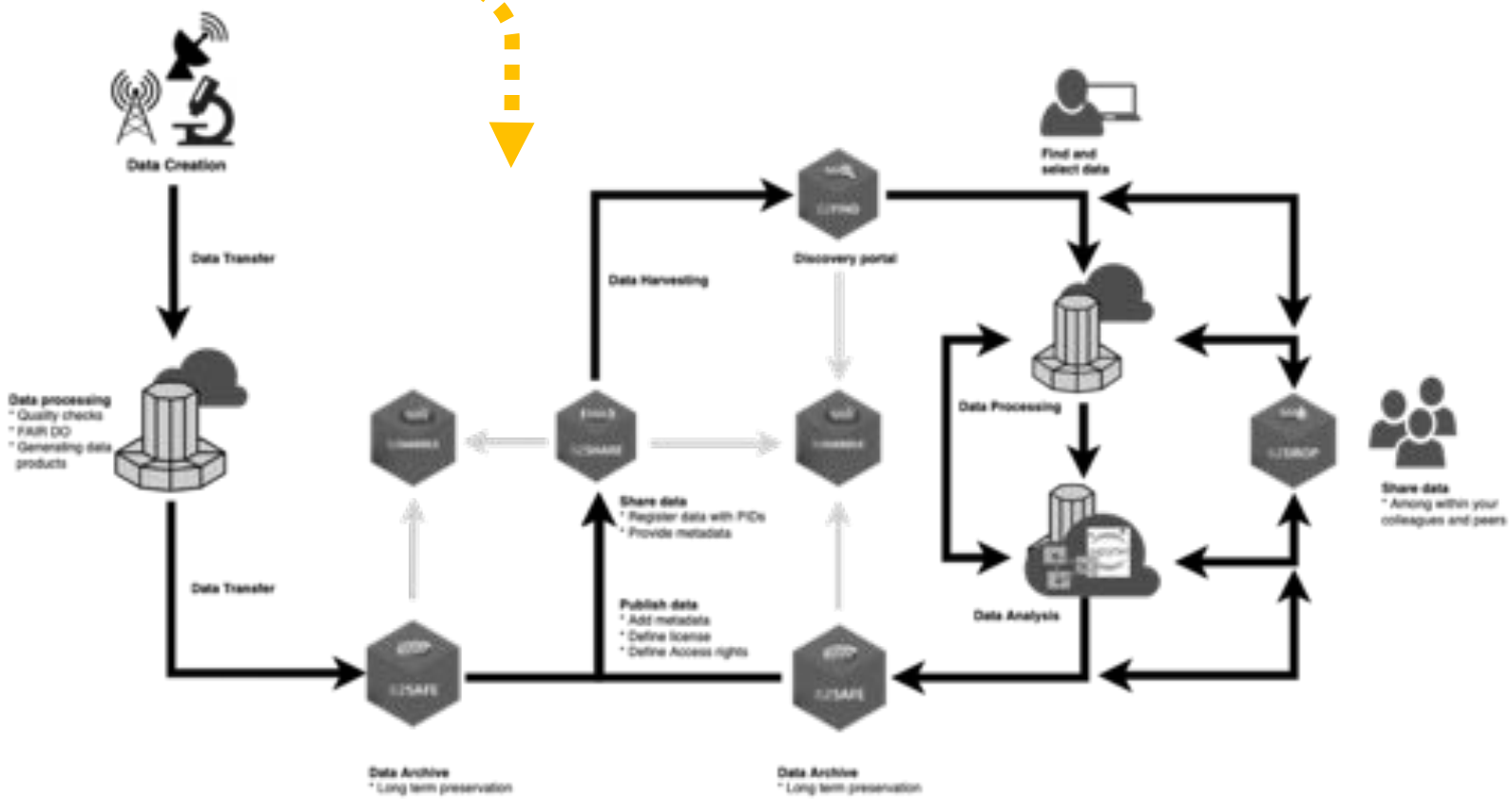

now
data persistency
PIDINST schema compliance
SURF (PoC)
GWDG (PoC)
Alpha
Beta
Prod
Prod + Training
B2ACCESS
ITSM (IT Service Management)

Service Management

The EUDAT Service Suite



Research Data Life Cycle



Management and Strategy

ITSM process

- We are implementing a new service under the umbrella of EUDAT
  - This requires us to follow some ITSM processes
- B2INST is now in Service Portfolio Management (SPM) process
  - SPM is one of the most strategic SM process
  - It defines how new services are introduced & created
  - Describes a managed way how EUDAT implements its B2INST service
- Working on a business model template
  - Example on the right →
- *Not all actions are obvious for us*
  - *"It was quite some time ago when a new service was added to the EUDAT portfolio."*

Business Model Template

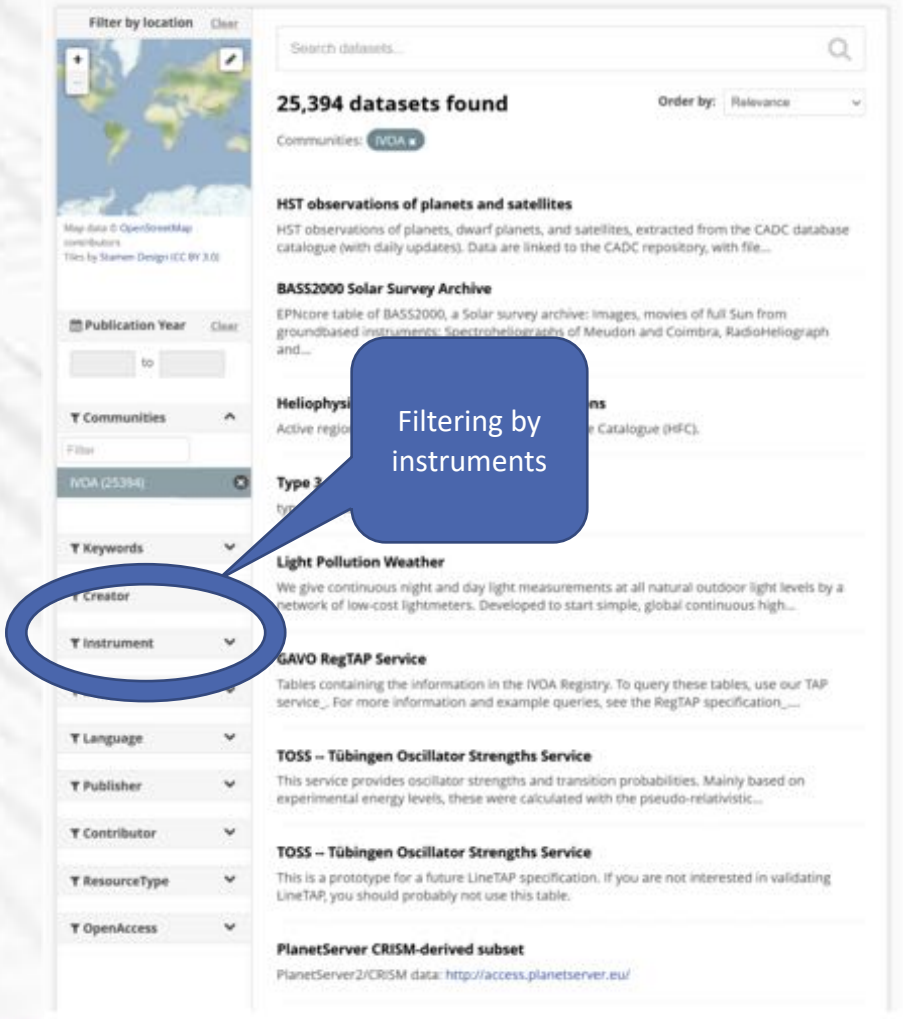
- *[…]*
- *Customers and Users*
- *[…]*
- *Value proposition and service description*
- *[…]*
- *Required partners*
- *Technology and Architecture*
- *Risk analysis*
- *[…]*
- *Communication and Marketing*

Indexing and Discovery

B2INST meets B2FIND

B2INST → B2FIND → EOSC / OpenAIRE

- Goal:
  - make use of registered instruments' metadata
- B2FIND
  - Provides metadata indexing service and joint metadata catalogue
  - Powerful discovery portal, which allows users to find data collections, objects, etc based on relations
  - Possible gateway to EOSC
- Prerequisite: Objects (data sets) in data repositories are linked to instruments (PID)
- What to show? Approaches being evaluated:
  - Link to instruments' landing page
  - Search in B2INST



Source: IVOA (Interop)

Ongoing activities

- Service provider partner has been changed
  - SURF → GWDG
- ITSM
  - PoC, alpha and beta successful
  - Towards a production service
  - Still waiting for B2SHARE v3
  - Providing a Production and a Test (Training) instance
  - Refine Service Level & agree on Terms of Use
- Schema
  - DataCite v4.5
  - Checking full compliance with the PIDINST metadata schema
  - Options for community extentions
- Keeping data persistent from Beta onwards
- PIDs
  - DOI prefix provided by DataCite
  - PID prefix provided by ePIC
- Discovery service
- Types & Type registries

Technology & Development

- Current B2INST is based on B2SHARE (“Invineo”, “Zenodo”)
  - Still committed to B2SHARE
- But… our development roadmap& timeline differs from B2SHARE’s
  - Handling of metadata & schemas
  - Supporting testbeds

Requirements

- AAI
- Payloads / additional objects
- OAI-PMH
- Central / distributed approach
- [DOIP]

JOIN OUR COMMUNITY

DICEosc

/company/diceosc

Tibor Kálmán
tibor.kalman@gwdg.de
GWDG - Gesellschaft für wissenschaftliche Datenverarbeitung mbH
Göttingen

PID (Handle):
21.11101/0000-0007-F961-3

Resolvable URL:
https://hdl.handle.net/**21.11101/0000-0007-F961-3**

SCAN ME

Backup slides

Research Data Life Cycle