

# B2SAFE

Johannes Reetz, RZG – Max Planck Society

B2SAFE @ 3<sup>rd</sup> EUDAT conference

Amsterdam

Sep 24-25, 2014



# What is B2SAFE?

**B2SAFE** is a service which allows repositories to **implement data management policies on their research data** across multiple administrative domains in a trustworthy manner.

- abstraction layer of large scale, heterogeneous data storages
- guards against data loss in long-term archiving
- allows to optimize access for users (e.g. from different regions)
- brings data closer to facilities for compute-intensive analysis

# Who uses B2SAFE & why?

- **Repositories** which do not have the capacity or funding to offer reliable storage and access services over longer periods in time.
- **Repositories** without adequate compute capacity for data-intensive computational services based on their data.
- **Data producers** who need to be sure that trusted centers are taking care of their data.
- **Data consumers** to get access to optimized services on data sources of interest to them.
- **Data consumers** who want to apply interdisciplinary data-intensive methods using data collected from various communities.

# B2SAFE features

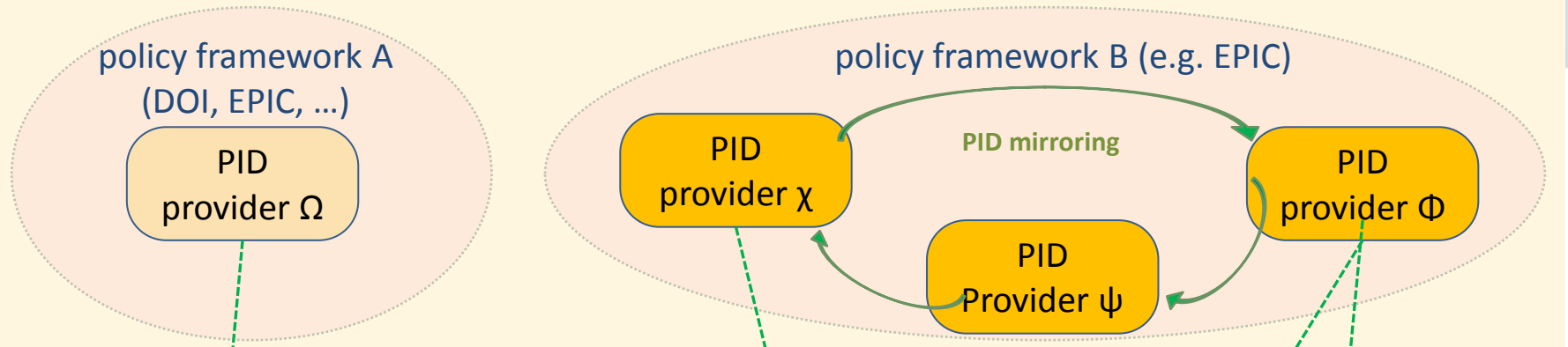
- Execution of **auditable data policy rules**
- Use of **persistent identifiers (PIDs)** for different purposes
- Respects data owners' rights to **define the access rights** for their data
- Respects the data owners' rights to decide **how and when their data is made public referenceable**
- Centrally managed **Data Management Policies**
- The **policy rules are distributed** and implemented site-locally
- Support for **repositories packages** (e.g. DSPACE, FEDORA)
- Data transfer from repositories to B2SAFE also via B2STAGE facilities (GridFTP, HTTP).



# B2SAFE for replication

*logical layer*

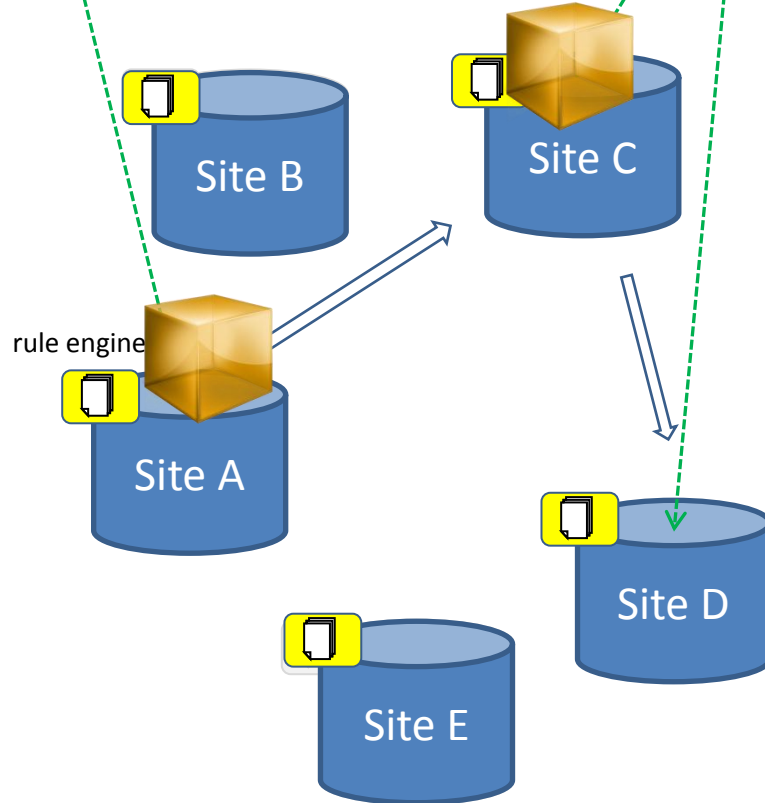
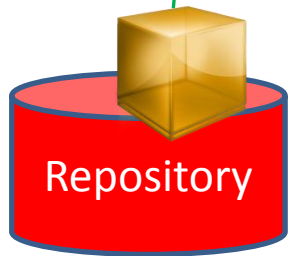
**Handle PID infrastructure**



Handle prefix managed via  $\Omega$

Handle prefix managed via  $\chi$

Handle prefixes managed via  $\Phi$





## B2SAFE

allows repositories to implement data management policies on their research data

# EUDAT Collaborative Data Infrastructure



## Browse

- o My Sites
- o Admin Domains
- o Projects
- o Sites
- o Service Groups
- o Services

## Add

- o Add Site
- o Add Service Group
- o Add Service
- o Add Downtime

## Downtimes

- o Active & Imminent

## About GOCDB5

- o Doc, Help & Support

## Search

## User Status

Registered as:  
Johannes Reetz

[View Details](#)  
[Manage Roles](#)



## Service Groups

All Service Groups in GOCDB.

[What is a service group?](#)

Filter (clear)

Scope: EUDAT

Extension Name: (none)

23 Service Groups

Name	Description	Scope(s)
» EUDAT_COLLABORATIVE_SERVICES	EUDAT Collaborative Services	EUDAT
» EUDAT_CORE_AAI_SERVICES	Service endpoints which belong or are related to the EUDAT AAI services	EUDAT
» EUDAT_CORE_B2FIND	EUDAT Joint Metadata Service	EUDAT
» EUDAT_CORE_B2SAFE	All service endpoints which belong or are related to the EUDAT B2SAFE service	EUDAT
» EUDAT_CORE_B2SHARE	EUDAT Simple Store Services	EUDAT
» EUDAT_CORE_B2STAGE	Service endpoints which belong or are related to the EUDAT core service B2STAGE	EUDAT
» EUDAT_CORE_EPICPID	The EUDAT coordinated core PID service using the EPIC service	EUDAT
» EUDAT_CORE_MYPROXY	EUDAT coordinated core MyProxy services	EUDAT
» EUDAT_CORE_SR4CLARIN	Service endpoints which belong to the EUDAT B2SHARE for all CLARIN repositories	EUDAT
» EUDAT_CORE_SR4CLARIN_CUNI	Service endpoints which belong to the EUDAT core service "Safe Replication" for CLARIN project: CUNI (Repository: LINDAT)	EUDAT
» EUDAT_CORE_SR4CLARIN_EKUT	Service endpoints which belong to the EUDAT core service "Safe Replication" for CLARIN project: EKUT (Repository: EKUT)	EUDAT
» EUDAT_CORE_SR4CLARIN_REPLIX	Service endpoints which belong to the EUDAT core service "Safe Replication" for CLARIN project: REPLIX (Repository: MPI-PL)	EUDAT
» EUDAT_CORE_SR4ENES	Service endpoints which belong to the EUDAT core service	EUDAT
» EUDAT_CORE_SR4EPOS	Service endpoints which belong to the EUDAT core service "Safe Replication" for EPOS	EUDAT
» EUDAT_CORE_SR4INCF	Service endpoints which belong to the EUDAT core service	EUDAT

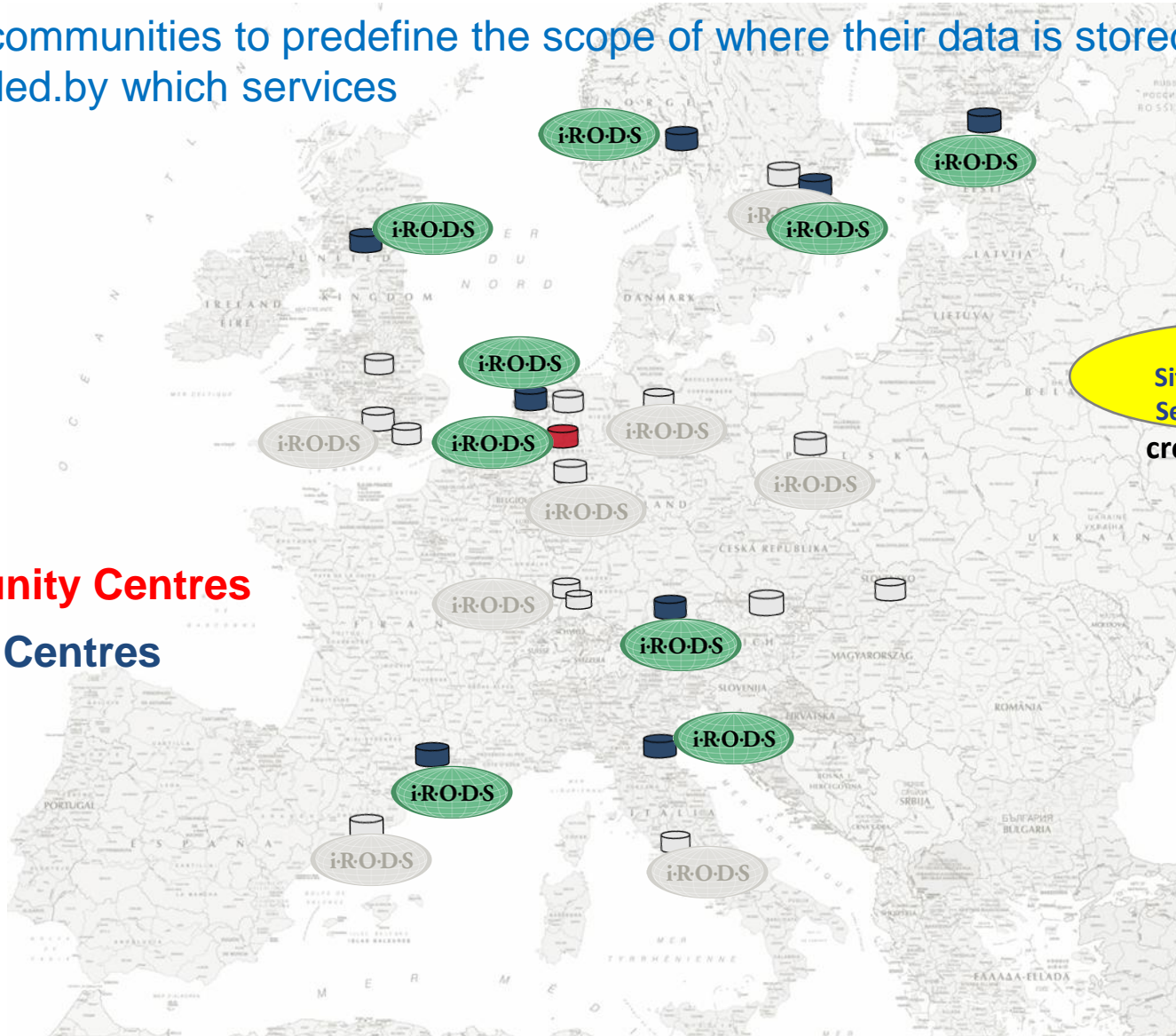
REGISTRY  
es, Services,  
ServiceGroups  
eg.eudat.eu





# Service Groups define *Islands of Trusted Service Endpoints*

Enables communities to predefine the scope of where their data is stored and handled by which services

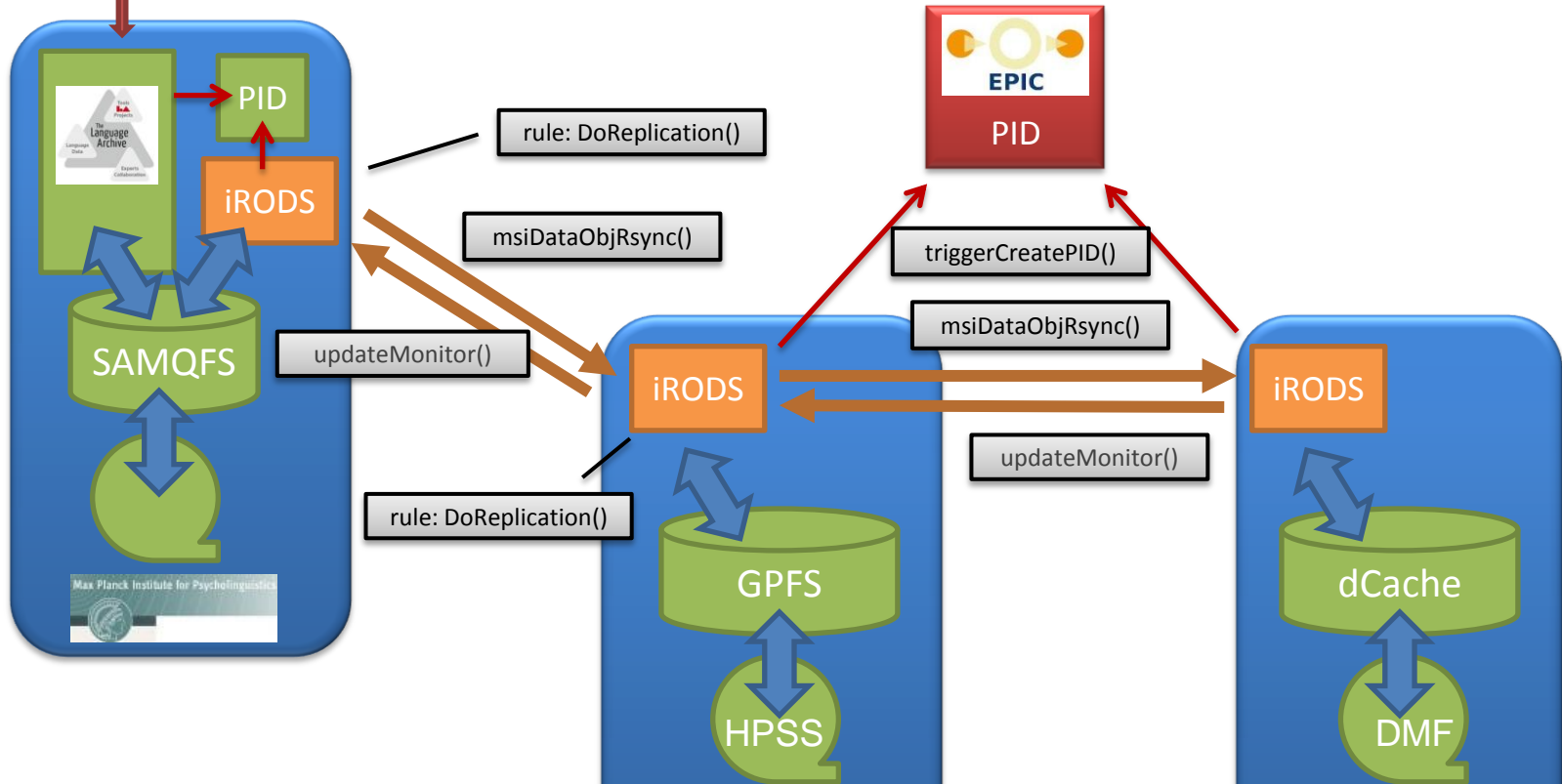


 **Community Centres**

 **EUDAT Centres**



# CLARIN B2SAFE use case



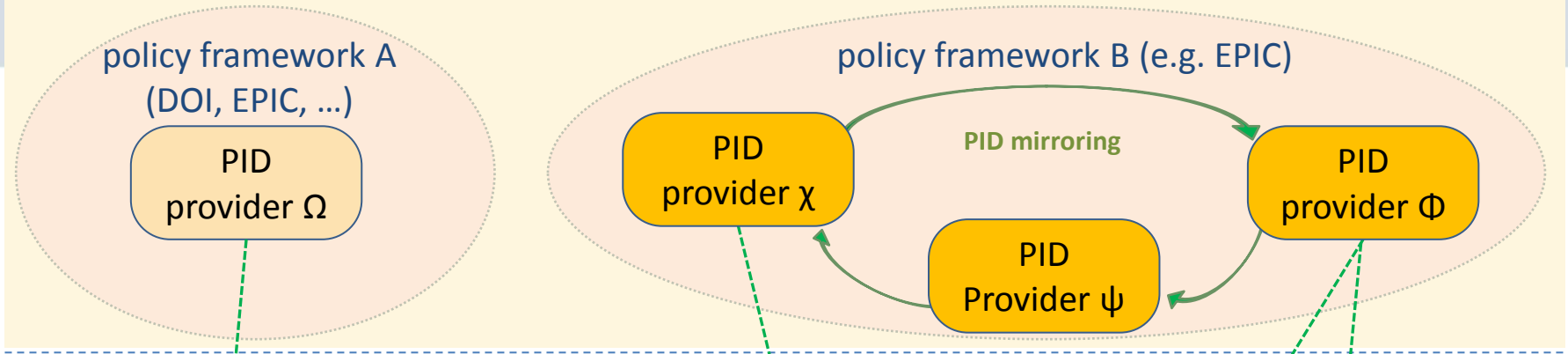
More details: see Poster

*B2SAFE Production Instances - some examples*

Elena Erastova et al., RZG

logical layer

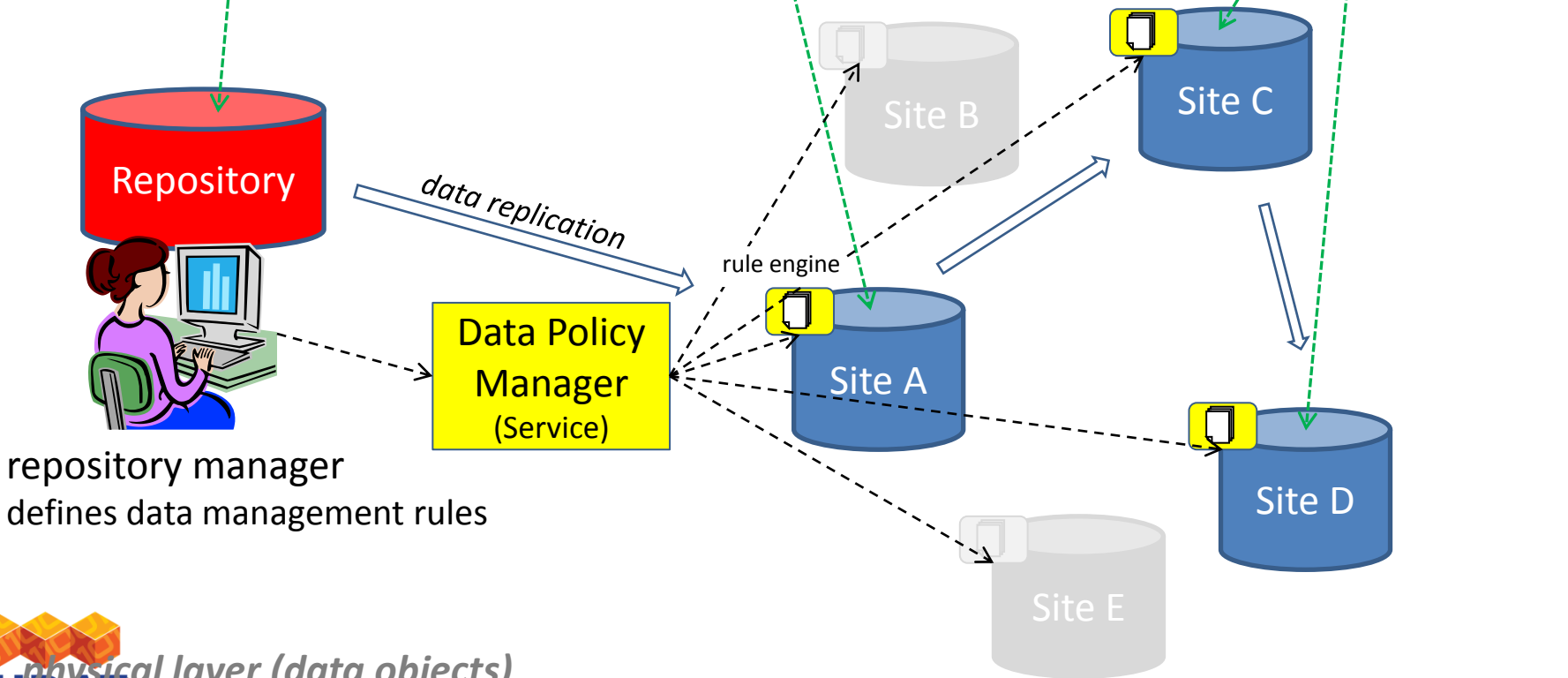
Handle PID infrastructure



Handle prefix managed via  $\Omega$

Handle prefix managed via  $\chi$

Handle prefixes managed via  $\Phi$

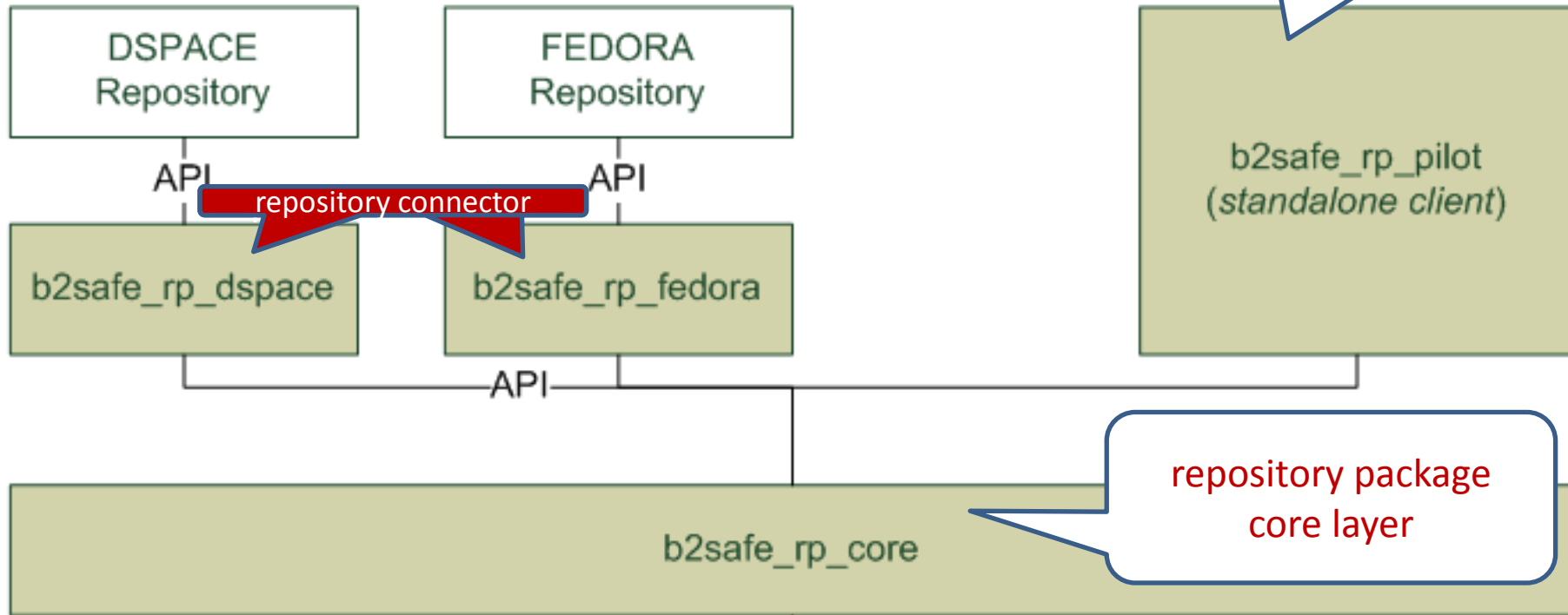




# B2SAFE Repository Package

connect widely used repository software to B2SAFE in a generic way

# Repository package components and interfaces



**More details: see Poster**

*Using EUDAT B2SAFE safe replication from a DSpace digital repository*

Josef Mitsuka et al., CUNI



# B2SAFE-related developments

B2SAFE core

<https://github.com/EUDAT-B2SAFE/B2SAFE-core/releases/latest>

B2SAFE Data Policy Manager

<https://github.com/EUDAT-B2SAFE/B2SAFE-DPM>

B2SAFE Repository Package

<https://github.com/ufal/lr-b2safe-dspace> (Dspace plugin)

<https://github.com/EUDAT-B2SAFE/B2SAFE-repository-package> (B2SAFE package)

# Posters about implementation details

*Using EUDAT B2SAFE safe replication from a DSpace digital repository*

Josef Mitsuka et al., CUNI

*B2SAFE Production Instances - some examples*

Elena Erastova et al., RZG

# References

Claudio Cacciari, CINECA

Stephane Coutin, CINES

Francesca Iozzi, Adil Hasan, UiO

Willem Elbers, MPI for Psycholinguistics

Johannes Reetz, RZG/MPS

B2SAFE core (Rules,..)

B2SAFE/Repository packages

Data Policy Manager

Data Management Policies

B2SAFE product

Dejan Vitlacil, KTH/PDC, Robert Verkerk, SURFsara

PID Services in EUDAT

Giuseppe Fiameni, Giacomo Mariani, CINECA

Shaun de Witt, STFC

John Kennedy, RZG/MPS

Mark van de Sanden, SURFsara, NL

B2STAGE (w/ B2SAFE interface)

HTTP/B2SAFE interface

EUDAT Site and Service Reg.

Technologies, Service Building



## Agenda 13:30 – 15:00

15min: *Introduction*

Johannes Reetz, RZG

25min: *The Data Policy Manager for B2SAFE*

Willem Elbers, Max Planck Institute for Psycholinguistics

Adil Hasan, University of Oslo

15min: *Proteomics repositories integration using EUDAT resources*

Rafael C Jimenez, ELIXIR

15min: *IPSL Climate Data Management backed by B2SAFE*

Franck Corsini, Institut Pierre Simon Laplace

Stéphane Coutin, CINES

15min: *B2SAFE at ODC and future EIDA/EPOS-S plans within EUDAT2020*

Luca Trani, KNMI and Orfeus Data Centre

5min: *Questions*