Facilitating Data Upload, Access and Deployment

Rainer Stotzka, Marcus Hardt,

LSDMA and DARIAH Collaborations
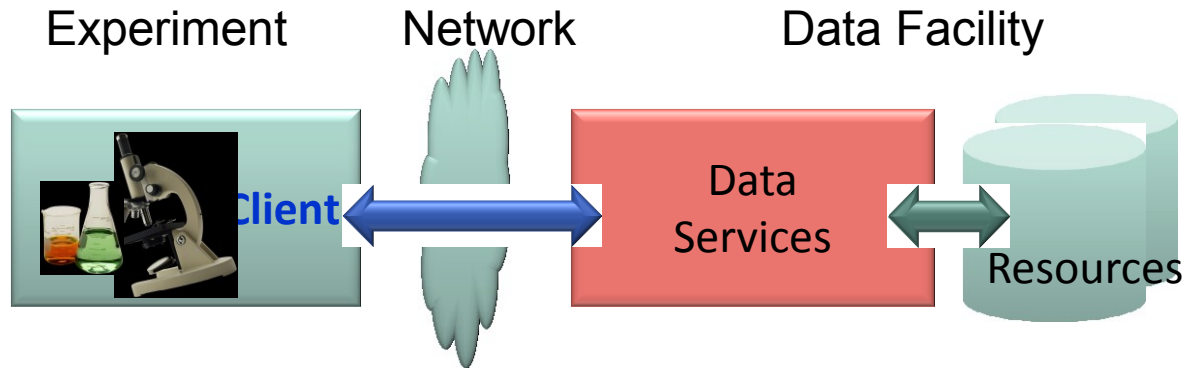
Large Scale Data
Management and Analysis

Digital Research Infrastructure for
the Arts and Humanities

Data Upload and Data Access

Experiment   Network   Data Facility

Client ⟷ Data Services ⟷ Resources

Data Services:
- Ingest / Access
- Storage virtualization
- Meta data mgmt.
- Archive service + PID
- Federations
- Workflow mgmt.
- Community services

Facilitation from the users perspective:
Interfaces and **clients** are easy-to-use, reliable, stable, sustainable
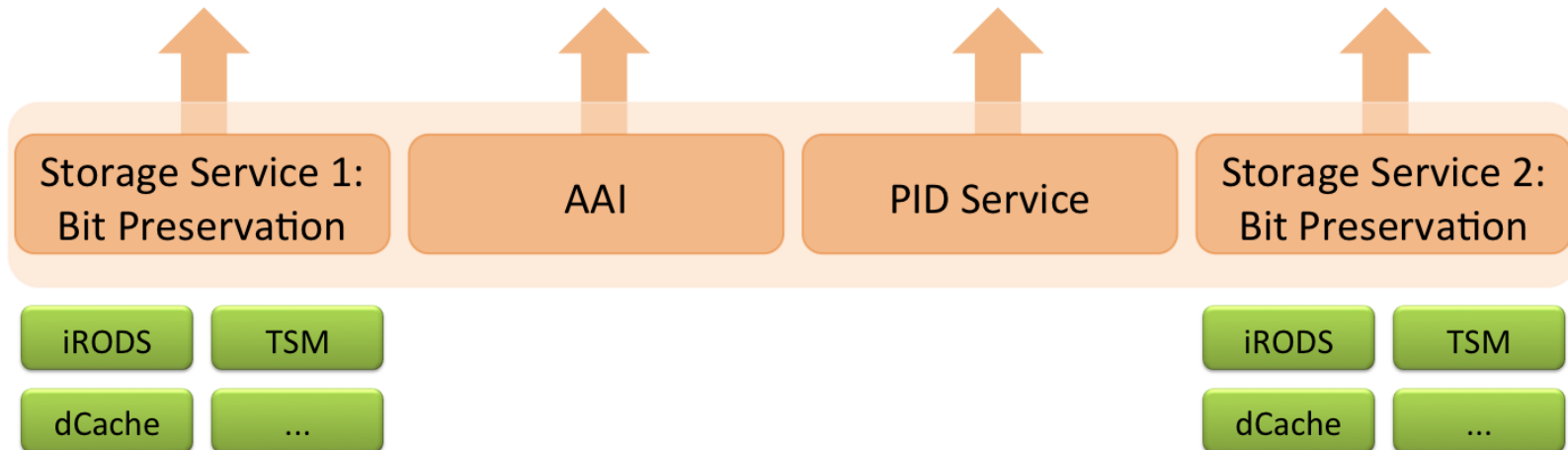
Challenges:
- Federated environment, not static
- Gaps between existing components, no interoperability
- Technology shifts expected, missing components

Required:
→ ***Comprehensive framework*** to handle Big Data
→ ***Seamless integration*** of the components, *interoperability*

Example: Archive Service

Portal and Web

Data Archive Service

Create
Read
Update
Delete
Authenticate
Authorize
Info
Reference

Storage Service 1: Bit Preservation
AAI
PID Service
Storage Service 2: Bit Preservation

iRODS
TSM
dCache
...

iRODS
TSM
dCache
...

Sharing Data: Generic Search

Data without meta data is useless!

Collection Registry: Registration of data collections
Schema Registry: Description of internal structure

Access to contents!

1. Selection of collections:
   Important to solve the research problem

2. Crosswalks:
   Analysis of selected collections and their schemas

3. Facetted search and filter:
   Parametrization

Selected Collections