

DLCL Key Technologies: October 2016

Rainer Stotzka, Swati Chandna, Richard Grunzke, Volker Hartmann, Michael Hausmann, Jürgen Hesser, Thomas Jejkal, Ralph Müller-Pfefferkorn, Francesca Rindone, Danah Tonne, Eberhard Schmitt, Margund Bach, Ajinkya Prabhune, Torsten Schrade



- ***High Throughput Microscopy:***
 - Selective Plane Illumination Microscope (Karlsruhe)
 - Gen Scans (Dresden: TU + MPI CBG)
- ***Nanoscience foundries and fine analysis (NFFA Europe)***
EU
- ***Dariah & eCodicology & MASi***
Arts & Humanities, ESFRI DARIAH EU + BMBF DARIAH DE,
Metadata Management for Applied Sciences (MASi)
- ***Light Optical Nanoscopy*** (Heidelberg, Mannheim, Mainz)
- ***CRC 980***
INF & Guest projects

Build up an **NFFA Information and Data Repository**

Distributed repository system for (meta-)data located at different nanoscience facilities all over Europe

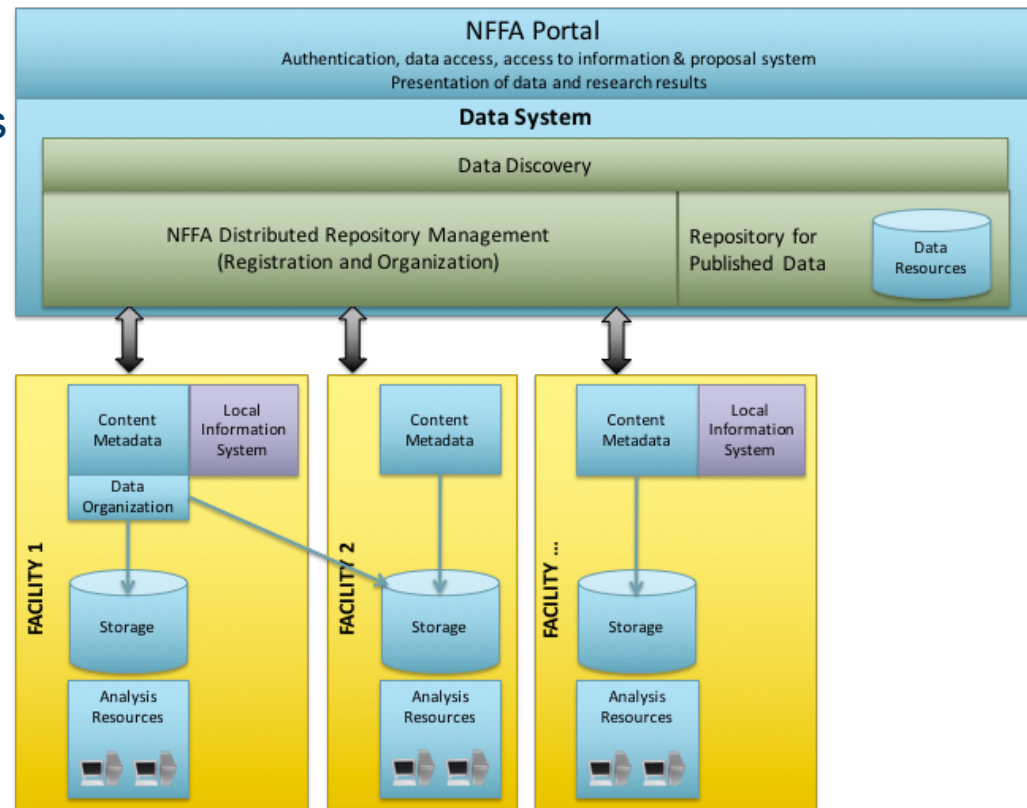
Common **AAI** supported by all facilities

Access via NFFA Web **Portal**

- Proposal creation
- Dataset registration, retrieval, update and sharing
- Dataset publication

Publishing data repository for publicly available datasets

- Registered using a nanoscience metadata schema defined in RDA WG



NFFA Europe



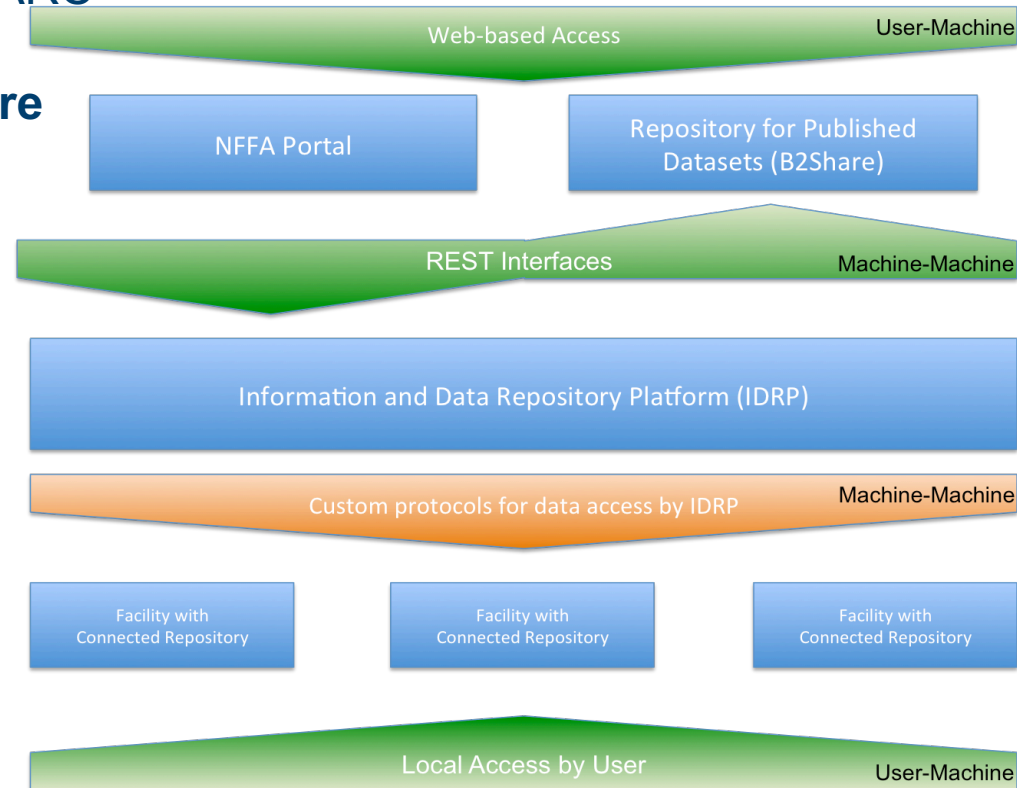
Deliverable on **AAI** concept

- Primary access is Web-based via NFFA portal
- NFFA portal serves as identity provider
- Umbrella (optional)
- Investigation of outcomes of e.g. AARC

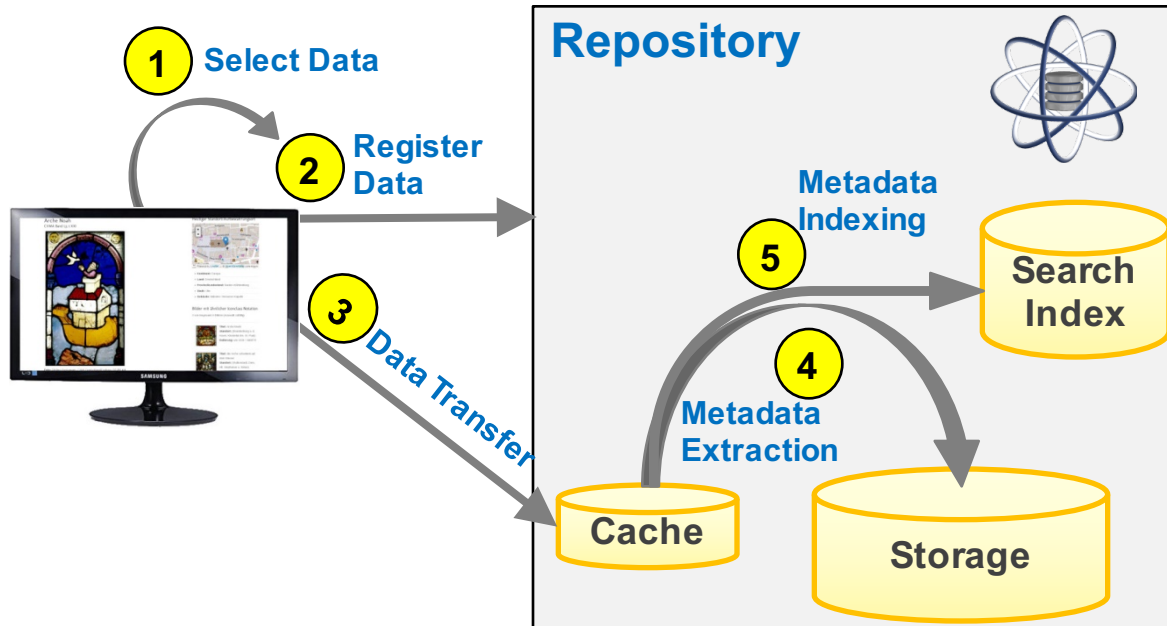
Deliverable on **repository architecture**

- Implementation of metadata standard for nanoscience
- Top-down approach with minimal influence to local facilities
- Integration of B2Share for publishing datasets

Testbed with first prototype available in **March 2017**



MASi – Metadata for Applied Sciences: Corpus Vitrearum Medii Aevi



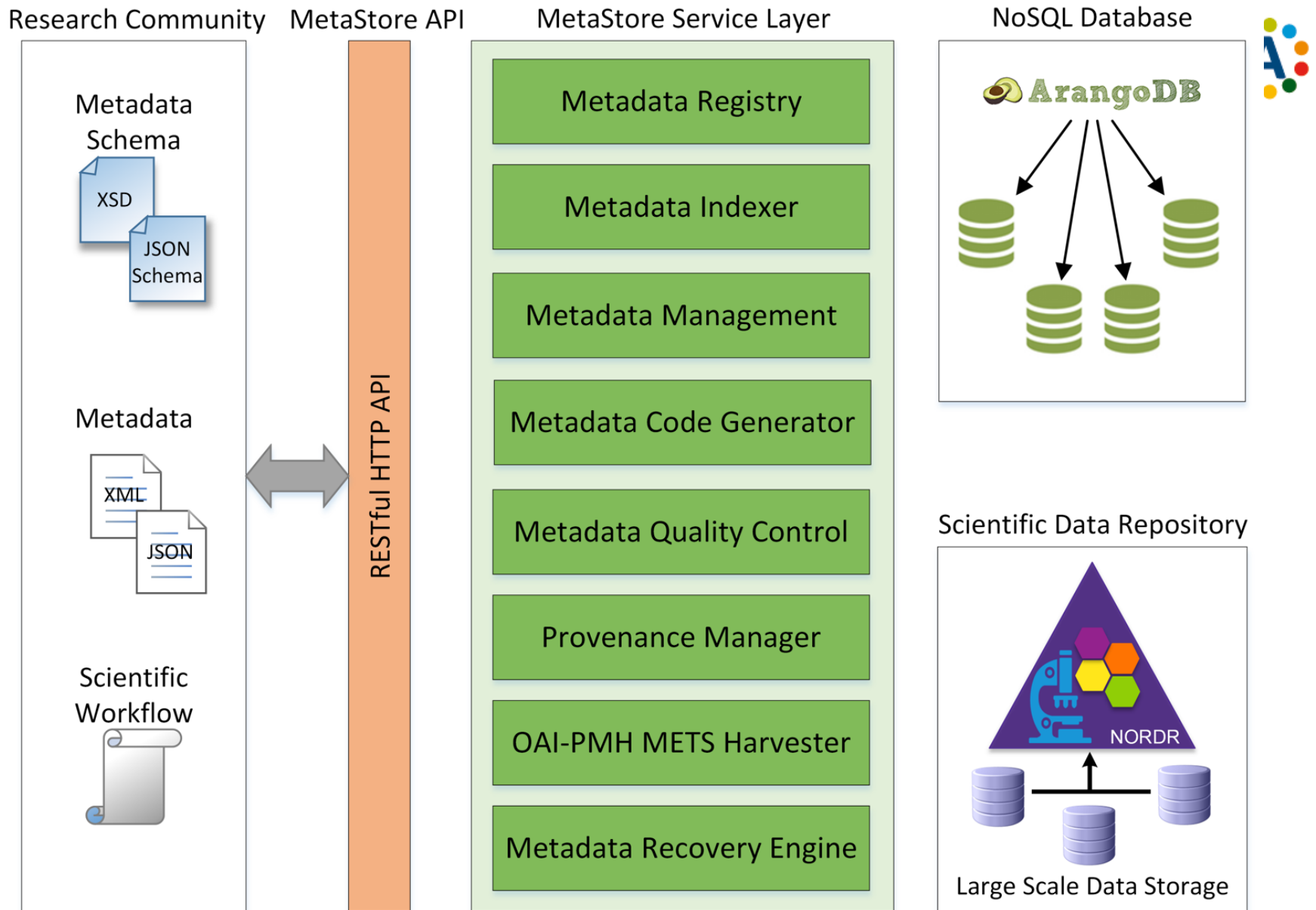
Ingest to repository:

- 4 Automatic metadata extraction: 'exiftool' + transformation to XML**
- 5 Metadata indexing**



4

```
<cvma xmlns=http://www.corpusvitrearum.de/cvma/1.0/>
  <dc:Title>Arche Noah</dc:Title>
  <dc:Type>Glasmalerei</dc:Type>
  <lptc4xmpExt:City>Ulm</lptc4xmpExt:City>
  <xmp:CreateDate>1420-01-01</xmp:CreateDate>
  ...
</cvma>
```

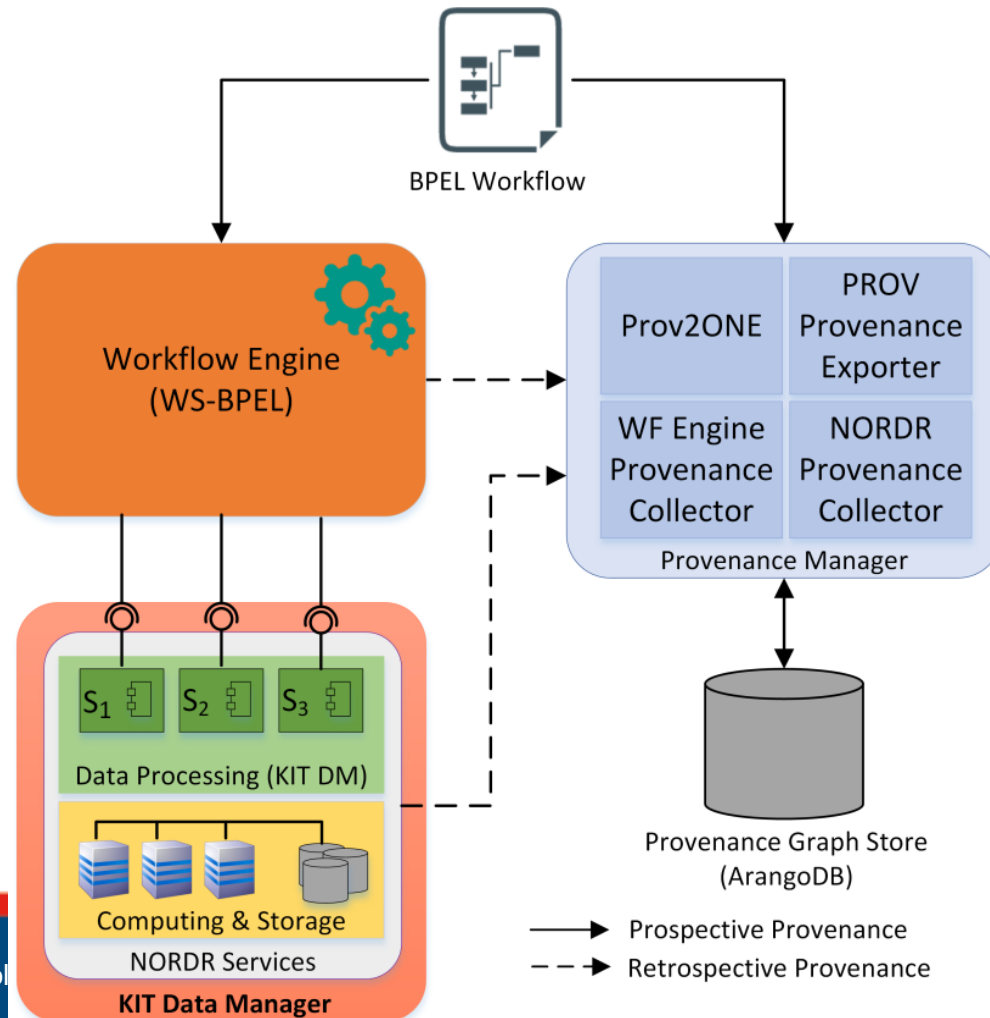
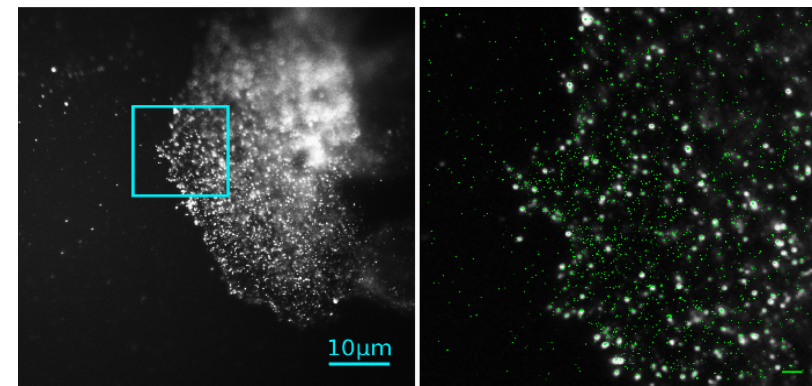


- Metadata schema registry
- Metadata quality check, schema validation
- On the fly generation of services for handling registered Metadata schema
- REST API for OAI-PMH metadata harvesting
- Automated metadata indexer for full-text search
- Single database system for heterogeneous metadata models (provenance, contextual, technical)

NORDR: Scientific Workflows, Provenance and Metadata

Comprehensive data repository for handling the lifecycle of nanoscopy research datasets

- Storage and access of large datasets (~150-200 TB per experiment)
- Integrated workflow engine
- Automated handling of provenance in **ProvONE** (reproducibility)
- Graph DB for storing ProvONE provenance graphs
- Provenance **interoperability** between PREMIS and ProvONE (in review)



DFG Collaborative Research Centre “Episteme in Bewegung”



Sonderforschungsbereich 980
**EPISTEME IN
BEWEGUNG**

Knowledge transfer from the Old World to the early Modern Age

Free University Berlin, 1.7.2016 – 30.6.2020

Guest Project „Manuscripts in Motion”, 2 Years

- Tools for documentation, analysis and visualization of texttopographic data
- Continuation of eCodicology
- Application to Aristoteles Archive data

INF Project „Travelling Books”, 4 Years

- Build-up of information infrastructure
- Research data repository
- Seamless integration of data analysis
- Management of dynamic meta data



Episteme Web Portal



Episteme Tools

Intelligent Search	Visual Analysis	Extraction of Dynamic Meta Data	Tools for Automatic Layout Analysis	Categorization of Relations
--------------------	-----------------	---------------------------------	-------------------------------------	-----------------------------

Episteme Repository

Standardized Interfaces

Model of Dynamic Meta Data	Model of Stable Meta Data
Dynamic Triple Store	Repository of Stable Data

Resources (Server and Data Storage)



Local Episteme Resources

Repository Pilot Projects

De Interpretatione - Aristoteles

- Scans of Aristoteles texts
- Re-use and enhancement of eCodicology image analysis tools



Egyptian Witnesses

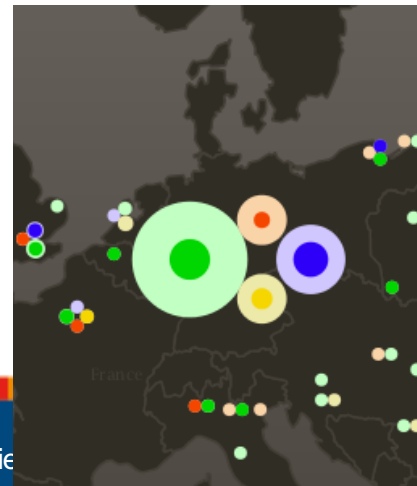
- Transliterations of pyramid texts
- Study of possibilities of tools for automatic text analysis

Siut I:

[223] *ir rmč.t nb.t shj(.w) nb rh-(i)h.t nb nčš nb twš nb*
224 irj.t(i)=šn šf-hrw m is pn
hčj.t(i)=šn shj(.w)=fnšš.t(i)=šn n hn.tiw=f
hr=[š]n n čnt nj Čhwti špł wn im(.i) nčr.w
iw=šn n č.t n.t srš.w(?)⁸²⁶ 225 hm.w bit im(.i)w hw.wt wr.wt

Change of Greek Libraries

- Bibliography of library stocks
- Visualizations of temporal and spatial relations



Conclusions (& LL)

- Repositories for scientific disciplines
→ preservation, description & access
- Sustainability and interoperability
- Metadata workflows and tools → researcher
- Provenance & reproducibility
- Communication (!)
- Seamless integration of analysis workflows