

# DLCL Key Technologies: März 2016

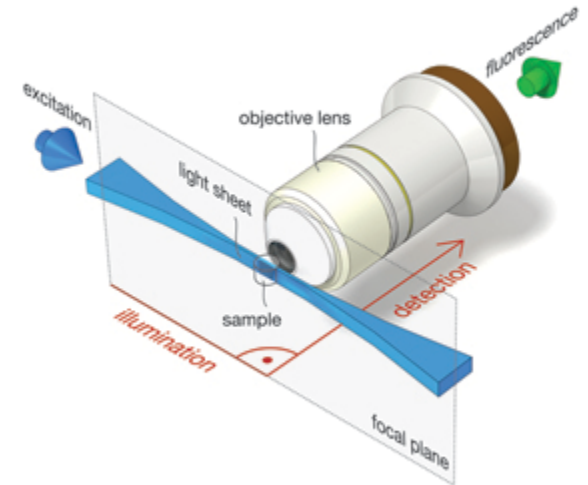
Rainer Stotzka, Swati Chandna, Richard Grunzke, Volker Hartmann, Michael Hausmann, Jürgen Hesser, Thomas Jejkal, Ralph Müller-Pfefferkorn, Michelle Pfeiffer, Francesca Rindone, **Danah Tonne**, **Xiaoli Yang**, Eberhard Schmitt, Margund Bach, Ajinkya Prabhune, Armin Volkmann, Hjalte Raun, Kevin Geggus, Anil Keshav, Hasebullah Ansari



- ***Light Optical Nanoscopy*** (Heidelberg, Mannheim, Mainz)
- ***High Throughput Microscopy:***
  - Selective Plane Illumination Microscope (Karlsruhe)
  - Gen Scans (Dresden: TU + MPI CBG)
- ***ANKA Tomography***  
Ultra Fast Tomography
- ***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)

# High Throughput Microscopy

- Finished: Prototype HPC integration on workflow-level of KNIME via UNICORE data oriented processing – paper submitted
- Finished: Prototyp implementation of fine-granular UNICORE-KNIME integration – paper planned
- Planned: Metadata-enabled microscopy data pipeline in cooperation with MASi project



Build up an NFFA Information and Data Repository

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

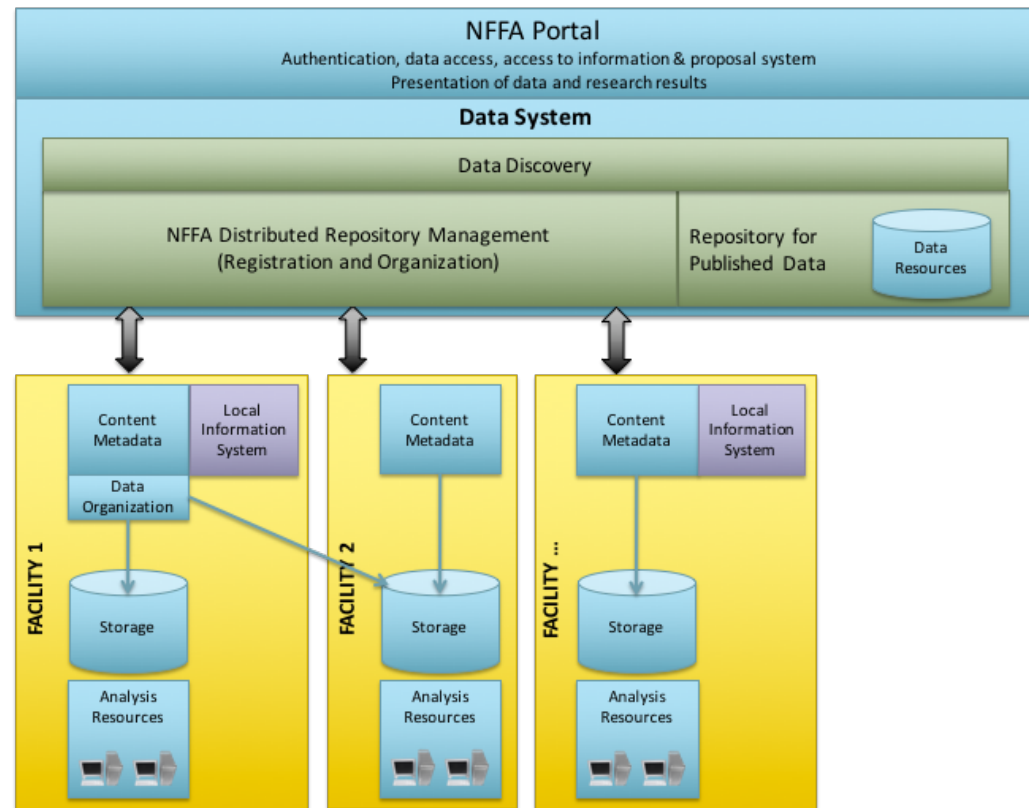
Common AAI supported by all facilities

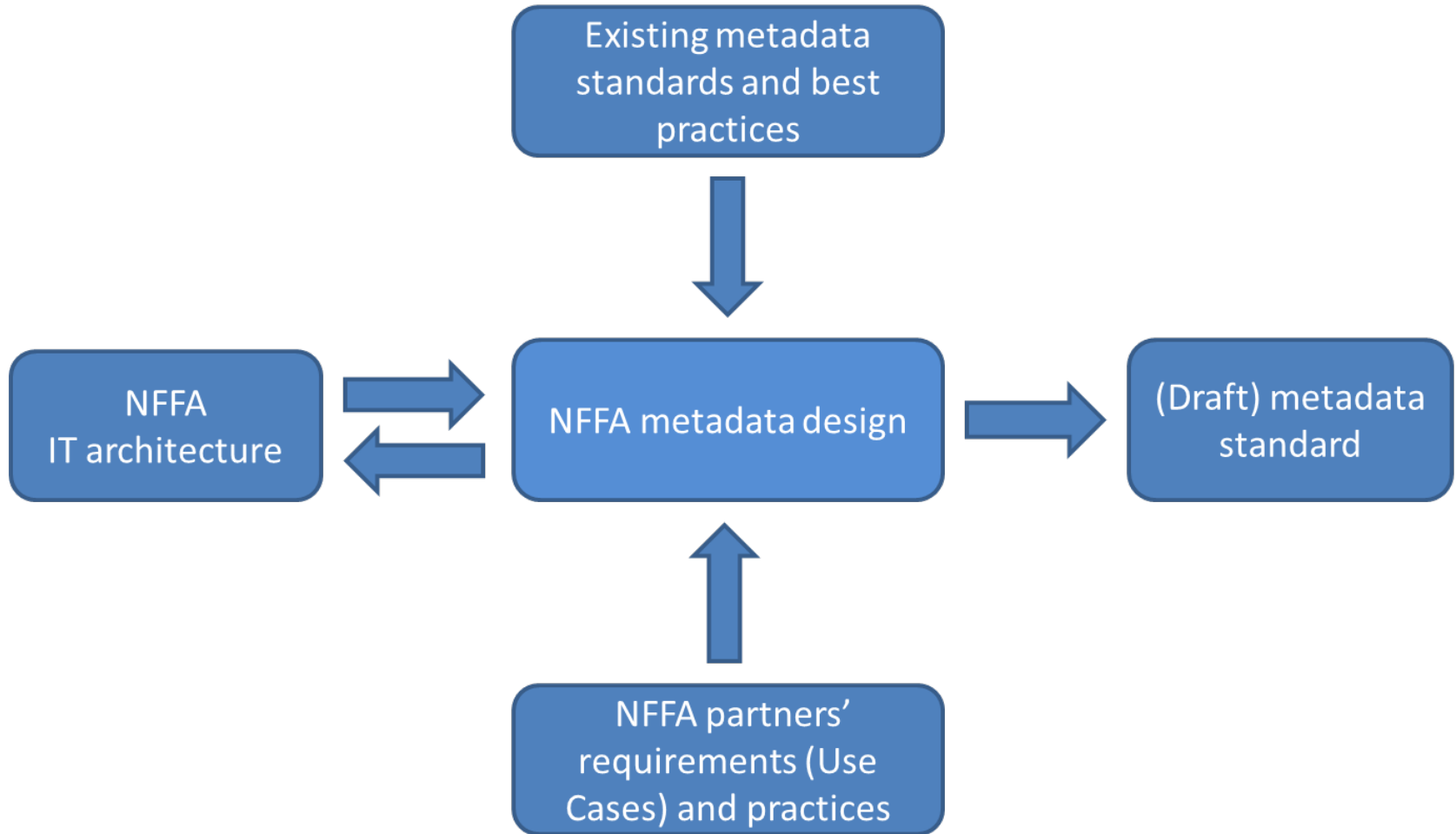
Access via NFFA Web Portal

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

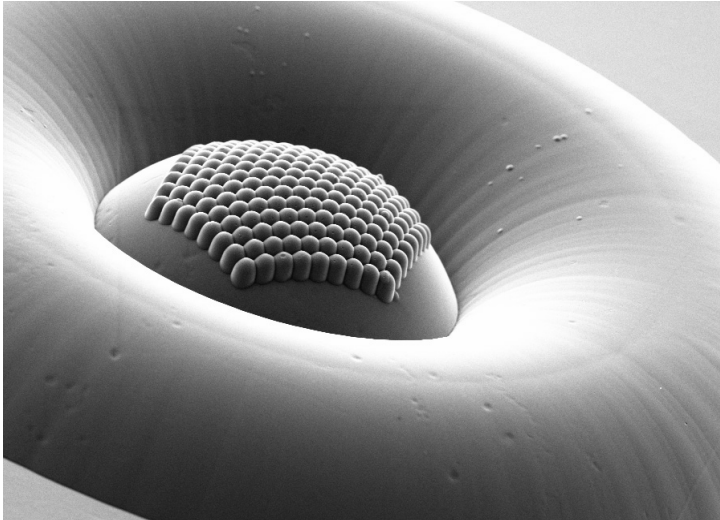
Published data repository for publicly available datasets

- Registered using a nanoscience metadata schema defined in RDA WG









## HOW TO APPLY

TO GET FREE ACCESS

**1**

### **BROWSE & CHOOSE**

Visit [www.nffa.eu](http://www.nffa.eu), browse the offer & select the tools you need

**2**

### **SUBMIT YOUR PROPOSAL**

on our single-entry point

**3**

### **HAVE IT EVALUATED**

& ranked by an international peer-review panel

**4**

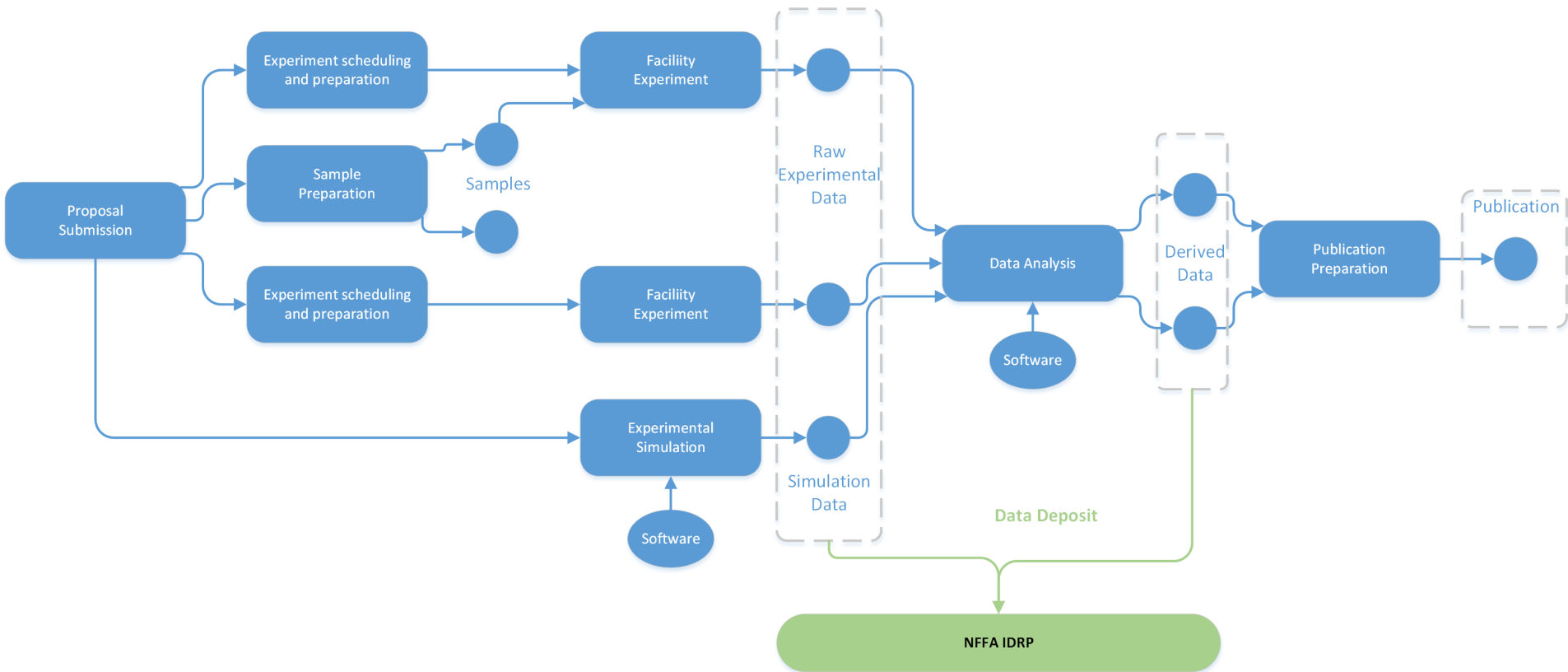
### **GET FREE ACCESS**

and receive a contribution for travel & accommodation costs



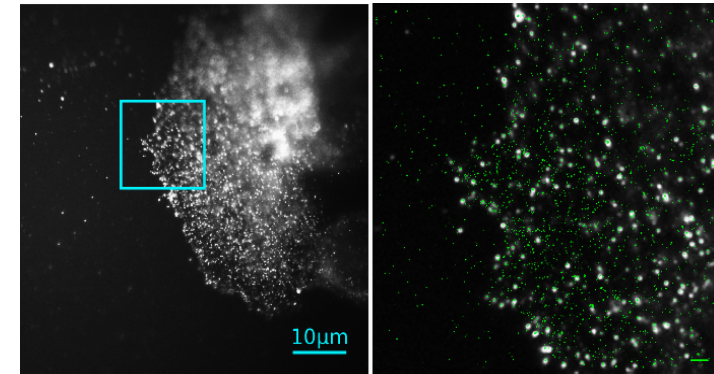
You can also access our comprehensive research information on the first  
**DATA REPOSITORY PLATFORM  
FOR NANOSCIENCE**

# NFFA Europe Experiment Workflow



## Nanoscopy Open Reference Data Repository (NORDR) for managing the complete life-cycle of extremely large datasets (~100 TBs)

- Allow storage, reuse and curation of data
- Manage heterogeneous data and metadata
- Enable scientific workflow and provenance

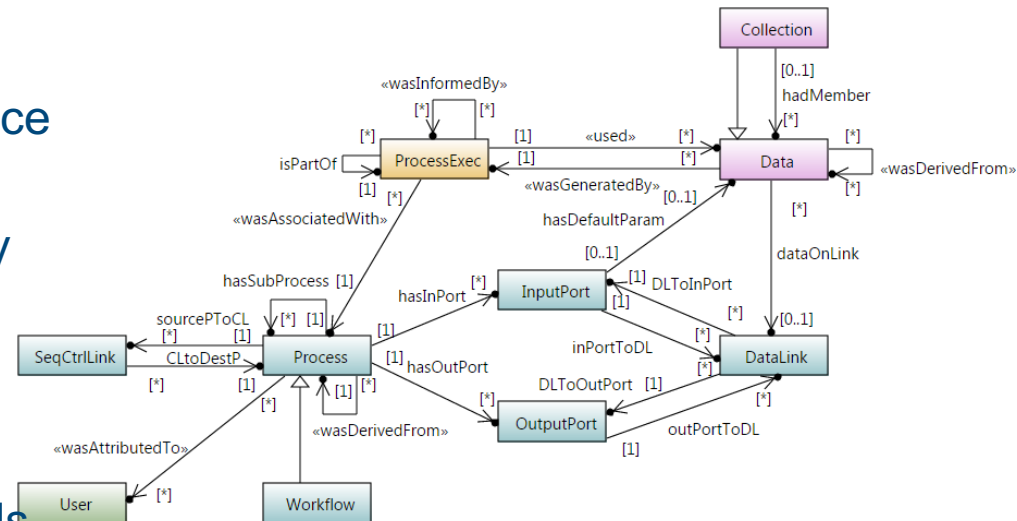


Microscope image of breast cancer cell

Localization image of breast cancer cell

## ProvONE Provenance Model

- Graph based modeling of provenance (ProvONE)
- Enable scientific data reproducibility
- Capable of capturing both
  - Prospective provenance
  - Retrospective provenance
- Interoperable with existing standards OPM/PROV



ProvONE Provenance Model



# NORDR: Scientific Workflows, Provenance and Metadata

## Scientific Workflow(WF) Engine

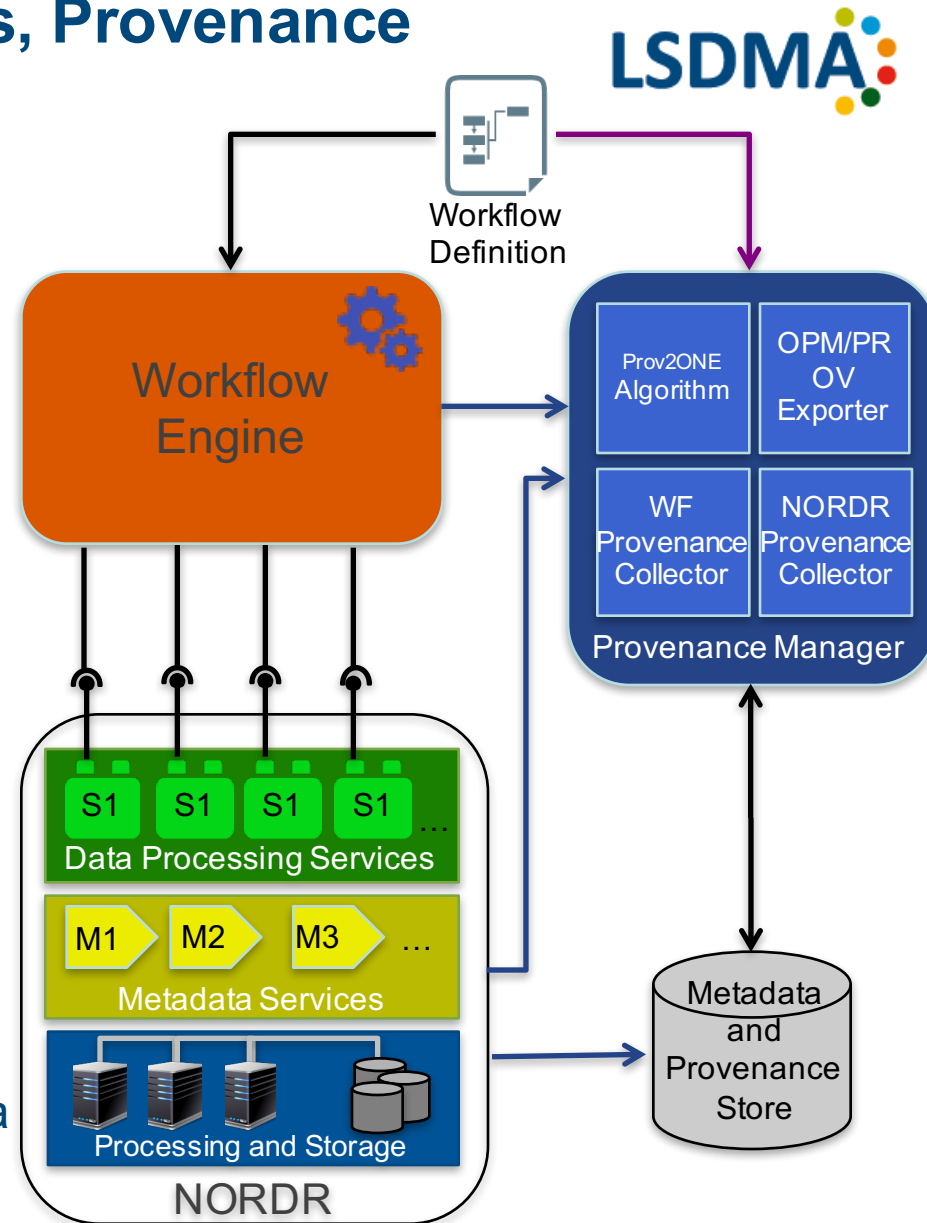
- Automated execution of workflows
- Integrated with data processing services of NORDR

## Provenance Manager

- Prov2ONE algorithm for creating provenance graphs (ProvONE)
- Provenance collector, WF engine and NORDR
- Export retrospective provenance in OPM and PROV standards

## Metadata and Provenance Store

- Common data store for context metadata as well as provenance graphs



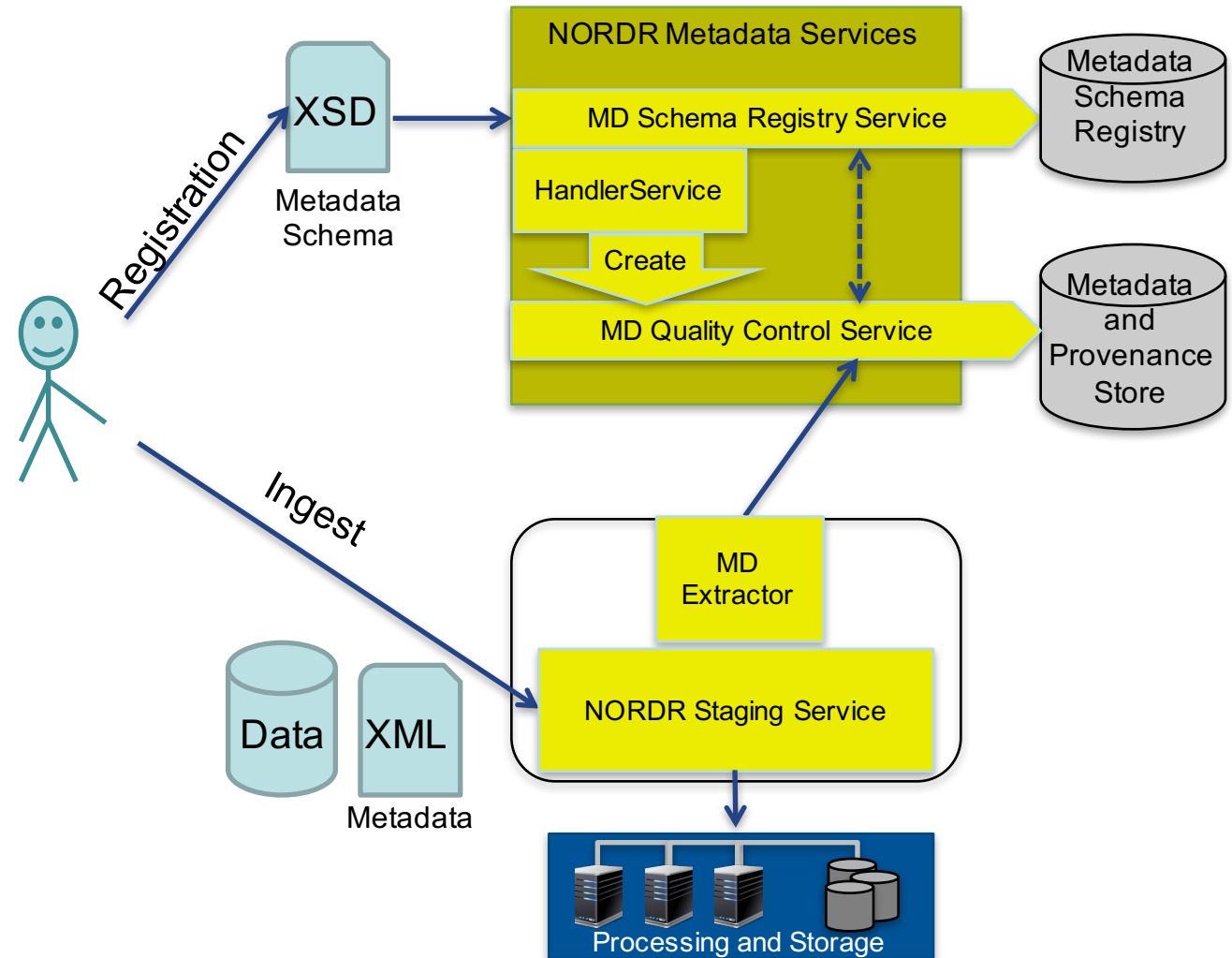
# NORDR Data Ingest & MD Quality Control

Registration of new MD schemas by user

Automatic creation of MD Quality Control Service

Ingest of data and MD

Automatic MD quality control



# Conclusions

- Joint research on solving community specific data problems
- R&D on generic data technologies (bottom-up)

