

Update from HFZ meeting and NFFA metadata

Vasily Bunakov and Brian Matthews
Science and Technology Facilities Council, UK

RDA 11th Plenary joint meeting:
IG RDA/CODATA Materials Data, Infrastructure & Interoperability,
WG International Materials Resource Registries,
IG Research data needs of the Photon and Neutron Science community
Berlin, 23 March 2018

File Edit View History Bookmarks Tools Help

Scientific Data Management for X

https://www.helmholtz-berlin.de/event

Search

HZB Home Imprint Data protection

HZB Helmholtz Zentrum Berlin

- Program
- Participant List
- Registration

Scientific Data Management for Photon and Neutron Facilities

– Associated workshop to the [RDA 11th Plenary Meeting](#) –


Berlin, 19th and 20th March 2018

Users of photon and neutron sources tend to move between facilities and instruments, using and combining different techniques, together with simulation data, to advance their science. Furthermore, as open science becomes a standard, it will become more common for scientists to use data not collected by themselves in their analysis or to compare their own findings with other peoples data.

In order to be able to combine data from different sources or to make use of third party data, it is a requirement to complement the data with metadata at a sufficient level of detail to understand the

Local organizer:

Rolf Krahl



(030) 8062-12122

[Email](#)

[Business card](#)

Topics under discussion

- Data/Metadata catalogues and data sharing
 - Existing data catalogues
 - Existing and emerging data management systems: ICAT, Materials Data Facility by NIST, SciCat by PSI & ESS
- Data representation for facilities data
 - Evolution of the NeXuS format (based on HDF5)
 - How to best represent different metadata definitions
- Persistent Identifiers
 - Instruments and equipment – participation in the new RDA group
 - FREYA project
- Relationship with Materials projects and activities
 - Materials Genome Initiative
 - RDA (IMRR International Materials Resource Registries) WG
 - NFFA-Europe

NFFA project

- Horizon 2020 project
- Distributed infrastructure to perform growth, nano-lithography, nano-characterization, theory and simulation and fine-analysis with synchrotron, FEL and neutron radiation sources
- “Virtual research enterprise” with proposals system and data management obligation



NFFA approach to metadata design

- Overall focus: data lifecycle in nano-facilities
- To satisfy both experimental and computational research
- Top-down route: consider existing data management recommendations and best practices
- Bottom-up route: consider NFFA partners Use Cases and practices
- “Neighbouring” route: consider NFFA IT Architecture
- Synthesis: amalgamate top-down and bottom-up with some “side” (IT architecture) considerations

Metadata artefacts and a potential for their common acceptance

- Common Vocabulary for entities
- Entity-Relationship diagram
- Lists of attributes for entities (with data type and cardinality)
- Particular metadata serialization



Relatively easy to agree upon



Variations inevitable

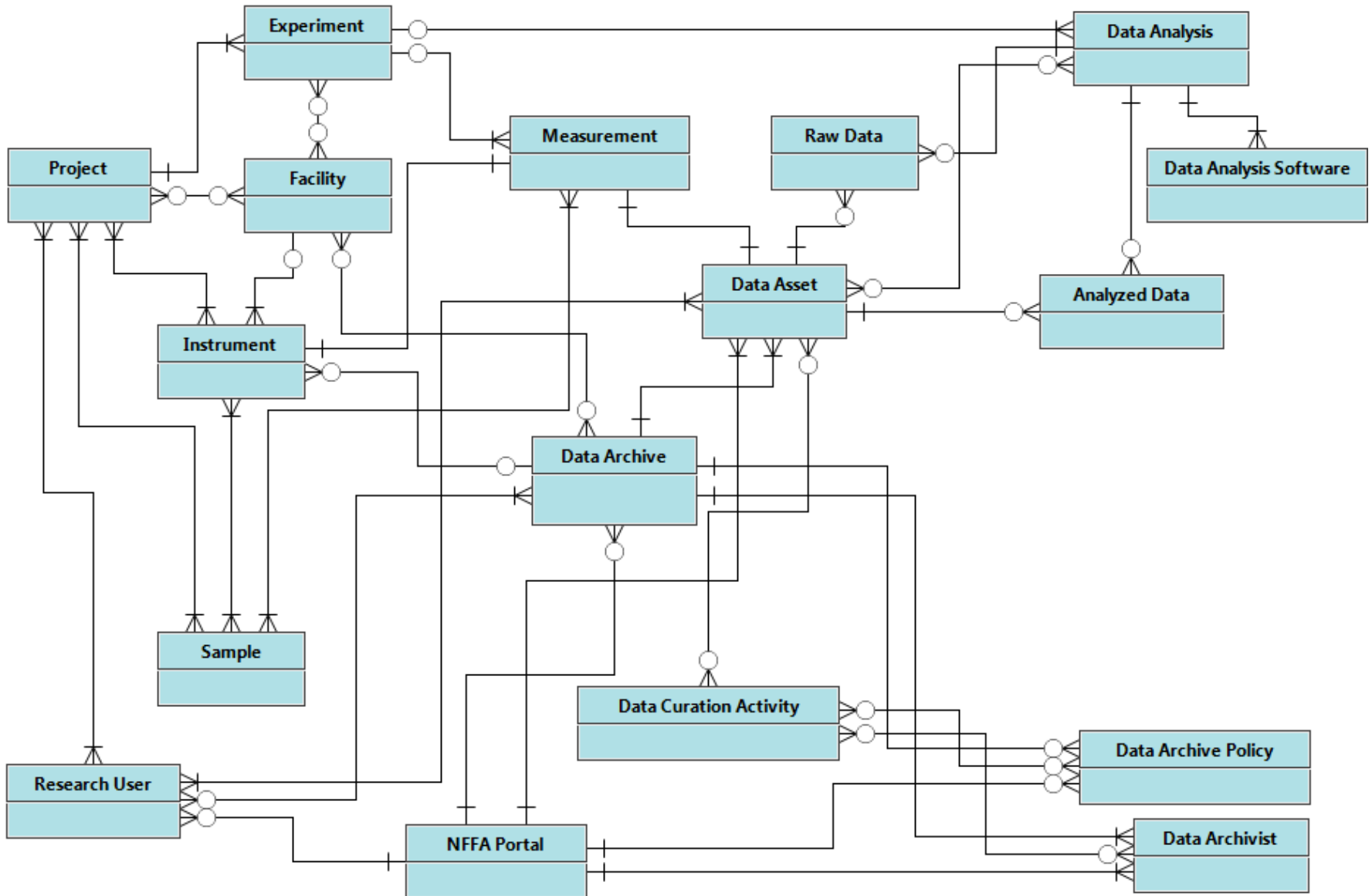


Can be specific to software platform

(A fragment of) the Common Vocabulary

- **Research User.** A person, a group of them, or an institution (organization) who conduct Experiment on a nanoscience Facility using a nanoscience Instrument in order to collect and analyze Raw Data, or is interested in data collected or analyzed by other Research Users on the same or other Facilities.
- **Project.** An activity, or a series of activities performed by one or more Research Users on one or more Facilities using one or more Instruments for taking one or more Measurements of one or more Samples during one or more Experiments. Facility, Instrument, Measurement and Sample can refer to computer simulation environment.
- **Facility.** An institution (organization), or a division of it that operates one or more nanoscience Instruments for Research Users. For computer simulation, Facility can be a software platform that allows to order and manage computational experiments (so that the software platform serves the purpose of managing software modules that can be considered virtual Instruments).
- **Instrument.** Identifiable equipment (such as a device or a stand or a line) that allows conducting an independent nanoscience research, perhaps without involvement of other Instruments. Instrument is hosted by Facility and used by Research User. Instrument produces Raw Data in the course of Experiment. Instrument can be in fact a software for computer simulation (a software module or/and a particular configuration of it).

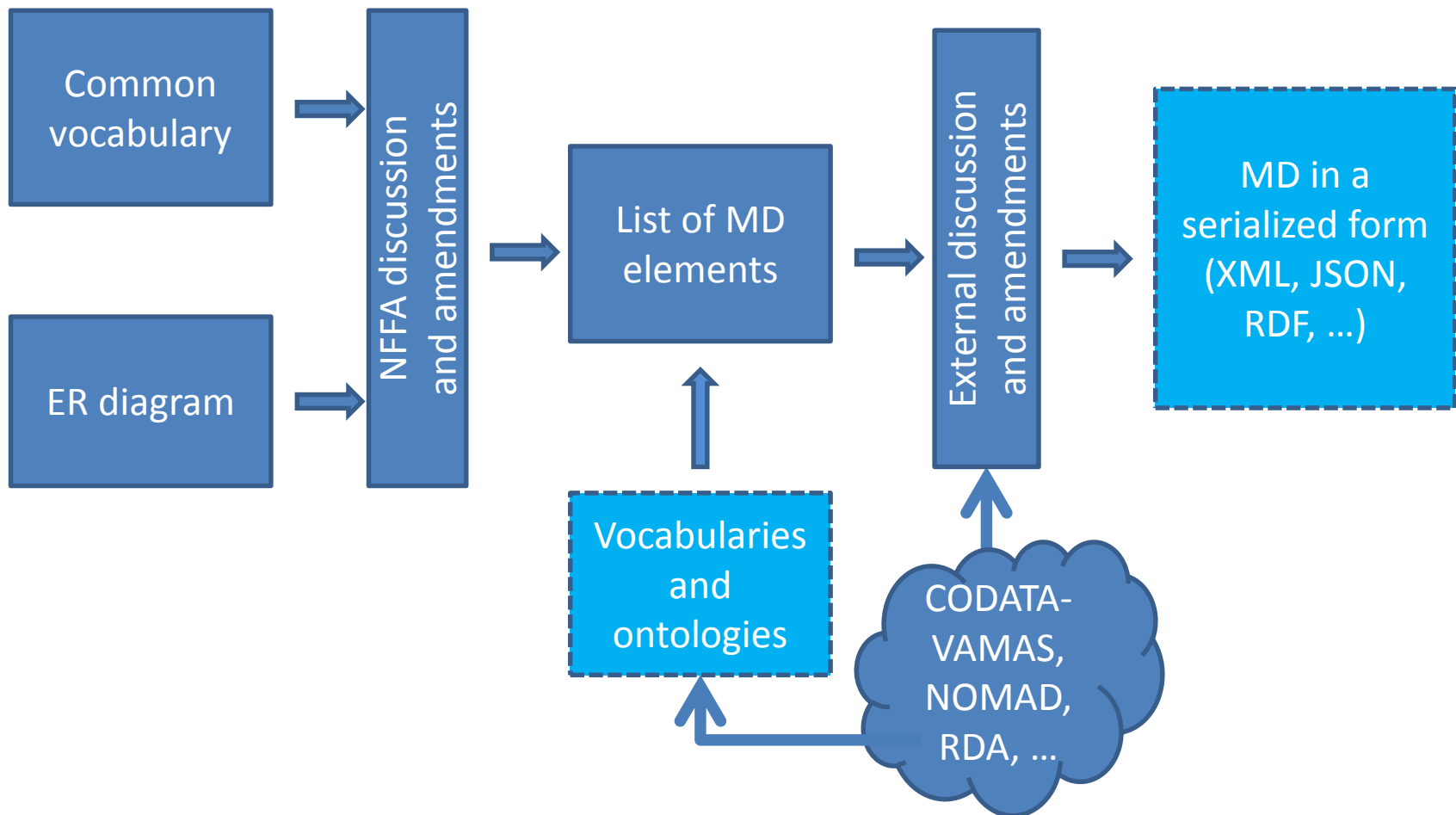
ER diagram



Cross-walks across NFFA, CODATA-VAMAS and NOMAD metadata models

NFFA concept	CODATA-VAMAS concept	NOMAD concept
Experiment	Nano-object production steps	Series of s/w runs
Measurement	Nano-object testing steps	S/w run
Sample	Nano-object or collection of objects	Input data
Data Asset		Output data

Stages of metadata design ("the process")



Cases for NFFA metadata use and collaboration with other projects and organizations

- Can be applied in any facility research (not necessarily nanoscience) that is interested in having a structured description of data lifecycle
- Can be applied to physical or / and computational experiments, and include data analysis phase Ω
- Can facilitate data exchange and interoperability of software platforms: KIT Data Manager ¹⁾, EUDAT services ²⁾, ICAT ³⁾, ...
- Can be used with controlled vocabularies and ontologies
- Can offer sensible placeholders for more detailed modelling

1) <http://datamanager.kit.edu/>

2) <https://eudat.eu/services>

3) <https://icatproject.org/>

Ω The most advanced (and most promising) case

Further reading & Contacts

NFFA project deliverables (available upon request)

Conference papers:

"Metadata for Experiments in Nanoscience Foundries"

<https://zenodo.org/record/1175958>

"Metadata for nanotechnology: interoperability aspects"

<https://zenodo.org/record/1175964>

NFFA project: <http://nffa.eu>

Vasily Bunakov vasily.bunakov@stfc.ac.uk

Brian Matthews brian.matthews@stfc.ac.uk

Thank you!

