e-infrastructure Requirements and Needs for Nano foundries and Fine Analysis (NFFA) scientific community

Stefano Cozzini CNR/IOM Trieste, Italy



Agenda

- Introducing the NFFA project
- The NFFA DATA repository efforts
 - The identified needs
 - The first prototype
 - Some week points
- Conclusions

Role of different e-infrastructure

The NFFA project (www.nffa.eu)

- A EU-funded design study for distribute research infrastructure in nanoscience.
- Leaded by CNR/IOM (Trieste) it involves several large scale facilities
- Involves many scientists and users, mainly experimentalist but NOT only



NFFA

Nanoscience Foundries and Fine Analysis Giorgio Rossi

Project coordinator (http://www.nffa.eu)





Research Infrastructures: Nanoscience Foundries and Fine Analysis a Design Study for a European Research Infrastructure enabling access to nanoscience, atomic precision and fine analysis with a unified metrology.



THE NFFA MISSION (www.NFFA.eu)

A DISTRIBUTED INFRASTRUCTURE LINKED TO ANALYTICAL LSFs

The NFFA Design Study supports the construction and operation of an ERIC consisting of Nanoscale Science Research Centers at European sites that already host Large Scale Facilities for Fine Analysis of matter.







Elettra, FERMI@Elettra and IOM-TASC



Data Repository

- NFFA has addressed the creation of the first Data Repository (DR) in nanoscience.
- Goal:
 - Store all the data produced in the NFFA centres
 - make them accessible by means of appropriate search tools (registered syntax and semantics)

Not only data...

- The DR needs to includes the relevant information
 - for data analysis
 - for the full reproducibility of preparations and experiments

The DATA REPOSITORY should enable the exploitation of truly complementary data as obtained by different methods on true replica samples and environment conditions

Data Repository Requirements

- Less invasive as possible for users: maximize automatic input from instruments and to supply smart and well scheduled forms whenever user is requested to operate
- Useful for external references (in particular for industries and technological districts): smart and easy access, keyword management, data readability, view sorting and priorities
- Useful for NFFA management: user access, technical liaison, reviewing, proposal support, IPR issues
 Useful for the "common metrology & standard"
- Useful for the "common metrology & standard" implementation in comparing data and protocols and in calibrating instruments
- Easy to connect to other e-infrastructures



Guidelines for NFFA- DR

- The data reside in file archives. The metadata are managed using a database
- Metadata management and search criteria will be performed by using both keywords and semantic search
- Extensive support for metadata management and data processing
- A flexible architecture that can be adapted to multiple scenarios (transparently extendible to other e-infrastructure)

The NFFA demonstator

- A prototype implemented in the design study
- Developed in collaboration with Elettra
- Based o european Grid infrastructure middleware tools
- Currently under revision for a further expansion



Demonstrator initial requirements

- Leverage the existing tools and middleware
- Concentrate on concepts, overall workflow and leave details for future developments
- Single sign on to all the resources
- Integrated components that can easily scale
- Follow other standardization activities (VEDAC, PaNDATA)
- Metatada and data curation by means of ICAT tools and possibly automatized via the e-Infrastructure approach.

Experimental facilities linked

- The demonstrator collects data from:
 - a scanning electron microscope (SEM) instrument,
 - a synchrotron radiation beamline for spectroscopy
 - an open package for first principle quantum simulation (Quantum Espresso)





Home Tunnel Application Manager	anoscience fo nd fine analys Applications Workflow	oundrie Sis Managen	Welcome , Stefano Cozzini Profile Home Logout	
Available Applications			plication Monitor	
Refresh List			2011-01-17-15_57_42_open-dataportalstep0	
Application		Open		
submit-proposal Open check-my-proposal Open		Open	Tunnel opened to nffa.grid.elettra.trieste.it Next> Clear Application	
		Open		
open-dataportal Open		Open	L	
start-investigation Open		Open		
quantum-expresso Open		Open		
submit-dataset Open		Open		
browse-nffa-storage		Open		
Active application for use Refresh active applications	er cozzini		ACTIVE	
Application	Date	View		
submit-dataset	2011-01-17-12_46_39	View		

List of application:

- Data entry applications for proposal submission and check status
- An application integrating the ICAT data portal
- An Applications to inject investigation for the integrated instruments
- Application to inject datasets acquired by the integrated instruments
- An application to browse the NFFA-storage
- An application integrating
 Quantum Espresso package

Available Applications				
Refresh List				
Application	Open			
submit-proposal	Open			
check-my-proposal	Open			
open-dataportal	Open			
start-investigation	Open			
quantum-expresso	Open			
submit-dataset	Open			
browse-nffa-storage	Open			

Advantages

- Data repository integrated in existing einfrastructure
- Data repository follows standard (ICAT)
- Data repository widely available (it requires just a browser and a network connection)
- Data repository does not change too much the user approach
 - Use is simple and not so complex

Weak points

- A two steps procedure to insert data:
 - A. load data on the nffa-storage
 - B. submit metadata to the ICAT tool
- Lack of automatic procedures
- ICAT likely not the right tool:
 - lack of flexibility
 - not the right tool for all the needs
 - lack of semantic support
- HPC infrastructure link missing

What next ?

- Integrate/bring the NFFA community and the needs within the European DATA initiatives
 - EUDAT and others.
 - Leverage on the experience already done
- Coordination with e-infrastructure initiative (EGI and PRACE)
 - Examples:
 - integration of the nano-simulation techniques as standard "experimental" tool requires HPC infrastructure
 - Available and already developed data services can be evaluated and used

Conclusions

- Data repository needs for NFFA identified
- Prototype built with the scope of validating the approach by user communities
- The prototype deployment and early users usage identified strong and weak features
- a new enhanced version of the prototype is under development
- Coordination with other communities and einfrastructure services (data and others) is to take into account