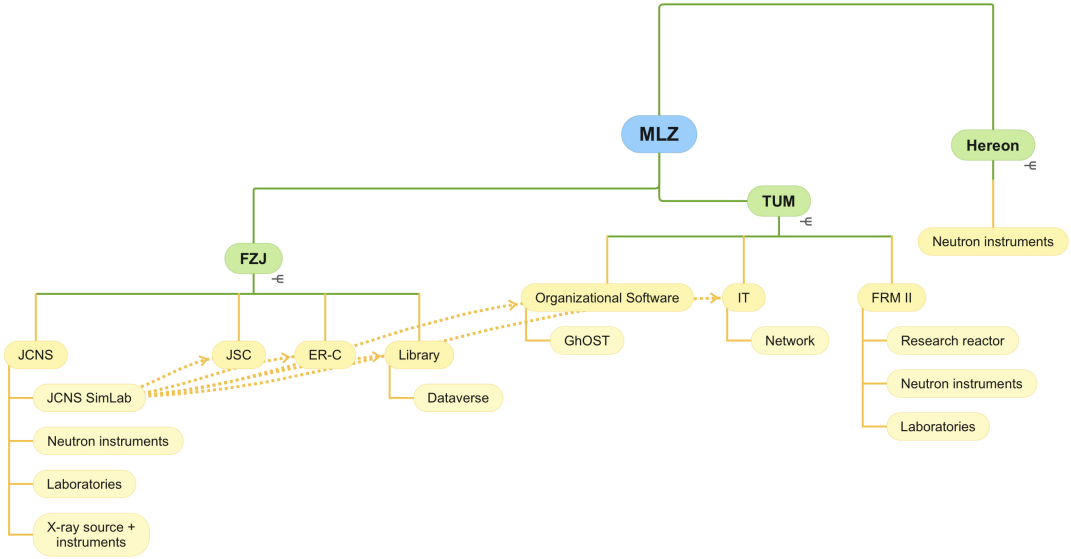
FZJ QUESTIONNAIRE PRESENTATION

DMA ST1 Synergy preparatory workshop

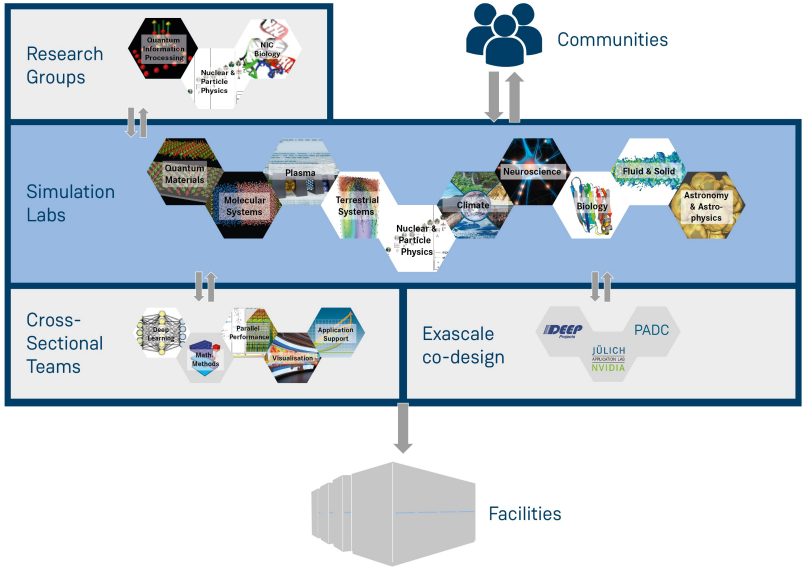
June 14, 2023 | C. Felder, G. Brandl, M. Ganeva | JCNS

MLZ Structure



SimLab Introduction

Simulation and Data Laboratories at Forschungszentrum Jülich



- have been founded in response to the challenges posed by Petascale and Exascale computer architectures
- assist user groups in enhancing performance efficiency and scalability of domain-specific computational and data-intensive applications

JCNS Neutron SimLab: Context

- We work in a close cooperation to scientists in Jülich and Garching
- Our groups support internal and external facility users
- We cooperate with other institutes of Forschungszentrum Jülich
- We participate in national and European consortia

JCNS Neutron SimLab groups

Groups in Jülich

- Instrument Technology (K. Lieutnant, Vitess)
- JCNS TA (J. Heinen)
- JCNS-1,2,4 Simulation and Data Analysis projects

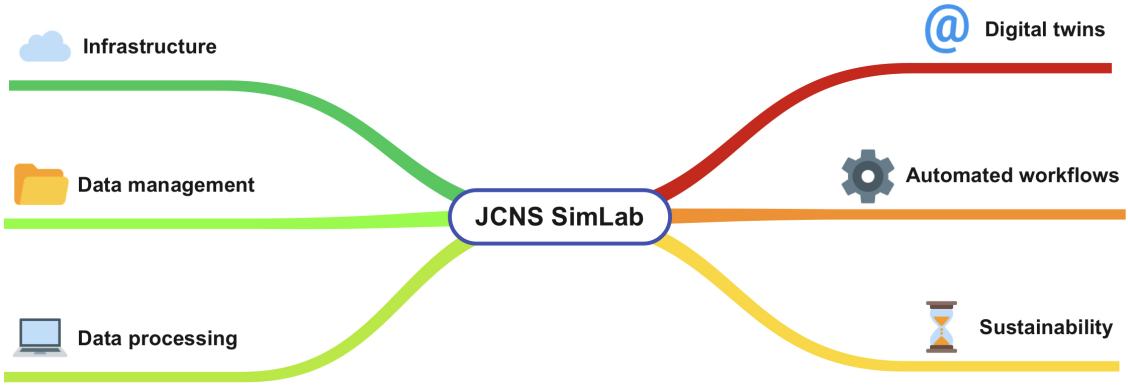
Groups in Garching

- Data management
- Cloud Infrastructure
- Data-driven Discovery
- Instrument control
- Scientific Computing

+ MD Simulations

JÜLICH Forschungszentrum

SimLab construction sites



+ data understanding

JCNS Neutron SimLab is a dynamic structure

Infrastructure

Cloud services: C. Felder et. al.

- Access to metadata catalogs, electronic laboratory books
- Web-based remote access- and online training systems
- Data Analysis as a Service: reproducible data analysis environments
- For developers: virtual development environments, git, CI/CD

Hardware: C. Felder, J. Heinen, E. Westphal et. al.

- Kubernetes Cluster (Garching)
- Data storage hardware
- Computational servers/clusters

IT services

Garching: Paulo Innocente

Jülich: JCNS Scientific IT systems, Josef Heinen (Markus Consoir)

Data management

SampleDB (F. Rhiem) as metadata catalog is implemented at following instruments:

- (Cryo-)Ultramicrotome Leica UC7
- Apreo-Volumescope
- D8
- Ganesha SAXS
- Huber 4 Circle Diffractometer
- JEOL JEM-F200
- MBE-MLZ
- MBE-MLZ AES
- MBE-MLZ LEED
- MBE-MLZ RHEED
- Mirror Furnace
- MPMS SQUID
- OMBE I
- OMBE I AES
- OMBE I LEED
- OMBE I RHEED
- Powder Diffractometer
- PPMS
- PPMS Dynacool
- Single Crystal Diffractometer Supernova
- SynergyS-MLZ Single Crystal Diffractometer

See `https://iffsamples.fz-juelich.de/instruments/` for more details.

Work in progress

GALAXI, Leica SP-8, JEOL JSM-6510LV

Data management

JCNS at MLZ in cooperation with DAPHNE (A. Schneidewind)

NICOS – Data Uploader	Instrument Control Groups (MLZ)	Cloud Infrastructure (JCNS)
Buffer & Stream	Instrument Control Groups (MLZ)	Cloud Infrastructure (JCNS)
Ingestion Deamons	Cloud Infrastructure (JCNS)	Instruments (TUM)
Applications	Cloud Infrastructure (JCNS)	Instruments (TUM)

Instrument Control Groups
G. Brandl
K. Kholostov
S. Rainow
E. Faulhaber
J. Krüger

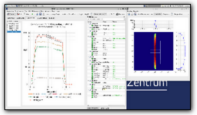
Cloud Infrastructure
M. Hannemann
O. Mwongera
C. Felder

Instruments
J. Baudisch
M. Ouaki
B. Pedersen

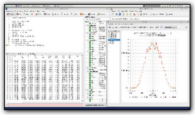
More DAPHNE@MLZ
Y. Tymoshenko (KIT)
C. Lobo Silva (JCNS)
N. Zec (hereon)
W. Lohstroh (TUM)
S. Busch (hereon)
A. Schneidewind (JCNS)

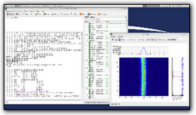
Digital twins

Groups of G. Brandl and C. Felder



TREFF


PANDA


STRESS-SPEC


KWS

- Electronic clones of existing instruments
- Implemented with a combination of NICOS and Monte-Carlo simulation
- Goal: emulate realistic behavior of the real-world instrument within identical user interface
- Application areas:
  - learning tool for new scientists (used in Praktikum!)
  - measurement planning tool for users
  - experiment planing: validation of expectations regarding possible instrument parameters, flux, etc.
- Currently implemented for 4 instruments

McStas

VITESS

Automated workflows

G. Brandl, A. Zaft, SECoP: Sample Environment Communication Protocol

Goal: standardize control of sample environment across facilities for easy transfer of equipment between institutes.

G. Brandl, K. Kholostov: Robotics

Several projects in different states.
Aim: develop a standard toolkit of suitable models for different use cases.

Data-driven discovery group: AI-assisted approaches

- Autonomous TAS experiments (M. Parente et. al.)
- SAS/GISAS/GIWAS/Reflectivity/Spectroscopy data analysis
- Addressing ill-posed inverse problems

Data processing

Data reduction

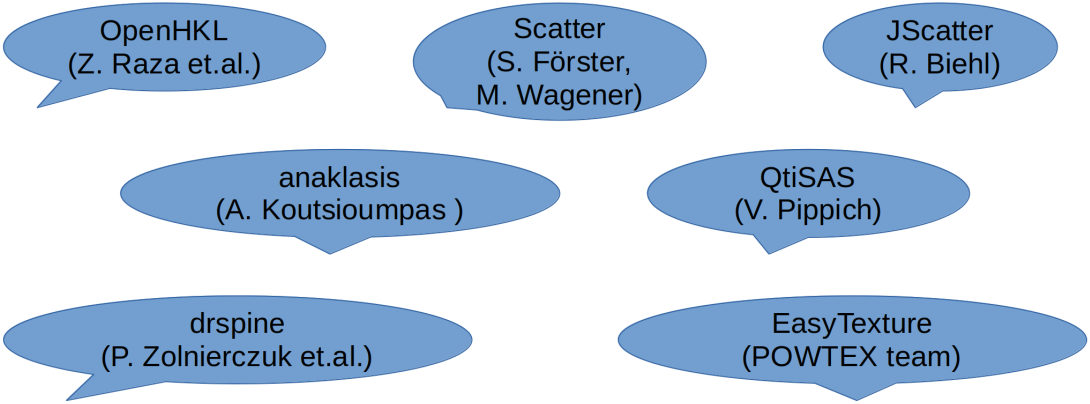
- Data-driven discovery group: Mantid for MLZ instruments (DNS, TOPAS, SANS-1, TOFTOF, POWTEX), EasyTexture
- Scientific Computing group: Steca, Frida, various libraries

Data analysis

- Data-driven discovery group: AI-assisted approaches, VIPR framework (just starting)
- Scientific Computing group: BornAgain, various libraries, EvalSpec software (maintenance after end of the project)

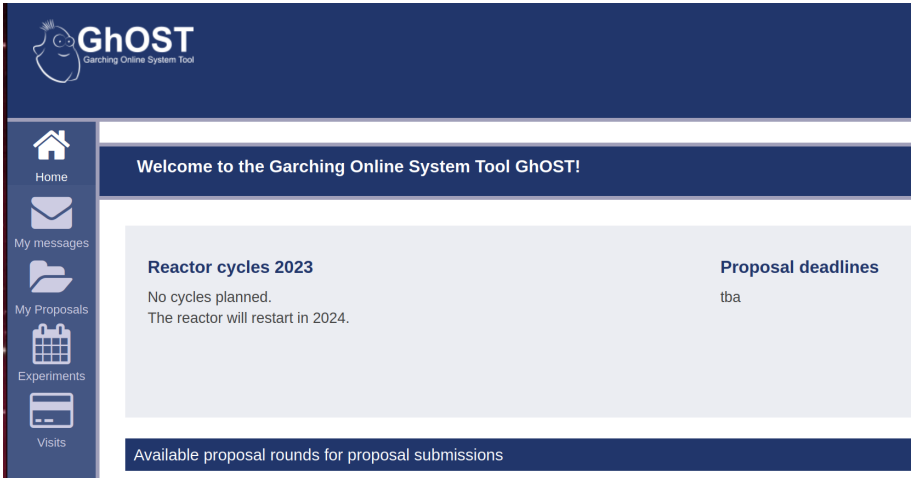
Data processing

Further software projects



Proposal management

Garching Online System Tool: ghost.mlz-garching.de



Neutron instruments only. Managed by TUM Organizational Software.

Data taking / detector

- Tango (device integration), Quango and Marche (N, PS)
- Unified Instrument control software: NICOS (N, PS)
- AI steered DAQ: ariane (integrates with NICOS) (N)
- data ingest: NICOS writes to NFS (N, PS)
- file formats: HDF5, NEXUS (N), TIFF, different ASCII formats (NICOS-Scanfile), compressed Numpy text, YAML (N, PS)
- Classical mode ‘first store, then process data’ (N, PS) Event/streaming mode to come (N)

Competences: Instrument control (TUM and FZJ/JCNS SimLab)

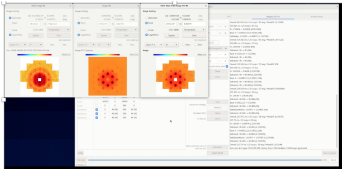
Start-To-End-Simulation

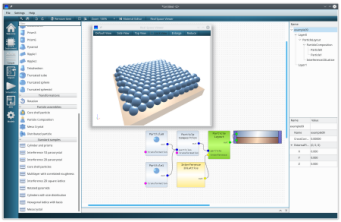
Tools:

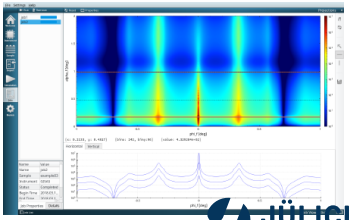
- Driven by scientific community
- Different simulation tools for different steps, depending on purpose and use cases
- N: tools for digital twins of instruments, e.g. McStas, VITESS

Competences:

- Content. SimLab: scientists, scientific computing group; community
- Infrastructure. SimLab: Cloud Infrastructure group, scientific IT systems







Online processing and online data reduction

- Driven by algorithms and software from the user community.
- Infrastructure operated by Cloud Infrastructure Group and Instrument-Control
- Data reduction: Instrument-specific software (N, PS) and community tools (Mantid, scipp) (N)
- Online event (pre-)processing: own development (N)
- Moving towards on the fly processing (no need for real time)

Competences: Scientific community and SimLab groups

Data storage

Tools:

- NFS, (SMB)
- Catalogues: SciCat, SampleDB

Infrastructure:

- Experiments: TUM/IT
- Central: Cloud Infrastructure group and Scientific IT systems

Catalogues:

- SimLab groups: Scientific IT systems, I-Control (FZJ/JCNS), Cloud Infrastructure group
- DAPHNE
- TUM groups: I-Control (TUM), TUM Organizational Software

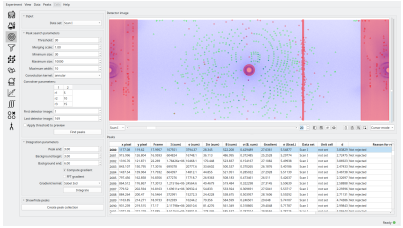
Offline data analysis

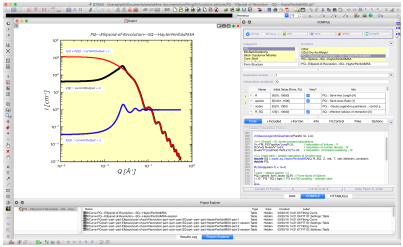
Tools

- Software: Community specific analysis software, FZJ Open source projects
- DAaaS: work in progress
- Infrastructure: HPC cluster (FZJ) SLURM, Jupyter, SciCat

Competences:

- SimLab scientists,
- SimLab groups: Scientific Computing group, I-Control (FZJ/JCNS), Cloud Infrastructure group, Scientific IT systems





FAIR data handling, publication and archiving

Tools:

- Metadata: SciCat (MLZ) and SampleDB (JCNS/FZJ)
- Archiving: LRZ Tape Archive, FZJ Tape Archive
- Publication: iMPULSE (MLZ) and JuSER (FZJ), dataverse (FZJ)

Competences:

- SimLab scientists
- DAPHNE at MLZ
- SimLab groups: Cloud Infrastructure group, Scientific IT systems

Step-overarching

Metadata handling and ELN

Tools:

- Data management and remote access: currently supervised access using TeamViewer and Apache Guacamole; unsupervised access (work in progress)
- SciCat
- Workbench (MLZ)

Competences:

- SimLab scientists
- DAPHNE at MLZ
- TUM Organizational Software
- SimLab groups: I-Control (FZJ/JCNS), Cloud Infrastructure group, Scientific IT systems

Step-overarching

Used data formats, Data flow and automatisation

Tools:

- Workflow engines based on e.g. CWL under investigation
- Efficient data formats for event mode under investigation (N)

Competences:

- Simlab/Cloud Infrastructure group
- DAPHNE at MLZ

Thank you four your attention!