DLCL Key Technologies: Status + Future

Rainer Stotzka, Richard Grunzke, Volker Hartmann, Michael Hausmann, Jürgen Hesser, Thomas Jejkal, Nick Kepper, Ralph Müller-Pfefferkorn, Halil Pasic, Rasmus Schröder, Danah Tonne, Xiaoli Yang
4 (+EU) locations

5 (+1) subprojects:

- Lightoptical Nanoscopy (Heidelberg, Mannheim)
- High Throughput Microscopy: Gen Scans (TU Dresden + MPI of Molecular Cell Biology and Genetics)
- High Throughput Microscopy: Selective Plane Illumination Microscope (Karlsruhe)
- Electron Microscopy (Heidelberg)
- Material Research: Tomography Beam Line at ANKA (Karlsruhe)
- ESFRI project DARIAH: Digital Research Infrastructures for the Arts and Humanities (Europe)
Status

Regular phone conferences:
• April
• June
• August

Publications:

[1] DLCL-Key Anforderungsanalyse, June 2012
[2] DLCL-Key Foliensatz, June 2012
Lightoptical Nanoscopy

Workplan:

2013  Definition of tool tree for image data evaluation, data storage and image visualization

2014  Data life cycle analysis, development data workflow for specific applications in biomedicine and radiation research

2015  Application analysis for meta data handling – multi user test

Most important requirements:

• Data storage facilities
• Data life cycle definition
• Application dependent evaluation
• Meta data handling

DSIT Services

- Federated Identity Management
- Federated Data Access
- Meta Data Catalogues and Repositories
- Archive Service
- Monitoring
- Data Life Cycle Support
- Data Intensive Computing

SPDM of nucleosomes
High Throughput Microscopy: Gen Scans

Workplan:

2013  Improvement data organization, data formats

2014  Data life cycle analysis, development data workflow, analysis and meta data handling

Most important requirements:
- Data life cycle definition
- Meta data handling
- Analysis workflow and seamless integration into the biologist’s working environment

DSIT Services
- Federated Identity Management
- Federated Data Access
- Meta Data Catalogues and Repositories
- Archive Service
- Monitoring
- Data Life Cycle Support
- Data Intensive Computing
High Throughput Microscopy: Selective Plane Illumination Microscope

Workplan:

2013  Development of an image analysis workflow based on ImageJ and HADOOP

2014  Graphical user interface for workflow management (editor, monitoring)

Most important requirements:
- Performant data access
- Meta data
- Data intensive workflows

DSIT Services
- Federated Identity Management
- Federated Data Access
- Meta Data Catalogues and Repositories
- Archive Service
- Monitoring
- Data Life Cycle Support
- Data Intensive Computing
Large Volume Electron Microscopy

Workplan:
2013  Define sample preparation and microscopy workflow, data recording and meta data scheme
2014  Automation of data recording workflow, 3D reconstruction pipeline, life cycle archiving

Most important requirements:
• Data life cycle definition
• Meta data definition and handling
• Analysis / 3D reconstruction workflow
• Performant data access
• Data intensive workflows

Material Research: Tomography Beam Line at ANKA

Data flow

DAQ → Online Analysis → Pre-Archival → Storing → Schedule Processing → Parallel Computing → Result Analysis

ANKA ← LSDF

Workplan:

2013 Development of data life cycle components within the beam line and the LSDF
Development of reconstruction algorithms

Most important requirements:
- Sustainable data life cycle definition
- Performant data access
- Meta data
- Data intensive workflows

2014 Automated workflow with data placement in the LSDF repository

DSIT Services
- Federated Identity Management
- Federated Data Access
- Meta Data Catalogues and Repositories
- Archive Service
- Monitoring
- Data Life Cycle Support
- Data Intensive Computing
DLCL-Key: Planned Publications

• A. Zyl, …, M. Hausmann, ..., J. Hesser, Deblurring Axial Tomography using Non-Linear Regularization
• J. Lux, …, M. Hausmann, Reconstruction and optical sequencing of a DNA repeat expansion unit by location nanoscopy.
• Xiaoli Yang, Halil Pasic et al., Data Intensive Computing of X-Ray Computed Tomography: Reconstruction at the LSDF, submitted to PDP2012
• Danah Tonne et al., Access to the DARIAH Bit Preservation Service for Humanities Research Data, submitted to PDP2012
• Volker Hartmann, Jens Otte, et al., Selective Plane Illumination Microscope
• Halil Pasic et al., Data Life Cycle Design at ANKA
• D. Haas, H. Pasic, et al., Status of the ultra fast tomography experiments control at ANKA, accepted at PCaPac 2012
Conclusions

a. DLCL-Key: *broad spectrum* of scientific projects

b. *Heterogeneous* team, *highly motivated*

c. Nearly all *DSIT services* will be required

d. Workplan (18 months) is defined

Next steps:
→ Overlaps and common tools need to be defined: imaging, large data volumes, …
→ Establish knowledge and software exchange?
International Activities

Research Data Alliance (www.daitf.org)
Data Access and Interoperability Taskforce

1.-3.10. Global Data Meeting Washington (WW)
  • Forum to achieve interoperability
  • Define building blocks of a data/information infrastructure
  • Definition of Working Groups (12-18 month)

22.-24.10. EUDAT conference (Barcelona)

3.-4.12. e-IRG Meeting on Data Issues (Amsterdam)

18.-20.3.2013 Global Meeting on Data Issues (Gothenburg)

Literature:
Riding the Wave, Knowledge Exchange: A Surfboard for Riding the Wave
Gemeinsame Wissenschaftskonferenz des Bundes und der Länder: Gesamtkonzept für die Informationsinfrastruktur in Deutschland, April 2011