Pilot integration of an electronic lab notebook and an open source research data repository as part of a modular biomedical research data platform

Harald Kusch, orcid.org/0000-0002-9895-2469
Sara Nußbeck, orcid.org/0000-0003-1223-6494
Péter Király, orcid.org/0000-0002-8749-4597
Rory Macneil
Mercè Crosas, orcid.org/0000-0003-1304-1939
Research Data Platform
Integrative Data Management

• Collection and interconnection of research data used in publications

• Persistent resolvable access and long-term re-usability of CRC 1002 data

• Data standardization for submission to public repositories

• Data access: Development of Use & Access Policies
RDP Modules are re-used and shared in **CRC 1190** Research Data Platform
Challenges

Academic research environments

• No routine workflows (basic research)
• Dynamic working structures (heterogeneous methodology)
• High dynamics of interdisciplinary team composition
• Data management/access policies rarely defined
Challenges
Introduction of ELNs

• Hard to follow rapid changes on overwhelming ELN market

• Acceptance of users (time and resource investment plus changing century old habits)

• Supplying 24/7 professional support

• Individual vs. institutional requirements
ELN data organization

Research group representation

CRC 1002
(Scientific Consortium)

Research Group 1
1. Community (optional)
2. PI
3. Lab Manager
4. User (Scientist)

Cooperation

Research Group 2
1. Community (optional)
2. PI
3. Lab Manager
4. User (Scientist)
ELN data organization

Structuring data in an ELN

a. Chronologically

Diary
  - Day 1
    - Exp 1a
    - Exp 2
  - Day 2
    - Exp 1b
    - Exp 3
    - Method 4

b. Project-oriented

Lab...
  - Project 1
    - User 1
      - Exp 1
    - User 4
      - Exp 1
    - Exp 2
    - Exp 2
  - Project 2
    - User 1
      - Exp 1
    - Exp 1
    - Exp 3
    - User 2
      - Exp 1

Lab...
  - Methods (SOPs)
    - PCR
    - Agarose-Gel
    - Cyto
  - Results
    - PCR
    - Agarose-Gel
    - Cyto

c. Method-oriented
ELN data organization

Unstructured documentation

- New methods
- Room for “creativity”
ELN data organization
Structured documentation

- Centralized protocols
- Learning tool
- Quickly shared
- Metadata recording facilitated
ELN data organization
Cross-linking data in a connected ELN

- Research Group Filestores
- Antibody Catalogue
- Published Data Registry
- Mouse Line Catalogue
- Cell Line Catalogue
- Web Resources (Databases, Catalogues)
- Research Group Databases
- RSpace
- Echo Data
ELN data organization

Example: Cross-linking data

The antibody used in the experiment is p115 (BD Biosciences Catalog no. 612261). It is highly specific and stored at -20°C.

Used Antibody: https://hdl.handle.net/11022/umg-sfb1002-antibody-primary-656
ELN data organization

Institutional filestore connectivity

• Cross-linking of primary research data
• Enables remote access to lab data
• Exports contain automatic lists of cross-linked data
ELN data organization

Sustainability

Implemented ORCID integration
ELN data organization

“Proof Security”

Audit Trail

Signing/Whitnessing
ELN data export
Sustainability

Easy additional re-usable archiving strategies:

Regular Data-Exports

Most similar to paper, partially interactive => PDF-Format

Most similar to ELN (Rspace), interactive => HTML-Format

Interoperability, human and machine readability => XML-Format

Interfaces to repository software (e.g. Dataverse, DSPACE)
ELN data publishing

Sustainability

Three-click action: Data export into Dataverse repository
ELN data publishing

Sustainability

Persistent identification via DOI and ORCID integration
ELN connectivity

Data & Files
- Lab Files
- Institutional Data
- Online Files
- Google Drive
- Dropbox

Tools
- eCAT
- eCAT
- RSpace
- APIs
- Apps
- Word Files
- GitHub

Publishers & Repositories
- Data
- Metadata
- File Links
- ORCID
- DataVerse
- Springer Nature
- Wellcome Open Research
- figshare
- Springer Nature
- Wellcome Open Research

DSPACE
Summary

• ELN facilitates digital data organization

• Metadata schema application easily configurable and re-usable

• Existing interfaces to repository software tools facilitate vendor-independent data sharing

• Dataverse repository enables easy publishing of persistently identifiable datasets
Developments Needed

• Extend and standardize APIs (Use case collection)

• Further improve sustainability (Use case collection)

• New / Improve interfaces with other RDM tools (rather than new ELN features)

• Establish ELN user community / communication platforms (product independent)

Please come to our demo for further discussions!
CRC 1002/1190 INF team

Dr. Sara Nußbeck
(TP INF-Leader IT)
orcid.org/0000-0003-1223-6494

Prof. Dr. Blanche Schwappach
(TP INF-Leader Research)
orcid.org/0000-0003-0225-6432

Dr. Evelina De Laurentiis
(PostDoc)
orcid.org/0000-0002-6469-1627

Dr. Harald Kusch
(PostDoc)
orcid.org/0000-0002-9895-2469

Markus Suhr
(B.Sc. Inf)
orcid.org/0000-0002-6307-3253

Georg Aschenbrandt
Luca Freckmann
Christian Henke
Sophia Rheinländer
Björn Hansen
Visit http://www.eresearch.uni-goettingen.de/content/eresearch-toolbox-electronicnote-keeping for further information.

All icons used were designed using Freepik from and licensed under Creative Commons BY 3.0. www.freepik.com; www.flaticon.com