ELN Workshop

# ELNs in the context of **Researcher and Lab Workflows** and **Institutional Infrastructure**


- Harald Kusch (Universitätsmedizin Göttingen)
- Rory Macneil (Research Space)


ELN Workshop Berlin – 14.-15. März 2016

# Who are ELNs for?

Researchers

Labs

and

Institutions

# ELN market has become commoditized

- More and more good (Standalone) ELNs
- Improving all the time as technology develops
- Similar features and capabilities (e.g. Fred Hutch requirements)
  - Organizing Work/Notes
  - Templates
  - Uploading/Exporting
  - Importing (Images Primarily)
  - Editing/Annotating Images
  - Search
  - Version Control/Timestamp
  - Sharing
- How to distinguish?

# What researchers, labs and institutions should look for when evaluating ELNs

Does the ELN possess 'best of breed' features and capabilities?

***AND ALSO***

Can the ELN connect to?

A. Files
B. Other things researchers deal with
   - Images
   - Documents
   - Specialist file types
   - Databases
C. Repositories and archives

Is the ELN a good tool for ***labs***?

A. Collaboration
B. Management
   - PI
   - Lab Admin
C. Change

Can the ELN be part of tomorrow's solution as well as today's?

A. As the infrastructure at the institution evolves
B. When researchers adopt new tools
C. When researchers moves between labs or the PI and lab move to another institution

# Research Data Management Tools

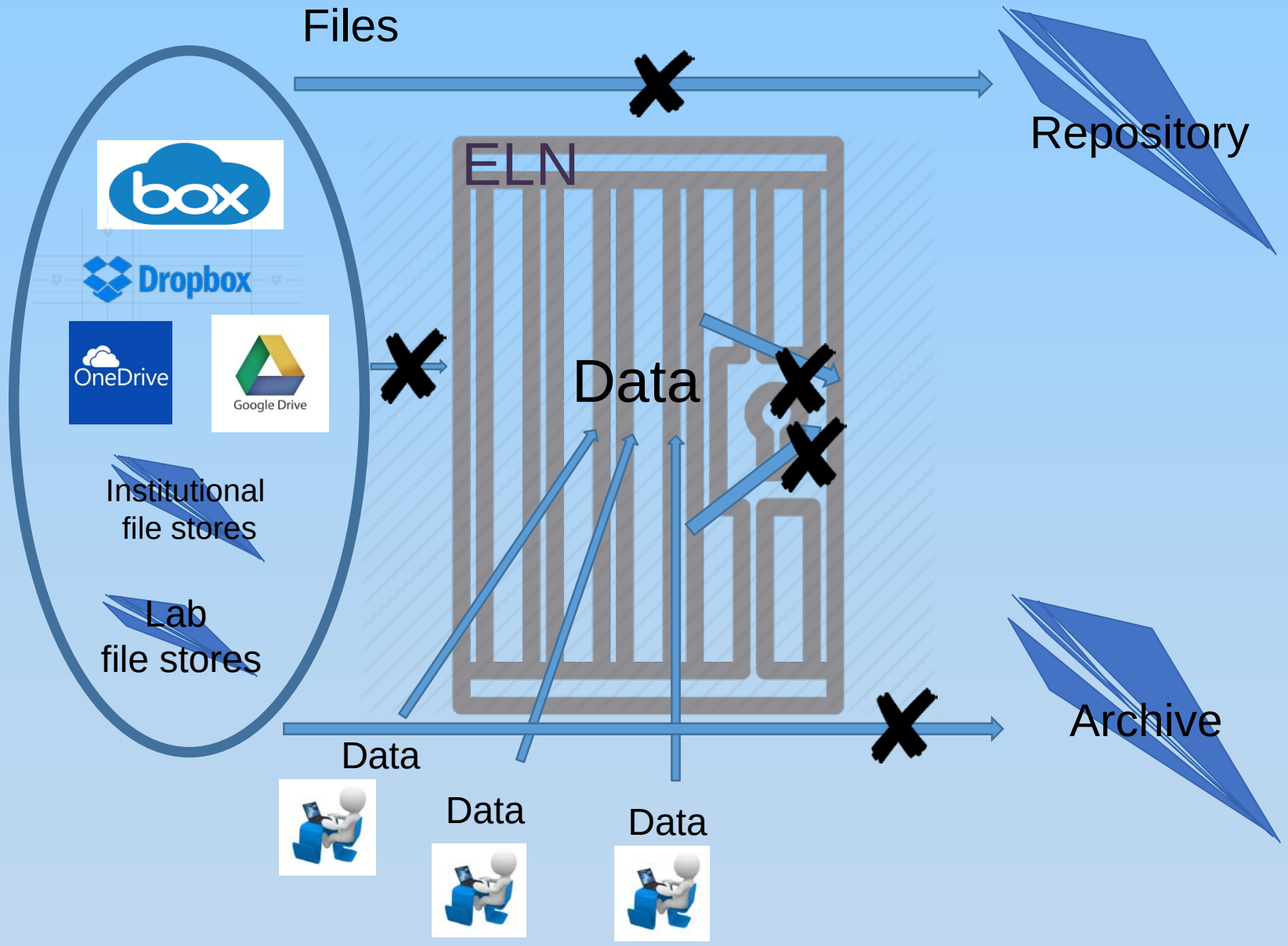
Electronic Lab Notebooks

Box/Dropbox

On Premises Storage

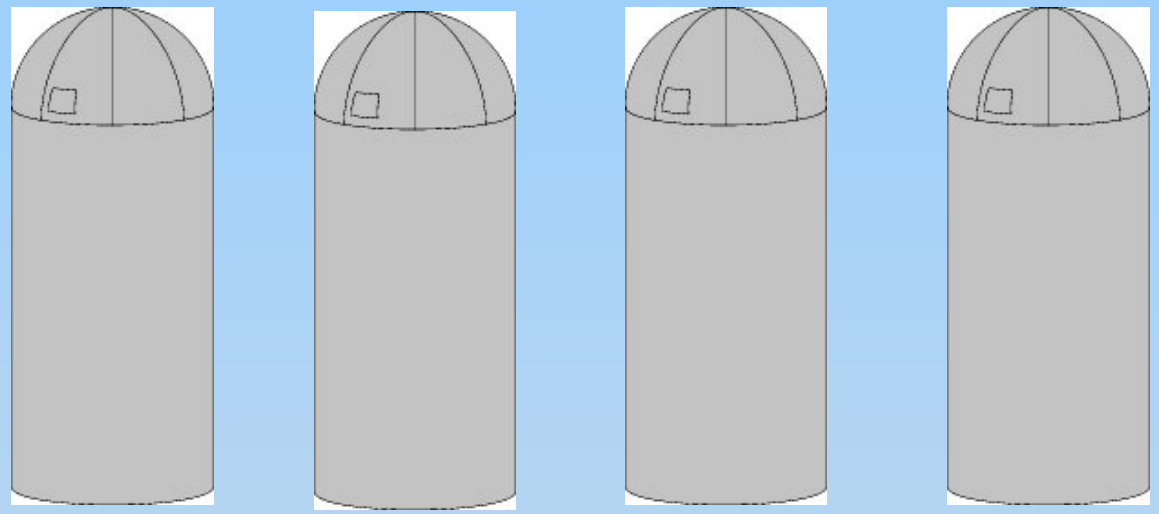
Institutional Repository

# Standalone ELN Workflow



# Result

**ELN just another silo**




**ELN does not integrate with other tools**

**Data does not get captured by the institution**

# RSpace:  the first Connected ELN

**Developed over 3+ years in partnership with**

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

THE UNIVERSITY OF EDINBURGH

**Released last November**

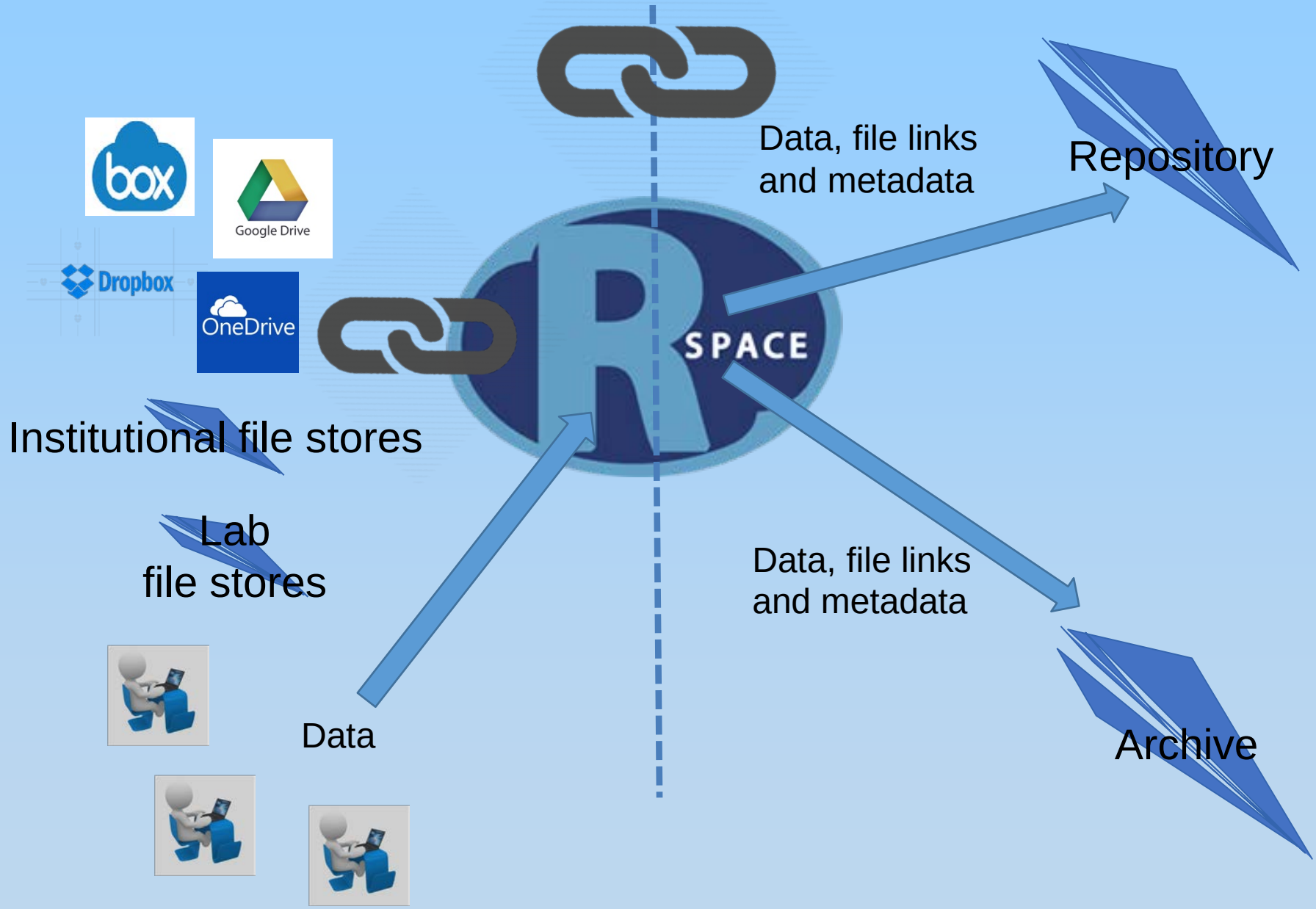
# RSpace: Best of breed user and lab ELN capabilities

- Fred Hutch requirements
  - Organizing Work/Notes
  - Templates
  - Uploading/Exporting
  - Importing (Images Primarily)
  - Editing/Annotating Images
  - Search
  - Version Control/Timestamp
  - Sharing
- Other features: signing and witnessing, snippets
- Integrations
- PI/Lab admin control

# RSpace:  The first Connected ELN

- √ Ability to link to external files
  - Commercial (OneDrive, Google Drive, Dropbox and Box)
  - Institutional and lab (sftp/cifs and samba protocols)
- Ability to track changes in locations of linked files
- √ Ability to export to other systems (e.g. archives and repositories) via  XML
- √ Ability to re-import data into other RSpace instances → portability and mobility
- √ Free Community version
- √ Enterprise support:  LDAP, Tiered Admin, Sysadmin features

# RSpace: the first Connected ELN



# Max Planck Institutes and Researchers

World leading research

Requires world leading research tools
for today
and tomorrow

# Collaborative Research Center (SFB) 1002

- „Modulatory Units of Heart Insufficiency“ (Funded by the DFG since 07/2012 with 10 Mio. €)
- 16 scientific subprojects: research groups from
  - University Medical Center Göttingen
  - MPI for Dynamics and Self-Organization
  - MPI for Biophysical Chemistry
  - with focus on basic research with mice and cell lines
- INF-Project:
  - Headed by medical informaticist (Sara Nußbeck) und molecular biologist (Blanche Schwappach)
  - Research focus on introduction and evaluation of an Electronic Laboratory Notebook (and a Research Data Platform)
  - Pilot ELN-introduction in four research groups

# CRC 1002 Research Data Platform

- Establishment of CRC 1002 Research Data Platform (modular approach)
- Collection and interconnection of research data used in publications
- Persistent resolvable access and long-term archiving of CRC 1002 data
- Data standardization for submission to public repositories
- Data access: Development of Use & Access Policies

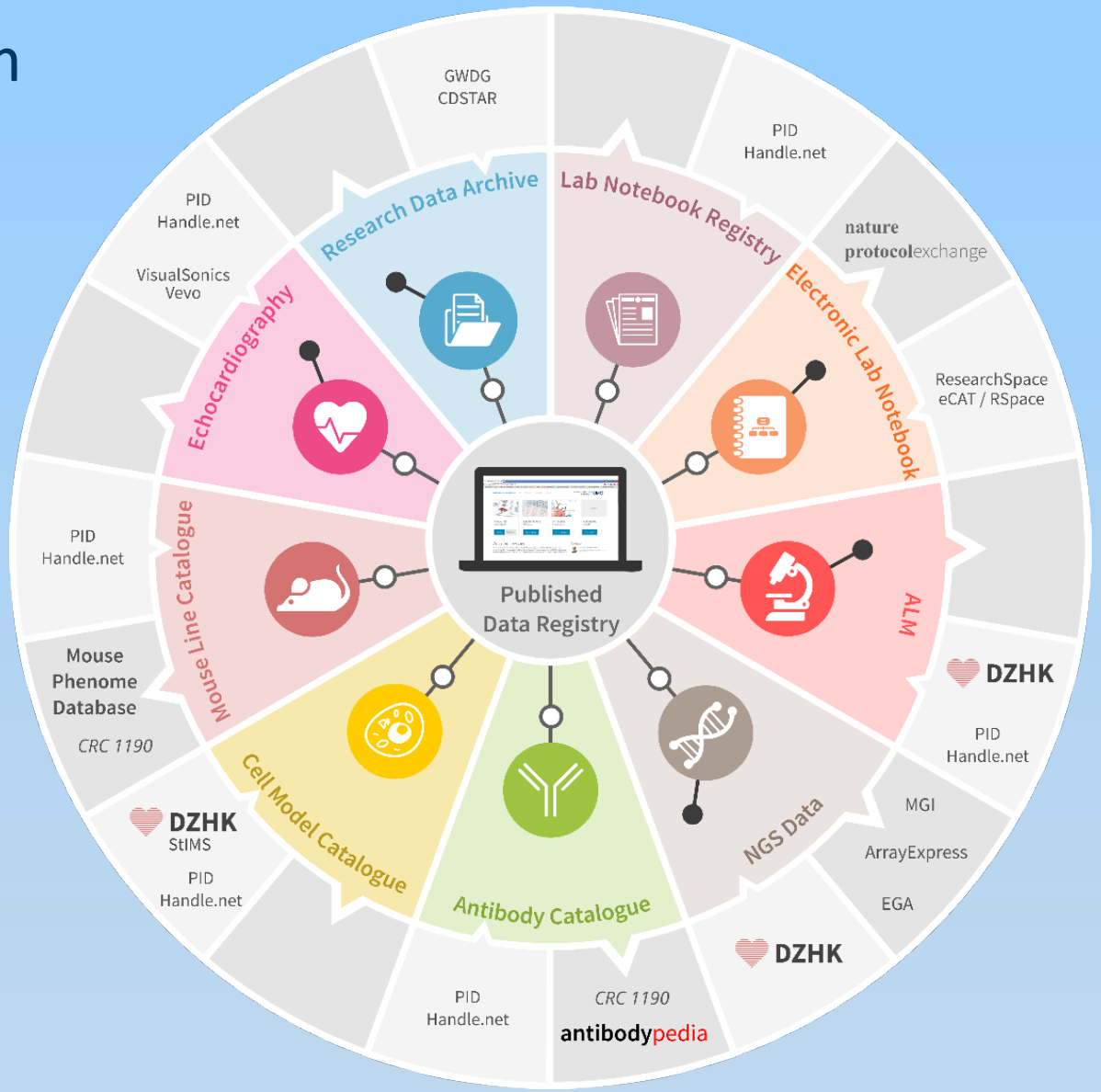



Figure design by Bartlomiej Marzec, University Medical Center Göttingen.
All icons used were designed using Freepik (www.freepik.com) from Flaticon (www.flaticon.com) and licensed under Creative Commons BY 3.0.

# Research questions

- Can a paper-based lab notebook be substituted by an ELN?
  - Which software and hardware is suitable?
  - Which framework / infrastructure is necessary?
  - What are aspects to consider?
  - What could be possible pitfalls and drawbacks?
  - Aim: Give recommendations!

EMBO Reports:
doi:10.15252/embr.201338358

eGMS:
doi:10.3205/mibe000162

# Challenges: Introduction of an ELN in 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 from an IT perspective

**Anticipated challenges**

- Acceptance of users (Anything new means investment [time and resource-wise] plus changing habits)
- Soft- und hardware decisions
- Development of metadata schemes→ for structured and standardized data documentation

**Not anticipated challenges**

- Delay in ELN introduction caused by differences in IT- facilities in different institutes
- Underestimation of the conceptualization of the structure for data documentation

# IT requirements – potential drawbacks

- Stable WiFi connection

- Sufficient storage capacity

- Linking net storage devices

- Long term archiving of data

# Aspects to consider: data structure in the ELN

- Should not be too complex to hinder sorting of data entries and therefore prevent writing them down.
- Should be easy and traceable for the PI
- The ELN offers many new possibilities (integrated search, linking of study objects, methods and results)
- Can support the writer in his/her thoughts (for theses writing, publications, grant proposals, reports by providing the structure -> intro, M&M, results, discussion, conclusion)

# How to document in the ELN (structure)?

- Paper-based lab notebook:
  - Chronologic like a diary

- ELN:
  - Chronologic like a diary
  - Project and work package-oriented
  - Sorted according to used methods

# How to document in the ELN (metadata)?

- Necessary for the management of large data volumes
  - Provide structure and harmonization for research data
  - Relevant for queries and speed of receiving query results

- Provide scientific context to research data

- Increase re-usability of data:
  - Use of ontologies and controlled vocabularies (standardization)
  - secondary and meta analyses (research networks, scientific communities)

# Importance of structured data export from ELN

- Integration in institutional Research Data Management Concepts
- Exchange of data between ELN systems
- Machine readability for future “big data” analyses

# Summary

- Do not underestimate the introduction of an ELN – a useful integration in the local research data management strategy is complex
- For achieving a high acceptance amongst users, it is useful to integrate them in all processes before starting to use the ELN
- Before starting an ELN, clarify local IT-requirements at partner sites (WiFi, server architecture)
- Start with only a few ELN users as an intensive coaching is recommended

# Acknowledgements

**Dept. of Medical Informatics, University Medical Center Göttingen, Germany:**
Sara Y. Nußbeck
Julia Menzel
Bartlomiej Marzec

**Dept. of Molecular Biology, University Medical Center Göttingen, Germany:**
Blanche Schwappach