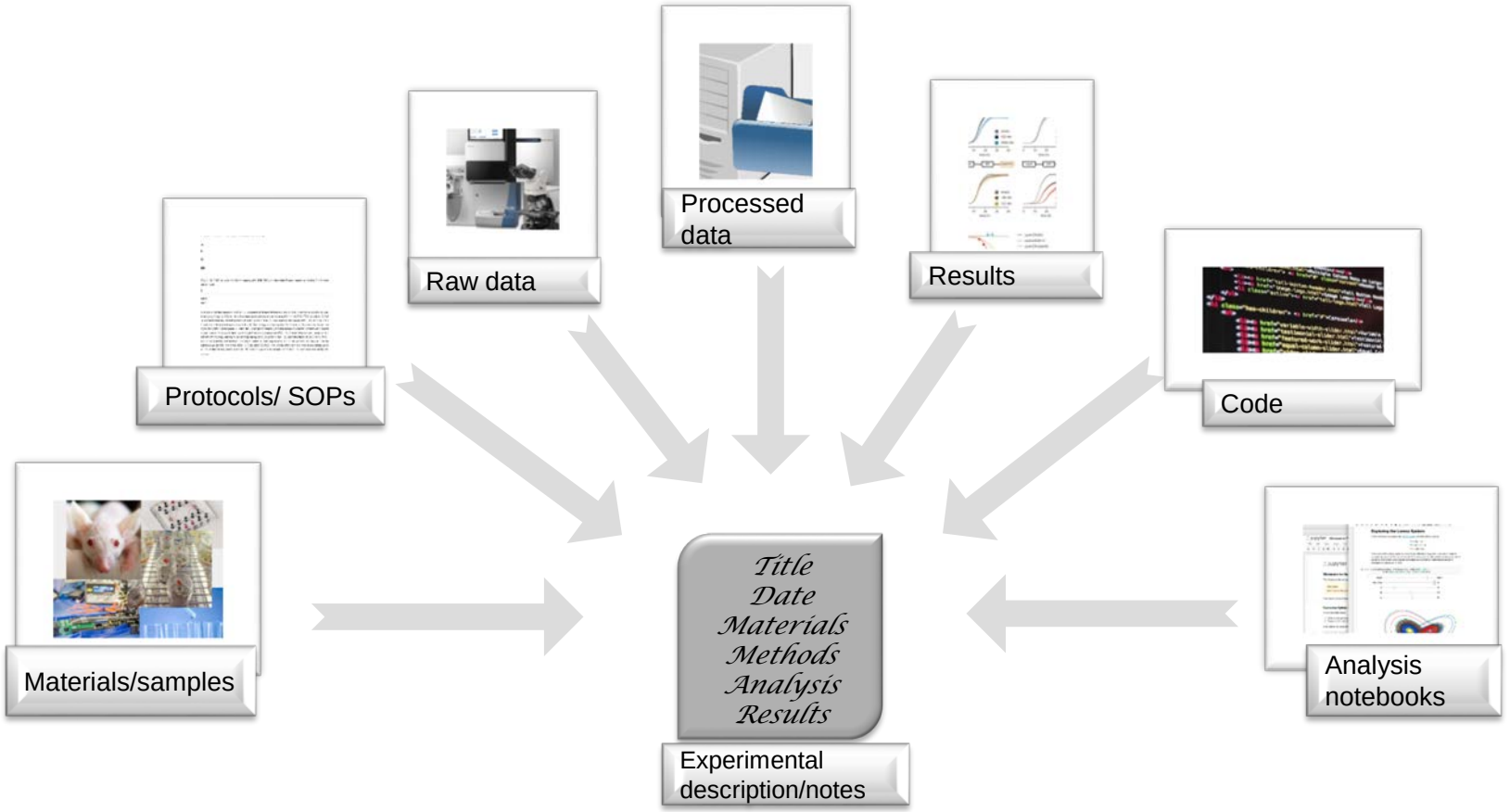
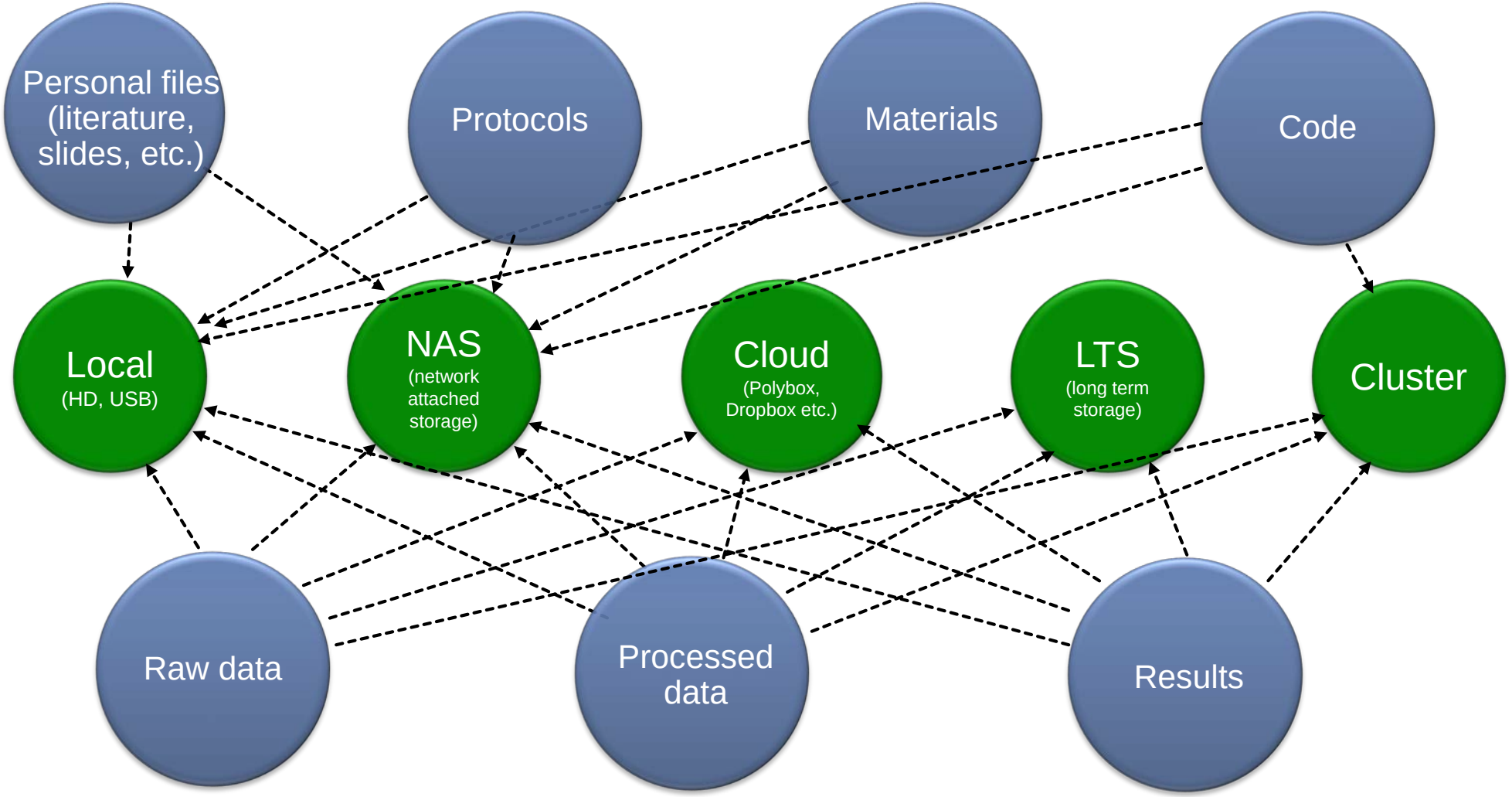
Active Research Data Management

What does it take to manage research data?

Complex process that requires tracking and linking different types of information



The Data Spread

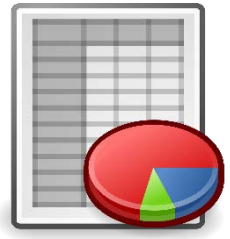


Management of Materials and Samples

What?

- Biological samples
- Chemical samples
- Materials
- Devices
- ….

How?



Spreadsheets

- *Not scalable*
- *No sharing*
- *No efficient search*
- *Easy to use*

WIKI

- *Scalable*
- *Sharing*
- *Search functionality*
- *Versioning*

Laboratory Information Management System (LIMS)

- *Scalable*
- *Sharing*
- *Search functionality*
- *Require time for set up and maintenance*

Management of Protocols

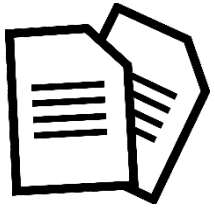
What?

- Step by step description of procedure
- Experimental/computational parameters (e.g. temperature, time, etc.)
- Machine used (experimental)
- OS, program, version, etc. (calculation)

How?

Paper notebook

- *Not scalable*
- *No sharing*
- *No search*
- *Easy to use*



Text files

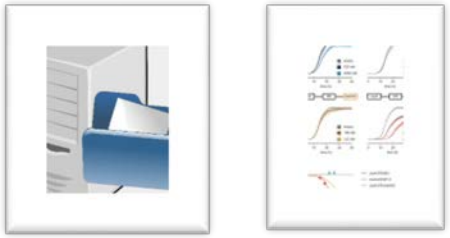
- *Not scalable*
- *No sharing*
- *No efficient search*
- *Easy to use*

WIKI

- *Scalable*
- *Sharing*
- *Search functionality*
- *Versioning*

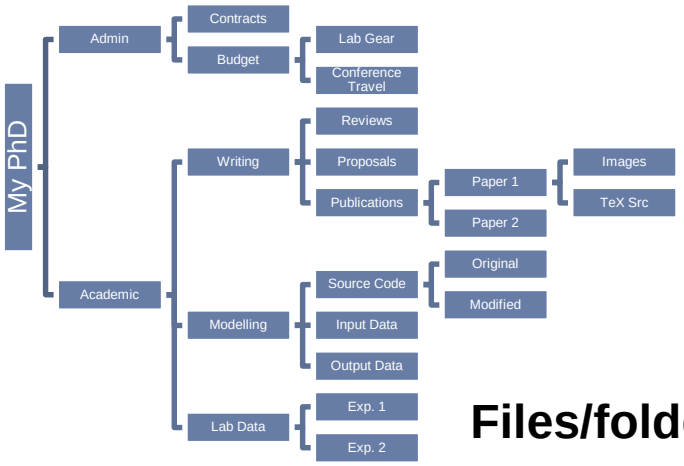
- *Scalable*
- *Sharing*
- *Search functionality*
- *Time for set up and maintenance*

Management of Research Data Files



What?

How?

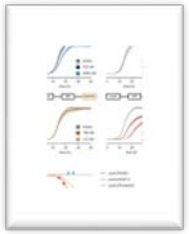


Files/folders naming conventions

- *Structured organization of data*
- *Data is annotated with metadata*
- *Searchable*
- *One central location*

Data management platform

Management of Research Data Files



Practical considerations for organizing files

- Keep stuff together that belongs together
- Keep path names short
  - < 255 characters
- File names should
  - Reflect content and be unique
  - Use only ASCII characters
  - No spaces & diacritic characters etc. (Ä, ö, è, ū, …)
- Careful: Not all systems are case sensitive!
  - UNIX: case sensitive
  - Win/Mac: mostly case insensitive
  - Assume that this, THIS and tHiS are the same.
- Document your structure and file naming conventions in a README text file

- Write **dates** like this: **YYYY-MM-DD**

Further file and folder organization tips:

http://www.data.cam.ac.uk/data-management-guide/organising-your-data

Metadata

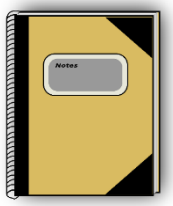
- **Metadata** is the *data about your data*
- Use of structured metadata facilitates data organization and searches
- Examples of metadata*:*
  - *Investigator*
  - *Date*
  - *Title & Description*
- Several metadata schemas are available – see: http://www.dcc.ac.uk/drupal/resources/metadata-standards
- Metadata can be supplied in the **file header** (e.g. TIFF, MP3) or as **sidecar files** (e.g. XML, JSON)
- Apply metadata as early as possible in your workflow (e.g. during acquisition)

Experimental Descriptions & Notes

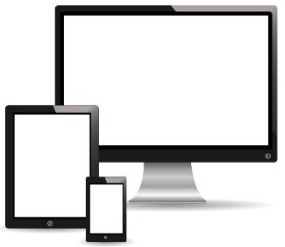
What to document?

- Goals
- Materials
- Methods
  - Experimental/computational procedure
  - Analysis procedures
- Results
- Links to data

How?



Paper laboratory notebook



Electronic laboratory notebook (ELN)

ELNs vs. Paper Notebooks

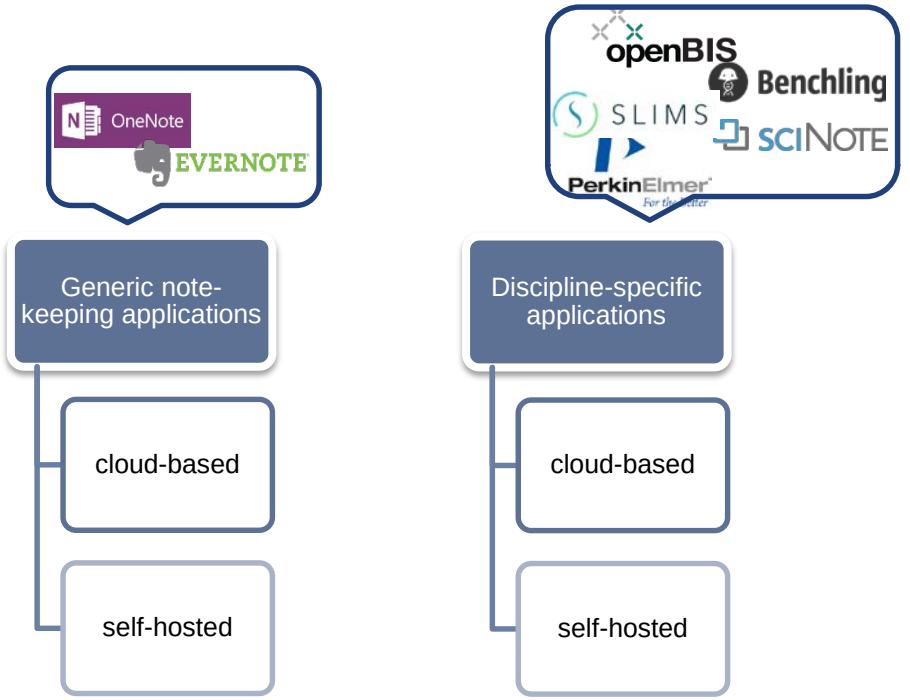
Advantages of ELNs over paper notebooks:

- Sharing
- Most ELNs have rights management
- Most ELNs keep track of changes
- Searching
- Easier to link digital data
- No issues with handwriting
- Can be backed up

Disadvantages of ELNs over paper notebooks:

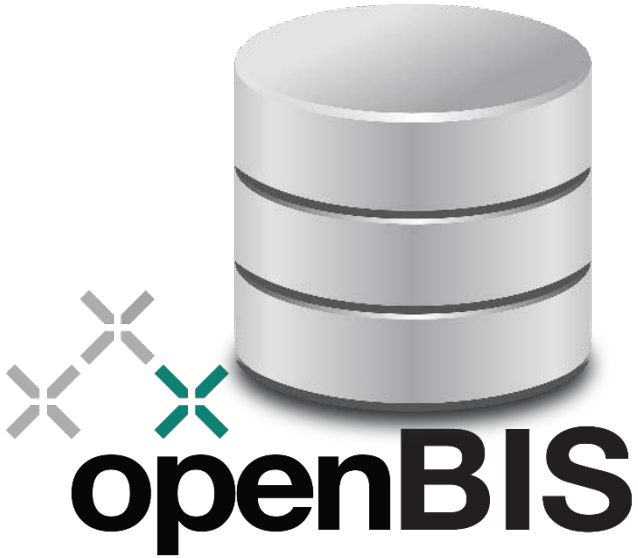
- Require change in working mode
- Have a learning curve

Different types of ELNs:



➢ *When choosing an ELN always make sure you can export your data in a common open format (e.g. .xml, .json, .html, .pdf, .txt, .docx)*

openBIS – A Comprehensive Data Management Solution



https://openbis.elnlims.ch