Research Data Management Plan

This Research Data Management Plan (RDMP) is valid for all projects in the Daumann Lab:

1. Data Summary

New data is generated during research projects by Students (Bachelor, student assistants, Master) and PhD, Postdoc or the PI. Data, Findings and Conclusions are included in research theses and publications as well as in grant proposals. Type and size of data is highly variable (CSV, CIF, PDB, QTOF and GC-MS data, NMR data, Gen5, Cary-Software, SpectraManager, Unicorn), below are specifications of best practices to ensure a FAIR (Findable, Accessible, Interoperable, Reusable) data policy in the Daumann Lab.

2. FAIR data

2. 1. Making data (and metadata) findable

Use of Electronic Lab Notebook (ELN) Chemotion for all projects https://chemotion.net/

Everyone uses ELN. New group members (including undergraduate students) are introduced to the ELN on their first day and given access to useful links/information (e.g. tutorial videos). If someone does not have a personal laptop or computer near the bench, the Daumann lab will offer a tablet specifically for the access (via google chrome) to the ELN during the time in our lab.

Requirements for Experiment-Identifiers/Naming conventions:

- Initials (LJD instead of LD when possible to avoid duplications)
- A clear numbering system (e.g. Lena used LJD-I-2, initials, labbook number I, page 2. Others may just number the samples sequentially from 1, 2, 3, do whatever is the best way moving forward with the ELN for chemical reactions)

What to include in ELN for reactions/experiments?

- Page numbers (not necessary for ELN, but for analog labbook if you use one in addition)

- Table of Content (not necessary for ELN, but for analog labbook if you use one)

- Water quality (e.g. distilled, MilliQ, LC-MS or HPLC grade)

- Chemicals and solvent quality

- Series of addition (e.g. what was added to what, when and how)

- Materials of vessels etc. used (e.g. 50 mL Round bottom flask, Scintillation vial, Falcon tube) Also for MS! We found differnet impurities coming from differnet vials, so it is crucial to note down what your sample came in contact with

- Note down a standard procedure, refer to it later, but also refer to differences! "overnight" is not acceptable (can vastly differ 8-12 h).

- Colour changes during the experiments.

- Temperature (measured outside in oil bath or inside?) "Room temperature" is not acceptable (can fluctuate)!

- Did everything dissolve or was it a suspension? Habitus of the reaction and product
- Additional notes for record keeping.

- If analog labbook in addition to ELN Chemotion, LABEL IT

- Include experiment identifier in EVERY measurement (name data files AND folders!)

- Differentiate between "washs" in GC-MS, LC-MS measurements (data does not need to be recorded, only necessary for running the instrument) and "blanks" (data needs to be recorded, necessary for understanding your experiments and as important controls!)
- Note down operators (e.g. who did the elemental analysis or solved the crystal structure)
- How to structure folders containing data:
- Year
- Name structure: date-identifier (2022-02-NJ189 OR 2022-02-28-NJ189)
- Where to store data:
- Sort by relevance, get rid of washs etc. regularly (1 year storage time!)
- For all data generated, provide clear version numbers (e.g, raw, converted, normalized, plotted)
- Make sure to safe your data regularly (use the 3-2-1 rule: Keep at least three copies of your data, and store two backups on different storage media, with one of them located offsite, use hard drives and Sciebo or BioAcNetwork folders etc.) For tips on refer to IT Admin Rachel or our Ak Wiki on IT.

2.2. Making data openly accessible

Multiple different types of data are produced for our projects (NMR, MS, crystallographic information, GC-MS, ITC, kinetics data reaction descriptions with yield and habitus of new compounds etc.). Data will be deposited on designated repositories CCSD, PDB, Chemotion repository and/or included in the supporting information of publications etc. Data will be converted in easily accessible files or standard file formats where available (cif, csv, etc.) and clearly linked to the respective publications.

Preference is given to certified repositories that support open access (Chemrxiv, Biorxiv, Chemotion). All synthetic data with accompanying raw analytical data in an easily readable format is deposited in the ELN Chemotion and upon publishing linked to the repository.

For easy linkage of Chemotion data with the associated repository please always upload the image of a spectrum and the file of your data to Chemotion. e.g. a png of the NMR as well as the .mnova or .fid file. Scan the spectra received from the MS facility and include picture of your ESI-MS. For UV-Vis use ASCII or CSV.

Use a dot for decimal separation! Global English notation for the entire AK Daumann in your data!

You may also deposit physical samples with the molecule archive https://compound-platform.eu/

2.3. Making data interoperable

Where defined, we will adhere to standards for formats compliant with available (open) software applications (e.g. CIF, PDB, ACSII/CSV).

2.4. Increase data re-use

Data will be included as supporting information in an open access format (BY, ND, NC licence). Manuscripts will be deposited to open access repositories prior to publication.

3. Allocation of resources

ERC Project: Costs for open access are covered in the ERC budget. Project DEAL (with certain publishers, Wiley, Springer) can make publications form LMU authors open access. LMU has also a fund to cover a part of open access publication fees. Lena Daumann and the team are responsible that the Research data management plan (RDMP) is implemented globally in the group.

4. Data security

4. Data security

Short term: 3-2-1- rule. Use our BioAC networkdrive or sciebo and group hard drive in addition to your computer during your time in the Daumann lab. ELN: transfer the collection with your experiments to PI when you leave the lab. During Home office: use VPN tunnel for transferring data.

Long term: Data safely stored in certified repositories for long term preservation and curation.

This RDMP was set up by the Daumann lab team (current version April 2024) is in continuously updated as needed.