

DAPHNE4NFDI TA2 Face to Face Workshop

(Meta)data repositories and catalogues

After a long time meeting solely in an online format, the members of DAPHNE4NFDI Task Area 2 (TA2), [\(Meta\)data repositories and catalogues](#), met in Erlangen, from September 29th to September 30th, to discuss sample IDs, metadata formats, as well as FAIR data and reproducibility. The workshop brought together sixteen people who had the opportunity to network and exchange views in person, with an additional handful of people having attended the event remotely.



Figure 1: ICSP institute, at the FAU university in Erlangen, Nürnberg. The host institute of the TA2 workshop.

Day 1 included updates from the several collaborators who contribute to Task Area 2 of DAPHNE4NFDI. Among the topics that were discussed on this day, were the metadata catalogues as well as their schema and file formats. The agenda also included updates from facilities that contribute to DAPHNE (MLZ, HZB), as well as discussions about the Open PMD file format, which is being used in simulation pipelines.

Day 2 was reserved for discussions, which started interestingly with the acknowledgement that a “sample” refers to, more generally, the subject(s) of analysis, which means that a sample is not always a physical object but may also be a process that is under investigation. However, a sample’s information may be updated as it undergoes investigation steps, which rises other questions, such as: where should the sample information be entered first? And how should updates to this information be propagated through the landscape of catalogues and ELN’s located at different sites? Neither of which, unfortunately, had an immediate answer, but will serve to guide our future steps within TA2.

It was noted, during the discussion time, that certain file types (NeXus, HDF5) are increasingly being adopted by large-scale facilities, but special formats still exist that are in use in special situations (e.g. high data-rate applications). Both must be kept in mind by the DAPHNE consortium, as this will require harmonization of the metadata schema and file converters in the future. Furthermore, there were some discussions on “linking or copying” information from ELNs and catalogues. However, participants agreed that a basic set of “catalogue-metadata” must be copied, with respect to the sample and experiment characteristics.

Discussions and results

Following the discussions, the few points below highlight the most important outcomes of this first TA2 (in person) meeting:

- Members of the DAPHNE4NFDI TA2 will work towards defining what kinds of samples (physical object, process) are investigated by the several groups working towards competition of DAPHNE's goals.
- Current and common issues with the available data formats will be identified and addressed.
- A decision regarding the metadata that must be included in the catalogues must be made.

If you are interested in joining DAPHNE4NFDI,
please contact our coordinator:
Lisa Amelung: Lisa.Amelung@desy.de