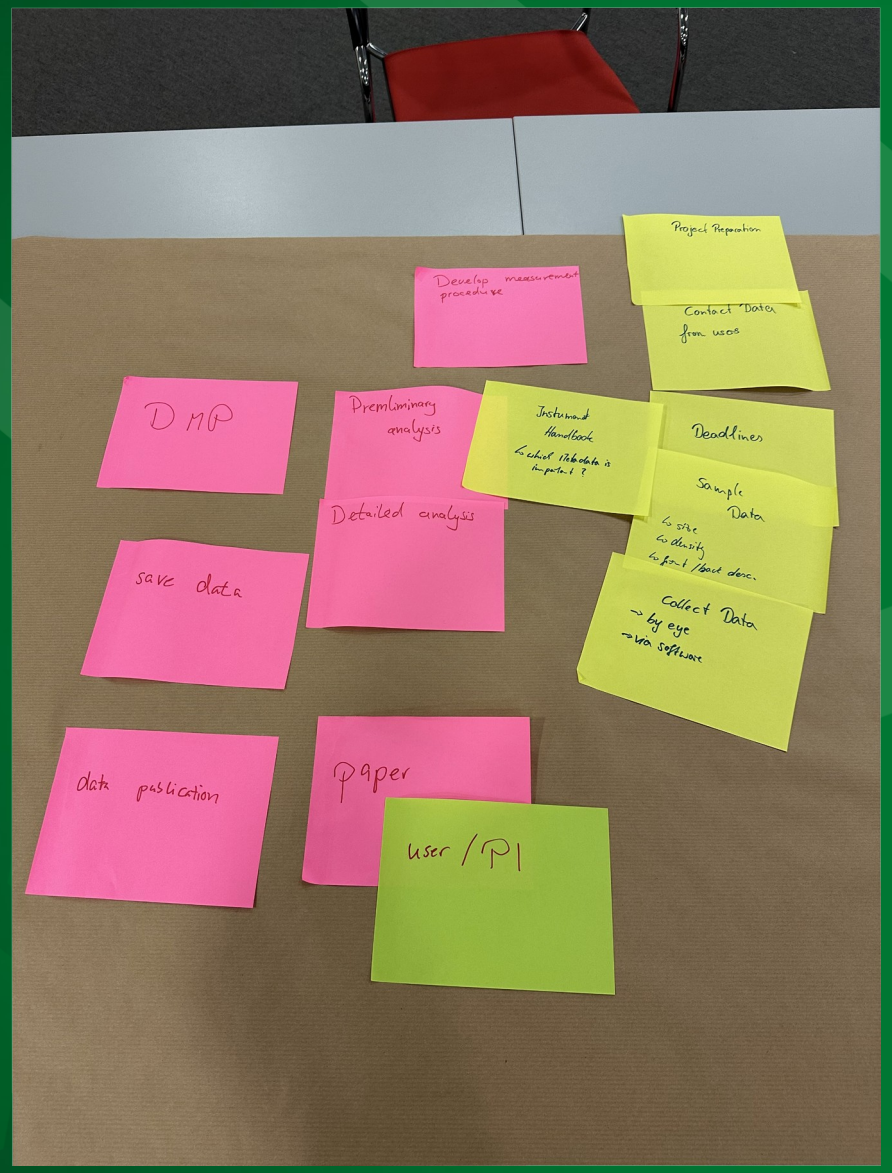
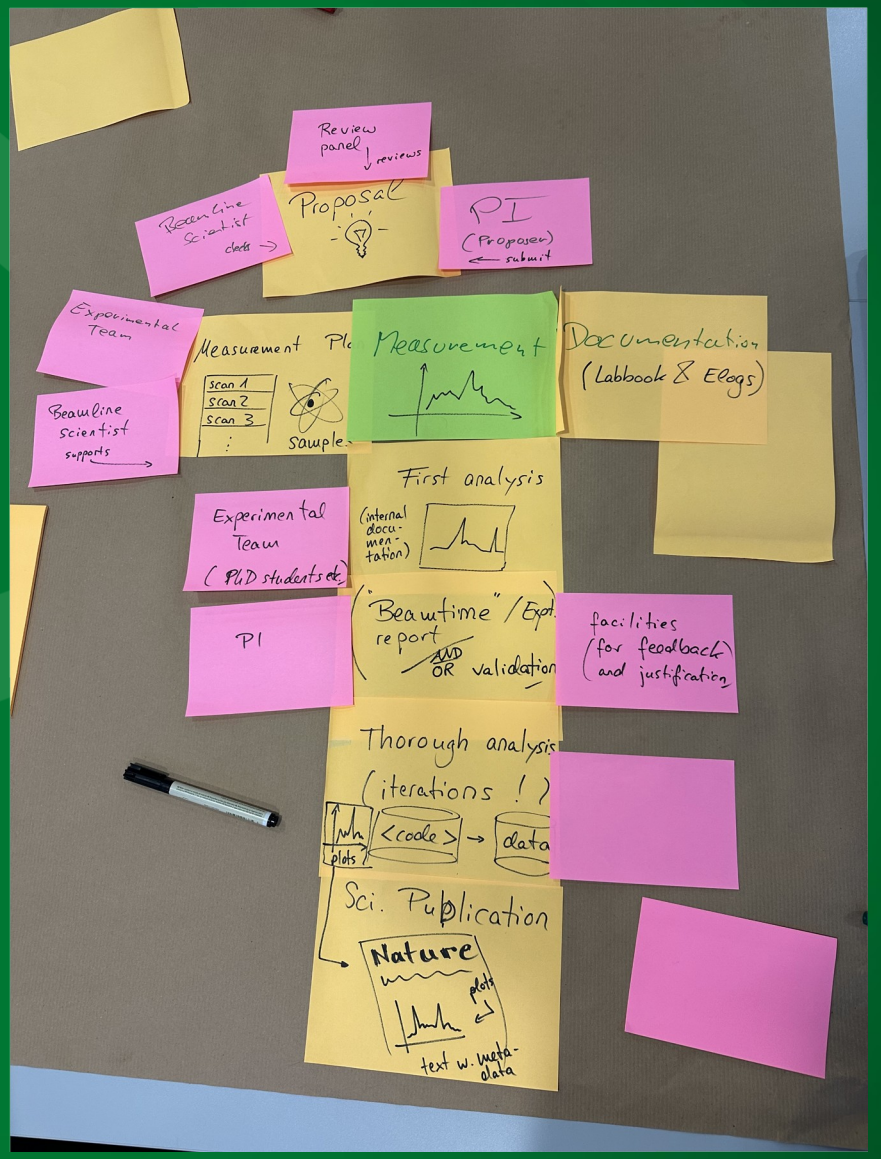
HZDR Data Management Days – Day 2

Workshop: Research Data Management in Practice: Identifying Needs and Developing Actionable Recommendations

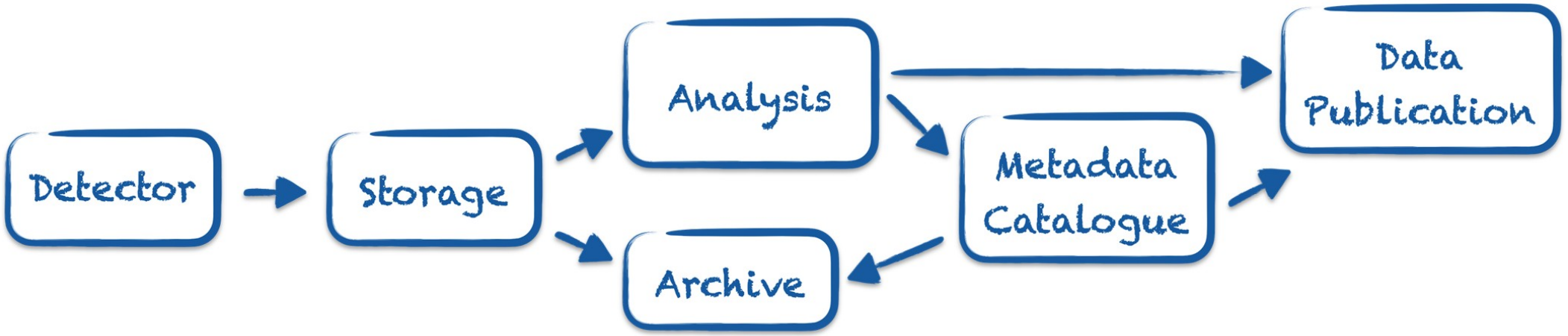
29/10/2024

Leibniz-Institut für Geschichte und Kultur des östlichen Europa
TECHNISCHE UNIVERSITÄT DRESDEN
ZIH
DRESDEN concept
Kontaktstelle Forschungsdaten
SaxFDM


Project Preparation
Develop measurement procedure
Contact Data from users
DMP
Preliminary analysis
Instrument Handbook
Deadlines
Sample Data
Detailed analysis
save data
Collect Data
by eye
via software
data publication
paper
user / PI


Review panel
reviews
Proposal
PI
(Proposer)
submit
Beamline Scientist
checks
Experimental Team
Measurement
Measurement
Documentation
(Labbook & Elogs)
scan 1
scan 2
scan 3
Sample
Beamline scientist supports
First analysis
(internal documentation)
Experimental Team
(PhD students etc)
"Beamtime" / Expt. report
AND
OR
validation
PI
facilities
(for feedback and justification)
Thorough analysis
(iterations !)
<code>
data
plots
Sci. Publication
Nature
plots
text w. metadata

„prototypical experiment“



Idea, Concept, Proposal

Every project at the HZDR usually starts with a proposal from external scientists, a project of our associates or simply with an idea.

Plan

The data that is to be collected or used to answer the research question is determined and data management is planned throughout the experiments life cycle. The phase in which data management-, as well as a software management plans, are created.

Experiment Cycle

In this phase, the actual data is collected, pre-processed, analysed, but also enriched with corresponding metadata and finally curated before further iteration takes place:

Collect

This is the phase in which experiments are conducted, observations made, surveys carried out, secondary material obtained, etc. This includes the documentation of the setup and the data generated data along with the associated metadata.

Process

Once data has been collected, it must be processed (e.g. cleaned, combined, converted, validated) to make it usable. All data processing must be documented so that the final result can be reproduced using the raw data.

Analyse

The analysis interrogates the raw material of the research to gain the insights that make up the research findings that are written up and published in the research results. The methods used for the analysis should be documented; the code written for the analysis and visualisation may need to be retained and made available.

Curate

In the curation phase, the data is provided with additional metadata and a decision is made as to whether the quality of the data and the scientific results are sufficient or whether changes to the structure are necessary in order to collect additional data.

Preserve and Archive

Preservation activities may include to store or archive data that substantiate your research findings and are of long-term value. A data storage/data centre is suitable for digital data. For metadata and software, suitable measures such as metadata catalogues or software repositories should be selected. The cross-linking of all data products is important in this context.

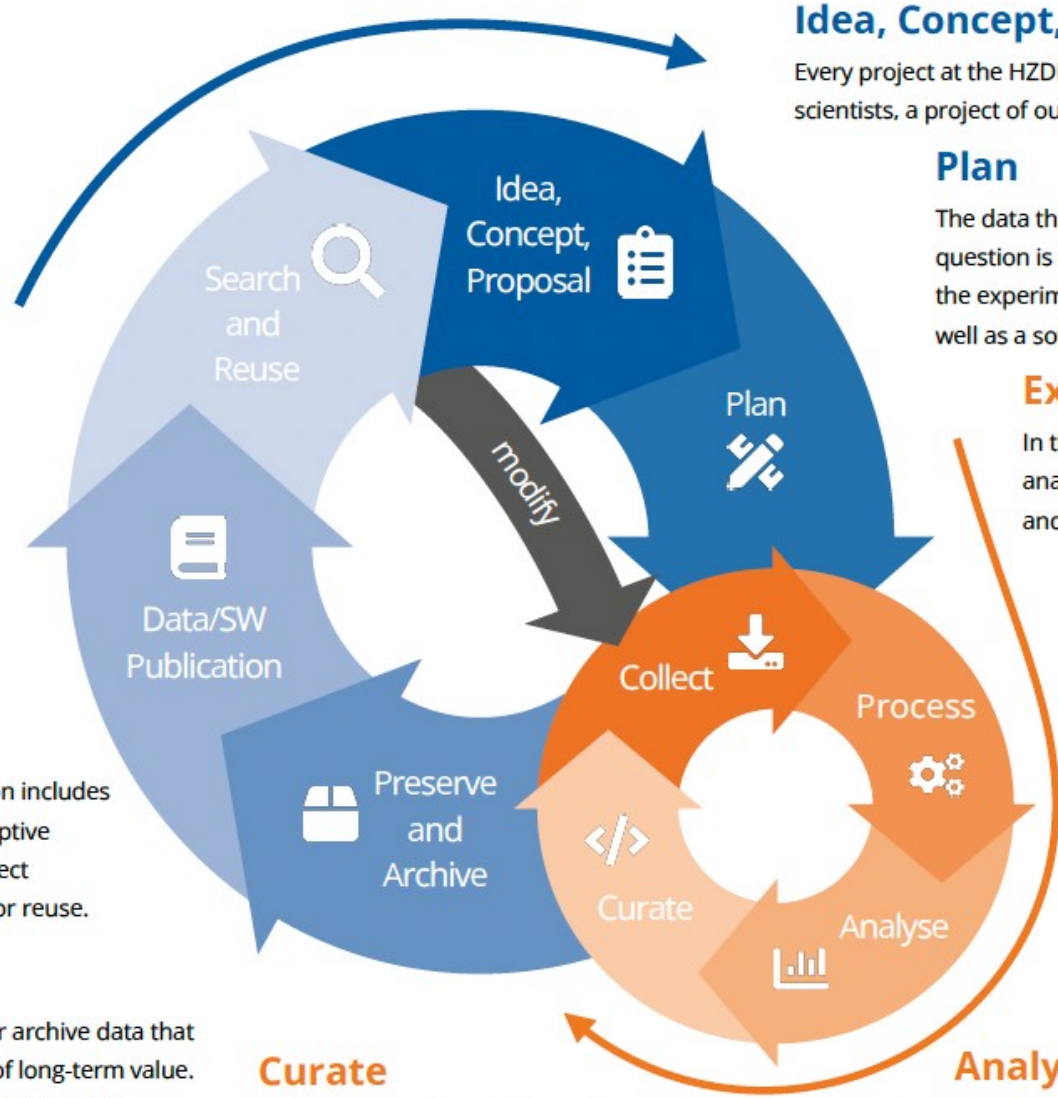
Data/Software Publication

The publication of data, software and documentation includes the conversion of file formats, the creation of descriptive metadata records with the assignment of digital object identifiers (DOIs) and the licensing of data records for reuse.

Share, Reuse and Search

Scientific Publications based on data and software should cite them or include a statement indicating where and under what conditions the data, supplementary metadata, documentation, software, are accessible and reuseable.

Data, digital artefacts available for discovery and access may be reused by other researchers, either to corroborate the findings of the original research or to gain new insights through further interrogation and analysis. At this stage, the data can become raw material collected for a new research cycle.



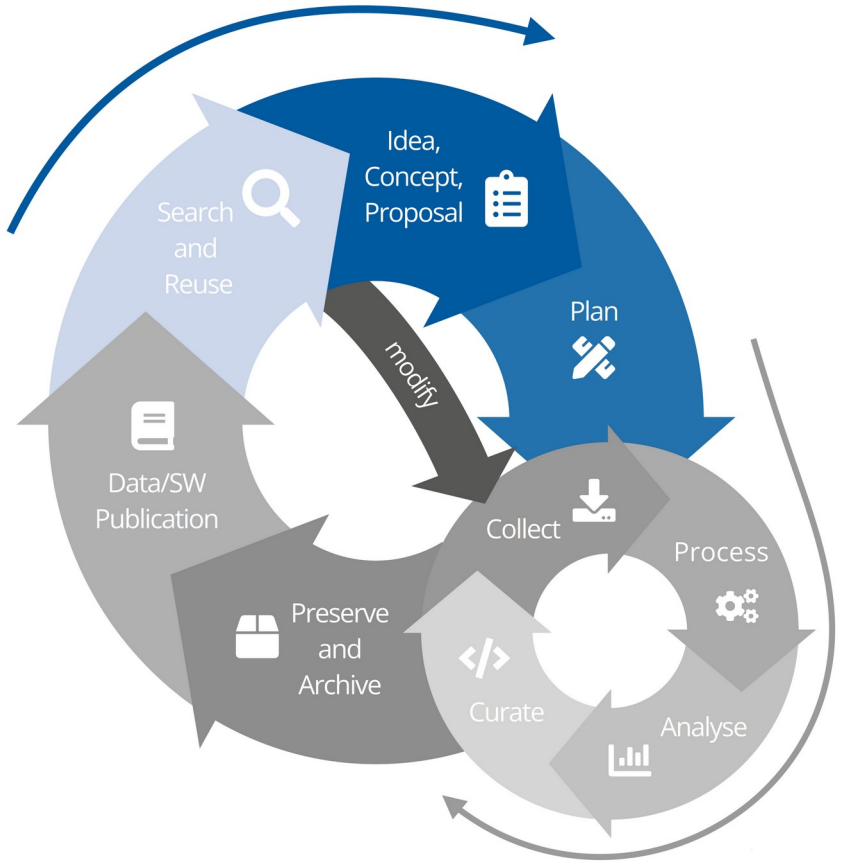
Group work instructions

Form three equally big groups

3 posters, each dedicated to 2-4 phases of the data life cycle (and the data flows in those phases)

Groups rotate as one from poster to poster and fulfill the respective tasks

Poster I – phases of the data life cycle

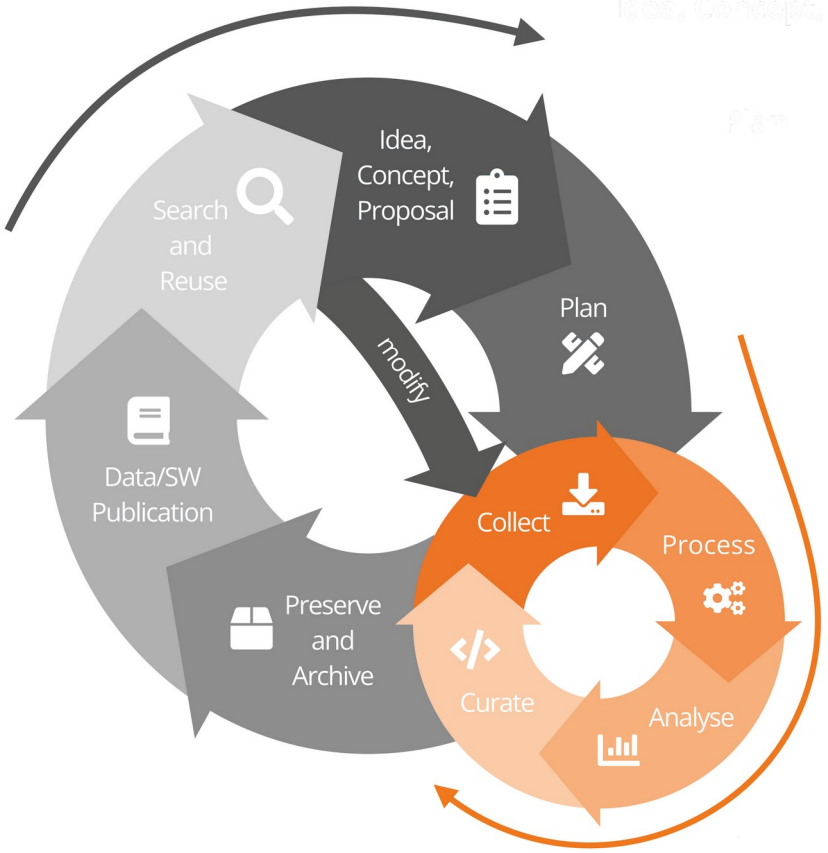


SEARCH AND REUSE

IDEA, CONCEPT, PROPOSAL

PLAN

Poster II – phases of the data life cycle



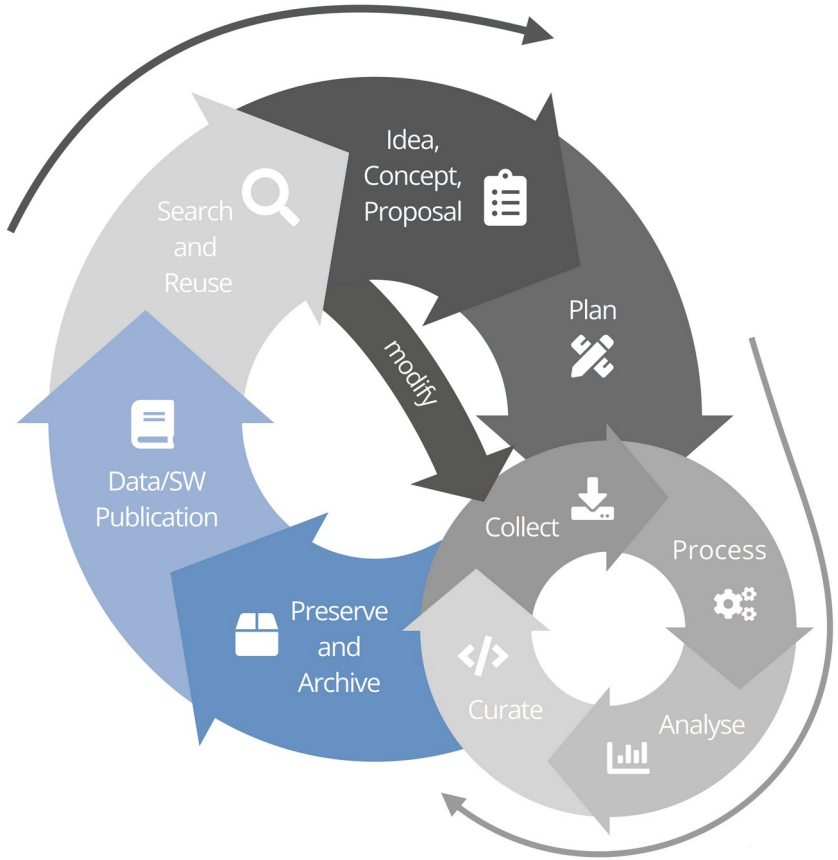
COLLECT

PROCESS

ANALYSE

CURATE

Poster III – phases of the data life cycle



PRESERVE & ARCHIVE

DATA/SOFTWARE PUBLICATION

Tasks

Part I:

Brainstorm how you implement these phases and data flows in your research

Make one post-it per thought about these steps

	Pre-experiment poster(A)	Experiment poster (B)	Post-experiment poster C
Tasks part I	Group 1	Group 2	Group 3
Tasks part II	Group 3	Group 1	Group 2
Tasks part III	Group 2	Group 3	Group 1

Tasks

Part II:

- Read the pre-existing notes
- Categorize the existing notes from the previous groups into **levels of difficulty in implementation (easy, difficult, almost impossible)**

	Pre-experiment poster(A)	Experiment poster (B)	Post-experiment poster C
Tasks part I	Group 1	Group 2	Group 3
Tasks part II	Group 3	Group 1	Group 2
Tasks part III	Group 2	Group 3	Group 1

Tasks

Part III:

- Mark in colour (traffic light system) the challenges/ difficulties for which solutions/services already exist in your working environment
- Add what resources are still missing to overcome the challenges within the present data life cycle phases

	Pre-experiment poster(A)	Experiment poster (B)	Post-experiment poster C
Tasks part I	Group 1	Group 2	Group 3
Tasks part II	Group 3	Group 1	Group 2
Tasks part III	Group 2	Group 3	Group 1

Open discussion

https://www.menti.com/alotqysgjc26

Thank you for your participation & input!

Leibniz-Institut für Geschichte und Kultur des östlichen Europa

TECHNISCHE UNIVERSITÄT DRESDEN

ZIH

DRESDEN concept

Kontaktstelle Forschungsdaten

SaxFDM