Plan Overview

A Data Management Plan created using DMPonline

Title: Dynamics of strongly correlated systems

Creator: Maciej Dendzik

Principal Investigator: Maciej Dendzik

Data Manager: Maciej Dendzik

Affiliation: KTH Royal Institute of Technology

Funder: Swedish Research Council

Template: Swedish Research Council Template

ORCID iD: 0000-0002-4179-0040

Project abstract:

This document presents a data management plan for the BALTAZAR time- and angle-resolved photoemission spectroscopy (trARPES) facility at the Applied Physics department of KTH.

ID: 60434

Last modified: 11-01-2022

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Dynamics of strongly correlated systems

General Information

Project Title

Dynamics of strongly corelated systems

Project Leader

Maciej Dendzik

Registration number at the Swedish Research Council

Version

1

Date

06.07.2020

Description of data – reuse of existing data and/or production of new data

How will data be collected, created or reused?

Data containing dynamic electronic bandstructures of investigated materials will be collected directly from a photoemission analyzer into a local data storage. Raw datasets will be stored in a year-month-day hierarchical file structure. Each collected dataset will be dressed with appropriate metadata describing the investigated material and conditions of the experiment. Furthermore, each dataset will documented in a electronic laboratory notebook (ELOG). Processed data will be stored separately in a project-related file structure. A complete analysis of each processed dataset will be recorded by means of computer scripts. Data will be shared on the open-access NOMAD repository (https://nomad-coe.eu/).

What types of data will be created and/or collected, in terms of data format and amount/volume of data?

Data will be stored in the NXmpes data format (https://github.com/Big-ARPES-data/mpes-nexus) designed for fulfillment of FAIR principles among the photoemission scientific community. Typical datasets of raw and processed data will be approx. GB in size. Such a size of datasets will pose no challenge for sharing and transferring data.

Documentation and data quality

How will the material be documented and described, with associated metadata relating to structure, standards and format for descriptions of the content, collection method, etc.?

Open-access NXmpes file format contains exhaustive metadata information necessary for a meaningful evaluation of the

photoemission data. The complete list of metadata fields can be found online: https://github.com/Big-ARPES-data/mpes-nexus.

How will data quality be safeguarded and documented (for example repeated measurements, validation of data input, etc.)?

Data quality will be safeguarded by repeated measurements and validation of data input parameters. Data will be documented through metadata collection and description in the ELOG notebook.

Storage and backup

How is storage and backup of data and metadata safeguarded during the research process?

Raw data and metadata will be collected without the possibility of modification by a user. Backup of data will be performed on the NOMAD repository.

How is data security and controlled access to data safeguarded, in relation to the handling of sensitive data and personal data, for example?

Not applicable.

Legal and ethical aspects

How is data handling according to legal requirements safeguarded, e.g. in terms of handling of personal data, confidentiality and intellectual property rights?

Not applicable.

How is correct data handling according to ethical aspects safeguarded?

Not applicable.

Accessibility and long-term storage

How, when and where will research data or information about data (metadata) be made accessible? Are there any conditions, embargoes and limitations on the access to and reuse of data to be considered?

Data will be freely-available immediately after publication of the scientific manuscript based on it.

In what way is long-term storage safeguarded, and by whom? How will the selection of data for long-term storage be made?

Long-term storage will be safeguarded by the NOMAD repository.

Will specific systems, software, source code or other types of services be necessary in order to understand, partake of or use/analyse data in the long term?

Open-source software package mpes (https://github.com/mpes-kit/mpes) will enable to analyze and understand the data.

How will the use of unique and persistent identifiers, such as a Digital Object Identifier (DOI), be safeguarded?

DOI numbers will be allocated by the NOMAD repository.

Responsibility and resources

Who is responsible for data management and (possibly) supports the work with this while the research project is in progress? Who is responsible for data management, ongoing management and long-term storage after the research project has ended?

Maciej Dendzik, Oscar Tjernberg and Magnus Hårdensson Berntsen will share the responsibility of the overall data management.

What resources (costs, labour input or other) will be required for data management (including storage, back-up, provision of access and processing for long-term storage)? What resources will be needed to ensure that data fulfil the FAIR principles?

Data management will carried using the existing laboratory infrastructure and resources of the Applied Physics department.