



DataSTaMP

Scientific collaboration on large datasets

2019-12-09

The MAX IV 3 month data retention policy is to be reconsidered



4.2. Curation of Experimental Data

- 4.2.1. All Raw Data will be curated in well-defined formats, for which the means of reading the data will be made available by MAX IV.
- 4.2.2. Metadata automatically captured by instruments will be curated either within the Raw Data files, within an associated On-line Catalogue, or within both.
- 4.2.3. Experimental Data will be read-only while stored at MAX IV.
- 4.2.4. Each data set in the On-line Catalogue will have a persistent identifier (PID). Anybody publishing Results must quote the PID of the data set (and related publications if available and appropriate).

4.3. Access to Experimental Data

- 4.3.1. Experimental Data will be stored in the MAX IV database for a maximum time of three months. MAX IV reserves the right to restrict the storage time, for high data rate instruments. Such restriction of storage time will be decided by MAX IV after consultation with the respective user communities and will be communicated to Users and User Institution in advance of each experiment.
- 4.3.2. Experimental Data will be made available to the User and/or User Institution. MAX IV assumes no responsibility for long-term storage of Experimental Data.
- 4.3.3. Experimental Data will be available via a searchable On-line Catalogue. The On-line Catalogue will enable the linking of Experimental Data to Experimental Proposals via the Experimental Proposal ID. Access to Experimental Proposals and Experimental Data will only be provided to the Experimental Team and MAX IV Staff who needs access, unless otherwise authorised by the PI.
- 4.3.4. The Experimental Team is responsible for ensuring that the Experimental Data collection uses the correct Experimental Proposal ID. Failing to do so may result

Experimental Data Policy

1. Policy purpose
This policy applies to Experimental Data that does not apply to diagnostic data.
2. To whom does this policy apply?
This policy applies to all Experimental Data for which the following will prevail over other policies.
3. Who is responsible?
The directors of the Experimental Team and the

This Policy
may be revised by MAX IV.

Research Council's guidelines for Open Access:
[English/aboutus/activities/analysis/evaluationandfollowup/nationalguaranteedaccessstoresearchinformation.4.18f425dd146e9437d292db46.html](#)
Information Security legislation
[Sakerhet.se/forfattningar/policy](#)
[s/staff.lu.se/files/policy-on-open-access-publication-at-lund-](#)

The following definitions shall apply:
Experimental Data that has been manually or automatically processed and associated Metadata originating from requests.
Document comprising the description of the proposed experiment, which may include the PI, who carry out an experiment, and at MAX IV.

Background

KAW has awarded MAX IV funding towards data storage and data management. Hence, MAX IV has designed a project called DataSTaMP - Data Storage and Management Project set out to improve and **extend the data storage infrastructure**, and to **elevate the research data services** offered to the user community at MAX IV.

- Duration: Aug. 2019 - Dec. 2023
- Total Budget: 75 MSEK
- Work effort: 8.7 FTE (in average)

*Knut and Alice
Wallenberg
Foundation*

User Research Data Services



Common Data Acquisition Service

Generic service for collecting experimental data which is capable of handling data flows from the next generation of high-speed detectors.



Long-term Storage Service

3-tiered data storage service with cold storage of inactive data on cost efficient tape storage and/or in the cloud.



Pre-Processing Service

On-the-fly visualization and inspection of large data volumes generated by high-speed detectors at a beamline.



Remote Analytics Service

Collaborative platform based on JupyterHub, to remotely perform interactive scientific analyses on collected data.



Research Data Service

Portal to search, browse, download data, metadata, publications and any other digital assets related to an experiment.



Advanced User Office Service

Microservices to assist users in the various stages of the user workflow, incl. samples handling, mail-in of experiments and block sessions.

ExPaNDS <-> DataSTaMP

European Open Science Cloud (EOSC) Photon and Neutron Data Service

- Aims to provide coherent FAIR data services to the research community
- Standardize access to all scientific data collected at the photon and neutron facilities
- MAX IV involved in...
 - ExPaNDS-WP4: EOSC data analysis services for PaN national RIs
 - ExPaNDS-WP3: EOSC data catalogue services for PaN national RIs



Remote Analytics Service

Collaborative platform based on JupyterHub, to remotely perform interactive scientific analyses on collected data.



Research Data Service

Portal to search, browse, download data, metadata, publications and any other digital assets related to an experiment.



Advanced User Office Service

Microservices to assist users in the various stages of the user workflow, incl. samples handling, mail-in of experiments and block sessions.

MAXIV

HZDR
HELMHOLTZ ZENTRUM
DRESDEN ROSENDORF

PAUL SCHERRER INSTITUT
PSI

HZB Helmholtz
Zentrum Berlin

SOLEIL
SYNCHROTRON

Partners

ExPaNDs

egi

UKRI

q
Elettra Sincrotrone Trieste

ALBA

DESY.



diamond

MAXIV

MAXIN

The image shows the word "MAXIN" in a bold, grey, sans-serif font. A bright yellow, curved swoosh underline starts under the 'M', loops under the 'A' and 'X', and ends under the 'N'. The swoosh has a slight curve and a thin tail at the end.