Data management policy at ICTS-IABA (CNA-CMAM)

This document describes the scientific data management policy *that* applies to the distributed ICTS IABA, composed by two different infrastructures, CNA (Centro Nacional de Aceleradores) and CMAM (Centro de Microanálisis de materiales). CNA is a mixed centre of the Universidad de Sevilla, the Junta de Andalucía and CSIC. CMAM is a centre of the Universidad Autónoma de Madrid.

This data management policy will be implemented gradually, on a best effort basis, and according to available resources.

1. General principles

1.1 This policy on data management outlines the rights to access and ownership of the experimental raw data and metadata collected and/or maintained by CNA and CMAM.

1.2 Agreement with this policy is required for the allocation of public beam time.

1.3Users are prohibited from accessing, exploiting, or distributing data or metadata as specified in this policy unless they have the proper authorization.

1.4 Intentional violations of the policy may result in the loss of access to data and metadata and exclusion from future proposal opportunities.

1.5 All data and metadata will be managed in compliance with Spanish Data Protection legislation (Ley Orgánica de Protección de Datos, LOPD).

1.6 The approval and modification of this document correspond to the government bodies of CNA and CMAM. The legal representative of CNA is the vicerrector of research of the University of Seville (US). The legal representative of CMAM is the vicerrector of research of the Universidad Autonoma de Madrid (UAM).

2. Definitions

Experiments performed at CNA or CMAM can be divided into several categories: public access (which includes regular user calls, friendly users, training experiments,

in-house research, access through formal collaboration agreements, and contingency beam time managed by the facility), access obtained through specific partnership

agreements, and proprietary access (which includes contracts with companies or particulars stating proprietary access).

This policy applies to experiments with public access and those conducted under specific partnership agreements. Experiments under proprietary access are conducted with confidentiality regarding the proposal, raw data, and results, and are therefore not included in this policy.

For the purposes of this policy:

2.1 The CNA and the CMAM are considered as two distinct facilities.

2.2 "Raw data" refers to the information obtained from experiments conducted using beamlines or other instruments at CNA or CMAM. This includes data produced either automatically or manually through facility-specific software or expert personnel to support the subsequent analysis of the experimental results, unless specified otherwise.

2.3"Metadata" refers to the information related to the data collected from the instruments. This includes, but is not limited to, details about the experiment's context, the experimental team, the conditions under which the experiment was conducted, and other logistical aspects. The information provided in the experimental proposal will be part of the metadata.

2.4 The "Principal Investigator" (PI) is the person listed as the PI on the proposal of the experiment.

2.5 The "experimental team" consists of the PI and any other individuals who are granted permission by the PI to access the resulting raw data and associated metadata.

2.6 When the PI is not a researcher in CNA or CMAM, a “contact person” (CP) of these institutions will be designated, which will be a member of the “experimental team”.

2.7 The "online catalogue" is a digital database of metadata that includes links to raw data files. It can be accessed through various methods, such as web-based browsers, among others.

2.8 "Results" refers to the data, intellectual property, and findings produced from the analysis of raw data.

2.9 "Open access" refers to content that is available to the general public, not protected by copyright or patents, and can be used by anyone. It implies unrestricted, but not anonymous, access at no cost. An identification procedure may be required,

Data considered “Open Access” will be provided under the CC-BY license (Creative Commons BY, http://creativecommons.org/licenses/by/4.0/legalcode).

3. Raw data and associated metadata

3.1 Access to raw data and associated metadata

3.1.1 Raw data and metadata from public access or specific partnership agreements experiments, as described in section 2, will be made available to the public after an initial embargo period. During the embargo period, access will be restricted to the experimental team, led by the PI.

3.1.2 US and UAM will provide custody of the raw data and associated metadata of CNA and CMAM, respectively, for at least 5 years after the embargo period.

3.2 Curation of raw data and associated metadata

3.2.1 Raw data and metadata may be curated at CNA and CMAM facilities or elsewhere. When curated, the data will be organized in clearly defined formats, and

CNA and CMAM will provide the necessary tools or methods for accessing the curated data.

3.2.2 Raw data and metadata will be read-only for the duration of their lifetime.

3.2.3 Each experiment conducted at the CNA or CMAM is expected to receive a unique persistent identifier. Anyone publishing results derived from open access data must reference the corresponding identifier and any related publications, if available and necessary.

3.2.4 High-level metadata from CNA or CMAM proposals, including Title, Authors, Abstract, and Beamline, will be publicly accessible once the experiment is completed. This information will be available on a specific experiment site on CNA or CMAM web pages, along with the persistent identifier and the subsequent publications.

3.3 Access to raw data and metadata

3.3.1 Access to raw data and metadata after the embargo period will be facilitated via appropriate listing and search tools.

3.3.2 Access to data will be limited to registered users. Retrieving data may involve accessing permanent storage, which could result in potential delays.

3.3.3 Access to raw data and associated metadata from CNA or CMAM experiments is limited to the experimental team for a period of 3 years (embargo period). Once this concludes, the data and metadata will be made publicly accessible. If a PI wishes to extend this period and keep their data private for a longer duration, they must submit a special request to the Director of the correspond infrastructure.

3.3.4 The PI, assisted by the CP, is responsible for ensuring that data is stored in the designated directories and that the experiment number is accurately recorded in the metadata for each raw data set. Failure to do so may result in the experimental team being unable to access the data or other users accidentally gaining access rights to it.

3.3.5 Authorized staff from CNA and CMAM have access to curated data and metadata for purposes related to facility operations. The staff will ensure the confidentiality of CNA or CMAM data during the embargo period.

3.3.6 The PI has the possibility to transfer parts of the totality of her/his rights during the embargo period to another registered person.

3.3.7 The PI has the possibility to create and distribute copies of the raw data.

4. Results

4.1 Ownership of results

4.1.1 The ownership of all results (intellectual property) derived from the analysis of raw data is governed by the contractual obligations of those conducting the analysis and the copyright agreements established at the time of publication.

4.2 Storage of results

4.2.1 The PI may request to have the results of the analysis stored at CNA and CMAM facilities. However, CNA or CMAM will not be responsible for ensuring that the necessary software for reading or manipulating the results obtained from the analysis of data is available.

4.2.2The US or the UAM cannot be made liable in case of unavailability or loss of data, results or data analysis software.

4.3 Access to results

4.3.1During the embargo period, access to the results of analysis performed on raw data and metadata is limited to the experimental team responsible for the analysis, unless the PI request otherwise.

4.3.2Authorized facility staff, such as instrument scientists and computing group members, have access to facility-curated data and metadata for operational purposes. CNA and CMAM will ensure the confidentiality of this data is maintained throughout the embargo period.

5. Good practice for data and metadata capture and usage

5.1 The experimental team is encouraged to make metadata as comprehensive as possible. This will improve their ability to search for, retrieve, and interpret their data in the future.

5.2 CNA and CMAM commit to making best efforts to provide tools for capturing metadata items that are not automatically recorded by instruments. This ensures the most comprehensive description of raw data is recorded.

5.3 Researchers intending to analyze openly accessible raw data and metadata should, whenever possible, reach out to the original PI to inform them and propose a collaboration if necessary. They must also acknowledge the data source, cite its unique identifier, and reference any related publications.

5.4 PIs and researchers conducting analyses on raw data and metadata are encouraged to utilize the available on-line tools to link their analysis results with the corresponding raw data and metadata. Additionally, they are encouraged to make these analysis results publicly accessible.

6. Publication information

6.1 Publications associated with data from experiments conducted at CNA or CMAM must include the persistent identifier of the experiment and the data in their citations.

6.2 References for publications stemming from experiments at the CNA or CMAM must be submitted to the experiment site on CNA or CMAM web pages, within six months of publication or upon any new beamtime application, whichever occurs earlier. Failure to comply may affect future proposal time allocations.

IABA
ACCELERATOR-BASED
APPLICATIONS INFRASTRUCTURE

ICTS
INFRAESTRUCTURAS CIENTÍFICAS
Y TÉCNICAS SINGULARES

Accelerator-Based Applications Infrastructure ICTS (IABA)

Unique Scientific and Technical Infrastructure, according to the latest ICTS map
from the Ministry of Science and Innovation