Experimental Data Retention Policy of Public Experiment Data Systems at SPring–8/SACLA

February 22, 2020

Amended March 3, 2020

SPring–8 Data and Network Committee

1. Purpose

At SPring-8 and SACLA, many kinds of research and development are conducted using the world's highest performance synchrotron radiation. Experimental data is expected to increase dramatically following the upgrade, which includes the enhancement of detector performance. RIKEN is responsible for developing Public Experiment Data Systems at SPring-8/SACLA and managing data appropriately in order to efficiently utilize experimental data, which is important intellectual property.

For these reasons, the SPring–8 Data and Network Committee (hereinafter the "Committee") has established the following basic policies regarding the retention of experimental data in Public Experiment Data Systems at SPring-8/SACLA.

2. Target Data and its Ownership

The data covered by this policy shall be recorded in Public Experiment Data Systems by electromagnetic means. The data types and their ownership are defined as follows:

(1) Data Types

Data Type A: Data measured by the monitoring equipment of the facility, such as light source performance.

Data Type B: Measured data derived from samples brought in by the Users, including meta-information related to samples.

Data Type C: Secondary data obtained by Users processing Data Type B with computer resources.

(2) Data Ownership

The ownership of Data Type A belongs to the facility, and the ownership of Data Types B

and C belongs to the relevant User. However, Users can record and retain Data Type A in the data formats of Data Types B or C.

3. Basic Principles of Public Experiment Data Systems

(1) Responsibility for Data Retention

All responsibility for managing data provided by the facility shall be borne by the User in accordance with the data management policies of the organization to which the User belongs. It is practically impossible to guarantee data retention due to the human and financial resource constraints of the facility. Therefore, the User should be aware that the facility cannot guarantee data retention while using Public Experiment Data Systems.

(2) Public Experiment Data Systems at SPring–8/SACLA

The facility must develop various Public Experiment Data Systems depending on the human and financial resources at hand in order to promote efficient utilization of data. The facility can set a data retention period, capacity limit, usage period, etc. for each Public Experiment Data System. When establishing various restrictions and conditions, the facility must disclose them on the official website or by similar means. The facility may change these restrictions depending on the human and financial resources at hand, as well as the usage situation of the system, etc. Upon changing restrictions and conditions, the facility must promptly inform the experiment Users. If the Committee approves a data system as a system of Public Experiment Data Systems that handles large amounts of data, the facility shall operate the data system in accordance with the guidelines provided in the annex.

4. Revisions to This Policy

This policy may be revised by a majority vote of the Committee. However, if one or more members of the Committee deem it urgent, this policy may be revised by obtaining the approval of a majority of the members by e-mail or other means.

5. Effective Date for This Policy

Application of data retention under this policy becomes effective from May 1, 2020.

Annex

Operational guidelines for Public Experiment Data Systems that handle large amounts of data

1. Basic Principles

For a system of Public Experiment Data Systems that handles large amounts of data, the data system must have a hierarchy in order to efficiently process data. For that reason, the facility shall operate the system on the following principles:

i. Hierarchy in the data system

The system shall consist of three layers: short-term storage with high-speed access (disk devices, etc.), medium-term storage with medium-speed access (disk devices, etc.), and long-term storage with low-speed access (tape devices, etc.). The Operation Manager should appropriately manage Data Type A and Data Type B using these three layers. The User may retain Data Type C in medium and low-speed storage by submitting an application to the Operation Supervisor.

ii. Rules regarding Retention Periods

The data retention period for medium-speed storage shall be determined by the Operation Manager in consideration of the period of Publication of Results. The retention period for high-speed and low-speed storage should be determined by the Operation Manager depending on the capacity of the storage. Data that has exceeded the retention period must be deleted without notice by the Operation Manager.

iii. Limits on Capacity

The data capacity that can be used by each User in medium-speed and low-speed storage shall be determined by the Operation Manager depending on the total capacity.

iv. User Accounts

The User can access data in medium-speed storage from an external network. The expiry date for the account of the experiment User for this purpose is determined by the Operation Manager in consideration of the period of Publication of Results.

2.   Revision of These Guidelines

These guidelines may be revised by a majority vote of the Committee. However, if one or more members of the Committee deem it urgent, these guidelines may be revised by obtaining the approval of a majority of the members by e-mail or other means.