

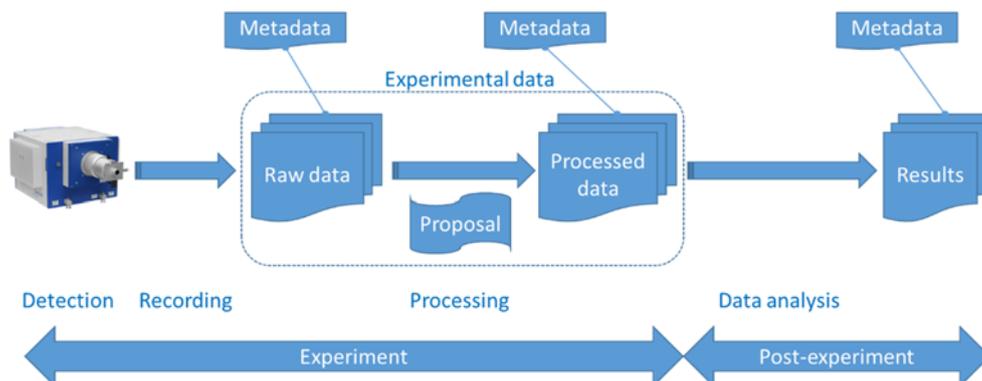
The APOLLON Data Policy

The *Laboratoire pour l'Utilisation des Lasers Intenses* aims to implement a Data Policy for its APOLLON laser Research Infrastructure starting as soon as possible in 2022. It provides a framework for storage and access to experimental data and associated metadata collected at the facility in order to facilitate production by its users of FAIR (Findable, Accessible, Interoperable, Reusable) results.

This policy follows recommendations issued by the Photon and Neutron Open Science Cloud (PANOSC) EU project (INFRAEOSC-04-2018).

1. Definitions

- 1.1. For the purposes of the present policy, the term **facility** refers to the APOLLON laser research infrastructure.
- 1.2. The term **experimental data** refers to data collected from peer-reviewed and in-house experiments conducted at the facility. It covers raw and processed data as defined below, as well as the experiment's proposal.



- 1.3. The term **raw data** refers to experimental data directly recorded from diagnostics or instrumentation owned by the facility, and not derived from any other persistent data or recorded from user-supplied equipment.
- 1.4. The term **processed data** refers to experimental data derived from raw data through any processing process, performed automatically or manually by facility software and/or staff, in order to facilitate subsequent analysis.
- 1.5. The term **metadata** describes information pertaining to experimental data collected at the facility, including (but not limited to) the context of the experiment (as the abstract of the proposal), the experimental team, the experimental conditions, the electronic logbook generated during the experiment (if any), and any other contextual

information (such as sample information) possibly useful for subsequent data treatment.

- 1.6. The term **results** pertains to intellectual property and outcomes, incl. algorithms, arising from the analysis of experimental data. This does not include publications.
- 1.7. The term **principal investigator** (PI) refers to the proposer identified on the experiment proposal as the key person interacting with the facility staff (for the peer-reviewed experiments). For the in-house experiments, the PI is the scientist initiating and performing the experiment
- 1.8. The term **experimental team** includes the PI and any other person to whom the PI assigns the right to access the experimental data.
- 1.9. The term **beam time** refers to the period of time when the experimental team has access to the facility resources to conduct an experiment.
- 1.10. The term **shot sequence** refers to a series of consecutive laser shots without any change of the experimental setup.
- 1.11. The term **users** refers to the members of all the experimental teams having obtained access to beam time.
- 1.12. The term **embargo period** refers to the period during which the users have exclusive access to the experimental data.
- 1.13. The term **public research** refers to publicly funded research which has been allocated peer-reviewed or in-house access to the facility and which is intended to lead to publication(s).
- 1.14. The term **proprietary access** refers to research done through purchased (commercial) or classified access. The term **client** refers to the structure (institution, company, etc.) on behalf of which the PI requests proprietary access.
- 1.15. The term **short-term** means a maximum of 1 year. The term **long-term** means a minimum of 5 years and a maximum of 10 years.
- 1.16. The term **curation** denotes the long-term storage, backup and protection of experimental data in a manner that respects and guarantees the prescribed access rights.
- 1.17. The term **open data** means that the data are made available for use by the community at large, without fees, copyright or patent restrictions.
- 1.18. The term **DMP** refers to a Data Management Plan enabling the clarification of all aspects of the data management life cycle between the facility and the users before the experiment takes place. It describes which data are collected, processed and/or generated, which methodology is applied, how data are curated, which data format is used, etc.
- 1.19. The term **custodian** refers to the facility storing, curating and providing access to experimental data and associated metadata.

2. General Principles

- 2.1. The present data policy governs the ownership of, the curation of and access to experimental data - and associated metadata - collected and/or stored at the facility.
- 2.2. Acceptance of this policy is a condition for the allocation of beam time.
- 2.3. Users must not attempt to access, exploit or distribute experimental data, and associated metadata, unless they are entitled to do so under the terms of this policy.
- 2.4. Deliberate infringements of the policy may lead to denial of access to experimental data, and associated metadata, and/or denial of future beam time requests.
- 2.5. All experimental data, and associated metadata, shall comply to the General Data Protection Regulation (GDPR)¹.
- 2.6. The facility will release open experimental data under a CC-BY-NC license (<https://creativecommons.org/licenses/by-nc-sa/4.0/>). Such a license prevents commercial use and obliges citation of the facility and of the permanent identifier of the associated experiment.

3. Ownership

- 3.1 Experimental data, and associated metadata, resulting from proprietary access is owned by the client. They are exclusively accessible by the experimental team, represented by the PI, and are not covered by the present policy. They are removed from the facility storage capacities within the month following the experiment, unless otherwise agreed with the facility management before the start of the experiment.
- 3.2 Experimental data, and associated metadata, resulting from public research, become open data after an embargo period during which access is restricted to the experimental team.
- 3.3 The intellectual property rights for results derived from analysis of experimental data are determined by the contractual obligations of the person(s) performing the analysis.

4. Roles and Responsibilities

- 4.1 Depending on his/her funding agency and/or institution, the PI may be required to establish a DMP. The facility shall provide any necessary information not specified in the present policy.
- 4.2 The PI shall ensure experimental data be collected with accurate metadata, including sample description, such that they fulfill the FAIR principles.
- 4.3 The facility provides means to collect such metadata items if not automatically captured in order to facilitate recording the fullest possible description of the experimental data.

¹ Some descriptive metadata (e.g. authors, date, facility, instrument, and abstract of experiment) are considered necessary for the publication of data and therefore covered by Article 85 of GDPR, which exempts them from the right to be removed unless they are incorrect.

-
- 4.4 The PI is strongly encouraged to provide a complete log of the protocol carried out and what happened during the experiment. The logs must be entered in the facility electronic logbook, if the facility provides one. In the absence of a facility electronic logbook, the experimental team must use other means (electronic if possible) and ensure that this logbook be linked to the experimental data.
 - 4.5 The PI has the possibility to transfer, at any time, parts or the totality of her/his rights to any member of the experimental team.
 - 4.6 The facility applies all reasonable efforts to ensure an accurate curation of experimental data as well as an uninterrupted access to them. However, failures caused by technical or human mistakes cannot be ruled out. The facility cannot warrant an absolutely accurate storing and curating. Access to data might be temporarily limited or impossible, especially due to necessary maintenance, service updates or failure of third-party service providers.²
 - 4.7 The facility cannot be made liable in case of unavailability or loss of data or data analysis software.
 - 4.8 In case of data breach, the facility shall timely notify the impacted experimental teams and cannot be made liable if reasonable security measures have been implemented.

5. Persistent identifiers

- 5.1. Each experiment has a unique persistent identifier, for example an Digital Object Identifier (DOI). Such persistent identifier is created and provided by the facility. It incorporates at least the year of acceptance, the experimental area in which the experiment will be conducted and the proposal number.
- 5.2. Users shall cite the persistent identifier in any publication that refers to the experiment data (or a subset of them).
- 5.3. High-level metadata, such the title, the PI, the summary and the persistent identifier, are made public through a dedicated landing page on the facility website, as soon as the experiment is accepted (<https://apollonlaserfacility.cnrs.fr/en/selected-projects/>).

6. Curation of experimental data

- 6.1. Experimental data, and associated metadata, are migrated from short-term facility servers (on which analysis can be performed by the experimental team) to long-term archival devices. However, depending on the type and volume of experimental data concerned, and the economic consequences associated with long-term data storage, the facility reserves the right to restrict the curation period or the set of experimental data archived in consultation with the users.
- 6.2. Experimental data are curated in well-defined formats, for which the means of reading the data are made available by the facility.
- 6.3. Only experimental data with associated metadata are archived.

² All data management and processing services are provided on a 'best effort' basis.

-
- 6.4. Experimental data, and associated metadata, are read-only for the duration of their life time.

7. Access to experimental data

- 7.1. The embargo period, during which access to experimental data, and associated metadata, is restricted to the experimental team, is set to 3 years, unless otherwise decided by the PI, in agreement with the facility management.
- 7.2. The embargo period begins at the end of the experiment.
- 7.3. Access details to the experimental data, and associated metadata, are provided on the landing page associated to the experiment's persistent identifier.
- 7.4. Authorized facility staff have access to experimental data, and associated metadata, curated by the facility in order to provide support to users. The facility reserves the right to use data still under embargo to improve facility processes and performance, however preserving their confidentiality.

8. Results

- 8.1. The facility does not provide curation of results.
- 8.2. The experimental team carrying out analyses of experimental data is encouraged to link complementary data, collected from user-supplied instrumentation, software tools used and results produced to the experiment's logbook.
- 8.3. PIs and researchers are encouraged to make software tools and results openly accessible.
- 8.4. A researcher who aims to re-analyze open experimental data is encouraged to contact the original PI to inform her/him and suggest a collaboration, if appropriate. The source of the experimental data must be duly acknowledged and the permanent identifier of the corresponding experiment cited, as well as and any publication linked to the same experimental data.

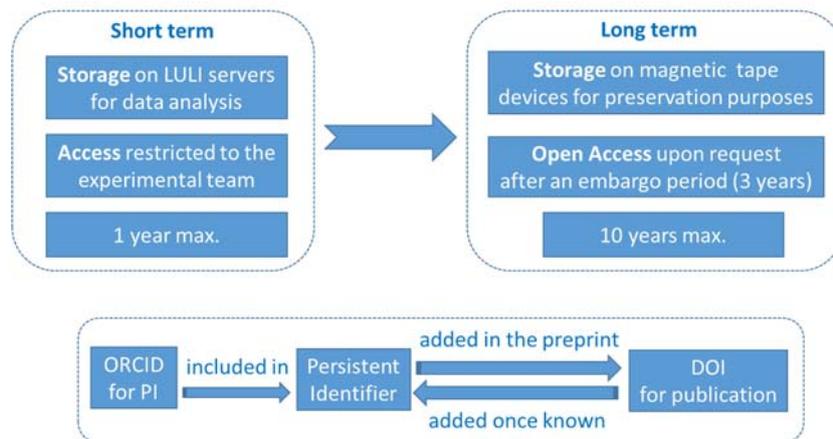
9. Publications

- 9.1. The permanent identifier of the experiment must be cited in any publication that refers to related experimental data. In addition, the facility must be duly acknowledged.

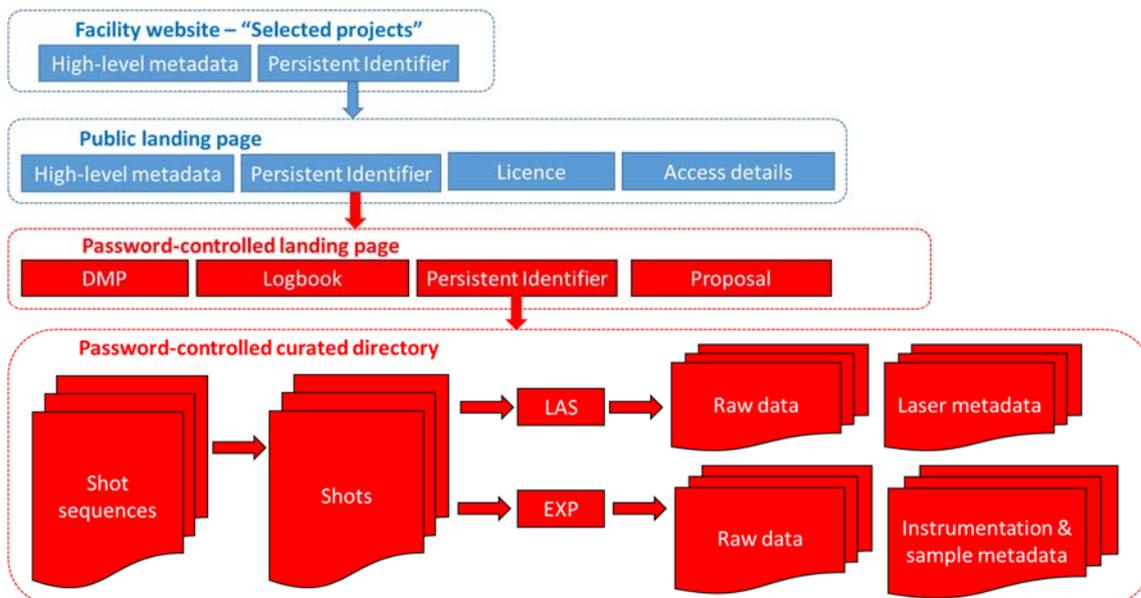
Preferred version: "The authors acknowledge the facility and the technical assistance of the national Research Infrastructure APOLLON."
- 9.2. The DOI of any publication that refers to experimental data acquired at the facility shall be sent to the facility in order to be posted on the corresponding landing page.
- 9.3. Users are strongly encouraged to follow best practices adopted by many journals concerning citing the software used and developed for the analysis of the experimental data.

10. Termination of custodianship

10.1. If the facility may decide to stop acting as custodian, it shall inform the PIs concerned in a timely manner and provide them with effective means to export their experimental data and associated metadata. To ensure a fair termination, if decided, the facility shall be aware of any change in the contact information (e.g. email addresses) of the PIs.



Synopsis for data curation & citation.



Access route to experimental data and associated metadata.