

CREMLINPlus WP8.6:

Access to scientific data at Russian RI

ICISTE, DESY, NUST MISIS, NRC KI

Self-assessment questionnaire on access policy

21 January 2022



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 871072.



CREMLIN P^LUS

QUESTIONNAIRE MAIN GOALS

- ◆ **1. Determine the current state of a data lifecycle on Russian RI**
 - Data ownership;
 - Data access;
 - Data processing;
 - Metadata access;
 - Publications policy;

- ◆ **2. Learn the specific features and limitations of data lifecycle in Russian RI**
 - Please report unavoidable data access restrictions on your RI

- ◆ **3. Establish a baseline to work towards the objectives of Task 8.6**
 - Develop harmonized data lifecycle approaches that are convenient for users and Russian RI.
 - Develop a common understanding on using the FAIR principles.
 - Finding ways for collaborations with EOSC and European H2020 projects (ESCAPE, PaNOSC).



GENERAL RECOMMENDATIONS

- ❖ **1. Have a Data Management Plan with defined strategy to**
 - Data collection;
 - Data storage, structure and organization;
 - Data access;
 - Measures to preserve data integrity in case of technical failures;

- ❖ **2. FAIR Guiding Principles (<https://www.go-fair.org/fair-principles/>)**
 - Findable. Metadata and data should be easy to find for both humans and computers.
 - Accessible. Once the user finds the required data, it is known how can they be accessed.
 - Interoperable. The data need to be integrated with other data. The data need to interoperate with applications or workflows for analysis, storage, and processing.
 - Reusable. To optimise the reuse of data, metadata and data should be well-described.

- ❖ **3. Online access to the data and aim for open-access.**



Description of the data ownership policy

- ❖ **1. Does your research infrastructure have a data ownership policy?**
 - Yes/No;
- ❖ **2. Who owns raw data obtained during experiments on your research infrastructure?**
 - Research Infrastructure;
 - Users
- ❖ **3. Who owns data produced during analysis of raw data obtained on your research infrastructure?**
 - Research Infrastructure;
 - Users
- ❖ **4. What obligations do users assume after completing an experiment on your research infrastructure?**
 - Report;
 - Acknowledgements in publications;
 - Include RI staff as authors;
- ❖ **5. Does the data become publicly available as open access after a specified period after the experiment?**
 - No;
 - Yes, immediately;
 - Yes, after embargo period;



Description of the data proposal management

- ❖ **1. Does your research infrastructure have a unified system for initiating, managing and tracking experimental proposals?**
 - Yes;
 - No;
 - Other;

- ❖ **2. How does your research infrastructure track publications based on experiment results?**
 - Does not track: my research infrastructure does not track publications;
 - By users: users must register publications in the unified system;
 - Manually: publications are tracked manually by research infrastructure staff;



Description of the data access policy

- ❖ **1. Where is the experimental data stored on your research infrastructure?**
 - Separately: experimental data are stored separately in storage systems of experimental stations;
 - Central storage system: experimental data are stored in a central storage system;
- ❖ **2. For how long are copies of the experimental data kept in storage systems in your research infrastructure?**
 - No specified period: experimental data are removed when free space on the storage system runs out;
 - No specified period: experimental data are kept "forever";
 - For a specified period: experimental data are removed after a specified period;
- ❖ **3. What measures are taken to ensure the safety of experimental data in storage systems?**
 - Multiple copies: several copies of the experimental data are saved on different storage systems;
 - Reliable storage systems are used that allow hardware failures without data loss;
- ❖ **4. Do experimental datasets have unique and persistent identifiers?**
 - Yes: experimental datasets have unique and persistent identifiers on my research infrastructure;
 - No: experimental datasets don't have unique and persistent identifiers on my research infrastructure;



Description of the data access policy

❖ 5. How experimental data are transferred to users?

- Portable storage: experimental data are copied to the user's portable storage after the experiment;
- Temporary link: users download experimental data from a temporary online link;
- Permanent link: experimental data are available for download at the permanent online link;

❖ 6. Whether other users can find and access experimental data?

- No: only the experiment team has access to experimental data;
- Cannot find but can access: users can download experimental data if they know the exact link;
- Yes: experimental data can be found and downloaded from a public website;

❖ 7. What protocols are used to transfer experimental data?

- FTP;
- SSH/SFTP;

❖ 8. What data formats are used to store experimental data? (Please provide a list)



Description of the data processing policy

- ❖ **1. Does your research infrastructure have standard processing workflows that apply to all experiment data?**
 - No: there are no standard processing workflows on my research infrastructure;
 - Other (Yes): please briefly explain;
- ❖ **2. Does your infrastructure have automatic validation and integrity checks for experimental data?**
 - No: there are no automatic validation and integrity checks on my research infrastructure;
 - Other (Yes): please briefly explain;
- ❖ **3. Are users given access to computing resources to analyze experimental data?**
 - No: users analyze data on their own computing resources;
 - Yes: users are given access to computing resources for analysis;
 - Other (Yes, but with restrictions): please briefly explain;
- ❖ **4. If users are given access to computing resources to analyze experimental data, is the processed data managed/stored the same way as the experimental data?**
 - No;
 - Other(Yes): please explain



Description of the metadata policy

- ❖ **1. How is metadata collected on my research infrastructure?**
 - Manually: metadata values are collected and stored manually;
 - Automatically: metadata values are collected and stored automatically by the experiment management system;
- ❖ **2. How is metadata stored on my research infrastructure?**
 - Separate files: metadata is stored as separate files;
 - Database: metadata is stored in a dedicated database;
 - With data: metadata is stored along with data in the form of archives;
- ❖ **3. How is metadata transferred to users?**
 - With data: metadata is transferred along with data;
 - Web: metadata can be found on a dedicated webpage;
- ❖ **4. Does your research infrastructure use standard metadata schemas?**
 - No;
 - Other(Yes): please briefly list used schemas;



THANK YOU FOR YOUR ATTENTION

ICISTE, DESY, NUST MISIS, NRC KI



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 871072.