Data Management Plan of the Institute of Scientific Instruments, Czech Academy of Sciences

1. Data Summary

For the purposes of this Plan, we consider as "data" relevant to this Plan only those datasets that must be made available under the rules and principles of open access to scientific data ("open data"), data that may become mandatorily available data under the current rules in the future, and data created for the benefit of non-institutional entities in the context of contract research or infrastructure activities. All other data collected or processed in the course of the research activities of the Institute of Scientific Instruments of the CAS, v. v. i., (hereinafter referred to as "ISI") are considered to be internal data, which are either of a non-important nature (i.e. they do not have to be disclosed to support the conclusions of the research) or are data for which there is a reason preventing disclosure, i.e. a valid exemption.

Due to the technological and instrumental focus of the institute, an important part of the research is design work in the field of electrical engineering, mechanical engineering, vacuum technology, etc. The output, often the main output, is therefore data (information) of the nature of design solutions. This may be engineering drawings, electrical schematics, software, etc. The actual data in the usual sense of the term data can then be data measured by means of apparatuses and their primary purpose is to verify the functionality and parameters of the designed instrumentation or subsystems. Such data are of fundamental importance for the further instrumental development of the ISI, but their relevance for external stakeholders is very limited, mainly due to their origin in specific instrument platforms of our own design. Their comparability and reproducibility to off-site origins is therefore very limited.

The data formats worked with at the Institute can be very diverse, reflecting the wide technological range of the Institute. The data can be categorised broadly as follows:

1. Data of traditional formats, which can be processed by common, widely used software. This category includes numeric and text files, image data in common formats, etc.

2. Data obtained by proprietary software that is part of special but commercial instruments operated at the Institute.

3. Data recorded by instruments constructed at the ISI and stored in a format specific to the particular apparatus.

4. Data of a design nature.

The use of data in the internal environment and its management for the research needs of the research teams of individual projects is under the control of the departments, groups and researchers.

2. FAIR data

Making data findable, including provisions for metadata

Metadata is an essential part of the stored data files and carries information about the time of creation of the data file, the circumstances of its recording, the type of experiment and the project within which it was taken. Metadata related to their interpretation may in many cases include intellectual property elements and their disclosure or transfer to external parties is subject to the rules for disclosure of data with application potential (see below).

Due to the broad disciplinary scope of the ISI, other details of the metadata structure are specifically linked to the practices of the respective disciplines. They are not, and cannot be, nominally addressed in this data management plan simply because of the variability of individual projects and their specific requirements.

With few exceptions, ISI data are not of the nature of data collected on sets of objects or samples that would have relevance for external machine processing (data harvesting). They will therefore not be modified for this purpose. In exceptional cases, on request, selected data sets may be so modified if required by an external entity for a meaningful purpose and no risk of intellectual property compromise or compromise of application potential is identified.

Making data accessible

The forms of access to the data of the UPT can be categorised according to the nature of the subject to whom they are to be made available:

1. Access to the ISI research team

    Access to data for the internal needs of the research teams of projects carried out at the ISI is full and unrestricted. Access to data repositories shall be protected by standard normal data security procedures. Metadata reflects the clarity needed internally within the research team.

2. Accessibility for ISI staff outside the project team

    Disclosure of data to persons outside the specific project team within the ISI shall be full as in point 1, unless the project team is contractually bound by confidentiality (NDA).

3. Disclosure to members of a project consortium with which the ISI is bound by a project agreement

    In the case of data generated by projects involving external co-investigators, the forms of data storage, the nature of the metadata and the forms of data sharing are determined by the contracts with the partners or the terms of reference of the individual consortium projects. These project and contractual requirements within these specific projects have priority over this data management plan.

4. Access to external bodies

    Data that are identified as data for which disclosure to external parties is permitted are disclosed on a "upon reasonable demand" basis - upon a meaningful request from an external party. The actual procedure for submitting a data transfer request and satisfying

it uses a web form on the ISI website. The request is directed to the Data Steward who will forward it to the relevant Institute staff member. The decision as to whether the request is 'reasonable' rests with the project investigator of the project in which the data was generated. If necessary, the researcher shall consult with the group or department head.

Disclosure of data to external parties is not anonymous; the request for data is subject to filling a questionnaire including the name, affiliation and e-mail address of the applicant.

The forms of access to ISI data can also be categorised according to the nature of the project within which it was produced and its outputs:

1. Publication-related data release

   Publication-related data are subject to the rules of the publishers and are made available in the appropriate form requested by the publishers.

2. Access to data generated by projects of different providers

   Also the tender documents and rules of the providers of project funding in the form of projects lay down their own rules for handling data. If these rules are not in line with this data management plan, for specific individual projects they override this plan.

Repository

The trustworthiness of data repositories and repositories is a subject of great attention and care. The internal data repositories, which are part of the Institute's computer network infrastructure, are essential for the management of the Institute's data. It is the responsibility of the Information Technology Department of the ISI to ensure proper backup, security and operational reliability. Individual scientific departments have dedicated data storage space and use it in a mode appropriate to their own needs. The further development and strengthening of internal data storage facilities, both in terms of increasing the volume of data to be stored and in terms of strengthening security, is one of the priorities of the ISI.

In addition to the internal data repositories, the Institute uses a number of other repositories and storage facilities. These can be categorised as follows:

1. Domain-specific data repositories, widely used by the international community, with credibility given by the breadth of their use in science. For example, arXiv or ZENODO.

2. Trusted national repositories offered by national infrastructures for science and research, namely EOSC-CZ, KNAV-ASEP, CESNET.

Repositories of publishers of scientific publications, whose use for publication-related data is an obligation for authors.

4. Domain-specific repositories dedicated to the narrower international scientific community, the use of which is usually necessary in the context of specific consortium projects.

The use of other data repositories and repositories is possible after due verification of their credibility and with the approval of the head of the relevant scientific department.

Data

With regard to the technological focus of the institute, many of the results have application potential and/or are created directly in collaboration with partners from the application sphere. Therefore, there is a possibility of future commercialisation or transfer into practice in a broader sense. Access to the data is then restricted or embargoed to ensure the protection of intellectual property. Understandably, it is necessary to deny access to any research data that may be used by a competitor to undermine IP protection or to prevent protection from being established (e.g. preventing patenting or licensing). This therefore means that all research data relating to potentially commercially useful and applicable outputs are considered to fall outside the category of research data covered by the open access principles.

Non-disclosure of data - data shall not be made available in the following cases:

1. Applied research project data

    This restriction does not apply to data related to publications produced by these projects.

2. Data from non-publicly funded contract and collaborative research projects and data generated under "Other Activity" (In Czech: “Jiná činnost”).

3. Data with application potential and data involving intellectual property elements

    The decision as to whether a dataset has application potential or contains an intellectual property element rests with the investigator of the project within which the data was generated, and consultation with the group or department leader is worthy of recommendation.

4. Data bound by non-disclosure agreements (NDAs)

5. Data of constructional nature - engineering drawings, electrical schematics, software (except software distributed in the form of Open Source)

6. Data that has been created by proprietary software that is bound by restrictions in the licensing arrangements for that software and associated metadata

7. Data with ethical constraints

8. Data that violates the principles of personal data protection (GDPR)

Making data interoperable

Data interoperability is essential for sharing, exchanging and reusing data. The primary imperative is to ensure interoperability within project teams for internal use. Therefore, the data are provided with adequate metadata and an understandable description. The formats and standards are consistent with the practices of research fields and the needs of the research teams. Project investigators are encouraged to use the standard metadata suggested by The Research Data Alliance, unless otherwise specified in the project specifications or contractually in project collaboration agreements. Standard vocabularies will be used for all types of research data generated in the project.

Interoperability of data relevant to scientific publications will be technically ensured by depositing a machine-readable electronic version of the published publication or the final version of a peer-

reviewed manuscript accepted for publication in the relevant repositories. In legal terms, open access to research data deposited in these repositories under the terms of the latest available version of the CC BY public licence or its equivalent.

Increase data re-use

The internal usability of the data is ensured by storage in internal data repositories, storage in standard formats used by research teams, laboratory apparatuses in operation, related software and provision of adequate metadata to ensure readability.

Data open for direct external use are data linked to specific publication outputs in the relevant repositories. These include information on partial research results or the tools and instruments needed to reuse or validate research data. Open access to the data will be provided in these repositories, subject to compliance with the terms of the latest available version of the CC BY public licence or its equivalent.

Data to be shared within consortium projects shall be available to consortium members for use by project partners. The quality of the research data will be assured by the researcher's codes of ethics applicable in the consortium institutions.

3. Other research outputs

As already described in the introduction, the technological and instrumental focus of the institute is directed towards outputs that have the character of design solutions in the form of engineering drawings, electrical schematics, software, etc. In most projects, these outputs, including realized apparatuses, instrumentation units, or subsystems, are the primary type of output, while data are secondary.

The management of tangible and intangible results of a design nature and the nature of the associated know-how have historically been the subject of due diligence at ISI, with tangible outputs being operated, maintained and repaired in the laboratories until they reach the end of their useful life.

Their sharing with external parties is and can be governed by a number of rules, determined by the projects in which they were created, contractual provisions in relation to providers and partners in the case of collaborative research.

Infrastructures and their sharing in Open Access mode are a special case, which ensures open access to the infrastructures and must also provide users with adequate protection of the data provided by them and of the data produced in the ISI in order to protect the intellectual property of the users. The handling of this data is governed by a specific data transfer agreement.

4. 4. Allocation of resources

The resources allocated to data care at the ISI are mainly directed towards investments in computer network infrastructure and data storage. An adequate annual budget is allocated for regular upgrades of network, computer and data resources. Investments are included in the ISI's instrument investment plan as appropriate.

Dedicated human resources for data management: data management and care at project level is the responsibility of project leaders. Data interoperability at scientific group level is entrusted to

group leaders and at department level to heads of departments. The care of the computer network, the Institute's data repositories, data security and data backup is the responsibility of the Information Technology Department. The role of the Data Steward, whose tasks are defined in the chapter 'Data access', is assigned to the Institute's library staff.

5. Data security

The security of data stored on internal data repositories is ensured by the IT Department of the ISI. This includes security against hostile attacks, against technology failure and against other risks of data loss.

Storage of data on external data repositories and repositories is ensured by selecting repositories and repositories that are trusted and properly certified.

6. Ethics

The data produced at the ISI are in most cases data on inanimate nature and are not subject to any ethical constraints.

Research data related to medical research (medical instrumentation research at ISI) are subject to the regime regulated by Regulation (EU) 2016/679 of the European Parliament and of the Council, Act No. 110/2019 Coll. on the processing of personal data, as amended, and Act No. 101/2000 Coll. on the protection of personal data, as amended. Therefore, this data cannot be published and shared. However, it may be re-provided to researchers and other bodies within the EU who have been involved in its collection or processing in the past. In accordance with the legislation in force, the data may also be made available to legal or natural persons carrying out, for example, auditing activities.

Ethical restrictions also apply to data related to animal model research conducted at the ISI. Individual experiments on animal models are subject to the approval process of the Ethics Committee for Work with Laboratory Animals. The handling of data resulting from these experiments is specified in the respective experimental plans.