

Data management plan – wymagany jest we wnioskach o finasowanie badań np. NCN, ESRF itp.

Poniższe przykłady są pomyślane jako inspiracja, w żaden sposób nie mogą być uznane za kompletne, mogą zawierać błędy itp. oparte są na formularzu NCN z połowy roku 2022.

Wymagania DMP zawarte są w opisie Narodowego Centrum Nauki:

https://ncn.gov.pl/sites/default/files/pliki/regulaminy/wytyczne_zarzadzanie_danymi_ang.pdf

Data description and collection or re-use of existing data

1. How will new data be collected or produced and/or how will existing data be re-used?

Data will be collected during the measurements with Spectrometer XXX in ASCII, yyy format on the computer dedicated to this instrument...

Data will be collected on a dedicated PC, Apple etc. with LabView, Matlab... data acquisition system, DAQ hardware with USB, PCI, PCI Express, PXI, PXI-Express devices... connected to own build measurement device...

The outsourced measurement data will be transferred to our system in encrypted files and data format will be homogenized with our standards...

Data will be collected within the joint experiment on the large scale facility, where data acquisition procedures are defined (www...) and we will strictly follow these procedures...

Data will be obtained from numerical calculations with software like Quantum Espresso, LAMMPS etc. The standard output files for a given software will be generated and collected... post processing will be performed with XXX software and the set of simplified data will be collected as well...

The existing crystallographic structural databases like Cambridge Crystallographic Data Centre (CCDC) will be used...

Existing literature, open data resources will be used during the project...

2. What data (for example the types, formats, and volumes) will be collected or produced?

The measurements will produce data in the ASCII, xml, proprietary... data format. Data will be stored in the individual directories/folders for each measurement/sample... data compression



will be done with xx. version yy software available for MS Windows, Linux, IOS, android... operating systems...

The standard for data transfer between our Laboratories will be Matlab proprietary format... (version x.x or earlier),

Each dataset will include date, time of the measurement, the sample description, thermodynamic conditions...

Graphical data will be stored in TIFF, jpg, png... format. Audio files in flac format...

the numerical data from the measurement consists of frequency, phase and intensity for each temperature... with an average size of 10kB in 120 files per measurement...

The calculations output data consists of ASCII and binary datasets that amount between 1 kB to 10 GB for an individual calculations. Details of the file formats are related to particular software used... reproducible binary not be preserved due to large storage space requirements...

The laboratory logbook will be stored in .doc format... (version Microsoft Office Word 2010), latex format .tex (with XXX package extensions)...

Data analysis will be done with OriginLab software (version ...) and proprietary data format will be stores together with numerical data in format .xlc (MS Excel 2007), csv...

Documentation and data quality

1. What metadata and documentation (for example methodology or data collection and way of organising data) will accompany data?

The data will be stored in a structured folders with appropriate laboratory logbooks for each folder...

The folder structure /Sample_descriptor/Date/time/... will be used as a standard...

The file naming convention will be: sample_measurement_time_set_...xxx

The xx version of .md file format will be used as a standard in the project. Format conversion, format updates will be provided in xxx file together with conversion software for Debian Linux operating system (version x.x.x).

Structural crystallographic data will be stored in Crystallographic Information File standard (cif) including metadata.

The data intended for a long term preservation will be stored at IFJ-PAN Repository, Zenodo Repository, Nomad Repository, Uniprot Repository... it will be reduced to a minimal set enabling reproduction of the results...

Note – data not intended for the long term preservation will not be shared, unless reasonable request is received.

Istnieją standardy metadanych, których można użyć do katalogowania danych, patrz poniżej.

2. What data quality control measures will be used?

The measurements will be done on XXX spectrometer calibrated with YYY standard, recalibration will be performed each month, day, year...

The sample quality will be assessed prior to any measurement according to the procedure...

The XXX software will be used to assess data quality, reduce noise etc...

The long distance connections will be recalibrated on daily basis in order to avoid artifacts...

The data will be produced by a high quality numerical software and the industry standards...

Data analysis done by the participants of the project, it will reveal any data irregularities...

Storage and backup during the research process

1. How will data and metadata be stored and backed up during the research process?

The data obtained during the project will be stored on the local PC computer, on the iphone, on local disc servers...

A daily data backup on the redundant storage will be performed at a cloud infrastructure...

Data from the analysis software or a portable storage devices will be weekly/monthly archived on the local served backup system.

Backup will not/be encrypted... the password will be known to...

In the event of an incident data restoration will be performed with XXX software ...

2. How will data security and protection of sensitive data be taken care of during the research?

Sensitive data will not be processed during the project.

Legal requirements, codes of conduct

1. If personal data are processed, how will compliance with legislation on personal data and on data security be ensured?

The personal data will not be collected during the project.

Any formal data related to employment will be handled according to IFJ-PAN rules and GDPR. (dpo@ifj.edu.pl)

2. How will other legal issues, such as intellectual property rights and ownership, be managed? What legislation is applicable?

The participants of the project will follow the intellectual property rights management according to the "Zmiana Regulaminu zarządzania prawami autorskimi i prawami pokrewnymi oraz



prawami własności przemysłowej oraz zasad komercjalizacji wyników badań naukowych i prac rozwojowych w IFJ PAN wprowadzonego Zarządzeniem nr 2/2015 z 9 marca 2015 r (Nr29/2017)”.

The published data shall be available according to Creative Commons license with the consent of all authors.

Data sharing and long-term preservation

1. How and when will data be shared ? Are there possible restrictions to data sharing or embargo reasons?

The research data elaborated in the project will be shared monthly, quarterly, after the end of the project...

The data will be shared no earlier than the final publication of the results receives DOI number...

The research data elaborated in the project will not be shared because they are secret...

Each publication with DOI identifier shall contain information relevant for data reproduction.

Data elaborated within the project which directly contribute the research publications will be considered as a long-term value and it will be preserved...

If appropriate a Publisher's embargo policy will be applied...

2. How will data for preservation be selected, and where will data be preserved long-term (for example a data repository or archive)?

The long term data storage will be related to reduced data sets. It will contain data which contribute to the research publications published during and shortly after the project accomplishment.

The long term data preservation will be applied to all/selected raw, processed, ... data produced within the project...

Data repository with DOI or ulr access will be used for the long term (10 years) data preservation, for example IFJ-PAN repository (<https://rifj.ifj.edu.pl/regulamin.html>), RepOD repository (<https://repod.icm.edu.pl/>), NOMAD repository, Zenodo repository (<https://zenodo.org/>), The Universal Protein Resource (<https://www.uniprot.org/>) ...

3. What methods or software tools will be needed to access and use the data?

ASCII, csv, xml, ... files will be stored they can be accessed with any text or xml editor...

Textual files saved with .doc format will require MS Office (version 2007 or later), LibreOffice (version xx or later), PDF files will be stored in PDF/A-3 format based on PDF 1.7 (ISO 32000-1:2008) ...

Pictures will be stored in TIFF format (TIFF 6.0 Baseline TIFF standard)...



If data is stored on NOMAD a proprietary data manipulation and AI systems are provided by NOMAD Laboratory, NOMAD API etc..

4. How will the application of a unique and persistent identifier (such as a Digital Object Identifier (DOI)) to each data set be ensured?

Data stored at recognized repositories like IFJ-PAN repository, RepOD, NOMAD, Zenodo etc... will possess unique DOI number or ulr identifier.

Data management responsibilities and resources

1. Who (for example role, position, and institution) will be responsible for data mangement (i.e the data steward)?

The data steward will be employed as a part time participant of the project. She/he will be responsible for the research data management and long term data acquisition, (estimated costs 75 PLN).

The PI will be responsible for data management...

For data stored on data sharing platforms the regulations of these platforms will hold (<https://rifj.ifj.edu.pl/regulamin.html>, <https://nomad-lab.eu/>, <https://zenodo.org/>).

2. What resources (for example financial and time) will be dedicated to data management and ensuring the data will be FAIR (Findable, Accessible, Interoperable, Re-usable)?

At present IFJ-PAN repositories are free of charge.

The local storage space will be purchased from the project resources (estimated cost 1000000PLN).

We will follow and ensure FAIR (Findable, Accessible, Interoperable, Re-usable) data standards.

Data stored on NOMAD are under the CC BY 4.0 license, enables FAIR sharing and use of materials science data.

Przydatne linki:

Przykłady, opisy DMP:

<https://www.ifj.edu.pl/library/open-access/plan-zarzadzania.php>

https://pg.edu.pl/files/2021-08/DMP_engineering.pdf

<https://www.biblos.pk.edu.pl/en/science/open-science/data-management-plan-dmp>

https://ruj.uj.edu.pl/xmlui/bitstream/handle/item/155290/Szafranski_Leszek_DMP_examples_2020.pdf



<https://www.ucl.ac.uk/library/open-science-research-support/research-data-management/policies/writing-data-management-plan>

<https://datamanagement.hms.harvard.edu/plan/data-management-plans>

<https://libraries.mit.edu/data-management/plan/write/>

opisy Komisji Europejskiej:

https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm#A1-template

Repositories

Regulamin Repozytorium IFJ PAN: <https://rifj.ifj.edu.pl/regulamin.html>

<https://www.ifj.edu.pl/library/open-access/info-repozytoria.php>

Zenodo: <https://zenodo.org/>

NOMAD Laboratory: <https://nomad-lab.eu/>

RepOD: (<https://repod.icm.edu.pl/>)

Repozytoria ICM: <https://drodb.icm.edu.pl/>

OpenAIRE: <https://www.openaire.eu/>

EIFL: <https://www.eifl.net/>

The Universal Protein Resource: <https://www.uniprot.org/>

FAIR: doi: 10.1038/sdata.2016.18 (2016)

MetaData standards: <http://rd-alliance.github.io/metadata-directory/>

<https://www.loc.gov/marc/marcdocz.html>

<https://ddialliance.org/Specification/>

<https://tei-c.org/release/doc/tei-p5-doc/en/html/index.html>

DMP online: <https://dmponline.dcc.ac.uk/>