

Data management plan in National Science Centre project

1. Data description and collection or re-use of existing data

1.1 How will new data be collected or produced and/or how will existing data be reused?

- This section should include a description of the context of the research.
- How will new data be collected? Standards, methods, software.
- What kinds of already existing data you will use? Own, held by a third party?
- What quality assurance processes will you use? Calibration processes, repeated samples or measurements, data recording and capture standards, usage of controlled vocabularies, data entry validation, data peer review, etc.
- How will you organize your files and handle versioning?

Notice: The description should include clear details of where the secondary data came from and how the primary data was collected or created.

The project may not generate new data → in which case an appropriate justification should be provided in the PZD

1.2 What data (for example the kinds, formats, and volumes) will be collected or produced?

- What type, format and volume of data will you collect, generate or reuse?
- What volume of data will you generate?
- Give details on the data format (preference to open and standard formats).

Notice: The section should clearly describe or list the types of data that were generated and the associated data formats, including data conversion strategies where necessary.

Explain why specific formats were chosen and indicate whether they are in an open and standard format. If a proprietary format was used, explain why

2. Documentation and data quality

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

- What metadata will the data be described with?
- What international standards or schemes (Dublin Core, Data Documentation Initiative DDI) will be used to structure metadata?
- What information is required for users (computer or human) to read and interpret the data in the future?

Notice: At this section, you need to clearly define the documentation needed to enable reuse of the data - indicate where the information will be saved (e.g. database with links to each item, README text file, file headers, code books or lab notebooks)

2.2 What data quality control measures will be used?

- How you will be control quality of data?
- How the data collection, analysis and processing methods used may affect the quality of data?
- Indicate the existing mechanisms to prevent unauthorized changes in the institution?
- How measurement error and bias will be eliminated?
- How you will minimise the risks related to data accuracy?
- How can data collection, analysis and processing methods affect data quality?

3. Storage and backup during the research proces

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

- What will be the method and procedures for making backups, and who will be responsible for them?
- How data will be recovered in the case of loss/damage?
- How data will be protected?
- Are special measures needed to transfer data from mobile devices, from fieldwork sites or from home equipment to a central work server?

Notice: This section needs to clearly describe the location where data and backups will be stored during the project, how often they will be performed, the data storage infrastructure provided by the home institution or alternative

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

If the project does not concern sensitive data (no research involving human) the following notation should be included:

Example answer: „The project does not provide the creation, storage and processing of sensitive data”.

- Where sensitive data will be stored?
- How the data will be recovered in the event of an incident?
- Who will have access to the data during the research and how access to data will be controlled, especially in collaborative partnerships?
- Which institutional data protection policies has been implemented?

Example answer: AGH University of Krakow authorities have implemented the Information Security Management System (ISMS), which goal is to protect data and information. A structure of information

security management includes, among others: DPO (Data Protection Officer) and IT System Administrator (Central and Local Officers). The Data Protection Officer has been appointed in 2018. ISMS is based on the following documentations: Information Security Policy at AGH University of Krakow (Regulation No. 9/2017), declaration of application, book of guidelines and security procedures, implementation and functioning of management control at AGH University of Krakow (Regulation No.40/2012) and Personal Data Protection Policy (Regulation No. 26/2019).

4. Legal requirements, codes of conduct

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

- If the project does not concern personal data (no research involving human) the following notation should be included::

Example answer: „The project does not provide for the processing of personal data”.

- Will personal data be collected and processed in the project? If so, how will the protection of this data be ensured?
- Do you need to use anonymisation throughout a data collection? Data encryption?
- Has an access procedure been defined for authorised users of personal data?

Example answer: The documentation of the personal data processing consist of:

- a) Register of processing activities
- b) Register of processing activities category
- c) Register of personal data breach
- d) Report of personal data breach
- e) Entrustment agreement for the processing of personal data

Management procedure of the personal data access is specified in the Personal Data Protection Policy of AGH University of Krakow (Regulation No. 26/2019)

Notice: Clearly indicate whether personal data will be collected/used as part of the project and, if applicable, how compliance with applicable legislation will be ensured (e.g. by obtaining informed consent, considering encryption, anonymisation or pseudonymization)

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

- Who will be the owner of the data (indicate all owners)?
- Which licences will be applied to the data?
- What restrictions apply to the reuse of third-party data?
- Do you need to seek copyright clearance before sharing data?

Example answer: In the field of intellectual property rights will be applied: the Act on Copyright and Related Rights and the Act on Industrial Property Law, as well the Regulation for the management

of copyright and related rights, industrial property rights and the principles of commercialization in AGH University of Krakow introduced by Regulation No. 18/2015, and the Regulation of RODBUK AGH Cracow Open Research Data Repository. Data will be shared under Creative Commons license - *please indicate on which one precisely. NCN allows for research data CC 0 (transferring the dataset to the public domain, allows unlimited and unbounded use of the work) or CC BY (allows users to copy, modify, distribute and create new works or collections based on the licensed work, provided the authorship of the work is indicated, allows commercial use).*

5. Data sharing and long-term preservation

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

- When data will be accessible: during or after the research? How will potential users find out about them?
- For how long will the data be stored in the repository?
- Are there any barriers and constraints to making the research data fully or partially accessible?
- Will journal publishers require deposit of data supporting article findings?
- Does the sharing of data require the consent of the participant?

Example answer: Data will be discoverable and shared in RODBUK AGH Cracow Open Research Data Repository. Principle Investigator will decide about sharing data between co-investigators and outside the research group. Moreover, data sharing could be postponed because of protect intellectual property, patent procedure or before publishing – all restriction will be made on the Principle Investigator demand and data sharing will be limited in time.

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

- What data must be retained or destroyed for contractual, legal, or regulatory purposes?
- How it will be decided what data to keep?
- What procedure would be used to select data to be preserved?
- Does the repository in which the data will be long-term stored respect the FAIR principles?

Notice: This point should detail what data collected or created as part of the project has been preserved in the long term and clearly indicate how long and how the selection of data for preservation was made

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

- Do data need to be converted to a standard or open format with long-term validity for long-term preservation?
- What mechanism will be used for data sharing (request handled directly, repository)?

Notice: This point must clearly indicate which tools or software (e.g. specific scripts, codes or algorithms developed by the project, software version(s)) may be needed by potential users to access, interpret and (re)use the data

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

Example answer: Data will have a digital object identifier (DOI)

6. Data management responsibilities and resources

6.1 Who will be responsible for data management (i.e. data steward)?

- Who will be responsible for data management during the research?
- Who will be responsible for data management after the research is finished? Will it be someone from the project or the repository where the data will be deposited
- Division of roles regarding data management: who produces the data, who is responsible for its quality, who for archiving and long-term management

Example answer: The person responsible for data management during the project will be the person realizing the scientific activity (name, Institution, Faculty). Once the data has been made available in the repository, RODBUK administrators will be responsible for managing it.

Notice: Clearly define data management roles and responsibilities (e.g. data acquisition, metadata creation, data quality, data storage and backup, data archiving and data sharing), identifying responsible individuals where possible

6.2 What resources will be dedicated to data management and ensuring that data will be FAIR? (What costs will be associated with ensuring FAIR standards in the project? How will they be paid for?)

- What are the costs for making data FAIR in your project?
- If the additional resources will be needed to prepare data for deposit?

Example answer: The project does not provide separate funding for preparing data for release and storage. The cost of research data management after the project is completed is carried by AGH University of Krakow

Consultation of the DMP Marta Urbaniec

murbaniec@agh.edu.pl

12 617 32 21