An example of a Data Management Plan for a project aimed at transliterating, translating and additional descriptions of the first volume of 18th-century manuscripts describing the several-year journey of Nathaneal Jacob Gerlach from Gdańsk around Europe.

The Project assumed that the work would be done with the use of a manuscript already digitized and made available in an electronic version. The project will result in ready-made transliterations, translations and other manuscript-related studies.

DATA MANAGEMENT PLAN – FORM FOR GDAŃSK UNIVERSITY OF TECHNOLOGY EMPLOYEES

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 re-used?

Source (raw) data: Scans in jpeg format from the Pomeranian Digital Library available under the public domain (CC0 license). Unique URI for the source data:
 https://pbc.gda.pl/dlibra/publication/10113/edition/6021
New data generated during the project: electronic documents with transliterations, translations, and descriptions of additional page elements such as seals, notes, etc. Each page of transliterations, translations will be corresponding to each scanned image in source data. The descriptions will include references - links to specific pages in the source data.

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

Source (raw) data: jpeg images collected in one folder. Each image has a unique name consisting of a prefix and a sequence number. Number of files: 363. Total size of files: 300MB

New data generated during the project: documents in .docx and .txt and .pdf formats. Each *.docx document will be converted to open formats (.rtf, .odt) . Pdf files will contain images join with translations and transliterations. Estimated number of files: 726 txt files, 3 .docx files, 3 pdf files. Estimated size of files: 900MB

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?

The data will be stored in an ordered structure of folders: separated folders for source data, translations, transliterations, and descriptions. Names of txt files will be corresponding to source files names. Each final document docx, and pdf will be described with metadata in the Dublin Core format. and will be share in Pomeranian Digital Library with a unique URI specified for the digital library system (dLibra).

Based on all the files collected during the project, will be generated the dataset in zip format. The dataset will be made available through the open research data repository - MOST Wiedzy Open Research Data Catalogue witch support different metadata standards and schemas (eg. DataCite, DublinCore). The authors will be identified and authorized by ORCID number.

2.2. What data quality control measures will be used?

 The specificity and scope of the project does not require the development and application of defined quality control procedures.
In the individual stages of the project will participate People with knowledge and experience in a given field (eg in translation and transliteration of texts). Finally, the consistency of the data compiled with the source data will be verified

3. Storage and backup during the research process

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

 Data generated during the project will be stored on the computers of the team members and each day the data will be backed up in a dedicated external disk space based on disk arrays with the automatic backup functionality.

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

During the project, We will not collect process any personal and sensitive data.

The access to the computers of the team members and to the external disk service will be restricted only to authorized users and will be protected by a strong password.

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?

During the project, We will not collect process any personal and sensitive data.

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

In the case of obtaining data or results that may constitute the matter of intellectual property, the rights to them are acquired by the person who obtained them.

The data generated during the project will be made available under a CC-BY license.

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?

All data obtained as a result of the project will be made available immediately after the end of the project.

The dataset with all orderly data will be made available through the open research data repository - MOST Wiedzy Open Research Data Catalogue that fulfills FAIR requirements.. Additionally, single studies and documents will be share in Pomeranian Digital Library.

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

All data obtained as a result of the project will be qualified for long-term preservation.

Both The MOST Wiedzy Open Research Data Catalogue, and The Pomeranian Digital Library guarantee that the data will be available for a minimum of 10 years.

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

The open file format allows using of commercial software as well the free, standard, and open software (e.g. txt reader, PDF viewer, open office). There are no requirements for the operating system.

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

The datasets made available through the MOST Wiedzy Open Research Data Catalogue will have the DOI assigned. The MOST Wiedzy Open Research Data Catalogue also supporting assigning DOI for each version of the existing dataset.
The Pomeranian Digital Library guarantee a unique URL for each resource.

6. Data management responsibilities and resources

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

The Project PI (Imię Nazwisko) will be responsible for the overall data quality.

The Open Science Competence Center established by PG (pg.edu.pl/openscience) will be responsible for support in the preparation and maintenance of the Data Management Plan and for support in the process of preparing and sharing datasets in the MOST Wiedzy Open Research Data Catalog.

6.2. 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)?

There will be no additional financial costs of data management involved in the project.