Data Management Plan Checklist

After answering each of the questions below, you should be able to easily create a data management plan.

1 What type of data will be produced?

- ☐ How will data be collected?
- ☐ What would happen if the data were lost or became unusable later?
- ☐ How much data will your project produce, and at what growth rate? How often will it change?
- ☐ What is your data storage and backup strategy?

2 What standards will be used for documentation & metadata?

- ☐ What directory and file naming conventions will be used?
- ☐ What project and data identifiers will be assigned?
- ☐ Is there a community standard for data sharing or integration?

3 What steps will be taken to protect privacy, security, confidentiality, intellectual property, or other rights?

- ☐ Who controls the data (e.g. PI, student, lab, University, funder)?
- ☐ Are there any special privacy or security requirements to uphold (e.g. personal or high-security data)?
- ☐ Are there any embargo periods to follow?

4 If you allow others to reuse your data, how will the data be accessed and shared? How will it be archived for preservation and long-term access?

- ☐ Are there any data sharing requirements?
- ☐ How long should the data be retained?
- ☐ What file formats are you using? Are they proprietary, like .xls or .docx, or are they open, like .csv or .rtf? Are there any special tools and/or software needed to work with or view the data?
- ☐ Are there data repositories that are appropriate for your data?
- ☐ Who will maintain your data for the long term?

guides.lib.olemiss.edu/datamanagement

Table 5.1. Malaysia Open Science Platform FAIR Data Management Plan (DMP) Template

SECTIONS	KEY ITEMS TO BE INCLUDED IN A DMP
1. Data Summary	• Specify the objective of the data collection and its relation to the overall aim of the project.
	• Specify the types and formats of datasets that will be collected or generated.
	• Specify if existing data is being re-used (if any). If yes, has consent or licensed use been obtained from the data owner, or is the data open access or openly licensed?
	• Specify the origin (source) of the data.
	• Specify the expected size of the data (if known).
	• Specify to whom will the data be useful.
2. FAIR Data	
2.1 Making data findable	**Explain how research data can be made findable.**
	• Specify the standard identification mechanism of data. For example, the use of persistent and unique identifiers such as Digital Object Identifiers.
	• Specify naming conventions and versioning that will be used.
	• Specify search keywords that will be used.
	• Specify standards for metadata creation. In case metadata standards do not exist in your discipline, outline what type of metadata will be created and how.
2.2 Making data openly accessible	**Explain how research data can be made openly accessible.**
	• Specify which research data can be made openly available. If certain research data cannot be shared or can be shared but with certain restrictions applied, describe reasons for the restriction by clearly citing whether it is legal or contractual reasons.
	• Specify methods or software tools needed to access the data.
	• Specify if the documentation about the software is needed to access the data. If it is a yes, is it included?
	• Specify if it is possible to include the relevant software, such as via open source code.
	• Specify the location where the data and associated metadata will be stored and be made accessible. Preference should be given to certified repositories which support open access where possible.
	• Specify how access will be granted for research data that are restricted.
	• Specify how the identity of the person accessing the data will be ascertained.
2.3Making data interoperable	**Explain how research data can be made interoperable.**
	• Specify data, metadata vocabularies, standards or methodologies to facilitate interoperability.
	• Specify whether standard vocabulary will be used for all data types to allow inter-disciplinary interoperability.

2.4 Making data reusable	**Explain how and when the research data will be made available for further re-use:**
	• Specify plans for licensing arrangements for the use of shared research data to allow the widest reuse possible.
	• Specify when the research data will be made available for reuse. If applicable, specify plans for embargoing (i.e. a time delay that is applied to research data, codes and other materials before they can be made available, accessible and usable by others) as needed. Be sure to explain the reasons why the embargo period is necessary, whether it is due to research funding policy, patent and clinical trial reasons and for what period an embargo is needed.
	• Specify whether the shared research data is usable by third parties, especially after the end of the project. If legal and ethical restrictions on access and use of sensitive data is applicable, explain the reasons (Is it due to the use of human subjects, research data containing information with national security risks etc.)
	• Specify data quality assurance process.
	• Specify the length of time for which the data will remain reusable.
3. Allocation of resources	• Estimate the costs for making research data to be FAIR. Specify plans to cover the costs.
	• Specify roles and responsibilities for data management in the project.
	• Specify costs and potential value of long-term preservation.
4. Data security	• Specify the level of openness of data, whether it is confidential, restricted or public, and describe how to secure and dispose of the created data. • Address data recovery as well as secure storage and transfer of sensitive data.
5. Data storage and backup	• Specify data storage mechanism (i.e. where do you store the data during and after the research is conducted, and how long will the data be stored?)
	• Specify data backup mechanism (i.e. how many times will a backup on the data be done, will backup be performed on all data or just some of it?
6. Ownership	• Specify who will be responsible to collect the data.
	• Identify if there is a joint ownership between an organisation with an external body or organisation.
	• Identify if there are any contractual agreements that may affect copyright ownership.
	• Identify if the data is collected by an employee of the University in the course of their employment.
	• Identify if the data collected or compiled is in Malaysia or elsewhere.
	• List those involved in the creation of data.
7. Ethics	• Specify if ethics review and clearance is needed. If yes, describe action plans for ethics application.
	• Specify if informed consent for data sharing and re-use and long-term preservation will be sought when dealing with personal data.
8. Other	• Specify if there are any other procedures for data management that will be used (if any).

5.4.2. Metadata Management

a) Good metadata is key for research data access and re-use.

b) Metadata fields selected for the digital repository must match metadata standards including different naming schemes for domain-specific repositories.

c) All metadata for raw research data will be stored at the Malaysia Open Science Platform. Anyone may access the metadata free of charge.

d) The metadata may be re-used in any medium without prior permission from the data originator for not-for-profit purposes provided the persistent identifier or a link to the original metadata record is given.

e) Metadata fields selected for the digital repository must match unqualified Dublin Core metadata fields, as well as including publication and refereed status.

f) The list of metadata fields following Dublin Core are:

g) Title, creator, subject, description, publisher, contributor, date, type, format, identifier, source, language, relation, coverage, rights, audience, provenance, rights Holder, instructional method, accrual method, accrual periodicity and accrual policy.

5.4.3. Research Data Management (RDM) in Public Funded Research

a. Ownership

(a) Malaysian Universities, Research Institutes, and respective Government Entities own all research data produced by research projects and activities conducted at or under the auspices of the University, Research Institute and Government Entities.

(b) All new contractual agreements for joint projects, studentship agreements and any other type of collaborative agreements with external bodies must comply with the Guidelines.

(c) Exceptions:

i. Where the research funder retains ownership of the research data.

ii. In joint projects, the research data management plans (DMP) shall address the creation, management, confidentiality, retention, and publication of data both digital and non-digital.

(d) The PI and his/her designated researchers have their rights to use and publish research data arising from their project, unless specific terms of sponsorship, other agreements, institutional policies or other relevant national laws and policies supersede these rights.

b. Data Management Plan (DMP) in Public Funded Research

(a) All funded research must include a DMP that records how the research data arising from the research project will be handled during and after the project is completed, describing what data will be shared and/or made open, and how it will be curated and preserved.

(b) DMP must comply with relevant laws which regulate access to and use of data.

c. Deposition

(a) Research data is deposited in the University, Research Institute or Government Entity owned data repository which complies with the FAIR principle (DOI: 10.1038/sdata.2016.18) and will be linked to the MOSP.

(b) The MOSP Platform accepts research data used in establishing and validating research findings, pre-print and post-print materials, publications and reports.

[14] https://www.dcc.ac.uk/guidance/standards/metadata/list

d. Data sharing

(a) The research data shall be made available for sharing via the MOSP unless there are prior formal agreements with external collaborators, funding bodies and parties on nondisclosure or proprietary use of the data.

(b) In the following circumstances, several additional criteria to (a) must be applied:

i. Consent must be obtained from all data subjects for all human data collected and must be anonymised before being deposited and published. The consent form must indicate the use of the data, if it is to be published and reused, and the type of third-class party who may have access to the data.

ii. For human data collected from data subjects under the age of eighteen years, consent must be obtained from the parent, guardian or person who has parental responsibility for the data subject concerned.

iii. For data containing information intended for commercialisation, it must not be deposited until the patent has been filed.

iv. For data that concerns national security matters, it must receive clearance from an authorised body prior to deposition. Access to the data will be completely restricted.

e. Storage and Retention

(a) All research data shall be stored in locations or devices on the Institutional Repository.

(b) Reasonable steps shall be taken to ensure the security and integrity of all research data under retention.

(c) All research data related to a research project or an activity shall be retained not more than ten (10) years after publication or after the completion of the project or last access to the dataset, whichever is later. A longer period of retention may be specified by external research funders. Under both circumstances, the period of retention is subject to legal and regulatory requirements.

(d) Material can be withdrawn from the MOSP, if it is proven copyright violation or plagiarism or falsified research. Withdrawn material is not deleted per se but is removed from public view. Withdrawn items' identifiers are retained indefinitely.

e) If there are major changes to work in the MOSP, an updated version may be deposited as a separate item and can be linked to the first deposited material.

f) Errata and corrigenda lists may be included with the original record if required. If necessary, an updated version may be deposited.

g) In the event of the MOSP being closed, the information stored will be transferred to respective institutions.

f. Data reuse

(a) If the data that will be reused is licensed, the conditions of that licence regarding data sharing i.e. redistribution, must be followed.

(b) If redistribution is permitted, the data may be shared and must be attributed to the originator of the research data.

(c) Data stewards will be able to guide researchers on the information on licensing terms.

(d) If data is re-used, the original author is not implicated with the consequences from the activity if the original author is not involved.

(e) Metadata may be re-used in any medium without prior permission from the data originator for not-for-profit purposes provided the persistent identifier or a link to the original metadata record is given. Anyone may access the metadata free of charge.

g. Disposal

Beyond the period of retention specified here, all research data must be disposed of. Any destruction of the research record, either whole or part shall follow the Guidelines.

CHAPTER 6

IMPLEMENTATION OPEN SCIENCE IN PUBLIC FUNDED RESEARCH

6.1. IMPLEMENTATION AUTHORITY

6.1.1. National Level

The Ministry of Science, Technology and Innovation (MOSTI) will be the implementation authority at the national level.

6.1.2. Institutional Level

At the institutional level, the Public Research Institutes (PRI) and Institutions of Higher Learning (IHL) will respectively be implementing Open Science Guidelines in their entities. For PRI, the Deputy Director General will be responsible for overseeing and managing the operation of the Guidelines while in IHL, this responsibility is undertaken by the respective Deputy Vice Chancellor (Research).

6.2. IMPLEMENTATION REQUIREMENTS

The fundamental principle underlying the Guidelines as mandated is that the results of research carried out with public funds shall be openly available for both other researchers and the interested public and companies. This is based on the accepted arguments that open access to scientific information is good to create better conditions for scientific research and for society at large. In addition to its research value, there is also a clear emphasis on the social benefits of open science.

In implementing the Guidelines on Open Science in Funded Public Research the followings are required:

a) Malaysia Open Science Alliance is responsible to ensure that the guiding principles in the Guidelines on Open Science in Public Funded Research is observed and implemented holistically across all levels of relevant stakeholders. Malaysia Open Science Alliance, as well as Institute of Higher Learnings, Government Research Institutes and Non-government Research Institutes may use assessment indicators for implementation of Malaysia Open Science in the form of "Checklist for Implementation of Raw Research Data Repositories at Institute of Higher Learnings, Government Research Institutes and Non-government Research Institutes" and "Checklist for Researchers Readiness to Share Raw Research Data" as in **Appendix 6.1 and Appendix 6.2**, respectively.

b) This Guidelines on Open Science in Public Funded Research is subjected to reviews and amendments at any time, as it deems necessary in the interest of technological changes, applications, procedures, legislations and societal benefits.

c) All beneficiaries of publicly funded research activities must follow this Guidelines on Open Science in Public Funded Research to achieve an effective and efficient implementation of Malaysia Open Science Platform that will empower planning and management of research, development, commercialisation and innovation.

CONTACT INFORMATION

For queries relating to this document please contact:

Guidelines Owner:	Academy of Sciences Malaysia (ASM)
All Enquiries Contact:	Academy of Sciences Malaysia (ASM) Level 20, West Wing, MATRADE Tower, Jln Sultan Hj Ahmad Shah, Off Jln Tuanku Abd Halim, 50480 Kuala Lumpur.
Telephone:	03-6203 0633
Email address:	admin@akademisains.gov.my