Open SCIENCE
Malaysia Open Science Platform (MOSP)

KEMENTERIAN SAINS, TEKNOLOGI DAN INOVASI
MINISTRY OF SCIENCE, TECHNOLOGY AND INNOVATION

THE LANDSCAPE OF *OPEN SCIENCE* IN MALAYSIA

REPORT

THE LANDSCAPE OF OPEN SCIENCE IN MALAYSIA

ACADEMY OF SCIENCES MALAYSIA,
MINISTRY OF SCIENCE, TECHNOLOGY AND INNOVATION

THE LANDSCAPE OF OPEN SCIENCE IN MALAYSIA
This landscape study is conducted by the Malaysia Open Science Alliance Working Group on Guidelines to **assess the landscape of Open Science in Malaysia.**

Academy of Sciences Malaysia
Level 20, West Wing, MATRADE Tower
Jalan Sultan Haji Ahmad Shah off Jalan Tuanku Abdul Halim
50480 Kuala Lumpur, Malaysia

Perpustakaan Negara Malaysia Cataloguing-In-Publication Data
The Landscape of Open Science in Malaysia

ISBN 978-983-2915-62-1
Keywords related to publication
provided by National Library during
ISBN application process
1.Information storage and retrieval system—Science.
2.Information storage and retrieval system—Science and state.
3.Science and state— Malaysia.
4.Research—Malaysia.
5.Government publications—Malaysia.
I. Akademi Sains Malaysia.
025.065

TABLE OF CONTENT

FOREWORD

Malaysia Open Science Platform (MOSP), an initiative from the Academy of Science Malaysia (ASM) supported by the Ministry of Science Technology and Innovation (MOSTI), has been tasked to conduct a landscape study in order to analyse and gauge the level of readiness of local institutions on open science. The top management, researchers and librarians from five local research universities were selected as the subject in this preliminary research. Surveys and interviews from targeted respondents were carried out to harvest their view and perspective on open science. Though the country was hit by a pandemic, the landscape study was successfully done, thanks to the advancement of technology that we have nowadays.

The findings from the researchers' surveys and interviews indicated that the majority of the respondents were still reluctant to subscribe to open science and that they are more comfortable with conservative approaches in disseminating research data. The main concerns voiced by researchers on open science is regarding data security and safety. However, researchers seem to also be lacking in their awareness on open science and what open science entails. Therefore, it is important for MOSP to provide awareness and develop a proper guideline for the researchers to understand, protect and avoid data abuse when sharing their data.

The study also indicated that the majority of librarians in Malaysian IHLs are still required to perform their traditional role. As data repository and management activities increase, librarians should reskill themselves, perhaps as data stewards, to enable them to better manage knowledge and data in repositories. The lack of awareness seems to be the main factor contributing to the lag of open science in Malaysia. Fortunately, there is strong support for open science by the top management from the five research universities in Malaysia.

Analysis on the level of readiness for open science in Malaysia says a lot about the current situation of research culture in Malaysia. Most researchers in Malaysia are undoubtedly still way behind their counterparts in other developed nations with regards to data sharing. To be on par with most other developed nations, Malaysia should emulate best practices from these nations, particularly on their data sharing and open science practices. The National Policy for Science Technology and Innovation (NPSTI 2021-2030) has included data sharing and open science as an important component in supporting the National STI agenda to create a new research data ecosystem that will benefit the various stakeholders in Malaysia.

Finally, for Open Science and MOSP to become a reality in Malaysia, several issues and concerns need to be addressed. MOSP needs to be holistically implemented nationwide. The open science vision and mission must explicitly be mentioned and highlighted in the national policy. Indicators for open science readiness should be developed. Funding should be made available for its sustainability, capacity building and, last but not least, open science policy should be positioned into the government institutional framework.

Prof. Dr. Noorsaadah binti Abd Rahman,
Chairperson,
Malaysia Open Science Alliance.

ACKNOWLEDGEMENT

We would like to extend our appreciation and thanks to all those who have participated in the study. We acknowledged with thanks the comments and suggestions imparted and guidance extended during the series of engagement with the Malaysia Open Science Alliance and Malaysia Open Science Alliance Working Group on Guidelines, the Academy of Sciences Malaysia, the top management, researchers, librarians and representatives from institutional repositories of the selected Malaysian public and private institutions of higher learning (IHL) and the public and private research institutions, government agencies and ministries, industrial players, legal units and representatives from global Open Science initiatives.

MALAYSIA OPEN SCIENCE ALLIANCE

1. Professor Dr Noorsaadah Abd Rahman FASc (Chairperson, UM)
2. Professor Datuk Dr Rohana Yusof FASc (ASM)
3. Academician Professor Emerita Datuk Dr Mazlan Othman FASc (ISC-ROAP)
4. Professor Dr Shaliza Ibrahim (UM)
5. Professor Dr Ismanizan Ismail (UKM)
6. Professor Dr Zulkifli Idrus (UPM)
7. Professor Dr Rosni Abdullah (USM)
8. Professor Dr Abrizah Abdullah (UM)
9. Puan Roshana Alma Mohd Ali (MOSTI)
10. Puan Noraziah Sharuddin (UTM)
11. Dr Liew Chee Sun (UM)
12. Dr Salina Abdul Aziz (MOH)
13. Encik Azih Yusof (MAMPU)
14. Encik Kamaruhzaman Mat Zin (MASTIC)

MALAYSIA OPEN SCIENCE ALLIANCE WORKING GROUP ON GUIDELINES

1. Professor Datuk Dr Rohana Yusof FASc (Chairperson, ASM)
2. Emeritus Professor Dato' Dr Rahmah Mohamed FASc (ASM)
3. Dr Khasiah Zakaria (UM)
4. Dr Sik Cheng Peng (UM)
5. Dr Nasuha Lee Abdullah (USM)
6. Dr Syed Norris Hikmi Syed Abdullah (UTM)
7. Puan Roshana Alma Ali (MOSTI)
8. Encik Kamaruhzaman Mat Zin (MASTIC)
9. Encik Kamarulzaman Kamaruddin (MOSTI)
10. Professor Dr Ismanizan Ismail (UKM)
11. Professor Dr Samsilah Roslan (UPM)

ACADEMY OF SCIENCES MALAYSIA - MALAYSIA OPEN SCIENCE PLATFORM

1. Puan Hazami Habib (Chief Executive Officer)
2. Dr Nurzatil Sharleeza Mat Jalaluddin (Senior Executive)
3. Puan Nurlizah Abu Bakar (Senior Executive)
4. Shyama Sundari Devi (MOSP Executive)
5. Nur Hanisah Ismail (MOSP Executive)
6. Nur Faizah Azhari (MOSP Executive)
7. Shaanthana Subramaniam (MOSP Executive)
8. Muhamad Fariz Muhamad Suhaimi (MOSP Executive)

LIST OF FIGURES

Chapter 7

Chapter 10

LIST OF TABLES

LIST OF BOXES

ABBREVIATION AND ACRONYMS

AI	Artificial Intelligence
AIM	National Innovation Agency Malaysia
ANDS	Australian National Data Service
AOSP	African Open Science Platform
APCs	Article Processing Charges
API	Application Programming Interface
ARDC	Australian Research Data Commons
CD	Compact Disc
CODATA	Committee on Data for Science and Technology
CSIRO	Commonwealth Scientific and Industrial Research Organisation
DANS	Data Archiving and Networked Services
DBD	Database Driver
DBI	Database Interface
DC	Dublin Core
DEVL	Data Enhanced Virtual Laboratory
DMS	Document Management System
DOAJ	Directory of Open Access Journals
DOAR	Directory of Open Access Repositories
DOI	Digital Object Identifier
DVD	Digital Versatile Disc
e-IRG	e-Infrastructure Reflection Group
EOSC	European Open Science Cloud
ETL	Extract, Transform, Load
ExPaNDS	EOSC Photon and Neutron Data Services
FAIR	Findable, Accessible, Interoperable and Reusable
FOSTER	Facilitate Open Science Training for European Research
FRGS	Fundamental Research Grant Scheme
GDP	Gross Domestic Product
GERD	Gross Domestic Expenditure on R&D
HTML	Hypertext Markup Language
HTTPS	Hypertext Transfer Protocol Secure
ICT	Information and Communications Technology
ID	Identification/Identity
IHLs	Institute of Higher Learnings
IP	Intellectual Property
IR	Institutional Repository
ISC	International Science Council
ISO	International Organization for Standardization
ISP	Internet Service Provider
IT	Information Technology
KNAW	The Royal Netherlands Academy of Arts and Sciences
KPI	Key Performance Indicator
LRGS	Long Term Research Grant Scheme
MARC	Machine-Readable Cataloguing
MASTIC	Malaysia Science and Technology Information Centre
MCO	Movement Control Order
MESTECC	Ministry of Energy, Science, Technology, Environment and Climate Change
MOHE	Ministry of Higher Education
MOSP	Malaysia Open Science Platform
MOSTI	Ministry of Science, Technology, and Innovation
MRDCS	Malaysian Research and Development Classification System
MyLAB	Malaysia Laboratories for Academia-Business Collaboration
MyREN	Malaysian Research and Education Network
NARCIS	National Academic Research and Collaborations Information System
Nectar	National eResearch Collaboration Tools and Resources project
NGO	Non-Government Organisation

NI4OS	National Initiatives for Open Science in Europe
NPSTI	National Policy on Science, Technology and Innovation
NWO	Netherlands Organisation for Scientific Research
OA	Open Access
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
OECD	Organisation for Economic Co-operation and Development
OpenAIRE	European Open Science Infrastructure
ORCID	Open Researcher and Contributor ID
PLOS	Public Library of Science
PRGS	Prototype Research Grant Scheme
PRIMS	Putra Research and Innovation Management System
PRIs	Public Research Institutes
PTM	*Pusat Teknologi Maklumat*
RAM	Random Access Memory
RAND	RAND Corporation
RDA	Resource Description and Access
RDA	Research Data Australia
RDBMS	Relational Database Management System
RDC	Research Data Cloud
RDF	Resource Description Framework
RDM	Research Data Management
RDMP	Research Data Management Plan
RDS	Research Data Services
RIs	Research Institutes
RMC	Research Management Centre
ROARMAP	Registry of Open Access Repository Mandates and Policies
ROI	Return on Investment
RU	Research University
S&T	Science and Technology
STI	Science, Technology and Innovation
SURF	Organisation for Open Science promotion in the Netherlands
ToT	Training of Trainers
TRGS	Transdisciplinary Research Grant Scheme
TRUST	Transparency, Responsibility, User Community, Sustainable, and Technology
TU Delft	Delft University of Technology
UKM	Universiti Kebangsaan Malaysia
UM	Universiti Malaya
UNESCO	United Nations Educational, Scientific and Cultural Organization
UPM	Universiti Putra Malaysia
US	United States
USM	Universiti Sains Malaysia
USMDESC	USM Digital Exhibition and Special Collections
UTM	Universiti Teknologi Malaysia
VSNU	Association of Universities in the Netherlands
XML	eXtensible Markup Language

GLOSSARY

Article Processing Charges
Fees that some scholarly publishers charge authors of academic papers to publish their work in open access.

Citizen Science
Citizen Science is defined as active involvement of the laymen or public in a research project. Through Citizen Science, researchers are encouraged to interact more with the society, thus offering the public with fertile opportunities for scientific learning. An example of a citizen science project is “Citizen Science in the Surveillance and Monitoring of Mosquito-borne diseases” and in Malaysia, is “Peninsular Malaysia Butterfly Count”.

Data curator
Responsible for organising and integrating data collected from various sources. It involves publication, presentation, reuse and preservation of the data.

Data Governance
A cross-functional management programme that treats data as an organisation asset through collection of policies, standards, process, people and technology to achieve a set of goals.

Data management plan
A data management plan 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.

Data steward
Protects the integrity and quality of data, adherence or compliance to standards and protocols, governance and advocacy. The role of data stewards complements that of data curators in the aspects of both metadata management activities and data governance.

Embargo
Embargo is the period during which a publication can be ‘closed’ while deposited in the repository (i.e. the publication is not openly available).

FAIR Data
FAIR Data Principle is the abbreviation of Fair, Accessible, Interoperable and Reusable, and is essentially a concise and measurable set of principles to support data sharing practices under the notion of Open Science.

Metadata
Metadata means "data about data". Metadata are the descriptors used for describing, tracing, use and management of the deposited item. Metadata describes characteristics such as content, quality, format, location and contact information. It can be used to describe physical items as well as digital items (documents, audio-visual files, images, datasets, etc.). Metadata can take many different forms, from free text (such as read-me files) to standardised, structured, machine-readable content. Some useful references that describe metadata can be found in ANDS and OECD websites.

Open Access
Open Access stands for free and unlimited access to research outputs including scholarly publications and books. There are several routes to making research outputs freely available such as gold, green and diamond routes.

Open Collaborative Platform
Open Collaborative Platform used by scientists to discuss potential solutions to a problem. An example of an open collaborative platform is The Polymath project.

Open Data
Open data is defined in essence, as data that can be freely used, re used and redistributed by anyone.
Besides commonly associated with Open Government Data, Open Data also refers to Open Business Data and Citizen Generated Data. The main criteria for Open data are complete, primary, timely, accessible, machine-processable, non-discriminatory, non-proprietary and license-free. An example of Open Data in the Open Science context is using or releasing the software code in the public domain for further re use, also known as Open Source Software.

Open Science
There is no general consensus on the definition of Open Science. Broadly, Open Science is an initiative to make research output (such as data, publications, etc.) more transparent and accessible. It is about extending the principles of openness to the whole research cycle based on cooperative work and new ways of diffusing knowledge through digital technologies and new collaborative tools. Open Science represents the necessary paradigm shift in scientific practices involving transformation in scientific culture, methodologies, institutions and infrastructures. Some examples of Open Science include Open Data, Open Access, Open Research Data, Open Collaborative Platform and Citizen Science.

EXECUTIVE SUMMARY

INTRODUCTION

Open Science has gained traction in recent years. It is a growing movement in making science open and making research output (such as data and publications) more easily accessible. This science movement demands that science be done in an open, and reproducible fashion. Various definitions of Open Science are found in the literature and they are based on the FAIR principle (i.e. Findable, Accessible, Interoperable and Reusable).

In this regard, Malaysia has embarked on this Open Science movement through the launching of the Malaysia Open Science Platform (MOSP). MOSP is aimed to gather and consolidate Malaysia's research data in a platform that allows accessibility and sharing of these research data. It is a strategic transformative initiative in strengthening country's STI collaborative ecosystem. Managed by the Academy of Sciences Malaysia through the Malaysia Open Science Alliance, MOSP is a two-year pilot project (2019 to 2021) involving five Research Universities in the country. The task is to look into the initial three main areas, namely, (1) Guidelines; (2) Capacity Building and Awareness; and (3) Infrastructure. One of the key deliverables for MOSP is to conduct a landscape study on Open Science in Malaysia. Inputs from the Open Science landscape reports are important to formulate the National Guidelines for Open Science.

MALAYSIA'S SCIENCE, TECHNOLOGY AND INNOVATION (STI): STI POLICY AND OPEN SCIENCE INITIATIVES

Malaysia has an overarching goal of becoming a high tech and advanced nation that is inclusive and sustainable. From this perspective, Science, Technology and Innovation (STI) is unquestionably important. It will assist in ensuring the establishment of a scientifically advanced and progressive society. In this context, Malaysia endeavours to strengthen and mainstream STI in all sectors and levels of national socio-economic development. As a vital ingredient to increasing productivity and ultimately raising the competitiveness of the economy, STI will enable the exploitation of new ideas under the new normal the country is currently adapting. A new STI policy has been formulated to fulfil the goal of a higher technological country by 2030. More importantly is making Malaysia not only a consumer of technology but also a contributor to scientific and technological advancement.

The promotion and role of Open Science in supporting the new STI Policy are indicated in some of the strategies. The direct ones are a) S6 (Open Data Sharing); b) S16 (Increased Effectiveness of STI Outreach Programme); and c) S17 (Mainstreaming Science Communication). There are also other indirectly related strategies associated with Open Science such as a) S1 (Enhanced STI Governance); b) S7 (Ventures and Collaborations in Overcoming National Challenges); c) S8 (Driving Social Innovation); d) S14 (Expands STI Talents Participation In Workforce); e) S18 (Coordinate International STI Collaborations); and f) S20 (Strengthened Network of International Collaborations).

Some of the Open Science and MOSP activities that have been implemented in the country include, among others, the following:

i. **Training Module Evaluation Workshop (July 24th & August 25th, 2020)** to discuss, consolidate and finalise contents for training module development for Open Science skills and Data Stewardship.
ii. **Success Stories Sharing Session on Data Sharing Platform (July 15th, 2020)** conducted virtually to exchange views about data sharing platform.

iii. **First Stakeholder Engagement Workshop for Open Science (June 4th, 2020)** which was conducted virtually with the objective to elicit and consolidate inputs from participants for the formulation of National Guidelines for Raw Research Data Sharing in Malaysia.

iv. **Participation in an Online Data Stewardship Training (June 2020 – September 2020),** to train 12 "masters trainers" who will be the certified instructors for the Training of Trainers Program for Data Stewardship on Open Science. The To program aims to train 200 data stewards in Malaysia by the end of 2021.

v. **Policy writing workshop (May 20th to 21st 2020), which was** conducted virtually to consolidate and finalise the first findings for landscape study of Open Science in Malaysia.

vii. **UNESCO Virtual Ministerial Dialogue (March 30th, 2020)** where YB Khairy Jamaluddin shared his view on COVID-19 and on how Open Science can assist through the creation of an international repository not only for medical purposes but also for the development of technological innovations.

viii. **Open Science Forum for Asia and the Pacific Region (February 13th, 2020)** aimed to develop common Open Science understanding between the Asia Pacific countries through engagement and knowledge sharing among the professionals.

ix. **Launching of MOSP and i-Connect initiatives (November 7th, 2019)** of which the MOSP initiative was launched by the Minister of Energy, Science, Technology, Environment and Climate Change (MESTECC).

The **MOSP's Future Activities** were also planned after the completion of the 2-year pilot project of which MOSP will connect with existing infrastructure and data sharing platforms such as those located outside five Research Universities, Open Government Data by Malaysian Administrative Modernisation and Management Planning Unit (MAMPU) and Raw Data for Research and Sciences (RADARS) under MOSTI. In the longer term, MOSP will connect with platforms in other countries and the Asia Pacific region and be equipped with big data analytics and real-time analysis features.

INTERNATIONAL PRACTICES FOR OPEN SCIENCE

Besides mapping out the existing Open Science activities and players in Malaysia, gauging the level of readiness of local institutions in adopting Open Science apart from identifying factors enabling and hindering Open Science adoption, a benchmarking exercise was conducted to identify global indicators to measure Open Science readiness. A number of countries and their agencies were selected such as the Australian National Data Service (ANDS), European Open Science Cloud (European Commission), National Open Science Platform (Netherlands), CODATA (Committee on Data of the International Science Council of ISC) and the African Open Science Platform (AOSP). Based on the benchmarking exercise, Australia's ANDS is regarded as the best model for Malaysia's MOSP to learn from in the perspective of the government, researchers, capacity building and infrastructure.

SOME IMPORTANT FINDINGS ON THE LANDSCAPE STUDY

Open science represents a change in the way key actors or stakeholders in research, education and knowledge exchange communities create, store, share and deliver the outputs of their activity. For Open Science principles, in order for policies and practices to be fully embraced in universities, there needs to be a holistic approach to organise each Open Science initiative, which ultimately is integrated and must be coherent with the overall goal of Open Science, is understood across all levels especially by researchers, and involves everyone's participation.

PERSPECTIVES FROM TOP MANAGEMENT OF RESEARCH UNIVERSITY

This part of the study sets out the perspectives on Open Science from Top Management Research University (RU), delving into institutional interests and activities in Open Science, and Vision for participation in Open Science initiatives including MOSP. Some of the key takeaways are:

i. Publishing Open Access articles and in Open Access Journals is a common practice for Research Universities in Malaysia. However, other essential elements of Open Science such as Research Data Management Plan is new to the universities; and

Ii. The Top Management of the five Research Universities are fully supportive towards implementing Open Science in Malaysia.

PERSPECTIVES FROM RESEARCHERS

Efforts to promote and implement Open Science practices are to large extent driven by the willingness and motivation of researchers to contribute and participate in data sharing. There is, in fact, ample evidence and success stories demonstrating that the success implementation of Open Science is characterised by researchers' free willingness in sharing their data. The study provides an examination of the current awareness, practices and required skills and supports as well as concerns related to Open Science from the perspective of researchers. The key takeaways are as follows:

i. Researchers have different levels of understanding about Open Science practices.

ii. Researchers are not familiar with the Findable, Accessible, Interoperable and Reusable (FAIR) principle.

iii. Researchers perceived Open Science and data sharing practices will greatly benefit their research endeavours.

iv. Researchers are not familiar with, and generally are not aware of Open Science related policies in the institution that they are working at.

v. Majority of researchers have published their research works in Open Access journals, as well as sharing their research findings as supplementary materials in journals, conferences and when receiving requests from emails, direct contacts or sharing via websites.

vi. Majority of researchers prefer to store their research data on personal storage.

vii. Majority of researchers described training for research data management plan is the skill they need the most

viii. Researchers produced a wide spectrum of datasets across multi research disciplines

ix. When sharing research data, researchers are worried about their research being scooped, ownership of the created research data and misuse and misinterpretation of shared data by third parties.

x. Researchers were also concerned about current traditional metrics for research performance that are not supported for Open Science and data sharing practices.

PERSPECTIVES FROM LIBRARIANS

In an academic entity or research institution, librarians are actively involved in providing various research support services for the entire research process to researchers and students. For example, librarians play a critical role in organising training programmes, scheduling webinars, classes or training on data management as well as providing consultation services for students and researchers to conduct research, collect data, and write theses or manuscripts, and to provide advice on publishing. Librarians are also involved in managing Open Access publications and institutional repositories. Key findings are:

i. The role of data curators and data stewards is very important in Open Science as they provide support to researchers in various aspects especially management of research data lifecycle;

ii. Majority of librarians in Research Universities are still performing their traditional roles as a librarian with few of them having started engaging in performing duties as a data steward;

iii. Training for data stewardship is needed to equip librarians with the essential skills supporting Open Science; and

iv. A strong support from the top management is necessary to drive the transformative role of librarians as data stewards.

PERSPECTIVES FROM INSTITUTIONAL REPOSITORIES

Institutional repository is defined as a library of digital objects and associated metadata from a single institution. Generally, repositories can hold structured, unstructured or even semi structured data. Structured data is data that adheres to a pre-defined data model and is therefore straightforward to analyse. Structured data, which usually is processed data, resides in relational databases or data warehouses and are easily searchable. On the other hand, unstructured data does not have a predefined data model or is not organised in a pre-defined manner. Text, raw images, audio and video files are some examples of unstructured data which reside in application or noSQL databases.

Recent interviews with Head of Data Centres or representatives from the five Research Universities in Malaysia, have established well-defined institutional repositories that are maintained by Library units. A majority of institutional repositories in research universities were developed using EPrints open-source repository software. The real strength of using EPrints lies in the flexibility and easiness of use for data depositors and system administrators. The process of depositing materials into EPrints is simple and easy as data depositors only have to key in metadata, for example type of document, title, name of authors using the form available on the website. Furthermore, knowledge about HTML/XML is not required for the purpose of data deposition. EPrints also allow depositors to submit suggestions to administrators if they would like to update or delete deposited materials. Lastly, EPrints offers some flexibility in metadata customisation that allow system administrators to tailor the standard according to their needs.
Some key takeaways are:

i. Most institutional repositories in five Research Universities in Malaysia use Dublin Core as the metadata standard;

ii. Most institutional repositories house structured data and the establishment of research data repository is quite new in these Research Universities; and

iii. From consultations with researchers and other stakeholders in Malaysia, they emphasise the importance of establishing a trusted repository to encourage data sharing in the country.

CONCLUSION AND RECOMMENDATIONS

The mandate and strategic thrusts embodied by the National Policy on Science Technology and Innovation (2021 – 2030) are pushing towards the full implementation of Open Science and data sharing practices in Malaysia. Realising this goal will require holistic and joint interventions by multi-stakeholders such as researchers, top management universities, research funder organisations, publishers, government agencies, industries and learned societies, aimed to nurture and encourage a research data sharing ecosystem in the country.
This study has provided some useful insights into the current landscape of Open Science in Malaysia based on six indicators: (1) Relevant acts, policies and guidelines locally and internationally, (2) Number of repositories, (3) Skill capacity, (4) Infrastructure capacity, (5) Awareness, and (6) Current Open Science activities from perspectives of five target groups, namely (1) Deputy Vice Chancellors (Research & Innovation), (2) chief librarians and librarians, (3) Head of data centres or

representatives of institutional repositories, and (4) junior and (5) senior researchers. The inputs were validated with engagements with a broader set of stakeholders including Institutes of Higher Learnings, Public and Private Research Institutes, government agencies and ministries, legal units, libraries, research funders, and industries.

Based on the consolidated inputs solicited from those involved in the landscape study and engagements, this document further sets out 9 recommendations to support a rapid and effective implementation of Open Science in Malaysia through coordinated actions within and among these stakeholders.

a) The provision of National Open Science Policy;

b) Guidelines for Implementing Open Science and MOSP;

c) Empowering the Role of Funding Body for Open Science;

d) Building a trusted and interoperable research data sharing platform;

e) Identifying funding streams to sustain MOSP operation;

f) Reform existing academic rewards system to incentivise Open Science and data sharing practices;

g) Training for Open Science knowledge and skills, including data stewardship;

h) Effective communication about Open Science and its incentives; and

i) More resource allocation for conduct of research projects.

WAY FORWARD

In making Open Science a reality for the country in terms of its readiness apart from making Open Science more visible in the National Policy for Science, Technology and Innovation (NPSTI) is enhancing the effectiveness of the Malaysian Open Science Platform (MOSP). As a way forward for Open Science, some proposed implementation actions needed to be done and are as follows:

a) Making Open Science more visible in the National Policy;

b) The need to develop Open Science Readiness Indicators to monitor the country's state of readiness on Open Science;

c) Provision of strong funding to support Malaysian Open Science Platform;

d) Competent Human Capital and Top Management;

e) Right Positioning in the Government Institutional Framework.

CHAPTER 1 INTRODUCTION

Open Science is an initiative or movement to make research output such as data and publications more transparent and accessible. It is about extending the principles of openness to the whole research cycle based on cooperative work and new ways of diffusing knowledge through digital technologies and collaborative tools.

Box 1.1. Various definitions of Open Science

1. **OECD[2]:** "...efforts by researchers, governments, research funding agencies or the scientific community itself to make the primary outputs of publicly funded research results – publications and the research data – publicly accessible in digital format with no or minimal restriction as a means for accelerating research; these efforts are in the interest of enhancing transparency and collaboration and fostering innovation."

2. **FOSTER[3]:** "...the practice of science in such a way that others can collaborate and contribute, where research data, lab notes and other research processes are freely available, under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods."

3. **RAND Corporation[4]:** "Open Science refers to ongoing changes in the way research is conducted: for scientists themselves, through increasing the use of open access scientific publishing and open data, and for the public, through increasing their understanding of and participation in science ... Open Science is one of three priority areas for European research, science and innovation policy."

Source: Compiled by ASM (2020)

Open Science allows research outputs to be available to access, use and share with certain conditions, depending on types of data. In other words, for those research outputs to be accessible and can be shared by "everyone", they should be properly managed and curated, meeting the principles of Findable, Accessible, Interoperable and Reusable (FAIR). FAIR data means that data is not always open, but it should be as open as possible, and as closed as necessary. With FAIR data, researchers can create, share and reuse quality, valuable, high integrity and responsible data, fueling scientific progress to its fullest potential.

- **Findable** means that data and metadata are easily found by both humans and computers. Usually this task is enabled by machine-readable persistent identifiers and metadata.
- **Accessible** means that data can be retrieved using the outline protocols by appropriate people, at an appropriate time and in an appropriate way. Data can be FAIR even if the data has different levels of accessibility, such as:

  (1) Data is completely private;

  (2) Data is accessible by a defined group of people; and

  (3) Data is accessible by everyone

[1] European Commission. (2016) Open Innovation, Open Science, Open to the World – A vision for Europe. Page 33.

[2] OECD. (2015) Making Open Science a Reality. OECD Science, Technology and Industry Policy Papers. No. 25. OEDC Publishing, Paris.

[3] FOSTER. Open Science. (Online) Available from https://www.fosteropenscience.eu/taxonomy/term/100

[4] RAND Corporation. Monitoring Open Science Trends in Europe. (Online) Available from https://www.rand.org/randeurope/research/projects/open-science-monitor.html

[5] FAIR data is similar to Open Data as both follow the philosophy of data sharing and encouraging collaborations among data users. Despite the similarities, the main difference between FAIR data and Open data is the access, use and sharing terms to shared data. Open Data means that by default, data should be available to everyone to access, use and share without any implications and restrictions on patents, copyrights and licences (Source: Open Knowledge Foundation 2011).

[6] It includes reliability and accuracy of data

- **Interoperable** means that the terminology system, protocols, standards and formats built and employed for datasets that are stored in a platform can be used and can communicate with other tools or platforms.

- **Reusable** means that data is well-defined and can be used for different purposes and in different settings, and the legal use is regulated by different terms and conditions. Data can be FAIR even if the data has different levels of reusability, depending on the stipulated licensing terms (E.g. acknowledgement, access and methods of data reuse, charges, exemption use of personal, sensitive and restricted data and proprietary information)

Box 1.2. FAIR guiding principle

To be Findable:
F1. (meta)data are assigned a globally unique and persistent identifier
F2. data are described with rich metadata (defined by R1 below)
F3. metadata clearly and explicitly include the identifier of the data it describes
F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:
A1. (meta)data are retrievable by their identifier using a standardized communications protocol A1.1. the protocol is free, open and universally implementable
A1.2. the protocol allows for an authentication and authorization procedure, where necessary
A2. metadata are accessible, even when the data are no longer available

To be Interoperable:
I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation
I2. (meta)data uses vocabularies that follow FAIR principles
I3. (meta)data include qualified references to other (meta)data

To be reusable:
R1. (meta)data are richly described with a plurality of accurate and relevant attributes
R1.1. (meta)data are released with a clear and accessible data usage license
R1.2. (meta)data are associated with data provenance
R1.3. (meta)data meet domain relevant community standard

(Wilkinson, M. et al, 2016)

Several scientific, civic, economics and societal reasons underpin the progress of Open Science today:

1. Sharing scientific knowledge makes research more discoverable, more visible, and less redundant.
2. Open Science changes the way research is conducted, promoting collaborative research across disciplines and research actors, strengthening research integrity, as well as stimulating active participation of citizens in data collection. In a wider society, Open Science helps restore trust between citizens and scientists.
3. Accessibility to datasets and results changes the way to the use of artificial intelligence (AI) powered tools such as text and data mining, and automatic learning, to discover useful information within a database.
4. Researchers are gaining control over costly pay-to-publish and pay-to-read systems by sharing and disseminating research data on a repository platform.

Ultimately, Open Science leads to Open Innovation where the fruits of Open Science become more interconnected and are more readily geared towards rapid translation of R&D discoveries. Through Open Science and Open Innovation, a broad and more effective engagement and participation of

[7] Scope of research data to be shared need to be defined clearly i.e. published R&D data, R&D project management data, raw data and structured data in database.

stakeholders in the quadruple helix can be achieved, which then facilitates effective commercialisation of new knowledge to benefit the society and the country as a whole.

Open Science initiatives have been launched and implemented in many countries such as Europe, Australia, United States, Japan, to name a few. Additionally, Open Science has been championed and advocated by international bodies such as International Science Council (ISC), Organisation for Economic Co-operation and Development (OECD), United Nations Educational, Scientific and Cultural Organization (UNESCO) and the World Bank.

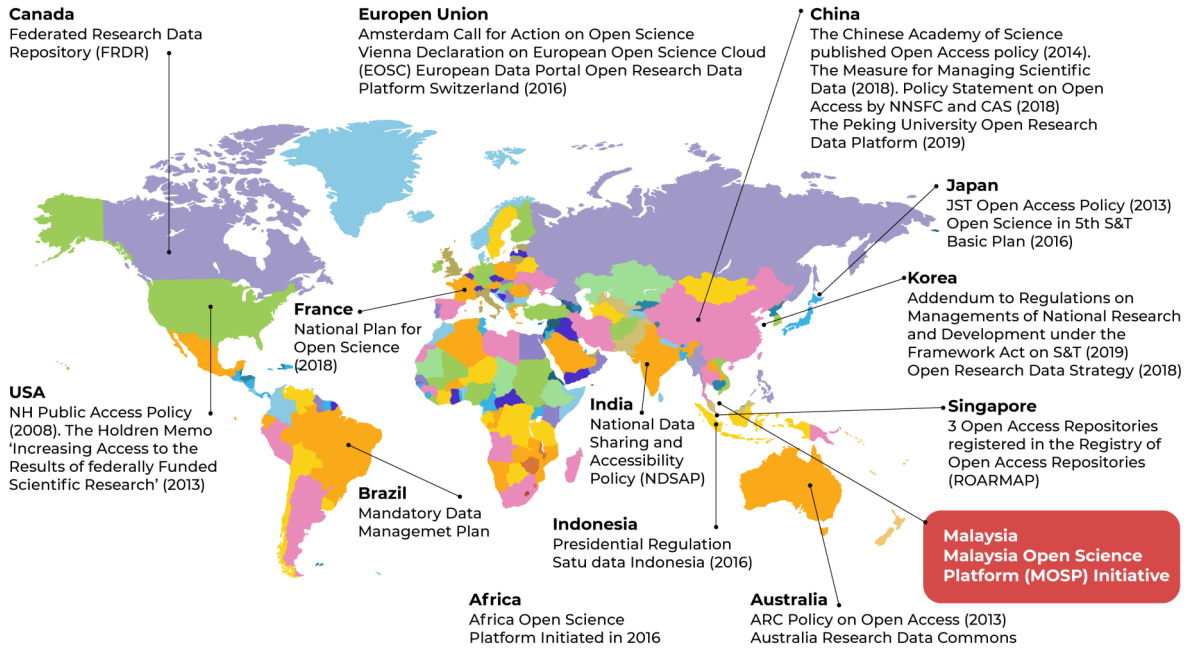


Figure 1.1 Open Science initiatives from the global perspective. Source: Compiled by ASM, 2020.

BOX 1.3 The Economic Benefits of Open Science

A key political driver of open access and open science policies has been the potential economic benefits that they could deliver to public and private knowledge users. Since Open science is just good science it will benefit those (including the public, government, and the private sector) that benefitted from good science. Some of these benefits are as follows.

Efficiency: Efficiency gains occur since academics, companies, policymakers, etc. do not have to pay to access findings and data, and access will become less time-consuming that lead to savings on labour costs. Further, the effectiveness and productivity of the research system is improved through greater access to scientific inputs and outputs by 1) cutting down duplicates and the reducing costs of "creating, transferring and reusing data"; 2) enabling more research to be conducted using the same data; 3) augmenting prospects or chances in the research process in terms of participation at the domestic and global levels; and 4) assist in stepping up the efficiency of research and of its diffusion by users of open search tools (The Royal Society, 2012).

Economic benefits: will foster spill overs to scientific and innovation systems with increased access to research results besides increasing awareness to consumers. In today's knowledge economies, science assumes a significant role (The Royal Society 2012:19) in terms of higher efficiency associated with it which benefits both the developed and the developing economies.

Innovation and knowledge transfer: Open Science promotes a smoother pathway from research to innovation to produce new products and services since it helps to hasten the re-use of results of scientific research including articles and data.

Public disclosure and engagement: Since Open Science helps to promotes citizen's engagement and active participation in scientific activities and data collection it will promote awareness among citizens and assisted in building trust and support for public policies and investments. Moreover, if the outcomes of public funded research provide evidence of benefits to them.

Global benefits : As Open Science is a global movement it encourages worldwide concerted endeavours besides allowing much speedier knowledge transfer that demands or necessitate coordinated international efforts and actions to tackle global challenges such as climate change, global warming, poverty or the ageing population. It could support in providing a more immediate and effective solutions in overcoming these challenges.

Source: (OECD, 2015:18) https://www.fosteropenscience.eu/content/what-are-benefits-open-science

Malaysia has recently embarked on this Open Science movement. Launched on 7th November 2019 by the Minister of Science, Technology, and Innovation (MOSTI, then MESTECC), the Malaysia Open Science Platform (MOSP) aims to gather and consolidate Malaysia's research data which are valuable national assets in a platform that would enable accessibility and sharing of these research data in accordance with the FAIR principle. MOSP represented as a strategic transformative initiative to strengthen Malaysia's STI collaborative ecosystem towards achieving Shared Prosperity Vision 2030 and addressing the United Nations Sustainable Development Goals.

MOSP is a two-year pilot project (2019 to 2021), managed by the Academy of Sciences Malaysia through the Malaysia Open Science Alliance. The pilot project will be involving the five Research Universities in Malaysia, i.e., Universiti Malaya, Universiti Kebangsaan Malaysia, Universiti Putra Malaysia, Universiti Sains Malaysia and Universiti Teknologi Malaysia.

The task is to look into the initial three main areas, which are:

1. Policy and Guidelines;
2. Capacity Building and Awareness;
3. Infrastructure

The development of National Policy and Guidelines for Open Science is a commitment for the Malaysia Open Science Alliance Working Group on Policy and Guidelines. Upon the cabinet decision on August 14th 2020, a key deliverable of Malaysia Open Science Platform was changed to National Guidelines for Raw Research Data Sharing, and no longer Policy. As the work carried out by the Working Group is intricate by nature, the formulation of the Policy document must consider and consolidate inputs of all relevant stakeholders including researchers, top management universities, government agencies, libraries, research funder organisations publishers, legal units, industries and research managers. This must happen in strong coherence with the mission, the vision and strategic thrusts outlined in the National Policy on Science, Technology and Innovation (NPSTI 2021-2030).

As the first step towards progressing the formulation of National Guidelines for Raw Research Data Sharing, the Working Group has studied the landscape and awareness of Open Science in Malaysia. Specifically, the Working Group has conducted a literature study, surveys, interviews and engagements with targeted groups on the following items:

1. Relevant acts, policies and guidelines locally and internationally
2. Number of repositories
3. Skill capacity
4. Infrastructure capacity
5. Awareness
6. Current Open Science activities

CHAPTER 2 OPEN SCIENCE LANDSCAPE STUDY

2.1. Background and Objectives

2.1.1. Background

Under the overall supervision of the Academy of Sciences Malaysia and Malaysia Open Science Alliance, and support by Ministry of Science, Technology and Innovation (MOSTI), the Malaysia Open Science Alliance Working Group on Guidelines conducted a landscape study on Open Science in Malaysia apart from formulating a document on National Guidelines for Open Science.

This landscape study is conducted by the Malaysia Open Science Alliance Working Group on Guidelines to assess the landscape of Open Science in Malaysia. This landscape study is aimed for researchers, librarians, top management of Research Universities and heads of data centres, who are among the primary stakeholders of the Malaysia Open Science Platform. The study gauged the respondents' knowledge, awareness and participation in Open Science activities. The study is also part of the Malaysian Open Science Platform (MOSP) Pilot Initiative that involves 5 Research Universities (Universiti Malaya, Universiti Sains Malaysia, Universiti Putra Malaysia, Universiti Kebangsaan Malaysia and Universiti Teknologi Malaysia).

2.1.2. Study Objectives

Four main objectives:

a. To **map existing Open Science** activities and players in Malaysia;
b. To **identify factors enabling and hindering** Open Science Adoption;
c. To **gauge the level of readiness** of local institutions to adopt Open Science; and
d. To undertake **benchmarking** exercises to identify global indicators to measure Open Science readiness.

2.2. Methodology

This Study adopted a Sequential Exploratory Strategy approach as depicted in Figure 2.1. It began with an initial exploratory phase where data (both qualitative and quantitative) were collected primarily through survey interviews with relevant stakeholders. The interviews conducted allowed for validation of quantitative data as well as provided more opportunity in understanding issues in great detail. Both of these exercises were supplemented with workshops and meetings to further validate and investigate specific challenges aside from obtaining additional inputs especially on international best practices relating to the Study. Briefly, this approach tried to offset the significance of capturing a variety of topics and ideas from the stakeholders with the need for statistical rigour in order to make valid inferences and conclusions.

2.2.1 Benchmarking exercise and desk study

Benchmarking exercise was done by undertaking desktop research on global Open Science initiatives to learn about their current status of Open Science and initiatives involved, as well as their motivations, challenges and strategies towards implementing Open Science in the short and long run. Engagements with the global Open Science champions have also been carried out to further explain their Open Science efforts from various aspects. This exercise has identified the following global indicators to be used for the Landscape Study:

a. Available Policies
b. Number of Repositories
c. Skill Capacity
d. Infrastructure Capacity
e. Awareness
f. Current activities

In addition to benchmarking analysis, this Study also reviewed relevant acts, policies and guidelines related to data sharing in Malaysia.

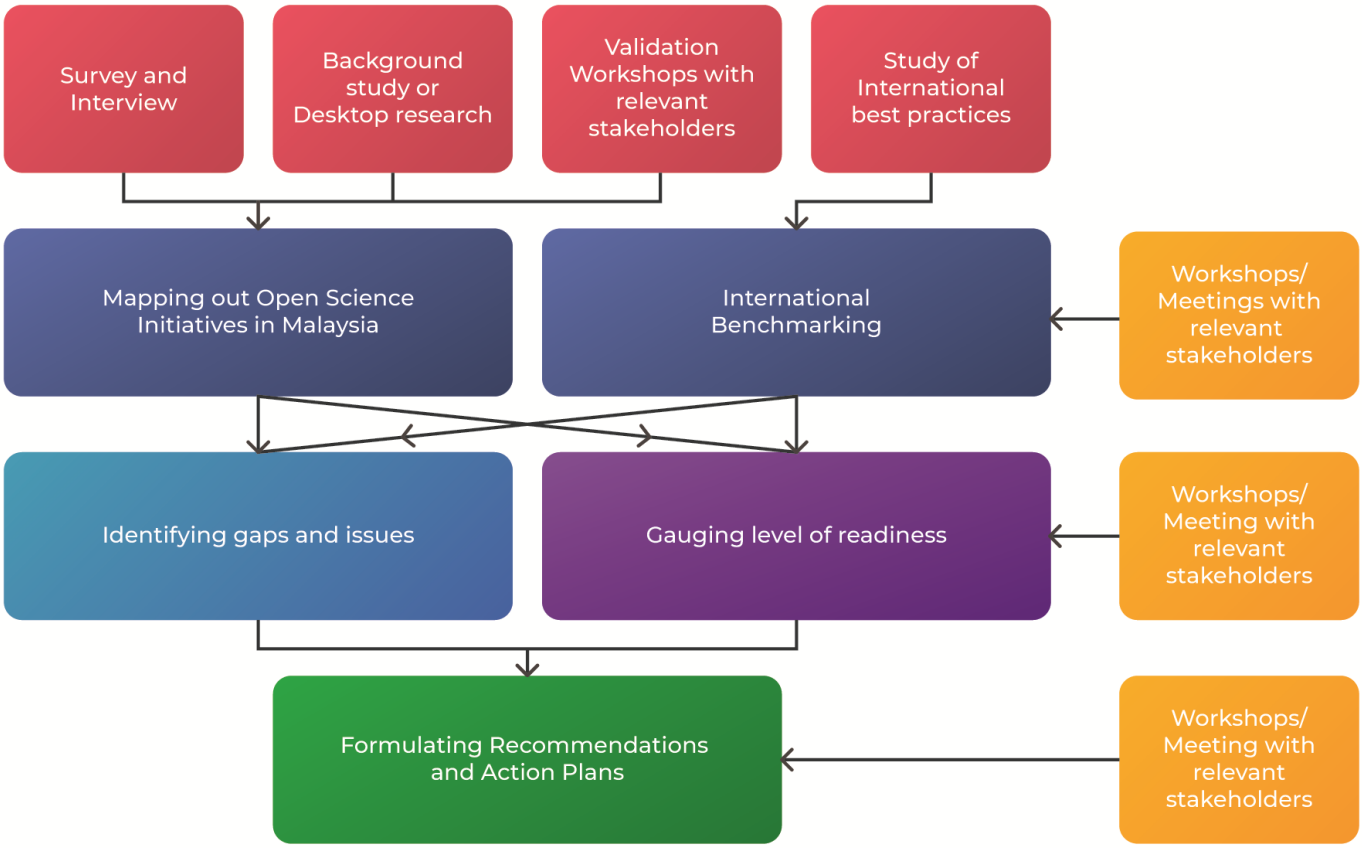


Figure 2.1 A Sequential Exploratory Strategy Approach to Overall Study (Source: ASM, 2020)

2.2.2. Web-based questionnaires

Survey was conducted from 15th April 2020 to 8th June 2020. Questionnaires were drafted and approved by the Malaysia Open Science Alliance and the Working Group on Guidelines. Survey questions were then developed into an online survey using the SurveyMonkey application. Two independent sets of questions were developed for each target group (i.e. researchers and librarians).

a. Questionnaire for researchers consists of 24 questions probing on familiarity with open science generally and in their institution, deposition and sharing of research data, and support for open science **(Appendix 2.1)**.

b. Questionnaire for librarians consists of 15 questions on open science, data management plan, and data steward ship or curation **(Appendix 2.2)**.

The survey link for researchers was sent to the Deputy Vice-Chancellors (Research & Innovation) for Institutes of Higher Learnings (IHLs), Head of Research Division or equivalent for Public and Private Research Institutes while the survey link for librarians was sent to chief librarians of respective IHLs and RIs. A total of 300 respondents were targeted which involved researchers, chief librarians and librarians. The list of target Institutes of Higher Learning (IHLs), Public and Private Research Institutes for survey as in **Appendix 2.3**.

2.2.3 Interviews

Several sets of interview questions were designed and used for intended groups as follows:

a) Top Management Universities (**Appendix 2.4**);
b) Junior Researchers (**Appendix 2.5**);
c) Senior Researchers (**Appendix 2.6**);
d) Chief Librarians (**Appendix 2.7**); and
e) Head of Data Centres or Representatives from Institutional Repositories (**Appendix 2.8**).

Semi-structured interviews were then carried out with Deputy Vice Chancellors (Research & Innovation), chief librarians, head of data centres, and junior and senior researchers from 5 Research Universities. The universities involved were Universiti Malaya, Universiti Sains Malaysia, Universiti Kebangsaan Malaysia, Universiti Putra Malaysia and Universiti Teknologi Malaysia. In the conduct of the interview, participant information and consent were obtained (see **Appendix 2.9** and **Appendix 2.10**).

Interviews were conducted via online video call mediums such as Zoom, Skype and Google Meet. The interviews were recorded (with consent) and were transcribed verbatim. Interview transcripts were analysed using thematic analysis.

2.2.4 Validation workshops/discussions

A series of validation workshops and discussions has been conducted virtually or through face-to-face meetings with various individuals and institutions, among others:

a. To explain about the study on Open Science and data collection process as well as deliberating and getting some key information from participants.
b. To validate initiatives and readiness level of institutions on Open Science.
c. To exchange views on Open Science practices in various countries and on current Malaysia Open Science Platform programs and activities.

2.3. Scope and Limitations of study

2.3.1 Scope

It is expected, among others, that the landscape report will:

a. Reflect (or make a reference to) the readiness level of Open Science in Malaysia, by incorporating perspectives of top management, researchers, librarians and head of data centres from public and private Institute of Higher Learnings, Public and Private research institutes.
b. Emphasise, among others, the current state of research data sharing policy and other open-science related policy, capacity building and infrastructure.
c. Incorporate Open Science initiatives and practise in other countries, such as Australia, Africa, Europe, to benchmark with.
d. Provide sets of recommendations and suggest ways forward for Open Science implementation in Malaysia in the future.

2.3.2 Limitations of Study

Some of the limitations identified were:

a. The analysis and scope of the study were limited to exploring the important general issues related to policy, especially as they relate to the objectives outlined above. This limited inquiry was mainly to public Research Universities as opposed to covering all higher education institutions and public research institutions in Malaysia.

b. Open Science actors such as management leaders in IHLs, Public and Private Research Institutes and researchers, along with policymakers and stakeholders, were very generous with their time, exemplified during the interviews and their willingness to complete surveys (the research team received a total of 302 responses). However, we acknowledged that there were circumstances when respondents had to fill out surveys or were interviewed in the past and have shown the so-called “survey fatigue” which is a real problem not only in eliciting responses but may also affect the quality of the responses themselves. Because the survey was one of the first of its kind, there are few other sources to compare the results with.

c. Though the survey was administered to a randomly chosen group of respondents with relatively high response rates, there remains the possibility of selection bias.

d. The surveys themselves were designed based on the data collected in a relatively few number of respondents and interviews. These methods and techniques adopted may be well-accepted and appropriate for this kind of preliminary study though there remains the possibility that some elements of Open Science were, though not Intentionally, missed or left out.

e. One of the biggest challenges faced during the study was trying to combine the perspective of the stakeholders on Open Science based on interviews and surveys. As discussed in the findings, Institutes of Higher Learning, Public and Private Research Institutes were asked a list of questions on Open Science that may also need other stakeholders' inputs with whom they work. While this situation has caused some concerns, the research team members overcame it by organising a series of stakeholder meetings and workshops to obtain additional input and information on Open Science.

f. Unfortunately, since the study was affected by the Movement Control Order (MCO), all interviews, and some engagements were done virtually. As the activities were carried out online, the research team acknowledged technical difficulties that have occurred during the engagements which may have halted participants from actively participating in conversations. The research team also expected for better and more effective engagements with participants if the interviews and workshops were done physically, leading to more meaningful discussions among participants and the research team.

CHAPTER 3

OVERVIEW OF MALAYSIA'S SCIENCE, TECHNOLOGY AND INNOVATION (STI): STI POLICY AND OPEN SCIENCE

3.1NATIONAL POLICY ON SCIENCE, TECHNOLOGY AND INNOVATION

3.1.1. Background

Malaysia has achieved remarkable macroeconomic success since attaining independence in 1957. From an economy that was entirely dependent on primary commodities, Malaysia has successfully transformed over the course of its brief history to become a multi-sector economy with manufacturing and services propelling its growth to its current emphasis on innovation-led economy. The country has recognised the importance of science, technology and innovation (STI) in contributing to its past successes in terms of its socio-economic development and sustainable growth for the last few decades.

Moreover, Malaysia has an overarching goal of becoming a high income and advanced nation that is inclusive and sustainable by 2020. From this perspective, science, technology and innovation (STI) is unquestionably important in ensuring the establishment of a scientifically advanced and progressive society as embodied in Vision 2020. Malaysia would like to see a society that is creative and innovative not only as a consumer of technology but also a contributor to the scientific and technological advancement of the country.

In this context, Malaysia endeavours to strengthen and mainstream STI in all sectors and levels of national socio-economic development. STI is a vital ingredient to increasing productivity and ultimately raising the competitiveness of the economy. Through STI and the exploitation of new ideas, additional values can be captured from the same base of capital and human resource.

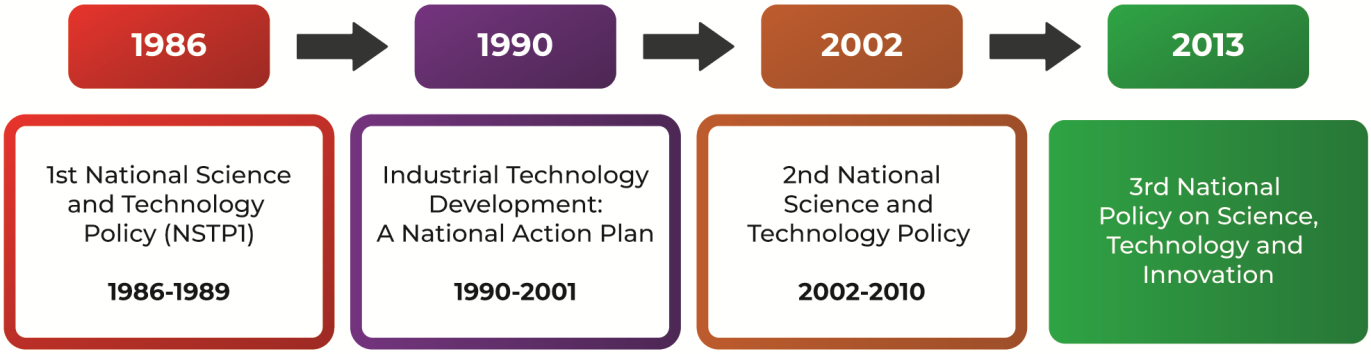


Figure 3.1 Key milestones in Malaysia's STI

Public awareness of Science and Technology (S&T) is critical for the overall development of society and nation. The challenge is the enculturation of STI to create a society that is not only comfortable with S&T, but also able to understand, absorb and utilise STI to its highest potential.

Inculcating a culture of STI at all levels is critical to enhance the scientific, creative and innovative thinking among Malaysians. STI should be imbued naturally and practised in their daily life. A strong commitment by the stakeholders is critical to promote, support and popularise STI related programmes.

In order to make informed decisions, the community should embrace S&T and be exposed to the advancements and issues in S&T through various means including dialogues and public forums. This will enable the community to become informed users responsible for the impact of S&T in society.

The Ministry of Science, Technology and Innovation (MOSTI), in collaboration with its strategic partners, has been and will continue to lead and implement various programmes to stimulate and cultivate creativity and innovation. This effort is in line with the declaration of the "Decade of Innovation", announced by the Deputy Prime Minister on 5th November 2012 in conjunction with the "World Innovation Kuala Lumpur Forum" in 2012. The Decade began in 2010 with the launching of several programmes such as the Year of Creativity and Innovation, Malaysia 2010 or "Innovative Malaysia 2010".

To ensure inclusive development, various programmes and knowledge-intensive activities, creativity and innovation continued to be encouraged and mainstreamed in all sectors and all walks of life. Through Budget 2013, this inclusive initiative was taken up by MOSTI in close collaboration with Malaysia Innovation Agency (AIM) and the NGOs. Performance of Malaysia's STI has been high among developing economies though it still lagged behind the advanced economies such as the OECD and the US.

3.1.2. Snapshots of Malaysia's Achievements and Performance

3.1.2.1. Domestic Performance and Achievements

Gross Expenditures on R&D (GERD)

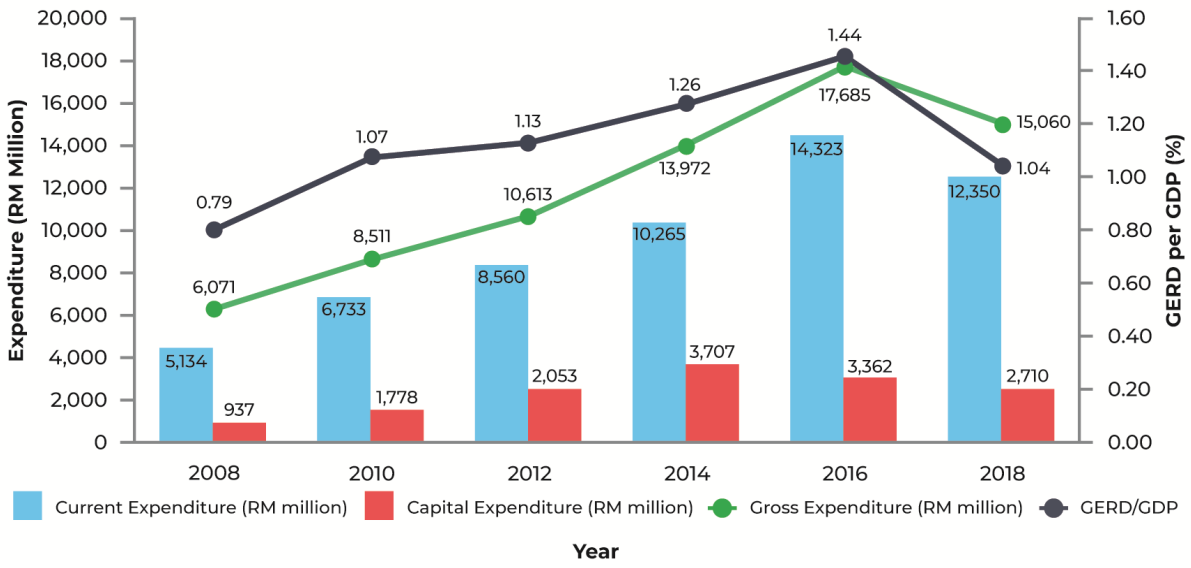


Figure 3.2. Expenditure on Research and Development (2008 – 2018)
Source: National Research and Development Survey in Malaysia 2019, MASTIC (2020)

Gross Expenditures on R&D (GERD)

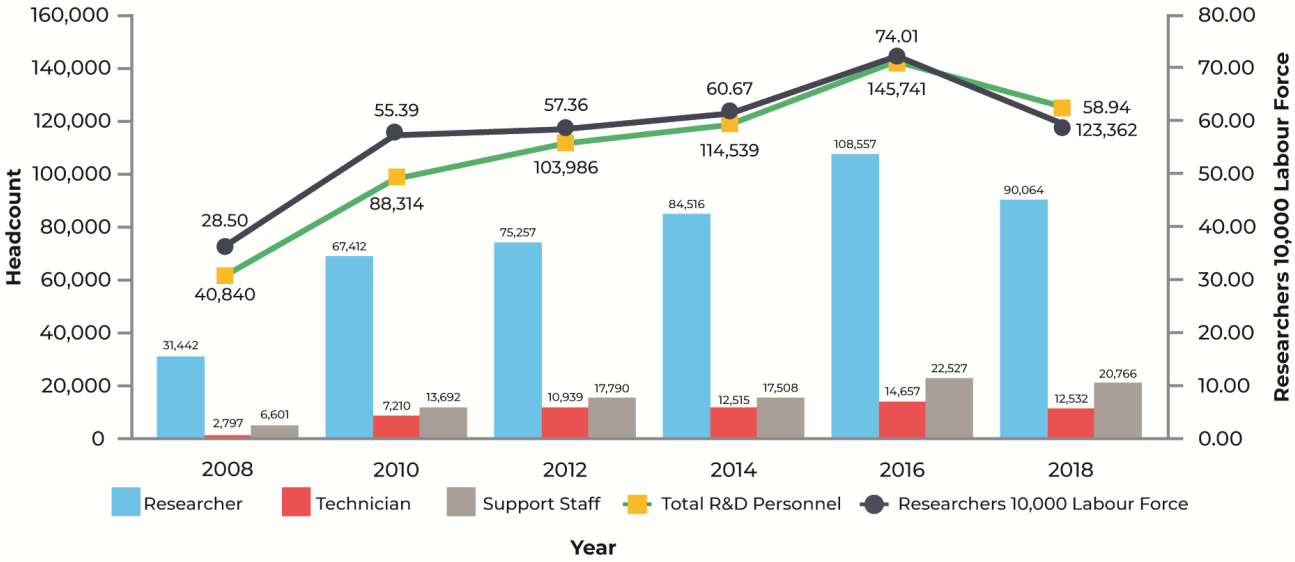


Figure 3.3. R&D Personnel and Researchers (2008 – 2018)
Source: National Research and Development Survey in Malaysia 2019, MASTIC (2020)

Publications

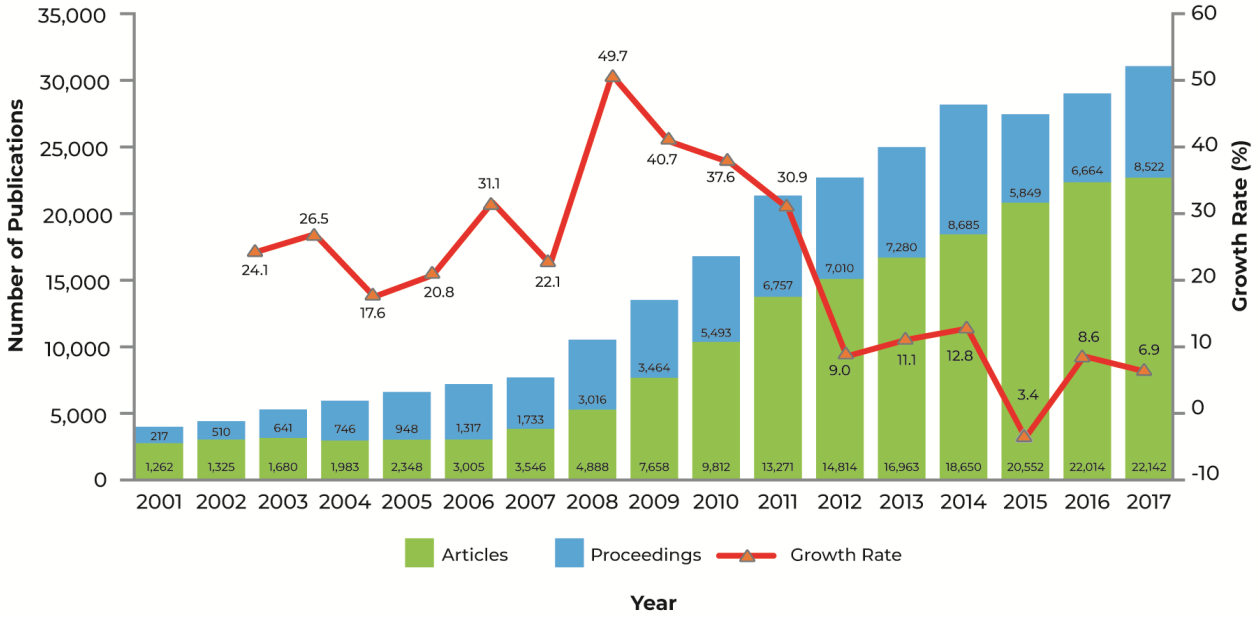


Figure 3.4. Yearly Publications Output and Growth Rate (2001 – 2017): Scopus
Source: National Bibliometric Study 2001-2017, MASTIC (2020), as of 24th August 2018

Citations

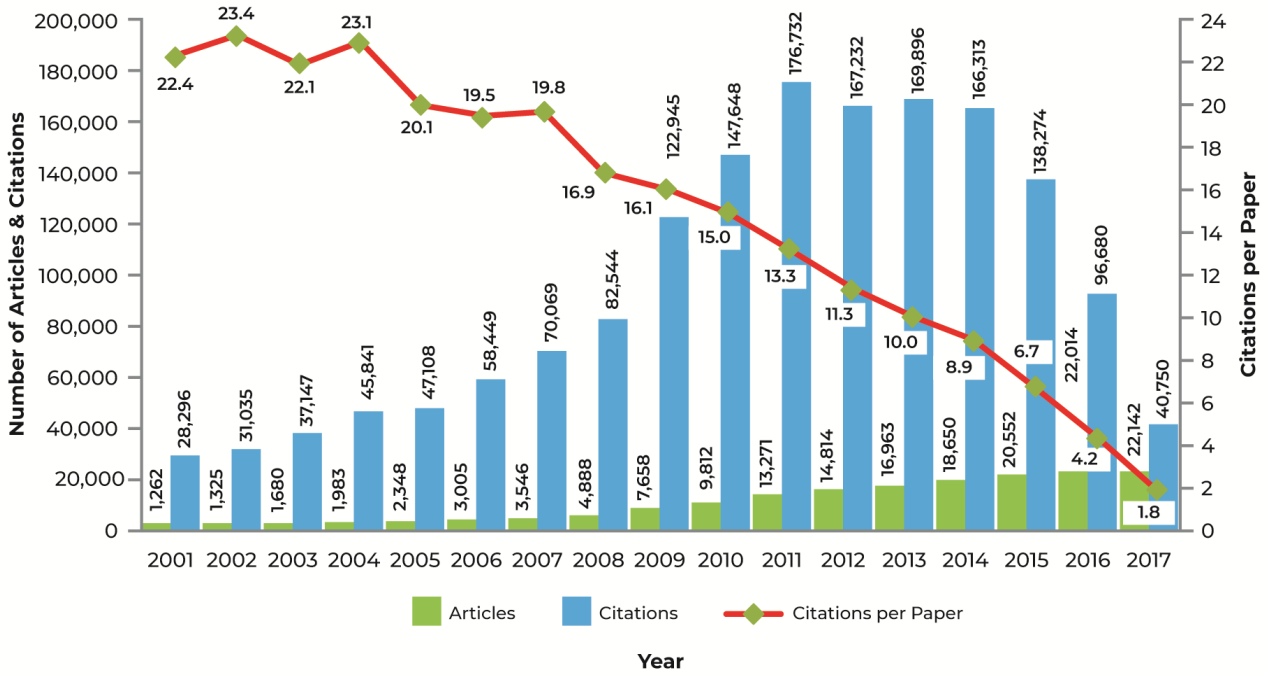


Figure 3.5. Citations Growth (2001 – 2017): Scopus
Source: National Bibliometric Study 2001-2017, MASTIC (2020), as of 24th August 2018

3.1.2.2. International Comparison

GERD (% of GDP)

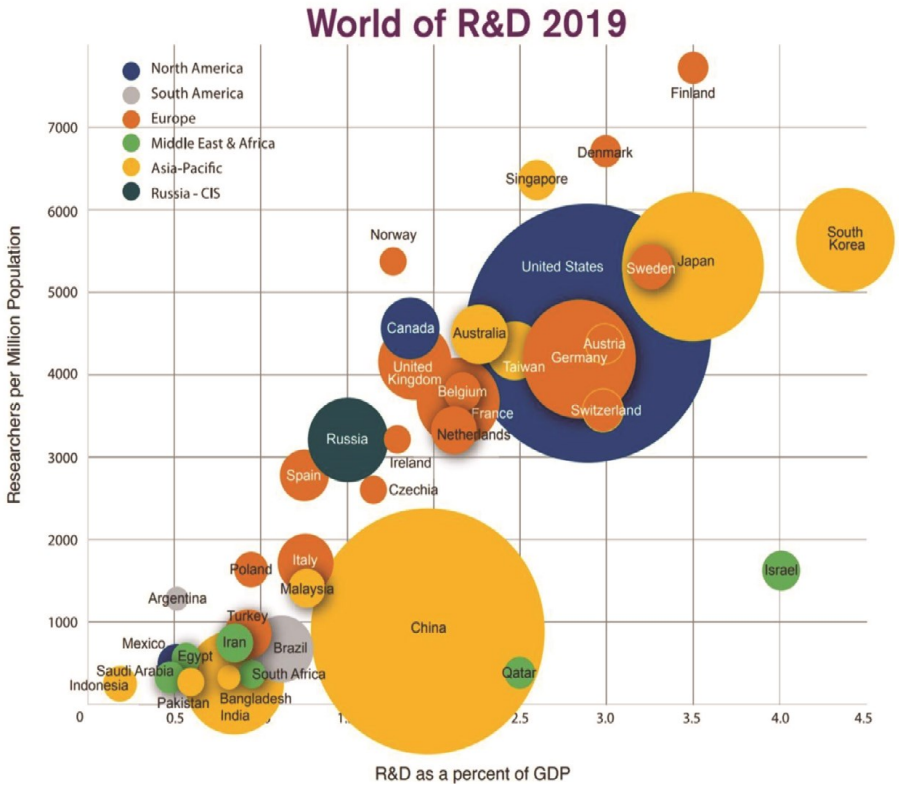


Figure 3.6. GERD (% of GDP)
Source: 2020 Global R&D Funding Forecast, R&D Magazine (2020)

R&D Spending (USD Billion)

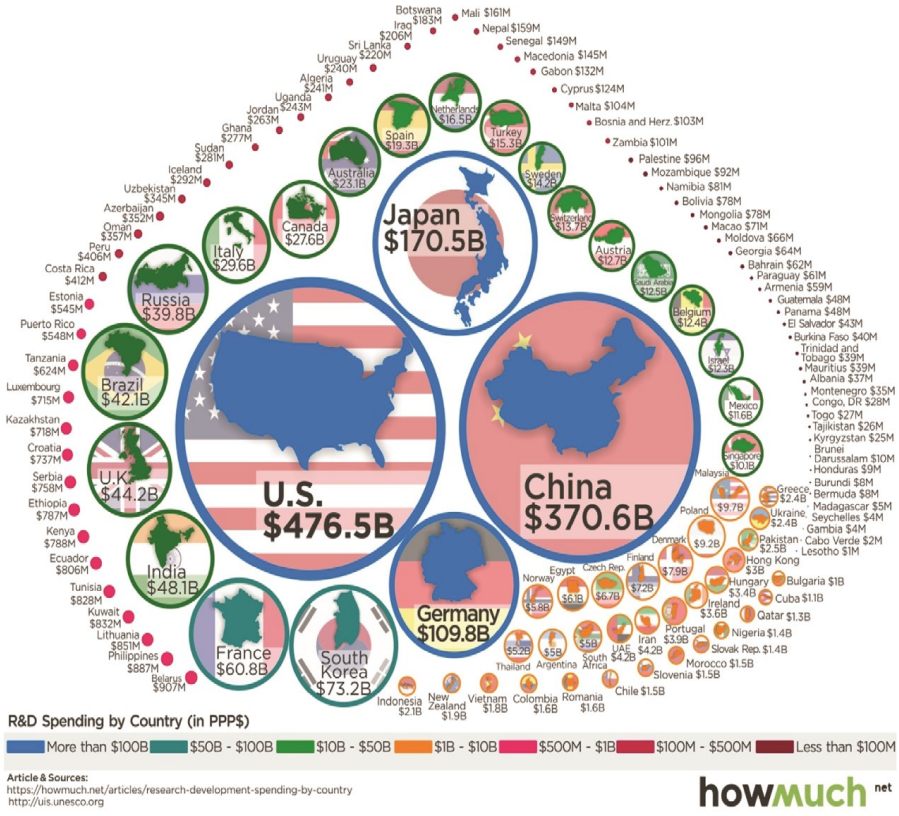


Figure 3.7. R&D Spending (USD Billion) at absolute level
Source: Desjardins, J. Visualizing How Much Countries Spend on R&D (2018)

Leadership in Innovation

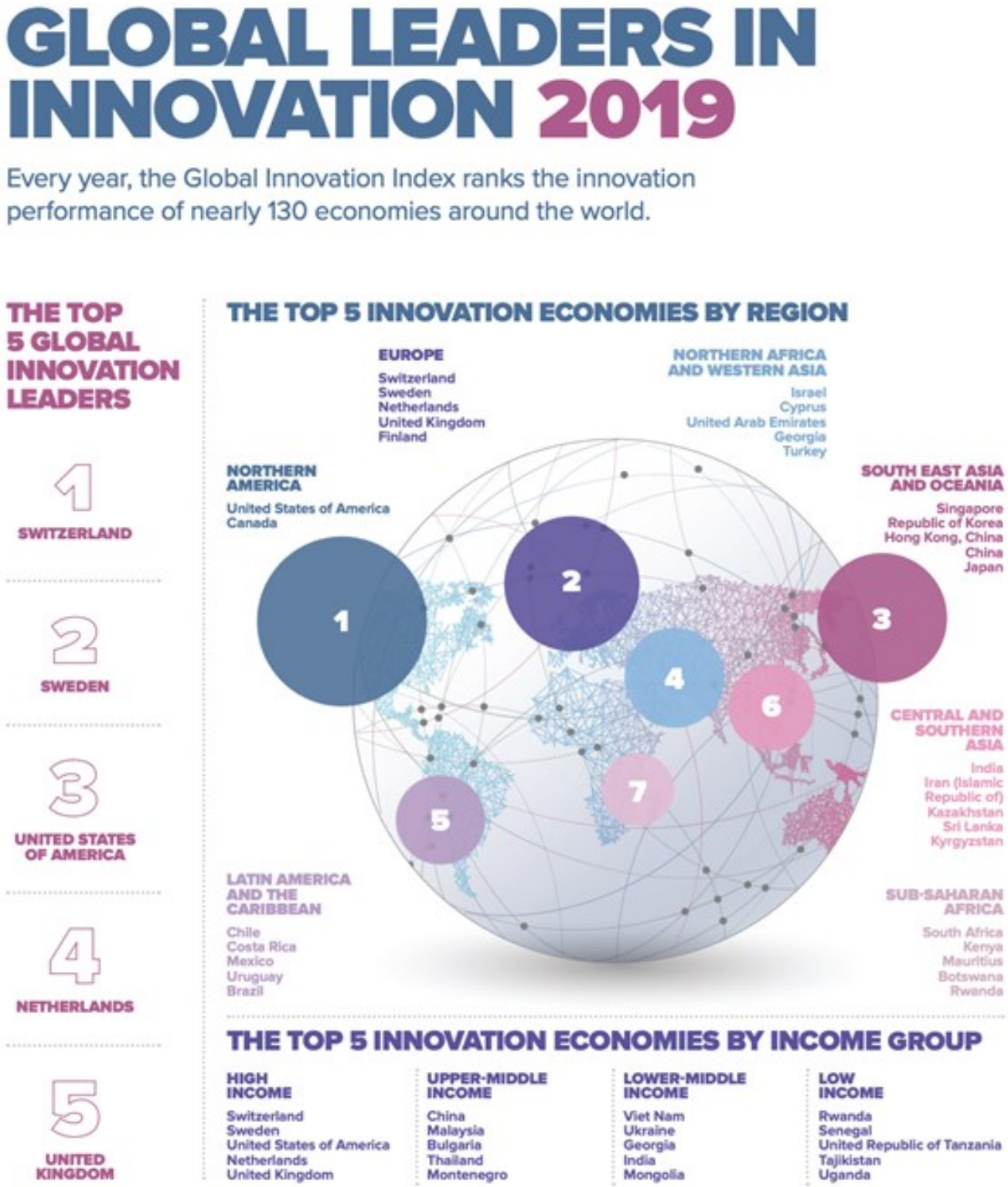
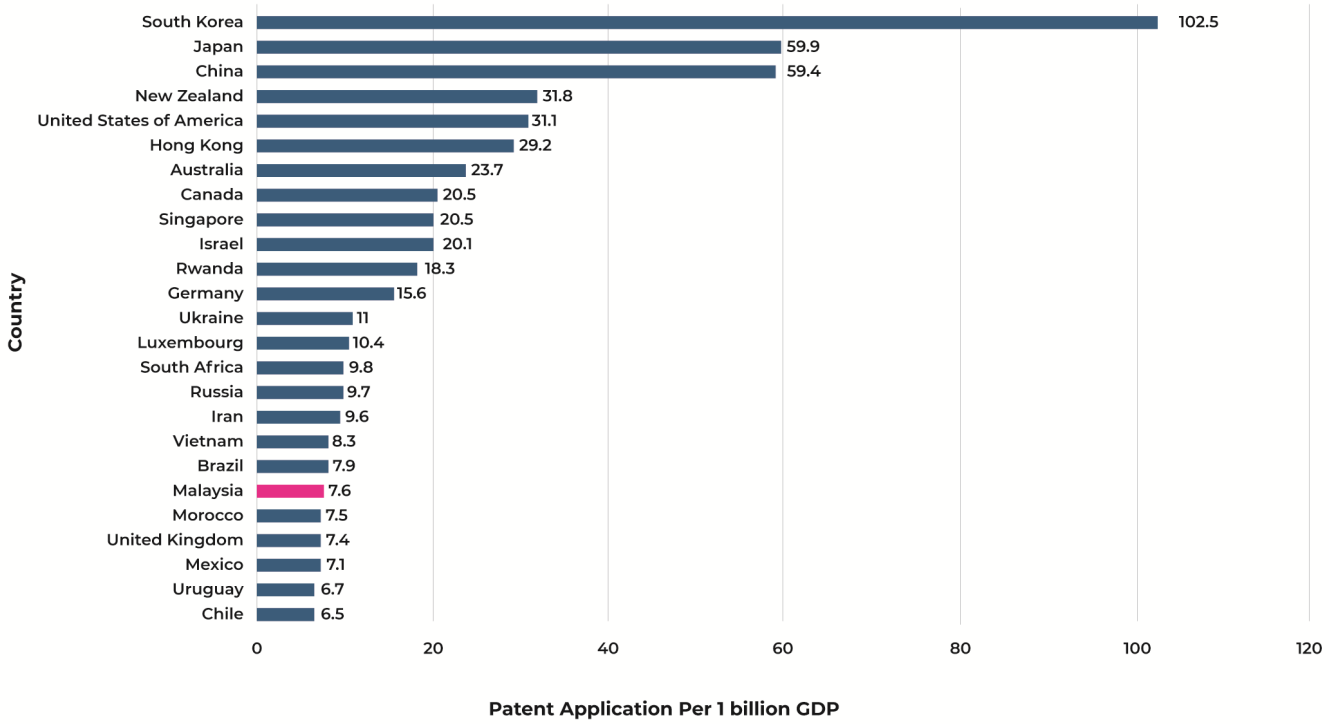


Figure 3.8: Global Leaders in Innovation
Source: World Intellectual Property Innovation (2019)

Patents



Note: Patent application data from the WIPO database, GDP data from World Development Indicators database

Figure 3.9. Patent Applications for Top 25 Countries
Source: National Survey of Innovation 2018, MASTIC (2020)

Publications

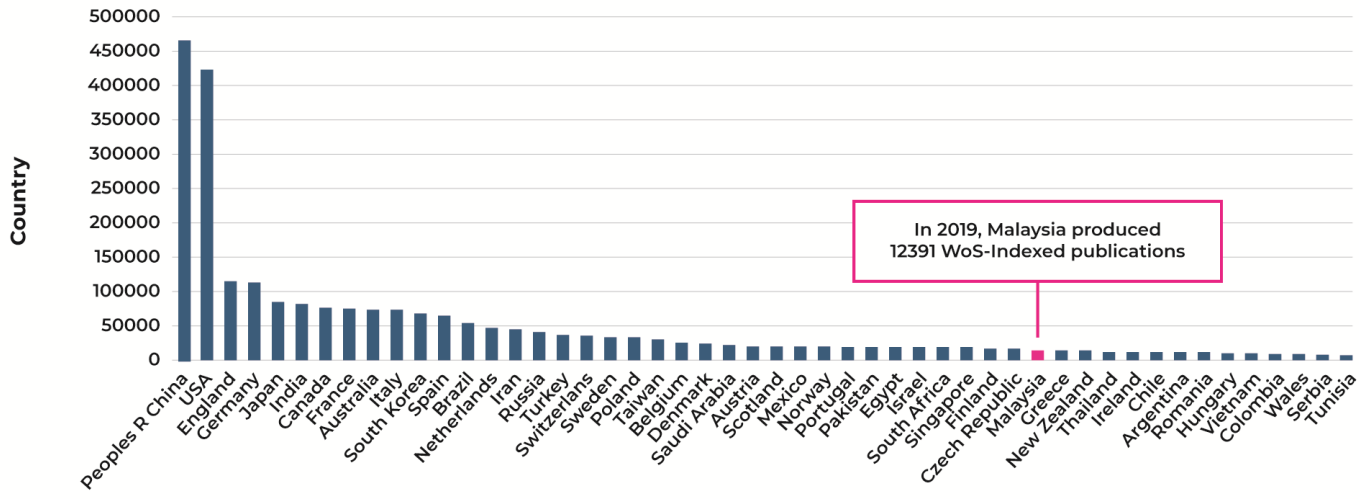


Figure 3.10. Global overview of total of publications in 2019
Source: National Survey of Innovation 2018, MASTIC (2020)

3.2. National Policy for Science, Technology and Innovation, NPSTI (2013-2020)

Moving ahead requires that the country needs to adopt bold and creative solutions especially under the new normal. The business as usual approach will not work in an era fraught with uncertainties and intense global competition. The general thrust of the NPSTI is to further strengthen the country's basic foundations, namely, its competency in generating and deploying knowledge through STI, strengthening its STI human capital, elevating the innovative capacity of industry and strengthening its STI governance as well as developing a STI culture through engagement with our community.

Hence, a fourth STI-+ policy i.e. the National Policy for Science, Technology and Innovation, NPSTI (2021-2030) has been formulated. It represents the country's commitment towards revitalising the nation's STI ecosystem under the new normal so that the nation can be more competitive, innovative and creative in advancing STI to achieve the nation's goal. The Policy will set out a new and focused approach to discover, utilise and optimise the full potential of STI to achieve our long-term economic, social and environmental goals. It has taken stock of the successes, challenges, lessons learnt and prospects of STI development and implementation plans both at national and international level. The Policy positions Malaysia to be a high-tech nation by 2030.

The 4th STI Policy is based on 6 thrusts as follows:

a) Responsive STI Governance (Tadbir Urus Yang Responsif);
b) Technology Development via R,D,C & I (Pembangunan Teknologi Melalui R,D, C & I);
c) Industry Based on Homegrown Technologies (Industri Berasaskan Teknologi Tempatan);
d) Adaptive STI Talents (Bakat STI yang adaptif);
e) Enculturation and Application of STI (Pembudayaan dan Pengaplikasian STI); and
f) Global Excellence (Keunggulan Global)

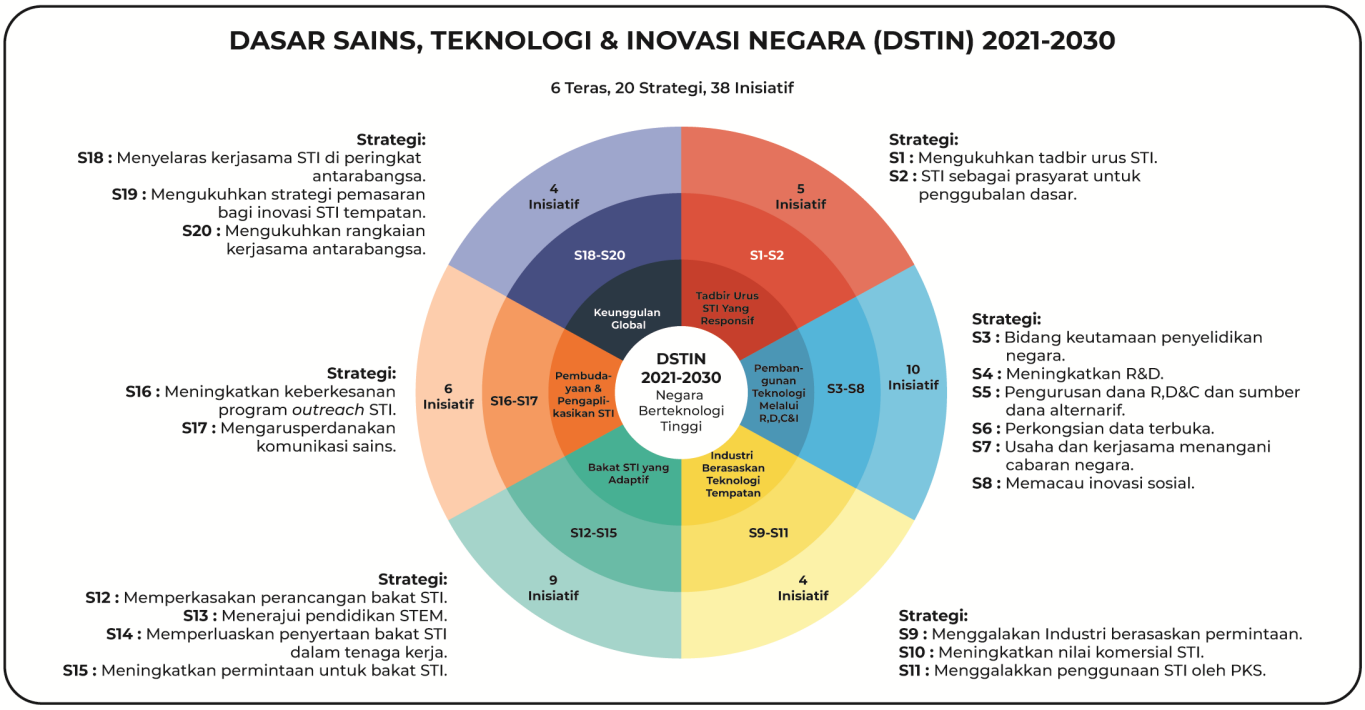


Figure 3.11. DSTIN (2021-2030) Framework, Launched on 8 December 2020
Source: Dasar Sains, Teknologi dan Inovasi (DSTIN) 2021-2030, MOSTI

Further, these six thrusts are supported by 20 strategies which in turn have identified 38 initiatives as depicted in the National STI (DSTIN) Framework (**Figure 3.11**). There are 10 main targets of DSTIN as shown in **Figure 3.12**, which will be supported by 10 Technology-Leap Programmes.

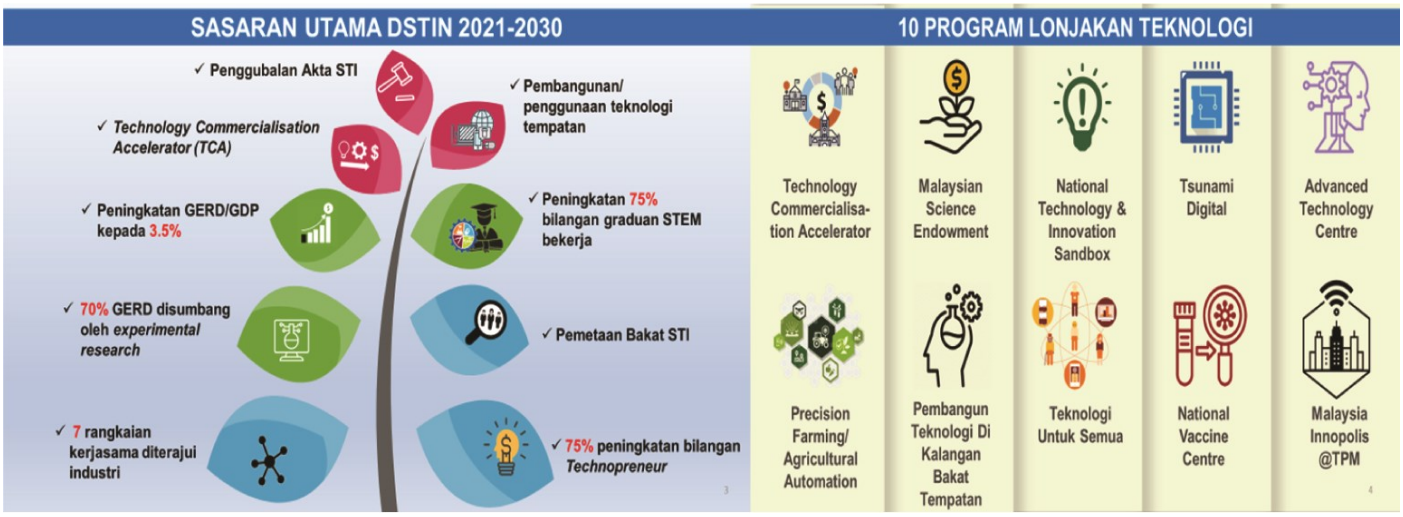


Figure 3.12. Main Targets and Technology-Leap Programmes
Source: Dasar Sains, Teknologi dan Inovasi (DSTIN) 2021-2030, MOSTI

Under the new STI Policy, a couple of strategies that are directly related to Open Science are:

a) S6 (Open Data Sharing);
b) S16 (Increased Effectiveness of STI Outreach Programme); and
c) S17 (Mainstreaming Science Communication)

Other related strategies which will indirectly assist in promoting Open Science are:

a) S1 (Enhanced STI Governance);
b) S7 (Ventures and Collaborations in Overcoming National Challenges);
c) S8 (Driving Social Innovation);
d) S14 (Expands STI Talents Participation In Workforce);
e) S18 (Coordinate International STI Collaborations); and
f) S20 (Strengthened Network of International Collaborations)

3.3. Open Science in Supporting National STI Agenda

In recent years, the Open Science movement is gaining traction in many countries with international bodies such as the International Science Council (ISC), Organisation for Economic Co-operation and Development (OECD), United Nations Educational, Scientific and Cultural Organization (UNESCO) and the World Bank are championing and advocating Open Science globally. The international principle of making research data Findable, Accessible, Interoperable and Reusable (FAIR) will not only democratise knowledge, it reinforces open scientific inquiry and integrity, enables better research management and promotes open innovation, citizen science as well as data intensive research. Integrating the diverse data streams and huge datasets across multiple disciplines offers unprecedented insights and solutions towards local, regional and global complex challenges. The rationale for engaging Open Science is well described in **Figure 3.13.**

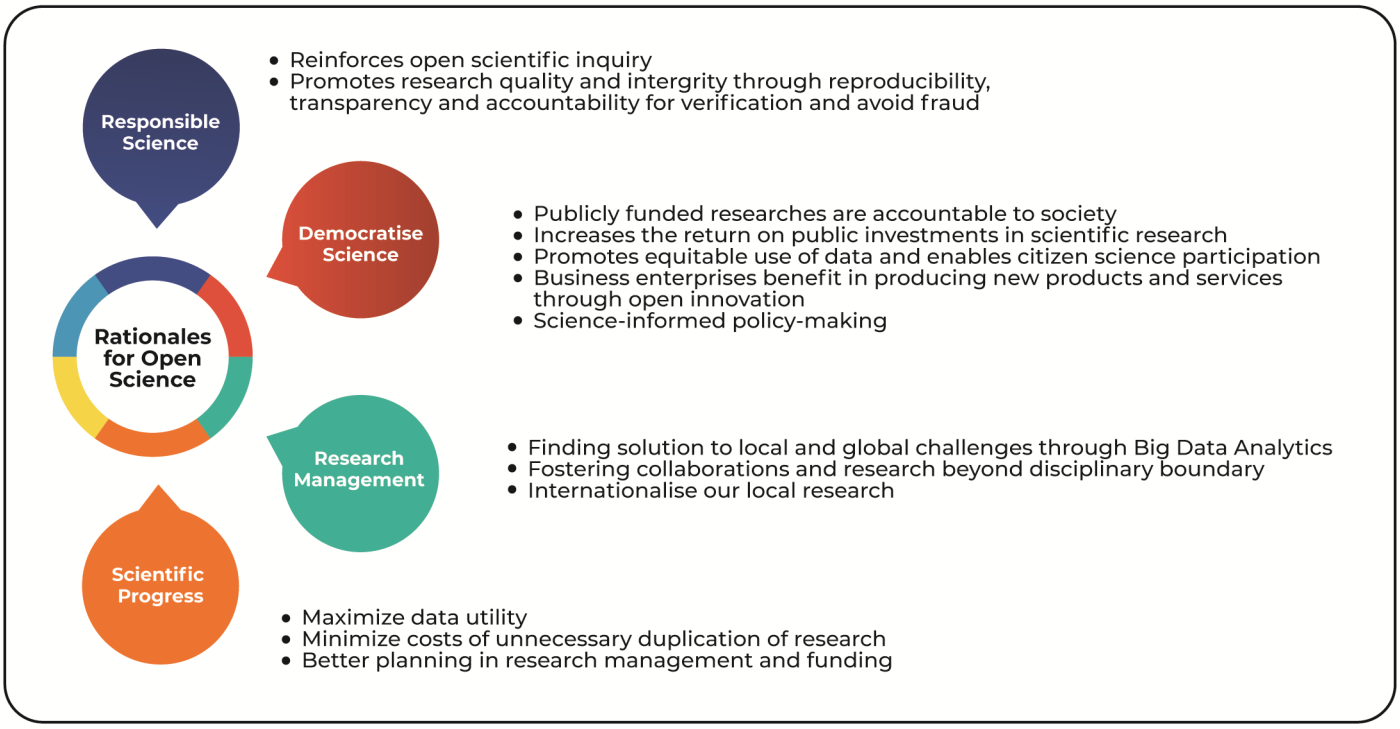


Figure 3.13: Rationales of Open Science Source: Analysed by ASM (2020)

3.3.1. Key Actors in Open Science

A country's economic structure lies on four pillars/helices: Academia, Firms, Government and Civil Society, and economic growth is generated by the clustering and concentration of talented and productive people. This is the premise of the Quadruple Helix Innovation Theory (QHIT). Based on a Quadruple Helix Model the main actors of STI are generally the same actors in Open Science. These actors in the Quadruple Helix model context are:

a. Researchers – They are at the forefront in promoting STI and Open Science too.

b. Government ministries and agencies which are involved primarily in the development of national strategies as well as funding in STI including open science.

c. Research funding agencies are key actors in the promotion of Open Science efforts, as they are responsible for defining the mechanisms and requirements to benefit from grants and funding for research.

d. Public Institution of Higher Learnings (IHLs) such as the public universities having some degree of autonomy and have their own policies to support Open Science and implementing the policies of funding agencies apart from the role in training of researchers in developing skills related to open science such as "from basic skills related to the use of online repositories, to the ones needed to implement data cleaning, curation and management."

e. Public Research Institutions (PRIs), just like IHLs, also assume a similar role in promoting and supporting Open Science.

f. Libraries, repositories and data centres are "key actors for and fundamental enablers of open science.". For example, the main role of Libraries will be "in the preservation, curation, publication and dissemination of digital scientific materials, in the form of publications, data and other research-related content.". Libraries together with repositories will form "the physical infrastructure that allows scientists to share, use and reuse the outcome of their work, and they have been essential in the creation of the open science movement.".

g. Private non-profit organisations including foundations are playing a significant supporting role "in developing, raising awareness and encouraging an open science culture.". Not only do they fund open research, they also develop and facilitate the creation of networks of stakeholders nationwide and worldwide too.

h. Private scientific publishers "offer a broad range of Open Access publishing (for example via the gold route or publishing in hybrid journals) and related services such as the maintenance of digital repositories and data sets or other scientific material, or the development of text and data mining (TDM) tools".

i. Businesses that form part of the demand side in Open Access publications and data will allow them to apply and develop new products and services. This is done through the public-private partnerships with universities through endowment contribution and funding of projects.

Above all, there is a need for a supra agency to super head the Open Science initiatives. The formation of such a body is crucial in ensuring the promotion and development of Open Science. Such a body may reside in the government.

3.3.2. Malaysia Open Science Platform

Recognising the importance of harnessing the potential impact of Open Science, and to address the global challenge of managing the growing data deluge in a collaborative manner, it became clear that implementing Open Science in Malaysia is the way forward to thrive the National Science, Technology & Innovation ecosystem. This motivation and optimism has led the Ministry of Science, Technology and Innovation (MOSTI) to initiate the 2-year Malaysia Open Science Platform (MOSP) pilot project. MOSP is managed and entrusted to the Academy of Sciences Malaysia (ASM), spearheaded by the Malaysia Open Science Alliance, to drive the National STI ecosystem towards wealth creation and societal wellbeing.

Objectives, key areas and targets

MOSP is a strategic transformative initiative to strengthen Malaysia's STI Collaborative Ecosystem towards achieving the Shared Prosperity Vision 2030 and addressing the United Nations Sustainable Development Goals. The goal of MOSP is to gather and consolidate Malaysia's research data, which are valuable national assets in a platform that would enable accessibility and sharing of these research data in accordance with the FAIR principle. The pilot project will involve the five Research Universities in Malaysia, namely, Universiti Malaya, Universiti Sains Malaysia, Universiti Kebangsaan Malaysia, Universiti Putra Malaysia and Universiti Teknologi Malaysia. The objective of the pilot project is to look into the initial three main areas:

a. Guidelines
b. Capacity Building and Awareness, and
c. Infrastructure

Table 3.1. Key areas and its respective targets to be achieved by end of 2021

Key areas	Targets
Guidelines	• To study the landscape and awareness of Open Science • To draft National Guidelines for Raw Research Data Sharing by end of 2020
Capacity Building and Awareness	• To develop a specialised training module and provide training for 200 data stewards by 2021 • To reach approximately 500,000 people by 2021
Infrastructure	• To produce one technical specification document by 2020 • To establish and deploy the MOSP Pilot Platform involving 5 Re search Universities by 2021

Structure

Following is a schematic presentation of the structure of Malaysia Open Science Platform (MOSP) initiative:

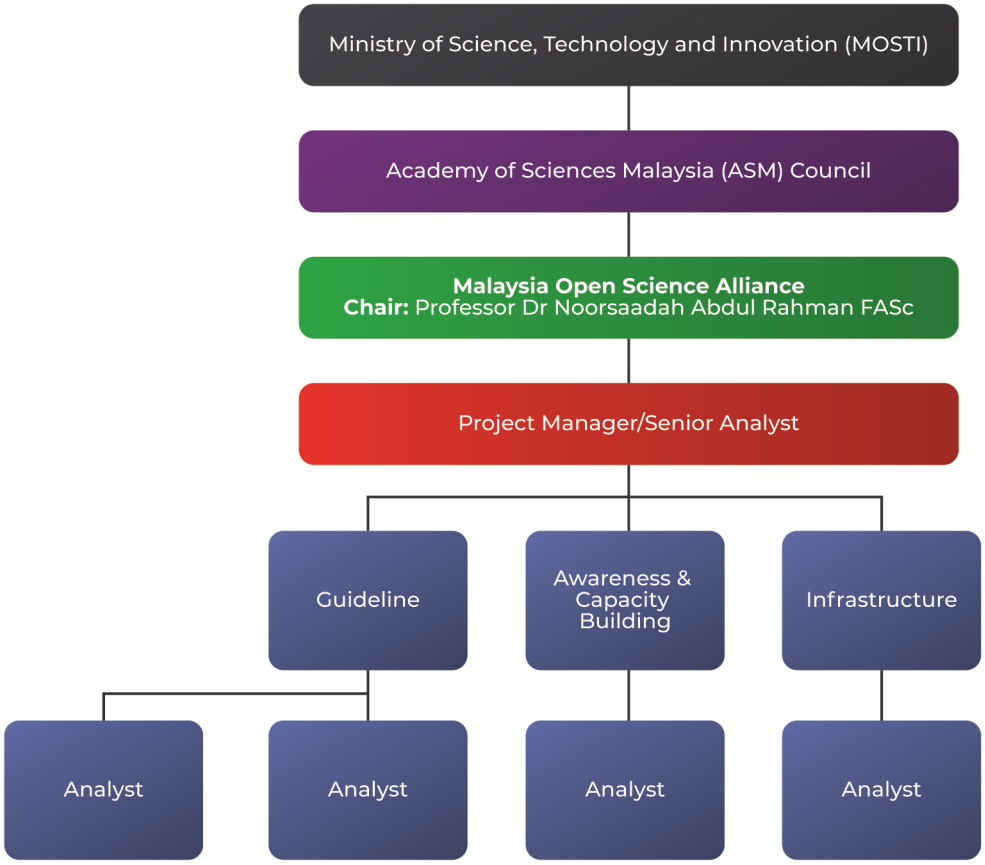


Figure 3.14. MOSP Structure Source: ASM (2020)

The 2-year pilot initiative MOSP is implemented under the purview of the Ministry of Science, Technology and Innovation (MOSTI). The project is managed by the Academy of Sciences Malaysia through the Malaysia Open Science Alliance. Three working groups are established to look into three key areas: Guidelines, Capacity Building and Awareness and Infrastructure.

3.3.3. Some Highlights of Open Science and MOSP activities

i. Training Module Evaluation Workshop, July 24th & August 25th, 2020

The Academy Sciences of Malaysia, all members of Malaysia Open Science Alliance Malaysia Working Group on Capacity Building and Awareness and the first-cohort Certified Data Stewards convened a one-day workshop to discuss, consolidate and finalise contents for training module development for Open Science skills and data stewardship. The Workshop was held twice, the first one in July and the second one in August 2020.

ii. Success Stories Sharing Session on Data Sharing Platform, July 15th, 2020

The Academy of Sciences Malaysia, all members of the Malaysia Open Science Alliance Malaysia Working Group on Infrastructure and selected industry and government agency representatives met in a virtual meeting to exchange views about data sharing platforms. Representatives from Microsoft Malaysia, Telekom Malaysia R&D, SIFULAN, Amazon Web Services, MIMOS and MAMPU shared their success stories with ASM and the Working Group members.

iii. First Stakeholder Engagement Workshop for Open Science, June 4th, 2020

a. The Academy of Sciences Malaysia and the Malaysia Open Science Alliance Malaysia Working Group on Guidelines conducted its inaugural Stakeholder Engagement Workshop for Open Science, which was held virtually. The objective of the workshop was to elicit and consolidate inputs from participants for the formulation of National Guidelines for Raw Research Data Sharing in Malaysia.

Over 60 stakeholders from various backgrounds, including researchers, top management universities, librarians, research managers, legal units and industry shared their thoughts and inputs to strategise methods and approaches for developing a more effective implementation for the Malaysia Open Science Platform initiative to result in an increased program impact.

b. The Workshop was officiated by the Chair of Malaysia Open Science Alliance, Professor Dr Noorsaadah Abd Rahman FASc. To encourage a more effective discussion, the workshop participants were grouped into three parallel sessions for two rounds. The Workshop was concluded by discussion in plenary, which summed recommendations and key points from the parallel sessions. The list of discussion topics is as follow:

- Formulation of National Guidelines for Raw Research Data Sharing in Malaysia
- Building responsibilities, support and policy for the development of appropriate skills for data stewards and data curators
- Towards building a trusted data-sharing platform
- Ownership and Reuse in Data Sharing Practices
- Why do Open Science & Incentivising Data Sharing Matter?
- Imagining the look of MOSP architecture and approaches towards a sustainable framework

iv. Participation in an Online Data Stewardship Training, June 2020 – September 2020

Twelve (12) candidates, two from each Research University and MASTIC (MOSTI) participated in an online training Certified Data Stewardship program to be certified as data stewards. They are "master trainers" who will be the certified instructors for the Training of Trainers (ToT) program for Data Stewardship on Open Science. The ToT program aims to train 200 data stewards in Malaysia by the end of 2021.

v. Policy writing workshop, May 20th to 21st 2020

Academy of Sciences Malaysia and the Malaysia Open Science Alliance Working Group on Guidelines participated in a virtual half-a-day workshop on two consecutive days to consolidate and finalise the first findings for landscape study of Open Science in Malaysia. The landscape study was carried out to, i) map existing Open Science activities and players in Malaysia; ii) gauge the level of readiness of local institutions to adopt Open Science; iii) identify factors enabling and hindering Open Science adoption and iv) undertake benchmarking exercises to identify global indicators to measure Open Science readiness. Inputs from this Workshop were presented in the First Stakeholder Engagement Workshop for Open Science, which was held on June 4th, 2020.

vi. Sharing session on the data repository platform architecture at the institutional level, May 15th and 22nd 2020

Academy of Sciences Malaysia and the Malaysia Open Science Alliance Working Group on Infrastructure held an insightful sharing session on the institutional data repository platform architecture at the five Research Universities in Malaysia. The objective of the sharing session was to provide the Academy of Sciences Malaysia and members of the Working Group to share, learn and understand in detail the existing platform architecture of institutional repositories at the five Research Universities in Malaysia.

vii. YB Khairy Jamaluddin shared his view on COVID-19 and Open Science during the UNESCO Virtual Ministerial Dialogue, March 30th, 2020

The Science, Technology and Innovation (MOSTI) Minister YB Khairy Jamaluddin shared his view in the UNESCO Virtual Ministerial Dialogue on COVID-19 and Open Science. During the Dialogue, YB Khairy Jamaluddin described current initiatives taken by Malaysia to control the pandemic and emphasised the urgent need for creating an international repository not only for medical purposes but also for the development of technological innovation. Data sharing is important to develop applications relevant to the control of the COVID-19 crisis, such as contact tracing application and diagnostic kit development.

viii. Dialogue on Open Science at Universiti Malaya, February 14th, 2020

Academy of Sciences Malaysia has organised the Dialogue on Open Science in Universiti Malaya on 14th February 2020. The event was attended by about 60 participants from various research institutions around Malaysia. The dialogue session was led by Professor Dr Barend Mons who spoke on the future of data stewards and FAIR Principles. The session discussed the challenges and issues faced in research and data sharing in their countries and ways forward to tackle the related problems mainly focusing on infrastructure and capacity building. It has provided an avenue for sharing best practices, experiences, knowledge as well as providing an opportunity for Malaysia to learn best practices on pursuing the Open Science agenda at the national level.

ix. Open Science Forum for Asia and the Pacific Region, February 13th, 2020

Academy of Sciences Malaysia has organised the Open Science Forum for Asia and the Pacific Region in conjunction with the 15th APEC's Policy Partnership for Science, Technology and Innovation (PPSTI) meeting on 13th February 2020 at Zenith Hotel Putrajaya. This forum aimed to develop common Open Science understanding between the Asia Pacific countries through engagement and knowledge sharing among the professionals. This is also an opportunity for Malaysia to learn the best practices on pursuing the Malaysia Open Science Platform (MOSP) agenda to accelerate development in science, technology and innovation.

x. Launching of MOSP and i-Connect initiatives, November 7th, 2019

The official launching of the Malaysia Open Science Platform (MOSP) Initiative took place at Putrajaya International Convention Centre (PICC) on 7th November 2019. The event was a great success and was attended by more than 200 participants from all around Malaysia. The MOSP initiative was launched by YB Yeo Bee Yin, the former Minister of Energy, Science, Technology, Environment and Climate Change (MESTECC). The event proceeded with a forum held in the afternoon. Entitled 'Exploratory Discourse: Charting the Way Forward on Open Science', the forum was moderated by Academician Professor Emerita Datuk Dr Mazlan Othman FASc, the Director of the International Science Council (ISC) Regional Office for Asia and the Pacific.

3.3.4. MOSP's Future Activities

Following the completion of the 2-year pilot project, MOSP will connect with existing infrastructure and data sharing platforms such as those located outside the five Research Universities, Open Government Data by Malaysian Administrative Modernisation and Management Planning Unit (MAMPU) and Raw Data for Research and Sciences (RADARS) under MOSTI. In the longer term, MOSP will connect with platforms in other countries and the Asia Pacific region and be equipped with big data analytics and real-time analysis features.

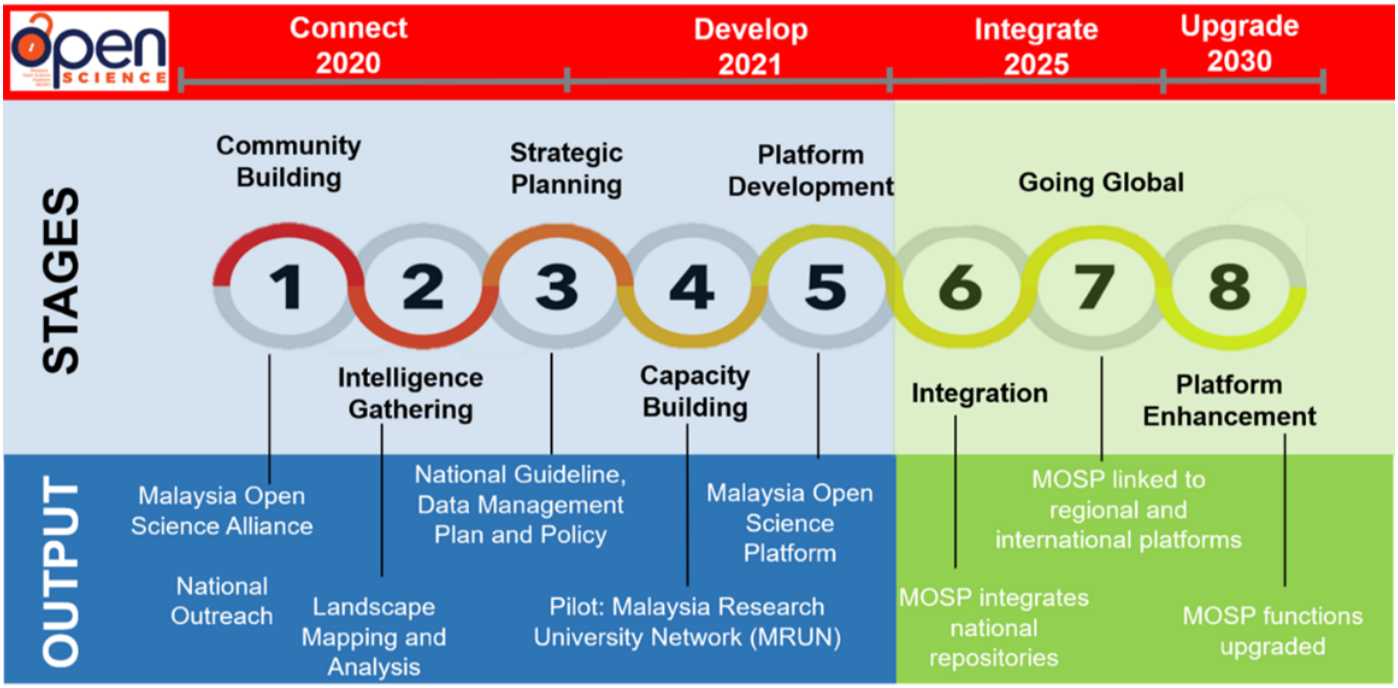


Figure 3.15. MOSP Activity and Timeline
Source: ASM (2020)

3.4. Acts, policies and guidelines related to data sharing in Malaysia

Data sharing is generally understood as a practice of making data available for use by individuals and entities. While it is important to make data as openly as possible to a wide range of audience, there are certain types of data that warrant several limitations and protective measures due to its varying nature of sensitivity and confidentiality. Such non-public data may require different access controls of data and data security measures such that necessary precautionary steps must be taken to minimise impacts to the Institution should that data be disclosed, altered or destroyed without authorisation. Sensitivity and confidentiality of data dictates the level of accessibility and reusability of such data, as stipulated in common laws, existing acts, policies and guidelines in Malaysia (**Appendix 3.1**).

Factors that influence limitations on accessibility and reusability include but are not limited to:

a. Privacy and data protection
b. Intellectual Property Rights;
c. Data in relation to endangered or threatened species, sovereign genetic resources, traditional knowledge, valuable artefacts;
d. Data containing information pertaining to national security risks,
e. Data produced during preclinical and clinical trials

Table 3.2. Impacts from legal issues in the various phases of the data cycle
Source: ASM (2020)

Phases of data cycle	Implications to data sharing practices
Data storage	• Personal data must be anonymised. • Consents or clearance is necessary for sensitive, restricted or personal data. • Data with commercial value shall not be deposited until a patent is filed. Intention to file the patent must be expressed when designing a Research Data Management Plan. • Data containing information pertaining to national security risks must be stored in a secure enclave with closed access. • Exemptions may be applicable for data arising from preclinical and clinical research.
Data processing & analysis	• Creative common licence may be applied to non-sensitive data. • Specific licensing terms should be in place, depending on the types (i.e. sensitivity) and use (involved profits or not) of datasets.
Data access	• All metadata, non-sensitive, non-restricted and non-personal data are accessible to the public. • Other types of data may have different levels of accessibilities. For example, some datasets may only be accessible between organisations or researchers, or need approval from relevant authorities.
Data re-use	• Reusing and creation of new data is governed by licensing terms associated with the original data. • Original data contributors must be acknowledged and credited. • Access to certain data may impose additional fees, as stipulated in the Fees Act 1951.

3.5. Open Access Options for Scientific Publications

"By 'open access' to the literature, we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself."

- Budapest Open Access Initiative -

Open Access (OA) refers to the practice of providing online access to scientific information that is free of charge to the end-user and is reusable with permissions. In the UK, The Research Council has adopted a 'Golden Road' policy, mandating researchers to publish results under a Creative Commons Attribution Licence 4.0. Choice of OA licences are dependent on publishers as they have the responsibility and right to choose licences that fit the type of data that they host and the needs of their clients. A few types of Open Access routes available for researchers include:

a. The "gold" Open Access route: Authors pay the publisher for Open Access and for which a licence that usually authorises the reusing of the publications, is applied. The publication costs are known as "Article Processing Charges (APCs)".
b. The "green" Open Access route: The original work is still licenced by the publisher but after an embargo period, it is made available under an Open Access scheme.
c. The "diamond" Open Access route: Publications via diamond route usually do not charge author-facing publication fees. Diamond OA journals are usually funded via library subsidy models, institutions or societies. Both green and diamond routes are offered by hybrid journals.
d. The "hybrid" Open Access route: Sometimes called as paid open access, this route refers to subscription journals with Open Access to each individual article in which a fee is usually paid to the journal or publisher by the author, the researcher funder or the research organisation.
e. The "bronze" Open Access route: Journal articles published under this route are freely accessible but with no clear or lack of identifiable licensing agreements in place, which makes it difficult for data to be reused.

Embargo period refers to a time delay which is applied to an article before it can be published or made publicly available. Embargo period is journal specific. During an embargo period, access to academic journals are not allowed to non-paid subscribers. However, after an embargo period has expired, authors can publicly post their published manuscripts on nominated repositories.

3.6. Readiness of funding body policy for Open Science

The role of funding bodies in promoting Open Science has been a major topic of discussion in other regions and countries such as Europe and the US. The view is concurred by researchers in Malaysia that it is important for research funders to make Open Science and data sharing practices a non-binding requirement to award grant funding to promote Open Science. In reference to this, the US National Science Foundation model could be a reference for public research funders in Malaysia to learn from.

In Europe, under the European Open Science Infrastructure (abbreviated as OpenAIRE) initiative, a model policy on Open Science for Research Funding Organizations was developed to assist research funders to develop policies for Open Science or Open Access. In addition to the model plan, an Open Science Policy Checklist for Research Funding Organizations was also created to assess Open Science readiness based on several indicators. The indicators were used to evaluate the readiness of the Ministry of Higher Education

[8] The Public Access Plan, Today's Data, Tomorrow's Discoveries, is a testament to the commitment of the US National Science Foundation to expand public access to the results of its funded research. The document can be found in this link: https://www.nsf.gov/pubs/2015/nsf15052/nsf15052.pdf

(MOHE) and the Ministry of Science, Technology and Innovation (MOSTI), who are among the key research funding organisations in Malaysia, to include Open Science and Open Access elements in their grant evaluation process. The analysis is tabulated in **Table 3.3**.

Table 3.3. Preliminary checklist of selected local research grant funders
Source: Analysed by ASM (2021)

Open Science Readiness Indicators	Ministry of Higher Education (MOHE)	Ministry of Science, Technology and Innovation (MOSTI)
1. Policy on Open Science or Open Access (e.g. inclusion of Research Data Management Plan, publishing in Open Access Journals)	The Guideline requires approval from relevant bodies/authorities (e.g. Research Ethics Committee for Human/Animals, if relevant) must be submitted together with the first report. Researchers are required to submit a research data sharing plan including raw data and unpublished data. Researchers are required to publish their research in an indexed journal with no restriction to an Open Access or subscribed journal."	The Guideline requires approval from relevant bodies/authorities related to the nature of projects (for example, Research Ethics Committee for Human/ Animals, Institutional Biosafety Committee Approval, if relevant) must be submitted. Researchers must comply with the Malaysian Code of Responsible Conduct in Research.
2. Publications and sharing (e.g. mandatory deposit, locus of deposit, time of deposit, provision of open access and embargo periods, licences and copyright)	The Guideline does not mandate open access to publications. However, researchers are required to publish in indexed journals. Open Access for the publications depends on the researchers and publishers.	The Guideline does not mandate open access to publications. However, researchers are encouraged to publish the results of their projects in local and renowned international publications only after all measures have been taken to protect any IPR generated from these projects.
3. Research data – specific provisions on data sharing, possible opts-outs, data archiving, long term preservation etc.	Researchers are required to submit a detailed research data sharing plan. This includes the acquisition, storage and sharing of data. All research outputs are also required to be submitted in MyGRANTS.	The Guideline does not mandate open access to research data.
4. Infrastructure –recommends the use of repositories that meet FAIR principles and provide guidance to grantees	MyGRANTS is a platform where any data pertaining to the researcher and research projects can be accessed by the public. MyGRANTS is currently the main repository used for research project management funded by MOHE. However, all raw data/analysed data is also kept by the research institutes or universities according to their repositories' requirements.	The Guideline does not provide guidance regarding the selection of repositories.

[10] For this landscape report, the assessed research grants under MOSTI are (1) R&D Fund, (2) International Collaboration Fund, and (3) Malaysia Social Innovation Fund

5. Inclusion of Open Science as a criterion in proposal or project evaluation	The “Data Sharing Plan” criteria was introduced in 2019 where researchers needed to submit a data sharing plan together with the proposal of TRGS and LRGS.	The Guideline does not include any evaluation criterion linked to “openness”.
6. Open Access Publication Fees	The Guideline provides funding for publication costs, including Open Access journals if it adheres to the publication requirements.	The Guideline does not specify allocation/vote for funding for open access publication costs.
7. Monitoring and Compliance	The Guideline has set up a mechanism for monitoring policy compliance by its grantees. This includes project and financial performance reports, site visits and audits. The Guideline also specifies sanctions for no compliance such as control actions and penalties.	The Guideline has set up a mechanism for monitoring policy compliance by its grantees, including sanctions for no compliance.
8. Revision and updates (contains a specific time plan for its review (and possible update)	The Guidelines are updated annually during the proposal phase for researchers’ and institutions’ reference.	There is no provision in the policy for its review/ update.

CHAPTER 4 INTERNATIONAL PRACTICES FOR OPEN SCIENCE

4.1. Australian National Data Service (ANDS)

ANDS was established in 2008 and was led by Monash University with collaboration from the Australian National University and the Commonwealth Scientific and Industrial Research Organisation (CSIRO). ANDS also works closely with Research Data Services (RDS) and National e-Research Collaboration Tools and Resources project (Nectar). ANDS's ultimate aim is to make Australia's research data assets more valuable for researchers, research institutions and the nation. ANDS provides the Research Data Australia discovery portal where users can find, access and reuse data for research from Australian research organisations, government agencies and cultural institutions. ANDS does not store data by itself but it displays description of, and links to, the data held by their data publishing partners or contributors.

4.1.1. Available Policies and Licensing

a. Research data rights management

Research data rights management provides information on the copyright basics and guidance in decision making about data licensing and data reuse to the data owners, users and the suppliers.

b. Institutional policies and procedures

Institutional policies and procedures cover the example of data management policies, approaches to data management policies, etc.

c. Research data policy and the Australian Code for the Responsible Conduct of Research (The Code)

All Australian universities have signed up to the Australian Code for the Responsible Conduct of Research (The Code). The Code delineates the principles and responsibilities that both researchers and institutions are expected to follow when conducting research. Researchers will upload the principles of responsible research conduct in all aspects of their research. On the other hand, institutions have an obligation to encourage and support responsible research conduct. They are accountable to funding organisations and the Australian community for how research is conducted.

4.1.2. Number of repositories

ANDS does not have a centralised repository that stores data but it provides a data discovery service, named as the Research Data Australia (RDA) registry, which stores and manages the metadata records that are provided to ANDS and were discovered via Research Data Australia. All repositories are located at research institutions and data publishing partners.

4.1.3. Skill capacity

ANDS works coherently with researchers and institutions to enhance the data skills and capacity of Australia's research system by providing resources and training like webinars, workshops and written guidelines. Below are the skill development programs conducted by ANDS:

a. Domain specific training
b. For data trainers
c. FAIR data
d. Using data in classes and tutorials
e. 10 and 23 (research data) Things
f. Technical data skills
g. Training resources
h. Multi-day data and research skills training

4.1.4. Infrastructure capacity

The national infrastructure of ANDS supports standard definitions, reference values, identifiers for these common entities and contributors to use when describing research data collections. A checklist for assessing IT infrastructure capability for data management is also provided to evaluate current infrastructure supporting the management of institutional research data assets.

4.1.5. Completed programmes

Since its establishment, a number of projects have been executed through partnerships between ANDS and research institutions and other counterparts. The projects include:

a. Trusted Research Outputs (2016-2017)
b. Trusted Data Repositories (2016-2017)
c. High Value Collections (2016-2017)
d. High Value Collections meets NCRIS (2016-2017)
e. Collection Enhancement Partnerships (2016-2017)
f. eResearch Infrastructure Connectivity Projects (2015-2016)
g. Institutional Research Data Capabilities (2015-2017)
h. Australian Research Data Commons Application (2012-2014)

ANDS, Research Data Services (RDS) and the National eResearch Collaboration Tools and Resources (Nectar) have integrated to form the Australian Research Data Commons (ARDC) on 1st July 2018. The following are the major programs under ARDC:

a. Research Domain Program
b. Research Data Cloud (RDC)
c. Data Enhanced Virtual Laboratory (DEVL)
d. Research Data Platforms
e. Cloud and Storage Infrastructure
f. Data Publication Services
g. Sector-Wide Engagement

4.2. European Open Science Cloud (EOSC)

European Open Science Cloud is a European Commission initiative aiming to build a trusted virtual environment to store, share and reuse research data seamlessly for Europe's 1.7 million researchers and 70 million science and technology professionals. The initiative was initiated in 2015 and was officially launched in 2018.

EOSC is governed by three main bodies, as defined by the European Commission Staff Working Document Implementation Roadmap for the European Open Science Cloud:

a) The Executive Board
b) The Governance Board of EOSC
c) The Stakeholder Forum

The development of EOSC is supported by multiple projects including:

a) EOSC-PILLAR - Coordination and Harmonisation of National Initiatives, Infrastructures and Data services in Central and Western Europe (July 2019 - June 2022),
b) EOSC Nordic (September 2019 to August 2022),
c) NI4OS-Europe - National Initiatives for Open Science in Europe (September 2019 - August 2022)
d) EOSC-Synergy (September 2019 - February 2022)
e) ExPaNDS - EOSC Photon and Neutron Data Services (September 2019 - August 2022)
f) e-Infrastructure Reflection Group (e-IRG) (duration: to be updated)
g) FAIRsFAIR - Fostering FAIR Data Practices in Europe (March 2019 - February 2022)

These projects have conducted landscape studies to assess several indicators that indicate Open Science readiness in the participating European countries. The summary is as below:

Table 4.1. Indicators used to assess Open Science readiness in Europe

Projects	Participating countries in the landscape analysis	Indicators
EOSC - PILLAR	Australia, Belgium, France, Germany and Italy	• Business models (for funding bodies) • Service level agreements by e-infrastructures and service providers. • University policy regarding research data management, open research data and compliance with the FAIR data principle. • Familiarity with EOSC and FAIR data principle. • Effects of EOSC on organisation and individual plans, and benefits gained from EOSC.
EOSC – Nordic	Iceland, Denmark, Norway, Netherlands, Germany, Sweden, Finland, Estonia, Latvia and Lithuania	• Open Science policies and access policies in EOSC – Nordic countries. • Open science resource provisioning in the countries: infrastructure of interest, willingness to be federated, and maturity.
NI4OS- Europe	Albania, Armenia, Bosnia-herzegovina, Bulgaria, Croatia, Cyprus, Georgia, Greece, Hungary, Moldova, Montenegro, Republic of North Macedonia, Romania, Serbia and Slovenia	• Awareness about EOSC. • Familiarity with FAIR principle. • Areas of training, support or advice to make data FAIR. • Organisational support and training.
EOSC Synergy	Spain, Portugal, UK, Czech Republic, Slovakia, Poland, and the Netherlands	• Experience, effort and resources of national publicly-funded e-infrastructures. • National plans/roadmaps for e-infrastructure and how these national plans feed the needs of EOSC.
ExPaNDS	10 National Photon and Neutron Research Infrastructures, as well as EGI.	• Data Policy – FAIR principle, raw data, metadata, identifiers. • Motivations for FAIR data. • Data Management Plan and Practices. • FAIR data training. • Data catalogues.
e-Infrastructure Reflection Group (e-IRG)	32 Member States and Associated Countries represented in e-IRG	• National e-infrastructure organisations and the coordination mechanisms among these, their governance, funding sources, access policy and coordination with domain-specific infrastructures.
FAIRsFAIR	Australia, Belgium, Cyprus, Czech Republic, Denmark, Spain, Finland, France, Germany, Ireland, Lithuania, Netherlands, Portugal Slovenia, United Kingdom, Norway, Switzerland, Serbia	• FAIR Policies and practices: Levels of maturity regarding FAIR practices among disciplines, policies and sources of support available to researchers. • Semantics and interoperability: Formats, semantic artefacts, identifiers and software practices. • Research data and FAIR data principles - State of FAIR data: Policy and support services, awareness of FAIR principle, teaching of research data competences and awareness and views of EOSC.

4.2.1 National Open Science Platform (Netherlands)

The Netherlands is one of the European countries that has made significant progress in implementing Open Science and data sharing practices. In the country, Open Science has been strongly supported and promoted by SURF, the Association of Universities - the Netherlands (VSNU) and the Ministry of Education, Culture and Science. Open Science is also supported by research funders, for example, Netherlands Organisation for Scientific Research (NWO), ZonMw and Royal Netherlands Academy of Arts and Sciences (KNAW) and other Open Science advocates such as Netherlands Data Archiving and Networked Services (DANS).

Funder Policy Supporting Open Science

a. Netherlands Organisation for Scientific Research (NWO): research results paid for by public funds should be freely accessible worldwide. Implemented data management policy in all NOW funding instruments.
b. Royal Netherlands Academy of Arts and Sciences (KNAW): All publications by KNAW researchers are freely available, preferable immediately, but at the latest after 18 months.

Institutional Policy

a. Dutch Universities (e.g. Erasmus University Rotterdam, Eindhoven University of Technology, TU Delft and University of Groningen) have adopted an Open Access policy for their institution.
b. Universities of Twente, Utrecht and Delft encourage Open Access publishing through a special fund.
c. VU Universities Amsterdam and Utrecht University and TU Delft support researchers who would like to set up Open Access journals.

Infrastructure and support

a. All universities in the Netherlands have their own repository that primarily stores Open Access publications.
b. The national portal NARCIS provides access to Open Access publications, datasets of a number of data archives and descriptions of research projects, institutes and researchers. DANS with the data archive EASY and 4TU.ResearchData provides durable storage and access to research data.
c. The national Open Access website www.openaccess.nl provides information on open access publishing for various stakeholders.
d. DataverseNL is a partnership that jointly manages and deploys the Dataverse Network open source application for the archiving, citing and sharing of research data. A number of universities and research institutes joined this partnership.

Capacity Building (Data Steward Competencies) activities/initiatives

a. Training sessions focusing on the FAIR creation of data, are also offered by DTL
under the auspices of ELIXIR.
b. ZonMw Project: The aim is to professionalise the data steward function within the life-science domain with a special focus on implementation of the FAIR principle. Many parties actively use open science in their initial research phase.
c. TU Delft Data Stewardship Centre.
d. Training sessions for research are offered through SURF. In addition to the training sessions, SURF offers consultancy and focuses on the connection between research support at an institutional level and the support from SURF under the name Support4research.
e. The eScience Centre offers training sessions such as data carpentry and software carpentry, wherein the focus lies fully on open science and open source.
f. Research Data Netherlands, which is a collaboration between 4TU.Centre for Research Data, DANS and SURFsara, offers researchers the Essentials 4 Data Support course.
g. Professionalising data stewardship in the Netherlands. Competences, training and education. Dutch roadmap towards national implementation of FAIR data stewardship was produced in March 2021.

4.3. Committee on Data of the International Science Council (CODATA)

CODATA is the Committee on Data of the International Science Council (ISC). Its mission is to connect data and people to advance science and improve our world. CODATA supports that data produced by research should be as open as possible and as closed as necessary. CODATA works also to advance the mnemonic FAIR (Findable, Accessible, Interoperable and Reusable) principle in data sharing practices. By promoting the policy, technological and cultural changes are essential to promoting Open Science., CODATA helps advance ISC's vision and mission of advancing science as a global public good.

Some initiatives that have been carried out in line with the mission of CODATA include:

a. The CODATA Report on Current Best Practice for Research Data Management Policies (2015).
b. The 2014 Joint Declaration of Data Citation Principles.
c. Legal Interoperability of Research Data: Principles and Implementation Guidelines.
d. Series of workshops held globally to promote use of citation in a standardised format for reused data, by bringing together international experts and national stakeholders including researchers, journal editors, scholarly societies, research institutions, funders, publishers and data repositories.

These initiatives were made possible under the leadership of a number of standing committees and strategic executives, through its task groups and working groups. CODATA also contributes to the Data Science Journal and actively collaborates in conferences such as SciDataCon and International Data Week.

4.4. African Open Science Platform (AOSP)

4.4.1. Background

The African Open Science Platform initiative (AOSP) was implemented and managed by the Academy of Science of South Africa and is funded by the South African Department of Science and Technology through the National Research Foundation. AOSP is a three-year project launched by the then Minister of Science and Technology in 2016 during the Science Forum South Africa.

The vision of AOSP was laid as follow:

1. A federated system that provides scientists and other societal actors with the means to find, deposit, man age, share and reuse data, software and metadata in pursuing their interests. This will comprise three strands:
   Strand 0: Promotion and registration of trustworthy African data collections and services.
   Strand 1: Cloud Computing;
   Strand 2: Practices and tools of Open Research Data Management.

2. A network of dispersed participants working though activity nodes in pursuit of shared and overlapping open science goals delivered through:
   Strand 3: The African Data Science Institute;
   Strand 4: Interdisciplinary Programmes (e.g. infectious disease; biodiversity; agriculture; resilient cities; disaster risk reduction; precision medicine; open innovation);
   Strand 5: Network for Education & Skills;
   Strand 6: Network for Open Science Access and Dialogue.

During its pilot phase (2017 -2019), the priorities of AOSP fell into four specific areas:

a. Establish the foundations for a functioning AOSP.
b. Map the current landscape of data/science initiatives in Africa, assessing several indicators including Afri can research policy or legislation framework, Research Data and Open Science initiatives (e.g. DOAJ-listed journals, Institutional Repositories registered on OpenDOAR, Open Access policies on ROARMAP, and re pository certified with CoreTrustSeal), Institutional Research Data Management Plan, research culture and ICT infrastructure.
c. Build a Pan-African Open Science community and encourage national Open Science forums.
d. Create a roadmap for the Platform development.

Following the completion of the pilot phase project, AOSP Project Office is now hosted by the National Research Foundation for the next 3 to 5 years. AOSP is also further supported by South Africa's Department of Science and Innovation, key institutions in Africa and in the International Science Council.

4.4.2. Available policies

There is no open science policy that covers the African region. However, there are 44 open access institutional policies that are registered in the Registry of Open Access Repository Mandates and Policies (ROARMAP). Ethiopia is the only country which has a national open access policy established in 2019.

A few progresses on Open Data/Open Science Policies are:

- Botswana –Draft White Paper on Open Research Data Strategy.
- Madagascar –Lobbying for Open Data policy.
- South Africa –White Paper on STI.
- Uganda –Draft Open Data Policy.

4.4.3. Number of Repositories

There are 174 Institutional repositories registered on the Directory of Open Access Repositories (OpenDOAR). 24 data repositories are registered on the Registry of Research Data Repositories (re3data), while the African Open Science landscape study identified more than 66 repositories in the African region. In all the repositories that were identified, only 1 data repository was assigned the CoreTrustSeal.

4.4.4. Skill Capacity

Various training activities have been conducted for open access publishing in Africa. For example:

- National Research and Educational Networks (NREN) provides training on subjects such as network management, blockchain, programming, data processing
- Online Open Science courses such as Open Science MOOC, Coursera, FOSTER, MANTRA, AIMS
- CODATA-RDA School of Research Data Science (Rwanda, 2018) – AOSP sponsored 3 participants to enrol in the courses
- AOSP School of Research Data Science in collaboration with CODATA & RDA (Ethiopia, 2019) with more than 20 participants

4.4.5. Infrastructure Capacity

Infrastructure is a challenge for the African region. Academic and research-intensive institutions in Africa rely heavily on NRENs (National Research and Educational Networks). However, many countries have low awareness on the importance of NRENs. Many NRENs are non-operational with low to no budget. Private Internet Service Providers (ISPs) in Africa also influence the cost for a person to access the internet. They monopolise the market, making internet connection inaccessible for some. Power outages also make it difficult for Open Science to prosper in Africa. Other than that, from the landscape study, 20 African governments applied some form of Internet censorship 45 times since 2001, of which 36 times the shutdowns were related to anti-government related protests.

4.4.6. Awareness

There are a few open science and data sharing activities that have been carried out throughout the African region. For example:

a. 12 Open Science-related (Open Access/Open Data/Open Science) declarations and agreements endorsed or signed by African governments (Academy of Sciences of South Africa, 2019);
b. 196 Open Access journals from Africa registered on the Directory of Open Access Journals (DOAJ);
c. 174 Open Access institutional research repositories registered on OpenDOAR (Directory of Open Access Repositories);

The existing research culture in Africa is a significant prohibiting factor that discourages data sharing practices among research communities in the region. Research performance evaluation for institutional metrics and funding systems rely heavily on publishing in high impact factor publications. Data sharing is not acknowledged for career promotional purposes or performance appraisals and this has disincentivised researchers from depositing their data in repositories. Since researchers are pressured by the necessity to publish their work in journal articles, they would maximise the use of created research data for publication purposes rather than sharing them on a research data platform.

Some researchers are also at a disadvantage in terms of research partnerships resulting in disproportionate benefits that may discredit their contributions in a research or disregard them to assume more prominent authorship positions in joint publications. They are also often excluded from being part of a research project that is happening in their localities and are not well supported to pay for publication fees. All these factors may discourage the researchers from participating in research data sharing activities.

4.4.7. Current activities

The first consultative stakeholder meeting was held during 2-3 September 2019 in Alexandria, Egypt. The National Research Foundation (NRF) of South Africa will host the African Open Science Platform (AOSP) Project Office for the next 3 to 5 years. The AOSP Project Office, based at the NRF in Pretoria, South Africa, with selected staff appointments, is due to be launched in 2020.

4.5. Benchmarking with other Open Science Initiative

By reviewing the past, current and future activities, planning and other initiatives for several Open Science in multiple regions/countries, this study attempts to evaluate and benchmark global Open Science initiatives against Malaysia based on several indicators, as illustrated in Table 4.2.

Table 4.2. Benchmarking the status of Open Science initiatives in Australia, Africa, the Netherlands and Malaysia

Indicators	Australia	Africa	Netherlands	Malaysia
Formal Open Science Policy				
Open Science Repositories				
Skill Capacity				
Infrastructure and Networking				
Awareness				
Current activities				
Funding				
Expertise				
Public Engagement				

Note:

Available	In Progress – Requires more commitments (funds, expertise, capacity building, etc)	Not available

4.5.1. Australia's Open Science as Best Model for Malaysia

The Netherlands has been committed to establishing Open Science ecosystem in the country, primarily towards ensuring all publicly funded scientific publications to be published in Open Access form. Similarly, Malaysia, through local research funders and institutions, has been supporting researchers to publish in Open Access journals in the past years. In addition to that, Malaysia also recognise the importance of gathering and consolidating raw research data into a trusted and reliable platform, which the country is currently lacking.

Australia, through the Australia National Data Service, has embarked on Open Science initiatives to make Australia's research data assets more valuable for researchers, research institutions and the nation, by creating platforms and tools that enable the discoverability of research datasets, establishing relevant Open Science policies, capacity building, incentivising researchers to share data and creating awareness about Open Science. Therefore, Australia is regarded as the best model for Malaysia Open Science Platform to learn from in the perspective of the government, researchers, capacity building and infrastructure.

a. Framework

The structure of the Australian Constitution delineates the traditional role of the Commonwealth Government that sets the government's central direction and policy settings including its responsibility for funding the higher education section. The unique environment of Australia provides an opportunity for the Australia National Data Service to encourage appropriate data management practise nationwide, such as with the release of the Australian Code for the Responsible Conduct of Research (NHMRC/ARC/UA2 2007). Malaysia is similar to Australia in the sense that the majority of research funding allocated to public universities are supported by public funding. This similarity exemplifies the role of MOSP to encourage best data sharing practices in the country for accountability to the public and to promote greater engagement between the public and researchers.

b. Open Science Policy

Australia does not have a formal national Open Science policy per se, however ANDS outlined national guidelines (e.g. data storage, metadata, file formats, data management plans, responsible conduct for research etc.) that serve as a template for local institutions to develop institutional policies adhering to the national standards of data sharing and Open Science practices. The visibility of Open Science in Australia is prominent as the country has identified Open Science in its National Science, Technology and Innovation Agenda. The incorporation of Open Science will help steer the country towards creating a knowledge-based economy that thrives on many aspects including Open Science and Open Innovation.

c. Platform architecture

Australia assume research data are stored in institutional and data publishing partners' repositories while ANDS provides a data discovery service for these repositories. This is different than UK's discipline centre approach or the Dutch national research data repository (i.e. DANS), but is consistent with the recommendations made by NHMRC (2007). MOSP, on the other hand, must consider existing institutional policies regarding arrangements for research data storage as well as its mission of enabling discoverability of Malaysia's research data when designing its platform architecture.

d. Funding

Funding has always been a constraint for a nationwide involvement in Open Science, therefore financing highly targeted programs is necessary for the ANDS program to move forward. ANDS started with consolidating existing useful content into repositories and identifying repositories housing useful content, and making all contents discoverable through the Australia Research Data Commons (ARDC). ANDS progressed by providing financial support to selected projects in which the outputs were also made discoverable through ARDC. MOSP should also be selective and strategic when piloting the national Open Science initiative. For example, the first phase of MOSP can be deployed by involving the five Research Universities in Malaysia, (i.e. Universiti Malaya, Universiti Sains Malaysia, Universiti Kebangsaan Malaysia, Universiti Putra Malaysia and Universiti Teknologi Malaysia), which are the primary public funding recipients and research data producers in the country. MOSP can consolidate all research data from these universities and make the content discoverable to all levels of communities.

e. Building capabilities

ANDS serves as the national coordinator of capacity building programs, and works with institutions, both research-intensive organisations and government agencies holding data or interest to researchers, to develop national curricula and organise institutionally based training activities. Data stewards are trained to support Open Science practices in the country, specifically providing retention and access services to research data, while researchers and research support staff are trained to improve data management practices.

f. Governance Structure

The Australian government has been successful at swiftly responding to any shortcomings identified in its existing policies and initiatives. This is possible through continuous reviews that assess performance and identify the improvements necessary to improve any particular programmes in question. For example, the government made improvements to R&D Tax Incentives after an all-encompassing review on its effectiveness and impact to businesses.

Another approach that can be applied to the Malaysian context would be Australia's whole-of-government approach in identifying issues and challenges across industries through the Department of Industry, Innovation and Science, that coordinates and consolidates feedback across industries, which are then fed into the policy-making process for consideration. This approach minimises, if not eliminates, redundancy in processes and administration which can increase the cost of compliance in delivering the initiatives. It also ensures consistency in processes, assessments, and compliance with initiatives; and avoids overlapping initiatives that allow for grant/funding shopping to occur.

g. Considerations for Competitive Research Grants vs Matching Grants

Some of the initiatives offering grants and funding in Australia require matched funding from other sources. This ensures that each application is genuine and that the programme owner is just as committed to reaching the project objectives. It is also an excellent tool that provides incentives for collaboration.

There are, however, several key considerations and critical operational practices which determine success/failure of matching grant initiatives. Grants targeted to service buyers limit the risks of undercutting the service providers of the services that fall under allowed expenses, as compared to subsidies to state agencies (Phillips, 2001).

Impacts and externalities are another important consideration in determining the type of funds that cater to the different types of research and stages in product development and commercialisation. It is important that these types of grants be complemented with more sustainable funding programmes ensuring sustained sector development in the longer term and are in alignment with the national agenda and policies. Some key operating principles include (Rajalahti & Farley, 2010):

i. Transparent procedures and timely and accurate dissemination of information to the public.
ii. Criteria for selection have to be clearly defined, and selection has to be performed by qualified scientific peers.
iii. Alignment with the national needs is crucial in ensuring the grants, directly and indirectly, benefit the intended beneficiaries.

h. Operational Efficiencies

i. Grant Database

A database containing an exhaustive list of grants and initiatives is available online with up-to-date details on programme launches and documentation required for fund applications (business.gov.au, 2017).

ii. List of Funding Offers

An exhaustive list of funding offers is available online and includes the details of applicant names, titles of projects, and grant amounts. This provides more transparency and instils confidence in the governance and administration of the funds.

CHAPTER 5

PERSPECTIVES ON OPEN SCIENCE FROM TOP MANAGEMENT OF RESEARCH UNIVERSITIES

Open Science represents a change in the way key actors or stakeholders in research, education and knowledge exchange communities create, store, share and deliver the outputs of their activity. For Open Science principles, policies and practices to be fully embraced in universities, there needs to be a holistic approach to organise each Open Science initiative, which is ultimately integrated and is coherent with the overall goal of Open Science, is understood across all levels especially by researchers, and involves everyone's participation.

This chapter sets out the perspectives on Open Science from Top Management of Research UniversitiesUniversity (RU), delving into institutional interests and activities in Open Science, and Vision for participation in Open Science initiatives including MOSP.

5.1. Institutional Interests and Activities in Open Science

Based on interviews with top management of the five Research Universities in Malaysia, the top management personnel were asked to describe the university's interests and motivation to participate in Open Science. The first point made was the importance of democratising knowledge that allows all inquiring societies to seamlessly access information, knowledge, research data and findings. Data sharing practices were also viewed to have a profound impact on society at large through generation of new knowledge and development of impactful end-products.

In addition, a University has interests in Open Science because through Open Science, a University can greatly reduce costs paid for journal subscriptions, which may range from several hundred thousand to several million Ringgit Malaysia per year. There is also an opportunity for local universities to consider eliminating the need to subscribe entirely, as practised in some universities in Europe and the US. Open Science and data sharing also provides an opportunity for other researchers to view, access, and reuse the deposited research data, which will ultimately increase the visibility and citations of researchers at universities. Data sharing platforms allow researchers in other countries that cannot afford to pay expensive fees for subscription-based journals to access research data.

Open Science is also perceived as a data sharing ecosystem that improves research integrity and transparency since research data will be peer-reviewed by platform users. The peer-review process will enhance data reusability and consequently optimise the use of data for validation or other purposes. By maximising the use of produced data and avoiding redundancy in providing funding for similar projects, it can prevent unnecessary duplications of research or 'reinventing the wheel', thus avoiding waste of monetary and human resources while increasing research efficiency. Due to the various benefits Open Science that it could offer, it has been accepted as a new culture in research universities with several Open Science initiatives having been implemented at the institution.

[11] An interviewee described Linux as a successful case study that illustrated impacts of shared R&D and open source innovation in revolutionising the IT industry. More information about Linux can be found in this link: https://www.linuxfoundation.org/

5.1. Examples of current Research University's Open Science Efforts and Initiatives

.Policy and incentives:

1.Implemented policy that requires researchers to publish their findings, and all research must be compiled into research reports which is then stored in an institutional repository.

2.Implemented university's research ethics policy requiring that all data be curated and stored (longer than the mini mal time) so that the produced data is available for scrutiny when necessary,

3.Policy implemented by library units and RMC offices related to institutional repositories

4.Ongoing development of the Data Management Plan Policy

5.Providing funding for researchers to publish in Open Access journals or publish Open Access scholarly articles.

Infrastructure:

1.Establishment of institutional repositories for digital collection of University's publications under management of Library units.

2.Engagement with stakeholders to establish institutional research data at an institution. For universities that have the repository, researchers are encouraged to deposit their research data.

3.Voluntary practices for data management plans among researchers, with supports from Libraries and Re search Management Centre (RMC) offices.

Capacity Building and Awareness:

1.Advocation towards promoting Open Science to researchers, such as promoting the use of data repositories for the purpose of data comparison and validation.

2.Empowering development of new Open Science skills for librarians.

(Source: Analysed by ASM, 2020

The following graph (Figure 5.1.) illustrates the percentage of Open Access publications by researchers from the five Research Universities in Malaysia over the period of 2001 to 2020.

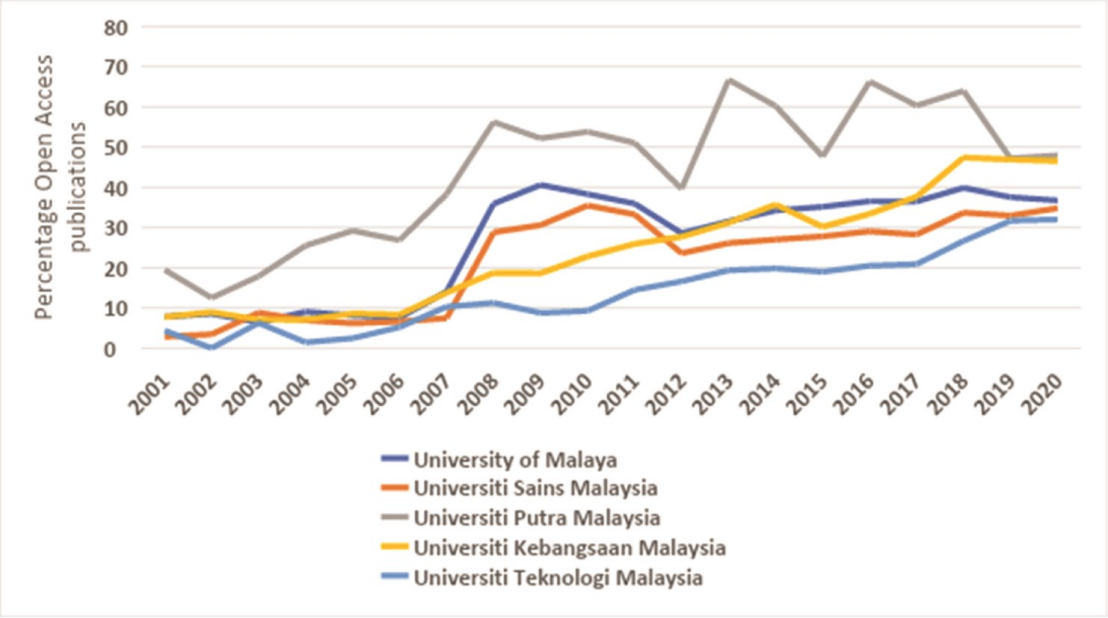


Figure 5.1. % of Open Access publications from the five Research Universities in Malaysia (From 2001 to 2020) This figure is accurate as of 18th July 2020. Analysed by ASM (2020)

.

Policy in supporting Open Science practices and implementation

"We do not have a formal policy for Open Science yet but the mechanism that we have been practising is planning for any Open Science-related activities which must be discussed and approved by the Top Management"

Institutionalising policy related to Open Science, such as the Research Data Management Plan Policy is a key determinant to Open Science readiness in an academic entity. Based on the interviews conducted with top management at the five Research Universities, only two universities have the Policy drafted but none of them have the Research Data Management Plan Policy approved and implemented. However, the universities do have policies related to institutional repositories, research ethics, and policies requiring researchers to publish their findings and depositing final research reports on institutional repositories.

Support in building talents capacity (i.e. data curators and data stewards)

"Top Management provides supports to staffs especially librarians to enhance research strategies, such as learning SciVal, which can be imparted to researchers"

Support in capacity building related to data curatorship and stewardship has been a central discussion agenda and initiative of Research Universities in Malaysia. For example, regular in-house training workshops, seminars, courses, or training workshops about data management by top publishers were held to give exposure to university staff, especially librarians to learn new methods, techniques or softwares. Librarians are also encouraged to upskill their IT skills through training on virtual technology skills and Big Data management. Staff are also entitled to financial support and sponsorship to attend paid training for data curation and stewardship. Those who have been trained will become trainers for the Training of Trainers (ToT) sessions to impart acquired knowledge to their colleagues.

5.2. Vision for participation in Open Science initiatives including MOSP

Five Research Universities i.e. Universiti Malaya, Universiti Sains Malaysia, Universiti Putra Malaysia, Universiti Kebangsaan Malaysia and Universiti Teknologi Malaysia are involved as the primary research data contributors for the pilot-scale implementation of Malaysia Open Science Platform (MOSP) which has beenis running from 2019 to 2021. The selection was made on the basis that these universities are research intensive institutions where the majoritywhere majority of research activities conducted at these institutions are funded by the Malaysian government, withand they have some experience in managing in-house repositories. Therefore, framing the execution and implementation of MOSP will need to consider and incorporate visions for participation in Open Science of the five Research Universities:

i. The value of a tertiary education provider is derived from its embodied core of responsible science, in which new research data should be shared to enable scientific progress and accelerate benefits gained by beneficiaries. Under this notion, all Research Universities play a critical role in supporting Open Science to create a great leap in the country's research and innovation, subsequently imparting positive spillovers to the national economy and society.
ii. Open Science is a key initiative that can contribute towards advancing national science and technology to becoming on par with developed countries. All university staff should be involved in Open Science related activities.
iii. The MOSP initiative should be under a custodianship of a trusted entity such as MOSTI to ensure data security and to build confidence for stakeholders.
iv. The MOSP Platform must be user-friendly and financially sustainable in the long term.
v. As a way forward, a long-term roadmap for MOSP is necessary to change the current "data sharing" culture, instilling researchers with the importance and benefits of Open Science. Additionally, Open Science and MOSP should be led by strong leadership that will coordinate efforts from multiple stakeholders and define each responsibility and contribution.

[12] Currently R&D&C funding by MOSTI is very much aligned to undertake experimental development towards commercialisation and to stimulate the growth of STI industries. Most R&D activities (basic and applied) are funded by MOHE. Therefore, research data repositories

5.3. Key takeaways

i. Publishing Open Access articles and in Open Access Journals is a common practice for Research Universities in Malaysia. However, other essential elements of Open Science such as the Research Data Management Plan is new to the universities; and

ii. The Top Management of the five Research Universities are fully supportive towards implementing Open Science in Malaysia.

are very much held by IHLs which needs MOHE and MOSTI as strategic custodians of MOSP in order to ensure data security and build confidence for stakeholders.
(basic and applied) are funded by MOHE. Therefore, research data repositories

CHAPTER 6

PERSPECTIVES ON OPEN SCIENCE FROM RESEARCHERS

Efforts to promote and implement Open Science practices are to large extent driven by the willingness and motivation of researchers to contribute and participate in data sharing. There is, in fact, ample evidence and success stories demonstrating the successful implementation of Open Science which is characterised by researchers' free willingness in sharing their data. This chapter provides an examination of the current awareness, practices and required skills and supports as well as concerns related to Open Science from the perspective of researchers.

6.1 Awareness about Open Science & FAIR Principle

There is no universal standard definition of Open Science, but the key principle promoting Open Science is the "openness" of a wide spectrum of datasets, publications, materials and methods for multiple groups of people, and at multiple levels and geographies. The diverse definitions of Open Science may contribute towards different levels of understanding and "openness" practices among the key actors, especially researchers. Awareness about data"FAIRness" is the core of Open Science, as FAIR determines the usefulness of shared data. Despite the importance of FAIR in data sharing practices, only 12% of researchers in Malaysia are "familiar with FAIR principles", according to the recent survey conducted with 208 respondents. The survey further unravelled clear disparities in understanding of the Open Science concept, depicting the various views of Open Science in daily research works.

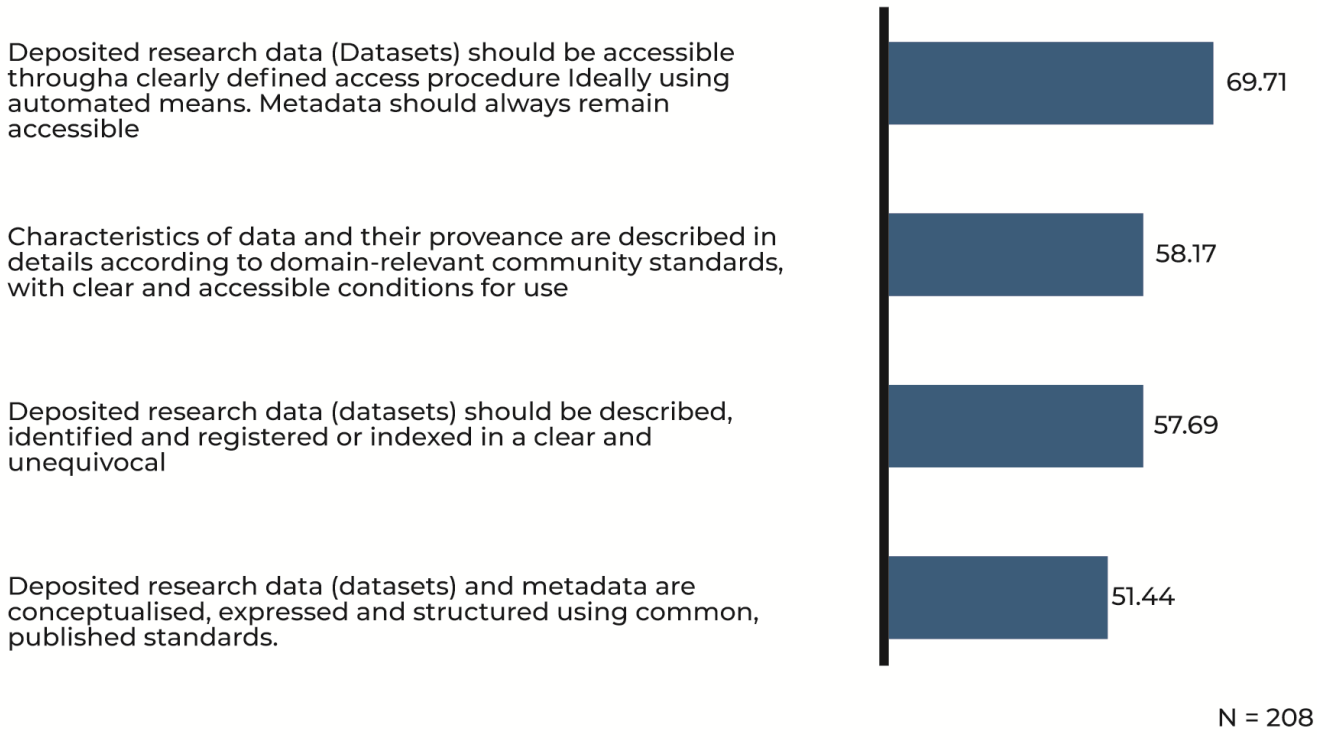


Figure 6.1. Which statement(s) is/are relevant and applicable about Open Science? You may choose more than one option

[13] Demographic background of survey participants (Researchers) as in **Appendix 6.1.**

Interviews with researchers found that data sharing and Open Science practices are generally viewed positively in multiple aspects. Firstly, Open Science can contribute towards research integrity and maximise the use of resources. Secondly, Open Science is regarded as a tool for researchers to generate new research ideas and to develop a wide array of research variables. For example, Open Source, which often refers to software and source codes, is very much perceived as an essential tool for researchers to use and modify for different purposes, which otherwise access to such data is impossible since they are not published in journals. Secondly, shared data on open platforms is also valuable for benchmarking exercises, allowing researchers to compare and contrast proposed solutions with benchmarked datasets. Thirdly, by making research data available on a sharing platform, researchers also perceive that this could increase visibility and citations of their research work, and more importantly, facilitate establishment of new collaborations either intra or inter-disciplines, and networking among research peers. The last point made by researchers is that Open Science helps in translating research for societal benefits by allowing various stakeholders such as policymakers and industrial players to re-use research data for more impactful outcomes to the public.

Box 6.1. Three examples of Open Science practices from researcher's perspectives

Case study 1.

Researcher A is a Professor from a business school at a Research University in Malaysia. One of her current research interests is exploring societal returns and evaluating project impacts from monetary and non-monetary perspectives. She used benchmarking or baseline data from similar projects that were deposited on international platforms to design objectives, and research methods, as well as to develop effective measures for evaluation of program impacts. To her, having access to valuable data allows cross-country comparisons and enriches her understanding on a research topic when she is able to compare and contrast findings resulting from different research settings.

Case study 2.

Researcher B is a Senior lecturer from an education school at a Research University in Malaysia. Her motivation when publishing research is not about citations but to disseminate and share research findings with practitioners i.e. teachers. This is because teachers will not read scholarly articles and therefore, she prefers to publish her work in modules or textbooks, which are commonly written in a more understandable manner and in layman terms. Unfortunately, such publications have a lot less weightage in the KPI system in comparison to scholarly articles, and this can somehow be discouraging. She hopes this situation can change so that researchers like her can employ more practical ways to disseminate knowledge.

Case study 3.

Researcher C is a Professor with a research background in climate change. The Research Institute that he is working at has established a database that store forestry-related datasets, which has been valuable and is shared with other research institutes and government agencies upon requests. Such data sharing practices have facilitated initiation of new research collaborations, and often results in joint publications.

For the Open Science initiative and FAIR principle to be fully embraced among research communities, researchers must not only be familiar and have a good understanding of Open Science and FAIR principle concepts, but they must also have knowledge about Open Science-related policies established at the institution they are working at. Zooming in on Open Science-related policies, a majority of researchers were not actually aware if such policies existed or vice versa, thus indicating that clear communication and application of policies must be established to ensure a holistic implementation of Open Science in the future

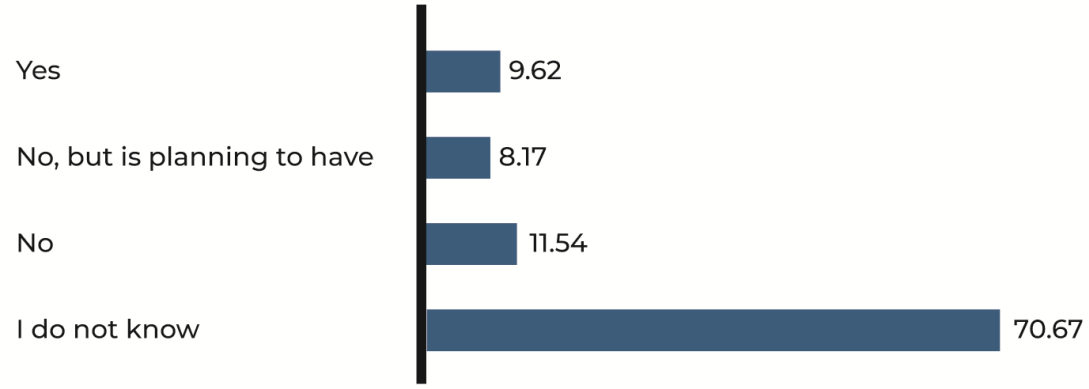


Figure 6.2a. Does your institution have any Open Science Policy?

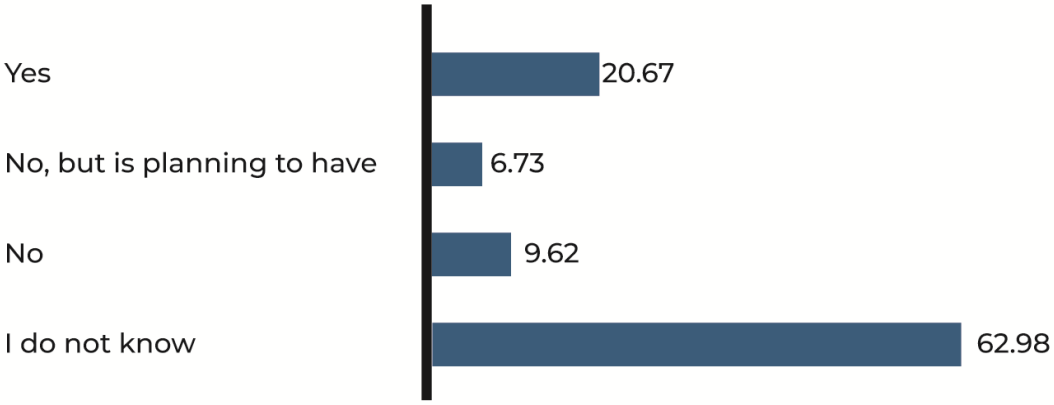


Figure 6.2b. Does your institution have any Open Access policy?

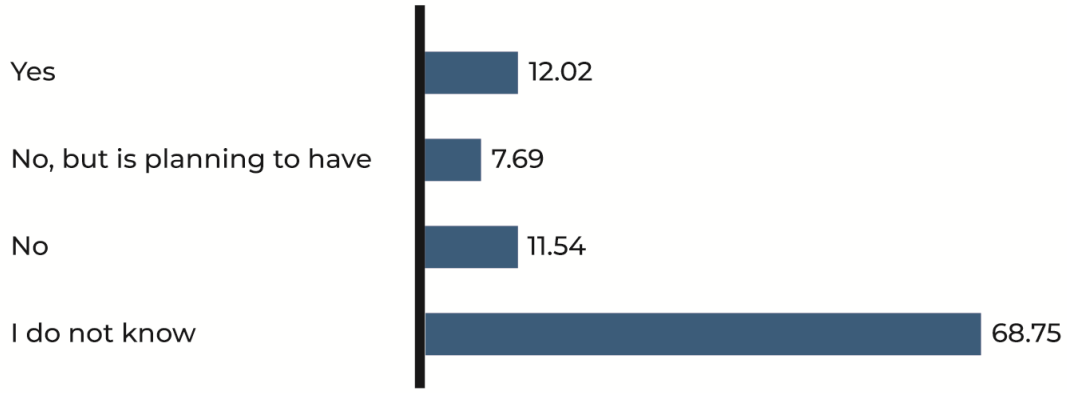


Figure 6.2c. Does your institution have any Open Data Policy?

6. 2. Practising Open Science

Open Science takes in different forms. It can involve (1) researchers publishing scholarly articles in Open Access journals and proceedings such as Nature Communications, PLoS and Frontiers, or publishing Open Access articles in Journals such as Nature, Cellulose and Topics in Current Chemistry; (2) depositing underpinning scientific research results on data sharing platforms such as in general repositories Dryad, Figshare and Zenodo, or discipline-specific repositories; (3) discussions to find potential solutions to a problem on a collaborative platform; (4) active involvement of the public in a research project and (5) use or release software codes in the public domain for further reuse.

Based on the survey carried out recently, a majority of researchers in Malaysia have been involved in Open Access publications, as indicated in Figure 6.3. The finding resonated with opinions from the Top Management Universities on the University's strategies and initiatives that encourage researchers to publish in Open Access journals.

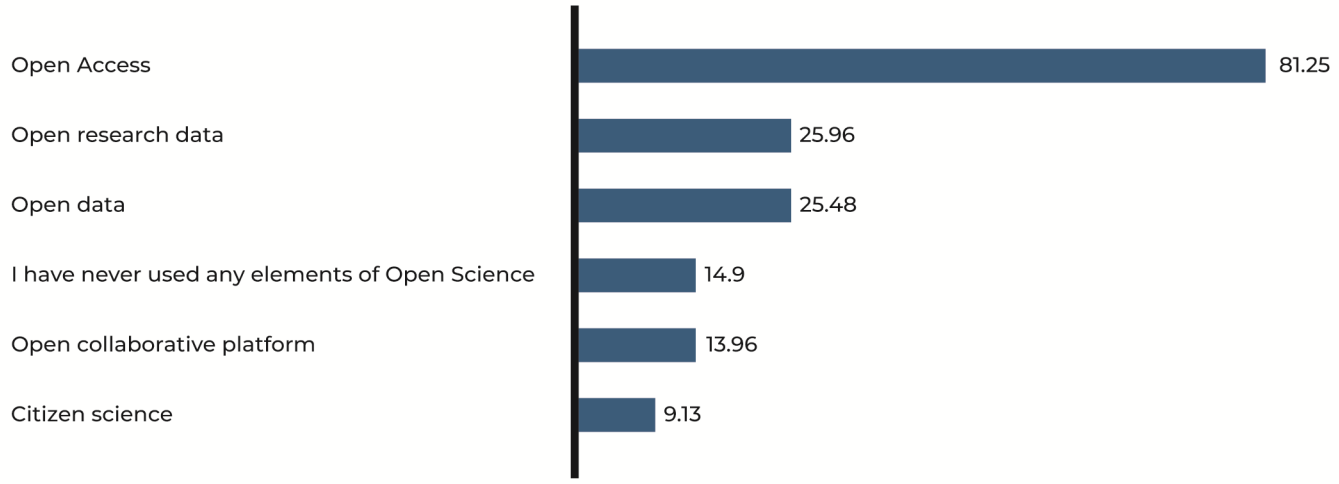


Figure 6.3. What elements of Open Science have you used? You may choose more than one option.

Besides publishing in Open Access journals, researchers in Malaysia have shared their processed research data using many ways such as publishing in journals as supplementary materials, presenting results in conferences and through informal means either via emails, websites or upon personal requests (Figure 6.4). While researchers are more open to sharing processed data, very few of them have actually shared raw research data on sharing platforms, including discipline-specific, general and institutional repositories. Sharing raw research does not seem to be a common practice among local researchers since a majority of them prefer to keep raw data on their personal storage, which is often inaccessible to others, either on local computers, thumb drives or hard disk drives (Figure 6.5). This situation poses a serious challenge to universities seeking to create a data sharing culture that aligns with the spirit of Open Science.

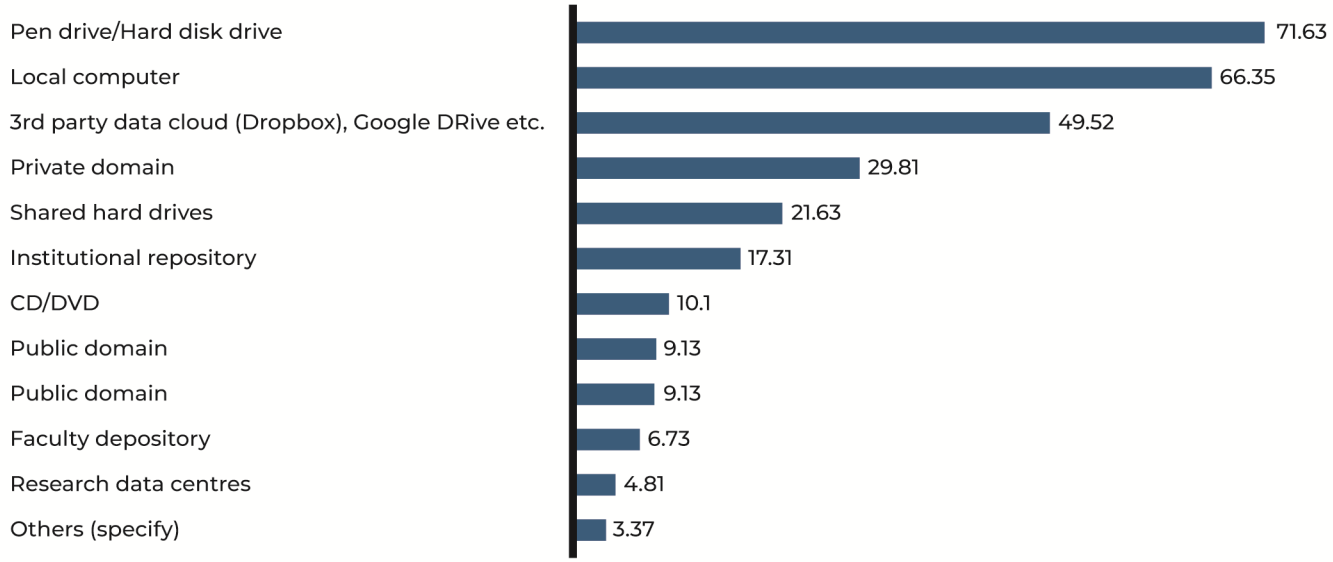


Figure 6.4. Where have you shared your data? You may choose more than one option

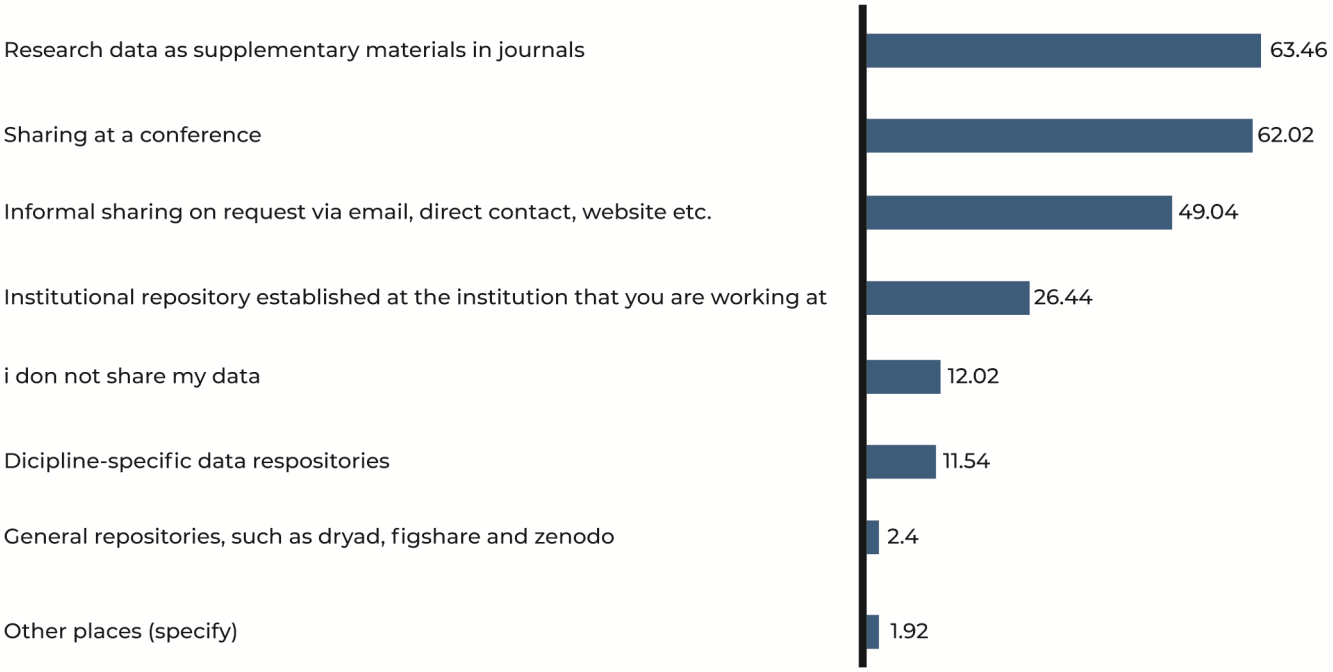


Figure 6.5. Where do you store your data? You may choose more than one option

6.3. Supporting Open Science Practices

A majority of researchers have not actually received formal training or participated in courses on research data management, which is the skill that they felt they needed to be trained on the most. According to some researchers, they spent at least 20% of their research time reorganising, reformatting or trying to remember details about data. This finding is not surprising since most researchers are neither aware of the availability of training courses nor does the institution itself offer such programs. Besides research data management, at least half of the surveyed researchers perceived that they require training support in storing,

reusing and analysing information, ethical issues about sharing data, legal and policy knowledge on Open Science and procedures to create quality metadata. These findings illustrate that researchers not only need support in the "know-how" when practising Open Science, but they also require legal and general support, especially when dealing with research data with commercial, potential and confidential data derived from human research.

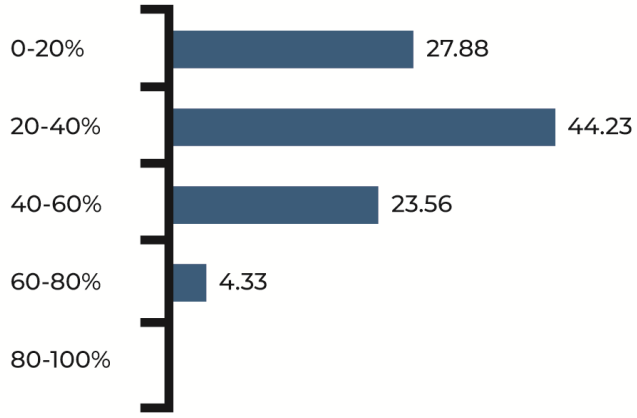


Figure 6.6. Can you estimate, as a percentage % of your research time, how much time you have lost reorganising, reformatting or trying to remember details about data?

Figure 6.7. Did you receive any training to manage research data from your institution?

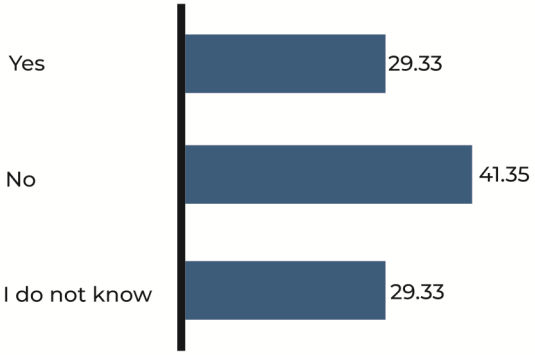


Figure 6.8. Does your institution provide any training to manage research data?

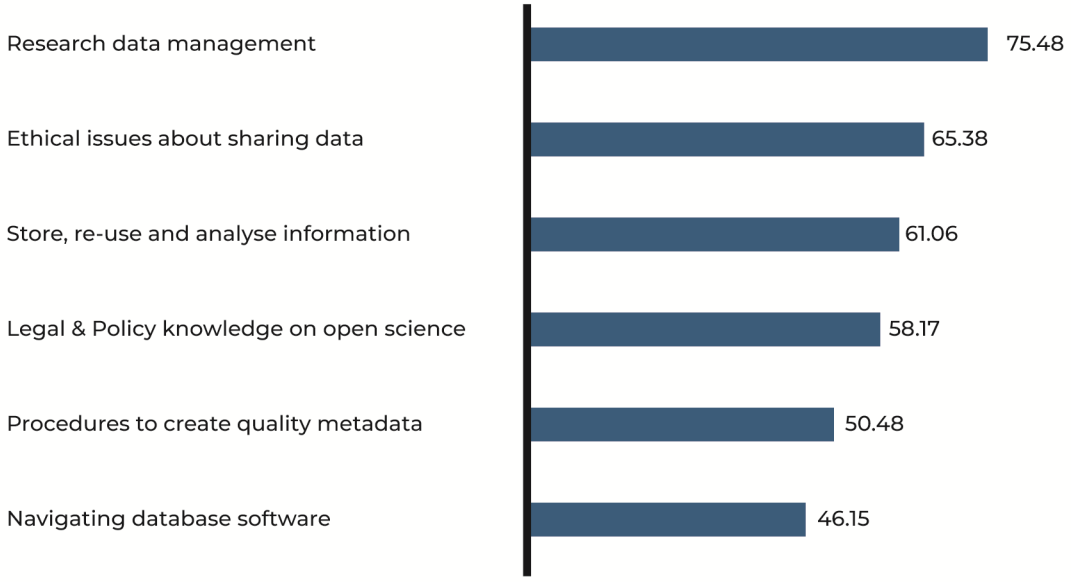


Figure 6.9. Which aspects do you feel you need support the most? You may choose more than one option.

To support researchers in creating and sharing quality research data, it is important to know the types of digital content that they have created. Researchers produce various types of research data, mainly research methodologies and workflows, fieldwork notes and sketches, and data collected from sensors or instruments. Genome and exome sequences, omics and transcriptomes, textual and tabular data were also created to a much lesser extent. Given the diverse types of research data produced, there is a clear need for universities to establish appropriate infrastructure, i.e. research data repositories, with enough capacity to store these research data in a safe and secure manner.

An important aspect of nurturing Open Science and data sharing practices is by educating students in university courses as well as doctoral students. In particular, the students must be taught about what entails Open Science and the FAIR principle, in particular the essentials of responsible conduct of research, research integrity, data management and transparency, to name a few. Based on recent interviews that were carried out, there is no formal education being offered to teach students on elements of Open Science and the FAIR principle. Nevertheless, students have been taught about ethics when handling research data, validity and reliability of research methodologies and outputs (e.g. qualitative research students were taught about the importance of descriptive audit trail to ensure reproducibility of research). Research students were also taught about the importance of planning, recording, safekeeping, reporting standards for research data and measures to be undertaken when dealing with sensitive data (e.g. obtaining consent from subject and anonymisation to protect privacy). While it is notable that students have been taught about Open Science and the FAIR principle informally to a certain extent and some values are already in practice by researchers, a minimal emphasis was given in classrooms to highlight and educate the importance and "know-how" of the Research Data Management Plan, ethics and legal aspects, the broad concept of Citizen Science (where researchers interact with the public to enhance research impacts) and other relevant elements of Open Science and the FAIR principle. Researchers need help and training in coping with these topics to support Open Science practices.

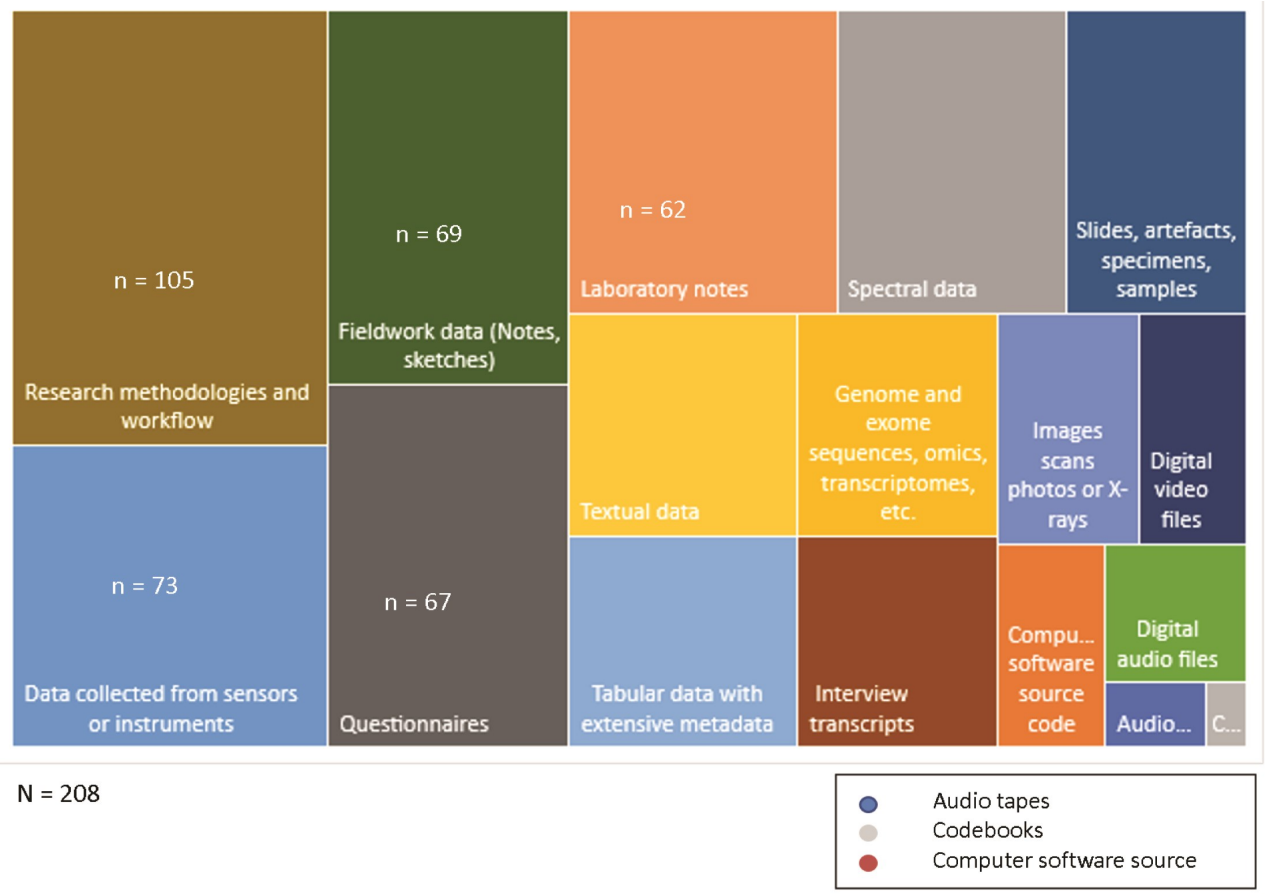


Figure 6.10. What types of research data do you produce? You may choose more than one option.

.

6.4. Challenges – Researchers' Concerns with Data Sharing

On an individual level, researchers may be concerned about ownership and entitlement over data they have created. However, it is important to highlight that in most universities in Malaysia, research data is owned by the institution and researchers only have the moral right and the non-exclusive right to use it at the mercy of the institution.

When discussing data sharing practices, early career researchers were worried if their research works would be scooped. To them, research is a continuous process, and by being open to sharing data, they may not be the "first to publish", hence losing the "first-mover" advantage. The "first-mover" advantage is extremely important for early career researchers to build his or her publication records in order to advance their career. In this context, a robust policy and appropriate mechanisms should be in place to protect early career researchers who want to share their data on a data sharing platform. An embargo period for data sharing can be applied to allow researchers to share data after the project, i.e. the grant funding, ends. Additionally, the data sharing platform should be supported by a strong policy that outlines appropriate acknowledgement and credits to data originators for data reuse, as well as governing shared data under several licensing mechanisms, including creative commons licensing.

In addition, some researchers are still resisting to openly share their data because of multitudes of issues related to data misinterpretation and misuse by third parties [14] For example, the interviewed researchers described that shared data may be manipulated by unethical researchers to stir sensational and unfounded headlines, which ultimately creates public confusions rather than educating the public. Researchers may also be concerned if their data will be reused for purposes they did not intend to, for example for commercial exploitation, or for supporting weak hypotheses, misleading or inappropriate secondary analyses.

[14] It is important to note that a majority of researchers preferred to keep raw data on personal storage which is often inaccessible to others, either on local computers, thumb drives or hard disk drives. This could be the greatest hindrance of data sharing and accessibility for MOSP. This challenge needs to be tackled with good data sharing policy and best governance.

Investing in skills, training and awareness for researchers and data stewards, as well as infrastructure could be possible solutions to the said problems. For researchers, they should be made aware of the ethical principle of research integrity, which is core to Open Science. Researchers must undertake research activities, deposit research data or publication results and interpret research findings in an open, honest, transparent and accurate manner to produce high quality data. Shared data must be FAIR, and all data originators and data users must be clear on their responsibilities and understand ethical rules when using the platform.

To build greater trust in data sharing practices, researchers should be supported by data stewardship, and operation of the data sharing platform must be regulated and monitored appropriately. For the latter, accountability and verification of proper data usage plays an important role in the data usage monitoring and thus, by keeping systematic records on essential information such as who used the shared data, the purpose of reusing the shared data and what datasets will be reused will be valuable to ensure that the shared data will used ethically by authorised parties and for approved purposes. This feature is exceptionally important for research datasets created in certain research fields such as the medical field since the produced data is not open by default.

6.5. Challenges – Use of traditional metrics that do not encourage Open Science or open scholarly practices

The current use of quantitative metrics, such as number of publications and arbitrary indicators (e.g. h-index) to assess research outputs as well as for career promotion and security tenure is one of the biggest barriers towards the adoption of Open Science. The traditional evaluation method must change and embody a broader set of Open Science values and practices to encourage researchers to participate in open scholarly practices. Examples include considering other types of Open Science activities, such as publishing research data, preprints and other forms of research products as measures of research outputs.

6.6. Key takeaways

i. Researchers have different levels of understanding about Open Science practices.
ii. Researchers are not familiar with the Findable, Accessible, Interoperable and Reusable (FAIR) principle.
iii. Researchers perceive Open Science and data sharing practices will greatly benefit their research endeavours.
iv. Researchers are not familiar with, and generally are not aware of Open Science related policies in the institution that they are working at.
v. A majority of researchers have published their research works in Open Access journals, as well as sharing their research findings as supplementary materials in journals, conferences and when receiving requests from emails, direct contacts or sharing via websites.
vi. A majority of researchers prefer to store their research data on personal storage.
vii. A majority of researchers described that training for the research data management plan is the skill they need the most.
viii. Researchers produce a wide spectrum of datasets across multi research disciplines.
ix. When sharing research data, researchers are worried about their research being scooped, ownership of the created research data and misuse and misinterpretation of shared data by third parties.
x. Researchers were also concerned about current traditional metrics for research performance that do not support Open Science and data sharing practices.

[15] An example of a sensational headline about mutant coronavirus published in a newspaper article, which was written based on a preprint article that was not peer-reviewed. Source:https://coronavirus.medium.com/whats-the-deal-with-the-l-a-times-story-about-the-mutant-coronavirus-a61d2ec11bc2 .

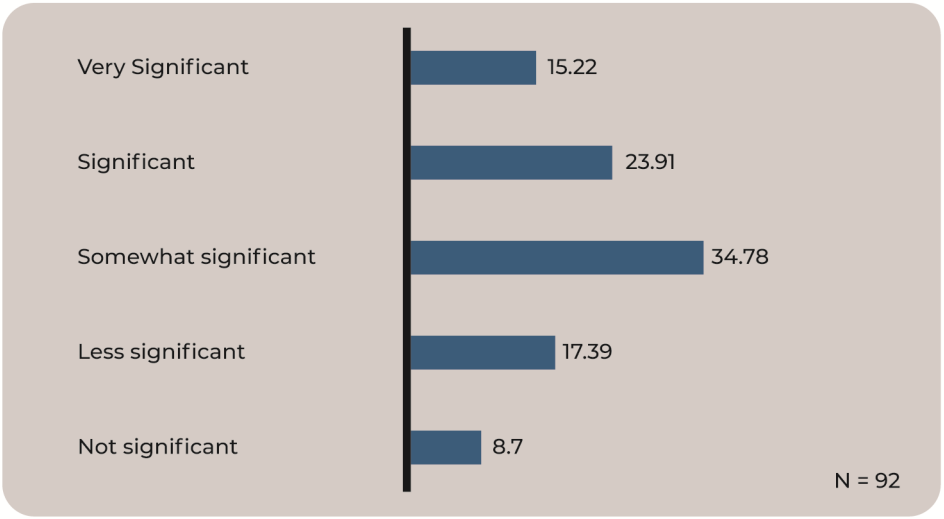
CHAPTER 7

PERSPECTIVES ON OPEN SCIENCE FROM LIBRARIANS

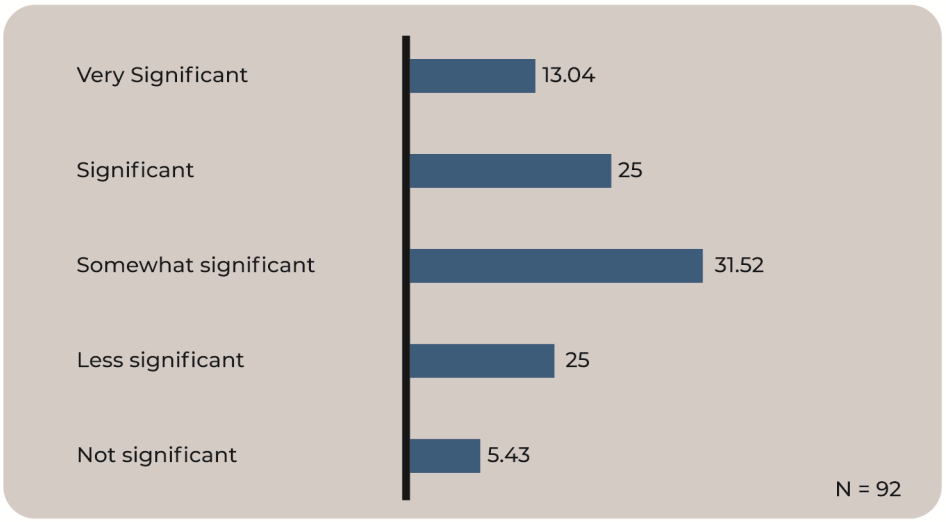
In an academic entity or research institution, librarians are actively involved in providing various research support services for the entire research process to researchers and students. For example, librarians play a critical role in organising training programmes, scheduling webinars, classes or training on data management as well as providing consultation services for students and researchers to conduct research, collect data, and write theses or manuscripts, and to provide advice on publishing. Librarians are also involved in managing Open Access publications and institutional repositories.

Box 7.1. A snapshot of survey on Open Science with librarians in Malaysia

Contribution in managing Open Access publications



Contribution in managing Open Access publications



[16] Demographic background of survey Librarians as in **Appendix 7.1.**

"Currently, we mostly work with structured data. We check if the data is annotated with names, authors etc. However, we do not deal with unstructured data or data cleansing."

With the emergence of Open Science in the digital age, the librarians' role has dramatically changed over the past several decades. Librarians have been recognised as a key enabler in Open Science and have been discussed on multiple platforms endorsed publicly by international organisation and stakeholders1. For example, academic libraries took up the role of managing institutional repositories which typically house numerous digitised collections of theses and dissertations, journal articles, conference papers, bulletins, newspaper clippings, to name a few. Librarians perform the curation task involving indexing, cataloguing and, classification of structured data. The indexing task, which is basically indexing the material by names or subjects for example, is carried out with a reference to the Resource Description Framework (RDF). The cataloguing task is to provide more specific and descriptive content of the material, and is based on the Resource Description Access (RDA).

7.1. Librarians as Data Curators and Data Stewards

"The strength of librarians is the skills that they have as data curators. They are familiar with naming standards, the curation process, data management plans and data flows, to name a few".

The involvement of librarians as data curators and data stewards derives from the support needed by researchers in management of research data lifecycle. Data curators provide metadata and ontology support for datasets, provide support and expertise to researchers engaged in research data management planning, data acquisition, data sharing and long-term data stewardship and support deposit and stewardship of datasets in digital repository platforms. They are responsible for organising and integrating data collected from various sources, which involves annotation, cataloguing, indexing, publication and presentation of the data such that the value of the data is maintained over time, and the data remains available for reuse and preservation. They also perform outreach activities focusing on promoting open access and data sharing through data repositories and consulting on metadata standards, data formats and citation standards.

On the other hand, data stewards play an important role to advise, support and train researchers on data life cycle and good data management practices, from initial planning to post-publication. This includes storing, managing and sharing research outputs such as data, images, models, programs and codes. Data stewards also advise and educate researchers on the practices that support open science and reproducibility of research, ethical, policy and legal considerations during data collection, processing and dissemination. They have multiple responsibilities including being held accountable for the quality of the organisation's data assets (in terms of completeness, consistency, uniqueness, validity and accuracy) and the curation process, i.e. data cleansing and digital preservation. Data stewards also work with researchers to develop naming standards, data definitions, and metadata to be used, and provide advice on or assistance with writing research data management plans for the researchers. Data stewards assist researchers in meeting data management and sharing requirements of funders and publishers and work collaboratively with other organisational units supporting research data management.

Because of the diverse roles and responsibilities of data curators and data stewards, possessing essential skill set, which have been the foundation skill that has been practised over the years by librarians, such as data cataloguing and indexing, data sharing and reusing, metadata management, knowledge on taxonomy and ontology and familiarity of subject databases is a fundamental criterion for both professions. However, in the era of digital data and Open Science, librarians' roles are changing, so too are the skill sets. Future librarians who will assume data stewardship roles are expected to be equipped with emerging skills such as knowledge on software and coding, data analytics, data management plan, computational skills, data mining, and subject matter knowledge.

7.2. Challenges – Revolutionising the traditional role of librarians for Open Science

While the core roles of academic librarians have been centred on data curatorship, academic libraries in Malaysia have started incorporating stewardship roles in their daily jobs. For example, librarians from Universiti Teknologi Malaysia (UTM) have started engaging with researchers from the chemical engineering field, who have agreed to deposit raw research data, to work on the data management plan. In Universiti Sains Malaysia (USM), librarians have organised roadshows around the USM campus to raise awareness about data management plan [17]. The librarians have engaged with groups of USM researchers, especially those from the engineering campus, and researchers may request a one-on-one consultation with the librarians on a case-by-case basis. Besides UTM and USM, librarians in Universiti Putra Malaysia (UPM) have also played data stewardship roles in assisting researchers to reuser secondary data [18] to compute new analysis, such as by providing advice on choosing the right software for data analysis.

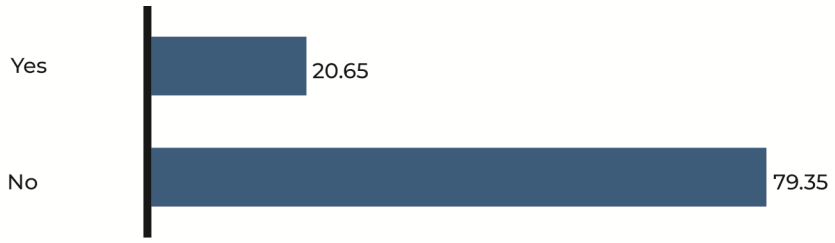


Figure 7.1. Engagement in writing data management plans

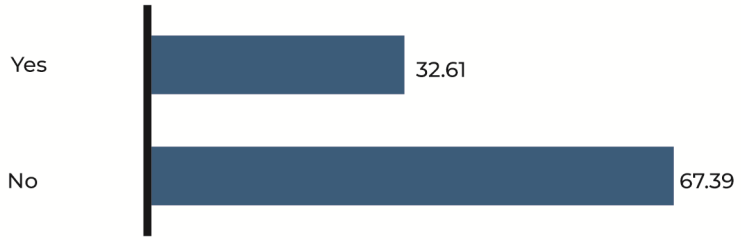


Figure 7.2. Reaching out to researchers or research students to support data management plans

In short, future roles of librarians are envisaged to change to facilitate the implementation of Open Science and FAIR data ecosystem. Librarians are also expected to work together with technical staff, researchers and the broader community to develop and manage several aspects including identifiers, metadata, standards, vocabularies and ontologies. Key roles of librarians may include:

[17] The standard format for data management plan that is currently being used in USM is based on Digital Data Curation (DCC), United Kingdom. More information about DCC can be found in this link: https://www.dcc.ac.uk/about. Source ASM.

[18] The secondary datasets were sourced from the Department of Statistics Malaysia (DOSM). The arrangement was made under a Memorandum of Understanding (MoU) between UPM and DOSM, which allows DOSM's raw data to be shared with UPM and to be reused by UPM researchers. Source: ASM.

Table 7.1: Library support services throughout the entire research process
Source: Adapted from LIBER Open Science Roadmap 2018-2020.

Research process	Support services offered
Planning	1. Develop Research Data Management Plans (RDMP) and assist researchers in implementing the RDMP. 2. Develop and provide tools for FAIR data management. 3. Help researchers to manage their personal identifiers, for example ORCID. 4. Provide information about research funding possibilities.
Supporting	1. Support access to information through portals and databases. 2. Ensure that the library can be used as a general help desk or a one-stop-shop to support researchers in all queries related to Open Science.
Managing	1. Ensure research outputs are interoperable in the sense that the shared research data must be tagged with identifiers and metadata that meet the designated standard. 2. Provide training in managing research datasets, either for usage of programming languages, statistics and high computing power. 3. Develop necessary infrastructure, e.g. ontologies and other tools to describe the content deposited on repositories.
Publishing	1. Encourage researchers and research students to deposit and use institutional repositories for publishing. 2. Provide training in Open Access publishing and raise awareness about requirements for publishers.
Assessing	1. Participate and review the development and adoption of metrics to measure impacts of Open Science.
Reusing	1. Raise awareness about reuse requirements. 2. Promotes the use of licensing agreements such as Creative Commons Licence.

7.3. How can we learn from other Open Science initiatives?

In other regions or countries where Open Science has embarked for quite some time, library units were recognised to have a significant leadership role in the Open Science movement. A case study was described in **Box 7.2**.

Box 7.2. Examples of leadership roles in Open Science manifested by libraries

Case study: Delft University of Technology (abbreviated as TU Delft)

The library unit at TU Delft was expected to provide a robust and high-quality infrastructure support through the creation and development of high-quality advice on research data management plan, and they closely worked together with faculties to give advice and training in developing good data management. The library unit is an integral component of data stewardship as it offer a variety of services to assist with the management of raw and processed data, and coordinated a network of data stewards who were embedded within the faculties. The library unit also served as a certified archival service that offers at least 15 years of long-term curation for research data. In the aspects of data curation and raw and processed data management, the library unit worked closely with the ICT department.

TU Delft Research Data Framework Policy (2018).

In Malaysia, not many librarians have attended a data stewardship course or are affiliated with any local or international network for data stewards. However, there is a great opportunity to train the librarians to assume data stewardship roles for Open Science practices. For that purpose, it is important it important for the librarians to receive training in many aspects such as research skills, knowledge on metadata and, legal knowledge on Open Science. More importantly, data stewardship training should also provide librarians the exposure and familiarity with specific research disciplines that produce major datasets in repositories. As for now, familiarity with research disciplines are skewed towards social science and humanities, with a very small number of librarians who are familiar with science subjects.

Figure 7.3. Linked to any local or international network for data stewards

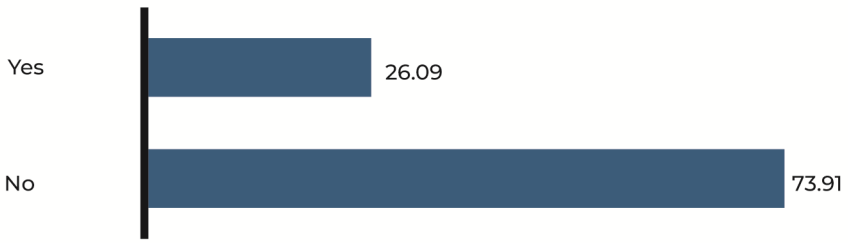


Figure 7.4. Attendance at a data stewardship course

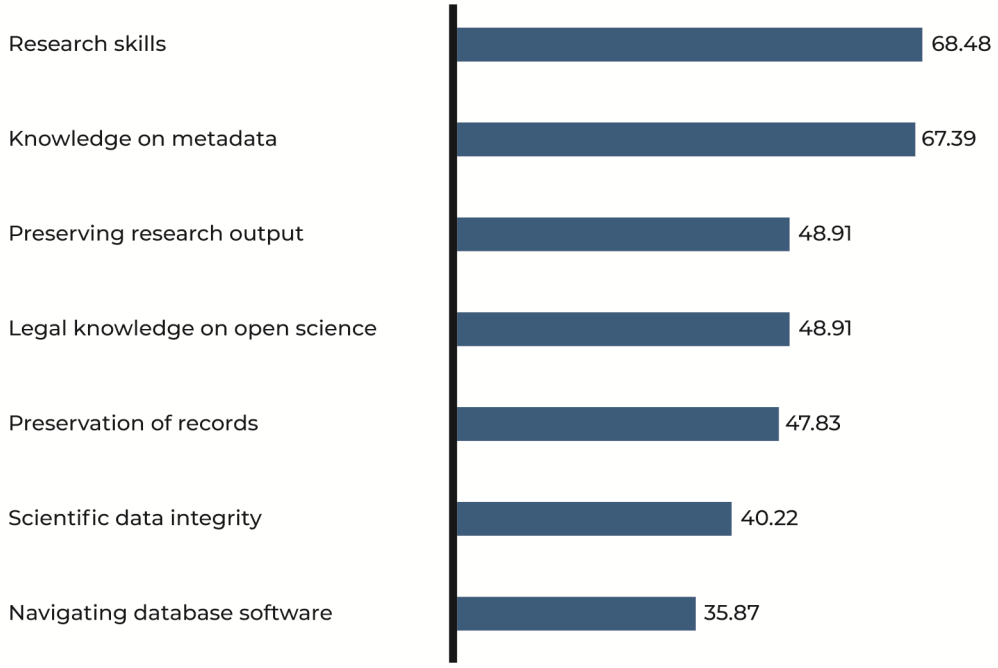


Figure 7.5. Main interests in data stewardship training sessions. You may choose more than 1 option

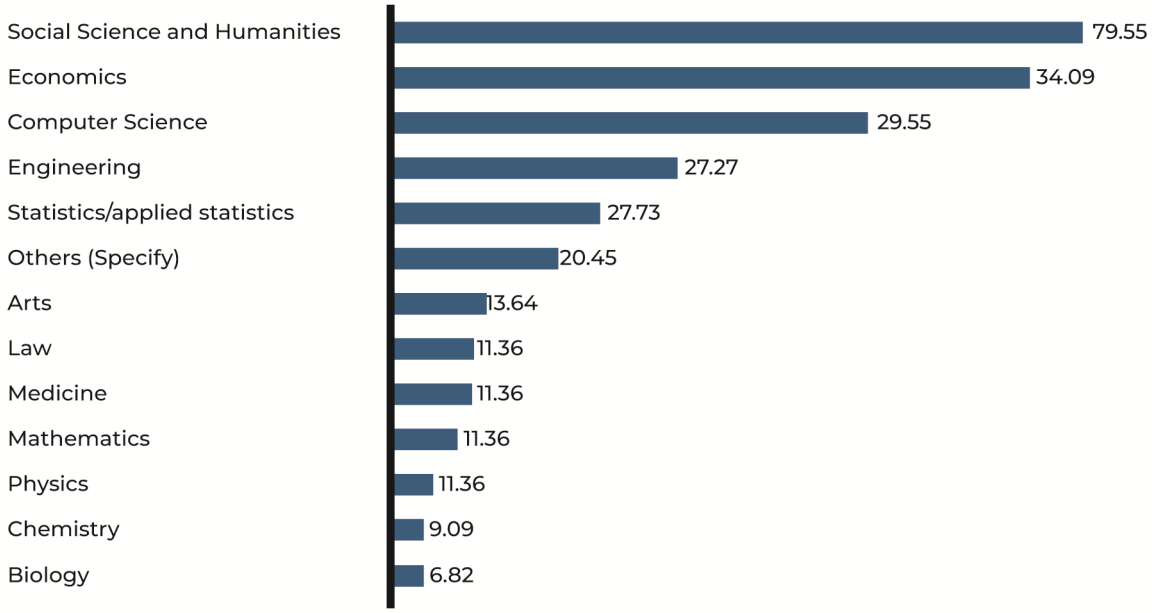


Figure 7.6. Research disciplines that you are familiar with as a data librarian

7.4. Key takeaways

i. The role of data curators and data stewards is very important in Open Science as they provide support to researchers in various aspects especially the management of research data lifecycle.
ii. A majority of librarians in Research Universities are still performing their traditional roles as a librarian with few of them having started engaging in performing duties as a data steward.
iii. Training for data stewardship is needed to equip librarians with the essential skills supporting Open Science.
iv. A strong support from the top management is necessary to drive the transformative role of librarians as data stewards.

CHAPTER 8

PERSPECTIVES ON OPEN SCIENCE FROM INSTITUTIONAL REPOSITORIES

Institutional repository is defined as a library of digital objects and associated metadata from a single institution. Generally, repositories can hold structured, unstructured or even semi structured data. Structured data is data that adheres to a pre-defined data model and is therefore straightforward to analyse. Structured data, which usually is processed data, resides in relational databases or data warehouses and are easily searchable. On the other hand, unstructured data does not have a predefined data model or is not organised in a pre-defined manner. Text, raw images, audio and video files are some examples of unstructured data which reside in application or noSQL databases.

Recent interviews with Head of Data Centres or representatives from the five Research Universities in Malaysia have established well-defined institutional repositories that are maintained by Library units. A majority of institutional repositories in research universities were developed using EPrints open-source repository software. The real strength of using EPrints lies in the flexibility and easiness of use for data depositors and system administrators. The process of depositing materials into EPrints is simple and easy as data depositors only have to key in metadata, for example the type of document, title or, name of authors using the form available on the website. Furthermore, knowledge about HTML/XML is not required for the purpose of data deposition. EPrints also allow depositors to submit suggestions to administrators if they would like to update or delete deposited materials. Lastly, EPrints offers some flexibility in metadata customisation that allow system administrators to tailor the standard according to their needs.

Table 8.1. Summary of institutional repositories at the five Research Universities in Malaysia

Case Study 1. Universiti Malaya
1. Number of Repositories: 2 2. List of Repositories: a. UM Research repository (contains published and unpublished research work produced by the UM researchers), b. UM Students repository (contains scholarly exercises, dissertations and theses produced by students and staff of Universiti Malaya) Note: UM Common repository (contains printed items, old copies of items such as gazettes, magazines, journals and books) and UM Memory (contains pictures, books, unpublished papers) are digital libraries of UM 3. Accessibility of Repositories: All are accessible by the public except contents dealing with copyright issues, embargoed or request by author not to make it available in full text
Case Study 2. Universiti Putra Malaysia
1. Number of Repositories: 5 2. List of Repositories: a. UPM Institutional Repository (UPM-IR) (contains postgraduate theses – accessible only to a certain extent i.e. abstract, Chapter 1 and references; project reports, journal articles, journals and bulletins, conference papers, books, book chapters, monographs, UPM news, articles from newspapers, patents and inaugural lectures), b. MyAGRIC (contains information of Malaysian agricultural science and technology via links and union catalogues. Examples of content includes acts, annual reports, bibliographies, bulletins, statistics and speeches) c. AGRIS (bibliographic database for agricultural sciences and technology consisting of publications from Malaysia)

d. Memory@Serdang (contains a collection of images, videos and audios that have academic and historical values at UPM; Open Access)

e. UPM Etheses (contains Dissertations & Theses)

*Note: UPM also established a research data repository but the accessibility levels of the repository is yet to be decided.

3. Accessibility of Repositories: UPM Institutional Repository (UPM-IR), MyAGRIC, AGRIS and Memory@Serdang are open to public to access while UPM Etheses is private access whereby it is only allowed for UPM staffs and students who have user accounts and are authorised to login to the repository.

Case Study 3. Universiti Kebangsaan Malaysia

1. Number of Repositories: 3
2. List of Repositories:

   a. e-repositori penerbitan (contains articles by UKM authors, Policy papers, Slides; 400,000 items deposited)

   b. UKM Journal Article repository (Articles developed locally and is published in UKM journals; Open Access), and

   c. Research and Learning repository (Theses, Books, Exam Papers, News articles; Seminar papers, Legislative documents)

3. Accessibility of Repositories: Both *e-repositori penerbitan* and the Learning and Research repository are semi-open. UKM Journal Article repository offers full text accessibility to the public

Case Study 4. Universiti Sains Malaysia

1. Number of Repositories: 6
2. List of Repositories:

   a. Repository@USM (contains Open Access scholarly publications articles, academic magazines, books, theses, examination papers, research reports, photographs and others and also includes Closed Access materials like thesis, textbooks and lecture notes)

   b. USM Digital Exhibition and Special Collections (USMDESC) and USM Library Light Letters Digital Collection (contains archival records like photos and letters)

   c. eDocs system and Microsoft Office 365 in Knowledge Management Collection (contains ISO documentation)

3. Accessibility of Repositories: Repository@USM provides both open and closed access. Both USMDESC and USM Library Light Letters Digital Collection are accessible by the public. Knowledge Management Collection is a private access repository.

Case Study 5. Universiti Teknologi Malaysia
1. Number of Repositories: 3 2. List of Repositories: a. UTM-Institutional Repository (UTM-IR) (Postgraduate thesis: Table of Content to Chapter Introduction, Journal article, Proceedings, Books, Books chapters, Research monographs and Conference papers), b. Document Management System (DMS) (Newspaper cuttings, theses, exam papers, journal articles, indexed and non-indexed papers, conference papers, project papers, photographs), and c. Research Data Management (RDM) *Note: Under development 3. Accessibility of Repositories: UTM-IR is accessible by the public, DMS is available on private access, exclusively to UTM staff and students. Private access rule also applies to Research Data Management (RDM).

Note: a. Open Access means anyone with public access is allowed to access materials; b. Closed, or Private access means only authorised users are allowed to access materials.

8.1. Detailed descriptions and specifications of selected institutional repositories

8.1.1. Universiti Sains Malaysia

Universiti Sains Malaysia developed their institutional repositories using EPrints (open access), established in 2005, and DSpace (closed access), established in 2016. Both librarians and material owners are responsible for depositing the materials into the repository. A depositor is allowed to deposit 7 items daily, weekly or monthly. Self-deposition is available via CampusOnline and Microsoft Office 365. Open Access materials from Repository@USM are crawled and aggregated by various index systems and use unique numeric persistent identifiers, such as:

1. Google Scholar
2. United States of America Library Company index system
   a. OCLC OAIster (http://oaister.worldcat.org/)
3. United Kingdom Higher Education Institutions (HEIs) - JISC index system
   a. CORE (https://core.ac.uk/)
   b. Registry of Open Access Repositories (http://roar.eprints.org/)
   c. OpenDOAR (https://v2.sherpa.ac.uk/opendoar/)
4. Germany HEI project index system
   a. BASE (Bielefeld Academic Search Engine) (https://www.base-search.net/)
5. ASEAN HEIs and academic library projects index system
   a. ASEAN Research repository (http://arr.in.th/)
   b. AUNILO Institutional Repository Discovery Service (http://aunilo.uum.edu.my/)
6. Malaysia Academic Library index system
   a. Malaysian Academic Library Institutional Repository (http://malrep.uum.edu.my/rep/)

During the establishment of repositories back in 2005, USM focused only on using standard metadata like Open-Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) and Dublin Core. However, with advancement in research databases like Google Scholar, Scopus and ORCID, the metadata standard in USM will be developed as required by Schema.org, EUROCRIS-CERIF and the Malaysian Research and Development Classification System (MRDCS).

Compliance of institutional repositories follow many policies and regulations including Act 629 – National Archives Act 2003, Act 709 – Personal Data Protection Act 2010 (PDPA), Act 332 – Copyright Act 1987, Dasar Pengurusan Rekod dan Arkib Elektronik (National Archive of Malaysia), Act 80 – National Library Act 1987, Dasar Pendokumentasian Bahan Perpustakaan (National Library of Malaysia), Research Policy and Ethics Universiti Sains Malaysia, Thesis Publishing Policy, ICT policies and guidelines and Creative Common Licence.

All historical documents of USM are kept permanently while documents related to USM staff are stored and retained based on the staff's availability. If the staff has retired, their data will be transferred to other storage but if the staff have passed away, then the data will be disposed of after a certain period of time. Some extremely important documents such as minutes of meetings and reports are retained for 50-60 years while Senate meeting documents are kept permanently.

The backup of the deposited data is done separately based on the front and back end. ICT handles the infrastructure backup while backup for all the frontend tasks are performed by the library. The backup process is performed daily and USM library has five cloud storages for data archival.

The internet bandwidth of the organisation network to the internet Service Provider (ISP) is 2 Gigabytes. MyREN, YES, TMNET and JARING are some of the ISPs and they have their specific territories to provide access.

8.1.2. Universiti Teknologi Malaysia

The UTM institutional repository was developed using EPrints (version 3.3.15) in 2007 and it is open for public use. Their database driver is MySQL (DBI 1.631; DBD: MySQL 4.028) and the database server version is MySQL 5.6.23-log. UTM uses Apache as their web server and CentOS for their open source operational system. UTM library also uses DSpace but it is not used for a wide spectrum as it is meant for internal use only.

UTM library uses Dublin Core and Machine-Readable Cataloguing Record (MARC 21) as metadata standards to organise and preserve the stored data. Bibliographic data which differs from other format data and is available in the library data is called the MARC data format.

Usually, all materials are stored permanently by the UTM library. The materials will undergo a backup process every 12 hours as per the SANDAR backup schedule. There are master and slave databases for data archival purposes. The backup server is located at the data centre and at the library. IT officers from the data centre and library are responsible for performing the backup process. All the backup materials are kept for one month and subsequent backup processes will be performed if there are no issues with previous backup materials. The internet bandwidth of the organisation network to the internet Service Provider (ISP) is 10 Gigabytes. The specification of the database server is 32 Gigabytes of RAM and 500 Gigabytes of storage.

8.1.3. Universiti Putra Malaysia (UPM)

The UPM institutional repository, UPM ETheses and MyAGRIC were developed using the EPrints software, AGRIS using DSpace while Memory@Serdang uses ResourceSpace. Data for UPM IR is obtained from PRIMS portal, UPM ETheses from thesis collection, UPM MyAGRIC and AGRIS from the internet while Memory@Serdang from archival collection. All digitised data were tagged with DOI as the persistent identifier. Information on AGRIS is monitored by the Food Agriculture Organisation and will be extracted and indexed into the FAO database.

The UPM library has several divisions that manage the repositories. Firstly, its Digital Resource Management Division officers, who are responsible for handling data deposition in UPM IR and UPM eTheses. Secondly are Malaysian Agriculture Resource Division Officers, who are responsible for UPM MyAGRIC and AGRIS and finally the Media, Archive and Preservation Division who takes charge over Memory@Serdang.

The metadata standard used to organise and preserve the stored data is the Dublin Core. Backup for the deposited data is performed everyday by the System and Information Technology division of the library. All the servers are located in the UPM Data Centre and they are in charge of managing all the data and the internet bandwidth of the library. The UPM Data Centre is also responsible for monitoring the performance of the server daily.

8.1.4. Big Data Repository at Universiti Putra Malaysia

UPM has developed a research data repository to support the Open Data Science initiative. The repository uses Cloudera Hadoop as the Big Data platform. The main focus of this repository is to store raw research data from research projects, either from internal or external sources and to allow analytics activities to be performed on these data. Since the research data policy is not in place yet, accessibility to this repository has not been decided by UPM as to whether it can be accessed by the public or with restrictions.

For research projects funded by UPM, the researchers are requested to deposit their research data into the repository. Researchers can decide whether the data can be shared publicly or privately but a designated Committee at the Research Management Centre will have the final decision on data accessibility. Nonetheless, discussions are still on-going as to whether the committee will continue their role to filter types of data that can be shared on this repository.

As of now, the repository receives raw data from a server hosted at UPM's Computer Centre and the repository is currently storing X-ray images created from medical research projects. The repository uses metadata standards such as titles, abstracts, keywords and domains for data retrieval. The repository currently runs at a storage capacity of 40 terabytes. The duration of data storage in the repository depends on the type of data. Some analytical data are kept longer for meaningful trend analysis (e.g. financial data is usually kept around seven to ten years in the system). If researchers would like to deposit newer versions of datasets, the repository will retain the older versions.

To deposit data from the library into the repository, the librarians are responsible for doing the data depositing task, while data from research projects will be deposited by the principal investigators. However, principal investigators can request to assign project members like Research Assistant or students to deposit the data.

The Internet Service Providers (ISPs) in UPM are Maxis, Time and MYREN and the Internet bandwidth is provided by these ISPs. The Internet bandwidth for UPM allocated by the ISPs is 4.8 gigabytes. Data backup operation is performed daily, weekly, monthly and yearly while maintenance of the server is done quarterly and yearly. Even though the entire process of backup and maintenance is automatic, the database administrator is responsible for monitoring the operation.

8.1.5. Data analytics at Universiti Kebangsaan Malaysia (UKM)

The Information and Communication Technology (ICT) centre in UKM has a data centre which uses an SQL server to perform data science analysis. The structured data is gathered from their informix database and data analytics is performed based on requests from UKM staff. These data are retrieved from UKM management, UKM business and data from Relational Database Management System (RDBMS). The SQL server uses the Extract, Transform, Load (ETL) feature to filter the transferred data to ensure the data is clean before providing it to users such as UKM staffs or the Ministry of Education. Data analytics were empowered by features in Power BI. To facilitate data analytics, UKM's ICT centre developed a dashboard for users to view data analytics from the SQL server. The internet bandwidth for UKM's ICT centre (Bangi Campus) is 8 Gigabytes. The backup process is performed by the Centre following SPKP UKM guideline and Information Security Management System (ISMS).

8.1.6. Universiti Malaya

Both UM Research Repositories and UM Students, as well as the digital library UM Common Repository use EPrints. An exception applies to the digital library UM Memory which uses Omeka. All materials which are deposited in the repositories will be kept forever. The metadata standards employed by the repositories follow Dublin Core standards. The database drivers used are MySQL and Postgres, and Apache is used as the web server.

The accessibility of the deposited materials is decided by data depositors or data originators. The data deposition process is as follow:

1. Data depositor fills up an online form, including some basic metadata
2. For student repository, students have to fill up forms manually at their faculty. The faculty will pass over the materials (i.e., saved in pendrive) to the library
3. Librarian will verify the form and metadata given before depositing the material. They will make necessary amendments

Besides storing published research works, the UM Research Repository also stores raw research data. Deposition of raw research data is voluntary and is not made compulsory by an institutional policy. The policy, which is the UM Research Data Management Policy, is being developed and is waiting for final approval from the top management.

As for the backup process, both Pusat Teknologi Maklumat (PTM) UM and UM library are involved in this task. PTM perform the backup process daily according to their schedule while UM Library performs the backup process less frequently than PTM. The internet bandwidth on the campus, according to PTM, is 18Gbps.

8.2. Challenges – Interoperability of Metadata Standards

Metadata has two fundamental components: semantics and content. Semantics refers to the definitions of the meanings of the elements and their refinements while content refers to declarations or instructions of

what and how values should be assigned to the elements[19]. There are many types of metadata available such as descriptive metadata, technical metadata, preservation metadata, rights metadata, structural metadata and mark-up languages, of which all of them have specific purposes. Interoperability among metadata standards that support data discovery, selection, access and use, will enable not only data integration, searches and cross-domain metadata harvesting, but also facilitates data-driven and automated analysis.

Metadata standards are an important tool to enable seamless integration and semantic linkages of data from single-and cross-disciplines. However, metadata standards have long been established tailoring to specific disciplines and datasets or applications, with little emphasis being given to support data integration and reuse of cross-disciplines. Additional investments and work are needed to establish more common metadata standards embodied in the metadata architecture that permits communication and exchange of metadata between repositories.

Mechanisms involved in interoperability of metadata are file exchange formats and API approaches. For the file exchange approach, the metadata must be written in a file of a standardised format that can be imported by other repositories. Although this file exchange approach provides an asynchronous access to external metadata, the metadata is only current at the time of import. Moreover, the method for translating between a tool's proprietary metadata format and the file format is hardcoded in the tool and it needs to be adapted when the file exchange format evolves. The second mechanism of metadata interoperability is using standard data-centric Application Platform Interfaces (APIs), that will enable interoperable data exchange across multiple repository platforms. While this option allows seamless integration between repositories, a detailed cost analysis is needed to understand the financial implications involved.

8.3. Challenges – Trustworthiness of Repository

Foreseeing the future change in current science and data sharing practices, as the world continues to shift towards Open Science, it is important for Open Science champions to recognise the criteria of a trustworthy repository to bring in confidence of researchers to deposit and use shared data. A trusted digital repository is viewed as a repository that is reliable and capable if it can manage stored data. Underpinning this notion, several attributes of a trusted and reliable digital repository are identified as: (1) Compliance with the Reference Model for an Open Archival Information System, (2) Financial Sustainability, (3) Administrative responsibility, (4) Organisational viability, (5) Technological and procedural suitability, (6) System security, and (7) Procedural accountability.

In recent years, discussions and communications about a trusted digital repository were guided by the TRUST principle, which described the main criteria for a digital repository to be considered trustworthy. The criteria are Transparency, Responsibility, User Community, Sustainability, and Technology. To establish a trusted repository that attains the principles of TRUST, data repositories should adopt a certification standard such as CoreTrustSeal Trustworthy Data Repository Requirements. The Australian Data Archive through Australia National Data Services has recently been certified as a trusted digital repository under the CoreTrustSeal certification.

8.4. Key takeaway messages

i. Most institutional repositories in the five Research Universities in Malaysia use Dublin Core as the metadata standard.
ii. Most institutional repositories house structured data and the establishment of research data repositories is quite new in these Research Universities.
iii. From consultations and engagements with researchers and other stakeholders in Malaysia, they emphasise the importance of establishing a trusted repository to encourage data sharing in the country.

[19] Interoperability can be defined as the ability of multiple systems with different hardware or software platforms, data structures and interfaces to exchange data with minimal loss of content and functionality (Source: National Information Standards Organization, 2004, Understanding metadata).

CHAPTER 9

CONCLUSION AND RECOMMENDATIONS

The mandate and strategic thrusts embodied by the National Policy on Science Technology and Innovation (2021 – 2030) are pushing towards the full implementation of Open Science and data sharing practices in Malaysia. Realising this goal will require holistic and joint interventions by multi-stakeholders such as researchers, top management universities, research funder organisations, publishers, government agencies, industries and learned societies, aimed to nurture and encourage a research data sharing ecosystem in the country.

This document was prepared to: (1) map existing Open Science activities and players in Malaysia, (2) identify factors enabling and hindering Open Science Adoption, (3) gauge the level of readiness of local institutions to adopt Open Science and (4) undertake benchmarking exercises to identify global indicators to measure Open Science readiness. In particular, this document provided insights into the current landscape of Open Science in Malaysia based on six indicators: (1) Relevant acts, policies and guidelines locally and internationally, (2) Number of repositories, (3) Skill capacity, (4) Infrastructure capacity, (5) Awareness, and (6) Current Open Science activities from perspectives of five target groups, namely (1) Deputy Vice Chancellors (Research & Innovation), (2) chief librarians and librarians, (3) Head of data centres or representatives of institutional repositories, and (4) junior and (5) senior researchers. The inputs wase validated with engagements with a broader set of stakeholders including Institutes of Higher Learnings, Public and Private Research Institutes, government agencies and ministries, legal units, libraries, research funders, and industries. Based on the consolidated inputs solicited from those involved in the landscape study and engagements, this document further sets out **9 recommendations** to support a rapid and effective implementation of Open Science in Malaysia through coordinated actions within and among these stakeholders. The proposed implementation actions shall provide clear directions for future activities of these stakeholders for areas that they may be lacking towards achieving the overarching goal of Open Science. These recommendations have been reviewed by the members of Malaysia Open Science Alliance and members of Malaysia Open Science Alliance Working Groups on Guidelines and were informed by engagements and consultations with relevant stakeholders.

9.1. Recommendation 1: The provision of National Open Science Policy

In order to bring about a holistic implementation of Open Science in Malaysia, such that they are considered as an integral element among research communities in the country, it is crucial to develop a national Open Science policy that will streamline data sharing policies among institutes of higher learnings, research institutes, government agencies and non-government organisations. The Policy shall be led by the Ministry of Science, Technology and Innovation (MOSTI) and is to be implemented in a top-down coordination approach.

The Policy should also:

i. Be supporting vision, mission and strategic thrusts related to data sharing as outlined in the National Policy on Science, Technology & Innovation (2021-2030).
ii. Be established and embody the Malaysia Open Science Platform guidelines that follow the FAIR (Findable, Accessible, Interoperable and Reusable) principle.
iii. Leverage and balance the power of secrecy acts in Malaysia and the interest of data sharing for Open Science.
iv. Contain a powerful policy statement that provides assurance and comfort to data originators.
v. Recognise and mandate the implementation of Research Data Management Plan practices across all Research Universities, other Institute of Higher Learnings and research organisations.
vi. Recognise the roles of data stewards and data curators, as well as research funding organisations for Open Science.
vii. Recognise the importance of a trusted and interoperable research data repository infrastructure.
viii. Following the provision of National Open Science Policy, each research university, Institute of Higher Learnings (IHLs) and research institutions must develop and approve its own Open Science related policies, such as data sharing and research data management policy to ensure an inclusive implementation of Open Science.

9.2. Recommendation 2: Guidelines for Implementing Open Science and MOSP

Best practice guidelines for applying Open Science should be established to facilitate Open Science and data sharing practices across different levels of stakeholders. The guidelines should be embodied by the National Open Science Policy and must address the following:

i. Outline data governance for Malaysia Open Science Platform initiative.
ii. Legal and ethical requirements, and codes of conduct, associated with data sharing, protection, retention and disposal issues, which may have different implications, depending on types and sensitivity levels of created data. Compliance with specific legislations and guidelines may be due to data arising from human research, data containing information intended for commercialisation, data that concerns national security matters, and data created from preclinical and clinical trials, to name a few.
iii. Provide details on core requirements for research data management plan, which can be used as a template for each research university, Institutes of Higher Learnings, other research institutions, government agencies and non-government organisations.
iv. Define features of the platform architecture of the repository, including a description for research institutions that do not have their own repository infrastructure as to where the data should be stored.
v. Define roles and responsibilities of data owners, data contributors, data stewards, data curators, data custodians and data users of the Open Science platform.

In addition, institutional policies, guidelines and strategies should be linked to the National Open Science Policy to ensure that institutional infrastructures are established and operated meeting the highest possible standards:

i. The Policy must clearly delineate roles and responsibilities for each stakeholder within the institution for all collaborative and non-collaborative publicly funded research projects.
ii. The Policy must recognise support systems for the development of Research Data Management Plan from the aspects of Policy, training, assistance and data protection.

9.3. Recommendation 3: Empowering the Role of the Funding Body for Open Science

To ensure a successful implementation of Open Science and data sharing practices, public funding bodies must be committed to promoting the principles of Open Science by implementing appropriate policies that encourage researchers to share their research datasets and outputs, as well as associated metadata as soon as possible. Evidence of planning to practice Open Science as written in a grant application should be considered as merits in the review of funding applications. Following are several suggestions for additional criteria for grant application reviews for public funding bodies in Malaysia:

i. A submitted grant application must include a detailed data management plan (DMP) that specifies how research data is managed throughout the data lifecycle. A DMP must describe the process of a researcher on how to acquire or generate data during the course of a research project, specifically how the researcher will manage, describe, analyse, and store these data, and the mechanism that will be used at the end of the research project to share and preserve research data.
ii. A submitted grant application must agree to make metadata and research data openly available at no costs as soon as possible after research data has been published. However, there are permissible circumstances that research data cannot be made openly available. For such cases, the metadata must be stored and be made available on the Malaysia Open Science Platform. The bibliographic metadata must follow a standard format and must include the grant number to ensure that the grant funding is acknowledged.
iii. A submitted grant application must include a publication plan that spells out whether the proposed research article to be published will be following either the green, gold or hybrid open access. In cases where the project will follow the gold or hybrid publishing routes, grant funding can be allocated to cover related publication costs.

In relation to the proposed criteria included in the grant review process, guidelines for funding bodies should incorporate the following provisions in its Policy and Guidelines:

i. Open Science elements should be included as a criterion in the grant review process. Such inclusion shall consider Open Science related policy and Open Access publications, such as Research Data Management Plan and publishing in Open Access journals.
ii. Specific provisions on research data sharing and reuse, including mandatory, locus and time of deposit, possible opts-outs for sensitive data, embargo periods, licences and copyrights, data archiving and long-term preservations etc. Careful considerations must be made when specifying the time period for funded publications and must be made freely accessible as the funder requirement must tally with the embargo period posed by publishers.
iii. Recommendations for the use of repositories that meet FAIR principles and are accompanied with guidance to grantees for data deposition.
iv. Some allocations in a grant application to bear costs for Open Access publication fees.
v. Some allocations in a grant application, for example under research costs vote, to bear costs associated with data sharing and re-use practices. This could be a possible option for funding bodies although the cost for maintaining and operating research data repositories will be borne as overhead costs at research institutions.

Additionally, it is equally important for public funding bodies in Malaysia to establish a monitoring and policy compliance mechanism and periodically review, revise and update its Policy and Guidelines and assess the overall program impacts.

9.4. Recommendation 4: Building a trusted and interoperable research data sharing platform

Acknowledging that there is a wide spectrum of datasets produced in Malaysia, it is recommended that repositories should be distributed at institutions and the role of the Malaysia Open Science Platform should focus on capturing and harvesting all metadata to ensure that the deposited data are publicly available and preserved. The Malaysia Open Science Platform should play its role to preserve and publish all metadata of research datasets and publications, irrespective of levels of sensitivity of created data.

Furthermore, MOSP and an institutional repository must:

i. Be user-friendly in terms of user interface, ease of use and navigation support.
ii. Be scalable and interoperable with other repositories across research universities, research institutes and government agencies. Additionally, the repositories must also be interoperable with international domain-specific repositories and other global repositories.
iii. Be equipped with appropriate data request handling process and metadata to enable the finding of data and cross-references with other publications. Metadata formats should follow standards that are broadly accepted, machine-retrievable and should be tagged to all types of data, even if the data is non-published, protected, retracted or deleted.
iv. Be supported with persistent and unique identifiers that allow data discovery, citation and retrieval.
v. Be able to ensure data retrieval, preservation and continued availability, and access to the deposited data and metadata under well-specified conditions. Terms for access and use of certain types of data including sensitive data must be highlighted clearly.
vi. Be able to ensure that deposited data is authentic and has integrity.
vii. Provide information about licensing and permissions for data re-use and ensure confidentiality and respect rights of data subjects and data originators.
viii. Be supported with a robust and transparent policy that details the mission and scope of the repository, governance aspects, financial sustainability, and retention period. The repository must also be equipped with a contingency plan that describes procedures to preserve data and metadata in the event of repository closure, for example, easy extraction and transfer of data and metadata to another repository.
ix. Adopt and be certified with a certification standard such as CoreTrustSeal Trustworthy Data Repository Requirements.

In addition, it is recommended for the repository development to consider data residency issues and thus, the repository shall be built on a hybrid cloud storage platform with on-premise infrastructure.

9.5. Recommendation 5: Identifying funding streams to sustain MOSP operation

With the fast paced, growing volume production of research data, Malaysia Open Science Platform needs to be ready financially to ensure that operation of the data sharing platform is sustainable in the long run. Several possible financial streams proposed include:

i. Data sharing and management costs should be included by research funding organisations as a part of project funding.
ii. Malaysia Open Science Platform to be supported by top-down funding, which is the Malaysian government. To convince the government, the projected Return on Investment (ROI) should clearly illustrate national saving and improved returns within the context of Gross Expenditure of Research & Development (GERD) and societal and economic benefits.
iii. Monetising non-sensitive research data, which can be used for paid, profit-driven data analytics service.

9.6. Recommendation 6: Reform existing academic rewards system to incentivise Open Science and data sharing practices

Research Universities and other Institute of Higher Learnings (IHLs) should explicitly support and reward the effort to facilitate the shift in Open Science culture. Traditional methods of assessing security tenure and career promotions must change for Open Science to thrive. Following are implementation actions to be undertaken by universities and other Institute of Higher Learnings (IHLs):

i. Universities and other IHLs should move towards evaluating published research data, preprints and other forms of research products in addition to published articles as a form of evaluating security tenure and career promotions. Measurements can be in the form of data attribution or data citation systems.
ii. Universities and other IHLs, together with relevant ministries should work together in developing new approaches to assess impacts of shared research data by developing and assessing against relevant metrics.

Researchers should also make use of opportunities to deposit their research datasets and publicise their Open Science practices either by updating them on online platforms or including it in their curriculum vitae so they can properly get credited and rewarded.

9.7. Recommendation 7: Training for Open Science knowledge and skills, including data stewardship

The importance of equipping researchers with Open Science knowledge and skills is considered and discussed in several places in the report, particularly Chapter 6 and 7. Research universities in Malaysia have formally and informally educated and trained researchers and librarians on the elements and practices of Open Science and the FAIR principle. The landscape report also identified the crucial role of library units in the promulgation and support of Open Science implementation. Building up on the findings, the following implementation actions are proposed:

i. Universities or other research institutions should provide training for undergraduate, graduate and postgraduate students to give them exposure and inculcate them with essential knowledge and skills about principles and best practices for Open Science including development of a written data management plan and reproducible Science. Open Science design of the course curricula should be tailored to specific research domains of which the students are enrolled in.

i. Universities or other research institutions should provide training for researchers, principal investigators and project main supervisors about principle practices for Open Science and preparation of a written data management plan.
ii. Universities or other research institutions should provide data stewardship training primarily for librarians and research data managers who have prior knowledge and experience in information management to upskill and be trained as data stewards. The training course can also be offered to MSc and PhD graduates who have knowledge in the subject. Courses related to data stewardship should also be offered in the course curricula for a library degree.
iii. To ensure coordinated and synchronised planning and consistent syllabus being taught in data stewardship training. The Ministry of Science, Technology and Innovation (MOSTI) should take the lead to oversee the overall implementation.

9.8. Recommendation 8: Effective communication about Open Science and its incentives

It is important to raise awareness about Open Science and the FAIR principle and highlight the difference between FAIR data and Open Data. Results of Open Science must also be portrayed as the basis that feeds into the development and translation of research discoveries through Open Innovation, which aims to accelerate commercialisation of new technologies through collaborations between institutional spheres in a quadruple helix.

Additionally, incentives of data sharing must be adequately communicated in a manner that resonates well with stakeholders especially Malaysian researchers:

i. Promotions through various platforms such as videos, newspapers, podcasts and workshops should be carried out to bring about better understanding about different forms of Open Science and the Malaysia Open Science Platform initiative.
ii. Content promotion should include several successful examples of Open Science in Malaysia and other countries to convince stakeholders, especially researchers, and to make them understand how the research cycle can happen in an open way.
iii. At the national level, the Ministry of Science, Technology and Innovation (MOSTI) and the Ministry of Higher Education (MOHE) can lead the initiative to promote and raise awareness about Open Science.
iv. At institutional levels, library units can actively contribute to raising awareness of Open Science.

9.9. Recommendation 9: More resource allocation for conducting of research projects

For Open Science and data sharing practices to be well embraced by researchers in Malaysia, it is important to implement strategies that will bring a change in existing data sharing culture among the researchers. A majority of researchers seem to be very reserved and protective over their data leading to a lack of enthusiasm and disinterest in sharing their data. The following suggestions are proposed:

i. More funding for research projects should be allocated to encourage them to produce more quality data, thus more data can be shared.
ii. Provide special or top-up grants for researchers. The grant recipient must share the created data on the Malaysia Open Science Platform.

CHAPTER 10

WAY FORWARD

One of the main objectives of this research is to gauge the level of readiness of local institutions on Open Science. Based on the preliminary institutional review based on interviews and focus group discussions (primary research), the study has identified several issues and concerns that need to be considered in the making of Open Science a reality for the country in terms of its readiness apart from making Open Science more visible in the National Policy for Science, Technology and Innovation (NPSTI) as well asin enhancing the effectiveness of the Malaysian Open Science Platform (MOSP).

10.1. Open Science in the National Policy

a. Open Science has to be more visible in the future National Policy for Science, Technology and Innovation, in which Open Science vision and mission should be mentioned explicitly in key strategies, initiatives and action plans to ensure a holistic implementation of the Malaysia Open Science Platform initiative nationwide.
b. Guidelines on Open Science for Research Performing Organisations (RPOs) and Research Funding Organisations (RFOs).

The aim of the guidelines is to assist Research Performing Organisation (RPOs) and Research Funding Organisations (RFOs) in Open Science/ Open Access which will also become a part of the Toolkit for policy makers on Open Science and Open Access in the context of the Public funded projects which supports Open Access/ Open Data mandates in the country. It will outline the “DOs” and “DON’Ts” as well as describe the responsibilities of the funders and performing organisation and other stakeholders such as fund recipients, researchers, data stewards, data curators and publishers.

10.2. Indicators for Open Science Readiness

There is a need to develop or adopt and adapt the available Open Science Readiness Indicators to monitor the country’s state of readiness on Open Science. For example, Figure 10.1 illustrates one such example of indicators that was developed by RAND Europe and partners in 2017 when it was requested by the European Commission (EC) to develop a monitor that tracks worldwide trends in open science. It is currently being hosted in EC’s website and has provided policy makers and other stakeholders such as researchers, funders, publishers and libraries with access to data related to Open Science practices in Europe and other parts of the world. Hence, using or adopting this model will be useful in benchmarking a country's state of readiness globally.

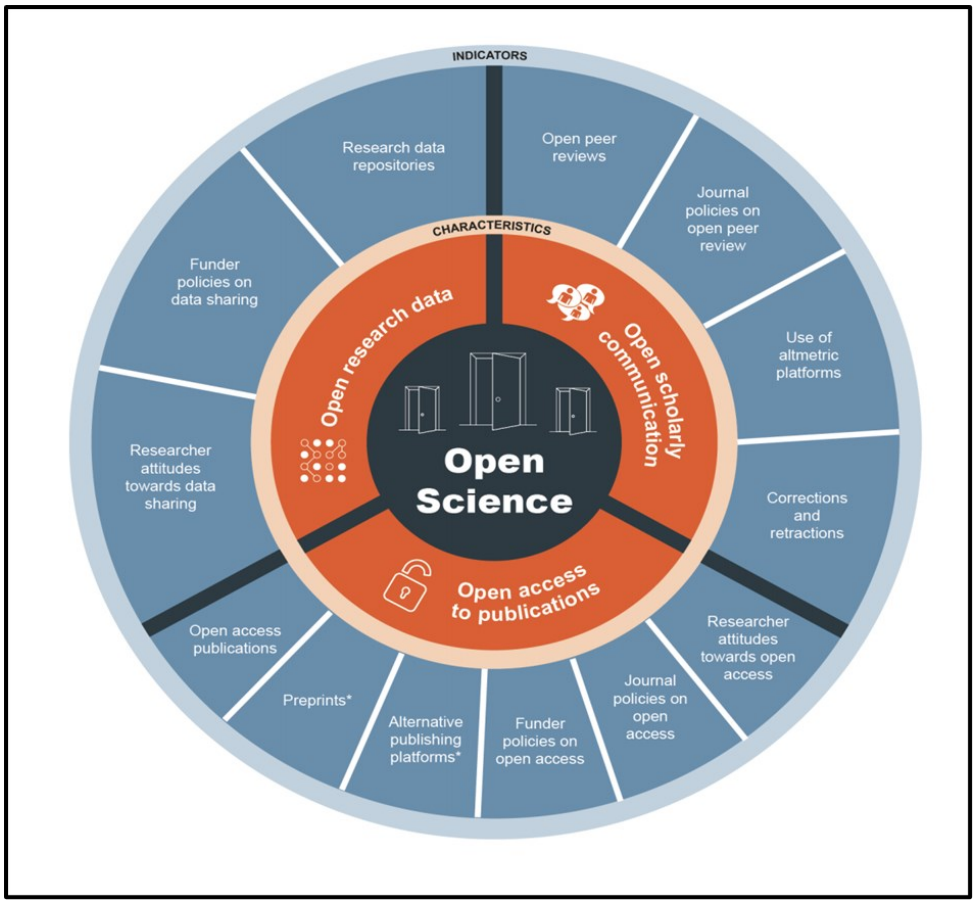


Figure 10.1. Open Science Readiness Indicators
Source: European Commission (2017)

10.3. Strong Funding to support Malaysian Open Science Platform

To develop and to sustain a Malaysian Open Science Platform will require a lot of funding. It cannot be disputed that money is of the essence to ensure the quality and breadth of MOSP. There are basically two reasons for this requirement: the high cost to purchase equipment and tools for MOSP and the cost to maintain or sustain it. The initial seed funding should be provided for by the government while the maintenance and sustainability of the MOSP will depend highly on the business model that will be developed and adopted.

10.4. Competent Human Capital and Top Management

The success or reality of MOSP will highly be dependent on the quality of human capital that will drive it to reach its goal and objectives. There are several important requirements in relation to human capital, namely: data curators and data stewards apart from the top management to be a person who has ample experiences in Open Science as well as having good networking with other stakeholders to promote Open Science and to establish research collaboration. MOSP will also require a blend of in-house experts that can lead in priority areas in Open Science activities such as data curators, data stewards, researchers and research fellows as well as an Advisory Board of which members are to include industry players, policy makers, international experts, academics and representatives from related government agencies and NGOs.

10.5. Right Positioning in the Government Institutional Framework

The present structure of the Malaysian Open Science landscape involves many funding agencies and ministries, public research institutions and institutions of higher learning as well as other stakeholders such as the NGO. The positioning of the new Malaysian Open Science Platform (MOSP) in the government institutional framework is very important to determine its effectiveness in implementing its roles and functions. The position of the MOSP must reflect its importance and authority to be able to give cross ministry directives in order to effectively manage different aspects of Malaysian Open Science. The location of the MOSP in the government structure will also signify its independency which will give the institute flexibility in charting the direction of the Malaysian Open Science Platform in accordance to national needs and global trends. It will also facilitate the funding and disbursement of financial resources in line with the direction set forth by relevant Guidelines for Open Science. Placing the MOSP at an appropriate hierarchy will also elevate the importance of the area that will generate support for public advocacy. It will also indicate the level of credibility of the MOSP to promote international networking and inter-government collaboration in many initiatives related to Open Science.

BIBLIOGRAPHY

GLOSSARY

1. Charalabidis, Y et al. 2018, "The World of Open Data: Concepts, Methods, Tools and Experiences". Cham, Switzerland, Springer International Publishing.
2. European Commission 2016, "Open Innovation, Open Science, Open to the World – A vision for Europe.", pp 33.
3. Heigl, F et al. 2019, "Opinion: Toward an international definition of Citizen Science", Proceedings of the National Academy of Sciences of the United States of America, vol 116, no 17, pp. 8089- 8092.
4. Kohl, HR & Clark, A "NASA Citizen Science App Tackles Mosquito-Borne Disease." NASA, 28 June 2017, www.nasa.gov/feature/nasa-citizen-science-app-tackles-mosquito-borne-disease/.
5. "Metadata." Australia National Data Service, https://www.ands.org.au/working-with-data/metadata
6. "OECD Glossary of Statistical Terms - Metadata Definition." OECD, stats.oecd.org/glossary/detail.asp?ID=5136.
7. "OpenUP Hub - Polymath Project." OpenUP Hub, www.openuphub.eu/component/k2/item/603-polymath-project.
8. "Open Definition 2.1 - Open Definition - Defining Open in Open Data, Open Content and Open Knowledge." Opendefinition.org, opendefinition.org/okd/
9. "Read the Budapest Open Access Initiative." 2002, https://www.budapestopenaccessinitiative.org/read
10. "The Annotated 8 principles of Open Government Data.", Opengovdata.org, https://opengovdata.org/
11. "The Open Source Definition." Open Source Initiative, https://opensource.org/osd
12. UNESCO 2020, "Towards a UNESCO Recommendation on Open Science: Building a Global Consensus on Open Science.", https://en.unesco.org/sites/default/files/open_science_brochure_en.pdf
13. University of Stanford Library. "About Data Management Plans.", https://library.stanford.edu/research/data-management-services/data-management-plans#:~:text=A%20data%20management%20plan%20
14. "What is Open Access" UNESCO, https://en.unesco.org/open-access/what-open-access
15. Wilkinson, Mark D., et al. "The FAIR Guiding Principles for Scientific Data Management and Stewardship." Scientific Data, vol. 3, no. 1, 15 Mar. 2016, 10.1038/sdata.2016.18.
16. Wilson, John-James, et al. "Citizen Science: The First Peninsular Malaysia Butterfly Count." Biodiversity Data Journal, vol. 3, 11 Dec. 2015, p. e7159, 10.3897/bdj.3.e7159.

CHAPTER 1

1. "Benefits of Open Science.", 2015, OECD, https://www.fosteropenscience.eu/content/what-are-benefits-open-science
2. European Commission 2016, "Open Innovation, Open Science, Open to the World – A vision for Europe.", pp 33.
3. FOSTER. "Open Science.", https://www.fosteropenscience.eu/taxonomy/term/100
4. OECD 2015, "Making Open Science a Reality." OECD Science, Technology and Industry Policy Papers. No. 25. Paris, OEDC Publishing.
5. RAND Corporation. "Monitoring Open Science Trends in Europe.", https://www.rand.org/randeurope/research/projects/open-science-monitor.html

CHAPTER 3

1. Desjardins, J 2018, "Visualizing How Much Countries Spend on R&D." Visual Capitalist. https://www.visualcapitalist.com/money-country-puts-r-d/
2. "Find out How Much Your Country Spends on Research & Development." 2018, HowMuch, howmuch.net/articles/research-development-spending-by-country.
3. Malaysian Science and Technology Information Centre (MASTIC) 2020, "National Bibliometric Study 2001 – 2017".
4. Malaysian Science and Technology Information Centre (MASTIC) 2020, "National Survey of Innovation 2018".
5. Malaysian Science and Technology Information Centre (MASTIC) 2020, "National Research and Development Survey 2019".
6. Ministry of Science, Technology and Innovation (MOSTI) 2020 "National Science, Technology and Policy 2021 – 2030".
7. Model Policy in on Open Science for Research Funding Organisations (RFOs), OpenAIRE, https://www.openaire.eu/model-policy-on-open-science-for-research-funding-organisations
8. National Science Foundation 2015, "The Public Access Plan, Today's Data, Tomorrow's Discoveries." https://www.nsf.gov/pubs/2015/nsf15052/nsf15052.pdf
9. Open Science: Open Access Publishing and Open Data. (Online) Available from https://www.aka.fi/en/funding/apply-for-funding/az-index-of-application-guidelines/open-science/
10. Organization for Economic Co-operation and Development (OECD) 2015, "Chapter 7: Legal aspects of Open Access to publicly funded research", Enquiries into Intellectual Property's Economic Impact.
11. "Read the Budapest Open Access Initiative." 2002, https://www.budapestopenaccessinitiative.org/read
12. "What is Open Access." Open Access Netherlands. https://www.openaccess.nl/en/what-is-open-access
13. World Intellectual Property Innovation 2019, https://www.wipo.int/export/sites/www/global_innovation_index/pdf/gii19_leaders_infographic.pdf

CHAPTER 4

1. National Health and Medical Research Council (NHMRC) 2018, Australian Code for the Responsible Conduct of Research
2. Australia National Data Service (Online) Available from https://www.ands.org.au/ and https://projects.ands.org.au/index.html
3. Australian Research Data Commons (Online) Available from https://ardc.edu.au/about_us/ and https://www.ands-nectar-rds.org.au
4. Committee on Data of the International Science Council (Online) Available from https://codata.org/
5. Hodson, S & Molloy, L 2014, "Current Best Practice for Research Data Management Policies. A Report."
6. Data Citation Synthesis Group: Joint Declaration of Data Citation Principles. Martone M. (ed.) San Diego CA: FORCE11; 2014. (Online) Available from https://doi.org/10.25490/a97f-egyk
7. RDA-CODATA Legal Interoperability Interest Group. (2016, October 20). Legal Interoperability of Research Data: Principles and Implementation Guidelines.

8. Grabus, S. & Greenberg, J 2019, “The Landscape of Rights and Licensing Initiatives for Data Sharing.” Data Science Journal, 18(1), p.29
9. African Open Science Platform (Online) Available from https://africanopenscience.org.za
10. The Dutch Research Council (NWO) (Online) Available from https://www.nwo.nl/en/policies/open+science
11. The National Programme Open Science (Online) Available from https://www.openscience.nl/en
12. The Netherlands: National Open Access Desk (Online) Available from https://www.openaire.eu/os-netherlands
13. European Open Science Cloud (EOSC) (Online) Available from https://www.eosc-portal.eu

CHAPTER 6

1. Working Group Working Group on Rewards under Open Science 2017, “Evaluation of Research Careers fully acknowledging Open Science Practices; Rewards, incentives and/or recognition for researchers practicing Open Science.” Luxembourg: Publications Office of the European Union. (Online) Available from https://ec.europa.eu/research/openscience/pdf/os_rewards_wgreport_final.pdf
2. Ali-Khan, SE et al. 2018 “Defining Success in Open Science.” MNI Open Research 2:2. doi:10.12688/mniopenres.12780.1
3. Willinsky, J 2005, “The unacknowledged convergence of open source, open access, and open science.” First Monday, 10(8). (Online) Available from http://firstmonday.org/article/view/1265/1185

CHAPTER 7

1. LIBER Open Science Roadmap 2018 - 2022. (Online) Available from https://publikationen.bibliothek.kit.edu/1000084369/15217908
2. TU Delft Research Data Framework policy 2018, (Online) Available from https://openworking.tudl.tudelft.nl/wp-content/uploads/sites/11/2018/06/tu-delft-research-data-framework-policy-version-for-cvb-18-june-2018.pdf
3. TU Delft Data Stewardship (Online) Available from https://www.tudelft.nl/en/library/current-topics/research-data-management/r/support/data-stewardship/
4. OECD 2015. Making Open Science a Reality, OECD Science, Technology and Industry Policy Papers, No. 25, OECD Publishing, Paris. http://dx.doi.org/10.1787/5jrs2f963zs1-en

CHAPTER 8

1. Chan, LM & Zeng, ML 2006, "Metadata Interoperability and Standardization – A Study of Methodology Part II." D-Lib Magazine. 12:6
2. Riley, J 2004, "Understanding metadata: What is Metadata, and What is it for?" National Information Standards Organization: Baltimore
3. Do, HH & Rahm, E 2000, "On Metadata Interoperability in Data Warehouses." Technischer Report, 01
4. Research Libraries Group (2002) Trusted Digital Repositories: Attributes and Responsibilities.
5. Lin, D et al. 2020, "The TRUST Principles for digital repositories." Scientific Data. 7:144.
6. Australian Data Archive. (2018) Australian Data Archive certified with the Core Trust Seal. (Online) Available from https://ada.edu.au/ada-core-trust-seal/).

CHAPTER 10

1. European Commission (2017) Providing researchers with the skills and competencies they need to practice Open Science. Report of the Working Group on Education and Skills under Open Science.

APPENDICES

Chapter 2

Appendix 2.1. Participant Information Sheet and Survey questionnaire for researchers

Title of Study: Landscape Study on Open Science - Researchers' Perspectives

The Ministry of Energy, Science, Technology, Environment & Climate Change (MESTECC), now known as the Ministry of Science, Technology and Innovation (MOSTI), has launched the Malaysia Open Science Platform (MOSP) Pilot Initiative in November 2019. Through the Academy of Sciences Malaysia (ASM), the Malaysia Open Science Alliance was formed to pave the way towards realising MOSP as a strategic transformative initiative to strengthen the STI Collaborative Ecosystem for Malaysia. The aim of this initiative is to make Malaysia's research data a valuable national asset by developing a trusted platform that enables accessibility and sharing of research data aligned with national priorities and international best practices.

This survey is conducted by the Malaysia Open Science Alliance – Working Group on Policy to assess the landscape of Open Science in Malaysia. This survey is aimed at researchers, who are among the primary stakeholders of the Malaysia Open Science Platform.

The survey will gauge the respondents' knowledge, awareness and participation in Open Science activity. The survey will take 15 minutes to complete. Participation is voluntary and you may exit the survey anytime. There is minimal risk in participating in this survey. Your participation will be strictly confidential and will only be available to the research team. The data will be published anonymously with no direct connection to participants.

If you have any questions, you can contact the Project Manager for MOSP, Dr Nurzatil Sharleeza Mat Jalaluddin (nurzatil@akademisains.gov.my), MOSP Executive, Nur Hanisah binti Ismail (hanisah@akademisains.gov.my) or MOSP Executive Muhamad Fariz Muhamad Suhaimi (fariz@akademisains.gov.my).

By clicking NEXT and completing the survey, you are indicating that you have agreed to take part in this research and give permission for us to gather and analyse the answers you provided. Once your answers are submitted, you will not be able to access them again.

1. Gender

☐ Male
☐ Female

2. Where do you work?

☐ Public Higher Learning Institution
☐ Private Higher Learning Institution
☐ Public Research Institute
☐ Private Research Institute

3. Do you work in any of these research universities?

☐ Universiti Malaya
☐ Universiti Kebangsaan Malaysia
☐ Universiti Sains Malaysia
☐ Universiti Teknologi Malaysia
☐ Universiti Putra Malaysia
☐ No, I am working at

Please specify: ☐

4. Primary role as a researcher

☐ Professor
☐ Associate Professor
☐ Assistant Professor, Senior Lecturer
☐ Senior researcher
☐ Junior researcher
☐ Postdoctoral research fellow
☐ Graduate research student
☐ Undergraduate research student

We would like to know your research field and your experience in research data management

5. What is your research field? Please tick which is applicable.

☐ Biology
☐ Chemistry
☐ Physics
☐ Mathematics
☐ Statistics/applied statistics
☐ Arts
☐ Economics
☐ Social Science and Humanities
☐ Medicine
☐ Law
☐ Computer Science
☐ Engineering
☐ Others

Others (please specify): ☐

6. What types of research data do you produce?

☐ Audio tapes
☐ Computer software source code (html, java, phyton, perl)
☐ Spectral data (such as chromatography, spectroscopy, and spectrometry, to name a few)
☐ Genome and exome sequences, omics, transcriptomes, etc.
☐ Data collected from sensors or instruments
☐ Digital audio files (for example: .mp3, .aif, .wav)
☐ Digital video files (for example: .mp4, .avi, .ogv, .ogg, .mj2)
☐ Interview transcripts

☐ Questionnaires
☐ Research methodologies and workflow
☐ Slides, artefacts, specimens, samples
☐ Fieldwork data (Notes, sketches)
☐ Images scans photos or X-rays
☐ Laboratory notes
☐ Codebooks
☐ Textual data (for example: .rtf, .txt, .xml, .doc/.docx, or some software-specific formats: NUD*IST, NVivo, ATLAS.ti)
☐ Tabular data with extensive metadata (for example, using statistical packages such as SPSS (.sav), Stata (.dta) and MS Access (.mdb/.accdb)
☐ Tabular data with minimal metadata (for example: .csv, .tab, .xls/.xlsx, .mdb/.accdb, .dbf, .ods)
☐ Others
Others (please specify): ☐

7. Where do you store your data?
☐ Public domain
☐ Private domain
☐ Research data centres
☐ Faculty depository
☐ Institutional repository
☐ Shared hard drives
☐ 3rd party data cloud (Dropbox), Google Drive etc.
☐ Pen drive/Hard disk drive
☐ CD/DVD
☐ Local computer
☐ Others
Others (please specify): ☐

8. Can you estimate, as a percentage % of your research time, how much time you have lost reorganising, re-formatting or trying to remember details about data?
☐ 0-20%
☐ 20-40%
☐ 40-60%
☐ 60-80%
☐ 80-100%

9. Does your institution provide any training to manage research data?
☐ Yes
☐ No
☐ I do not know

10. Did you receive any training to manage research data from your institution?
☐ Yes
☐ No

11. Does your institution have a department/centre or a group of professionals that specifically handles all the research data?

☐ Yes

☐ No

If yes, please specify: ☐

We would like to know more about your understanding about open science and experience in open science practice for research data.

12. Are you familiar with the Findable, Accessible, Interoperable and Reproducible (FAIR) principle?

☐ Yes

☐ No

13. Which statement(s) is/are relevant and is applicable about Open Science. You may choose more than one option.

☐ Deposited research data (datasets) should be described, identified and registered or indexed in a clear and unequivocal

☐ Deposited research data (datasets) should be accessible through a clearly defined access procedure, ideally using automated means. Metadata should always remain accessible

☐ Deposited research data (datasets) and metadata are conceptualised, expressed and structured using common, published standards

☐ Characteristics of data and their provenance are described in detail according to domain-relevant community standards, with clear and accessible conditions for use

14. What elements of open science have you used?

☐ Open data
(An example of open data is using or releasing the software code in the public domain for further re-use).

☐ Open access
(Open access means publication in open access journals & proceedings)

☐ Open research data
(Open research data means publishing the data underpinning scientific research results. An example is the deposition of sequenced genomic data on a data repository)

☐ Open collaborative platform
(Open collaborative platform is used by scientists to discuss potential solutions to a problem. An example of an open collaborative platform is The Polymath project)

☐ Citizen science
(Citizen science means active involvement of the public in a research project. An example of a citizen science project is Citizen Science in the Surveillance and Monitoring of Mosquito-borne diseases).

☐ I have never used any elements of Open Science

15. Where have you shared your research data?

☐ Research data as supplementary materials in journals
☐ Discipline-specific data repositories
☐ General repositories, such as Dryad, figshare and zenodo
☐ Institutional repository established at the institution that you are working at
☐ Sharing at a conference
☐ Informal sharing on request via email, direct contact, websites etc.
☐ I do not share my data
☐ Other places

Other places (please specify): ☐

Along with the Open Science movement, the role of a data steward has been recognised immensely. In Open Science, data stewards are accountable for the quality of the deposited data and the curation process, i.e. data cleansing and digital preservation. Data steward is also important in providing necessary support for researchers at an institution.

16. Which aspects below do you feel you need support the most? You may choose more than one.

☐ Store, re-use and analyse information
☐ Research data management
☐ Procedures to create quality metadata
☐ Navigating database software
☐ Ethical issues about sharing data
☐ Legal & Policy knowledge on open science

We would like to know more about open science-related policy at your institution

17. Does your institution have any open science policy?

☐ Yes
☐ No, but is planning to create one
☐ No
☐ I don't know

*18. Does your institution have any open access policy?

☐ Yes
☐ No, but is planning to create one
☐ No
☐ I don't know

19. Does your institution have any open data policy?

☐ Yes
☐ No, but is planning to create one
☐ No
☐ I don't know

20. Does your institution own a research data repository?

☐ Yes
☐ No, but is planning to create one
☐ No
☐ I don't know

*21. Does your institution use a shared research data repository?

☐ Yes
☐ No, but is planning to create one
☐ No
☐ I don't know

We would like to know about your knowledge on research data infrastructure at the Institution that you are working at

*22. Where does the institution deposit research data?

Public domain	☐ Yes	☐ No	☐ I do not know
Private domain	☐ Yes	☐ No	☐ I do not know
Research data centres	☐ Yes	☐ No	☐ I do not know
Faculty depository	☐ Yes	☐ No	☐ I do not know
Shared hard drives	☐ Yes	☐ No	☐ I do not know
3rd party data cloud (Dropbox), Google Drive etc.	☐ Yes	☐ No	☐ I do not know
Other (please specify)			

Malaysia Open Science Platform is a strategic initiative with the aim to make Malaysia's research data a valuable national asset by developing a trusted platform that enables accessibility and sharing of research data aligned to national priorities and international best practices.

*23. Would you give your consent to be contacted for future MOSP-related engagements, such as stakeholders' workshops, training as a data steward and other similar activities?

☐ Yes
☐ No

24. If yes, please fill in your information

Salutation: ☐ Name: ☐ Email Address: ☐

Appendix 2.2. Participant Information Sheet and Survey questionnaire for chief librarians and librarians

Title of Study: Landscape Study on Open Science - Librarians' Perspectives

The Ministry of Energy, Science, Technology, Environment & Climate Change (MESTECC), now known as the Ministry of Science, Technology and Innovation (MOSTI), has launched the Malaysia Open Science Platform (MOSP) Pilot Initiative in November 2019. Through the Academy of Sciences Malaysia (ASM), the Malaysia Open Science Alliance was formed to pave the way towards realising MOSP as a strategic transformative initiative to strengthen STI Collaborative Ecosystem for Malaysia. The aim of this initiative is to make Malaysia's research data a valuable national asset by developing a trusted platform that enables accessibility and sharing of research data aligned with national priorities and international best practices.

This survey is conducted by the Malaysia Open Science Alliance – Working Group on Policy to assess the landscape of Open Science in Malaysia. This survey is aimed at librarians, who are among the primary stakeholders of the Malaysia Open Science Platform.

The survey will gauge the respondents' knowledge, awareness and participation in Open Science activity. The survey will take 15 minutes to complete. Participation is voluntary and you may exit the survey anytime. There is minimal risk in participating in this survey. Your participation will be strictly confidential and will only be available to the research team. The data will be published anonymously with no direct connection to participants.

If you have any questions, you can contact the Project Manager for MOSP, Dr Nurzatil Sharleeza Mat Jalaluddin (nurzatil@akademisains.gov.my), MOSP Executive, Nur Hanisah binti Ismail (hanisah@akademisains.gov.my) or MOSP Executive Muhamad Fariz Muhamad Suhaimi (fariz@akademisains.gov.my).

By clicking NEXT and completing the survey, you are indicating that you have agreed to take part in this research and give permission for us to gather and analyse the answers you provided. Once your answers are submitted, you will not be able to access them again.

1. Gender

☐ Male
☐ Female

2. Years of professional working experience as a librarian

☐

3. Do you work in any of these research universities?

☐ Universiti Malaya
☐ Universiti Kebangsaan Malaysia
☐ Universiti Sains Malaysia
☐ Universiti Teknologi Malaysia
☐ Universiti Putra Malaysia
☐ No, I am working at

No, I am working at (please specify): ☐

We would like to know your perception about your roles in open science practice in the institution you are working at
Open science
Open science is a worldwide movement to make scientific research, data and dissemination accessible to all levels of an inquiring society, amateur or professional. Examples of Open Science activities include open data (for example, using or releasing the software code in the public domain for further re-use), open access (defined as publication in open access journals & proceedings, open research data (defined as publishing the data underpinning scientific research results), open collaborative platform used by scientists to discuss potential solutions to a problem. An example of an open collaborative platform is The Polymath project) and Citizen Science (defined as active involvement of the public in a research project. An example of a citizen science project is Citizen Science in the Surveillance and Monitoring of Mosquito-borne diseases).

4. Your contribution in managing open access publications.
☐ Very significant
☐ Significant
☐ Somewhat significant
☐ Least Significant
☐ Not Significant

5. Your contribution in managing research data repositories.
☐ Very significant
☐ Significant
☐ Somewhat significant
☐ Least Significant
☐ Not Significant

We would like to know your experience and key strengths in data management plan Research data management
A research data management 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.

The job scope for librarians in terms of RDM for Open Science is assisting researchers to handle unstructured data they produced, as well as providing advice on metadata standards for researchers. An example of research data management plan can be found from the University of Edinburgh
https://www.ed.ac.uk/files/imports/fileManager/DataManagementPlans.pdf.

6. Have you engaged in writing data management plans at the university that you are working at?
☐ No
☐ Yes

7. Have you reached out to researchers or students to support their data management plans?
☐ No
☐ Yes
Yes (please specify how frequent): ☐

8. Answer this question if Question 7 is YES.
Select a maximum of 5 research disciplines that you are familiar with as a data librarian.
☐ Biology
☐ Chemistry
☐ Physics
☐ Mathematics
☐ Statistics/applied statistics
☐ Arts
☐ Economics
☐ Social Science and Humanities
☐ Medicine
☐ Law
☐ Computer Science
☐ Engineering
☐ Others
Others (please specify): ☐

Open science promotes data sharing in communities under the tagline "As open as possible, as close where necessary". Along with the Open Science movement, the role of data librarian as a data steward has been recognised in many countries especially in Europe. In Open Science, data stewards are accountable for the quality of the deposited data and the curation process, i.e. data cleansing and digital preservation.

Data curators

Definition:
Data curators are responsible for organising and integrating data collected from various sources. It involves annotation, publication and presentation of the data such that the value of the data is maintained over time, and the data remains available for reuse and preservation.

Activities:
The duties of data curators include cataloguing (i.e. to describe the materials), which must follow the Resource Description and Access (RDA) instructions and indexing (i.e. make the material retrievable and indexed by name, subject etc.)

Data stewards

Definition:
The role of data stewards overlaps with data curators in the aspects of both metadata management activities and data governance. In addition to this, data stewards are accountable for the quality of the data (in terms of completeness, consistency, uniqueness, validity and accuracy) and the curation process, i.e. data cleansing and digital preservation.

Activity:
In the current situation in Malaysia, the job scope of librarians as a data steward is assisting researchers to identify and advise which metadata to be used according to the Resource Description Framework (RDF). Another task is providing advice on & writing the data management plan with researchers.

9. Have you attended a course that prepared you as a data steward?
☐ No
☐ Yes

10. If a data stewardship training session is to be conducted in the library that you are working at, which of the following aspects are your main interests? You may choose more than one.
☐ Scientific data integrity
☐ Research skills
☐ Knowledge on metadata
☐ Navigating database software
☐ Preserving research output
☐ Preservation of records
☐ Legal knowledge on open science

11. Are you linked to any local or international network for data stewards?
☐ No
☐ Yes

Lastly, we would like to know about your knowledge on research data infrastructure (e- infrastructure and capacity) at the Institution that you are working at

12. Where does the institution deposit research data?

Public domain	☐ Yes	☐ No	☐ I do not know
Private domain	☐ Yes	☐ No	☐ I do not know
Research data centres	☐ Yes	☐ No	☐ I do not know
Faculty depository	☐ Yes	☐ No	☐ I do not know
Shared hard drives	☐ Yes	☐ No	☐ I do not know
3rd party data cloud (Dropbox), Google Drive etc.	☐ Yes	☐ No	☐ I do not know

13. Besides data librarians, do you think researchers should also be trained as a data steward and be familiar with the essentials of research data management in the context of Open Science?

☐ No
☐ Yes

The Malaysia Open Science Platform is a strategic initiative with the aim to make Malaysia's research data a valuable national asset by developing a trusted platform that enables accessibility and sharing of research data aligned with national priorities and international best practices.

14. Would you give your consent to be contacted for future MOSP-related engagements, such as stakeholders' work-shops, training as a data steward and other similar activities?

☐ No
☐ Yes

15. If yes, please fill in your information

Salutation: ☐ Name: ☐ Email Address : ☐

Appendix 2.3. List of Target Institutes of Higher Learning (IHLs), Public and Private Research Institutes for survey questionnaire

Institutes of Higher Learning
1. Universiti Teknologi Malaysia
2. Universiti Malaya
3. Universiti Putra Malaysia
4. Universiti Sains Malaysia
5. Universiti Kebangsaan Malaysia
6. Universiti Malaysia Sabah
7. Universiti Tun Hussein Onn Malaysia
8. Universiti Islam Antarabangsa
9. Multimedia University
10. Open University Malaysia
11. Universiti Teknologi Mara
12. Universiti Malaysia Pahang
13. Universiti Malaysia Terengganu
14. Universiti Kuala Lumpur
15. Universiti Malaysia Perlis
16. Universiti Malaysia Sarawak
17. Universiti Tenaga Nasional
18. Universiti Sains Islam Malaysia
19. International Medical University
20. Universiti Utara Malaysia

Public Research Institutes
1. Malaysian Palm Oil Board
2. Forest Research Institute Malaysia
3. Institute for Medical Research
4. Malaysian Agricultural Research and Development Institute
5. International Science, Technology and Innovation
6. Malaysia Nuclear Agency
7. Maritime Institute of Malaysia
8. National Institute of Biotechnology Malaysia
9. Fisheries Research Institute
10. Department of Mineral & Geoscience Malaysia

Private Research Institute
1. World Wide Fund For Nature – Malaysia
2. Cancer Research Malaysia
3. Tropical Rainforest Conservation & Research Centre
4. MareCet Research Organisation
5. Marine Research Foundation
6. Malaysian Biotechnology Information Centre

Appendix 2.4. Interview instrument with Top Management Universities

INTERVIEW QUESTIONS

1. Why would the University have interest in Open Science?
2. Has the University established an Open Science initiative at an institutional level? Could you please describe it further?
3. What are University's policies supporting open science practices and implementation that are already in place?
    a. Probe: Data management plan
4. How did the University support building the talents capacity (i.e. data curators & data stewards) who can guide and advise researchers throughout the whole data management plan & data stewardship?
    a. Probe: What are the initiatives of the University to equip essential skills to train new data stewards or upskill the existing ones? Are there any fees for sponsorship for participation in training courses outside campus, or the establishment of your own training centre on-campus?
5. What is your vision for participation in Open Science initiatives including MOSP?

Appendix 2.5. Interview instrument with Junior Researchers

INTERVIEW QUESTIONS

1. How important are data sharing and re-use and Open Science practices in your daily work as researchers? Probe: collaboration, visibility & citation
2. What are the challenges regarding the sharing of research data in your field? Probe: IP and commercial conflicts, research being scooped,
3. As young researchers, how do you see your future? What are your objectives and how can MOSP support you in achieving them?
4. How do you see MOSP supporting the advancement of research for Malaysia and the World? Probe: E.g. Importance of Open Science in light of COVID-19; strengthening Malaysia's presence in the global research landscape.
5. How can Malaysia support the future of Research, in the context of Open Science?
    a. What is your perspective about challenges related to data sharing with third parties, especially commercial-oriented entities?
    b. What is your perspective about the role of funding bodies in promoting Open Science?

Appendix 2.6. Interview instrument with Senior Researchers

INTERVIEW QUESTIONS

1. How do you share your research with others, besides lectures, conferences, publications & discussions?
    a. Probe: Do you publish your research dataset in a repository? Please specify
    b. Probe: Do you encourage the public to participate in your research? Please specify
    c. Probe: Do you release software codes in the public domain (Note: this question is domain-specific)
2. How important are data sharing and re-use and Open Science practices in your daily work as researchers?
    a. Probe: collaboration, visibility & citation, interdisciplinary research

3. What elements are essential for open science & the FAIR principle that you have taught in class?
    a. Probe: The importance of designing a data management plan before starting a research project; The importance of conducting a reproducible method and analysis; The importance of obtaining consent from data subjects for sensitive data sharing and re-use

4. What is the best way to promote data sharing and open science to researchers and industries in Malaysia?
    a. Probe: Barriers, E.g. IP and commercialisation, plagiarism; Incentives – e.g. funding body, KPI for open science

Appendix 2.7. Interview instrument with Chief Librarians

INTERVIEW QUESTIONS

1. How many librarians are employed in the University?
    a. Full-time
    b. Part-time
    c. Casual
2. What are the library practices at (Name of University) around data management, including data curation and data stewardship?
    a. How many open access journals are subscribed?
    b. How many repositories are being established?
    c. For each (a) and (b): Probe: Are these aspects (i.e. persistent identifiers, metadata, licensing, IPR, data citation, archiving and back-up data) adhering to international best practices?.
    d. For (a) and (b): Probe: Are the aspects listed above tailor-made and for what purpose?
3. What is the frequency for each librarian to engage with a researcher when writing a data management plan?
4. What are the main research disciplines that the Library has engaged with as a data steward?
5. What are the current strengths of librarian skills that can contribute to the success of the Open Science initiative, such as MOSP?
6. What areas require improvement such that Open Science will be more successful in the future.

Appendix 2.8. Interview instrument with Head of Data Centres or Representatives from Institutional Repositories

INTERVIEW QUESTIONS

1. What is the data repository system that you are using?
    a. Is the system built in for public use or private access?
2. What is the type of data being deposited? E.g. omics data, software codes
3. How is the data being retrieved? What type of persistent identifier is being used?
4. How long can the data be stored? Is there any specific data management plan or policy being adhered to? If yes, what is that?
5. What is the Internet bandwidth of the organisation network to the Internet Service Provider (ISP).
6. What is the current capacity at your centre?
    b. How big is the size of data being deposited on a daily, weekly or monthly basis?
7. How are daily operations and maintenance of running the infrastructure sustained?

Appendix 2.9. Participant Information Sheet for semi-structured interviews

Participant Information Sheet for interview

Title of Study: Landscape Study on Open Science
The Ministry of Energy, Science, Technology, Environment & Climate Change (MESTECC), now known as the Ministry of Science and Technology (MOSTI), has launched the Malaysia Open Science Platform (MOSP) Pilot Initiative in November 2019. Through the Academy of Sciences Malaysia (ASM), the Malaysia Open Science Alliance was formed to pave the way towards realising MOSP as a strategic transformative initiative to strengthen STI Collaborative Ecosystem for Malaysia. The aim of this initiative is to make Malaysia's research data a valuable national asset by developing a trusted platform that enables accessibility and sharing of research data aligned with national priorities and international best practices.

This study is conducted by the Malaysia Open Science Alliance – Working Group on Policy to explore the landscape of Open Science in Malaysia from multiple perspectives' point-of-view. The interview is aimed at the main stakeholders of the Malaysia Open Science Platform which includes researchers, deputy vice chancellors (research and innovation) of higher learning institutions, head of data management centres and chief librarians.

Opinions will be elicited from a semi-structured interview, guided by a set of open-ended questions. Upon consent from the participant, the interview will be recorded and transcribed verbatim. The interview may last between 30 to 60 minutes. Participation is voluntary. There is minimal risk in participating in this survey. Your participation will be strictly confidential and will only be available to the research team. The data will be published anonymously with no direct connection to participants.

If you have any questions, you can contact the Project Manager for MOSP, Dr Nurzatil Sharleeza Mat Jalaluddin (nurzatil@akademisains.gov.my), MOSP Executive, Nur Hanisah binti Ismail (hanisah@akademisains.gov.my) or MOSP Executive Muhamad Fariz Muhamad Suhaimi (fariz@akademisains.gov.my).

Appendix 2.10. Participant Consent Form for semi-structured interviews

Consent Form for Interview

Project title: Landscape Study on Open Science
By the Malaysia Open Science Alliance – Working Group on Policy

I, ……, have been invited
to participate in the above study.

- I have received and read the attached 'Participant Information Sheet' and understand the general objectives, methods and demands of the study. I understand that the project may or may not be of direct benefit to me.

- I have read and understood in the attached 'Participant Information Sheet' describing the tasks that I may be required to perform and the possible risks.

- I understand my participation in this study will be kept confidential and access to the data will be limited strictly to the research team. I consent to the publishing of results from this study provided my identity is anonymous.

- I understand that I can refuse to consent or withdraw from the study at any time without explanation.

- I hereby voluntarily consent and offer to take part in this study.

Signature: …… Date: ……
Name ……
Position ……

Please return this completed form to the MOSP Project Manager/Executive

Chapter 3

Appendix 3.1. Relevant acts, policies and guidelines related to data sharing in Malaysia

COMMON LAW PRINCIPLES

1. Defamation
2. Trespass
3. Nuisance
4. Trade Secrets and Breach of Confidence

ACTS / REGULATIONS / ENACTMENT

5. Biosafety Act 2007
6. Computer Crimes Act 1997
7. Copyright Act 1987
8. Fees Act 1951 Act 209
9. Industrial Designs Act 1996
10. Layout-designs of Integrated Circuits Act 2000
11. Malaysia Communications and Multimedia Act 1998
12. National Security Council Act 2016
13. Official Secrets Act 1972
14. Patents Act 1983
15. Penal Code (Act 574)
16. Personal Data Protection Act 2010
17. Statistics Act 1965 Act 415
18. Trademarks Act 2019
19. Wildlife Conservation Act 2010, Act 716
20. Malaysian Palm Oil Board Act 1998, Act 582
21. Access to Biological Resources and Benefit Sharing Bill 2017

Environment

22. Environmental Quality Act 1974 - ACT 127
23. Exclusive Economic Zone Act, 1984, Act No. 311
24. National Forest Policy 1978
25. Malaysian Timber Industry Board (Incorporation) Act 1973
26. Research and Development Authority Act 1985 Malaysia's forests
27. Plant Quarantine Act 1976
28. Fisheries Act 1985
29. National Resources and Environment Ordinance, 1958, Sarawak

Biodiversity

30. Sarawak Biodiversity Centre Ordinance (Amendment) 2003
31. Sarawak Biodiversity Regulation 2004
32. The Sabah Biodiversity Enactment 2000

Land

33. National Land Code 1965

Water

34. Waters Act, 1920

POLICY & GUIDELINES

35. Garis Panduan Perkongsian Data Terbuka Bagi Data Mentah Saintifik dan Penyelidikan Melalui Platform Raw Data for Research & Science (RADARS)
36. Garis Panduan Mengakses Malaysian Health Data Warehouse (MyHDW)
37. Garis Panduan Perkongsian dan Penyebaran Maklumat Geospatial melalui Infrastruktur Data Geospatial Negara (MyGDI)
38. Intellectual Property Commercialisation Policy for Research & Development (R&D) Projects Funded by the Government of Malaysia
39. National Policy on Biological Diversity 2016 – 2025
40. UM Intellectual Property and Commercialisation Policy
41. UKM Intellectual Property Policy
42. Universiti Putra Malaysia (Research) Rules
43. UTM Intellectual Property Commercialisation Policy
44. UTM Intellectual Property Policy
45. Intellectual Property Policy USM
46. Malaysian Medical Council (MMC) revised guidelines on Confidentiality

CIRCULAR

47. General Circular Number 1 of 2015 Open Data Implementation
48. NPRA directive: Arahan Dibawah Peraturan 29, Peraturan-Peraturan Kawalan Dadah Dan Kosmetik 1984, Keperluan Mendaftar Penyelidikan Klinikal Yang Melibatkan Ubat¬Ubatan Dengan National Medical Research Register (NMRR)

TERMS

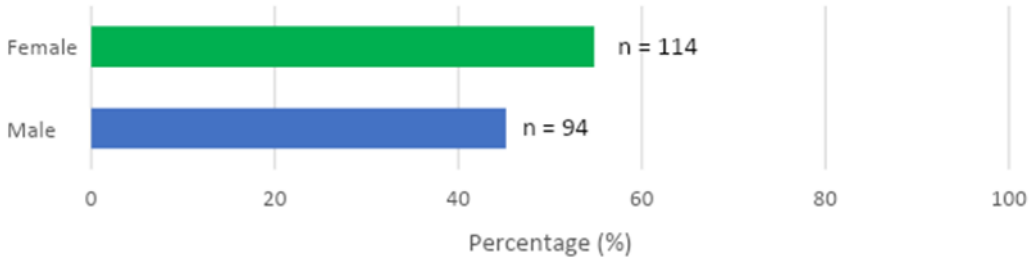
49. Terms for Government Open Data Sharing 1.0
50. Terms for Knowledge Resources for Science and Technology Excellence, Malaysia (KRSTE.MY)

**any reference to the list of acts/regulations/enactment/police hereinabove, shall be deemed to include a reference to any other written law, regulation or order relevant to data sharing*

Chapter 6

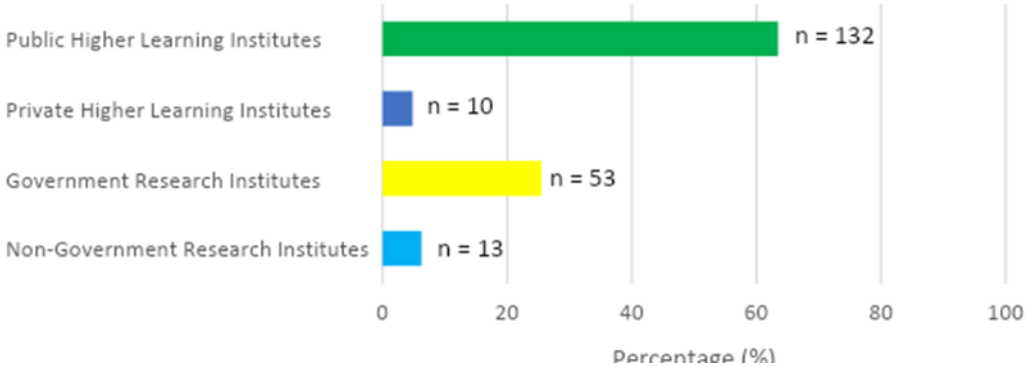
Appendix 6.1. Demographic background of survey participants (Researchers)

Gender



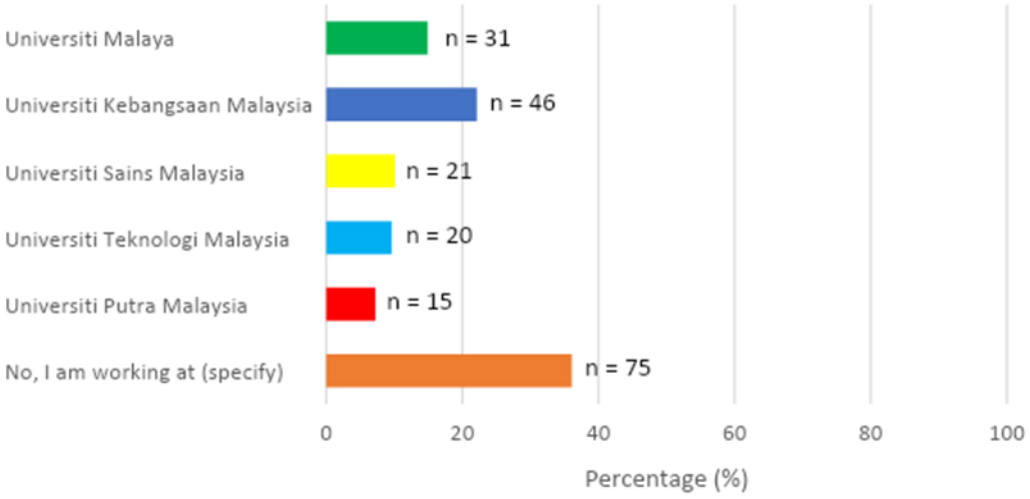
Where do you work

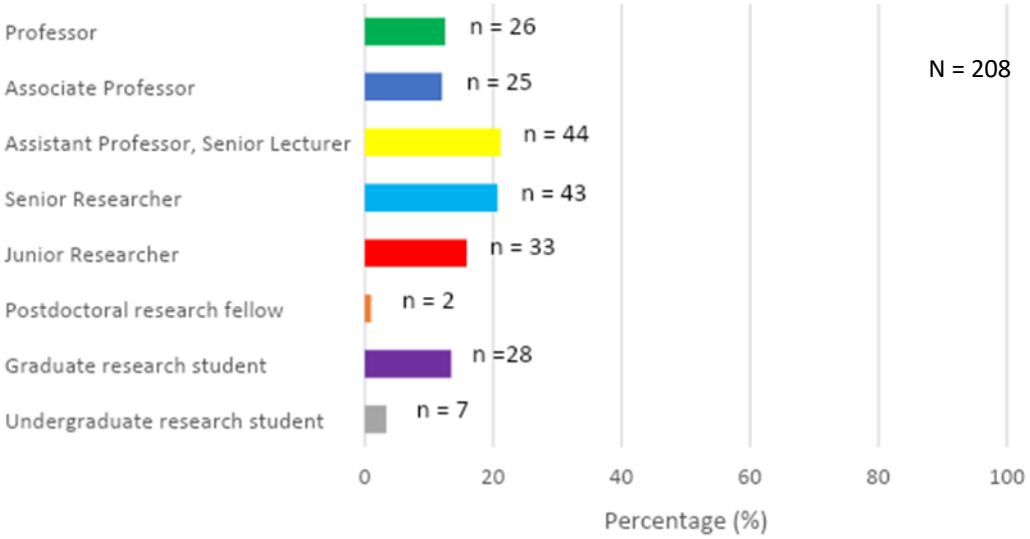
N = 208



N = 208

Do you work in any of these Research Universities





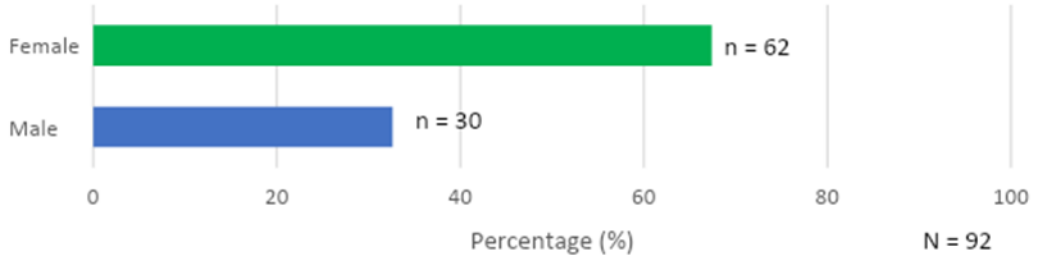
N = 208

N = 208

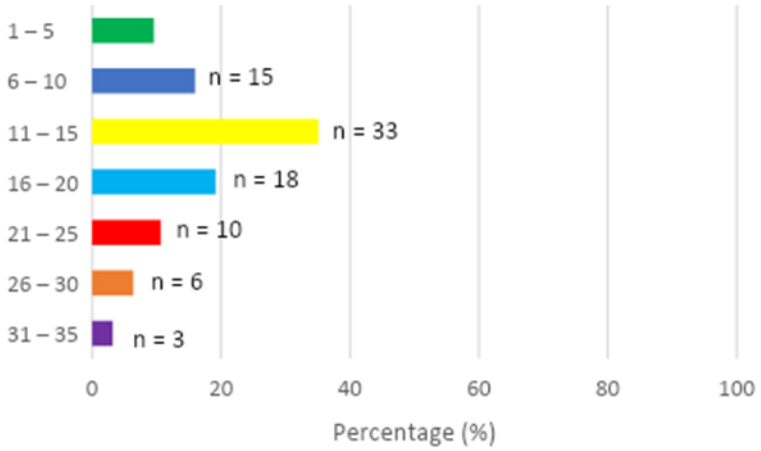
Chapter 7

Appendix 7.1. Demographic background of survey participants (Librarians)

Gender

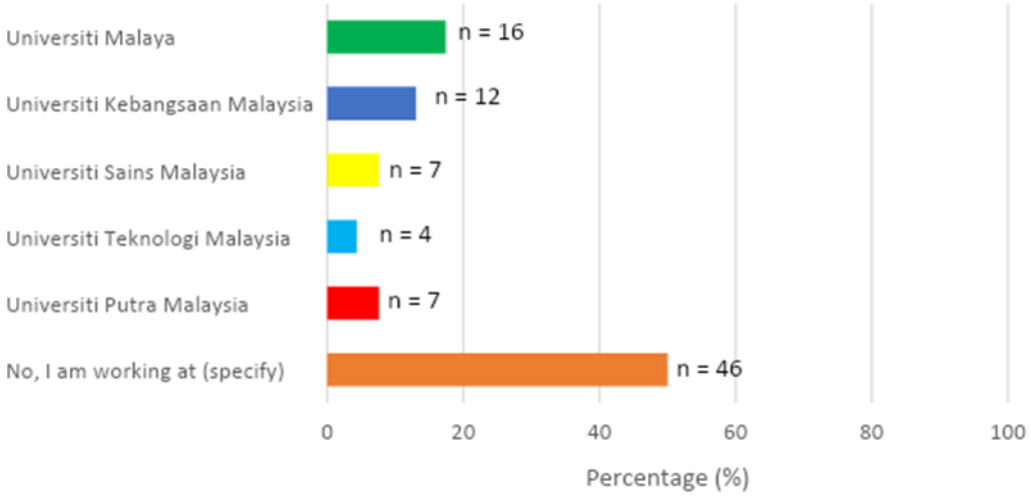


Years of Professional Working Experience as a Librarian



Do you work in any of these Research Universities

N = 92



N = 92

AKADEMI SAINS MALAYSIA

Think Science.
Celebrate Technology.
Inspire Innovation.

ACADEMY OF SCIENCES MALAYSIA
Level 20, West Wing, MATRADE Tower, Jalan Sultan Haji Ahmad Shah
off Jalan Tuanku Abdul Halim, 50480 Kuala Lumpur
T: +6 (03) 6203 0633 / F: +6 (03) 6203 0634
www.akademisains.gov.my

ISBN 978-983-2915-62-1
9 7 8 9 8 3 2 9 1 5 6 2 1