NIH Data Repository (NIH-DaRS)-Unlocking **Challenges and Possibilities in** Healthcare **Research Data** Sharing

Experience from NIH Malaysia

Dr. Muhd Zulfadli Hafiz Ismail Sector for Biostatistics & Data Repository, NIH

23/10/2023



# **NIH Overview**



### Scope of Institutes & Divisions







#### NIH Data Repository (NIH-DaRS)

# Understanding the landscape of open science in NIH

# Challenges and measures to overcome

Way forward and key points

# Agenda



# NIH Data Repository (NIH-DaRS)



### Research data lifecycle



- Data repository will be the missing piece of jigsaw puzzle to complete the data lifecycle
- Researcher role in data lifecycle- data creators
- As data creators, our role is to ensure the data produced will be of benefit to data owners (funder- MOH)- make data available
- Data more accessible, more utilisation, more evidence produced- betterment for nation's health

### What is NIH-DaRS?

- National Institutes of Health Data Repository System (NIH-DaRS) is a platform for curating primary research data produced in Ministry of Health Malaysia.
- Functions :
  - Deposit data for long-term preservation
  - Request data for secondary usage (eg. for publications, presentation, thesis or research projects)

### nihdars.nih.gov.my



About NIH-DaRS

Nanous institutes of Health - Dala Reports System (NH-Gal45) is a patituthe for contracting seasch data (p) Meahor (H-Health Matayaaa, Research data (a che sosten unavos filmeta dato forume data (personicita) de potentital for sociatifica data guarda (matanticulana usage) de matanticulare sociatifica data (a che sosten assel in supporting and facilitarigi deticulare), pelos perspectivos, and promoting new pacifications and research projects. NH-Gal45 is managed by the Societor for Boostandes Cala Respectivos (che Matayaa) and and che Matayaaa (che Matayaaa).



• Sector Biostatistics & Data Repository as gatekeeper

#### **NIH-DaRS** life stage



### Where we are at the moment



Total: 38 deposited data

Total: 161 requested data

# Understanding the landscape of open science in NIH

#### Understanding factors influencing the adoption of open science practices among Malaysia's NIH researchers

Information Development 1-20 © The Author(s) 2024 Article reuse guidelines: sagepub.com/journals-permissions DOI: 10.1177/02666669241248145 journals.sagepub.com/home/idv Sage

#### AN OPEN SCIENCE MODEL FOR RESEARCH DATA MANAGEMENT AT THE MALAYSIA NATIONAL INSTITUTES OF HEALTH

- By: Mahfooz Ahmed (G2115023)
- Main Supervisor: Prof. Datin Dr. Roslina Othman
- Co-Supervisors: Prof. Dato' Dr. Mohamad Fauzan Noordi
- Dr. Muhd Zulfadli Hafiz Bin Ismail (NIH)
- Dr. Zulkarnain Bin Abd Karim (NIH)
- Dr. Sharifah Nazeera binti Syed Anera (NIH)

#### Mahfooz Ahmed

Kulliyyah of Information and Communication Technology, International Islamic University, Kuala Lumpur, Malaysia

#### **Roslina Othman**

Kulliyyah of Information and Communication Technology, International Islamic University, Kuala Lumpur, Malaysia

#### Mohamad Fauzan Noordin

Kulliyyah of Information and Communication Technology, International Islamic University, Kuala Lumpur, Malaysia

#### Muhd Zulfadli Hafiz bin Ismail

Office of the NIH Administrator, National Institutes of Health (NIH), Ministry of Health Malaysia, Setia Alam, Malaysia

#### Zulkarnain bin Abd Karim

Office of the NIH Administrator, National Institutes of Health (NIH), Ministry of Health Malaysia, Setia Alam, Malaysia

#### Sharifah Nazeera Binti Syed Anera

Office of the NIH Administrator, National Institutes of Health (NIH), Ministry of Health Malaysia, Setia Alam, Malaysia

#### Abstract

**Background:** Open science practices, including research data sharing and reuse, are crucial for advancing scientific knowledge. However, the extent to which Malaysia's National Institutes of Health (NIH) researchers participate in open science remains a subject of inquiry.

**Methods:** This study employs an exploratory sequential mixed-method approach to investigate the factors influencing Malaysia NIH researchers' participation in open science practices. The research combines indepth interviews with key managers and administrators from NIH and a quantitative survey distributed to the researchers. Through in-depth interviews, administrators provide insights into the institute's perspective on open science. At the same time, the survey explores researchers' views on factors influencing their participation in open science and how these factors influence the data-sharing practices at the institute.

**Results:** The findings reveal key factors that influence the participation of Malaysia NIH researchers in open science, shedding light on both motivating factors and barriers. The study identifies strategies to promote researchers' involvement in open science practices through data sharing and highlights the practical implications for the Malaysia NIH and the broader scientific community.

### **QUALITATIVE RESEARCH FINDINGS**

What are the key components of an open science model for open science participation?



Figure 2.28 Conceptual Model of the Current Research

### **QUANTITATIVE RESEARCH FINDINGS**





### Challenges and measures to overcome



# Challenges

- Awareness on open science guidelines in NIH
- Nature of data- diverse and ability of platform to handle the diversity
- Cultural change- what is it for me?

## Challenges

#### • Awareness on open science guidelines in NIH

- Nature of data- diverse and ability of platform to handle the diversity
- Cultural change- what is it for me?

# Guidelines and policy



SE

#### AKU JANJI KETUA PROJEK National Institutes of Health (Institut Kesihatan Negara), Kementerian Kesihatan Malaysia



ahawa	asanya, saya	[Nama	Penuh]					[N	o Kad Pe	engenalan]
			seorang w	arganega	ara Malaysia da	alaysia dan kini berkhidmat di [Nama Institut/ Hospital/ dengan sesungguhnya dan sebenarnya men	ital/ Bah	Bahagian] gaku ianji		
eperti	berikut:					, in the second s	,		U	
а	. Bersetuju	untuk	menerima	Geran	Penyelidikan	Kernenterian	Kesihatan	Malaysia	(KKM)	berjumlah



- Bahawa saya akan sentiasa menjalankan projek penyelidikan dalam tempoh yang telah diluluskan dengan cermat, cekap, jujur, amanah dan bertanggungjawab;
- Bahawa saya akan membelanjakan peruntukan dengan berhemah mengikut undang-undang dan peraturan-peraturan kewangan yang berkuat kuasa dari semasa ke semasa;
- Mencalonkan pengganti sebagai Ketua Projek jika berpencen atau meninggalkan perkhidmatan dan akan berusaha memastikan projek berjalan dengan lancar sehingga projek tamat;
- Memastikan semua Laporan Kewangan, Laporan Kemajuan Projek dan laporan yang diperlukan dan ditetapkan oleh Institut Kesihatan Negara (*National Institutes of Health*) dikemukakan pada masa dan tarikh yang ditetapkan berdasarkan garis panduan yang berkuat kuasa;
- f. Menyimpan data penyelidikan yang telah lengkap ke dalam Sistem Data Repositori Institut Kesihatan Negara (NIH-DaRS) selewat-lewatnya satu tahun selepas projek tamat (tempoh satu tahun ini dikira bermula dari tarikh penyerahan laporan tamat projek);
- g. Mendapatkan kelulusan daripada Ketua Pengarah Kesihatan sebelum menerbitkan sebarang diseminasi atau sebaran hasil penyelidikan melalui penulisan, pengiklanan, pembentangan atau untuk disiarkan di media massa/ media sosial; dan
- h. Memastikan semua hak cipta dan harta intelek milik serta rahsia Kerajaan terpelihara walaupun selepas Projek tamat...



HOME SEARCH BROWSE POLICY FAQ USER MANUAL CONTACT US REGISTER LC

#### Data Access and Use Policy for NIH-DaRS

The NIH Data Repository System (NIH-DaRS) is operated under Sector for Biostatistics & Data Repository of the National Institutes of Health (NIH), Ministry of Health Malaysia. NIH-DaRS is a platform for centralizing research data to make data more FAIR (findable, accessible, interoperable, and reusable).

Research data can be stored in various formats and for long-term data preservation. These datasets are valuable and potential for secondary data usage and translational research. Therefore, this system will assist in supporting and facilitating decision-making, policy preparation, and promoting new publications and research projects.

The major function of NIH-DaRS is to support good practice in data management and completing the research data lifecycle, namely curation, discover and reuse of data. This function will be implemented through these two activities;

#### Deposit Data

Researchers are welcomed to deposit their research data in NIH-DaRS to make data more FAIR (Findable, Accessible, Interoperable, and Reusable) and to ensure long-term preservation.

#### Request Data

Researchers can explore and request datasets in NIH-DaRS for secondary usage (example for publications, presentation, thesis or research projects) according to data access and use policy.



### NIH-DaRS ROADSHOWS Total: 32







## Challenges

- Awareness on open science guidelines in NIH
- Nature of data- diverse and ability of platform to handle the diversity
- Cultural change

### **Preparation for Depositing Data**



### Data validation and standards

- Data validation
- Data standards and data harmonisation
  - OMOP CDM
  - ICD-11, LOINC, SNOMED



The Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM)

/ou may view the ⊞ <u>Field Types video (4 min)</u> . ✓
Use the Rich Text Editor ? Variable Name (utilized in logic rates and expecte)
condition         Enable auto naming of var based upon its Field Label           ONLY letters, numbers, and underscores
How to use [] Smart Variables Piping + Field Embedding
Validation? (optional) None V
Enable searching within a biomedical ontology ?
∠ ICD10 - International Classification of Diseases, Ve
Required?* ● No ○ Yes * Prompt if field is blank
Identifier?  No O Yes Does the field contain identifiant information (e.g., name, SSN, address)?
boos die neu condain lachdrying mormadori (cigi, name, son, address).

REDCap

Small reminder text displayed underneath field

## Challenges

- Awareness on open science guidelines in NIH
- Nature of data- diverse and ability of platform to handle the diversity
- Cultural change

# Balancing act- contextualising repository practice

- Incentive to data creators (researcher) vs data owner (funder) interest
  - All primary research data that was produced by ;
    - Researchers who received funding from MOH research grant are compulsory to be deposited in NIH-DaRS.
- Depositors may submit the research data latest by 1 year after completion of the research project as an incentive to data creators to mine the data and publish before depositing in NIH-DaRS.
- Depositors are encouraged to send list of potential publications (plan or ongoing) together with the research data to ensure the publications will not be produced by others upon depositing the data.
- FAIR- closed and restricted access

#### Depositor can choose access level for the data

No.	Title	Access Level	Reason	Size	Action	
1	BOOK1.XLSX	D	Select Reason	9 KB	× Remove	
2	BOOK2.XLSX	CLOSED	Select Reason Prior formal agreement with other parties on non-disclo National security issues Ethical reason Legal requirement	osure of	the data	
		CLOSED	Choose File No file chosen		wnload 🗠	
3	BOOK3.XLSX			9	× Remove	

**RESTRICTED** Research data are accessible only after approval of data request process, while the research information is available publicly.



The research data are not accessible by other user (other user cannot request

The research data are not accessible by other user (other user cannot request data), only the research information is accessible publicly.

#### **Received notification**

email

#### A researcher has requested data from your deposit file entitled National Health and Morbidity Survey (NHMS) 2018: Elderly Health in DaRS. You received this notification because you enabled notifications when someone requested your data deposit in your profile.

Details as below :

Good morning,

-Name : Hamimat -Email : <u>Hamimat</u>

Effiail . <u>Hamimat</u>

-Title : Air Pollution and the Health of Vulnerable Groups: Short- and Long-term Analysis of Exposure to Air Pollution in Children, Adolescents and Older Populations in Malaysia.

#### View Deposit



### Way forward and key points



# Way forward

- API- link with MOSP platform to increase visibility
- Source code, tools/instruments
- Federated approach computational resources, analytical tools, data storage (cloud)



# Key points

- NIH-DaRS upholds the FAIR principle
- Challenges- cultural, awareness, nature of data
- Enhancing open science in NIH- incentives, protect interests of data creators and data owners, data standards and harmonisation
- Way forward- increase visibility, federated approach



### m.zulfadli@moh.gov.my

