NII RCOS
Research Center for Open Science and Data Platform

NII Research Data Policy and Institutional Data Governance

Meeting with Professor Laurent Romary

September 4－5, 2023

Miho Funamori

National Institute of Informatics

Governance Session 1:
System Architecture Context

National context on research data policy

- May 2021
- Cabinet Office, "6th Basic Plan for Science, Technology and Innovation"

Sets the targets:

1. All Japanese universities with an institutional repository need to set up a research data policy by 2025.
2. All funding agencies need to introduce DMPs for CFPs 2023 and a system supporting metadata schema needs to be introduced as well.

Universities with a research data policy

- Kyoto University
- Nagoya University
- Tokyo Institute of Technology
- Tohoku University
- Kanazawa University
- Keio University
- Kobe University
- Tokyo University

:

Kyoto University Policy on Research Data Management and Sharing

(Preamble)

1. In this policy, research data is defined as the recorded information, both digital and non-digital,gathered or produced by the researchers of Kyoto University in the process of their research activities.
2. The researchers of Kyoto University should be aware that Research Data Management, defined as managing and preserving research data appropriately, is indispensable in conducting research of a high standard.
3. Kyoto University acknowledges that researchers who have gathered or produced research data in principle have the right and responsibility to implement Research Data Management. The researchers of Kyoto University implement Research Data Management in compliance with legal and ethical requirements in each research field for the purpose of retaining the value of research data.
4. Kyoto University will promote the utilization of research data by sharing it to the greatest possible extent unless otherwise specified, based on the understanding that research data, along with academic theses, serves as a foundation of knowledge that contributes to the current and future development of scholarship and society.
5. Kyoto University is responsible for providing the environment to support Research Data Management and sharing.

This policy will be reviewed as needed in response to changes in society and academic circumstances.

Approved by the Kyoto University Research Information Management Committee on March 19, 2020

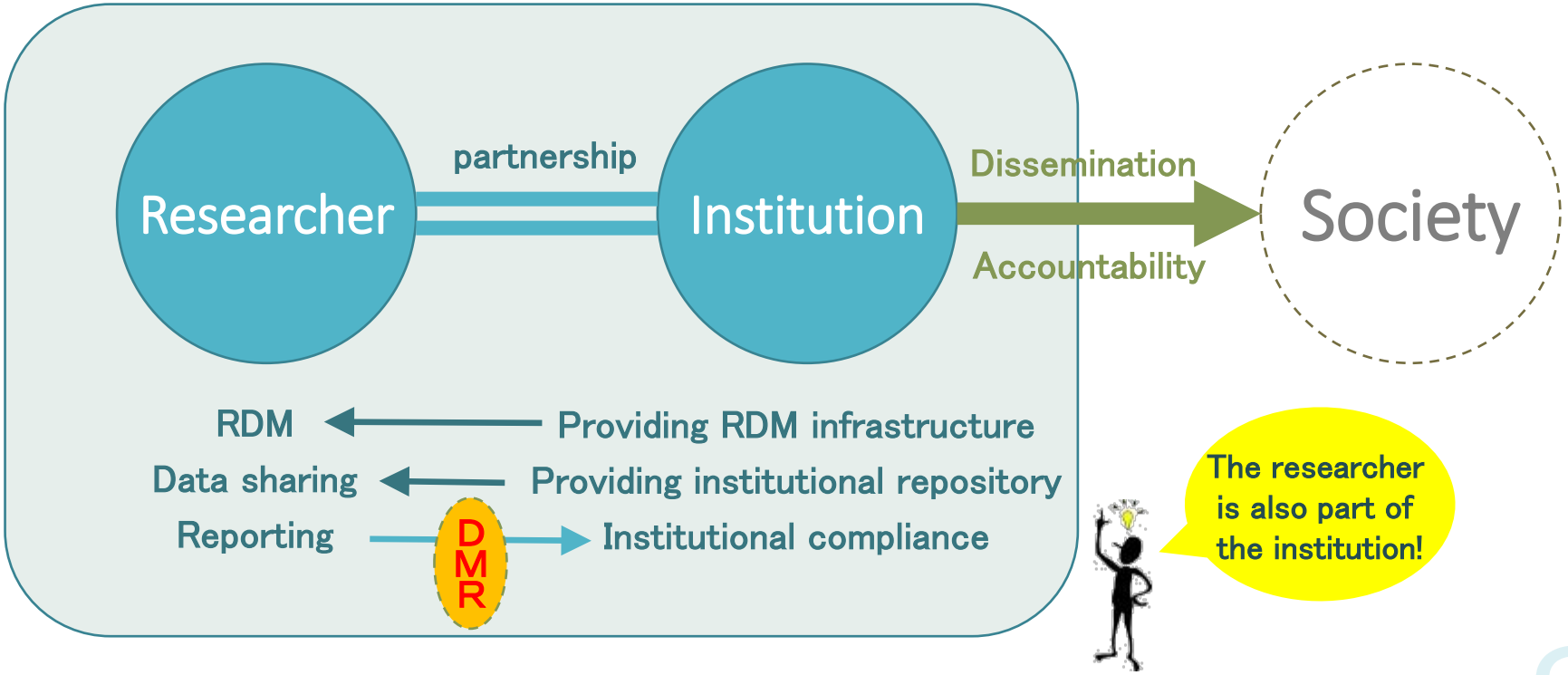
Features of NII Research Data Policy

1. Concrete and actionable
2. Defines institutional research data governance
3. To be implemented utilizing data governance function of NII RDC
4. To be provided to the Japanese universities as a policy template along with NII RDC

Why do we need institutional data governance?

- Currently, research data is managed by the researcher only and the institution does not have any clue what data the institution has, how it is managed etc.
- However, once something bad happens, such as research fraud, data leakage, the institution is the one who needs to take responsibility.
- In order to be accountable to the public, the institution needs to know what data the institution has and needs to manage them at all times.

Institutional research data management through partnership between institution and researcher

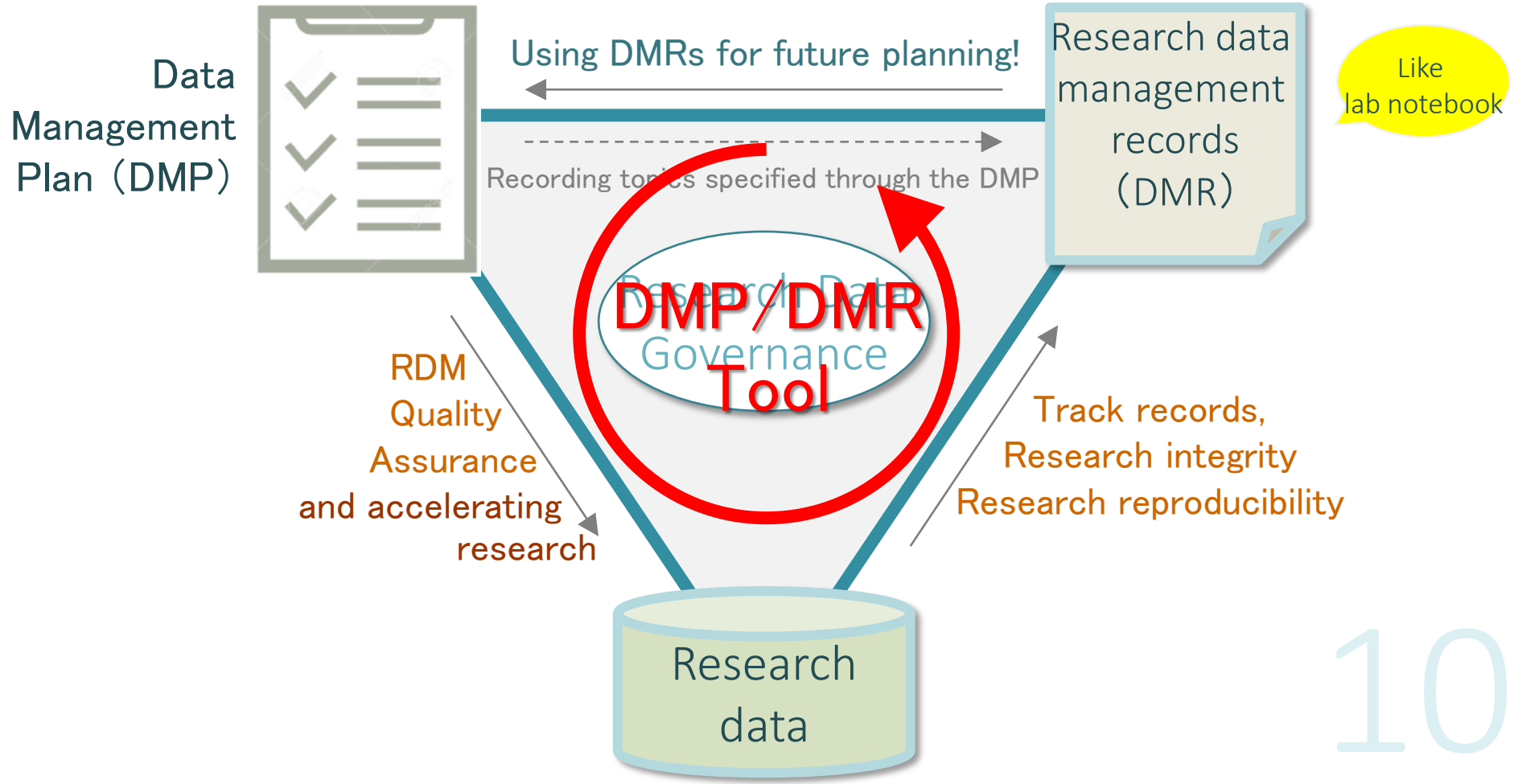


DMP → DMR

Data Management Record

Institutional research data governance should be underlined by DMRs, not DMPs!

Using DMP, DMR, and Research Data
for institutional research data governance



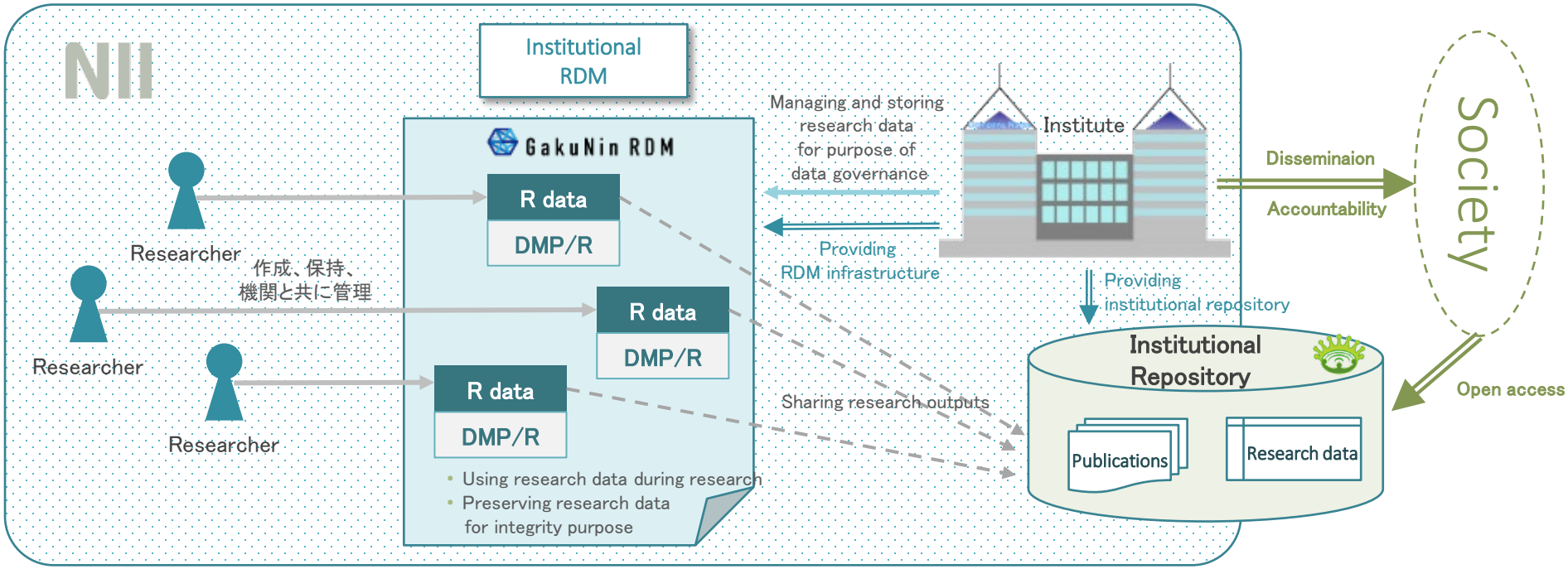
NII's Research Data Governance

Researcher responsibility

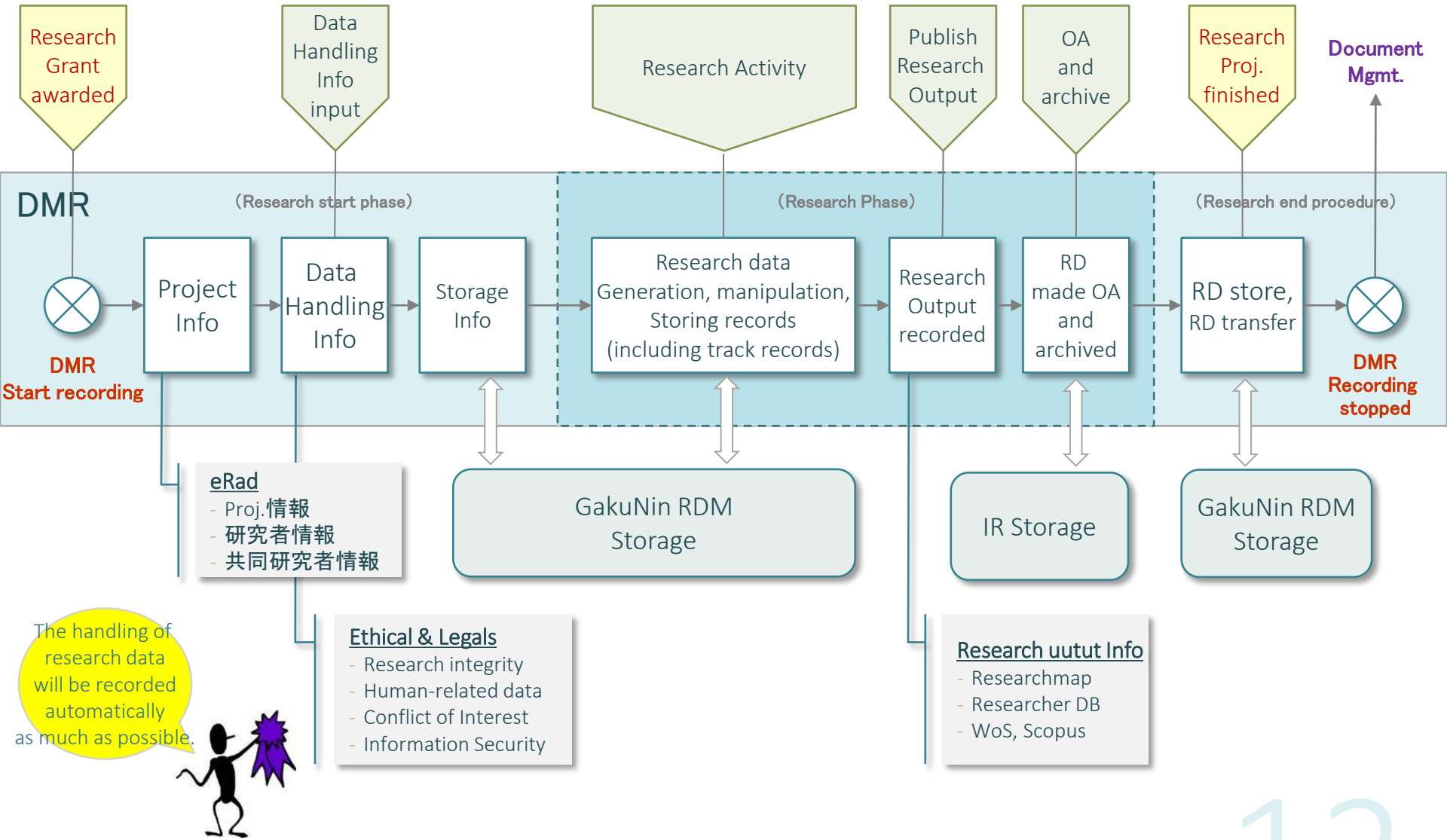
- Shares DMP, DMR, and research data with the institution
- Shares research data as open as possible

Institute responsibility

- Institutional RDM and compliance
- Providing infrastructures for RDM and data sharing



DMR—Records of Research Workflow



※ Inter-system data linkage should allow copying of data through single input.

How to align researcher workflow and institutional governance purpose

Institutions need minimum input on data management from researchers.

The data governance function, i.e., DMP/DMR tool, should work automatically to ease the burden on the researcher.

How can this be done?

1. Will various DMP templates for each discipline or research format work?
2. Is it possible to extract metadata for institutional governance purposes from research activity data?
3. Is it possible to develop a general data governance function, i.e. DMP/DMR tool, which is machine-actionable?

Governance Session 2:
Institutional Context

Why do we need institutional data governance?

- Currently, research data is managed by the researcher only and the institution does not have any clue what data the institution has, how it is managed etc.
- However, once something bad happens, such as research fraud, data leakage, the institution is the one who needs to take responsibility.
- In order to be accountable to the public, the institution needs to know what data the institution has and needs to manage them at all times.

What needs to be in place for institutional data governance?

- What data the institution will be responsible of.
- On what principles the data will be managed.
- The procedure of how the data will be managed.
- The organizational structure of the institution to be responsible to data.
- The handling of data in case the researcher leaves institution.

NII Research Data Policy (details)

NII-RDP: Title

Policy on Research Data Management and Sharing at NII

～Towards responsible management and sharing of research data

1. Purpose

This policy intends to promote the following two purposes in management and sharing of research data:

1. Responsible management and sharing of research data
2. Inheritance and development of scholarship

Research data to which the institution is responsible

Research data to be managed by the institution,
i.e., research data to which institution is responsible

The institute is responsible for the handling of all research data within research projects conducted under the name of the institution.

Research data to be published by the institution,
i.e., research data to be published on IR

The institute will make the greatest effort to make research data available as much as possible taking into consideration the sensitivity of data and potential to promote scholarship.

① Research data generated newly in the institution, i.e., original data

② Databases created and maintained by the institution

Principles and considerations in managing and publishing research data

PRINCIPLES

A Compliant to law and contracts, accountability to research data

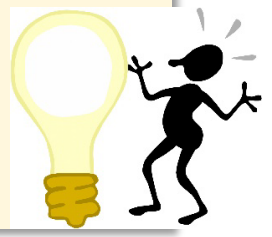
B Proper RDM in:

① Information security
② Protection of rights and interests of RD and related people
③ Adherence to the conditions of data provider
④ Making clear the various contributions to RD within research projects
⑤ Setting the conditions of RD publishing and reusing

C Research integrity and reproducibility

D Long term preservation and publishing of RD as much as possible

Points clear to which institution needs to hold!



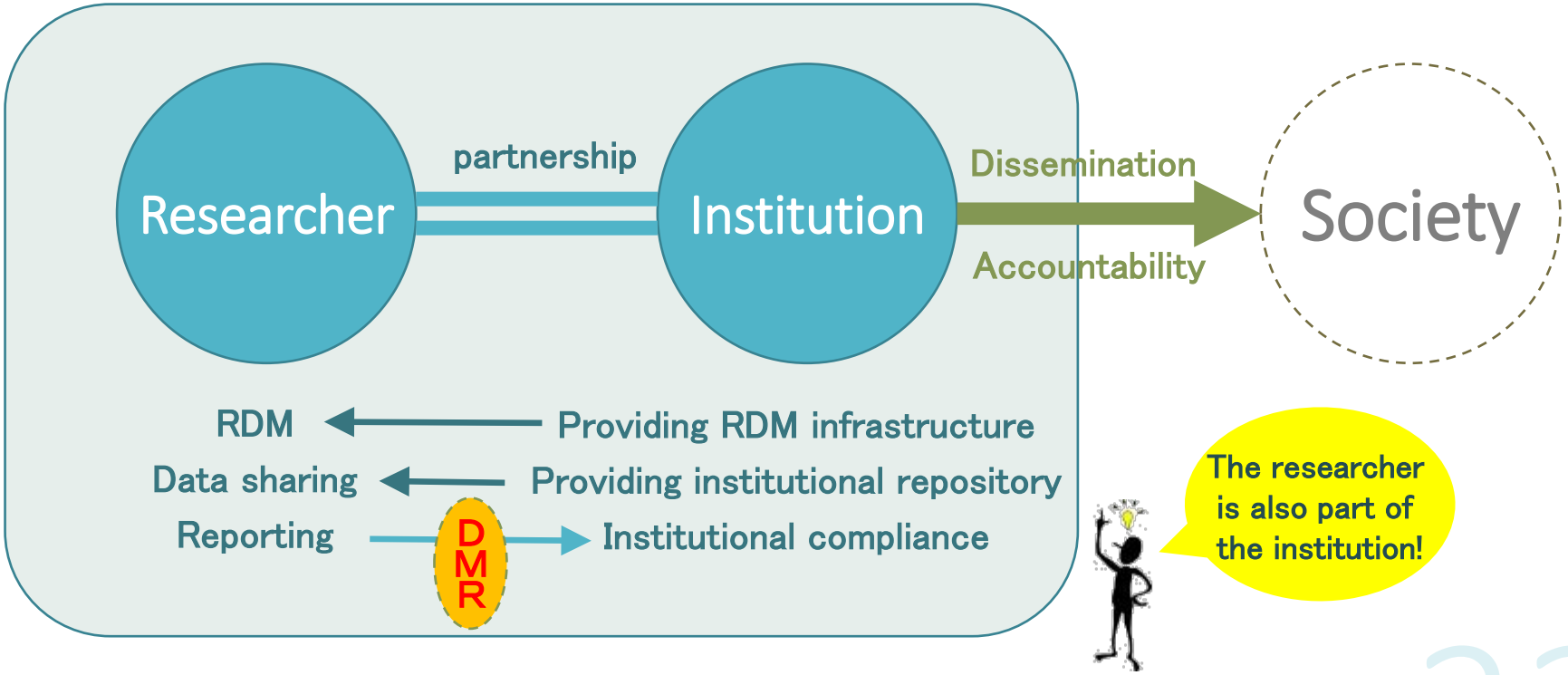
Considerations

A Promoting research—taking into consideration the burden on researcher in RDM.

B Cooperation and trust between institution and researcher

C Confidentiality and interests of research data, i.e., open & closed strategy

Institutional research data management through partnership between institution and researcher



8. Organizational structure to be responsible to RD

① Chief Research Data Officer (CRDO) :
Vice president for research

② Responsibility within research :

- Researchers are responsible in management and sharing
- PIs are responsible for management of research project and RD within project

③ Responsibility of strategic planning and devising framework for RDM environment

③-1
Responsible for strategic planning and research support: URAs

③-2
Responsible for RDM administration: Research office director

③-3
Responsible for RDM platform:
ICT center director

③-4
Responsible for IR:
Library director

Service unit for each function to be devised by each institution

Implementation of RD management and sharing

Framework for RD mgmt & sharing

MANAGEMENT of RD

Manage the following for each research project:

i. DMP (data mgmt. plan)
ii. Research data
iii. DMR (data mgmt. record) 11

SHARING of RD

Publish the following RD on IR as much as possible:

Research data generated newly in the institution, i.e., original data 12

PROVISION of INSTITUTIONAL DBs

DBs created by the institution should be:

i. In principle made available to the public.
ii. Promoted for re-use. 13

14 Environment for RD mgmt & sharing

14.1 Overview of environment
14.2 Environment for legal compliance and accountability
14.3 Environment for information security
14.4 Environment for protection of rights and interests of RD and related people
14.5 Environment for proper RDM
14.6 Environment for research integrity and reproducibility
14.7 Environment for long term preservation and publishing of RD as much as possible

Procedure in case of researcher leaving the institution

(Institutional RD management)

① Institution preserve and manage RD and DMP/DMR for defined time period after researcher leaves institution.

② RD and DMP/DMR may be managed by the new institution researcher moves to 1) upon the application and approval from researcher to the institution and 2) upon approval by the new institution researcher moves to.

(The RD use by researcher after leaving)

③ Researcher can take the copy of RD and DMP/DMR to the new institution upon the approval of dean the researcher belongs to and 2) upon the approval of institution given there is no restriction to do so through contracts or other condition set to the RD.

④ Institution will recommend researcher to publish RD generated or used by the researcher while working at the institute.

How to define Research Data Governance

1. How should institution manage and take responsibility on research data which is under control of researchers?
2. Is the concept of “ownership” on research data necessary? If so, how should it be?
3. How do you get the researchers collaborate on institutional research data governance?

RCOS

NII Research Data Policy and Institutional Data Governance (Background)

NII research data policy planned to be for Japanese universities

- Why NII works on institutional research data policy for Japanese universities
  - Need to position NII RDC in Japanese universities' context
  - Needs to embed NII RDC, especially GakuNin RDM, into researcher's workflow
- Strategy
  - Draft NII research data policy and provide it as policy template to Japanese universities
  - Use "data governance" as the link between the policy and NII RDC

University of Oxford
Policy on the Management of Data Supporting Research Outputs…ポリシーの目的

1.2 Purpose of policy

1.2.1 There is an increased requirement on researchers to manage and preserve their research data and to share it with as few restrictions as possible, while at the same time respecting concerns in relation to privacy, safety, security and commercial interests; this is under the assumption that the outputs of University research are a public good, produced in the public interest and for societal benefit. This policy aims to establish the measures needed to facilitate the appropriate curation and management of data, to secure its longevity and its potential to be shared.

1.2.2 This policy is intended to help promote good practice around research data, with the particular aim that it is:

a. stored securely and preserved in order to ensure its continuing utility;
b. appropriately identifiable, retrievable, and available when needed;
c. an accurate, complete, reliable and coherent representation of the materials collected;
d. kept in a manner that is compliant with legal obligations, including the Data Protection Act 1998 / The General Data Protection Regulation (GDPR) (Regulation (EU) 2016/679) and the Freedom of Information Act 2000, and, where applicable, the requirements of funding bodies and project-specific protocols approved under the University Policy on the Ethical Conduct of Research Involving Human Participants and Personal Data;
e. able to be made available to others in line with appropriate ethical, data sharing and open access principles, especially when the data underpins published research.

https://researchdata.ox.ac.uk/university-of-oxford-policy-on-the-management-of-data-supporting-research-outputs/

University of Oxford
Policy on the Management of Data Supporting Research Outputs…ポリシーの目的

2.0 Responsibilities of the University

2.1 The University will provide:

a. sufficient support, advice and information to researchers on all aspects of research data management
b. suitable infrastructure to facilitate the appropriate management of research data.
c. relevant training to promote best practice in research data management.

2.2 In order to support and enable this policy, the University is committed to engaging with researchers to fulfill the mutual responsibilities described herein, and through this engagement, ensure that the research data management services and infrastructure capabilities it provides are adequate and fit for the purpose of supporting researchers' compliance with the policy.

2.3 In the event that a research data management requirement cannot be addressed or treated via the advice, infrastructure and training noted above, then the University will be responsive to requests for bespoke support to ensure that researchers can meet their responsibilities under this policy.

3.0 Responsibilities of the researcher

3.1 Principal Investigators hold day-to-day responsibility for the effective management of research data generated within or obtained from their research, including by their research groups. This shall include understanding and complying with the requirements of any relevant contract with or grant to the University that includes provisions regarding the ownership, preservation and dissemination of research data.

3.2 Researchers will protect confidential, personal and sensitive personal research data in accordance with legal and ethical requirements related to the research they conduct.

3.3 Researchers will make every reasonable effort to keep an accurate and comprehensive record of their research, including documenting clear procedures for the collection, storage, use, reuse, access and retention or deletion of the research data associated with their records. Where appropriate, this should include defining and documenting protocols and responsibilities in collaborative research projects.

https://researchdata.ox.ac.uk/university-of-oxford-policy-on-the-management-of-data-supporting-research-outputs/

HARVARD UNIVERSITY
Policy Title: Research Data Ownership Policy

Responsible Office:	Office of the Vice Provost for Research
Effective Date:	February 19, 2019
Revision Data:	February 15, 2019

INTRODUCTION

Ultimately, the University is responsible for complying with laws, regulations, and requirements of its research sponsors, many of which pertain to research data. To ensure its ability to satisfy those requirements, the University asserts ownership over research data generated at Harvard for projects conducted at the University, under the auspices of the University, or with University resources. Although the University is the owner of all such research data, sound management practice and common-sense call for the University and researchers to work in partnership to fulfill these obligations. This policy defines "research data," assigns roles and responsibilities to key actors, and describes its relationship to other relevant University policies.

POLICY AND PROCEDURES

1. OWNERSHIP AND ROLES

A. The University: The University asserts ownership over research data for all projects conducted at the University, under the auspices of the University, or with University resources.

B. Principal Investigators: Principal Investigators (PIs) and other researchers are stewards and custodians of research data. However, if PIs choose to delegate responsibility within their research groups, the PIs remain accountable to the University for the stewardship of research data.

https://cpb-us-e1.wpmucdn.com/websites.harvard.edu/dist/6/18/files/2020/07/data_ownership_policy_08.06.19.pdf

2. RESPONSIBILITIES

Harvard's responsibilities with respect to research data include, but are not limited to:

i. Ensuring compliance with the terms of research agreements;
ii. Protecting the rights of researchers, including, but not limited to, their rights to access to data from research in which they participated;
iii. Securing intellectual property rights;
iv. Facilitating the investigation of charges such as research misconduct or financial conflicts of interest;
v. Maintaining appropriate confidentiality and security protections over research data; and
vi. Complying with applicable federal, state, and local laws and regulations.

The University's rights and obligations are not subject to negotiation and may not be altered in any agreement or proposal prepared by any faculty member or administrator.

The PI's responsibilities with respect to research data include, but are not limited to:

i. Ensuring proper management and retention of research data in accordance with this policy and the Guidance on Retention and Records of Research Records and Data.
ii. Establishing and maintaining appropriate procedures for the protection of research data and other essential records, particularly for long-term research projects;
iii. Ensuring compliance with program requirements;
iv. Maintaining confidentiality of research data, where appropriate;
v. Maintaining appropriate data use agreements for the sharing of research data, and
vi. Complying with applicable federal, state, and local laws and regulations.

3. DATA RETENTION

Harvard's policies and the retention of research data and materials are set forth in the Guidance on Retention and Records of Research Records and Data.

4. TRANSFER IN THE EVENT A RESEARCHER LEAVES HARVARD

If a PI leaves Harvard and a project is to be moved to another institution, ownership of the original data may be transferred from Harvard to the PI's new institution upon request from the PI subject to: (a) the prior written approval of the Vice Provost for Research; (b) written agreement from the PI's new institution that guarantees (1) its acceptance of ongoing custodial responsibilities for the data and (2) Harvard having access to the original data, should such access become necessary for any reason; and (c) relevant confidentiality restrictions, where appropriate.

When individuals, other than PIs, who have been substantively involved in research projects at Harvard leave the University, they may take with them copies of research data resulting from these projects, subject to relevant confidentiality restrictions and any requirements of the original research project, and conditioned upon the approval of the individual's Department Chair or Dean. In this event the ownership of the research data remains with the University and original data must be retained at Harvard by the PI.

In either of these instances, the remaining members of the research team retain the rights to use the original data. Any publications resulting from the data will be subject to the Harvard University Authorship Guidelines as well as the specific requirements of the journal in which publication occurs.

https://cpb-us-e1.wpmucdn.com/websites.harvard.edu/dist/6/18/files/2020/07/data_ownership_policy_08.06.19.pdf

How to spread NII Research Data Policy nation-wide

1. Too detailed and difficult to get the buy in of Japanese universities
   - Should we use the simple policy version?
   - Will two layered "basic policy" and "implementation policy" work?
   - What if we have a success story, i.e. a university adopts NII policy?
2. Does one research data policy fit all? What are the institution-intrinsic needs on research data policy?

Australian Case

豪州大学の学内DMP導入の経緯

1. 豪州研究助成機関が、研究者ではなく、機関に対して「責任ある研究実践」を求めた。
   - Australian Code for the Responsible Conduct of Research 2018
2. また、付随して、機関における研究データ管理のガイドを提示した。
   - Management of data and information in research
3. これらにDMPは言及されていないが、一部の大学はDMPを機関内に導入することにより、機関のデータガバナンスを構築しようとしている。
4. なお、現状では多くの場合、研究者ではなく、（研究開始の条件として）大学院生にDMP作成を義務化している。

https://www.nhmrc.gov.au/about-us/publications/australian-code-responsible-conduct-research-2018

Management of data and information in research

—a guide supporting the Australian Code for the Responsible Conduct of Research

2. Responsibilities of institutions

- 2.1 Provision of training for researchers
- 2.2 Ownership, stewardship and control of research data and primary materials
- 2.3 Storage, retention and disposal
- 2.4 Safety, security and confidentiality
- 2.5 Access by interested parties
- 2.6 Facilities

3. Responsibilities of researchers

- 3.1 Retention and publication
- 3.2 Managing confidential and other sensitive information
- 3.3 Acknowledging the use of others' data
- 3.4 Engagement with relevant training

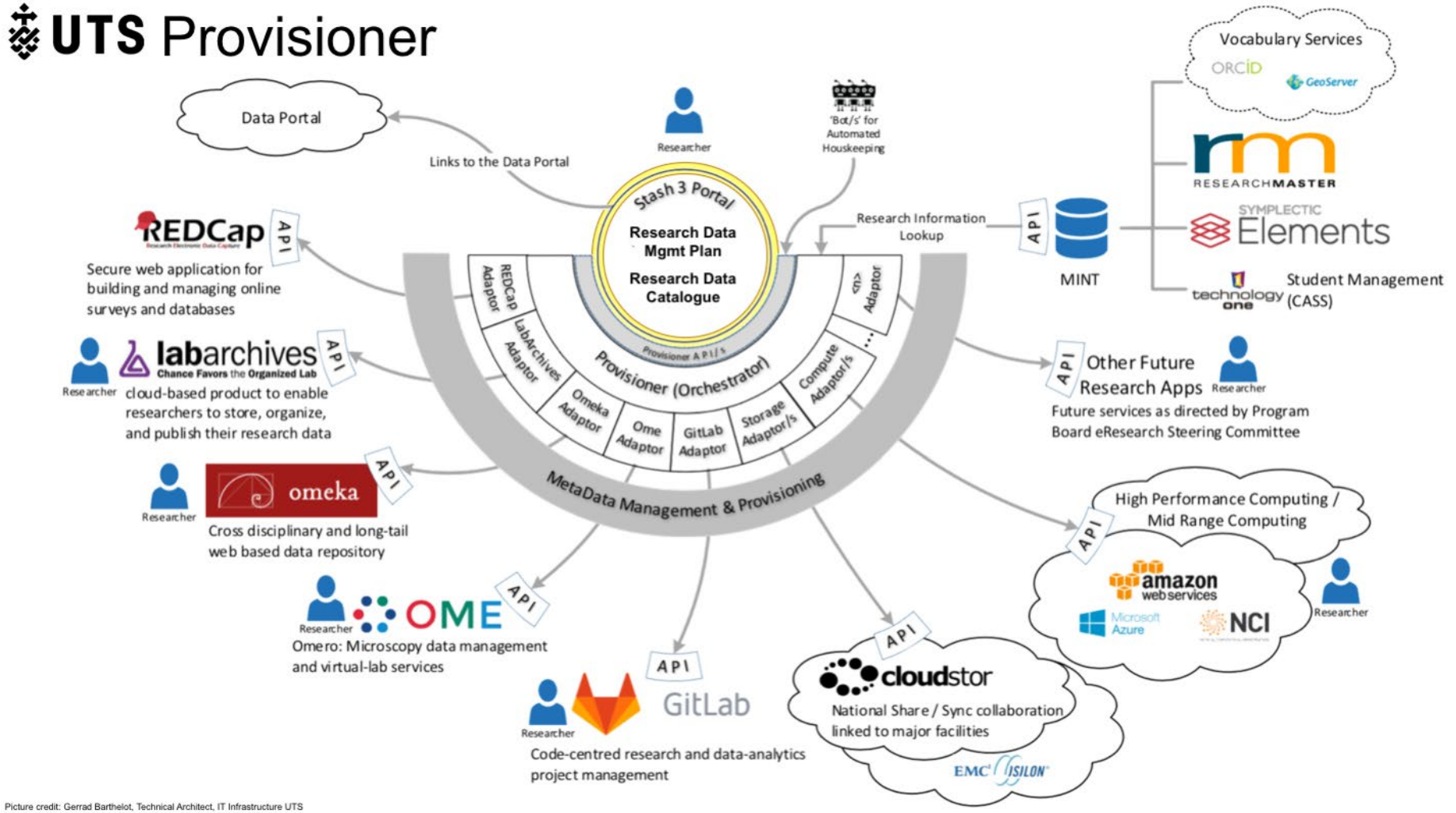
https://www.nhmrc.gov.au/about-us/publications/australian-code-responsible-conduct-research-2018

DMP required at University of Technology Sydney (UTS)

- Project
- People
- Ethics and Security
  - Information Security Classification (public/internal/sensitive/confidential)
  - Research involves:
    - ✓ *Human participant data*
    - ✓ *Use of animals*
    - :
  - Ethics approval needed (y/n)
- Data Collection and Storage
  - Data collection methodology
  - File type, file format
  - Storage location
- Data Retention and Disposal
  - Minimum retention period
  - Data steward
  - Commitments to destroy data prior to end of retention period
- Access and Rights
  - Copyright and intellectual property owners of data created in project
  - Access after the project
  - Use of secondary or third-party data
  - Attach ethics approval, licenses
- Research Workspace

https://www.lib.uts.edu.au/services/services-for-researchers/manage-your-information-and-data
https://www.youtube.com/watch?v=5uuvhPdcsqs&t=128s

UTS's RDM Platform linked to various other systems



Picture credit: Gerrad Barthelot, Technical Architect, IT Infrastructure UTS

https://eresearch.uts.edu.au/2018/07/04/APRI_2018_provisioner/index.html
https://zenodo.org/record/4817188#.YkWdbeczYQ8

UTS Storage options

Data workspaces for active research	Office 365 OneDrive	eResearch Store	CloudStor	Omero, git.research.uts.edu.au	eNotebooks	REDCap, Qualtrics	Limesurvey
	Storage			Repository		Data collection tool	
Suitable data classifications	C S i p	C S i p	S i p	C S i p	S i p	C S i p	C S i p
Stored in Australia? [1]	NSW or Vic	NSW	Australia	NSW	Australia	NSW	NSW
Mobile app available?	✓	✗	✓	✗	✓	✓ [2]	✗
Can restore user data? [3]	✓ (60 days)	✓ (14 days)	✗	✗	No deletion	✓ (30 days)	✗
Storage limit?	1TB [4]	On request [5]	1TB [6]	Unlimited	Unlimited [7]	n/a	n/a
Version control?	✓	✗	✓	✓	✓	✓	n/a
External Collaboration?	✓	✗	✓	✗	✓	✓	✗

Archival Storage

The data workspaces above are for when your research project is **in progress**. Before you use them create a **Stash RDMP** (Research Data Management Plan).

At the end of your project you should archive your data by creating an **Archival Data Record** in **Stash**. You can upload data within the **Stash** interface.

You can also use **Stash** to publish research data, providing it is UTS public, to **data.research.uts.edu.au**.

Please contact **eResearch-it@uts.edu.au** if you have any questions or need help.

[1] Data jurisdiction is important in the case of personal and health data due to privacy legislation. Health records and information should ideally be stored in NSW.

[2] REDcap is not available as an app but it does support app-based surveys for mobile platforms.

[3] All options above has disaster recovery and backup/replication, but not all allow restoration of a single user's data. Therefore we recommend you also make backups.

[4] Individual file size up to 10GB. You can apply for more storage in ServiceConnect.

[5] Allocation is based on justifiable needs.

[6] You can apply for an increase in storage in ServiceConnect.

[7] Individual file size up to 250MB but unlimited total storage.

https://www.lib.uts.edu.au/services/services-for-researchers/manage-your-information-and-data
https://www.lib.uts.edu.au/sites/default/files/2020-11/research_data_storage_options_for_uts_researchers.pdf