

(Annex 10)

## **Basic Policy on Data Management in BRAIN-commissioned Moonshot Research and Development Project**

As research data generated by research activities funded by public funds is common intellectual property shared by all citizens, while taking into consideration the characteristics of research content, it is necessary to manage by developing appropriate rules and infrastructure such as data sharing policies, data management plans, repositories, and others. In moonshot research and development, the Bio-oriented Technology Research Advancement Institution, National Agriculture and Food Research Organization (hereinafter referred to as the "BRAIN") promotes advanced data management based on the Operation and Evaluation Guidelines for the Moonshot Research and Development System (February 2020).

This basic policy establishes the basic data management principles in this project in line with this intent. Based on this basic policy, the Program Director (PD) and Project Manager (PM) will promote advanced data management, such as promoting information exchange among researchers and encouraging the storage, sharing, and disclosure of research data. In addition, the status of storage, sharing, and disclosure of research data will be subject to project evaluation.

Matters not described in this policy shall be determined as necessary by agreement between the PM and the project participants, in response to the purpose of each project.

### 1. Definition of terms used in this policy

#### (1) Research data

All data generated in the process of research and development that can be managed in electromagnetic form. This includes research notes and memos, data directly obtained from experiments, observations, simulations, and others, data processed from them, data that serves as evidence for papers, conference materials.

#### (2) Data subject to management

The data subject to management is the data that serves as evidence showing the results of research and development activities among the research data generated by the implementation of this project, and the PM shall determine the scope. Data subject to management is updated at the PM's discretion in accordance with the progress of research and development.

#### (3) Data to be unmanaged

This refers to research data other than the data subject to management.

#### (4) Metadata

This refers to data composed of information for explaining the data rather than the data subject to management itself. Metadata includes information such as data name, data description, data manager and contact information, data location, data storage/sharing/publication policies, and so on.

- (5) Data Management Plan (DMP)  
A plan that shows the concept of management and utilization of the data subject to management is called a Data Management Plan (DMP).
- (6) BRAIN specified data  
Among the research and development data, it refers to the data that should be managed by the consignor and designated as research and development data to be provided to the consignor.
- (7) Voluntary management data  
This refers to R&D data other than BRAIN-designated data that is voluntarily managed by project participants.

## 2. Basic items of research data in this research and development

Regarding the voluntary management data, the PM and project participants should primarily determine the management policy, but based on the open/closed strategy, clarify the categories such as storage, sharing, disclosure, and others, and research data to the extent necessary shall be made public, and efforts shall be made to promote utilization by the project participants themselves or provision to others for various purposes and uses.

## 3. Matters promised by BRAIN and the trustee

- (1) Creation and submission of Data Management Plan (DMP)  
After the start of the project, as of when it is assumed that data will be acquired or collected, the PM shall have the project participants create a draft of the Data Management Plan (DMP) related to the data subject to management, compile it, and submit it to BRAIN in a form separately designated by it.
- (2) Additional submission/correction of Data Management Plan (DMP)  
The PM shall appropriately update the scope of the data subject to management according to the progress of research and development, add or modify the Data Management Plan (DMP), and submit it to BRAIN in a form separately designated by it.
- (3) Submission of metadata  
After having the project participants create metadata according to the Data Management Plan (DMP), the PM aggregates the metadata and submits it to BRAIN in a form separately specified by it. It is assumed to be consent that the submitted metadata will be included in the public, shared and private data catalog to be created by BRAIN. Metadata is created for all the data subject to management.
- (4) Disclosure, and others of the data subject to management based on the open/closed strategy  
PMs and project participants store and share the data subject to management and metadata, and disclose them to the extent necessary, in accordance with the Data Management Plan (DMP). In the future, as soon as preparations for external use of NARO Integrated DB \*1 are completed, BRAIN is scheduled to

recommend PMs and project participants, who wish it, to store the data subject to management and metadata after understanding its advantages and disadvantages.

(5) Development of data management system

In accordance with this policy, to properly manage the data subject to management, the Intellectual Property Committee \*2 (Refers to the Intellectual Property Committee in II-2-(1) of Moonshot Research and Development Program for Agriculture, Forestry and Fisheries Implementation Outlines (December 14, 2020, No. 0909003), the same shall apply hereinafter) shall be given data management functions. The Intellectual Property Committee shall promote advanced data management by setting data management policies for Goal 5, designating the BRAIN-designated data, exchanging information among researchers, and promoting storage, sharing, and disclosure of research data.

4. Matters to be determined between project participants

(1) Development of a data management system in the Consortium

In accordance with this policy, to properly manage the data subject to management, the Intellectual Property Steering Committee for each consortium \*3 (Refers to the Intellectual Property Committee in II-2-(2) of Moonshot Research and Development Program for Agriculture, Forestry and Fisheries Implementation Outlines (December 14, 2020, No. 0909003), the same shall apply hereinafter) shall be given data management functions. The Intellectual Property Steering Committee shall consider research data utilization policies, identify research data to be managed, examine open/close strategies such as clarifying categories such as storage, sharing, and disclosure of the data subject to management, and coordinate conditions for data usage permission, and so on. The results of these coordination shall be summarized in Intellectual Property Form 6 and submitted to BRAIN.

(2) Prior approval for disclosure of research data of this project to third parties \*4

Of the research data obtained or collected through the implementation of this project, the data subject to management should not be disclosed or leaked to third parties other than the project participants without obtaining the approval of the Intellectual Property Steering Committee. However, with regard to research data approved by the Intellectual Property Steering Committee, efforts shall be made to promote wide-ranging utilization. The policy for each data shall be aggregated in Intellectual Property Form 6.

(3) Management and licensing of research data

PMs and project participants manage and utilize the data subject to management in accordance with the Data Management Plan (DMP).

Licensing of the data subject to management shall be made in accordance with the Data Management Plan (DMP). If discussions between the parties regarding the scope of research data to be licensed, licensing fees, and other matters are difficult, and there is a risk of hindering the commercialization of the results of

the project, the Intellectual Property Steering Committee shall make coordination and a reasonable resolution shall be sought between the parties.

- (4) Licensing of research data for research and development during the project period or for commercialization of project results  
Project participants shall, to the extent necessary, as a rule, grant licenses free or reasonable compensation for research and development activities by other project participants within the project period, or for activities to commercialize the results of the project. (This does not apply if there are separate agreements, such as licensing of the data subject to management for a fee, between project participants.) However, if licensing of the research data is expected to affect the existing or future project activities of the party who granted the license, the party can be refused. In addition, the scope of exceptions (especially research data brought in by project participants for the purpose of implementing the project) shall be clarified to the extent necessary based on an agreement between the project participants.

\*1 The National Agriculture and Food Research Organization (NARO) integrated database is designed to manage and preserve various research data related to agriculture, promote cross-disciplinary research through sharing and utilization among researchers safely and permanently. It is built for the purpose of advancing and speeding up data-driven agricultural research through artificial intelligence (AI) analysis support and promoting "smart agriculture" by providing data to the agricultural and industrial worlds. NARO Integrated DB has the following features.

- 3PB storage with up to 3 available backups and automatic version control
- Protection of research data by thorough network security measures
- Give customizable metadata to all data to visualize and catalog research data (primary DB)
- Various databases can be built in a format suitable for AI analysis (secondary DB)
- It is possible to freely set viewing and editing rights to data on an individual or group basis.
- Coordination function with external data disclosure system and NARO AI supercomputer (separate application required)
- Linkage function with agricultural data linkage platform WAGRI (separate application required)

As a result, users of the NARO Integrated DB can safely integrate and store data obtained in research and development, and the obligation and burden of data storage by researchers themselves can be reduced. In addition, by equipping appropriate metadata, the existence of useful research data can be visualized, and new research knowledge can be obtained through metadata searches and full-text searches. In addition, flexible access control enables appropriate sharing of data (access rights to data can be set by registrants).

NARO Integrated DB will be operated free of charge for the time being. The timing of the fee-based service is undecided, but if it becomes a charged service, we plan to notify you in advance with sufficient time to spare. In the future, we are

considering archiving the data if the fee will be unpaid for a certain period when the fee starts to be charged. Archived data can no longer be accessed online, and backups, and others are not performed. In addition, it may be deleted without notice after a certain period. The use of the NARO Integrated DB by persons other than NARO employees is scheduled to become possible after the terms of use are established. We will let you know when it is ready for use.

\*2 The Intellectual Property Committee consists of PD, PM, related ministries and agencies, and experts, and its duties are to publish papers on research and development results and to apply and maintain policies for intellectual property rights (patent rights, breeder's rights, and others). In addition to making decisions, and so on, it also coordinates licensing of intellectual property rights as necessary.

\*3 An Intellectual Property Steering Committee shall be established for each Consortium. The committee shall be chaired by the PM and shall consist of collaborative research institutions related to the subject intellectual property rights and may include external experts who have agreed to comply with the terms of Intellectual Property Rights Agreement regarding confidentiality as necessary. Missions related to research data include examination of open/close strategies, such as clarification of categories such as storage, sharing, and disclosure of the data subject to management.

\*4 When providing research data containing personal information to others, it is necessary to obtain the consent of the individual and process it so that it cannot identify specific individuals. And it should be noted that the project participants need to fully consider laws, regulations, guidelines, and others in the processing concerned. In addition, note that, in managing the data subject to management, to receive protection under the Unfair Competition Prevention Act, the data subject to management must be managed as a "trade secret" under the Unfair Competition Prevention Act.

(End of basic policy)

**(Reference 1)** Operation and Evaluation Guidelines for the Moonshot R&D Program (Moonshot Goals 1 to 6) (February 4, 2020 (Partially revised March 4, 2020, Partially revised January 18, 1998) Director General of Policy, Cabinet Office (Science, Technology and Innovation), Director General of Science and Technology Policy Bureau, Ministry of Education, Culture, Sports, Science and Technology, Director of Agriculture, Forestry and Fisheries Research Council Secretariat, Ministry of Agriculture, Forestry and Fisheries, Director General of Industrial Technology and Environment Bureau, Ministry of Economy, Trade and Industry) (excerpt)

1. Features of the systems

- To promote the utilization of research data generated by research activities and support for advanced research management based on it, advanced data management, such as utilizing the research data infrastructure system (NII Research Data Cloud) shall be promoted.

### 3. Promotion Systems for R&D

[Research promotion institutions]

- The research promotion institutions have the following duties.
  - 13 To promote advanced data management, manage a data catalog consisting of metadata submitted by PMs and researchers by using the research data infrastructure system, and so on.
- In carrying out its duties, the research promotion institutions shall work in cooperation with related ministries and agencies and other research promotion institutions and strive to create an environment where researchers can concentrate on research by improving the efficiency of evaluation work. In addition, towards ensuring the fairness of research, we shall make efforts to prevent misconduct in research activities and misuse of research funds at consignees, and others, while utilizing advanced data management.

[PM]

- The PM has the following duties.
  - 4 Formulate a Data Management Plan (DMP) that defines such as the scope of the data subject to management and based on this, collect the metadata of the data subject to management from researchers and submit it to the research promotion institution. In addition, using the research data infrastructure system, and others, we shall store, share, and disclose the data subject to management to the extent necessary.

### 4. Implemental method for research and development

[Implementation of research and development]

- PDs and PMs promote advanced data management, such as exchanging information between researchers and encouraging storage, sharing, and disclosure of research data.

**(Reference 2)** Implementation Guidelines for Moonshot Research and Development Program for Agriculture, Forestry and Fisheries (excerpt)

## VI Handling of Intellectual Property Rights

### 22 Handling of research results

#### (2) Data management

As research data generated by research activities funded by public funds is common intellectual property shared by all citizens, while taking into consideration the characteristics of research content, from the perspective of promoting open science, creating appropriate data sharing policies, data

management plans, repositories, and others, it is necessary to manage and publish the data.

To promote the advanced data management, BRAIN shall manage data catalog consisting of metadata submitted by PMs and researchers, such as utilizing the research data infrastructure system (NII Research Data Cloud).

As a PM's duty, we formulate a data management plan (DMP) that defines the scope of the data subject to management, and others, and based on this, it is necessary for the PMs to collect the metadata of the data subject to management from research implementers and to submit it to BRAIN. In addition, PMs shall create a data sharing policy and store, share, and disclose the data subject to management to the extent necessary using the research data infrastructure system, and others.