

## **Basic Plan for Digital Preservation of National Diet Library's Collections for 2021-2025**

### **1 Background**

In recent years, the National Diet Library (NDL) has come to hold more materials in digital form. The NDL has been working on digitization of its collections continuously since 2000. In addition to this, in the same year, the NDL started collecting “packaged electronic publications”, in which information is fixed in a tangible media such as CD-ROMs and DVDs, through the legal deposit system. Furthermore, in response to the development of the internet, the NDL has been involved in the project of collecting and preserving internet information since 2002, as well as having started institutionally collecting and preserving internet information disseminated by public institutions such as the government in 2010. Since 2013, the NDL has been collecting and preserving materials such as privately published e-books and e-magazines which are free of charge and without Digital Rights Management (DRM, technologically restrictive measures) through the E-legal Deposit of Online Publications system. It has also been considering collecting materials which are paid for or which have DRM.

However, materials in digital formats have problems such as the fragility of media, difficulty in obtaining playback equipment and obsolescence of technologies such as playback software. It is therefore necessary to take measures to ensure proper management and long-term availability of materials in digital formats.

Since 2017, the NDL has conducted a trial of preservation measures for packaged publications, and a study of the deterioration of optical discs. Based on the results of the trial, etc., the NDL is proceeding with relevant efforts.

### **2 Positioning of this plan**

This plan sets out the basic approach to be taken in relation to the preservation of materials in digital formats held by the NDL, based on the background set out in 1 above and on *Preservation of the National Diet Library's Collections* (国图収 No. 37 in 2003). In this plan, “preservation” means long-term preservation, unless otherwise stated, to ensure long-term availability of materials.

When specific measures are taken based on this plan, appropriate plans, etc., shall be prepared in advance as necessary.

The planning period of this plan shall be until FY 2025, and it will be revised as necessary.

### **3 Purpose**

As a national library, the NDL plays the role of collecting and preserving domestic publications as Japan's cultural properties and intellectual resources as well as passing them on to future generations. In order to

fulfil this role, the NDL will take the necessary measures to preserve its packaged electronic publications, internet materials, online materials, digitized materials, etc. (“digital materials”) and ensure their long-term availability.

#### **4 Materials covered by this plan**

Digital materials which are covered by this plan shall be as follows. See the Appendix for the number and volume of each material.

(1) Packaged electronic publications

Electronic publications, such as floppy disks, CDs, DVDs, Blu-ray discs, etc., collected by the NDL in the form of physical media under the legal deposit system, etc.

(2) Internet materials/Online materials

Internet materials such as websites collected by the NDL through the Web Archiving Project (WARP), and online materials collected through the E-legal Deposit of Online Publications system.

(3) Digitized materials, etc.

Digitized materials<sup>1</sup> (including data for preservation) created by the NDL or collected from other institutions, etc. and digital deliverables<sup>2</sup> whose contents are equivalent of materials held by the NDL, but not falling under (1) or (2), and copies of (1), which were prepared by the NDL for preservation.

#### **5 Basic guidelines**

(1) The NDL will promote organization-wide efforts to prevent data corruption and loss of digital materials as well as ensure their long-term availability.

(2) The NDL will consider media’s fragility, difficulty in obtaining playback equipment, and obsolescence of technologies such as playback software as important issues, which are related to fundamental roles of the NDL, and will work to address them.

(3) The priority of the digital materials for implementing preservation measures shall be decided comprehensively based on the following judgment criteria. In addition to decisions based on media and material types, preservation measures may be implemented on an individual basis, depending on state of deterioration, etc.

① Urgency of preservation depending on obsolescence of playback environment as well as fragility and deterioration of media

Materials that are difficult to obtain due to obsolete playback equipment/software, materials that are vulnerable due to characteristics of media, or materials that have severely deteriorated

② Uniqueness and rarity

Materials whose content does not overlap with other media or materials for which it is not possible

---

<sup>1</sup> Materials which are digitized from analog format in the form of images, audio, video, etc., for use as a substitute for analog media materials.

<sup>2</sup> Images, sound, video, text, etc. in digital formats.

to confirm whether they are held by other institutions

③ Social need for preservation

Materials that are socially needed to be preserved for the long term, and that are appropriate for the NDL to take measures in light of its mission, role-sharing with other institutions, etc.

④ Status of establishment of preservation measures and costs required for such measures

Materials for which technical measures for preservation and methods for providing access to have been established, as well as materials whose costs for implementing preservation measures are not excessive

(4) Based on the judgement criteria in (3), the NDL will implement preservation measures on packaged electronic publications first. As a precautionary measure, the NDL shall take into account preservation measures when preparing digitized materials, etc.

(5) In order to implement preservation measures, the NDL will select the most appropriate measure for media conversion<sup>3</sup>, file format conversion, etc., based on the characteristics of materials, the availability of playback environment and other circumstances.

In the preservation measures, media conversion, file format conversion, etc. are collectively referred to as “migration”.

(6) The NDL will keep a close eye on developments in international standardization in this field and will also participate in standardization activities to the extent possible. The NDL complies with international and other existing standards as much as possible.

(7) The NDL will be actively involved in discussions on preservation of digital materials and strive to raise awareness of the necessity of this initiative. The NDL will publicize its efforts widely.

## 6 Preservation measures

The NDL will implement the following measures to ensure preservation and availability of digital materials. The NDL will decide specific measures in the light of the research results, which are described in 7 below.

(1) Condition check of materials

The NDL will continuously check the condition of its collections<sup>4</sup> and assess the risks based on appropriate criteria. In particular, the NDL will check optical discs, which there are a large number of collections of and have a high degree of individual variation in deterioration, while examining checking methods and establishing checking systems.

(2) Metadata maintenance

The NDL will preserve metadata required for preservation together with contents to be preserved. In addition, the NDL will proceed with maintenance of desirable metadata (including metadata related to original materials before preservation measures are taken) while considering data elements, data storage

---

<sup>3</sup> Media conversion includes transfer to the same type of media.

<sup>4</sup> If migration is implemented by medium, migrated materials will be included in the collections.

methods, maintenance systems, etc.

(3) Choosing appropriate file formats

The NDL will choose appropriate file formats from the perspective of long-term preservation when preparing digital materials (including digitization of analog materials, content creation in digital formats such as online materials, etc.). When choosing a file format, the NDL will choose a widely-used and standardized format to the extent possible.

(4) Implementation of migration, etc.

The NDL will implement measures such as migration in earnest. If it is difficult to ensure reproducibility of materials after file format conversion, the NDL will give priority to ensuring data readability and will consider measures, such as emulation, which simulate the playback environment upon media conversion while maintaining the file formats of original materials.

(5) Development of preservation environment

With regard to digitized materials and the preservation of migrated data, the NDL will consider preservation media and methods which utilize standard technologies as well as develop reasonable and stable environment for preservation to prevent data corruption and loss.

(6) Maintenance of playback environment

The NDL will periodically check condition of the playback equipment to ensure its long-term availability and will repair or replace it, etc., as necessary.

(7) Development of operating environment, etc. after measures are taken

In order to provide access to data after measures (e.g., migration) are taken, the NDL will develop environments to that end such as system and hardware (including use through emulation, virtualization, etc.).

(8) Maintaining preservation environments of original materials, etc.

The NDL will preserve the following materials in an appropriate environment: original materials after migration and materials for which preservation measures have not been implemented.

(9) Publicize necessity of preservation and various information

The NDL will publicize the necessity of preservation and relevant information based on the results of efforts such as points to pay attention to when preparing digital materials from the perspective of preservation through the NDL's website and other public relations media as appropriate.

## **7 Technical research study**

In order to ensure long-term access to digital materials and promote specific preservation measures, the NDL will implement research studies on technologies that are considered to be highly applicable for the following matters in a planned way. Furthermore, in order to contribute to long-term consideration of preservation measures, the NDL will continue research on overseas technological developments, such as preservation systems.

- ① Checking methods of preservation media (optical discs, hard disks, etc.)
- ② Desirable file formats from the perspective of preservation
- ③ Metadata elements required for preservation and methods to implement them in preservation systems, etc.
- ④ Migration and emulation methods according to media and content (including implementation methods for a large number of materials)
- ⑤ Reasonable and stable preservation environment (media, systems, etc.), in light of standard techniques
- ⑥ Effective methods to provide migrated data

### **8 Collaboration, cooperation, and human resources development**

The NDL will send its staff to relevant conferences and events such as the International Conference for Digital Preservation (iPRES) and will continuously strive to collect information. At the same time, the NDL will promote information sharing and collaboration with relevant institutions, companies, etc. in Japan and abroad. Along with these efforts, the NDL will develop its staff's knowledge on preservation as well as improve its human resources and expertise in preservation.

### **9 Progress management, etc.**

Preservation of digital materials is a cross-departmental and organization-wide initiative at the NDL, and the Digital Information Department, the departments for collection, bibliographies, and material preservation, and departments for providing access to materials will collaborate to implement measures. The Digital Information Department will implement coordination and progress management of this plan.

With regard to research studies, the Kansai-kan will be in charge of the areas related to specific preservation measures, and the Digital Information Department will be in charge of the areas involving technical demonstration experiments, and both of them will cooperate closely to proceed with research studies.

(Appendix)

**The number of digital materials held by the NDL (as of the end of FY2019)**

	Type	Number of items (approximate)	Data volume
Packaged electronic publications	Magnetic disk (e.g., floppy disk)	12,400	-
	CD, DVD, Blu-ray Disc	873,400	
	Other kinds of optical disc (MD, etc.), semiconductor memory, etc.	4,600	
	Total	890,400	
Internet materials, online materials	Internet materials	177,000	1678.6 TB
	Online materials	1,247,000	1.5 TB
Digitized materials, etc.	Image	2,772,000	150.9 TB
	Audio/video (including DAISY)	75,000	
	Other (Braille data)	2,000	
	Total	2,849,000	

- The number of packaged electronic publications does not include extra copies, but does include appendices (combined materials, supplements, etc.).
- The number of internet materials indicates the cumulative preserved number of materials collected through WARP.
- The number of online materials and digitized materials indicates the number of items.

(Surveyed in July 2020)