

Survey of domestic institutions on the long-term preservation of digital materials (2021)

Executive Summary

(Research and Development for Next-Generation Systems Office, Digital Information Planning
Division, Digital Information Department, The National Diet Library)

1. Survey summary

The purpose of this survey was to ascertain the current status and issues relating to the long-term preservation of digital materials in libraries, museums, archives, universities, research institutions and local public bodies (prefectures and ordinance-designated cities) in Japan.

All public libraries, university libraries (including Inter-University Research Institute Corporations), registered museums, museum-equivalent institutions and archives, plus various institutions that met the survey objectives, were mailed a questionnaire on 2 November 2021, with responses requested by 10 December (online, email, fax, or post). After the distribution of the questionnaire, the final number of institutions surveyed was 5,409, following adjustments such as the identification of changes or discontinuation of institutions and the circulation of the questionnaire to relevant institutions within the same establishment. The number of valid responses was 2,921 (response rate 54.00%).

Next, based on the results of the questionnaire survey, interviews (face-to-face or online) were conducted from 12 January to 16 February 2022 with 12 institutions selected from among the respondents by type of institution, size of organization and specific initiatives.

2. Collection and possession of digital materials

- Around 70% (2,041) of all institutions responding to the questionnaire survey hold digital materials (Figure 2).
- The average number of digital materials held by each institution is 23,676 and the median is 3,000.
- The digital material most widely collected or possessed by the institutions that responded to the questionnaire survey is 'archive footage' (1,050 institutions). This is followed by content such as books, magazines, articles, films, photographs and musical pieces (Figure 4).

3. Digitization of materials

- There are 1,406 institutions that have digitized their collections, accounting for 48.13% of all responding institutions (Figure 11).
- One-third of the digitization was implemented in-house, one-third was partially outsourced, and one-third was fully outsourced (Figure 15).
- The institution types with the highest rate of digitization are, in descending order, archives (71.79%), museums (63.97%), and libraries (39.02%) (Figure 12).

4. Development of policies and plans for preservation of digital materials

- Only about 13% of institutions that hold digital materials have policies and plans for the digital

preservation. The total number of institutions that plan to develop policies and plans is less than 30% (Figure 33). On the other hand, the interview survey confirmed some cases where policies and plans played an important role (3.2.2(10)).

5. Implementation systems and human resources development related to the preservation of digital materials

- Only about 20% of institutions that hold digital materials have a person in charge of preservation assigned to digital materials (Figure 26). In most cases, the number of persons in charge is small (1 to 3) (Figure 30).
- These institutions provide training in digitization, research data management, image preservation, rights handling, etc. Some institutions have budgetary provisions for acquiring related qualifications (3.2.2(2)).
- In interviews, it was pointed out that collaboration and information sharing among them would be effective for the long-term preservation of digital materials. There are also cases where the persons in charge of practical work across different types of institutions regularly gather and hold study sessions (3.2.2 (10)).

6. Budgetary provisions for preservation and management of digital materials

- Only 26.75% of institutions that hold digital materials have budgetary provisions for preservation and management of digital materials. As the size of institutions grows, the number of institutions with budgetary provisions tends to increase. (Figure 31-32)
- Many of the interviewed organizations make budget requests based on the importance and development potential of a project, as well as responding to technological changes and advancements (discontinuation of equipment production, upgrades, etc.). In addition, an institution is investing in equipment to implement digitization and migration of materials within the institution with a view to the future (3.2.2(3)).

7. Recording media, file formats, preservation measures, etc. related to preservation of digital materials

- Responding institutions use not only normal optical disks, such as CDs and DVDs, and hard disks as storage media, but also various storage media for long-term preservation such as LTO, optical disks for long-term preservation, Optical Disc Archive (ODA), and M-DISC, as well as NAS and file servers that are easier to extract and back up data to (Tables 10, 13, 16).
- JPEG, TIFF, etc. are often used as file formats for preservation. Some institutions have adopted PDF/A, which is suitable for long-term preservation. In the interview survey, specifications specific to the materials (video materials, 3D data, etc.) handled by the responding institutions were also mentioned (3.2.2(8), 3.2.2(6)).
- Only less than 30% of institutions that hold digital materials have some form of backup. As a whole, museums and archives have a higher backup implementation rate than libraries (Figures 41 and 43). Some institutions have remote storage or backup to more than two medias.

- As preservation measures, considerably fewer institutions are taking actions such as inspection of material condition, migration (media migration, file format conversion), maintenance of playback environment, and use of external services, and more than 60% of institutions that hold digital materials answered that they are not implementing any preservation measures (Figure 44).

8. Cataloguing and metadata for digital materials

- It became clear that metadata schemas are being designed not only for standard elements such as those of Dublin Core, but also for a variety of fields depending on the materials in the collection. University libraries are making progress in metadata development (Figure 47).
- Additional elements for managing digital materials included file format, digitization date, rights information, license information, URL, and information about publication.
- In the interview survey, we confirmed cases in which metadata elements were designed according to the needs of each institution, such as the metadata of the video materials of the National Film Archive of Japan and the unification of metadata schemas through ColBase (National Institutes for Cultural Heritage Integrated Collection Search System), in which the Tokyo National Museum participates (3.2.2(8), 3.2.2(6)).

9. Collaborative activities and information sharing for the preservation and provision of digital materials

- Collaboration and information sharing with external organizations include cooperation with the National Archives of Japan and national research institutes, participation in international electronic resource sharing networks such as Internet Archive and ERDB-JP, regional collaboration and information sharing, and collaboration across administrative agencies and MLAs.

10. Services such as "digital archives" and "institutional repositories"

- The use of outside services is evident in public libraries and museums. University museums often develop and operate their own services. For university libraries and public archives, the ratio of use of external services and their own services is about half (Figure 52). Open source software was sporadically found as their own developed service to use. Most institutional repositories use external services, especially JAIRO Cloud.
- The National Archives of Japan has released the standard specifications of the digital archive system for archives nationwide (3.2.2(12)).

11. How to handle rights

- The results of the interview survey revealed various issues related to the handling of rights such as copyrights and portrait rights. Kobe University Library (3.2.2 (5)) and Toyonaka City Library (3.2.2(1)) gave specific examples.

12. Conclusion

The survey revealed that many organizations have digital materials in their collections, and that they are

promoting the digitization of materials. While some institutions are actively working on long-term preservation, most have not developed policies or plans for long-term preservation, and the lack of personnel and budget has emerged as a common issue.

The results of this survey revealed specific examples of advanced long-term preservation initiatives, collaborative activities, and information sharing, which will be helpful for future consideration of measures by various domestic organizations.