



The Migration of Optical Discs Book Supplements at the National Diet Library, Japan

KINOSHITA Takafumi, TAKEHANA Kazuo (National Diet Library, Japan)
iPRES2024 @Ghent & Flanders, Belgium (16–20 September 2024)



✉ lab@ndl.go.jp

Previous research on digital preservation at the NDL (in Japanese) →



1. Background

- The National Diet Library (NDL), as Japan's only national library, is responsible for the permanent preservation of its collections. Given this mission, the long-term preservation of digital materials such as optical discs is a new and important issue for the NDL.
- The NDL now seeks to develop a workflow for the efficient migration of diverse media into easy-to-use formats.
- This report describes our work in FY2023.

2. Target Discs and Results

Total	Successful	Error Sector	Unreadable	Noise	Others
8,840 (*1)	8,547 (*2)	168 (*3)	40 (*4)	55 (*5)	30

- *1: The discs targeted for migration were published in Japan no later than 2000 as book supplements.
- *2: 97% of the discs were migrated successfully.
- *4: Many of the unsuccessful discs were unreadable because vinyl packaging had deteriorate and adhered to the surface of the disc.

4. Results and Issues

- Detailed and accurate preservation of all data on discs might be better served by using procedure B for all disks.
- At the NDL, however, many cases to date have been performed using A or C on (4), and NDL staff are comfortable using data files and audio files in WAV format. During FY2023, we chose an eclectic approach and investigated the issues affecting each procedure.

A The data transfer is apparently slower when copying a large number of files and less efficient than B.

B Occasionally, defective sectors were found (*3) even when the error check found no defects. Additional investigation found only 15 discs (10%) with an actual error. Improved efficiency should eliminate the need for additional investigation.

C Occasionally, noise or silence were mixed with the audio. (*5) A partial listening inspection of 242 discs found no audible noise or silence was detected in discs with good error check value. So, efficient checking is possible by prioritizing discs with bad error check values (≠ Listening inspection (242 discs this time) is no longer necessary)

3. Workflow during FY2023

(1) Check physical condition

Visual check for obvious problems



(2) Check for Errors (for efficient inspection)

All discs were checked using error correction codes. (ISO/IEC 18630:2023 or equivalent)



(3) Categorize discs (for ease of use)

Discs were categorized according to type of data (audio, informational, etc.) and filesystem. Migration was performed according to category.

(4) Migration Procedure

A: Simple Data Disc (1,744 discs)

- A single file system with data readable by a Windows machine
- File-by-file copying of data

B: Mixed Data Disc (3,411 discs)

- A data disc that did not fit category A or that was at first categorized as A but failed to migrate.
- Create ISO image file (iso/cue, ISO Buster)

C: Audio CD (3,613 discs)

- A data disc comprising only audio tracks (CD-DA)
- Track-by-track conversion of all audio data to WAV format using Windows Media Player (WMP)

5. Future Study

- Procedure B is preferable even for discs that were migrated using Procedure A in FY2023
- We will continue to research and improve the workflow.