NDL Research Report No. 6

The Long-term Accessibility of Packaged Digital Publications

March 2006

National Diet Library

This report is a summary of results obtained through surveys conducted by external research organizations on behalf of the National Diet Library. It has been published with the aim of sharing these results with many other libraries and related organizations.

Preface

As a part of its digital library operations, the National Diet Library (NDL) conducts research on digital libraries. In FY2003 and FY2004, we commissioned external research organizations to conduct studies relating to the long-term preservation of and establishment of access methods for packaged digital publications such as CD-ROMs.

With dramatic advances in technology going on, digital information continuously assumes new forms, leading to the possibility that an increasing proportion of digital information stored in libraries and other places, even that which is only one generation old, will become obsolete and unplayable. Furthermore, there are different opinions in relation to the lifespan of the media on which digital information is recorded, and we are being urged to establish methods of preserving these media. These studies were initiated from a keen awareness that these issues are of vital importance to the NDL, whose duty is to collect, preserve and make available a large number of digital publications through the legal deposit system and other means.

The subject of these studies was packaged digital publications held by the NDL. However, we think that the results are relevant not only to the NDL, but also to other libraries and related organizations in the field of library and information science. It is for this reason that we decided to publish this research report to make these results available to many others, as Library Research Report No. 6, under the title "The Long-term Accessibility of Packaged Digital Publications."

The NDL will actively continue to invest in publishing Library Research Reports this year, and we ask for your continued support and cooperation.

March 2006

Toru Toyoda
Director
Library Support Division
Projects Department
Kansai-kan of the National Diet Library

CONTENTS

1. Introduction

This report is a summary of results of studies conducted in FY2003 and FY2004 in which the usability of packaged digital publications, such as CD-ROMs, was investigated. These studies form a part of research conducted by the National Diet Library (NDL) in relation to the "long-term preservation and use of digital information."

The use of digital information requires access to the medium on which the information is recorded, as well as playback equipment that is compatible with that medium. However, the lifespan of such media is said to be only about 20 to 30 years when preserved in a suitable environment, and the lifespan of playback equipment even shorter. Furthermore, standards for playback equipment and media change so frequently that it can be difficult to obtain items whose standards have become obsolete. Packaged digital publications only exist in media on which the information is recorded and as such they raise various issues relating to playback.[1]

The NDL is the only national depository library in the country, and as such is responsible for developing and preserving a comprehensive collection of domestic publications as valuable cultural heritage for future generations. In recent years the number of paper publications accompanied by supplementary floppy disks (FDs), CD-ROMs and similar digital media as well as packaged digital publications is increasing. In response to this trend, the NDL Law was amended[2] in 2000. Since then not only items published in the traditional paper form but also packaged digital publications domestically produced have been exhaustively collected under the legal deposit system.

In order to guarantee the long-term preservation and usability of packaged digital publications held by the NDL, there is a need to comprehend the current issues and establish measures necessary for long-term preservation. Therefore, in FY2003, a study was conducted to test the usability of digital publications in what was then the latest personal computer (PC) environment. In order to solve the usability problems revealed in the 2003 study, two methods which have been thought to be solutions to long-term preservation, namely migration and emulation, were tested, and the result was evaluated in FY2004.

The results of the FY2003 study show that approximately 70% (138 items) of the sample packaged digital publications (200 items in total) had usability problems in one way or another. The results of the FY2004 study indicate that migration as a means of copying to media of different types is effective, but that it would be difficult to claim that emulation and migration as a means of file format conversion are effective technological solutions at this point.

The results of both studies are detailed in the following chapters.

---

[1] The problems related to preservation of digital information were recognized in early stage by organizations such as national libraries in Europe, U.S., and Australia, and various projects have been implemented since the 1990s. In each country, the national library is leading the way in this field, and the NDL also began its own research in FY2002.

[2] National Diet Library Law (Law No.5 of February 9, 1948, as amended by Law No.37 of April 7, 2000 (effective from October 1, 2002)). Prior to this amendment, packaged digital publications were being collected through purchasing and other methods.

2. Results of FY2003 Study

Packaged digital publications held by the NDL were examined over a period of one month starting in December 2003. The investigation included counting the library's stock of these digital publications by operating environment and medium, and testing the usability of sample publications by playing them in the then latest operating environment.

2.1. Stock survey of packaged digital publications

In preparation for the playback tests, the library's stock was first counted by parameters including content, medium, playback environment and year of acquisition, and the result was summarized in tables.

The subject of the survey consisted of the following three groups of packaged digital publications:

- Digital material: Materials other than visual and audio materials. Materials produced for playback on equipment such as PCs and game consoles.
- Visual material: Materials containing recorded images, such as video tapes, video discs and DVDs.
- Audio material: Materials containing recorded sound (primarily music), such as CDs, DVD-Audio discs, Mini Discs and vinyl records.

Of these materials, items acquired up to FY2002 were subjected to examination. The number of acquisitions was taken from the library's existing catalogs and statistics.

In order to investigate the relationship between the variables such as year of acquisition, playback environment and media type, the result was summarized in three tables. As it is necessary to find about media type and trends for all the various playback environments of digital materials, tables were made up to show the number of materials in stock sorted by playback environment and medium, and the number of acquisitions sorted by playback environment and year of acquisition.

(1) Table 2.1-1 Year of Acquisition and Media Type
In this table the numbers of digital, visual and audio materials are shown for each media type and each year of acquisition.

Trends of each media type can be seen in this table.
- The number of magnetic discs acquired began to decrease after the peak in FY1996.
- The number of LDs acquired began to decrease after FY1996.
- Acquisition of DVDs began in FY1997, and a dramatic rise in the number of acquisitions began from FY2000.

DVDs accounted for 80% of the visual materials acquired annually in 2003. It seems that in the near future DVDs will become almost all the acquisitions.

(2) Table 2.1-2 Playback Environment and Media Type
In this table, the number of digital materials of each media type is shown for each playback environment.

The following general categories were used for playback environment: "PC," "game console," "electronic book reader," "DAISY[3]," "other" and "not specified." Materials for PCs

[3] DAISY stands for Digital Accessible Information System, an international standard for digital recorded books for the visually impaired and people who have difficulty in reading ordinary printed matters. Here it refers to materials that conform to this standard.

are further sorted into the following categories: “PC/AT compatible[4],” “FM-TOWNS[5],” “Macintosh,” “PC-98[6],” “materials designed to run on specific applications,” and “other.”

For statistical convenience, digital materials made for multiple playback environments were counted in one of the choices. For example, hybrid digital materials that work with both Windows and MacOS operating systems were classified under “Windows.” Likewise, materials designed to operate in a playback environment provided by a specific application were classified under that application, e.g., Acrobat. (The “win” category is for materials that were designed for a version of Windows, but where there is no indication of a specific version in the catalog.)

This table shows that how each type of media is used in each playback environment.

(3) Table 2.1-3 Playback Environment and Year of Acquisition
In this table, the number of digital materials acquired each year is shown for each playback environment.

Trends of rise and decline of each playback environment can be seen in this table.

[4] Also known as DOS/V machines, IBM compatibles and Windows computers. In this survey, because operating systems for these machines are limited to DOS and Windows, this name is used. “PC/AT” was originally the name of personal computers sold by IBM from 1984, but the specifications were made public and many compatible machines have been produced. It has become widely accepted as the global standard for PCs.
[5] A personal computer made by Fujitsu. Sales began in 1989.
[6] This is a generic name used to refer to personal computers sold by NEC from 1982 until the early 90s, and to those named the “PC-9801” series and its successor, the “PC-9821” series.

Table 2.1-1 Year of Acquisition and Media Type

Material	Media Type			Year of Acquisition																							Total
				1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	
Digital material	Optical disc	CD									1	6	5	18	19	8	70	119	142	260	274	126	479	551	2,666	6,614	11,358
		DVD																								103	103
		Other*1																					1	6	66	83	156
	Magnetic disk	FD	3.5"								1			2		8	73	28	45	56	18	6	4	1			242
			5"										2	1	5	1	1			1	1						12
			Not specified							1		1			5	8	7	21	2							36	81
		Other													1				4			2	10	9	19	18	63
	Other*2																1			5		1		14	122	126	269
Visual material	Optical disc	DVD																			11	22	36	1,523	4,540	4,299	10,431
		Video disc	LD		145	178	170	299	424	677	559	735	822	1,257	1,852	1,771	1,511	1,319	1,050	1,315	855	616	550	216	35	76	16,432
			LDs															24	29	23	41	47	49	57	68		338
	Magnetic tape	Video cassette	VHS				2	3		11	19	28	99	116	220	203	150	186	224	256	225	359	303	1,077	1,351	603	5,435
			Beta					3		2	1												1				7
			Other		43		32	22	1	20			5														123
Audio material	Optical disc	CD				20	291	1,322	2,425	3,833	5,335	6,412	6,566	9,705	16,340	13,571	13,260	14,092	11,775	13,234	13,422	11,968	10,223	14,506	9,405	8,418	186,123
		DVD-AUDIO																							64		64
		MD														65	48	165	445	184	130	214					1,251
	Other	Vinyl record	LP/EP/SP	11,244	9,172	8,812	10,681	5,957	11,046	4,457	3,923	2,716	785	342	143	267	207	71	84	110	99	94	87	200	56	1	70,554
Total				11,244	9,360	9,010	11,176	7,606	13,896	9,001	9,839	9,898	8,284	11,441	18,585	15,902	15,328	16,025	13,800	15,444	15,076	13,455	11,743	18,160	18,392	20,377	303,042

*1 Includes items listed only as “optical disc”
*2 Includes items whose media type is not specified

Table 2.1-2 Playback Environment and Media Type

Material	Media Type			PC																Game console	Electronic book reader / Electronic book	DAISY	Other	Not specified	Total
				windows									FM-TOWNS	MacOS	PC-98	Application specified				Other						
				DOS	win	win3.x	win95	win98	winME	winNT	win2K	winXP				Acrobat Reader	Web browser	EPWING	Other		*1					
Digital material	Optical disc	CD		39	699	134	693	133	88	269	295	262	3	135	32	179	59	47	39	453	522	51	2,629	59	4,538	11,358
		DVD																			32				71	103
		Other			11			1											1		132				11	156
	Magnetic disc	FD	3.5"	1	28	6	7							21	7										172	242
			5"																						12	12
			Not specified	4	2		2							1	6				3						63	81
		Other		2	5	1	2													1				5	47	63
	Other																				250			12	7	269
Total				46	745	141	704	134	88	269	295	262	3	157	45	179	59	47	43	454	936	51	2,629	76	4,921	12,284

Table 2.1-3 Playback Environment and Year of Acquisition

Material	Year of acquisition	PC																	Game console	Electronic book reader / Electronic book	DAISY	Other	Not specified	Total
		windows									FM-TOWNS	MacOS	PC-98	Application specified				Other						
		DOS	win	win3.x	win95	win98	winME	winNT	win2K	winXP				Acrobat Reader	Web browser	EPWING	Other		*1					
Digital material	1986																						1	1
	1987	1																					1	2
	1988	2											2										3	7
	1989																						7	7
	1990										1	1	8			1							10	21
	1991	2										3	1			1							23	30
	1992											4	1										20	25
	1993	1	3								2	4	27			4							111	152
	1994	3	6	7								7	3			11							131	168
	1995	17	69	10	3							23											71	193
	1996	3	81	13	16							23	3		1	3							179	322
	1997	1	50	21	101			17				9		1		9							84	293
	1998	1	2	17	47			24						3	1		6						34	135
	1999	7	4	59	261	19		39	1			6		7	12		15	1	9	6		2	46	494
	2000	2	42	8	156	41	6	47	36			36		23	4		7	1	87	2		19	64	581
	2001	3	235	4	62	53	63	114	179	11		13		71	16	17	9	22	379	2	300	35	1,285	2,873
	2002	3	253	2	58	21	19	28	79	251		28		74	25	1	6	430	461	41	2,329	20	2,851	6,980
Total		46	745	141	704	134	88	269	295	262	3	157	45	179	59	47	43	454	936	51	2,629	76	4,921	12,284

*1 Type of game: Dreamcast, NINTENDO64, PlayStation, PlayStation2, SegaSaturn, Xbox, Game Gear, GameCube, Game Boy.

2.2. Usability study

(1) Sample selection

The three tables produced did not provide sufficient information to select samples. Therefore, the following assumptions and facts were also considered in sample selection.

- Older materials are more likely to be unusable.
- Audio and visual materials (CDs, DVDs, and LDs, etc.) whose media and playback environment hardly change over time should not be treated in the same way as the other digital materials whose medium and playback environment come in variety and are continuously being replaced by newer versions (operating systems, applications, formats, etc.).
- In the case of audio and visual materials, where the medium and playback environment do not change, investigating usability by playing sample items would be equal to checking whether media deterioration has occurred or not.
- In the case of digital materials, there are various playback environments which are continuously being replaced by new versions. Therefore, it is necessary to check media deterioration, but a usability study of digital materials would be almost equal to investigating which materials have become difficult to use due to technological obsolescence.
- It can be assumed that most recent materials, not only audio and visual materials but also digital materials, could be played with almost no problems.
- CDs account for the majority of audio and visual material.
- Included among digital materials is software for game consoles, whose medium and playback environment do not change, as is the case for audio and visual materials. "Materials with various playback environments which are continuously being replaced by new versions" should be regarded as meaning digital materials that are played or operated on PCs and similar equipment.

Taking the above points into consideration, the samples were selected as described below.

- In regard to audio and visual materials, the study was limited to audio CDs. Five CDs were randomly selected from the acquisitions of each year starting from FY1982, when acquisition of CDs began, to FY1991. (50 CDs were selected in total.)
- In regard to digital materials, the survey was limited mainly to materials that were designed for playback on PCs. The samples were comprised of 29 materials acquired before FY1990, 20 each between FY1991 and FY1998, and 11 each from FY1999, all of which were randomly selected. (200 items were selected in total.)

(2) Testing method

The following procedures were as follows:

- For digital materials, file list information contained in the media was checked, and loading and running a simple operation was tested to see if it worked on one of the latest operating systems (Windows XP Professional or MacOS X v10.3).
- For audio materials (audio CDs), tests were done to see if proper playback is possible by playing the forepart of the first track, the entire length of one track in the middle, and the forepart of the last track of each CD.

However, we do not think that these testing methods were sufficient. Especially in the case of audio CDs, an error correction function is incorporated in the standard, so even if there is any error on the data recorded in the medium, it can be played without any apparent problems. Furthermore, it is impossible to tell whether such errors are the result of deterioration over time or if they have existed from the beginning. Therefore, in order to confirm the extent of deterioration of CDs caused by aging, we assume that it is necessary to measure the error rate of information recorded in a sufficiently large number of samples or examine the surface of discs by using a microscope. Therefore we regard the usability study of audio CDs in this research as providing no more than a rough

indication.

In the case of digital materials, especially those that contain programs, there may be cases in which a problem will only occur when a certain operation is carried out. However, verifying such a problem would be like the debugging process performed by computer manufacturers, involving a huge amount of work, which would make it impossible to perform an investigation. Therefore, in this investigation we chose only to test the readability of the media itself by displaying file list information and check to see if there were any obvious problems by performing simple operations.

(3) Test results
<Audio materials (audio CDs)>
All 50 items were playable.

Table 2.2-1 Usability of Audio Materials by Year of Acquisition

Year acquired	Number of items tested	Number of usable items	Percentage of usable items
FY1982 (April 1982 – March 1983)	5	5	100%
FY1983 (April 1983 – March 1984)	5	5	100%
FY1984 (April 1984 – March 1985)	5	5	100%
FY1985 (April 1985 – March 1986)	5	5	100%
FY1986 (April 1986 – March 1987)	5	5	100%
FY1987 (April 1987 – March 1988)	5	5	100%
FY1988 (April 1988 – March 1989)	5	5	100%
FY1989 (April 1989 – March 1990)	5	5	100%
FY1990 (April 1990 – March 1991)	5	5	100%
FY1991 (April 1991 – March 1992)	5	5	100%
Total	50	50	100%

<Digital materials>
Usability problems were found in a little less than 70% of the materials tested.

As we presumed, older materials were less likely to be usable, and some media were unreadable. Such media included not only 5-inch floppy disks (hereinafter 5" FDs) and 3.5-inch floppy disks (hereinafter 3.5" FDs) but also some CD-ROMs. The usability of materials acquired until FY1994 was particularly low.

Table 2.2-2 Usability of Digital Materials by Year of Acquisition

Year acquired	Number of items tested	Number of usable items	Proportion of usable items
FY1990 and earlier (– March 1990)	29	1	3%
FY1991 (April 1991 – March 1992)	20	2	10%
FY1992 (April 1992 – March 1993)	20	2	10%
FY1993 (April 1993 – March 1994)	20	3	15%
FY1994 (April 1994 – March 1995)	20	3	15%
FY1995 (April 1995 – March 1996)	20	8	40%
FY1996 (April 1997 – March 1998)	20	8	40%
FY1997 (April 1999 – March 2000)	20	16	80%
FY1998 (April 2001 – March 2002)	20	13	65%
FY1999 (April 2003 – March 2004)	11	6	55%
Total	200	62	31%

(4) Outline of factors that make materials unusable

The results of the usability study confirmed that there are many cases in which digital information past a certain age cannot be used on the latest operating systems. The causes are as follows:

- Cause: Operating systems or other fundamental PC software

  Applications usually only work with a particular version of a particular operating system. Applications designed for old operating systems will not necessarily function with the latest operating systems.

  Various problems arose from incompatibilities between the latest operating systems and obsolete playback applications required by digital materials.

- Cause: Applications

  In some cases, the use of digital information requires specific playback applications in addition to applications normally installed on PCs. In this study, some materials could not be used because the playback applications for these materials were not available or could not be identified.

  Furthermore, some materials did not function properly because of incompatibility between specific plug-ins needed to be installed in an application and the latest version of that application.

- Cause: Medium

  Some materials were unusable because they were contained on 5" FDs for which a suitable FD drive could not be obtained. As technology becomes obsolete and production/distribution of a medium ceases, as in the case of 8-inch and 5" FDs and their drives, it becomes difficult to continue using these media and drive systems.

  Also while the lifespan of media is said to be 20-30 years, some materials tested could not be read, even though they had been stored in the NDL stacks under relatively suitable conditions.[7]

- Other causes

  Some materials could not be played because FDs or CD-ROMs necessary for installation and setup were missing, or because the playback environment could not be identified.

[7] The stacks of the NDL are maintained at a temperature and humidity suitable for the storage of paper media. However, temperature and humidity recommendations for the preservation of digital media are lower than those for paper media. (UNESCO, MEMORY OF THE WORLD: Safeguarding the Documentary Heritage. A guide to Standards, Recommended Practices and Reference Literature Related to the Preservation of Documents of All Kinds.)
http://www.unesco.org/webworld/mdm/administ/en/guide/guidetoc.htm

2.3. Analysis of test results

Factors preventing use of all the digital materials tested are shown in Figure 2.3-1. Shifts in factors preventing use are shown in Figure 2.3-2.

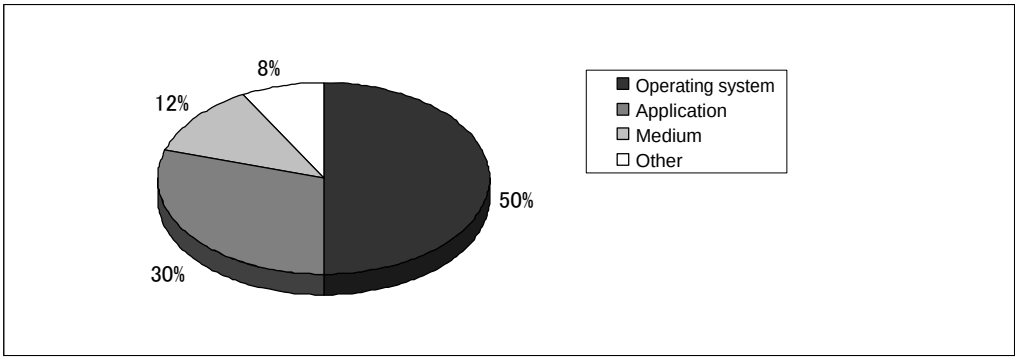

Figure 2.3-1 Factors preventing use (of all samples tested)

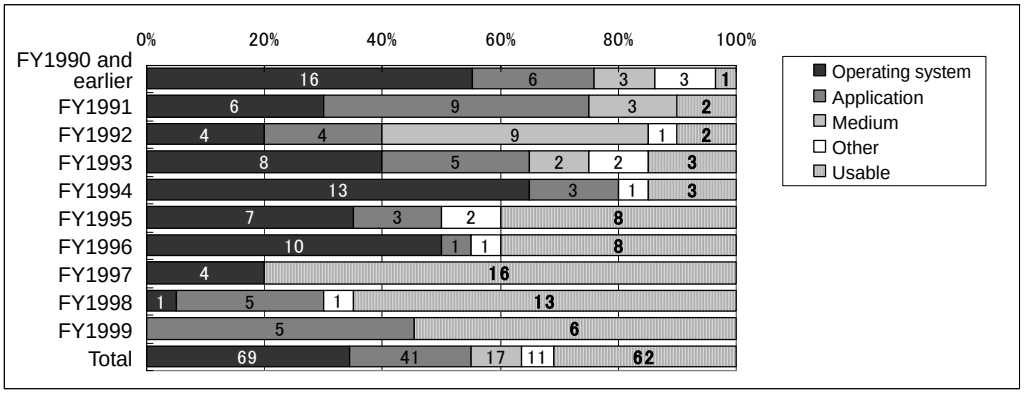

Figure 2.3-2 Factors preventing use and the proportion of usable materials by year of acquisition

In half the cases the factor preventing the use of digital materials lay in the operating system (incompatibility between operating systems and applications). In 30% of the cases it was application (unavailability of applications, incompatibility between applications and plug-ins), and in 10% of cases it was the media (technological obsolescence and deterioration of media).

The proportion of each factor has changed over the years, and some trends can be observed to a certain degree.

(1) The newer the materials, the fewer the cases where the factor preventing use is the operating system. However, it does not necessarily mean that this proportion increases with the number of years that have passed since acquisition.

The usability of digital materials is determined by functional continuity and levels of similarity between the latest and superseded operating systems on which the material was designed to play; by the timing of operating system upgrades and sales discontinuation of superseded operating systems; and by the functions of the operating systems used by the digital material. It is for these reasons that some of the digital materials held by the NDL for PC-98 machines (used to be called "the most popular computer for the Japanese people") can work on the latest operating systems, while some do not.

(2) There is no evidence that the proportion of cases in which the factor preventing use is application-related increases with the number of years that have passed since acquisition.

The proportion of cases in which the factor preventing use was related to applications decreases until FY1997, after which it begins to increase. However, consideration must be given to the small number of materials and samples upon which the study was conducted. There is also a possibility that the result may change significantly depending on the samples selected.

Among the unusable materials from FY1999 there were materials which use Acrobat Reader as their playback application. Four of these items could not be played due to incompatibilities between the latest version of Acrobat Reader and the plug-ins included in the media (where the total number of items that were unusable due to application was five).

Although they were not found to be unusable in this study, the NDL holds dozens of digital materials that use web browsers as their playback applications, and the number is estimated to grow in the future. Plug-ins are used not only by Acrobat Reader, but also by web browsers, and it is possible that similar problems (incompatibility between applications and plug-ins) will arise in the future.

(3) The proportion of cases in which the factor preventing use is the medium increases with the number of years that have passed since acquisition.

The media tested in this study were acquired approximately 5 to 18 years prior to the time of study (December 2003 to January 2004), and it is therefore unlikely that any materials have become unusable due to aging past their lifespan.[8] However, the number of media that become difficult to use because of change (or abolition) of standards is expected to continue to increase in the future. 3.5" FDs have already become inferior to other media currently in circulation in terms of data capacity and media safety, and it is unlikely that their use, production and circulation, and the production and circulation of compatible hardware and drivers, will continue over the next several decades.

At some point in the future, it will probably happen that all digital materials stored on 3.5" FDs will all at once become difficult or impossible to use. Although many PCs currently on the market are still equipped with 3.5" FD drives as standard, the falling costs of CD-ROM/R/RW drives and the increase of the data storage capacity of these media is causing FDs to be used less and less frequently. Some PCs are no longer equipped with 3.5" FD drives as standard. In a few decades, a similar situation will probably befall CDs and DVDs though they form the majority of the digital media of today.

[8] The condition of data contained in 5" FDs is unknown because no drives could be obtained to play them in this study.

2.4. Detailed analysis of factors preventing use

(1) Incompatibility with the operating system

Incompatibility with the operating system accounted for the majority of cases where materials were found to be unusable, and was thought to cause a variety of symptoms, including the following:

- Computer ceases to respond during installation,
- Installation is unsuccessful,
- Incorrect color are displayed,
- Application does not boot,
- Abnormal termination of application.

In order to play digital materials, both the correct equipment and playback application (in many cases the digital material itself is the application) are required, but such applications usually only work with a particular operating system. There are many different types of operating systems, and many different versions of each type have been developed. As a result, a significant number of operating systems have become obsolete. The release of a new operating system is always accompanied by the release of new versions of applications.

The number of digital materials that cannot be played because of incompatibility with operating systems is predicted to increase rapidly with the release of new operating systems and new versions of operating systems.

(2) Application-related factors

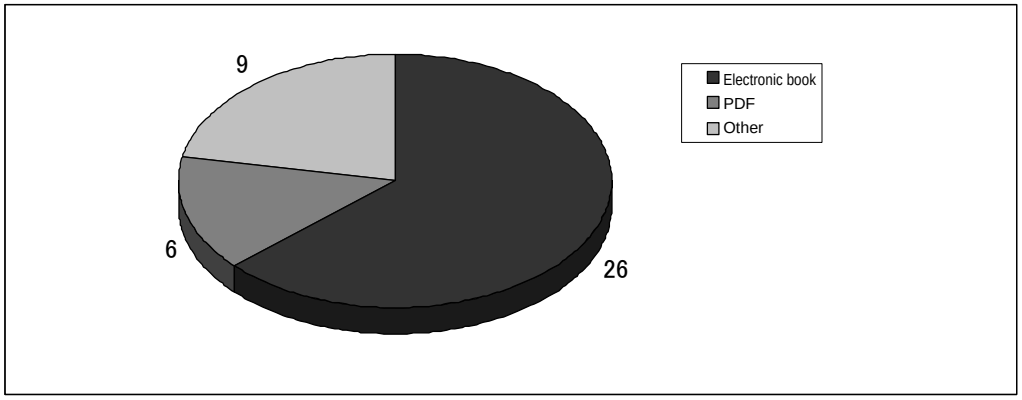

Figure 2.4-1 Application-related factors preventing use (number of materials)

There were 41 materials which could not be played due to application-related factors, and the majority of these were materials that required applications for playback of electronic books (26 items in total).[9]

It was noted that there are a number of electronic books playback applications, some of which are free and some of which are available for a fee, and that they can be obtained fairly easily. (Some are even available for download over the Web.) However, in some cases playback was possible using one application and not another, and in these cases the items were treated as being unplayable.

PDF[10] is sometimes recommended as a recording format for long-term preservation, but in this survey 6 items recorded in the format were found to be unusable. As already mentioned, the reason was incompatibility between the latest version of Acrobat Reader and plug-ins included on the media.

Materials listed in "other" were items that required specific word processing software, database software, spreadsheet software, special applications or kana-kanji conversion software. Although these playback applications are not necessarily difficult to obtain, materials that require them were treated as unusable because the applications they require are not currently widely used, or because they were designed for superseded versions of these applications.

[9] Electronic books and electronic dictionaries recorded in the EPWING format.
[10] The Portable Document Format is a format for electronic text, which includes not only text but also layout and image information. It was developed by Adobe Systems.

(3) Breakdown of media-related factors preventing use

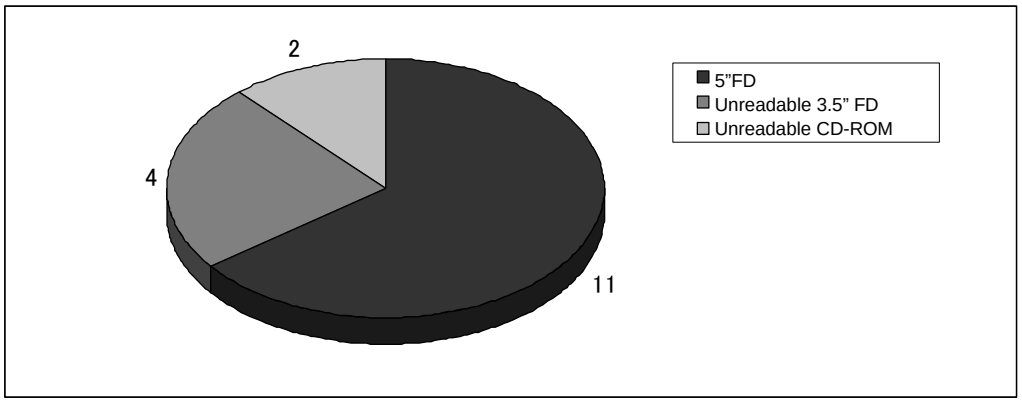

Figure 2.4-2 Media-related factors preventing use (number of materials)

Seventeen of the materials tested were unusable due to problems with the recording media. The majority (11 items) of these were 5" FDs, which could not be read because a compatible drive could not be obtained.

In addition, six materials (four 3.5" FDs and two CD-ROMs) for which compatible drives were available could not be read. In the case of the 3.5" FDs, it is likely that they could not be read because part (or all) of the data recorded on the disks had been lost. In the case of the CD-ROMs, both were domestic electronic materials acquired in FY1994, and it is unlikely that media deterioration or the use of special recording formats were responsible for their inability to be played. As both discs were published by the same publisher at the same time, and serious consideration must be given to the possibility that the items were defective from the time of manufacture. It follows that materials need to be tested at the time of acquisition.

(4) Other

As already mentioned, there were some materials that could not be played because FDs or CD-ROMs necessary for installation and setup were not included, or because the correct playback environment could not be identified. These problems occurred because at the time of acquisition the necessary media were not checked and metadata describing details of playback environments were not sufficiently created. These cases accounted for approximately 10% of all cases where materials could not be used.

Recent catalogs contain fairly adequate descriptions of playback environments based on information printed on packages, but where there is insufficient information on a package, the information in the catalog will also be insufficient.

Furthermore, descriptions of playback environments are based on the technological awareness level that prevailed at the time of manufacture. Therefore, when dramatic changes in technology (for example the increase in the size of internal memory from the order of kilobytes to the present order of gigabytes, and the decline of the floppy disk) cause changes in the general knowledge, there is a possibility that descriptions of playback environments will not be correctly understood.

3. FY2004 Study

3.1. Purpose

Results from the FY2003 study indicate that unless special measures are taken to ensure the long-term preservation and accessibility of digital materials, a significant number of these materials will become unusable. In general, procedures such as emulation and migration are considered necessary for the implementation of such measures.

(1) Migration
Migration is the movement and conversion of programs and data, which means the processes of copying to new media of the same type (FD to FD etc.), copying to different media (CD-ROM to DVD-R etc.), updating recording formats, and rewriting playback applications (in some cases the digital material itself is the playback application) to make them work in other playback environments.

There are several different methods of migration, and although there is no single classification system, we have chosen to classify them as follows.[11]

(i) Copying to new media of the same type
(ii) Copying to different media
(iii) Converting the data format
(iv) Transferring programs and systems to a new playback environment

Method (i), copying to new media of the same type, can be applied to all digital materials. As a short-term solution, this method is easy to implement in terms of both technology and labor. However, technical difficulties can be expected in many cases when attempting to transfer data from media that have copy protection features.

As a long-term solution, this method presents various technological issues (such as drive maintenance) that accompany the obsolescence of media standards and the associated difficulties in obtaining drives and media.

The process itself, although simple, is unavoidably labor-intensive, and the increasing workload involved in processing an increasing number of archived resources cannot be ignored. Furthermore, migration must be performed before media deterioration occurs, and variation in the lifespan of individual media means that migration would have to be performed on a much shorter cycle than the expected lifespan of the media.

Considering the fact that media standards change and become superseded, and that at some stage materials will need to be copied onto different types of media, this method may be effective in the short term, but its long-term validity is questionable.

Method (ii), copying to different media, like method (i), is easy to implement in terms of technology and labor, but technical difficulties can be expected in many cases when attempting to transfer data from media that have copy protection features.

Unless transferring materials to a single storage unit of sufficient size, this method, like copying to new media of the same type, becomes very labor-intensive and involves significant workload in the long term.

However, in contrast to the process of copying to new media of the same type, this method can be said to be valid over the long term provided there is a reasonable hope that the new

[11] Under the OAIS (Open Archival Information System: a technology standard that specifies abstract specifications for archiving systems for digital information, ISO 14721:2003) migration classification system, these types are referred to as refreshing, replication, repackaging and transformation.

media chosen will continue to be used in the medium-to-long term (although it is difficult to state an exact period).

Method (iii), converting the data format, is the process of converting data to a new format, for example from JPEG to JPEG 2000, or from RTF to PDF. Provided the new data format is a standard one that will continue to be widely used in the future, this method is an easy way to maintain the playability of materials. In some cases, this process can be performed easily using commercially available or free conversion programs.

Conversion is possible for data only; the conversion of digital materials that contain programs is generally not possible.

Method (iv), program migration, is the process of using specifications, design sheets and source code to rewrite programs, making amendments as necessary. Because it is difficult to obtain all the resources needed to migrate commercial software, it is likely that this method could only be used for programs such as those created internally by organizations. Although this method has been included in the migration category, compared with the other methods it is much more complicated and requires more advanced technology.

(2) Emulation

Emulation refers to the process of artificially recreating one operating environment in another so that software (operating systems and applications) designed for an old operating environment can be run on a new one.

In order to access digital materials, the correct operating environment is needed, i.e. specific hardware and software. However, because the lifespan of hardware is short[12], maintaining hardware in an operable condition is not realistic. Emulation solves this problem by recreating hardware through simulation. (In some cases even whole operating systems can be emulated.)

Conducting emulation on the latest operating system means running an application called an emulator on the latest hardware and on the latest operating system, which performs the same tasks as old hardware and an old operating system.

Normally, applications access hardware functions and functions of an operating system through the operating system itself. An emulator is a type of application that artificially recreates functions necessary for old software.

Even when hardware and operating systems change, simply creating an emulator can allow people to continue to use existing applications and operating systems.

Many such emulators are being made and distributed.[13] However, creating one requires technical knowledge, and it may not be possible to completely recreate a superseded operating environment.

These measures are only theoretical solutions, and their effectiveness and the existence of any related issues can only be revealed through their practical application. The FY2004 study was conducted with the aim of actually performing migration and emulation on packaged digital publications stored by the NDL, assessing the effectiveness of these measures and identifying any issues.

[12] Many of the electronic components used in hardware have short life spans.
[13] See Appendix 1.

3.2. Outline of Study

Although the aim of the study was to assess the effectiveness of and to identify the issues associated with migration and emulation by actually applying these methods, there are many different ways in which migration and emulation can be conducted. The following procedures were chosen and implemented after giving consideration to the current appropriateness of different techniques, while keeping in mind the envisaged future implementation of such measures at the NDL.

(1) Migration (transfer to hard disk)
   (i) Transfer digital materials to hard disk using a commercially available migration program.
   (ii) Create metadata from instructions and manuals accompanying digital materials.
   (iii) Measure the time required for migration. Investigate the effects of differences in CD read rate, CPU performance and other parameters.

(2) Emulation
   (i) Use commercially available emulators to artificially recreate superseded hardware environments, and then establish superseding operating environments by installing superseded operating systems in those hardware environments.
   (ii) Test the usability of digital materials that contain programs by playing each in the superseded operating environment stipulated.

(3) File format conversion
   (i) Use a commercially available file format conversion program to perform file format conversion on digital materials that contain only data, and then confirm the playability of converted files.
   (ii) Use a multi-file viewer to confirm the playability of digital materials that contain only data.

Regarding procedure (1), migration includes a number of different concepts, but in this case, considering its future implementation at the NDL, we chose to transfer data to large capacity hard disks – in other words migration to a different media type. For procedure (2), emulation, we chose to use a commercially available emulator because it was the simplest method available. Although procedure (3), file format conversion, is a concept that could be included in migration, because it is a way, like procedure (2), of maintaining accessibility over the long term, we decided to treat it as separate from procedure (1), migration. File format conversion was also performed using a commercially available application for simplicity.

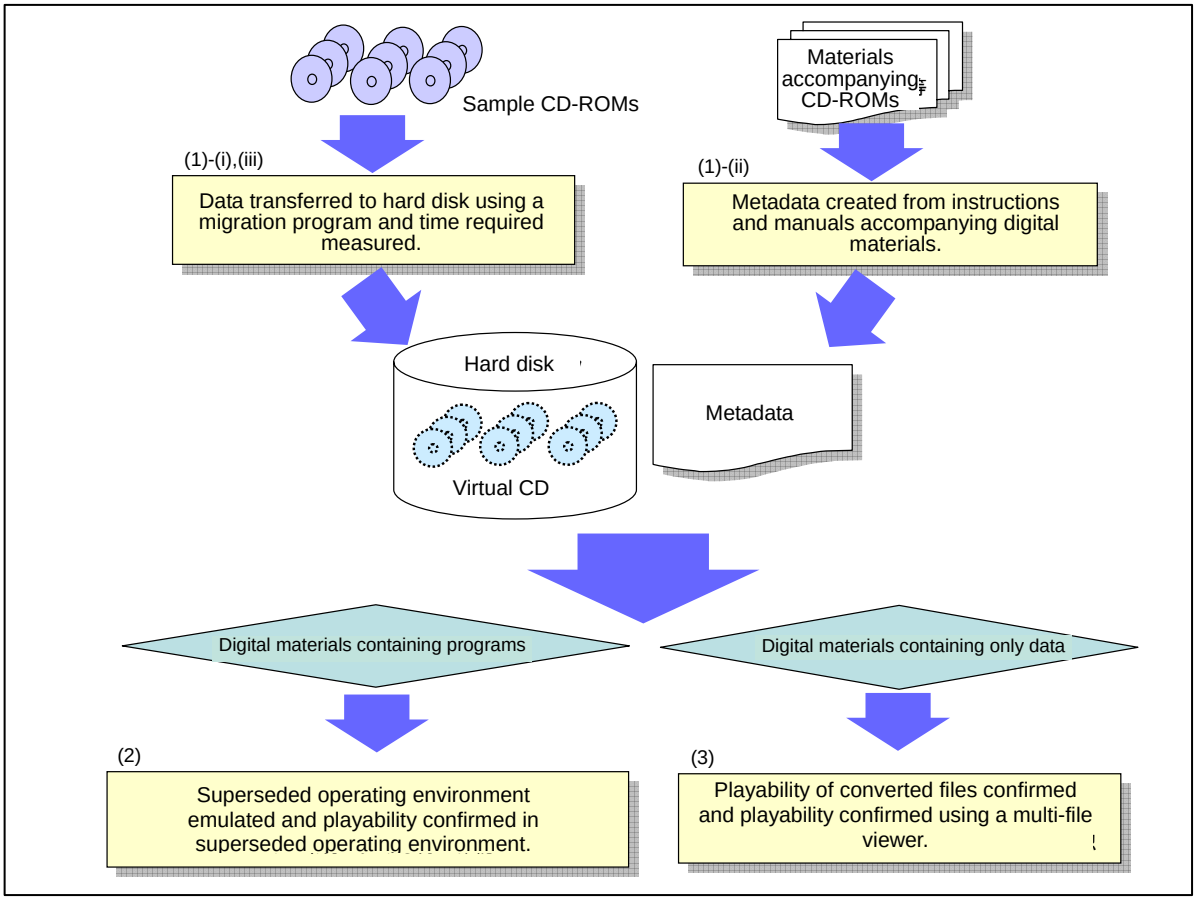

Figure 3.2-1 Pictorial representation of study outline

The numbers above the top left of each box in the above diagram correspond to the numbers in 3.2.

3.3. Migration to hard disk

3.3.1. Migration trial

Packaged digital publications designed to be played mainly on PCs were the subject of the FY2003 study, and these included software for Macintosh and PC-98 computers, as well as 5" FDs. Considering the fact that the majority of packaged digital publications for PC held by the NDL are for PC/AT compatibles, and the fact that the majority are recorded on CD-ROM, it would seem appropriate to limit the investigation to CD-ROMs for PC/AT compatibles. The subject of the investigation was therefore defined as follows.

[Subject of the investigation]
A sample of 354 discs[14] was selected from domestic digital materials acquired prior to 1999 that were recorded on CD-ROM/R/RW for Windows or DOS.

Only discs acquired before FY1999 were included in the investigation because we wanted to use materials from the same period as those used in the FY2003 study. The subject was limited to domestic publications because domestic publications should be given priority over overseas publications, and because overseas publications are sometimes affected by problems related to character codes, resulting in the possibility that the scope for problems would become unnecessarily great.

(1) Results
A commercially available migration program[15] was used to transfer the sample digital materials from CD-ROM to hard disk. The following results were obtained.

Table 3.3-1 Migration results

Year of publication	1991-1994	1995	1996	1997	1998	1999	Total
Number of items tested	16	23	31	48	143	93	354
Number of failures				1		1	2
Number of successes	**16**	**23**	**31**	**47**	**143**	**92**	**352**

- In both cases where migration failed, the process was aborted because migration had not finished after one hour of processing time.

[14] In order to provide enough material for later playback tests of digital materials containing programs and digital materials containing only data, it was necessary to migrate 100 items of each type. However, because it was difficult to judge from the catalog which materials included programs and which included only data, and the only sure way to distinguish is to actually play them, "materials *thought to* contain programs" and "materials *thought to* contain only data" were selected including a reserve of extra items.
[15] We used CD Kakumei Virtual Pro Ver. 8 (by ARK Information Systems). Our aim was not to evaluate the product, but to generate generalized results.

(2) Data compression

The migration program has a data compression function.[16] The results of compression are shown below. On average, files were compressed to approximately 60%.[17]

Table 3.3-2 Compression during migration

	Value
Total size before migration	123,575MB
Average size before migration	351MB
Total size after migration	84,358MB
Average size after migration	240MB
Average compression rate	65.8%

- Average compression rate was obtained by averaging the compression rates[18] for each digital material.

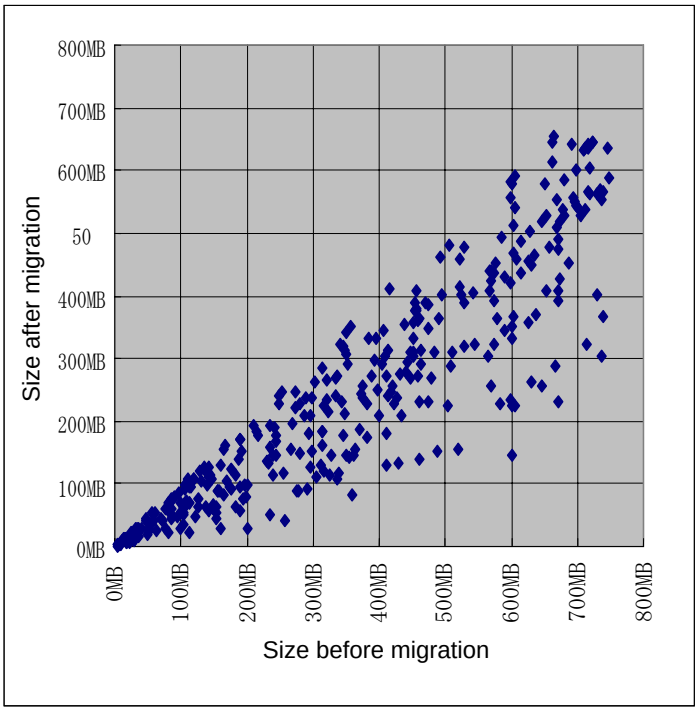

Figure 3.3-1 Distribution of data size before and after migration

[16] The compression method used has not been disclosed. It is thought to be proprietary technology belonging to the manufacturer.

[17] It is important to note that this figure may only be characteristic of the sample tested in this survey, or of the program used to perform migration.

[18] Compression rate was calculated as size after compression / size before compression x 100.

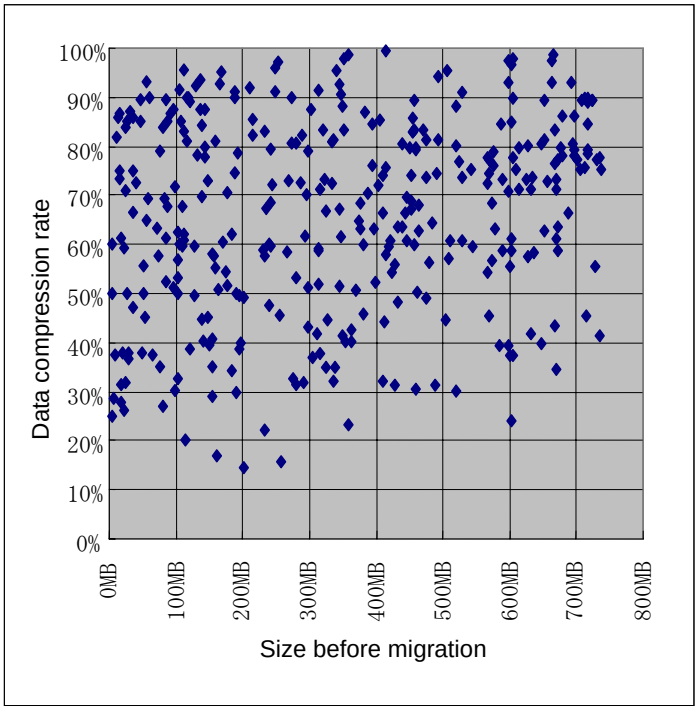

Figure 3.3-2 Data compression rate and size before migration

3.3.2. Time required for migration

The time required to complete the migration of 352 digital materials was approximately 16 hours.[19] On average, it took a little less than three minutes for one item.

Table 3.3-3 Time required for migration

	Migration time (hours : minutes : seconds)
Total	15:30:31
Maximum	0:14:53
Minimum	0:00:08
Average	0:02:39

However, the results depended on the environment[20] in which migration was performed. For future transfers, it is desirable to be able to estimate the time required from the number of materials and the migration environment. The migration time would be calculated by using information such as the number and size of items to be transferred, CD read rate, memory size, CPU performance, hard disk speed and bus speed.

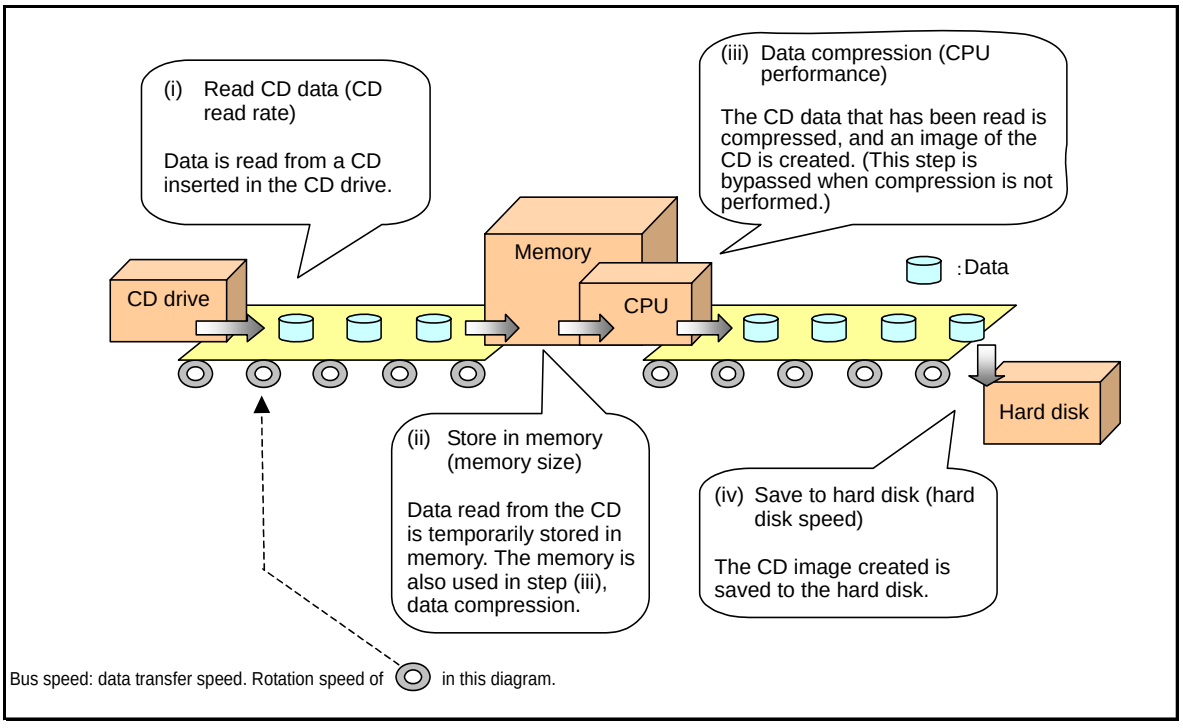

Figure 3.3-3 Pictorial representation of the migration process

[19] Total processing time required by the migration program. This figure does not include time required to change CDs and perform other tasks.
[20] The hardware used had a CD read rate of 24 x.

3.3.3. The effects of technological factors on migration time

Of the technological factors thought to affect migration time, those such as CD read rate, hard disk (hereinafter "HDD" in tables) speed and bus speed can be compared using transfer rates. Comparing these transfer rates tells us that CD read rates are extremely slow.

Table 3.3-4 Estimated transfer rates

Comparison of speed by device	Commonly used performance indicator	Standard specification	Estimated transfer rate
CD read rate	x speed	40x	6MB/s
HDD speed	rpm (revolutions per minute)	5,400rpm	50MB/s
Bus speed	MB/s	533MB/s (PCI)	533MB/s

- This is a general estimation based on the standard platform. Actual rates vary depending on manufacturer and model.

From the above results, we can speculate that of the technological factors that can be compared using indications of performance expressed as a transfer speed, the factor that has the greatest effect on migration time is a CD read rate.

Furthermore, because a bus speed is so much faster than a CD read rate, we surmised that a bus speed would not affect migration time. Likewise, a hard disk speed was thought not to affect migration time, but it was included in the investigation in order to confirm that migration speed is decided by the technological factor with the slowest transfer rate.

In addition to the transfer rates, both CPU performance and memory size may affect migration time. However, it is impossible to compare these technological factors by a single standard and it is difficult to imagine that either would cause a constant bottleneck for migration. Therefore, instead of selecting only one or the other, we decided to include both factors for the investigation.

Based on the observations made above, we decided to investigate the effects on migration time of the following four technological factors:

(i) CD read rate
(ii) Hard disk speed
(iii) CPU performance
(iv) Memory size

(1) Testing methods
In order to gain an understanding of how changes in these technological factors affected migration time, all other variables were fixed. Then the time required to complete the migration of 99 MB, 351 MB and 649 MB[21] CD-ROMs was measured while varying only the factors under investigation. Because the migration program allows the user to specify whether data is to be compressed during migration or not, the experiment was performed both with and without compression.

(2) Effects of the CD read rate
Migration time was measured with the CD read rate set at 24x and 40x.

[21] Small, medium and large samples were selected based on the distribution of actual data transferred during migration.

Table 3.3-5 Migration time versus CD read rate

		CD read rate	HDD	CPU	Memory	Compression	Processing time (seconds)		
							99MB	351MB	649MB
CD read rate	Environment 1	24x	5400rpm	1.7GHz	1024MB	No	50	142	250
						Yes	50	142	250
	Environment 2	40x				No	31	85	150
						Yes	31	86	180

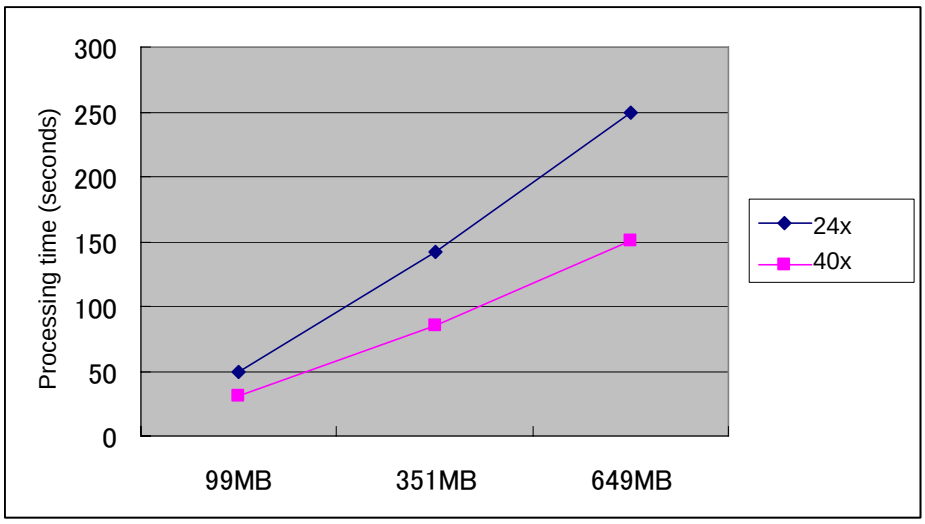

Figure 3.3-4 Migration time versus CD read rate (without compression)

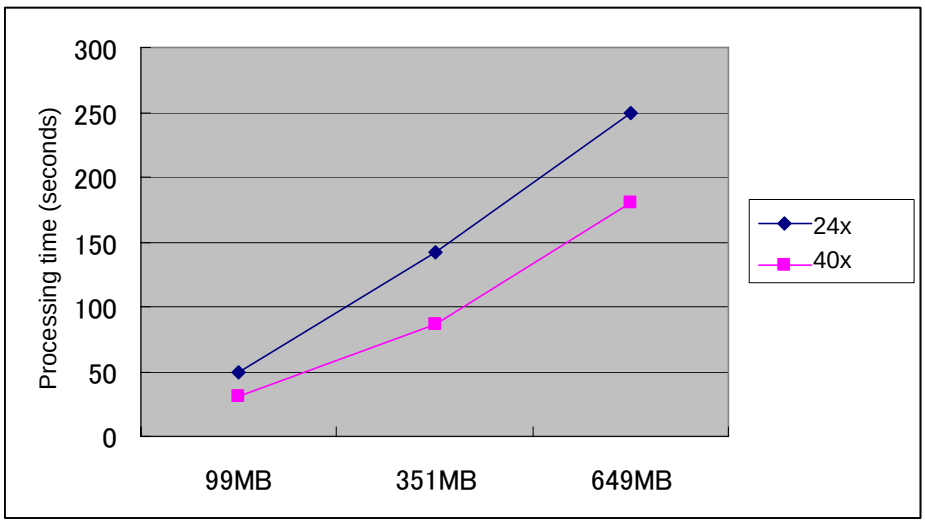

Figure 3.3-5 Migration time versus CD read rate (with compression)

Measurements showed that as the CD read rate increased, there was a clear reduction in migration time.

In relatively fast CD drives, data are often read using the constant angular velocity method.[22] When using this method, data transfer rates are slow at the center of a CD-ROM, but data are read faster towards the edge of the disc. Therefore, the average data transfer rate of a CD-ROM depends on the diameter of the area occupied by data on the disc.

[22] This is a method by which the CD is rotated at a constant speed while data is read.

The approximate processing time when using the constant angular velocity method was calculated based on the time taken to complete migration of a CD-ROM containing 351 MB of data, which was considered an average amount.

351 [MB] / 85 [s] = 4 [MB/s]	:	The amount of data that can be read per second when operating at 40x.
4 [MB/s] / 40 [x] = 0.1 [MB/s][23]	:	The amount of data that can be read per second when operating at the equivalent of 1x.

Migration time [s] =
CD-ROM data size [MB] / (0.1 [MB/s] x drive speed [x])

However, CD-ROMs containing little data will take longer to process than estimated with this formula[24], while it will take less time than estimated to process CD-ROMs containing a lot of data[25]. This is because average transfer rates are slower for smaller amounts and faster for larger amounts of data.

Migration of a CD-ROM containing 649 MB at 40x with compression took longer than estimated because the CPU could not keep up with processing tasks, resulting in a longer migration time. (See "(4) The effects of CPU performance.")

[23] Using 24x gives the same answer: 351 MB / 142 s = 2.47 MB/s and 2.47 MB/s / 24 [x] = 0.1 MB/s.
[24] In the case of a disc containing 99 MB, which is less than the average amount of 351 MB, the formula gives 99 MB / 4 MB/s = 25 seconds, whereas the actual figure is 31 seconds, or 6 seconds longer than the estimated time.
[25] In the case of a disc containing 641 MB, which is more than the average amount of 351 MB, the formula gives 641 MB / 4 MB/s = 160 seconds, whereas the actual figure is 150 seconds, or 10 seconds less than the estimated time.

(3) Effects of hard disk speed

The effects of hard disk performance on migration time were investigated. As predicted, changes in hard disk speed within the normal range had virtually no effect on migration time.

Table 3.3-6 Migration time versus hard disk speed

		CD read rate	HDD	CPU	Memory	Compression	Processing time (seconds)		
							99MB	351MB	649MB
HDD speed	Environment 2	40x	5400rpm	1.7GHz	1024MB	No	31	86	150
						Yes	30	86	179
	Environment 1		7200rpm			No	30	86	150
						Yes	31	86	181

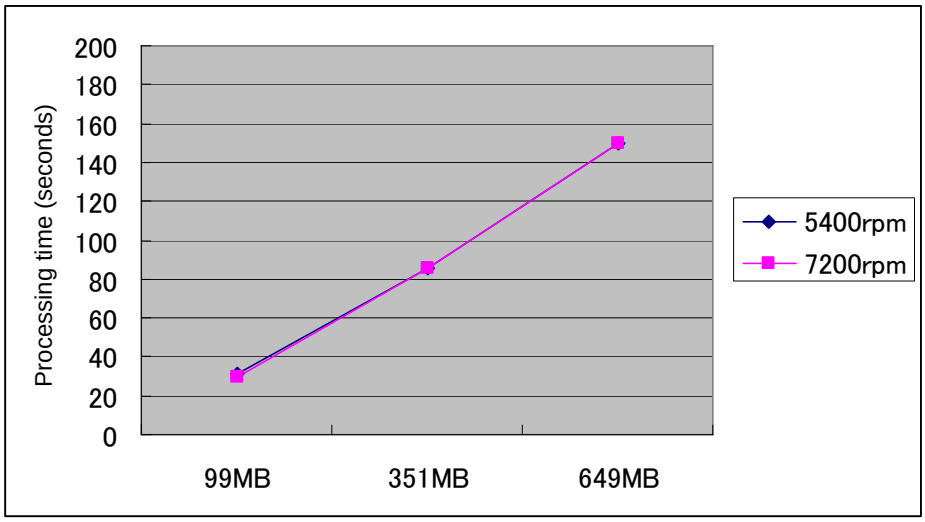

Figure 3.3-6 Migration time versus hard disk speed (without compression)

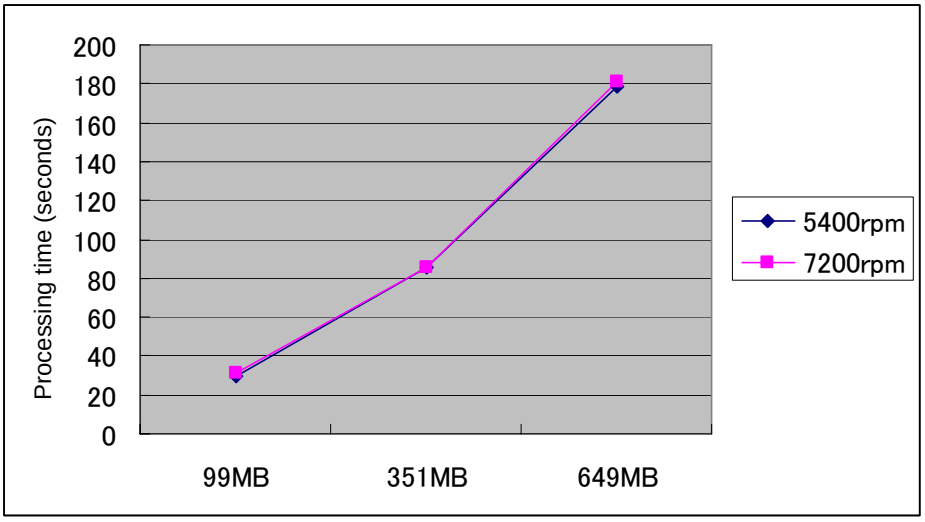

Figure 3.3-7 Migration time versus hard disk speed (with compression)

(4) Effects of CPU performance
The effects of CPU performance on migration time were investigated.

Table 3.3-7 Migration time versus CPU performance 1

		CD read rate	HDD	CPU	Memory	Compression	Processing time (seconds)		
							99MB	351MB	649MB
CPU	Environment 1	40x	5400rpm	1.3GHz	512MB	No	31	86	150
						Yes	30	95	217
	Environment 2			1.7GHz		No	30	87	150
						Yes	31	86	179

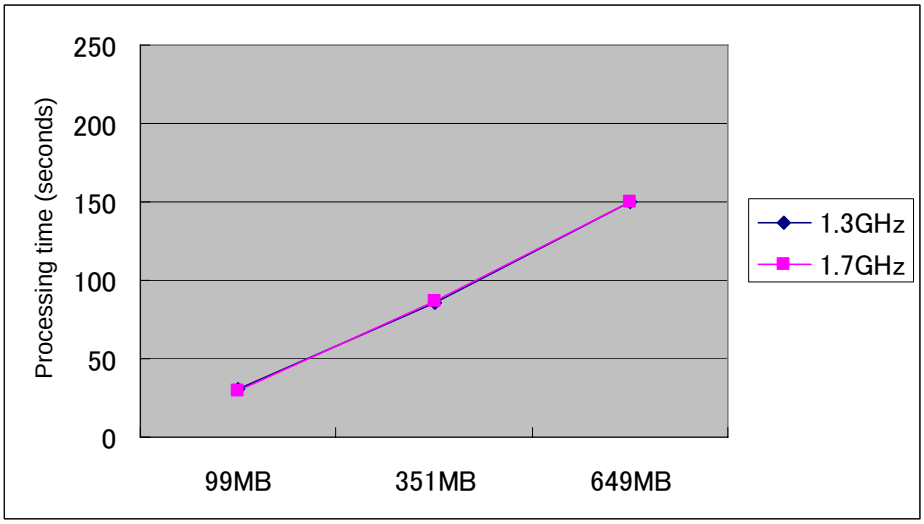

Figure 3.3-8 Migration time versus CPU performance 1 (without compression)

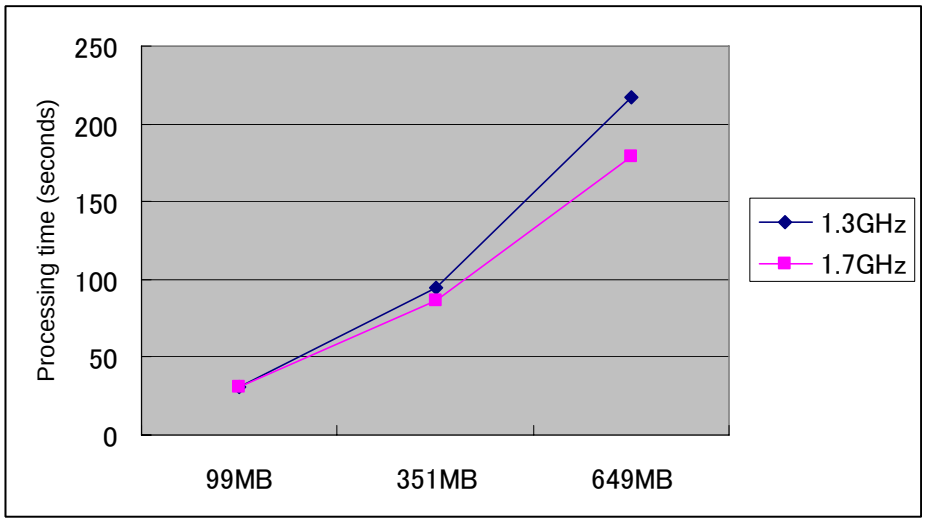

Figure 3.3-9 Migration time versus CPU performance 1 (with compression)

When a CD image is not compressed, there is little difference between migration time of low performance CPU and that of high performance CPU. This is probably because when compression is not performed, the CPU is placed under virtually no load and data read from the CD-ROM can be written to the hard disk without delay.

When a CD image is compressed, and when the amount of data per second being sent from

the CD drive increases, the results indicate a tendency for migration time to increase due to loading of the CPU when a low-performance CPU is used.

Because data is recorded from the center on a CD-ROM, data transfer rates are slower for CD-ROMs containing little data, when being read using the constant angular velocity method. In these cases, the presence or absence of compression is thought not to affect migration time because the rate at which data is read is lower than the rate at which the data can be compressed.

However, as the amount of data recorded on a CD-ROM increases, the area occupied by data approaches the outside edge of the disc, resulting in faster data transfer rates. Depending on CD read rates and CPU performance, the rate at which data is read may exceed the rate at which it can be compressed, thereby affecting migration time.

In order to confirm this, the rate at which data is supplied to the CPU was lowered by setting the CD read rate to a slow 8x, and the time required for migration (with compression) was measured.

Table 3.3-8 Migration time versus CPU performance 2

		CD read rate	HDD	CPU	Memory	Compression	Processing time (seconds)		
							99MB	351MB	649MB
CPU	Environment 1	8x	5400rpm	1.3GHz	512MB	Yes	73	261	553
	Environment 2			1.7GHz		Yes	73	260	554

The results confirm that when the rate at which data is supplied to the CPU does not exceed the processing capabilities of the CPU, then CPU performance does not affect migration time.

Based on these observations, the following conclusions can be drawn in relation to variations in CPU performance.

(i) When the CD image is not compressed:
CPU performance will not affect migration time.

(ii) When the CD image is compressed:
If the rate at which data is supplied to the CPU falls below the processing capabilities of the CPU, then CPU performance will not affect migration time. However, if the rate at which data is supplied to the CPU exceeds the processing capabilities of the CPU, then CPU performance will affect migration time.

(5) Effects of memory size

The effects of memory size on migration time were investigated. Changes in memory size within the normal range had virtually no effect on migration time.

Table 3.3-9 Migration time versus memory size

		CD read rate	HDD	CPU	Memory	Compression	Processing time (seconds)		
							99MB	351MB	649MB
Memory	Environment 1	40x	5400rpm	1.3GHz	256MB	No	30	86	150
						Yes	32	95	217
	Environment 2				512MB	No	31	86	150
						Yes	31	94	218

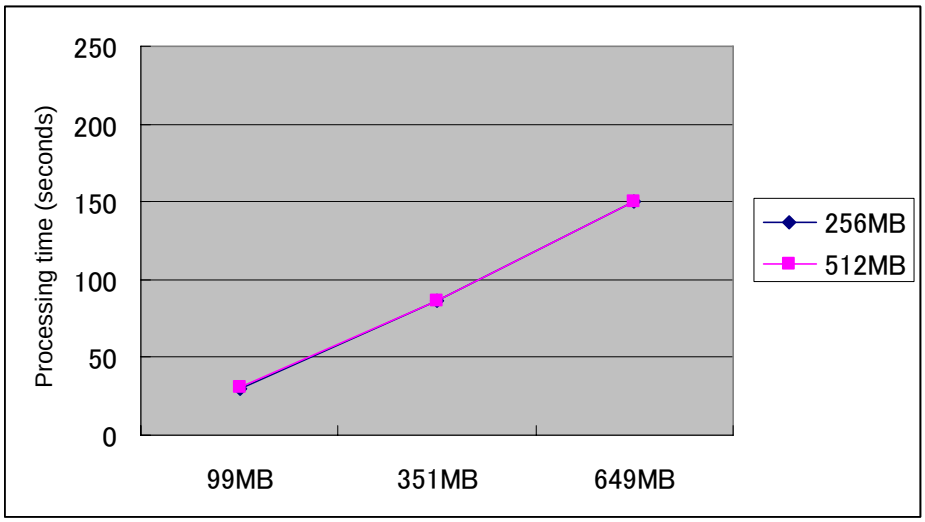

Figure 3.3-10 Migration time versus memory size (without compression)

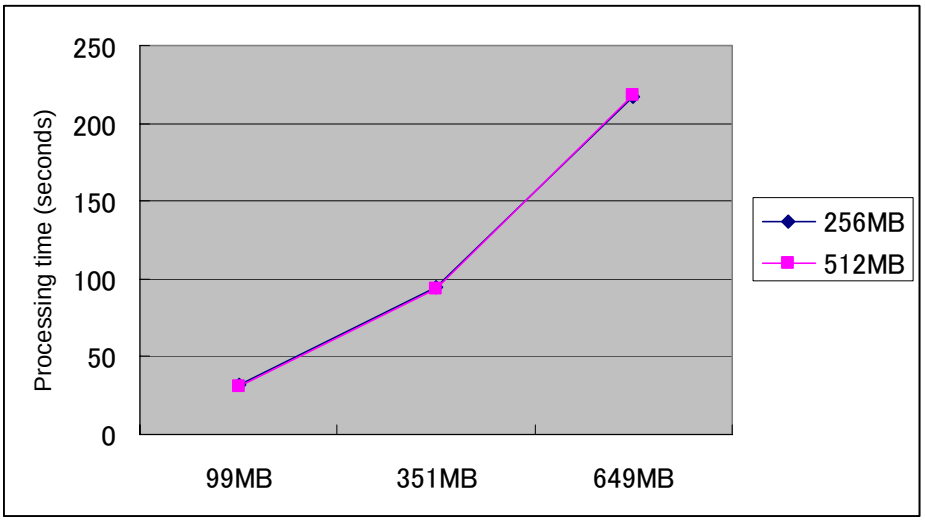

Figure 3.3-11 Migration time versus memory size (with compression)

3.3.4. Discussion

(i) Performing Migration
Migration of the digital materials was performed using a commercially available migration program. The fact that virtually all the digital materials were converted successfully suggests that migration itself can be performed with relative ease.

(ii) File Format after Migration
Because the file formats created on a hard disk during migration are specific to the migration program, problems can occur with respect to the long-term accessibility of digital materials after migration. Continued use of files recorded in these formats requires the continued use of the same migration program, but upgrades to operating systems and other changes may prevent the continued use of such a program. It is also possible that upgraded versions of a migration program will not support files created by previous versions, and that sales of the migration program being used will be discontinued.
Alternatively, migration might be performed without a migration program by making copies of individual files. However, with this method it may not be possible to overcome various copy protection schemes, and files may not be accessible from hard disks because they were designed to be played from a CD drive.

(iii) Time Required for Migration
The time required for the complete migration of one CD is dependent on the volume of data on the CD-ROM and the read rate of the CD drive. If data are compressed during migration, CPU performance may also be a factor. Compared with other technological factors, the extremely slow read rates of CD drives are a major issue. Due to physical factors involved in spinning media, such as vibration and air resistance, this read technology is believed to have already reached its limits and further significant progress is considered unlikely. Therefore, the CD read rate will continue to be the deciding factor in determining the time required for migration in the future.

(iv) Migration from Other Media
Although CD-ROMs were the subject of this investigation, the large number of DVDs being acquired and stored suggests that there is a need for a similar investigation of DVDs.

(v) Requirements for Migration Programs
Since the aim of migration is the long term preservation and accessibility of digital materials, migration programs must satisfy the following requirements.

- They must remain functional for a long period of time, through upgrading or other measures.
- The file formats created by migration must be usable over the long term. They must not be made unusable as a result of upgrades to migration programs.
- It must be possible to handle the files created by migration as if they existed on a CD-ROM.
- It must also be possible to migrate CD-ROMs that contain copy protection features.

However, at present, solutions are not yet available for all these issues, and continued investigation is required.

3.4. Playback testing

3.4.1. Emulation and file format conversion

Emulation and file format conversion are referred to as methods for ensuring the long-term accessibility of digital material. Both these methods were tested and their effectiveness and any associated problems investigated.

Programs contained in digital materials are designed to work with specific operating systems. These operating systems themselves are also designed to work on specific hardware. Therefore, when considering the long-term preservation of digital materials, it is vital to consider also the lifespan of hardware. Even if programs contained in digital materials could be preserved for decades or centuries, they would not be playable without the appropriate hardware. The process of emulation is generally considered to be an effective way of overcoming this problem.

Emulation involves installing an application called an emulator in a piece of hardware in order to artificially recreate an operating environment equivalent to that of another piece of hardware. In other words, by installing an emulator, we can make recent hardware able to imitate older versions. In this investigation we installed an emulator in the latest hardware and tested digital materials containing programs in this emulated operating environment to see if they would play properly or not.

On the other hand, digital materials that contain only data are often designed not for a particular operating system, but rather to be played using a specific application. If such applications cease to function in new operating environments in the future due to discontinuation of sales or support, such digital materials will no longer be playable. This problem can be avoided by converting file formats in advance so that they can be played using other applications. In this investigation, digital materials containing only data were subjected to file format conversion as well as playback in a multi-file viewer.

In this investigation, digital materials containing programs were defined as digital materials for which playback requires an application made only for that material or for materials from the same series.

Digital materials containing only data were defined as digital materials designed to be played using an application that is currently or was once distributed.[26]

[26] Resource collections etc. were classified as digital materials containing only data because although they contain proprietary applications for playback, they can also be played normally on general applications. Digital materials accompanied by programs that supply additional functions, such as search plug-ins, were also classified as digital materials containing only data because with the exception of add-on functions, they can be played normally using general applications.

3.4.2. Emulation playback tests

3.4.2.1. Playback testing method

A commercially available emulator[27] was used to create guest operating environments[28] for each superseded operating system to run in. The superseded operating systems were installed in these environments and the superseded operating environments necessary for the playback of the digital materials were constructed.

The sample for the playback tests consisted of 100 digital materials[29] containing programs selected from the digital materials that were created on hard disk as a result of migration.

Playback tests involved actually loading each material, establishing the presence or absence of the following functions, and checking if each function operates correctly.[30]

Table 3.4-1 Details of emulation playback tests

No.	Function	Presence or absence of function	Operation check
1	Installation	Is installation necessary for playback of the material?	Is installation completed correctly?
2	Start-up screen	Is there a screen that is displayed after booting?	Is the start-up screen displayed correctly after booting?
3	Response to instructions	What buttons and keys initiate actions?	Does the program respond correctly to the actions mentioned on the left?
4	Still images	Are there still images such as photographs, text or slides?	Are there any garbled characters? Are images free from disturbance?
5	Video	Is there any video such as animations or movies?	Are there any garbled characters? Are images free from disturbance?
6	Audio	Are there any sounds?	Is there any static or distortion?
7	Calculations and judgments	Are there any logic functions that perform calculations or make judgments?	Are calculations being made correctly?
8	Search and extraction	Are there any search or extraction functions?	Do search and extraction results correspond to instructions by a user?
9	Data records	Are there any functions that allow data to be saved until the next time the program is run?	Are the data saved?
10	Other	If there are any functions other than those listed above, establish an outline of them.	Do the functions mentioned at the left work correctly?

- Step 2, confirmation of start-up screen, was not performed for digital materials that could not be installed correctly in step 1. Likewise, steps 3 to 10 were not performed for materials that did not demonstrate correct operation in step 1, installation, and step 2, start-up screen.

[27] VMware Workstation 4.5.2 (by VMware Inc.) was used. Our aim was not to evaluate the product, but to generate generalized results.
[28] Simulated hardware environments created by the emulator and in which each operating system runs.
[29] 126 materials, including reserves, had already been selected at random for migration. Of these materials, those containing programs (those with sample ID codes stating with "P" in later sections) were the base from which this sample was selected.
[30] Correct playback can be confirmed in a non-emulated environment, but in this investigation testing was only carried out in the emulated environment.

3.4.2.2. Results of playback tests

The results of playback tests conducted on the sample of 100 digital materials are shown below. “Partially unplayable” refers to materials which, although playable, had problems with some functions, such as distortion in sound. “Unplayable” refers to materials which could not be installed and materials whose start-up screens did not display.

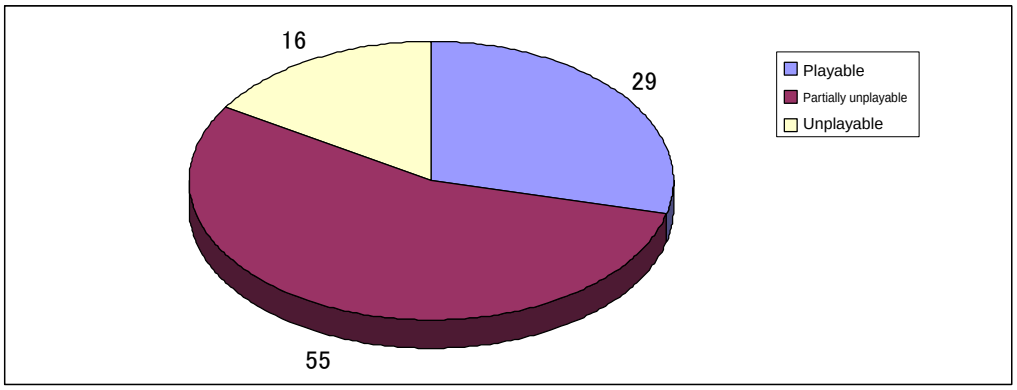

Figure 3.4-1 Results of emulation playback test

3.4.2.3. Playback results for each operating system

Information printed in instruction manuals and others accompanying each digital material was used to establish the appropriate operating system for its playback. In cases where more than one compatible operating system was listed, one operating system was selected according to the following rules.

(i) Choose the newest operating system.
For example, if Windows 95 and Windows 98 are listed, Windows 98 should be chosen.
(ii) Choose the most widely used operating system.
For example, if Windows 9x and NT systems are both listed, the Windows 9x system should be chosen.

The proportion accounted for by each operating system of digital materials treated in this investigation that contain programs is shown below.

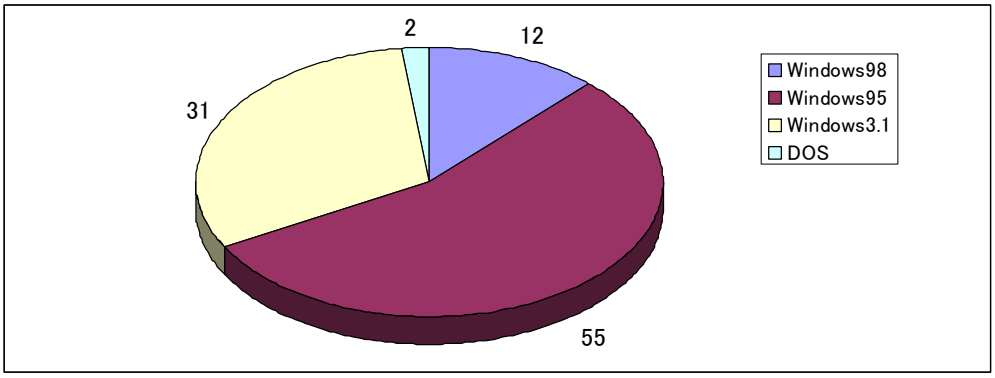

Figure 3.4-2 Distribution of operating systems that were emulated

Playback results for each operating system are described below.

(1) Windows 98

To construct the necessary playback environment, Windows 98 was installed in the guest operating environment and playback was tested. Probably because Windows 98 is a relatively new operating system, none of the materials tested were found to be unplayable, although one was found to be partially unplayable.

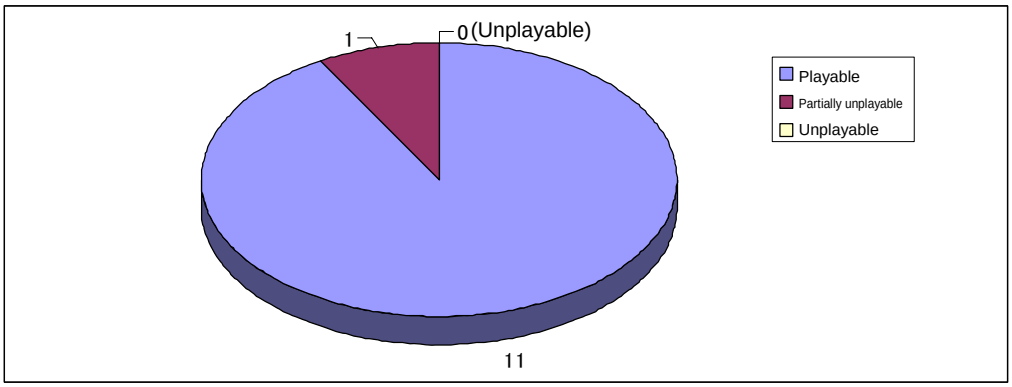

Figure 3.4-3 Playback test results for Windows 98

Table 3.4-2 Playback test results for each digital material in Windows 98

Sample ID	Installation	Start-up screen	Response to instructions	Still images	Video	Audio	Calculations and judgments	Search and extraction	Data records	Other	Playback result
P01	X	X	X	X							X
P02	X	X	X	X				X			X
P03	X	X	X	X	X	X		X			X
P04	X	X	X	X							X
P05	X	X	X	X		X	X		X		X
P06	X	X	X	X	X	X	X				X
P07	X	X	X	X	X	X					X
P08		X	X	X	X	P					P
P09	X	X	X	X		X		X		X	X
P10	X	X	X	X				X			X
P11	X	X	X	X				X			X
P12	X	X	X	X							X

X: Playable, P: Partially unplayable, U: Unplayable, --: Presence of function unconfirmed, Blank: Absence of function

Table 3.4-3 Problems with playback on Windows 98

Function	Number of materials with this function	Number of materials which have problem with this function	Percent of materials showing problem
Installation	11	0	
Start-up screen	12	0	
Response to instructions	12	0	
Still images	12	0	
Video	4	0	
Audio	6	1	17%
Calculations and judgments	2	0	
Search and extraction	5	0	
Data records	1	0	
Other	1	0	

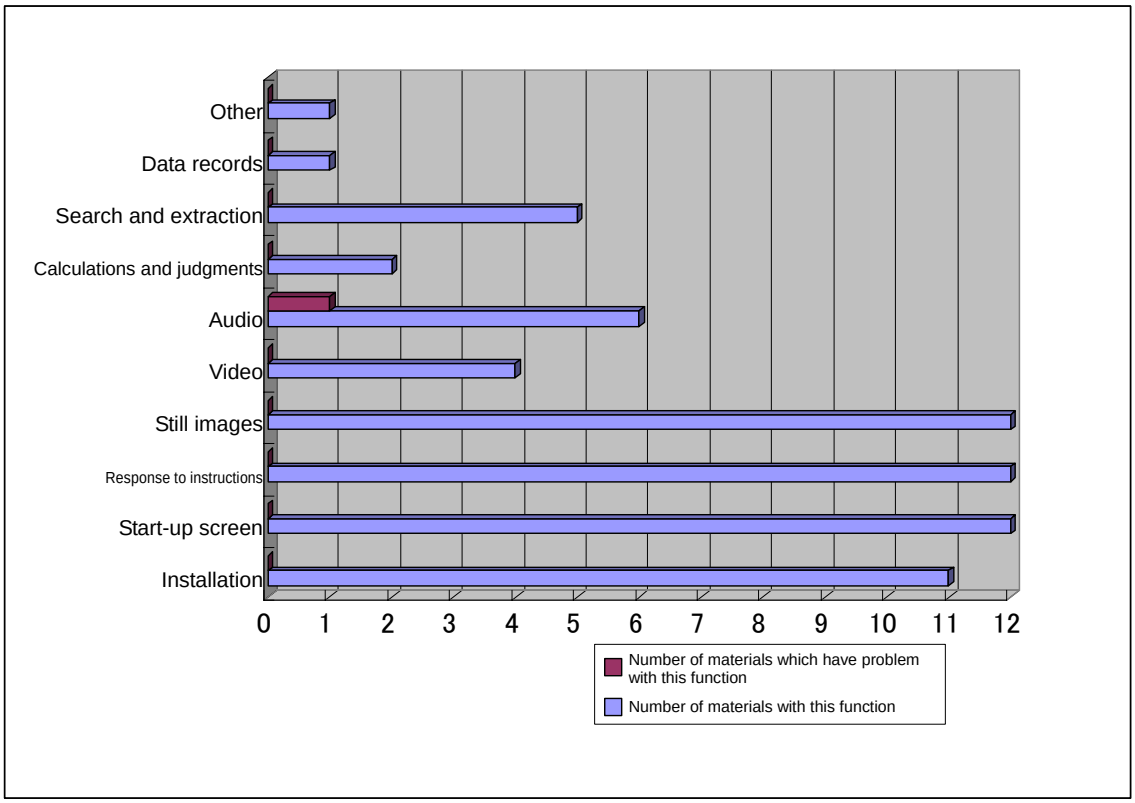

Fig. 3.4-4 Problems with playback on Windows 98

(2) Windows 95

To construct the necessary playback environment, Windows 95 was installed in the guest operating environment and playback was tested. Materials for Windows 95 were most affected by audio problems. In all cases where materials were unplayable, errors occurred just before the start-up screen appears. Most of the problems with materials that were partially unplayable were audio related (distortion, instability of playback speed).

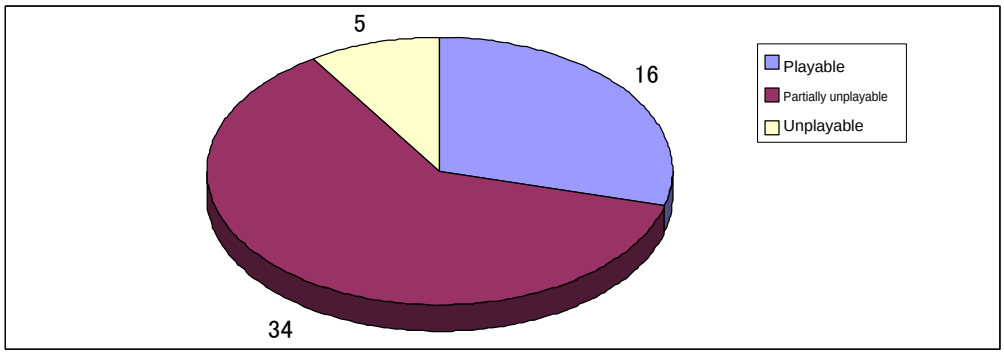

Figure 3.4-5 Playback test results for Windows 95

Table 3.4-4 Playback test results for each digital material in Windows 95

Sample ID	Installation	Start-up screen	Response to instructions	Still images	Video	Audio	Calculations and judgments	Search and extraction	Data records	Other	Playback result
P13	X	X	X	X	X	P					P
P14	X	X	X	X		P					P
P15		X	X	X	X	P					P
P16	X	X	P	X	X	X		X			P
P17	X	X	X	X		P					P
P18		X	X	X	X	P					P
P19		X	X	X	X	P					P
P20		X	X	X	X	P					P
P21		X	X	X				X			X
P22		X	X	X		P					P
P23	X	X	X	X		P					P
P24	X	X	X	P				X			P
P25		X	X	X		P	X				P
P26	X	U	--	--	--	--	--	--	--	--	U
P27		X	X	X	X	P					P
P28	X	X	X	X				X			X
P29		X	X	X		P					P
P30		X	X	X	X	P			U	U	P
P31	X	X	X	X	X	P					P
P32	X	X	X	X	X	P				X	P
P33	X	X	X	X	X	X		X	X		X
P34	X	X	X	X		X					X
P35	X	X	X	X		P		X			P
P36	X	X	X	X				X			X
P37	X	X	X	X	X	X					X
P38	X	X	X	X						X	X
P39		X	X	X		P	X	X			P

Sample ID	Installation	Start-up screen	Response to instructions	Still images	Video	Audio	Calculations and judgments	Search and extraction	Data records	Other	Playback result
P40	X	X	X	X		P					P
P41	X	U	--	--	--	--	--	--	--	--	U
P42	X	X	X	X	X	P					P
P43		X	X	X	X	P					P
P44	X	X	X	X		X					X
P45	X	X	X	X		X					X
P46	X	X	X	X			X				X
P47	X	X	X	X				X			X
P48	X	U	--	--	--	--	--	--	--	--	U
P49		X	X	X	X	P					P
P50	X	X	X	X		P	X		X		P
P51	X	X	X	X		X		X			X
P52	X	X	X	X		P	X		X		P
P53		X	X	X		P					P
P54	X	X	X	X	X	P					P
P55	X	X	X	X		X					X
P56	X	X	X	X				X			X
P57	X	U	--	--	--	--	--	--	--	--	U
P58		X	X	X		P					P
P59	X	X	X	X	X	P			U		P
P60	X	X	P	X		P					P
P61	X	X	X	X	X						X
P62		X	X	X	X	P					P
P63		X	X	X	X	P		X			P
P64		X	X	X	X	P					P
P65	X	X	X	X	X	X					X
P66	X	U	--	--	--	--	--	--	--	--	U
P67	X	X	X	X		P		X			P

X: Playable, P: Partially unplayable, U: Unplayable, --: Presence of function unconfirmed,
Blank: Absence of function

Table 3.4-5 Problems with playback on Windows 95

Function	Number of materials with this function	Number of materials which have problem with this function	Percent of materials showing problem
Installation	37	0	
Start-up screen	55	5	9%
Response to instructions	50	2	4%
Still images	50	1	2%
Video	22	0	
Audio	41	32	78%
Calculations and judgments	5	0	
Search and extraction	13	0	
Data records	5	2	40%
Other	3	1	33%

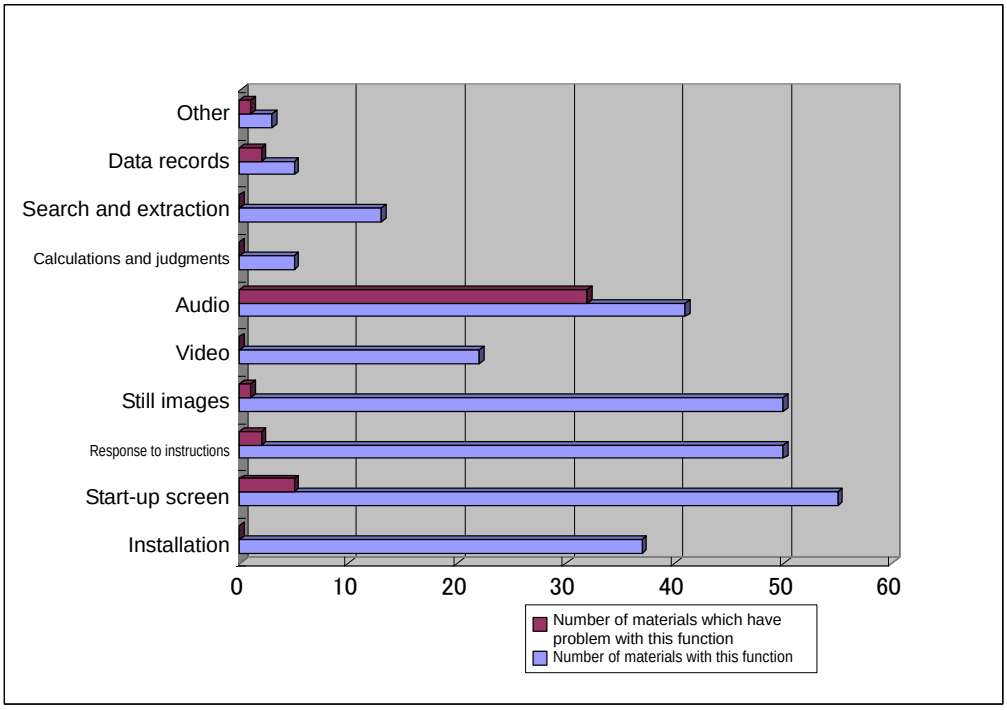

Fig. 3.4-6 Problems with playback on Windows 95

(3) Windows 3.1 and DOS

To construct the necessary playback environment, DOS (J6.1/V) and Windows 3.1 were installed in the guest operating environment and playback was tested. Because the following problems were encountered when constructing playback environments in Windows 3.1 and DOS, very few materials were able to be completely played.

(i) Screen problems

Some materials could not be played properly because in the guest operating environment for Windows 3.1 and DOS, the emulator only supported VGA (640x480 pixels in 16 colors).

(ii) Audio problems

Windows 3.1 and DOS drivers that were compatible with the sound card used by the guest operating environment could not be obtained. Therefore, sounds could not be played.

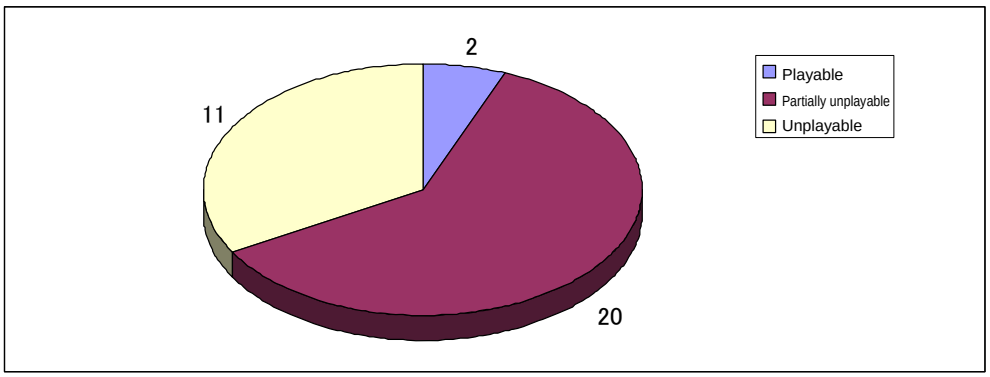

Figure 3.4-7 Playback test results for Windows 3.1 and DOS

Table 3.4-6 Playback test results for each digital material in Windows 3.1 and DOS

Sample ID	Installation	Start-up screen	Response to instructions	Still images	Video	Audio	Calculations and judgments	Search and extraction	Data records	Other	Playback result
P68	U	--	--	--	--	--	--	--	--	--	U
P69	X	X	X	X				X			X
P70	X	U	--	--	--	--	--	--	--	--	U
P71	X	U	--	--	--	--	--	--	--	--	U
P72	X	U	--	--	--	--	--	--	--	--	U
P73	X	X	X	P		U					P
P74	X	X	X	P		U					P
P75	X	X	X	P		U		X			P
P76	X	U	--	--	--	--	--	--	--	--	U
P77	X	X	X	P	P	U					P
P78	X	X	X	P	P	U					P
P79	X	X	X	P	P	U					P
P80	X	U	--	--	--	--	--	--	--	--	U
P81	X	U	--	--	--	--	--	--	--	--	U
P82	X	X	X	P	P	U		X			P
P83	X	X	P	P	P	U					P
P84	X	U	--	--	--	--	--	--	--	--	U
P85	X	X	X	P				X			P
P86	X	X	X	P	P	U					P
P87	X	X	X	P	P	U		X			P
P88	X	X	X	P	P	U					P
P89	X	X	X	P	P	U					P
P90	X	X	X	P	P	U					P
P91	X	X	X	P	P	U					P
P92	X	X	X	P	P	U					P
P93	X	X	X	P	P	U					P
P94	X	U	--	--	--	--	--	--	--	--	U
P95	X	X	X	X	X	U	X		X		P
P96	X	X	X	P	P	U	X		X		P
P97	U	--	--	--	--	--	--	--	--	--	U
P98		X	X	P	P	U					P
P99	X	X	X	X	X			X	X		X
P100	X	U	--	--	--	--	--	--	--	--	U

X: Playable, P: Partially unplayable, U: Unplayable, --: Presence of function unconfirmed,
Blank: Absence of function

Table 3.4-7 Problems with playback on Windows 3.1 and DOS

Function	Number of materials with this function	Number of materials which have problem with this function	Percent of materials showing problem
Installation	32	2	6%
Start-up screen	31	9	29%
Response to instructions	22	1	5%
Still images	22	19	86%
Video	17	15	88%
Audio	19	19	100%
Calculations and judgments	2	0	0%
Search and extraction	6	0	0%
Data records	3	0	0%
Other	0	0	0%

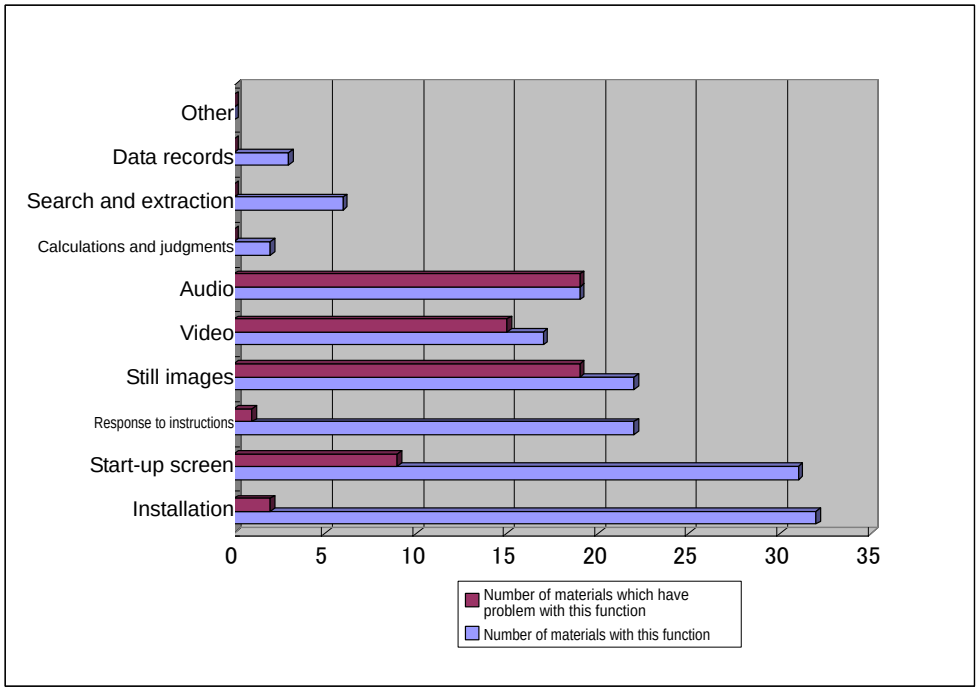

Fig. 3.4-8 Problems with playback on Windows 3.1 and DOS

3.4.3. Playback tests using file format conversion and a multi-file viewer

Of the digital materials migrated to hard disk, a sample of 100 digital materials[31] was selected from those containing only data. These materials were then subjected to file format conversion and playback tests, as well as playback tests using a multi-file viewer.[32]

3.4.3.1. Playback tests

(1) Types of digital materials
The results of the above tests revealed that the content of the sample of 100 data-only materials could be broadly divided into the following four categories:

(i) Resource collections etc.
Items in this category contain files of various different formats, but the files are not interrelated and are used independently. Users select a file they need from many contained in the material.

(ii) HTML documents for playback in browsers
Files in the HTML format are not played as independent HTML files, but rather they are combined with and linked to various types of data files necessary for playback. In addition to hyperlinks between HTML files themselves, many files use combinations of images, video, audio and data from other applications.

(iii) Electronic books and dictionaries
Electronic books comprise data and also control information. The accompanying control information is used to express the data correctly and allows the data to be searched. EPWING (JIS-X4081 Retrieval data structure for Japanese electronic publication) and J-BISC (NDL Catalogue) belong to this category.

(iv) Other
Items that do not fit any of the categories listed above. For example, there were some items that consisted of PDF data with attached video files.

[31] Of the 352 digital materials that had been migrated, there were 110 that contained only data. 100 items (those with sample ID codes stating with "D" in later sections) were selected at random from among these 110 materials.

[32] File format conversion was performed using Rich Text Converter 2004 for Windows (by Antenna House Inc.). The multi-file viewer used was Multi File Viewer Jizaigan 8 (by Antenna House Inc.). Our aim was not to evaluate the product, but to generate generalized results.

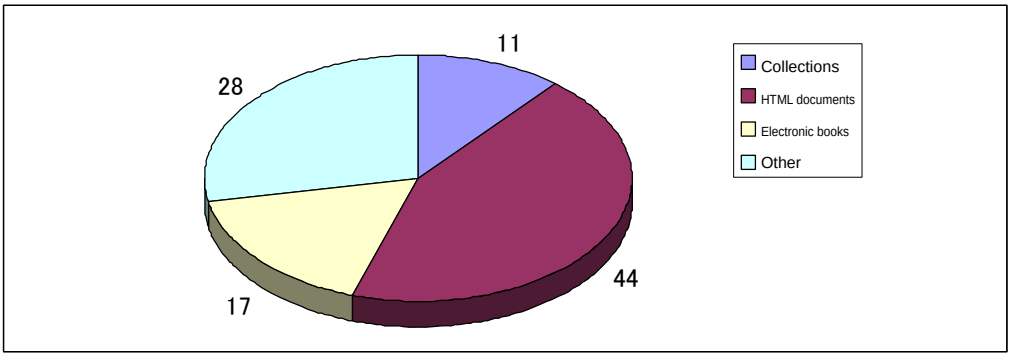

Figure 3.4-9 Types of materials that contain data only

File format conversion by file format conversion programs is conducted for one file at a time. Therefore, when converting files from category "(ii) HTML documents for playback in browsers," links to other connected files are lost. Even if they can be converted file by file, the content will no longer be playable in its original form. Furthermore, items from category "(iii) Electronic books" rely on close connections between data because control information is constantly referred to while the text data itself is played. Individual files thus cannot be extracted and used independently. In summary, file format conversion is difficult or impossible for items from categories "(ii) HTML documents for playback in browsers" and "(iii) Electronic books."

Multi-file viewers also display files one at a time. It follows that items from"(ii) HTML documents for playback in browsers" and "(iii) Electronic books" cannot be played in multi-file viewers for the same reasons they cannot be converted.

(2) Types of file formats contained in digital materials

There are many different types of file formats, and they can be divided into standardized and non-standardized. If a file format has been standardized and its specifications made public, it is likely that more than one compatible application exists, making it relatively easy to achieve long-term accessibility. It follows that if a format has been standardized, there is little need to perform file format conversion. In order to identify digital materials containing files suitable for file format conversion, the file formats contained by each material were investigated.

It would be difficult to list all the file formats contained in the digital materials examined, so in the following table file formats were classified as "HTML," "image," "video," "audio," "Microsoft Office document" (hereinafter "Office"), "text," "electronic book" or "other." Within each category a distinction was also made between file formats for which standards exist and those for which they do not exist. File formats with standards were defined as formats for which data and text formats have been made public, and for which a standard for creating data exists.

Figure 3.4-8 Types of file formats contained in digital materials

Sample ID	Category	HTML		Image		Video		Audio		Office	Text		Electro-nic book		Other	
		Standardized	Non-standardized	Standardized	Non-standardized	Standardized	Non-standardized	Standardized	Non-standardized	Non-standardized	Standardized	Non-standardized	Standardized	Non-standardized	Standardized	Non-standardized
D01	Electronic book													X		
D02	Collections			X												
D03	Collections			X							X					
D04	Electronic book													X		
D05	Collections			X							X					
D06	Collections			X							X					
D07	Collections			X			X									
D08	Collections			X							X					
D09	Electronic book										X			X		
D10	Collections			X							X					
D11	Electronic book										X		X		X	
D12	Browser playable	X		X				X								
D13	Browser playable	X														
D14	Browser playable	X								X					X	
D15	Electronic book												X		X	
D16	Browser playable	X														
D17	Electronic book			X							X		X			
D18	Electronic book										X		X		X	
D19	Collections			X				X			X				X	
D20	Other										X					
D21	Browser playable	X									X					
D22	Browser playable	X								X						X
D23	Browser playable	X									X					X
D24	Other			X							X					X
D25	Other			X							X					
D26	Other										X					X
D27	Browser playable	X				X										

Sample ID	Category	HTML		Image		Video		Audio		Office	Text		Electro-nic book		Other	
		Standardized	Non-standardized	Standardized	Non-standardized	Standardized	Non-standardized	Standardized	Non-standardized	Non-standardized	Standardized	Non-standardized	Standardized	Non-standardized	Standardized	Non-standardized
D28	Other										X				X	
D29	Browser playable	X		X	X					X	X				X	X
D30	Other			X			X				X					X
D31	Other									X	X				X	
D32	Browser playable	X	X								X					X
D33	Browser playable	X	X	X		X	X	X			X				X	X
D34	Browser playable	X	X								X					
D35	Browser playable	X		X			X	X		X	X				X	X
D36	Browser playable	X									X					X
D37	Other									X	X				X	
D38	Browser playable	X	X	X			X	X			X				X	X
D39	Electronic book										X		X		X	
D40	Electronic book												X			
D41	Browser playable	X	X													
D42	Other										X					X
D43	Browser playable	X								X	X					
D44	Other			X			X				X				X	X
D45	Browser playable	X		X							X	X				
D46	Browser playable	X									X					X
D47	Collections									X	X					
D48	Electronic book					X	X	X			X		X		X	
D49	Other						X				X				X	X
D50	Electronic book												X		X	
D51	Browser playable	X									X					
D52	Browser playable	X	X													
D53	Browser playable	X		X	X					X	X	X			X	X
D54	Electronic book			X			X	X			X			X		
D55	Other															X
D56	Browser playable	X	X								X				X	
D57	Other						X				X				X	X
D58	Electronic book												X		X	
D59	Browser playable	X														
D60	Other										X				X	
D61	Browser playable	X		X							X					
D62	Other							X		X	X					
D63	Electronic book												X		X	
D64	Collections			X				X			X				X	
D65	Browser playable	X														X
D66	Browser playable	X								X	X				X	X
D67	Other										X				X	
D68	Other						X				X				X	X

Sample ID	Category	HTML		Image		Video		Audio		Office	Text		Electro-nic book		Other	
		Standardized	Non-standardized	Standardized	Non-standardized	Standardized	Non-standardized	Standardized	Non-standardized	Non-standardized	Standardized	Non-standardized	Standardized	Non-standardized	Standardized	Non-standardized
D69	Electronic book										X		X		X	
D70	Other			X			X			X	X				X	X
D71	Browser playable	X	X	X			X	X	X	X	X				X	
D72	Browser playable	X		X						X	X	X			X	X
D73	Browser playable	X														
D74	Other			X							X					X
D75	Browser playable	X		X	X	X	X			X	X	X			X	X
D76	Browser playable	X								X	X				X	X
D77	Other										X				X	X
D78	Other										X					X
D79	Browser playable	X								X	X					X
D80	Browser playable	X									X					
D81	Other			X	X		X									
D82	Other					X					X					X
D83	Browser playable	X		X		X										
D84	Browser playable	X														X
D85	Other															X
D86	Browser playable	X								X	X					
D87	Browser playable	X													X	
D88	Electronic book	X									X		X		X	
D89	Browser playable	X		X				X			X				X	
D90	Electronic book			X										X		
D91	Browser playable	X														
D92	Browser playable	X	X				X				X				X	X
D93	Other						X				X					X
D94	Browser playable	X					X			X						
D95	Browser playable	X				X				X					X	
D96	Browser playable	X						X			X					
D97	Other										X					X
D98	Other									X						
D99	Collections			X				X			X				X	
D100	Other										X					X

Contained file formats are represented by X.

File formats corresponding to categories in Table 3.4-8 are shown below.

Table 3.4-9 Classification of file formats

Category	Standard	File formats	No. of materials containing the formats on the left
HTML	Standardized	HTML[33], CSS, VRML, JavaScript, MIDI[34] etc.	45
	Non-standardized	Director[35], Flash, QuickTime[36] etc.	9
Image	Standardized	BMP, GIF, JPEG, WMF, icon, Photo CD, PostScript etc.	33
	Non-standardized	AI[37], PCX, sgi	4
Video	Standardized	AVI, MPEG	7
	Non-standardized	QuickTime, Director[38], ScreenCam etc.	18
Audio	Standardized	WAVE, AU, MIDI etc.	13
	Non-standardized	EUP, FMB, PMB	1
Office	Non-standardized	DOC, XLS, PPT, PPS, MDB[39]	21
Text	Standardized	TEXT, C, CSV, SQL, TeX etc.	69
	Non-standardized	VBP, FRM	4
Electronic book	Standardized	Electronic book[40], EPWING	12
	Non-standardized	Expanded Book[41], DTONIC, J-BISC	5
Other	Standardized	FONT[42], CAB, HELP, ZIP, Rich Text Format, GZ, BinHex, Electronic Navigational Chart[43] etc.	41
	Non-standardized	PDF, AutoCAD, Ichitaro, Hanako, Lotus123 etc.	36

[33] Including linked GIF and JPEG images.
[34] Only files that play in browsers were included here. Files that play in Windows Media Player and similar applications were classified as “audio.”
[35] Only Director videos of the DCR format that use plug-ins to play in browsers were included here.
[36] Only QuickTime movies (MOV files) that use QuickTime plug-ins to play in web browsers were included here. Files for playback in QuickTime players and similar applications were classified as “video.”
[37] Although images recorded in the AI format are internally identical to Encapsulated PostScript images, they were treated as a separate format because they have a different extension and were designed for use with Adobe Illustrator.
[38] Only Director videos of the DIR and DXR formats intended for playback in Shockwave players or Director itself were included here.
[39] MDB files cannot be used if they are read-only. It is necessary to copy them to a hard disk or other drive and remove the read-only setting in the file properties.
[40] These files are commonly referred to as “electronic books”, but this name is actually a Sony trademark. The actual standards they are based on are called EB and EBXA.
[41] Expanded Book is a proprietary standard of The Voyager Company, but it is a very popular format which is used by Aozora Bunko (a free electronic library).
[42] Windows fonts are usually copied to the Windows System Folder where they are accessed via other applications, and therefore font data files are not normally viewed independently. However, they are often included in resource collections and others.
[43] The S-57 format prescribed by the International Hydrographic Organization (IHO).

(3) Choosing files for conversion and playback

The PDF format is widely used as a data format for electronic documents. They are easily accessed and problems are considered unlikely to develop in the medium-to-long term. For these reasons, we thought it appropriate to treat the PDF format as a standardized file format. Therefore, materials chosen for file format conversion were resource collections etc. and other digital materials which were composed of files with non-standardized formats (with the exception of the PDF format). The same types of material were used for playback using the multi-file viewer.

Table 3.4-10 Number of digital materials and materials selected for file format conversion

Category	Number of materials	Standardized or PDF materials	Materials for conversion and playback
Resource collections etc.	11	9	2
Other digital materials	28	15	13

As a result, 15 materials were selected for file format conversion and playback in the multi-file viewer.

(4) Playback test results

The results obtained by running these materials in the file format conversion program and the multi-file viewer are shown below.

Table 3.4-11 Results of file format conversion and playback using a multi-file viewer

Category	Materials for conversion and playback	Materials for which both file format conversion and playback in the multi-file viewer were successful	Materials for which both processes failed
Resource collections etc.	2	1	1
Other digital materials	13	1	12

The materials for which file format conversion was successful were also successfully played on the multi-file viewer. They contained DOC (Microsoft Word document) files and XLS (Microsoft Excel book) files. These files were converted to the PDF format using the file format conversion program and could be played normally after conversion.[44]

[44] The files subjected to conversion did not contain macros. If a data file containing a macro was converted, it is likely that the macro would cease to function.

3.4.3.2. Availability of applications

The results of playback tests using file format conversion and the multi-file viewer showed that there are very few digital materials to which these methods can be applied. It is obvious that the surest way to gain the results the creator intended when playing digital contents is to play them in the operating environment for which they were designed. With that in mind, we identified the applications needed for the playback of the 100 data-only materials chosen as samples and investigated their availability. There were some materials for which the playback application was not specified or was missing, but the necessary applications were identified by examining the contents of these materials.

(1) Testing method

For each of the 100 data-only digital materials, the necessary playback application was identified from information contained in accompanying instruction manuals or the actual content of the material, and the availability of the application was investigated. There were some materials for which the playback application was not specified or was not contained in printed information, but the necessary applications were identified by examining the contents of these materials.

Judgments about the availability of applications were based on information from manufacturers' websites. Applications whose sale and distribution could be confirmed on the manufacturer's website were deemed to be available. If there was no information about an application on the manufacturer's website, or if the application had been discontinued, it was deemed unavailable.[45] For freeware and materials with multipurpose file formats, the availability of playback applications was investigated by searching on an online software distribution site.[46]

(2) Testing results

(i) Availability of specified versions

The availability of specified versions of applications was as follows.

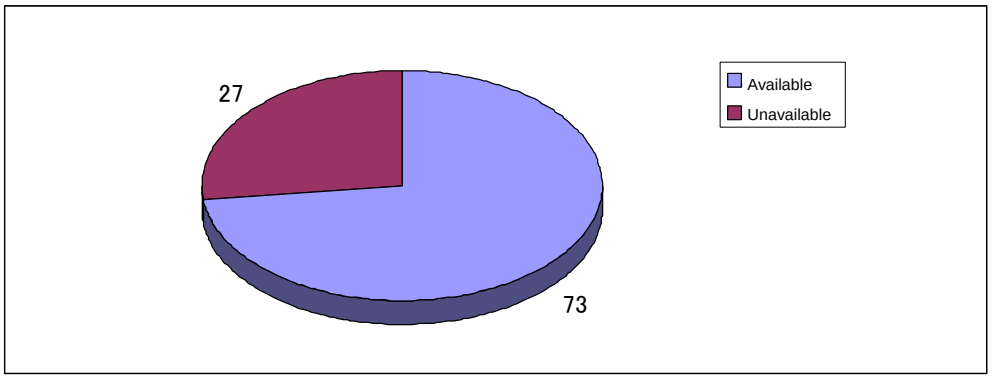

Figure 3.4-10 Availability of specified version

Approximately 3/4 of the applications were available. This includes the materials whose application is included on the CD-ROM.

[45] Unsold stock and the used market were not considered.
[46] We used Vector (http://www.vector.co.jp).

(ii) Availability of the most recent versions
The availability of the most recent versions of specified applications was also investigated because applications are often upward compatible. The results are shown in the figure below. The most recent versions of almost all the applications were found to be available.

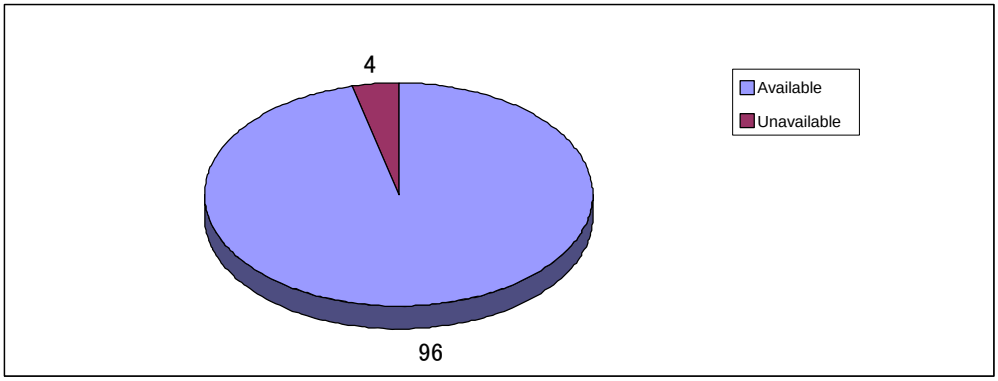

Figure 3.4-11 Availability of the most recent versions

(iii) Materials for which there is little chance of obtaining applications
There were four digital materials for which even the most recent version of the necessary application could not be obtained and which contained file formats which could not be played.

Table 3.4-12 File formats for which applications are difficult or impossible to obtain

File format	Name	Application	Explanation	Sample ID
Bamba Video	Bamba Internet Video	Bamba Player	Video file format developed by IBM for online distribution of video. Playable with the Bamba Player and Plug-ins etc.	D33
PAN	Panoram File	PanoramIX Plug-In	Panorama data file for PanoramIX, a piece of software developed by IBM for the creation and display of 360 degree panorama images. The PanoramIX Plug-In can be used to display panorama images on the Web.	D33
sgi	Silicon Graphics Image	Unknown	Image format developed by the American Silicon Graphics company for workstations manufactured by the same company.	D75
FMB[47]	Unknown	Unknown	Audio format used by FM-TOWNS. For FM sound source.	D71
PMB[48]	Unknown	Unknown	Audio format used by FM-TOWNS. For PCM sound source.	D71
C4	CANDY4 File	CANDY4	Drawing data created by CANDY4, a CAD application produced by the i4 Corporation.	D53
CWJ	Claris Works Document	Claris Works	Document file created by the word processing application Claris Works.	D53

[47] Although this is an FM-TOWNS file format, sample D71 was a digital material for Windows 95/98 that was published in 1999.
[48] See 47.

3.4.4. Review

(1) Emulation

(i) The effectiveness of emulation
Although emulation is probably necessary in the long term, our investigation showed that it is not the solution in all cases at present.

(ii) Quality of emulation produced by the emulator
In this investigation, only about 30 percent of the materials tested were able to be played perfectly; there were obvious problems such as audio distortion encountered on Windows 98 and Windows 95, and audio and visual problems on Windows 3.1 and DOS. These results could be characteristic only of the emulator used in this investigation, but it could be seen as an indication of how difficult it is to perfectly reproduce audio data and other elements of digital material. We can only conclude that at present, perfect emulation is not a realistic possibility.

Furthermore, because development of commercial emulators tends to be centered on the reproduction of operating environments for which there is greater demand, it is possible that newer versions of the same emulator will become poorer at emulating superseded operating systems. This tendency was observed in the results from this investigation in the fact that a smaller proportion of materials for older generations of operating systems were playable than of those for newer generations.

(iii) Requirements for emulators as a tool for maintaining long-term accessibility
In order to guarantee the long-term accessibility of digital materials, an emulator must satisfy the following requirements.

- It must be able to reproduce superseded hardware environments independent of hardware and operating system environments.
- All device drivers for each guest operating system must be obtainable and it must be able to reproduce superseded hardware environments.
- New versions of the emulator must continue support for superseded operating systems that were supported in previous versions.

However, because emulators themselves usually only work with specific operating systems, they cannot continue to be used over the long term as they are. In order to overcome this problem, efforts need to be made to make superseded emulators work in new emulators, and to develop emulators that operate in virtual machines that can be relatively easily implemented in a wide range of hardware.

(iv) The need to maintain superseded operating environments
Although the need for emulation cannot be avoided in the long term, it is much easier to continuously maintain all the elements necessary for playback such as hardware and operating systems; this can be considered a necessity until the arrival of a perfect emulator.

(2) File format conversion and multi-file viewers

(i) Applications for file format conversion
Contrary to our expectations, of the 100 data-only digital materials examined, only 15 were suitable for file format conversion and playback with a multi-file viewer. Furthermore, only two of these materials were successfully converted or played using these methods. The majority of digital materials for which these techniques were not applicable were materials containing multiple interconnected files, such as HTML files, that are needed for playback.

(ii) Conversion at the material level
The conversion of digital materials that contain multiple interconnected files, such as HTML files, needs to be performed at the material level and in such a way that the connections between the multiple files are maintained. In this case, judgments need to be made at the material level, for example the designation of the file which is to become the starting point.

(iii) Methods for maintaining accessibility that do not involve file format conversion
At present, applications necessary for playback of digital materials are easily obtainable. In this investigation we discovered that applications could be obtained for 96 out of 100 samples in the case where the newest versions of these applications were used. Because in most cases applications necessary for playback are sold by manufacturers which possess the specifications for the applicable file formats, we expect that these applications will provide complete and accurate playback. However, although upward compatibility is often supposedly maintained, it is uncertain whether old versions of data can be played correctly on the latest versions of these applications. Furthermore, applications require operating environments in which to operate; thus the continued maintenance of these operating environments is also an issue.

(iv) File format conversion as a method of maintaining long-term accessibility
Obtaining the necessary application is an effective method for playing digital materials containing only data. It allows the playback of the largest range of materials, and also allows complete and accurate playback. From the point of view of long-term preservation, which was the topic of this research, file format conversion programs and multi-file viewers were very restricted in their application, and their effectiveness as a method for maintaining long-term accessibility must be considered low.

3.5. Copy protection and metadata

Here we report on the creation of metadata and the situation relating to the use of copy protection as observed throughout this investigation.

3.5.1. Copy protection

Although examination of copy protection was not a primary aim of the FY2004 study, there were instances when copy protection became a problem.

(i) Copy protection by copy prevention
There were two cases where migration could not be completed[49] because of a type of copy protection that prevents the process of copying itself.

[Problem outline]
- Type name : SafeDisc V1
- Method : Technology for preventing copying by intentionally recording error data, thereby causing errors during the read process.[50]

[Solution]
Because the reading accuracy of drives has improved since the time when this technology was originally introduced, readout is currently said to be possible on almost all CD-R/RW compatible drives in circulation.[51]
Nevertheless, the time required varies greatly depending on the drive used, and there are some cases where an extremely long time is required.

(ii) Copy protection by user verification
There were some applications that could not be used unless the user entered information indicating that they were the correct user.

[Problem outline]
Some items required the input of a serial number during installation, while others required the input of user ID upon boot-up.

[Solution]
Installation and playback were made possible by entering the correct information from the instruction manuals accompanying these digital materials.

Although copy protection was discovered in only four of the materials examined in this survey, it is possible that many more actually had copy protection features. In order to gain an understanding of the problems that may arise during a future large-scale migration performed by the NDL, there needs to be a thorough investigation of copy protection.

The subject of this investigation was digital materials acquired prior to FY1999, but if an investigation were conducted into the effects of copy protection on long-term preservation, it would be necessary, considering the progress and diversification of copy protection technology, to include in the investigation digital materials acquired since FY1999.

---

[49] In both cases the migration process was aborted after an hour because in addition to the process not having been completed, the migration program displayed a required time estimation of close to 10 hours. When the CDs in question were examined with a program for identifying types of copy protection, they were found to contain copy protection features.

[50] Other methods of preventing the process of copying itself are: SecuRom – a method which uses special data areas; RingPROTECH – a method of embedding error sectors that resemble annual rings; and alpha-ROM – a sector duplication method that prevents copying using forward reading.

[51] From CD Kakumei Virtual Ver. 8.0 – Ver. 8.5, special protection investigation reports.
http://www.ark-info-sys.co.jp/protectrepo/index.html

3.5.2. Metadata

Under the NDL's present management system for digital materials, media, instructions and manuals are included in packages, making it easy to obtain information about the operating system, such as CPU performance and memory size necessary for playback. However, if materials are migrated to hard disk for the purpose of long-term preservation, they will become separated from these sources of information. Therefore, when migration is performed, metadata necessary for playback must also be recorded and the link between this and the saved information must be preserved.

(1) Creation of metadata
The following metadata were created when performing migration.

Table 3.5-1 List of metadata

Category	Playback requirement	Details
Operating environment (hardware)	CPU	Conditions such as CPU type and clock speed etc.
	Memory	Minimum memory size and recommended size
	Display	Resolution, number of colors
	CD-ROM drive	Conditions relating to CD read rates etc.
	Audio	Presence or absence, type of sound card
	Other	Hardware requirements other than those listed above
Operating environment (software)	Operating system	Compatible operating systems
	Application	Software necessary for playback
	Other	Software requirements other than those listed above

Metadata took approximately seven minutes per material to create.[52]

(2) Information required during playback tests
The above metadata were thought to include all the information needed for conducting playback tests, but some materials required serial numbers and user IDs, and some that were actually supplementary materials required the primary material.

(i) Serial number
Some digital materials required the input of a serial number during installation.
(ii) User ID
There were digital materials that, although they could be installed, would not run without the input of a user ID during loading.
(iii) Supplementary digital materials
Although installation was completed successfully, an error message was displayed during loading, requesting the primary publication.

(3) Metadata for long-term preservation and use
Metadata were created using instructions, manuals and other information accompanying digital materials, but there was considerable variation in the type of information available and the way it was presented.

Information such as serial numbers and user IDs was also required for use, but these requirements were only revealed during examination. Some digital materials in the investigation were made to be played while continuously referring to the paper media included in the package. It is likely that in some cases information also exists whose

[52] The creation of metadata was conducted simultaneously with the migration process. This figure was obtained by subtracting (i) the time required for changing CDs (= 1,068 min.), (ii) the waiting time due to copy protection (= 120 min.), (iii) the time spend waiting for the completion of migration (= 60 min.) and (iv) time spent returning materials etc. (= 400 min.) from the total time required (= 3,960 min.).

necessity will not be revealed until the digital material is actually used. It would be difficult to treat as metadata and record all information printed on packages and their contents.[53]

Many of the descriptions[54] of hardware operating environments are meaningless today because of dramatic increases in the performance of hardware environments. Because information printed on the instructions for digital materials describes only the operating environment that existed at the time of creation, it would not be appropriate to use this information directly as metadata whose purpose is to enable long-term usability. Descriptions of operating environments often specify essential technological elements (operating system, CPU, applications etc.) as well as minimum requirements (CD drive speed, required memory size etc.), but information such as which technological series elements belong to (for example, the fact that a CPU is an x86 series CPU) also needs to be included.

Given the diversity of descriptions of operating environments and other requirements printed on packages and accompanying materials, in order to ensure the long-term preservation and usability of publications, there is a need to consider how other information necessary for use should be handled, and to establish rules about essential metadata and their description, and about what part of packaged digital publications these descriptions should be based on.

[53] Therefore, it would be insufficient to offer publications for use supplying only the bit strings recorded on the electronic media and metadata. The packages and the paper materials contained also need to be provided together with the digital media.

[54] For example, the requirement that CD-ROM drive speed must be at least 2x.

4. Conclusion

The preceding chapters have given an account of the FY2003 and FY2004 studies relating to packaged digital publications. The results of the usability study conducted in FY2003 revealed that problems exist for the use of a considerable number of digital materials. This result prompted the FY2004 study, which revealed that migration and emulation, two methods proposed for long-term preservation, are not necessarily realistic solutions at the present. However, because these methods are heavily dependent on the testing environment, further research is necessary before deciding that these are not effective methods for the long-term preservation and usability of packaged digital publications.

In addition to packaged digital publications, there are other kinds of digital information, such as information from the World Wide Web found on the Internet, and the same problems and issues exist for the long-term preservation and use of these sources of information as for packaged digital publications.

The guaranteed long-term preservation and future usability of digital information requires not only that the data itself be preserved, but also that playback environments (playback applications, operating systems and playback devices etc.) be maintained over the long term alongside the data. One way would be to comprehensively collect and maintain ideal playback environments for each type of digital information, but given the fact that media and playback equipment have short life spans, this method can only be considered a short-term solution. Migration and emulation would seem to be essential for the long-term preservation of digital information with its many and diverse playback environments and rapidly changing media standards.

Furthermore, in order to maintain playback environments (or emulated operating or hardware environments), it is necessary to provide and maintain information about the attributes of digital information, or in other words, information about the file formats used and metadata relating to the playback environments required.

On the other hand, the problem of long-term preservation of digital information is not one that can be tackled solely by the organizations that carry out the preservation. Previously mentioned issues such as the maintenance and standardization of metadata, and the standardization of file formats, can only be solved through cooperative partnerships with related organizations, including the government, academic institutions, companies involved in information and telecommunications technology, and archiving organizations, as well as related individuals and groups, including manufacturers. In addition to these problems, the revision of systems such as those surrounding copyright must also be addressed.

In FY2005, based on research conducted so far, the NDL began development of the NDL Digital Archive System for the long-term preservation and use of digital information. The aim of this project is to construct a long-term preservation system based on the international standard OAIS, to design schemas for metadata required for use, and to preserve digital information in permanent storage.

In addition to the development of the NDL Digital Archive System, there is an urgent need to develop cooperative partnerships with related organizations in order to make the long-term preservation and usability of digital information a reality.

Appendix 1 Examples of emulators

Many emulators have been developed and distributed. While some are offered for sale, the majority have been created by individuals. Some are open source software. For reference, some examples of emulators are given below. We do not claim that the following list is exhaustive, nor do we recommend any of the emulators listed.

Table A-1 Examples of emulators

Environment emulated	Name of emulator	Emulator operating environment
Nintendo Entertainment System	Aphrodite	DOS
	basicNES 2000	Win, DOS
	BioNES	Win, DOS
	CMNES	Win
	DarcNES	Win, DOS, Unix/Linux
	DNE	Win
	DreamNES	Win, DOS
	DRR-NES	DOS
	EmilNES	Unix/Linux
	EmuSchool	Win
	FakeNES	Win, DOS, Mac, Unix/Linux
	FakeNES WIP	Win, DOS
	Famtasia	Win, DOS
	FCE Ultra	Win, DOS, Mac, Unix/Linux
	FCFAN98	Win
	FE	Win, DOS
	fwNES	Win, DOS, Unix/Linux
	fwNES98	DOS
	G-NES	Win, DOS
	GrayBox	Mac
	GrayBox Classic	Mac
	HyNES	Win, DOS
	iNES	Mac, Unix/Linux
	InfoNES	Win, DOS, Unix/Linux
	Jnes	Win, DOS
	LazyNES	Win
	LissNES	Win, DOS
	Little John	Win
	Little John New Generation	Unix/Linux
	loopynes	Win, DOS
	macFC	Mac
	madNES	DOS
	MarioNES	Win, DOS
	MESS	Win, Mac, DOS, Unix/Linux
	mIRCNES	Win, DOS
	NE	Win, DOS

	NES 9x	DOS
	NES496	Win
	NES4PC	Win
	NesEmu8	Win
	NESMac	Mac
	Nessie	Win, DOS
	NESten	Win, DOS
	nester	Win, DOS
	nesterJ	Win, DOS
	NESticle	Win, DOS
	Nestopia	Win, DOS, Mac
	Nestra	Unix/Linux
	NEStron	Win, DOS
	NextFCE	Unix/Linux
	Nintendulator	Win, DOS
	NinthStar NES	Win, DOS
	NNNesterJ	Win, DOS
	Nofrendo	Win, DOS, Unix/Linux
	Pasofami	Win
	PCNES	DOS
	pNesX	Win
	Pretendo	Win
	Reminesce	Win, DOS
	rew	Win
	RockNES	Win, DOS, Mac, Unix/Linux
	RockNES X	Win
	Shatbox	Win, DOS
	SleepNES	DOS
	SMYNES	Win
	SMYNES	Win, DOS
	Squeem	Win, DOS
	SwNES	Win
	SwNES	Win, DOS
	TextNES	Win, DOS
	Turbo Nes	Win
	TuxNES	Unix/Linux
	UberNES	Win, DOS
	Ultee	Win
	Ultimate Nes	DOS
	unofficial nester	Win
	VirtuaNES	Win, DOS
	Vortendo	Win
	WinNES	Win
	XNes	DOS, Mac, Unix/Linux

	YAME	Win
Super NES	CHAMPI SNES	Win
	ESNES	Win, DOS
	MESS	Win, Mac, DOS, Unix/Linux
	NLKE	Win, DOS
	NLKSNES	Win, DOS
	Peer's SNES 9X Tracer	Win, DOS
	Silhouette	Mac
	SNEeSe	Win, DOS
	SNem	Win, DOS
	SNEmul	Win, DOS
	SNEqr	Win, DOS
	SNES Professional	DOS
	SNES97	Mac
	SNES9x	Win, DOS, Mac
	Snes9xGL	Win, DOS
	SNESGT	Win, DOS
	SNEShout	Win, DOS
	Super Pasofami	Win, DOS
	TheSE	Win, DOS
	Unofficial Snes9x	Win
	USNES	Win, DOS
	Virtual Super Wild Card	Mac
	VSMC	Win, DOS, Mac
	ZSNES	Win, DOS
NINTENDO64	1964	Win, DOS
	Apollo	Win, DOS
	Blade64	Win, DOS
	Corn	Win, DOS
	Daedalus	Win, DOS
	Daedalus-Lkb	Win, DOS
	Dream64	Win, DOS
	Fake64	Unix/Linux
	Mac TrueReality	Mac
	Mupen64	Win, DOS, Mac, Unix/Linux
	N64 VM	Win, DOS
	NEmu	Win, DOS
	NEmu64	Win
	NIN64	Win, DOS
	NINCEST 64	Win, DOS
	NSFE	Win, DOS
	Pagan	Win, DOS
	PC64	Win, DOS
	Project 64	Win, DOS

	Project Unreality	Win, DOS
	Project64	Win
	Project64k	Win, DOS
	SixtyForce	Mac
	Sunset	Win, DOS
	SupraHLE	Win, DOS
	TR64	Win, DOS
	TrueReality	Mac, Unix/Linux
	TRwin	Win, DOS
	Ultra 64	Win, DOS
	UltraHLE	Win, DOS
	UltraHLE 2064	Win, DOS
GameCube	Dolphin	Win, DOS
	Dolwin	Win, DOS
	gcube	Win, DOS, Unix/Linux
	WhineCube	Win, DOS
Game Boy	Boycott	Mac, Unix/Linux
	cingb	Unix/Linux
	DBOY	DOS
	FondleGB	DOS
	GameBoy98	Win
	GameLad	Win
	GBE	DOS, Unix/Linux
	GBEmu	Win
	GBfanPlus	Win
	GBMac	Mac
	GBSIM	DOS
	Gngb	Unix/Linux
	gnuboy	Mac, Unix/Linux
	GooBer	Unix/Linux
	Hash	Win
	HeboWin	Win
	HelloGB	Win
	HyperBoy	Win
	K.G.B.	Win
	KiGB	Mac, Unix/Linux
	MESS	Win, Mac, DOS, Unix/Linux
	NO$GMB	DOS, Win
	PlayBoy	Mac
	prototype-D	Win
	rew.	Win
	SMYGB	Win
	TGB	Win
	TGB Dual	Win

	TKGBC	Win
	VenGaboyTK	Unix/Linux
	VGB	Win, DOS, Unix/Linux
	VGBMac	Mac
	Victoly Boy	Mac
	Virtual Gameboy	Mac
	VirtualGameBoyColor	Win
	VisualBoy	Win
	WINBOYCOTT	Win
	XGB	Win
	Yage	Win
	YAME	Win
Game Boy Advance	BoyCott Advance	Win, Mac, Unix/Linux
	DreamGBA	Win
	iGBA	Win
	PlayBoy Advance	Mac
	VGBA	Unix/Linux
	VirtualGameBoyAdvance	Win
	VisualBoy Advance	Win, Mac, Unix/Linux
Sega Master System	BRSMS	Win, DOS
	Calypso	Win, DOS
	CHASMS	Win, DOS
	Dega	Win, DOS, Unix/Linux
	eSMS	Win, DOS
	FoolsSMS	Win, DOS
	FreezeSMS	Win, DOS
	GGBoy	Win, DOS
	Massage	Win, DOS
	MasterGear	Win, DOS, Mac, Unix/Linux
	Masterlator	Win, DOS
	MEKA	Win, DOS
	MEKANIX	Unix/Linux
	MEKAW	Win, DOS
	MesaDX	Win, DOS
	MESS	Win, Mac, DOS, Unix/Linux
	MGX	Win, DOS
	NeoSMS	Win, DOS
	Past-O-Rama	Win, DOS
	Saya RX-SMS	Win, DOS
	SegaG3	Win, DOS
	SMaSher	Win, DOS
	SMS Plus	Win, DOS, Mac
	Wakalabis	Win, DOS
Sega Mega Drive	DGen	Win

	Genecyst	DOS
	GenEm95	Win
	Generator32	Win
	GENS	Win
	KGen98	DOS
	Megasis	Win
	MESS	Win, Mac, DOS, Unix/Linux
	SGE	DOS
	VEGAS	Win
	VGen	DOS
	WinAGES	Win
	Xega	Win
Sega Saturn	A-SATURN	Win, DOS
	Cassini	Win, DOS
	GiriGiri Debugger	Win, DOS
	Hyperion	Win, DOS
	PC-Saturn	Win, DOS
	Project Titan	Win, DOS
	Satan	Win, DOS
	Sat'On'Em	Win, DOS
	Satourne	Win, DOS
	Saturnin	Win, DOS
	Semu	Unix/Linux
	SSE	Win, DOS
	SSEmu	Win, DOS
	SSF	Win, DOS
	UltraSat	Win, DOS
	Yabause	Win, DOS, Mac, Unix/Linux
Dreamcast	Chankast	Win, DOS
	DreamEMU	Win, DOS
	Dreamer	Win, DOS
	Nightmare	Win
	Swirly	Unix/Linux
Game Gear	BRSMS	Win, DOS
	Calypso	Win, DOS
	CHASMS	Win, DOS
	Dega	Win, DOS, Unix/Linux
	eSMS	Win, DOS
	FoolsSMS	Win, DOS
	FreezeSMS	Win, DOS
	GGBoy	Win, DOS
	Massage	Win, DOS
	MasterGear	Win, DOS, Mac, Unix/Linux
	Masterlator	Win, DOS

	MEKA	Win, DOS
	MEKANIX	Unix/Linux
	MEKAW	Win, DOS
	MesaDX	Win, DOS
	MESS	Win, Mac, DOS, Unix/Linux
	MGX	Win, DOS
	NeoSMS	Win, DOS
	Past-O-Rama	Win, DOS
	Saya RX-SMS	Win, DOS
	SegaEMU	Win
	SegaG3	Win, DOS
	SMaSher	Win, DOS
	SMS Plus	Win, DOS, Mac
	Wakalabis	Win, DOS
WonderSwan	Cygne	Win, DOS
	Oswan	Mac
	WSCamp	Win
Neo-Geo	Danji	DOS
	Gekko	Win, DOS
	Gngeo	Unix/Linux
	Kawaks	Win, DOS
	KBMAME	Win, DOS
	NeoCD	Win, DOS
	NeoCD/SDL	Win, DOS, Unix/Linux
	NeoDanji	Win, DOS
	NEOGem	Win, DOS
	NeoGeo CD Emulator	Win, DOS
	NeoRage	DOS
	NeoRageX	Win
	Neo-RAGEx	Win, DOS
	NMAME	DOS
Neo-Geo Pocket	Koyote	Win
	LAME?	Win
	NeoPocott	Win, Mac, Unix/Linux
	NEOPOP	Win
	NeoPop-SDL	Unix/Linux
	NGPOCKET	Win
3DO	FreeDO	Win, DOS
PlayStation	4EverPSX	Win, DOS
	AdriPSX ILE	Win, DOS
	Bleem!	Win, DOS
	ConnectixVirtualGameStation	Win
	ePSXe	Win, DOS, Unix/Linux
	FlareStorm	Mac

	FPSE	Win, DOS
	Jackal	Win, DOS
	PCSX	Win, DOS, Unix/Linux
	pkemu	Win
	PSEmu	Win
	PSEmu/PSEmu Pro	Win, DOS
	PSInex	Win, DOS
	PSMac	Mac
	PSXeven	Win, DOS
	Psyke	Win, DOS
	Sope	Unix/Linux
	The Pi	Mac
	Virtual Game Station	Win, DOS, Mac
PlayStation 2	EMUtion Engine	Win, DOS
	neutrinoSX2	Win, DOS
	PCSX 2	Win, DOS, Unix/Linux
	PS2Emu	Win, DOS
XBOX	Cxbx	Win, DOS
	Xeon	Win, DOS
Arcade	Callus	DOS
	Callus95	Win
	CPS2Burn-JK	Win
	CPS2mame	DOS
	CPS2MAME32j	Win
	EmeraldMAME32jk	Win
	FinalBurn	Win
	FinalBurnHack	Win
	HauMAME	DOS
	HauMAME32	Win
	impactemu	Win
	Kaillera	Win
	Kawaks	Win
	MAME	DOS
	MAME32	Win
	MAME32j	Win
	MAME32jn	Win
	MAME32jp	Win
	NEBULA	Win
	RAINE	DOS
	RAINE32	Win
	StretchMame32	Win
PC Engine	DeePCE	Win
	HU6280-DirectX Beta Test...	Win
	Hu-GO	DOS

	MagicEngine	Win, DOS
	MESS	Win, Mac, DOS, Unix/Linux
	npce	Win
	pce	Win
	rew.	Win
	VPCE	Win
	xpce	Win
	YAME	Win
PC-6001	iP6 Plus	Win
	P6001V	Win
PC-8801	M88	Win, DOS
	MESS	Win, Mac, DOS, Unix/Linux
	PC88Win	Win, DOS
	Quasi88	Unix/Linux
	X88000	Win, DOS, Unix/Linux
PC-9801	ANEX86	Win, DOS
	Neko Project II	Win, DOS, Mac
	PC98E	Win, DOS, Unix/Linux
	T98	Win, DOS
	T98-Next	Win, DOS
	T98vmm	Win, DOS
	Virtual 98	Win, DOS
MSX	blueMSX	Win, DOS
	BRMSX	Win, DOS
	CJS MSX Emulator	Win, DOS
	fMSX	Win, DOS, Mac, Unix/Linux
	fMSX FAN	Win, DOS, Unix/Linux
	fMSX for Win95	Win
	fMSX SO	Win, DOS
	fMSX/MacOS	Mac
	fMSX98/AT (AT)	DOS
	fMSXMS-DOS	DOS
	fMSX-SDL	Win, DOS
	freeM	Mac
	FreeMSX	Mac
	MESS	Win, Mac, DOS, Unix/Linux
	MSKISS	Win, DOS
	MSX Emulator	Win, DOS
	MSXPLAYer	Win, DOS
	NLMSX	Win, DOS
	NO$MSX	Win, DOS
	openMSX	Win, DOS
	P ParaMSX	Win
	ParaMSX	Win, DOS

	Power MSX	Win, DOS
	RedMSX	Win, DOS
	RuMSX	Win, DOS
	SVGAMSX	Unix/Linux
	Virtual MSX	Win, DOS
	WOOM!	Win, DOS
	Zodiac	Win, DOS, Mac
X68000	AX68K	Win, DOS
	EX68	Win, DOS
	Keropi (WinX68K)	Win, DOS
	Libvm68k	Unix/Linux
	Virtual X68000	Unix/Linux
	Win68K	Win
	WinX68k High-Speed	Win, DOS
	X68000	Mac
	X68000 Emulator PPC	Mac
	XM6	Win, DOS
	Keropi	Win
Sharp MZ	EmuZ-2000	Win
	EmuZ-2500	Win
	MESS	Win, Mac, DOS, Unix/Linux
	MZ700Win	Win
	MZ800emu	Win
	Z80x	Win
Sharp X1	X1 EMU	DOS
	X1 Millenium	Win
	X1WIN	Win
FM 7	XM7	Win
FM TOWNS	UNZ	Win
Apple I	Cocoa Pom I	Mac
	MESS	Win, Mac, DOS, Unix/Linux
	Sim6502	Mac
Apple II	][In A Mac	Mac
	Apple //e Emulator	Win, DOS
	Apple II Oasis	Win, DOS
	ApplePC	Win, DOS
	Appler	Win, DOS
	AppleWin	Win, DOS
	Bernie][The Rescue	Mac
	Catakig	Mac
	Dapple	Win, DOS
	Dapple][	Win, DOS
	Gus	Mac
	iGS	Mac

	iGS-Mac	Mac
	IIe	Mac
	KEGS	Unix/Linux
	Kegs32	Win, DOS
	KEGS-OSX	Mac
	MacE	Mac
	MESS	Win, Mac, DOS, Unix/Linux
	OpenPhoenix	Mac
	OSXII	Mac
	Stop the Madness	Mac
	Virtual][	Mac
	XGS DOS	Win, DOS
	XGS/32	Win, DOS
Apple III	Sara	Mac
Macintosh	Basilisk II	Win, DOS, Unix/Linux
	Basilisk II/JIT	Unix/Linux
	EmMac	Win, DOS
	Executor	Win, DOS, Unix/Linux
	Fusion	DOS
	MESS	Win, Mac, DOS, Unix/Linux
	MOL	Unix/Linux
	PearPC	Win, DOS
	SheepShaver	Unix/Linux
	SoftMac XP	Win, DOS
	vMac	Win, DOS
Atari ST	Echo	Win, DOS
	FAST	Win, DOS
	Gemulator 2000	Win, DOS
	Gemulator Classic	Win, DOS
	Hatari	Mac, Unix/Linux
	NoSTalgia	Mac
	PaCifiST	Win, DOS
	PowerST	Mac
	SainT	Win, DOS
	Steem Engine	Win, DOS
	STew	Win, DOS
	STonX	Win, DOS, Unix/Linux
	TOSBOX	Win, DOS
	WinSTon	Win, DOS
	XSteem	Unix/Linux
Commodore Amiga	Akiko	Win, DOS
	Experimental UAE	Mac
	Fellow	Win, DOS
	MacUAE	Mac

	UAE	Win, DOS, Unix/Linux
	xFellow	Unix/Linux
Commodore 64	C64S	Win, DOS
	CCS64	Win, DOS
	Come Back 64	Win, DOS, Unix/Linux
	ec64	Unix/Linux
	Emu64	Win, DOS
	Free64	Win, DOS
	Frodo	Win, DOS, Mac, Unix/Linux
	GameBase 64	Mac
	Hoxs64	Win, DOS
	Mac64	Mac
	PC64	Win, DOS
	Pfau Zeh	Unix/Linux
	Power64	Mac
	SIDekick	Mac
	SIDPlay	Mac
	VB64	Win, DOS
	VICE	Win, DOS, Mac, Unix/Linux
	Win64	Win, DOS
TRS-80	TRS-80	Mac
	TRS-80 Colour Computer Emulator	Win, DOS
	CoCo 3	Win, DOS
	Virtual MC-10	Win, DOS
	xtrs	Unix/Linux
Hitachi Basic Master Level 3	Basic Master Level 3 Emulator	Win
PC/AT compatibles	Microsoft Virtual PC 2004	Win
	VMWare	Win, Linux

Appendix 2 Examples of conversion programs

Many conversion programs have also been developed and distributed. Unlike emulators, many are offered for sale. For reference, some examples are given below. We do not claim that the following list is exhaustive, nor do we recommend any of the programs listed.

Table A-2 Examples of conversion programs

Name	Marketer	Operating environment	Compatible file formats
GraphicConverter Carbon	Lemke Software GmbH	Mac	Compatible with over 175 file formats. PICT, Startup-Screen, PDF, MacPaint, TIFF, RIFF, PICS, 8BIM, 8BPS/PSD, JPEG/JFIF, GIF, PCX/SCR, GEM-IMG/-XIMG, BMP, ICO/ICN, PIC (16 bit), FLI/FLC, TGA, MSP, PIC (PC Paint), SCX (ColoRIX), SHP, WPG, PBM/PGM/PPM, CGM (only binary), SUN, RLE, XBM, PM, IFF/LBM, PAC, Degas, TINY, NeoChrome, PIC (ATARI), SPU/SPC, GEM-Metafile, Animated NeoChrome, Imagic, ImageLab/Print Technic, HP-GL/2, FITS, SGI, DL, XWD, WMF, Scitex-CT, DCX, KONTRON, Lotus-PIC, Dr. Halo, GRP, VFF, Apple IIgs, AMBER, TRS-80, VB HB600, ppat, QDV, CLP, IPLab, SOFTIMAGE, GATAN, CVG, MSX,PNG, ART, RAW, PSION, SIXEL, PCD, ST-X, ALIAS pix, MAG, VITRONIC, CAM, PORST, NIF, TIM, AFP, BLD, GFX, FAX, SFW, PSION 5, BioRad, JBI, QNT, DICOM, KDC, FAXstf, CALS, Sketch, qtif, ElectricImage, X-Face, DJ1000, NASA Raster Metafile, Acorn Sprite, HSI-BUF, FlashPix, etc.
VINC	Recosoft Corporation	Win, Mac	Compatible with over 60 file formats. Text, Recowrite, Egword, Clarisworks, AppleWorks, MacWrite II, MSWord, RTF, HTML, GIF, JPEG, TIFF, PNG, TGA, Photoshop, QuickTime Image, MacPaint, Silicon Graphics, etc.

Jizaigan 8	Antenna House Inc.	Win	BMP, FlashPix, JPEG, MacPaint, Kodak PhotoCD, Mac QuickDraw, PNG,Adobe PhotoShop5.0, PaintShopPro5.x, TIFF, CCITT Fax3 (G3:B/W), WordPerfect Graphics V1.0, 2.0, PICT, Excel, MIDI, Video for Windows, QuickTime Movie, MPEG 1 Video, MPEG 1 Audio, MPEG 4Video, Windows Media Audio, Windows Media Video, Access, Paradox, dBASE, Kiri, BinHex, LHA, PKZip, TAR, Z-compression, GZ-compression, UUencode, Ichitaro, MS Word for Windows, MS Word for Macintosh, RTF, Aurora Ace, Shinmatsu, Matsu, Lotus AmiPro, LANWORD for Windows (95/NT), Bungo, HTML, EGWORD, MacWrite II, Claris works4, AppleWorks, OASYS, OASYS5000, BungoJX, RUPO, BungoHYPER7, TOSWORD, BungoMini5, Shoin, BungoMini7, U1PRO, LANWORD (PTOS), Panaword, Canoword, U1B, Report, Sanword, Wordpal, Wordpro Ace, Casioword, Lotus 1-2-3, MS Excel for Windows, MS Excel for Macintosh, OASYS 1-2-3, FileMaker PRO, MS PowerPoint, MS PowerPoint98 Macintosh Edition, etc.
Rich Text Converter	Antenna House Inc.	Win	File formats used by word processors (OASYS, TOSWORD, Shoin, LANWORD, Bungo, U1PRO, Panaword, U1B, Sanword, Wordpro Ace, Report, Wordpal, Casio Word), file formats used by word-processing software (jxw, atr, ctl, jsw, jaw, jtw, jbw, juw, jfw, jvw, jtd, jtt, bun, doc, sod, aad, jtx, wps, rtf, sam, fmt, oas, dwd, cwj, txt, html, pdf, ClarisWorks) Apple Works, MacWrite II, EGWORD) and other file formats (xls, wj2, wj3, wk4, 123, fmj, fp5, jad).

* Win: Operates in Windows machines. Mac: Operates in Macintosh machines.