

デジタル情報の管理

～デジタル情報を継続的に保存・利用するために～

大城 博光†

はじめに ～迫られるデジタル情報への対応～

1 デジタル情報の問題点を理解する

- 1-1 デジタル情報に必要な利用環境
- 1-2 利用環境の移行期
- 1-3 記録メディアの耐久性

2 デジタル情報を管理する

- 2-1 記録メディア
- 2-2 データファイル形式
- 2-3 管理情報の記録
- 2-4 デジタル情報の保全

おわりに ～電子自治体制度の中での公文書館の役割～

はじめに ～迫られるデジタル情報への対応～

「デジタル情報は恒久的保存に適さない」と言われる。たしかに、数百年前の記録が特別な保存措置を施さなくても現在まで残っているのは、「紙と墨」等の保存性のよい記録メディアが使われていたからであろう。そのように言われながらもデジタル情報を全く保有していない資料保存機関は少ないのではないだろうか。それは、デジタル情報の利便性や流通性の良さが、情報を欲する側の要求に適合しているからである。更に行政組織においては、電子自治体政策の中で、文書事務の効率化のための電子文書システムも波及してきている。これからは、好むと好まざるとに関わらず、デジタル情報が複製資料として、またはオリジナル資料として増えてくる事が確実である。そのような状況において「デジタル情報は恒久的保存に適さない」ことを理由に、デジタル情報への対応を棚上げにしたままだと、まさに数年後にそのデジタル情報が利用できない状態になっており、記録管理者がその責任を果たしていないことになる。

本稿は、恒久的保存の面で課題の多いデジタル情報に対し、どのような管理を施せば継続的な利用が可能になるのか検討する。

1 デジタル情報の問題点を理解する

何故、「デジタル情報は恒久的保存に適さない」と言われるのか。デジタル情報の問題点を整理し、継続的な保存・利用のための方向性を考える。

1-1 デジタル情報に必要な利用環境

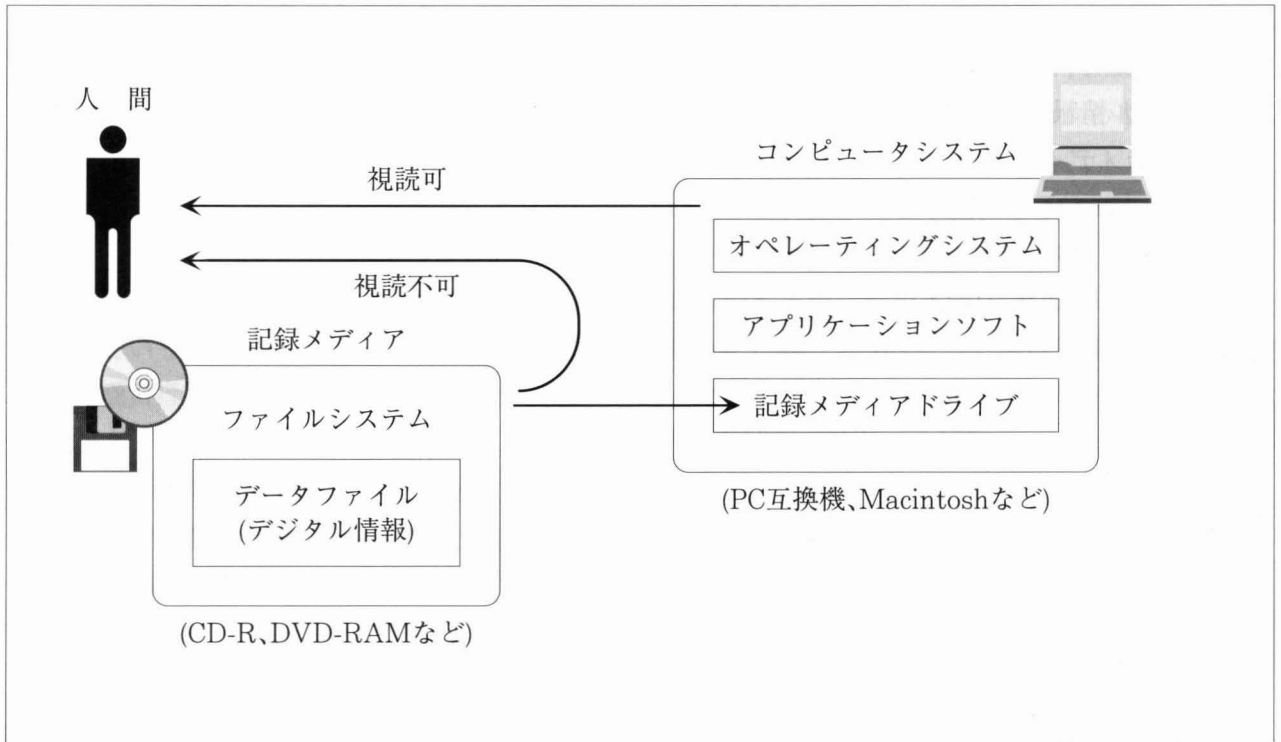
デジタル情報は、「0」と「1」の数値信号の配列が、さまざまな形式に準じてデータファイルとして構

†財団法人沖縄県文化振興会公文書管理部公文書専門員

成される。よって、デジタル情報を保存するには、その数値信号を記録するメディアが必要になるとともに、その内容を表現するのに、その記録メディアの再生装置（記録メディアドライブ）とデータファイルを解析するためのアプリケーションソフト、及びそれらが稼動するコンピュータシステムが必要になる。

すなわち、人間は、記録メディアを直接読み取ることができないので、その記録メディア及びデータファイルに対応した利用環境を必要とする。

図1 デジタル情報の利用環境



1-2 利用環境の移行期

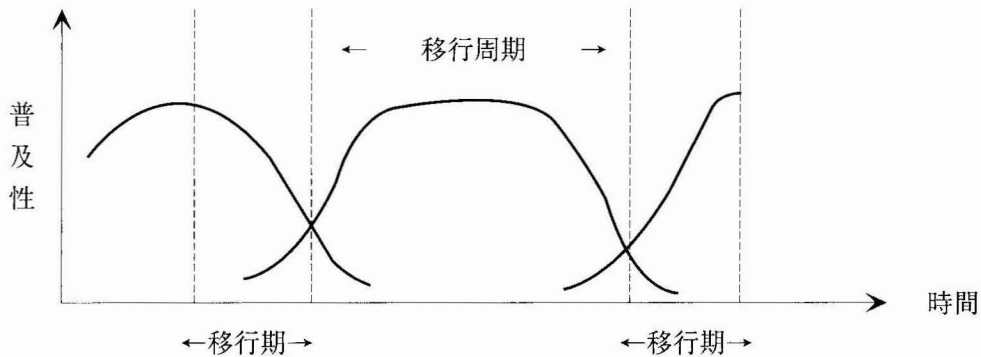
記録メディアとその利用環境は、一対を成して利用が可能となる。記録メディアだけを保存していても、対応した利用環境がなければその情報を利用することはできない。

しかし、デジタル情報の利用環境には、これまで主流として用いられたものが、技術の進歩等により徐々に、または急速に一般社会から姿を消していく"移行期"があり、この移行期を過ぎた後に、旧来の利用環境を継続的に維持管理していくことは、困難である。このことが冒頭に述べた「デジタル情報は恒久的保存に適さない」と言われる大きな要因である。

では、どのように対処すればよいのであろうか。デジタル情報を一定期間（数年程度）だけ利用するには、利用環境を維持管理していくことが可能かもしれない。しかし継続的な保存・利用を実現するには、その時代とともに変化する利用環境を維持するのではなく、移行期を適切に捉え、記録メディアと記録されているデータファイルの形式を、変化した新しい利用環境に対応できるものに変換する措置を施すことによって可能になる。

また、新しい利用環境に対応できるものは、移行期と移行期の間の"移行周期"が最も長く安定した記録メディアやデータファイル形式を選定することが、適正な管理を継続していく上で重要である。

図2 記録メディアとデータファイルの普及性



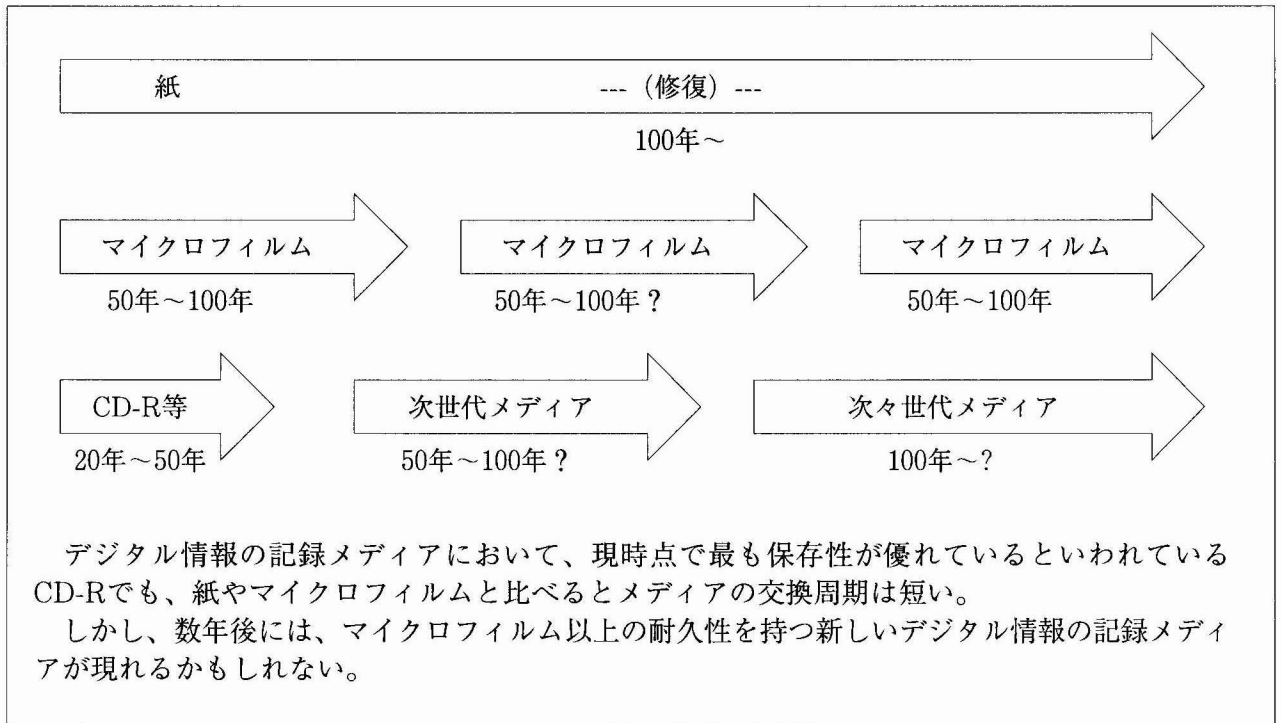
移行期には、新しい記録メディアやデータファイル形式が旧来のものによって普及する。移行期を過ぎると、旧来の記録メディアやデータ形式を利用することが困難な状況に陥る。

1-3 記録メディアの耐久性

記録メディアは物理的なものである限り劣化が免れない。現在のデジタル情報の記録メディアは、紙やマイクロフィルム等のアナログ情報の記録メディアと比べ、耐久性が劣る上、長期保存の実績もない。また、アナログ情報の記録メディアは徐々に、そして部分的に劣化するが、デジタル情報の記録メディアが劣化すると全ての情報が読めなくなる。この耐久性が劣るということは、管理の面からいえば記録メディアの交換周期が短いということである。しかし、デジタル情報は記録したデータファイルを劣化させることなくメディアを交換・変換¹することが可能である。このことは、デジタル情報に最適な記録メディアの選択権を与えていることになる。現時点での最適な記録メディアがその最終型ではない。デジタル情報は、データファイルをその時代の最も保存性の良い記録メディアへ複写を繰り返し、継続的に保存していくことができるのである。

¹ 同じ種類の別メディアへ移し変えることを"メディア交換"、別の種類のメディアへ移し変えることを"メディア変換"とする。

図3 記録メディアの交換周期



2 デジタル情報を管理する

前章では、デジタル情報の恒久的保存に関する方向性を述べた。本章では、具体的にどのような管理を行えばよいか、デジタル情報の管理方法を表1にまとめた。以下は、表1で記した事項の根拠等を補足する。

2-1 記録メディア（管理方法1.1）

現在、デジタル情報の記録メディアには、FD,HD,MO,ZIP,CD-ROM,CD-R,DVD-RAM...と様々なメディアが存在するが、今後も進歩を続け、次々と新しい記録メディアが誕生する。これらの記録メディアの中から継続的保存に適したものを選定するには、メディアの寿命が大きな要因になる。メディアの寿命とは、物理的耐久性だけでなく、移行周期の長さである。

[記録メディアの選定要因]

- △耐久性：劣化・損傷に強い
- △普及性：主流の記録メディアとして利用されている
- △汎用性：利用環境の制約が少ない
- △保護性：記録された情報を更新できない
- △記憶容量：多くの情報が記録できる
- △携帯性：記録メディアの出納が容易²
- △低価格：記憶容量に対する単価が安い

² 全てのデジタル情報を常時、直接的にアクセス可能な状態にできない場合、紙やマイクロフィルムと同様に保管庫に収めて出納することになる。

これらの事を総合的に判断し、現時点で最もデジタル情報の保存用記録メディアに適している記録メディアは、"CD-R"だと言える。

CD-Rは有機色素³へ記録することにより、磁気や電気による影響を受けることがなく、一度記録された情報を変更できない事が、保護性に優れているといえる。また、現在ほとんどのコンピュータに標準装備されているとともに、利用環境に依存しないファイルシステムを使用することができる⁴。

2-2 データファイル形式（管理方法1.2）

2-2-1 静止画像系のファイル形式

文書、図書、写真等のアナログ情報をデジタル化する際に保存するデータファイル形式には、BMP, TIFF, GIF, JPEG, PICT...などがあるが、記録メディアと同じように、選定する際の重要な要因は、普及性と汎用性である。

(1) TIFFファイル

ハードウェアやソフトウェアに依存せず、幅広い種類のアプリケーションがサポートしている。また、柔軟性が高く、かつ多くの情報を保有することができるデータ形式なので、データ容量は大きくなるが、生成時の情報損失が最も少なく、オリジナル情報を正確に記録することができるファイル形式である。

(2) JPEGファイル

JPEGファイルもTIFFファイル同様、ハードウェアやソフトウェアに依存しない。また、圧縮率に応じて画像が劣化するが、その分ファイル容量が小さく、最も普及性と汎用性があるインターネットブラウザ（Internet ExplorerやNetscape Navigator）で参照できる。また、インターネット上で利用できるということは、継続的に利用されるファイル形式ということである。

2-2-2 電子文書（テキスト系）のデータファイル形式

Microsoft-Wordやジャストシステム一太郎などのワープロソフトで作成された電子文書は、専用のアプリケーションソフトを必要とするため、そのままのデータファイル形式で保存していたのでは、データファイルの移行周期が短くなる上、利用できる環境も極端に制限される。よって、電子文書の作成時または受入時に、汎用的なファイル形式へ変換して保存しなければならない。

³ CD-Rは、記録層に使用される有機色素の種類によっても、耐久性が異なってくる。
・シアニン系：CD-Rの初期の頃から採用されており、それだけに信頼性がある。
・フタロシアニン系：外的要因に強く、通常の湿度や光度ではほとんど変化しない。
・アゾ系：最近使われるようになった色素で、耐光性や耐熱性にすぐれている。

⁴ CD-Rのサイズや特性などは、統一規格があるが、情報を記録する形式は、ISO9660, Joliet, ROMEO, HFS, UDFなど複数存在する。これらの記録形式の中でISO9660は、オペレーションシステム等に制約される事が無く、CD-Rの誕生時からの実績がある。

前述した静止画像系のファイル形式に変換することも可能だが、テキスト系の電子文書は文字データで構成される為、画像ファイルよりファイル容量が小さく、ネットワーク配信時の負荷が軽減される上、記録されている文字列を検索することも可能である。よって、その特性を維持できるファイル形式を選定する。

(1) PDFファイル

PDFファイルは、文字コードを扱いながらも⁵異なる利用環境においても、オリジナル文書とほぼ同様のレイアウトを表現することが可能である。また、参照用アプリケーションである「Acrobat Reader」も無償で配付されており、インターネット上の異種環境間でやりとりする際のファイル形式として最も主流となっている。しかし、PDFファイルはAdobe Systems社の独自規格であることからその継続性が不安に思われるが、規格自体が公開されており（利用は無償）、ここまで世界中で認められている規格が、今後、互換性なく放棄されるということは考えにくい。

(2) XL系ファイル

HTML,SGML,XMLは、いずれも電子文書として汎用性・普及性のあるファイル形式であり、環境や時代を越えて利用されうる電子文書の規格を求めて、国際的に進められている規格である。ただし、現時点ではオリジナルのレイアウトを保持する面においてPDFファイルには劣っている。

2-3 管理情報の記録（管理方法2.1）

所蔵するデジタル情報の移行期は、記録メディアやデータファイル毎に異なって訪れる。よって、移行期に適正な移行作業を行うために、記録メディア管理票（表2）⁶とデータファイル管理票(表3)を作成するとともに、記録メディア目録データベース（表4）に記録メディアに関する情報を記録する。また、デジタル情報の移行期は、現在の担当者の時に訪れるとは限らない。これら管理票は、どのようなデジタル情報を所有していて、どのような処置をいつ頃行えばよいのか、次の担当者が判断するためにも必要となる。

また、効率的な管理を継続していくには、デジタル情報を生成・受入時に、様々な記録メディアやデータファイル形式などを混在させない事が重要である。

⁵ PDFファイルについて

文字データを扱うファイル形式には問題が多い。コンピュータでの文字の取り扱いは、それぞれの文字にコードを割り当てて管理する。和文文字のコード体系にはJIS、Shift-JIS、EUCが存在し、それぞれの文字コード割り当てがこのコード体系によって異なる。よって、作成した環境とそれを表示する環境が違くと全く違う文字が表示されることになる。

⁶ 記録メディア管理票は、同一メディアでも製品毎に記述しなければならない。それは、デジタル情報の記録メディアは、安定期に達してないため、どんどん新製品が出現する。そこである日、「〇〇系の有機色素を用いているCD-Rは、書き込み後10年で劣化することが判明」などと発表されることもあるかもしれない。その際に、所蔵しているCD-Rの中から効率的に該当する製品を抽出できるためである。

2-4 デジタル情報の保全（管理方法2.2及び2.3）

デジタル情報の保全業務として、記録メディアの劣化・損壊時のメディア交換、移行期のメディア交換、及び移行期のデータファイル形式変換がある。

これらの保全業務にかかる経費、すなわちデジタル情報の維持経費は、高価なものだと思われるが、これは、物品的な面であれば、CD-Rは1枚100円以下の価格である。記録メディアの交換・変換にかかる処理は、コンピュータが行う為、人的労力はほとんどいない。同様に、データファイル形式の変換においても同じ事がいえるが⁷、データファイル形式の変換には、情報の劣化が起こる。よって、繰り返すが、移行周期の最も長いデータファイル形式を選定することが重要である。

おわりに ～電子自治体制度の中での公文書館の役割～

電子自治体制度の中では、電子文書が作成から保存まで統合管理されるとともに、情報公開制度に対応し、ネットワークを介した検索・公開サービスが確立される。そこでは、公文書館の主業務として存在した「引渡」、「評価選別」、「目録作成」という業務が効率化され、軽減あるいは必要がなくなる。しかし、「保存」という業務が効率化されたわけではない。そのような制度の中で公文書館はこれまで以上に情報の継続的保存にかかる比重は大きくなるであろう。

従来の紙文書は、保存スペースの制約から、文書の「保存期間」を設けて「廃棄の決定」を行い、「公文書館への引渡」が必要であった。紙文書が電子文書に変わることで、保存スペースの制約から解放され、「物」としての移動の必要性はなくなり、「現用文書」と「非現用文書」の境界も薄れていく。ネットワーク上で共有された電子文書の管理では、保管場所＝管理部署という関係は変化し、各共有段階（「課共有文書：各課管理」→「組織共有文書：文書課管理」→「県民共有文書：公文書館管理」）により管理体系が整備されていくであろう。その中で公文書館は、最終の共有段階の管理を担う。それは、継続的保存という重要な、そして困難な役割である。

また、電子文書を扱う公文書館に求められる役割の1つに、情報の「信頼性の確保」がある。電子文書は、それだけで「物」として存在しないため、「オリジナル」と「複製」の区別が物理的にはできない。電子文書は、完全な同一体が簡単に作れるため、「複製」というよりも同一体の「増殖」であり、紙文書特有のオリジナル性であった「この世に1点しかない…」という物としての希少価値は意味を持たない。電子文書に要求されるのは、その情報の信頼性である。情報は広く普及されてこそ生きるが、1次、2次、3次資料と複製が増殖すると、その情報の信頼性に不安がでてくる。特にインターネットの世界ではそうである。そのような状況において電子文書の信頼性は、電子的(技術的)保護だけでなく、情報を提供する側の情報管理体制が大きく影響する。

(おおしろ・ひろみつ)

⁷ 静止画像データファイル形式は、多種に存在するが、そのファイル形式の変換を一括で処理できるソフトウェアが多く存在する。

表1 デジタル情報の管理方法

デジタル情報の管理方法

デジタル情報を継続的かつ合理的に保存していくために必要な事項を記す。

1 デジタル情報の保存

デジタル情報は、次の記録メディアやデータファイル形式で保存する。

1.1 保存用記録メディア

- (1) 保存用のデジタル情報を記録メディアに保存する場合は、「CD-R」を使用する。
- (2) CD-Rへ記録するファイル形式を「ISO9960」とする。

1.2 保存用データファイル形式

(1) 静止画像ファイル

文書、図書、地図、写真などのアナログ情報をデジタル化する際のデータファイル形式は、「TIFFファイル」または「JPEGファイル」とする。

(2) 電子文書ファイル

電子文書のファイル形式を「PDFファイル」とする。

(3) その他のデータファイル

音声、動画、データベースなどについては、別途定める。

1.3 閲覧用デジタル情報

別にオリジナル情報を保有し、その利便性の向上を図った複製物としてのデジタル情報は、1.1項及び1.2項で定めたものとは別に、利用形態に即した記録メディア及びデータファイル形式を使用できるものとする。

2 デジタル情報の管理

保存したデジタル情報を次の要領で管理する。

2.1 管理情報の記録

- (1) 保存した記録メディアの製品別に「記録メディア管理票」を作成する。
ただし、既に作成済みの同製品を使用する場合は作成は必要ない。
- (2) 保存したデータファイルの形式別に「データファイル形式管理票」を作成する。
ただし、既に作成済みのデータファイル形式を使用する場合は作成は必要ない。
- (3) 所蔵資料目録データベースに、保存した記録メディアのメディア種別、メディア特性、記録日付等を入力する。

2.2 デジタル情報の保全

2.2.1 記録メディアの損壊防止

- (1) デジタル情報は、利用時の損壊防止のため、保存用と閲覧用の二つの記録メディアに保存する。
- (2) 閲覧用記録メディアが利用できなくなった場合には、保存用記録メディアから新しい記録メディアへ情報を複製し、常に、保存用と閲覧用の二つの記録メディアを保持する。

2.2.2記録メディアの劣化防止

- (1) 6ヶ月周期で、記録メディアの製品別及び保存環境別に部分抽出して再生を試み、劣化を検査する。また、その際に利用不能な記録メディアを発覚した場合には、同一条件の記録メディアのメディア交換を行う。また、「記録メディア管理票」のメディア交換周期を見直す。
- (2) 1年周期で、所蔵資料目録データベースから、記録メディア管理表の記録メディアとメディア交換周期をもとにメディア交換周期に達した記録メディアを抽出し、対象となる記録メディアに対し、メディア交換を行う。

2.2.3 デジタル情報の利用環境の変化

所蔵するデジタル情報の利用環境を維持する事が困難な場合、所蔵するデジタル情報の記録メディアやデータファイル形式の変換をおこなう。また、その際に使用する新しい記録メディアやデータファイル形式は、継続的保存に適したものを選定する。

3 デジタル情報の利用提供

デジタル情報は、館内コンピュータやインターネットを利用して利用者へ提供する。

3.1 デジタル情報の閲覧

- (1) 館内において利用者が、デジタル情報の閲覧を希望した場合には、「沖縄県公文書館管理規則」（以下規則という）第8条及び第9条の規程に準じ、閲覧用の記録メディアを館内コンピュータを用いて閲覧提供する。
- (2) ただし、専用の利用環境が構築されている場合は、別に定める要領に準じる。

3.2 デジタル情報の複写

- (1) 利用者が、デジタル情報の紙への複写(出力)を希望した場合には、「沖縄県公文書館管理規則」（以下規則という）第10条の規程に基づき定めた"簡易複写"の扱いに準じる。
- (2) 利用者が、デジタル情報のデジタル記録メディアへの複写(複製作成)を希望した場合には、規則第10条の規程に基づき定めた"複写許可申請"の扱いに準じる。

3.3 デジタル情報の出版掲載等

利用者が、複写により入手したデジタル情報を、出版物等へ掲載を希望する場合には、規則第11条の規程に準じる。

3.4 デジタル情報の貸出

利用者がデジタル情報の貸出を希望する場合には、規則第12条の規程に準じる。

表2 データファイル管理票

データファイル管理票		版	作成日	作成者	責任者	No
		初	H15/02/10			001
データ	データファイル形式	TIFFファイル				
	データタイプ	<input type="checkbox"/> 文字 <input checked="" type="checkbox"/> 静止画像 <input type="checkbox"/> 音声 <input type="checkbox"/> 動画 <input type="checkbox"/> その他[]				
	データ特性	<ul style="list-style-type: none"> ・ハードウェアやソフトウェアに依存しない ・幅広い種類のアプリケーションがサポートしている。 ・静止画像を最も精細に記録できる。 ・非圧縮データ ・情報量が多いため、ファイル容量が大きい 				
	データ圧縮	<input checked="" type="checkbox"/> 未使用 <input type="checkbox"/> 使用 圧縮方式[]				
プラットフォーム	オペレーティングシステム (バージョン等に制約がある場合は必ず記述すること。)	<ul style="list-style-type: none"> ・WINDOWS, ・MAC-OS 				
	アプリケーションソフト (バージョン等に制約がある場合は必ず記述すること。)	<ul style="list-style-type: none"> ・AdobePhotoShop (WIN,MAC) ・他、幅広いソフトが対応 				
特記事項						

表3 記録メディア管理票

記録メディア管理票		版	作成日	作成者	責任者	No
		初	H15/02/10			001
メ デ ィ ア	メディア種別	CD-R				
	製品名	MITSUI GOLD				
	規格	Orange Book Part II Ver3.0				
	記憶容量	650MB				
	構造	フタロシアニン系の有機色材を使用				
	交換周期(想定寿命)	15年				
	劣化検査(部分抽出)	6ヶ月周期				
	メディア特性	<ul style="list-style-type: none"> ・ CD-Rは有機色素に記録させるため、磁気や電気による影響を受けない。 ・ 有機色素に一度記録された情報は、変える事ができない。 ・ フタロシアニン系の有機色素は外的要因に強く、通常の湿度や光度ではほとんど変化しない。 				
記 録 形 式	ファイルシステム	ISO9660				
	ファイルシステム特性	<ul style="list-style-type: none"> ・ 国際標準化機構(ISO)による国際標準規格 ・ OSに依存しない(Windows,Macで動作可) ・ ファイル名半角8文字、拡張子3文字の制限有り 				
	記録速度	4倍速				
メディアドライブ		<ul style="list-style-type: none"> ・ CD-Rドライブ ・ CD-ROMドライブ 				
特 記 事 項	<p>CD-Rの物理的寿命は、20年～100年といわれているが、その実績がないため、交換周期を15年とする。また、定期的な劣化検査時に、15年に満たないメディアに劣化が発生していた場合には、メディア交換周期を見直す。</p>					

表4 記録メディア目録データベース

メディア コード	メディア 種別	記録日	用途 種別	利用 制限	データ形式	メディア 記録票No	データ 記録票No
0000012532	CD-R	1995/10/10	保存	なし	TIFF	001	001
0000012533	CD-R	1996/05/08	利用	なし	PDF	001	003

当館の所蔵資料目録のデータベースは、複製物や重複物の管理を適正に行うために、タイトル、年代等のコンテンツ情報と、上記のメディア情報の2つのデータベースで構成される⁸。メディア情報にはこの他に、所在情報(書庫、書架、収納等)や受入情報(受入先名、受入先との取決事項など)などが記録される。

参考文献

若林宏『最新ファイル形式と文字コードの基本と仕組み』(株)秀和システム
 デジタル著作権を考える会 『デジタル著作権』ソフトバンクパブリッシング(株)
 柿井弘『デジカメ写真「あとで困らない」整理と保存術』(株)技術評論社
 オレンジフォーラムホームページ <http://www.cds21solutions.org/main/osj/>
 総務省ホームページ <http://www.soumu.go.jp/>

⁸ 沖縄県公文書館の所蔵資料目録データベースの構造や制御について、『沖縄県公文書館研究紀要第2号』「公文書目録情報のデータベースモデル ～階層構造を持つ目録情報のリレーショナルデータベースでの実相～」に詳細を記述している。