

国立国会図書館における デジタル資料の長期保存に関して

国立国会図書館 関西館
事業部 電子図書館課
藤原 誠

デジタルアーカイブシンポジウム
ーデジタルアーカイブとその長期利用に関する研究会ー
2007.2.16

1



目次

1. 電子情報の脆弱性
2. 電子情報の長期的な保存と利用についての調査研究
3. NDLデジタルアーカイブシステムの開発

2



目次

1. 電子情報の脆弱性
2. 電子情報の長期的な保存と利用についての調査研究
3. NDLデジタルアーカイブシステムの開発

3



納本制度

- 国立国会図書館は、日本国内で刊行される出版物を納本制度により広く収集し、文化財として長く保存する。
 - 国立国会図書館法第24条から第25条の2まで
- 納本の対象となる出版物
 - 図書、小冊子、逐次刊行物、楽譜、地図、レコード等
- 平成12年10月からパッケージ系電子出版物が納本の対象
 - 電子的方法、磁気的方法その他の人の知覚によっては認識することができない方法により文字、映像、音又はプログラムを記録した物

4



パッケージ系電子出版物の問題

- 電子媒体の寿命は20～30年程度と短い
- 再生機器の寿命は、部品の故障・劣化などにより、さらに短い
- 電子媒体や再生機器の規格は頻繁に変更される

5



インターネット情報(ウェブ情報)

- 社会基盤として重要な情報資源
- 国内のウェブサイト
 - 約4億5千万ファイル、約18.4TB(2005年3月時点)
 - 毎年数TB単位での増加
 - * <http://www.ndl.go.jp/jp/aboutus/bulkresearch2005summary.html>
- 不安定な情報資源
 - 内容の更新
 - URLの変更
 - 公開の中止
- 学術論文引用URLは4年で40～50%がアクセス不可
 - D. Spinellis, "The Decay and Failures of Web References" (2003)

6



電子情報の長期保存についての 世界的な取り組み

- UNESCO 「デジタル遺産の保存に関する憲章」
(2003)
- NDIIPP (National Digital Information
Infrastructure and Preservation Program)
米国
- DPC (Digital Preservation Coalition) 英国
- PADI (Preserving Access to Digital
Information) オーストラリア

7



目次

1. 電子情報の脆弱性
2. 電子情報の長期的な保存と利用について
の調査研究
3. NDLデジタルアーカイブシステムの開発

8



電子情報の長期的な保存と利用についての調査研究（平成14年度～16年度）

- 平成14年度の調査研究
 - 海外の動向、長期利用保証技術

- 平成15年度の調査研究
 - 当館所蔵資料の再生可能性調査

<http://www.ndl.go.jp/jp/aboutus/preservation.html>

9



電子情報の長期的な保存と利用についての調査研究（平成14年度～16年度）

- 平成16年度の調査研究
 - 長期利用保証技術の試行
 - マイグレーション
 - エミュレーション
 - ファイル形式変換

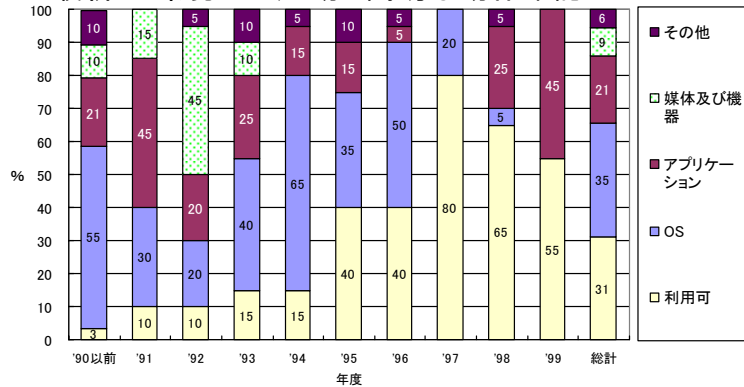
<http://www.ndl.go.jp/jp/aboutus/preservation.html>

10

当館所蔵資料の再生可能性調査

PC用電子資料200点(1990~1999年度受入分)

最新PC環境での起動と簡易な動作確認



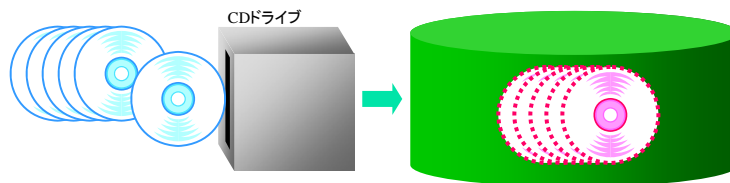
- 全体の7割弱の資料の利用に問題

11

マイグレーション

- 市販のマイグレーションツールにより、CD-ROMの内容をHDDに移行

- サンプル：354点



- 354点中352点成功(99.4%)

12



エミュレーション

- 市販のエミュレータにより旧式ハード環境を再現し、旧式OSをインストールし、電子資料の再生確認
- サンプル: 100点(プログラムを含む電子資料)

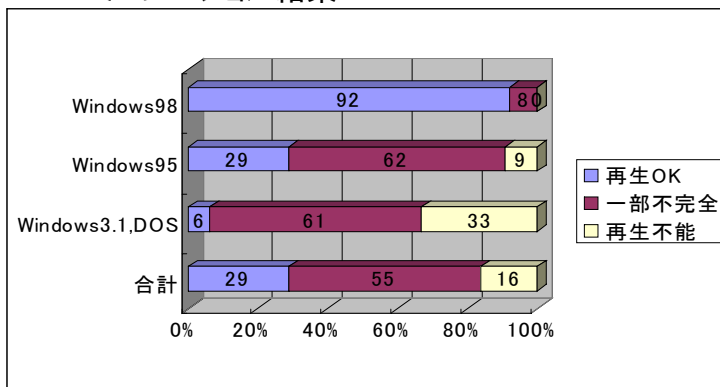
電子資料	電子資料	電子資料
DOS+Win3.1	Win95	Win98
エミュレータ		
WinXP		
新PC		

13



エミュレーション

- エミュレーション結果

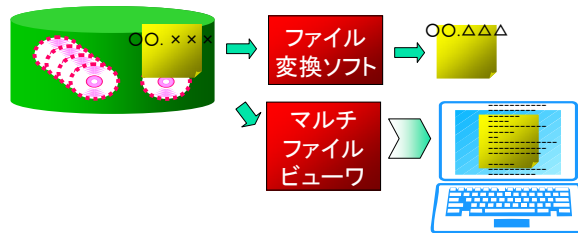


- 100点中29点再生成功(29%)

14

ファイル形式変換

- 市販のファイル形式変換ソフトによる変換後の再生確認、マルチファイルビューワによる再生確認
- サンプル: 100点(データのみの電子資料)中15点



- 15点中2点再生成功(13%)

15

長期利用保証のために

- 再生環境の維持管理
- マイグレーション、エミュレーションの技術は不可欠
- ファイル形式の標準化、規格化

16



電子情報の長期的な利用と保証についての調査研究（平成18年度～22年度）

調査内容

- マイグレーション及びエミュレーションに関する機能・技術等の要件調査
- 録音・映像資料のデジタル化に関する調査
- 電子情報保存に関するガイドラインの策定

17



目次

1. 電子情報の脆弱性
2. 電子情報の長期的な保存と利用についての調査研究
3. **NDLデジタルアーカイブシステムの開発**

18



NDLの電子図書館サービス

- 1998年「電子図書館構想」策定
 - 業務、システムの基盤を整備
- 2000年「電子図書館サービス実施基本計画」
- 2004年「電子図書館中期計画 2004」
 - 電子図書館サービスの高度化
 - 電子図書館サービスの具体的方向と実現に必要な枠組み

19



WARP (Web ARchiving Project)

インターネット情報選択的蓄積事業

- インターネット上のウェブサイトや電子雑誌を収集し文化資産として保存
- 国、地方公共団体、法人化前の国立大学、イベント等のHP、電子雑誌
- 平成18年度から本格事業化

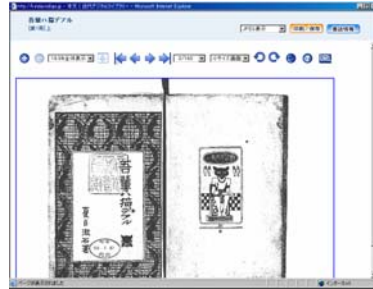


<http://warp.ndl.go.jp>

20

近代デジタルライブラリー

- 当館所蔵の明治期刊行図書を収録した画像データベース。
- 平成18年4月現在、約127,000冊を収録。
- 平成17年度から大正期刊行図書に着手



<http://kindai.ndl.go.jp>

21

NDL DAシステムの目的

- デジタル情報を広く収集する
- 長期保存し、かつ将来にわたって利用保証する
- 他システム等との連携・協働を図る
- システムの拡張性・移行容易性を確保する

22

NDL DAシステムの対象コンテンツ



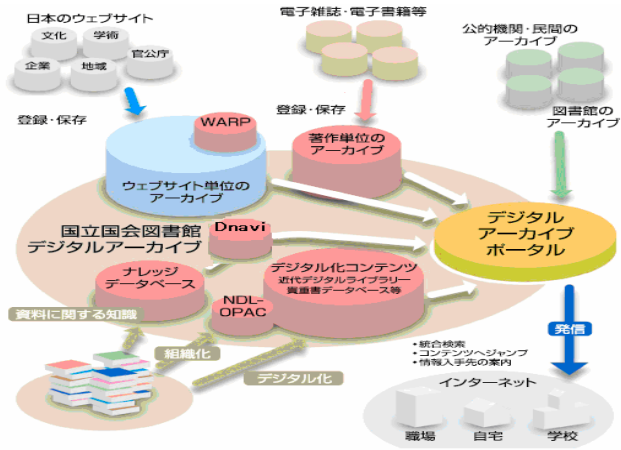
23

NDL DAシステムの基本方針

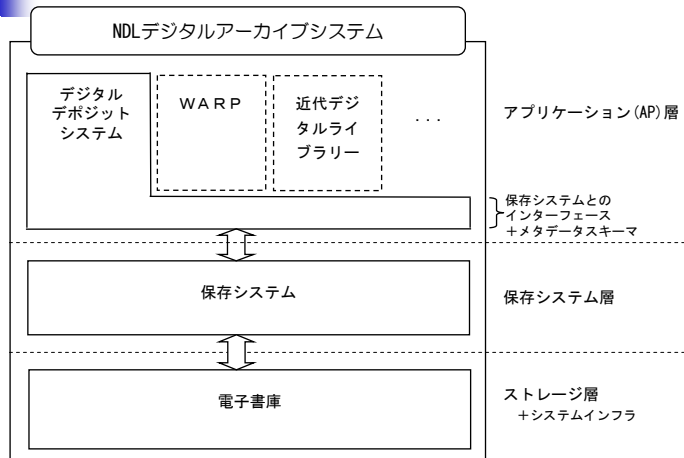
1. システム統合の容易性の確保
システムライフサイクルのトータルコスト削減
2. OAISへの準拠
3. ストレージ規模の拡張性
4. 館内他システムの統合
5. 他システムとの連携

24

NDL DAシステムのイメージ

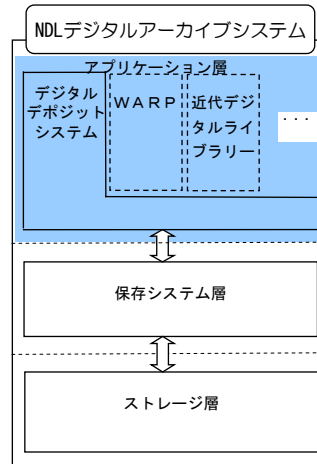


NDL DAシステムの構成



アプリケーション層

- 収集
 - クローラによる収集
 - 送信による収集
 - 媒体受付
- 組織化
 - メタデータ付与
- 提供
 - メタデータ検索
 - 全文検索
 - 統合検索

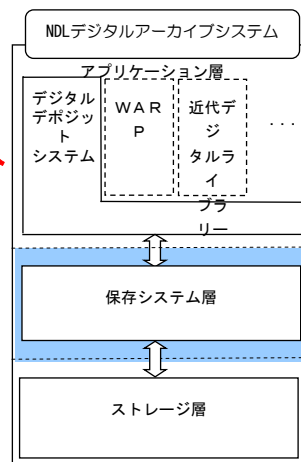


27

保存システム層

長期保存のための情報パッケージ管理
 <OAIS準拠>

- アプリケーション層からの情報の**受入**
- アプリケーション層への情報の**提供**
- **メタデータ**の管理
- **永続的識別子**の管理
- **長期利用保証**
 - マイグレーション/エミュレーション

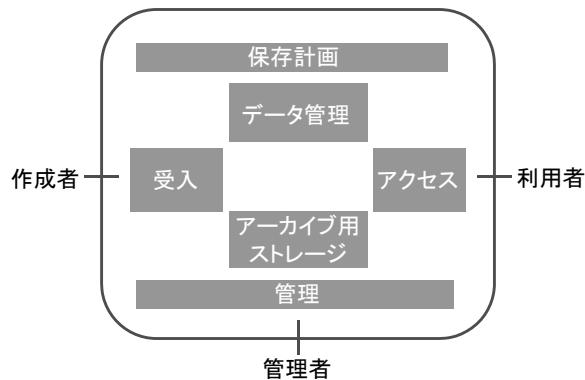


28

OAIS参照モデル

Open Archival Information System(ISO 14721:2003)

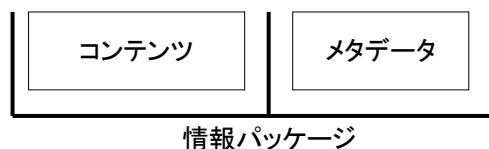
- 電子情報の長期保存システムの抽象的な仕様を規定した技術標準



29

情報パッケージ

- 情報パッケージ
 - メタデータとコンテンツをまとめて保存
 - METS (Metadata Exchange and Transmission Standards)



30



NDL DAシステムのメタデータ スキーマ設計

設計の基本方針

- 長期的な可読性、相互運用性
 - XMLベース
 - 標準的な既存のメタデータスキーマ
- 将来的な拡張性と柔軟性

31



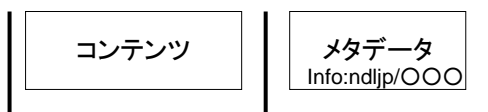
NDL DAシステムのメタデータ構成

- 記述メタデータ (MODS)
- 技術メタデータ
- 権利メタデータ
- 保存メタデータ
- 管理メタデータ

32

NDL DAシステムの永続的識別子

- コンテンツに付与
- 永続性
- 一意性
- 相互運用性
- info URIスキームを使用
 - info:ndljp/〇〇〇

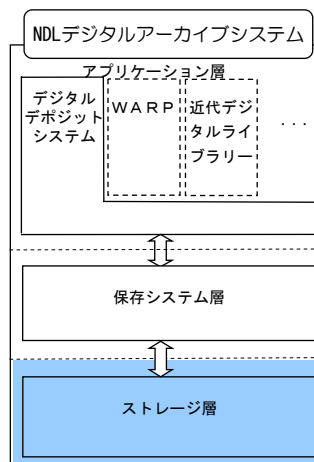


情報パッケージ

33

ストレージ層(電子書庫)

- AIP(情報パッケージ)を永久不変に蓄積
- 拡張性
- 今後の技術革新に対応
- 特定技術に依存しない
- 耐障害性



34



NDL DAシステムの課題

■収集・組織化・提供

- 方針策定
- 方法検討
- 計画等策定

■他機関との連携・協働

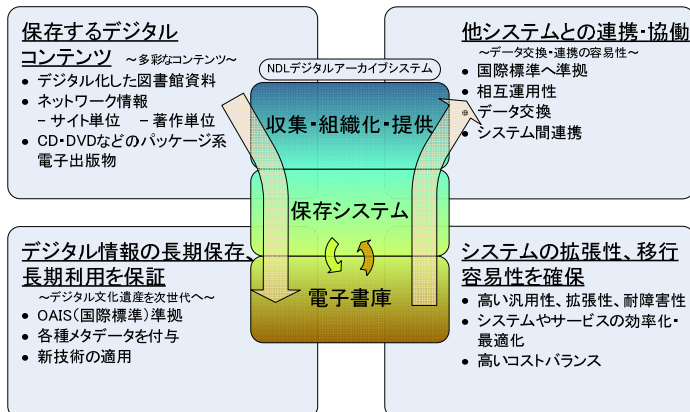
- 普及・広報活動
- システム連携・相互運用性
- 標準化、規格化

35



まとめ

NDLデジタルアーカイブシステム



36