

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

講演資料とまとめ

2007年3月

まえがき

本資料集は、平成 18 年度筑波大学図書館情報メディア研究科プロジェクト研究「デジタル情報資源の長期保存のための総合的アーカイブ環境」の一部として、2007 年 2 月 16 日(金)に筑波大学春日キャンパスにおいて開催した「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」の講演資料とパネルディスカッションでの議論をまとめたものである。

インターネットの爆発的な発展とともに非常に多くの情報資源が電子的に発信されるようになった。たとえば、国立図書館や大学図書館を中心として歴史資料や貴重資料の電子化が 90 年代から広く行われてきている。現在の大学図書館では電子ジャーナルや機関リポジトリにみられるような学術情報の電子的な出版や発信が一般化している。大きな情報発信減である行政機関では行政情報の Web 上での発信や公文書の電子化が進んでいる。また、玉石混合とは言え、現在の文化を将来に伝えるうえで有用なさまざまな情報が Web 上に置かれている。こうしたさまざまなネットワーク上の情報資源に加え、CD や DVD 等を用いて出版されてきたものも含めて現代の文化が作り出したデジタル情報資源である。このような資源を、将来に渡って利用できる状態で保存していくことが重要であることは疑えない。

デジタル情報資源の長期保存は、1990 年代に進められたデジタルライブラリの研究開発の中で、難しいが解決する必要のある問題として広く認められ、90 年代から欧米各国でも国立図書館や大学図書館を中心とした取り組みが進められてきている。アメリカでは、議会図書館に大きな予算がつけられ National Digital Information Infrastructure and Preservation Program (NDIIPP)が進められてきた。我が国においても国立国会図書館における Web アーカイブの実証実験への取り組みなどがある。

一方、デジタルアーカイブを長期に利用できるようにするにはデジタル情報資源そのものの保存だけではなく、ひとまとまりのアーカイブとしてどのようなサービスを提供することが求められるかといった問題や、知的財産権やプライバシーといった社会環境に関わる問題も合わせて考える必要がある。そこで、本シンポジウムはできるだけ広い視野での議論を目的として講演をお願いした。

本シンポジウムでご講演をいただいた皆様方に深い感謝の意を表したい。また、本シンポジウムが、デジタルアーカイブの長期利用に関する今後の研究開発に関していろいろな視点からを提供することができれば幸いに思う。

筑波大学・図書館情報メディア研究科
知的コミュニティ基盤研究センター
杉本重雄

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

目次

講演資料

・デジタルアーカイブシンポジウム講演資料(発表順)

- (1) 杉本重雄(筑波大学・図書館情報メディア研究科／知的コミュニティ基盤研究センター)
デジタルアーカイブとその長期利用に関して……………1
- (2) 藤原誠 (国立国会図書館関西館)
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……………13
- (3) 牟田昌平(国立公文書館)
公文書の電子的保存に関して-内閣府の取り組み- ……………32
中島康比古(国立公文書館)
公文書の電子的保存に関して-国立公文書館の取り組み-……………39
- (4) 和田光俊 (独立行政法人科学技術振興機構文献情報部)
学術雑誌のデジタルアーカイブと長期利用に関して……………49
- (5) 西岡貞一(筑波大学・図書館情報メディア研究科)
文化財のデジタルアーカイブと長期保存に関する問題 ……………63
- (6) 新保史生(筑波大学・図書館情報メディア研究科)
デジタルアーカイブと法……………79
- ・パネルディスカッション……………87

パネリスト

大場利康(国立国会図書館関西館)
田窪直規(近畿大学短期大学部)
中井万知子(国立国会図書館書誌部)
中島康比古(国立公文書館)
司会 杉本重雄

参考：講演の案内およびプログラム……………89

デジタルアーカイブとその長期利用に関して

筑波大学・図書館情報メディア研究科・知的コミュニティ基盤研究センター 杉本重雄

内容

1. はじめに
 2. デジタルアーカイブを俯瞰する
 3. デジタルアーカイブと長期保存、長期利用
 - (1) 「アーカイブ」ということば
 - (2) アーカイブと保存
 - (3) 何を保存するか？
 - (4) 保存のための文書機能、品質の低下はどの程度許されるか？
 - (5) アーカイブは長命か？
 - (6) 非デジタル資料のデジタル化とデジタル資料の保存
 4. Web 上の情報資源の保存 - Web アーカイブ
 5. 機関リポジトリとアーカイブ
 6. 公文書のアーカイブ
 7. デジタルアーカイブとメタデータ
 8. おわりに
- 参考資料

デジタルアーカイブとその長期利用に関して

杉本重雄

筑波大学・図書館情報メディア研究科・知的コミュニティ基盤研究センター

1. はじめに

デジタル情報技術の発展とともに多様なデジタルコンテンツが作られてきた。特に、90年代のインターネットの爆発的な広がりとともに、ネットワークを通じて発信されるデジタルコンテンツが飛躍的に増えた。それとともに、デジタルコンテンツを集積したデジタルアーカイブ(Digital Archive)が発展してきた。現在、デジタルアーカイブということば自体、かなり広い意味で用いられていると感じられる。一方、デジタルコンテンツの長期利用性が必ずしも保証されないことは広く知られており、デジタルアーカイブを長期に渡って使い続けることはデジタルアーカイブにとっての共通の問題であるといえる。そこで、ここでは、デジタルアーカイブについて、できるだけ広い視野から眺め、デジタルアーカイブの長期利用に関して考えてみたい。

2. デジタルアーカイブを俯瞰する

アーカイブということばは、将来に向けて、現在ある文書あるいは記録を集め、蓄積し、それを長期に渡って提供するサービス、あるいはそれを行う組織を意味する。デジタルアーカイブの場合、デジタル情報技術を用いて表現されたコンテンツを集め、蓄積して行うサービスととらえることができる。これまで「デジタルアーカイブ」として、蓄積されてきたコンテンツには、電子化された文書、あるいは電子的に作られた(Born Digital)文書や記録などが含まれるのみならず、建造物や遺跡、美術工芸品などの物体をデジタルコンテンツ化したものも含まれる。そこで、ここではこれらを一括に分類してみよう。

1. 貴重書や歴史的な文書などを提供するもの
2. 公文書(行政文書)を保存、提供するもの
3. 学術論文を収集し、提供するもの
4. Web上の資料を収集し提供するもの
5. さまざまな文化遺産を電子化し、提供するもの
6. 電子的に出版された図書や雑誌を提供するもの
7. そのほか、教育情報資源や特定の分野の情報資源などを収集、蓄積、提供するもの

デジタル情報資源はCDやDVDなどのパッケージ型のメディアに作りこまれて提供されるものと、ネットワーク上で提供されるものがある。また、一般に、デジタル情報資源は物理的

な資料を、スキャナなどを用いてデジタル化したものと、ワープロなどで作られた「生まれながらにデジタル形式の資源 (Born Digital)」がある。前者はデジタル化による付加価値を目的にデジタル化されるものといえる。現在の我々の日常的な活動の中でワープロを使わずに文書を作ることが少なくなってきたことに代表されるように、電子的に文書を作るとは日常の一部になっている。電子的な文書の場合、リンクや画像や音声の埋め込みなど、紙の文書では実現できない機能が多く取り入れられている。また、物理的な物体から仮想現実感機能を用いて作り出された資源の場合、これらの両方の性質を持つとって良いかもしれない。いずれの場合も、情報資源の利用には適切なハードウェアとソフトウェアが要求される。このことが、デジタル情報資源の保存を難しい問題にしている。

3. デジタルアーカイブと長期保存、長期利用

デジタルアーカイブを長期に渡って利用すること、あるいはデジタルアーカイブに蓄積された資源を長期に渡って保存することの重要性は疑えない。ここでは、デジタルアーカイブとは何か、保存に関わる基本的な問題はなにかといった、基本的な項目について考えてみたい。

(1) 「アーカイブ」ということば

アーカイブは、本来、収集した資料を長期に渡って保存することを前提としてきたと思われる。しかしながら、2 節に示した 1～7 に示すデジタルアーカイブということばの使われ方をみると、デジタル形式で蓄積・利用できるようにすることに必ずしも「長期の保存」を含意していない。しかしながら、実際にはそれなりのコストをかけて作成したアーカイブが短期間（たとえば、5 年あるいは 10 年）で利用できなくなるのはあまりに無駄が大きい。その意味では、デジタルアーカイブには、長期に渡ってサービスを提供し続けることが暗黙の内に求められていると考えたい。

(2) アーカイブと保存

収集して、保存ならびに提供するということがアーカイブにとっての基本機能であろう。一方、デジタルアーカイブの場合、これらを分離してとらえることができる。たとえば、図書館が持つ貴重資料をデジタル化し、デジタル化した資料の長期保存を任務とする別の組織に送って保存を依頼する一方、デジタル化資料を提供に適した形式に変えてネットワーク上で提供することができる。デジタル情報資源のコピーと処理のしやすさを考えると提供の際のさまざまな付加価値サービスが可能である。

(3) 何を保存するか？

いうまでもないが、デジタル情報資源は、紙やものと異なり、物理的に形を持つものではない。パッケージ型の資源であっても、パッケージ（たとえば、1 枚の CD）を残すこととその中身を残すこととは異なる。また、デジタル情報資源は利用する環境や利用者に応じて表示形式を変えることができるように設計されていることもある。一方、保存に際してデジタル資料をも

そのまま残すことは必ずしも用意ではない。こうしたことは、デジタル情報資源の保存には、デジタル情報資源の「何を」保存すべきかを定めることも含まれると理解できる。たとえば、みばえや使い勝手（Look and Feel）も含めた完全な保存、データとしての完全な保存、テキストだけの保存、スナップショットの保存などいろいろなケースが考え得る。

(4) 保存のための文書機能、品質の低下はどの程度許されるか？

デジタル資料、特に Born Digital 資料の保存の際に文書が持つ機能の停止や品質の低下を避けられないことがある。たとえば、ハイパーリンクを持つ文書の保存の場合、リンクが行き続けることを保証することは困難である。CAD システムで作られ、利用されている文書（たとえば、設計図）を、もとのまま保存するには CAD システムそのものの保存が求められることになる。こうした場合、保存のために機能の低下を避けることはできない。逆に、機能の低下を許すことで保存が可能になるとも言える。そこで求められるものは許し得る機能や品質低下の範囲を決めるガイドラインであろう。

(5) アーカイブは長命か？

最近の省庁の再編、自治体の統廃合や大学の統合など、従来はとても安定していると思っていた組織が簡単に变化してしまうことを経験してきた。こうした組織の改変はデジタルアーカイブにも影響する。アーカイブの高信頼化のために、蓄積した資源をアーカイブ間で移動したり、共有したりする必要がある場合もあろう。また、法律や規則の変更によりアーカイブの運用方法が変化することもある。そうした変化に対応しながらアーカイブを運営できないと、アーカイブはブラックホールようになってしまう可能性もある。

(6) 非デジタル資料のデジタル化とデジタル資料の保存

非デジタル資料をデジタル化することは、原資料の保存とアクセス性の高いデジタル化資料の提供という面から進められてきている。デジタル化にかかるコストとアクセス性を高めることによって得られるメリットに関する議論は必要ではあろうが、デジタル化によって我々がネットワーク越しに利用できる情報資源が豊かになることは間違いない。また、提供されるデジタル化資料を基礎として作り出されるさまざまな情報や知識によって、我々の情報環境はより豊かになっていく。その一方、デジタル化した資料であっても、適切にメンテナンスをしていかなければ利用できなくなっていく。デジタル化資料そのものだけが失われると、その上に作り上げられたいろいろな情報資源も影響を受けてしまう。

4. Web 上の情報資源の保存 - Web アーカイブ

Web は我々の情報基盤となっている。Web を介してさまざまな情報が発信されている。Web 上でしか発信されない情報も多くある。Web 上の情報はよく変化する。安定しない情報資源なので保存する価値がないという議論もあるが、逆にその時代を表す情報がとてもよく現れるから保存すべきであるとも言える。いずれにしても Web 上の情報資源の完全な保存は困難である。加え

て、Born Digital 情報資源に共通の保存の困難さ、違法コンテンツやウイルスなどの問題もある。

Web アーカイブは大別すると人手による選択的収集によるものと、ソフトウェアロボットを用いて行う網羅的収集によるものがある。前者の場合、内容の評価に基づいて収集されるが大規模化には困難を伴う。後者の場合、大規模化は可能であるが、基本的に内容に関わらず収集することになり、内容による評価を効率的に行うことが求められる。後者の代表として良く知られているものにインターネットアーカイブ (<http://www.archive.org/>) がある。これはインターネットを介して公開された Web ページを収集し、蓄積保存するものである。納本図書館である国立図書館の場合には、自国の、あるいは自国に関する Web ページを収集することになる。欧米の国立図書館を中心に Web アーカイブのためのコンソーシアムを作っている[1]。また、Web ページは従来の出版物とは性質が異なるため、各国で法律を作るなどして対応している。我が国の場合、国立国会図書館の納本制度審議会がネットワーク系出版物に関する答申を出しているが、現時点で法律改正にまではいたっていない[2]。

インターネット上の Web ページを網羅的に収集するシステムは、リンクが与えられないと Web ページを収集することができない。また、イントラネット内部のページのように外部からは見えないページも多くある。基本的に、こうしたページはインターネット上での Web アーカイブの対象外と考えることもできる。一方、そうしたページを発信している組織が組織自身の Web アーカイブを行う場合、いわば組織内 Web アーカイブを考えることもできる。インターネットだけではなく、イントラネットを利用した Web アーカイブも同様に考えることはできる。一方、組織のアーカイブポリシーに基づいて行う Web アーカイブも必要である。

5. 機関リポジトリとアーカイブ

現在、我が国のみならず欧米各国でも機関リポジトリ (Institutional Repository) の開発が盛んに進められている。機関リポジトリは学術雑誌の電子ジャーナル化の発展と学術雑誌の値段の高騰によって後押しされている。機関リポジトリでは論文や報告書などのコンテンツの収集蓄積と提供に目が行きがちであるが、保存の視点も忘れてはならない。

機関リポジトリは、一般に大学図書館のように安定した組織が提供している。一方、現在では、論文や紀要などの発信は誰であってもできる。また、一般の利用者にとっては Google あるいは Google Scholar のようなサービスの方が、特定のリポジトリを探し出して使うよりはなじみやすい。こうした環境を考えると、収集した資源の長期の保存は、安定した組織が提供する機関リポジトリの重要な要件であるように感じる。学習コンテンツについても同様である。

大学や研究機関では、論文にはのらないけれども面白い内容のページが提供される。こうしたページは一般には消えていく。(運よく Web アーカイブの対象になるものもあるが。) 機関リポジトリを組織として行うアーカイブとしてとらえれば、組織内の Web アーカイブと機関リポジトリの組み合わせで、より内容の豊富なアーカイブになることが期待できる。

6. 公文書のアーカイブ

公文書は歴史資料として将来に残さねばならない。紙の資料はこれまで同様の方法で残していくことができる。しかしながら、電子政府化や電子自治体化の進展にともない、Born Digital の資料が増えることは疑えない。そこで、デジタル情報資源の保存に共通の問題以外に、直感的に思いつく問題点をいくつか挙げてみたい。

- (1) 電子公文書とは何かに関する定義： 電子メールや Web ページ、掲示板への書き込み、ブログは保存すべき公文書か、といった問題以外に、データとテンプレートあるいは適切なスクリプトがあれば、印刷形式や表示形式を動的に作り出すことは簡単であり、その際、何が記録として残すべきものか（もとのデータなのか、作り出された印刷形式なのか）がはっきりしない。
- (2) 保存のために許される機能と品質低下の範囲： デジタル文書の保存には機能低下が避けられない。どのような機能や品質の低下が許容範囲であるのかについてのガイドラインなしの保存は困難であろう。
- (3) 原本性の保証： デジタル資料は簡単にコピーができる。信頼できるアーカイブ (Trusted Archive) から外部に提供されたコピーについて、そのものだけで原本性を保証できるようにする必要があるのかどうか、またあるとした場合に長期間有効な保証方法はどのようなものが適切なのかといった問題がある。

以上のような問題のほかに、資料のメタデータをいかに効率よく作り出すかと言った問題もある。資料の内容に関わるメタデータに関しては、保存の現場で作成するより、資料の作成現場で作成するほうが効率的であることは疑えない。公文書の保存は公文書のライフサイクル全体から考えなければならない問題である。

7. デジタルアーカイブとメタデータ

デジタルアーカイブにおいてメタデータは重要な要素である。デジタルアーカイブでは、資料の検索や内容に関する記述を与える記述メタデータ (Descriptive Metadata)、資料のアクセスや保存管理のための管理メタデータ (Administrative Metadata)、資料の (物理的、論理的) 構造を表すための構造メタデータ (Structural Metadata) といった、異なる側面からの記述が求められる。

デジタル資料のアーカイブと長期保存のためのメタデータとしては Open Archival Information System (OAIS)参照モデル[3]を基礎として提案されたもの、PREMIS ワーキンググループ[4]によって提案されたものがある。また、EAD (Encoded Archival Description)[5]や METS (Metadata Encoding and Transmission Standard)[6]といったメタデータスキーマも提案されている。実際には、アーカイブする資料の種類、アーカイブの目的などによって適切なスキーマが選択されることになる。また、一般に、保存のためのメタデータ記述は項目が多くなる。管理情報や構造情報など付与の自動化が行いやすい部分もあるが、メタデータ付与の効率化が求められる。

8. おわりに

ここでは、ごく簡単にデジタルアーカイブの長期利用に関して概観してみた。ネットワーク情報化社会は今後もますます進んでいくことは疑えない。電子出版物もますます増えていくことは疑えず、行政情報やサービスの電子化も進んでいく。我々の周りにデジタル情報資源がどんどん増えていくことは間違いないにもかかわらず、それらを適切に将来に残していくことに関しては心もとない状況であると思う。筆者は数年前にデジタルアーカイブについて解説記事を欠いたことがある[7]。そのころと比べると理解は進んできたように思う。その一方、90年代からデジタル情報資源の保存の重要性と難しさは言われてきたが、それが現実の問題として現れてきたという状況のように思う。

参考資料

- [1] International Internet Preservation Consortium, <http://netpreserve.org/>
- [2] 国立国会図書館納本制度審議会, ネットワーク系電子出版物の収集に関する制度の在り方について, 2004, http://www.ndl.go.jp/jp/aboutus/data/a_toushin_2.pdf
- [3] Reference Model for an Open Archival Information System (OAIS), <http://public.ccsds.org/publications/archive/650x0b1.pdf>
- [4] PREMIS Final Report, <http://www.oclc.org/research/projects/pmwg/premis-report.pdf>
- [5] Encoded Archival Description, <http://www.loc.gov/ead/>
- [6] Metadata Encoding and Transmission Standard, <http://www.loc.gov/standards/mets/>
- [7] 杉本重雄, Maria Luisa Calanag. デジタルアーカイブとメタデータ, 人工知能学会誌, Vol.18, No.3, pp.217-223, 2003

デジタルアーカイブの長期利用に関して

杉本重雄

筑波大学・図書館情報メディア研究科・知的コミュニティ基盤研究センター

デジタルアーカイブの概観

- 「デジタルアーカイブ」とは？
 - （長期に渡る）提供を目的として、デジタル資料を収集、蓄積したもの
- デジタルアーカイブのコンテンツは？
 - 図書やマニュスクリプト、文化遺産などをデジタル化して作成した資料
 - 電子文書や電子的記録（データ）など、もともと電子的に作られた資料

デジタルアーカイブの概観

- デジタルアーカイブ
 - 貴重書や歴史的文書などを提供するもの
 - 公文書(行政文書)を保存、提供するもの
 - 学術論文を収集し、提供するもの
 - Web上の資料を収集し提供するもの
 - さまざまな文化遺産を電子化し、提供するもの
 - 電子出版物(図書や雑誌)を提供するもの
 - そのほか、教育情報資源や特定分野の情報資源などを収集、蓄積、提供するもの

デジタルアーカイブと長期利用、長期保存

- アーカイブにもいろいろある
- アーカイブと保存
 - 収集・蓄積、保存、提供を分けることができる
- 資料の何を保存するか
 - 保存すべきものは何なのか
 - Look and Feelまで、それとも中身のテキストだけ
- 保存のための文書機能、品質の低減(degrade)
 - 機能や品質を100%保証することは難しい
 - degradeはどの程度許されるのか。ガイドラインの必要性

デジタルアーカイブと長期利用、長期保存

- アーカイブは長命か
 - アーカイブそのもののおかれる環境、制度が変わることもある
- 非デジタル資料のデジタル化とデジタル資料の保存
 - デジタル化のコストとメリット
 - デジタルアーカイブの上に作られる新たな情報と知識

いろいろなデジタルアーカイブ

- 貴重資料や文化遺産のアーカイブ
 - 非デジタル資料のデジタル化
 - 図書、マニュスクリプトから、自然、遺跡、無形文化財まで
- Webアーカイブ
 - Web上には現代のホットな話題を表した資料がいっぱいあるから将来に残したい
 - インターネットアーカイブ、イントラネットアーカイブ
 - 組織のポリシーに基づくアーカイブ
 - メタデータのコストはできるだけ下げたい

いろいろなデジタルアーカイブ

- 機関リポジトリ
 - 組織(機関)が発信する学術情報資源を収集提供することが注目されている。収集したものは保存しなければならない。
- 公文書のアーカイブ
 - 電子政府や電子自治体の進展とともに電子公文書や電子データの保存が求められることになる
 - Born Digital資料の完璧な保存は不可能
 - 保存の観点から、文書作りに制約をかけることは不可
 - 文書のライフサイクル全体を見たガイドラインの必要性

デジタルアーカイブとメタデータ

- メタデータは重要な要素
 - メタデータだけでは保存はできないが、メタデータなしには保存はできない
- OAIS
 - デジタルデータ、電子文書の保存のための参照モデルとメタデータ
- PREMIS
 - デジタル情報資源の保存のためのメタデータ

まとめ

- デジタルアーカイブということばで表されるものであっても、目的と内容に応じていろいろなものがあるので、ひとくくりにはしにくい
- それなりのコストをかけて作ったものは、長期に渡って利用できるようにしなければならない
- デジタル情報資源を100%もとのまま利用可能な状態で保存し続けることは不可能
 - 現実的な保存方法を合意する必要がある
- 技術面だけではなく、制度、利用者の側面も忘れてはならない

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

内容

1. 電子情報の脆弱性
 - ・ 納本制度
 - ・ パッケージ系出版物の問題
 - ・ インターネット情報（ウェブ情報）
 - ・ 電子情報の長期保存についての世界的な取り組み
2. 電子情報の長期的な保存と利用についての調査研究
 - ・ 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）
 - 当館所蔵資料の再生可能性調査
 - ◇ マイグレーション
 - ◇ エミュレーション
 - ◇ ファイル形式変換
 - ◇ 長期利用保証のために
 - ・ 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）
3. NDL デジタルアーカイブシステムの開発
 - ・ NDL の電子図書館サービス
 - ・ W A R P（Web ARchiving Project）
 - ・ 近代デジタルライブラリー
 - ・ NDL DA システムの目的
 - ・ NDL DA システムの対象コンテンツ
 - ・ NDL DA システムの基本方針
 - ・ 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
 - 保存システム層
 - OAIS 参照モデル
 - 情報パッケージ
 - ・ NDL DA システムのメタデータスキーマ設計
 - NDL DA システムのメタデータ構成
 - NDL DA システムの永続的識別子
 - ・ ストレージ層（電子書庫）
 - ・ NDL DA システムの課題
4. まとめ

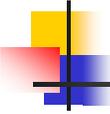


国立国会図書館における デジタル資料の長期保存に関して

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

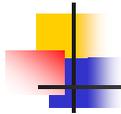
1



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

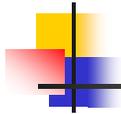
2



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

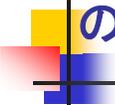
7



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



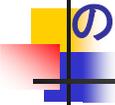
電子情報の長期的な保存と利用についての調査研究（平成14年度～16年度）

- 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

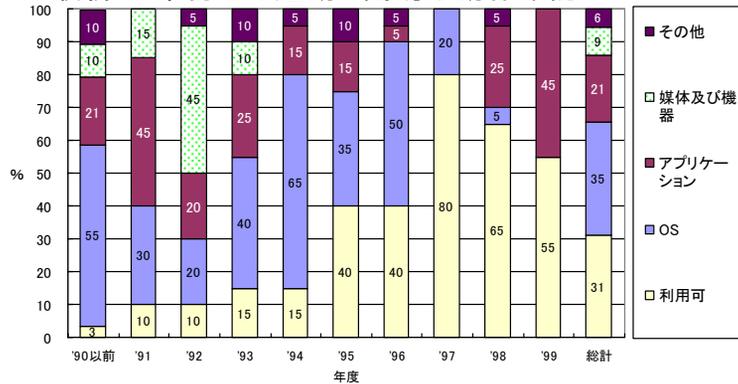
<http://www.ndl.go.jp/jp/aboutus/preservation.html>

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



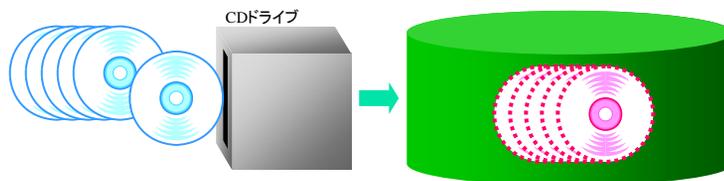
- 全体の7割弱の資料の利用に問題

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)

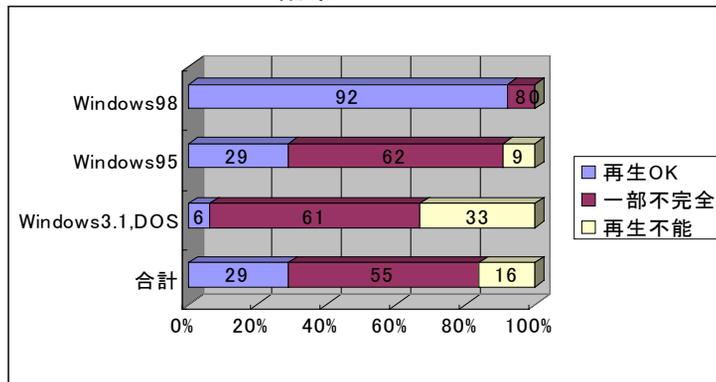
電子資料	電子資料	電子資料
DOS+Win3.1	Win95	Win98
エミュレータ		
WinXP		
新PC		

13



エミュレーション

- エミュレーション結果

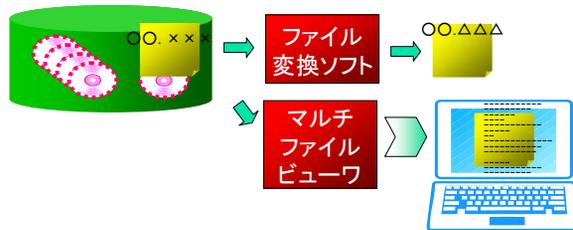


- 100点中29点再生成功(29%)

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

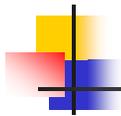
17



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

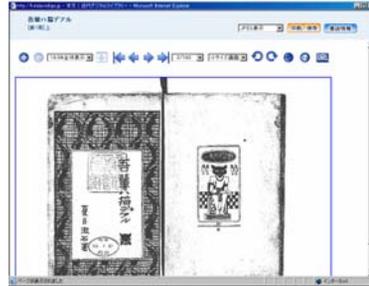


<http://warp.ndl.go.jp>

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



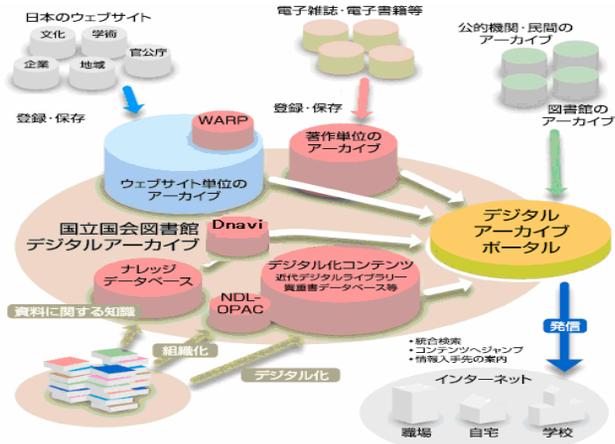
23

NDL DAシステムの基本方針

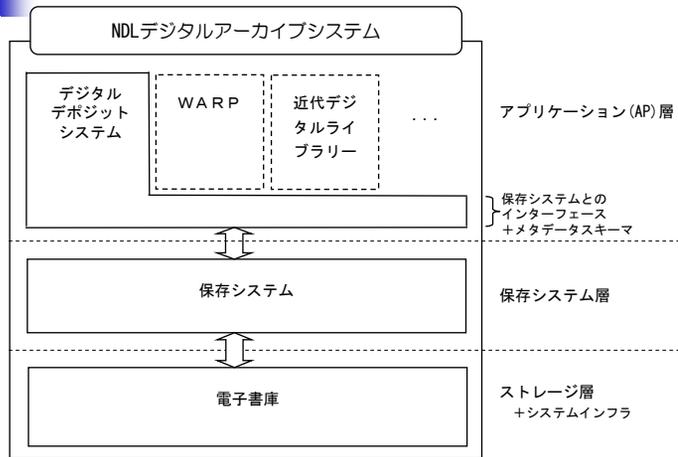
1. システム統合の容易性の確保
システムライフサイクルのトータルコスト削減
2. OAISへの準拠
3. ストレージ規模の拡張性
4. 館内他システムの統合
5. 他システムとの連携

24

NDL DAシステムのイメージ

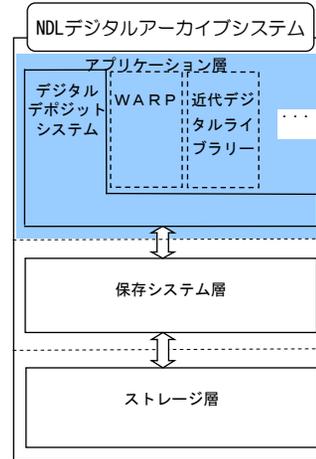


NDL DAシステムの構成



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

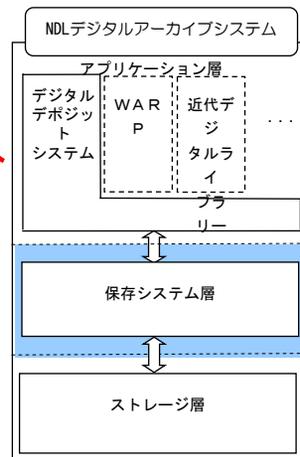


27

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

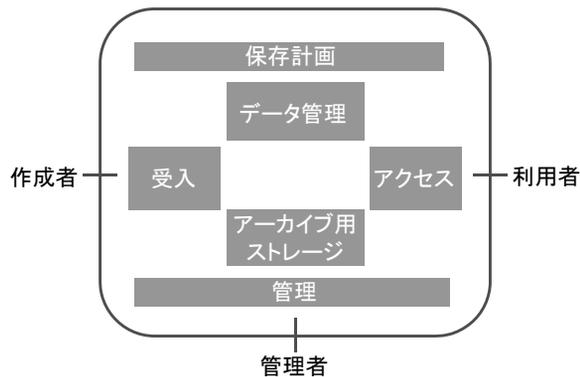


28

OAIS参照モデル

Open Archival Information System(ISO 14721:2003)

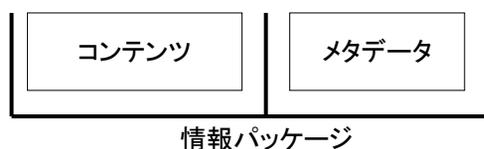
- 電子情報の長期保存システムの抽象的な仕様を規定した技術標準



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



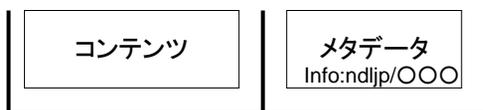
NDL DAシステムのメタデータ構成

- 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

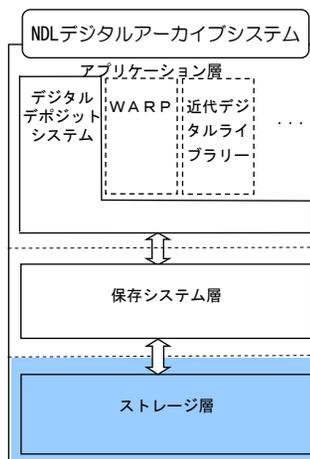


情報パッケージ

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

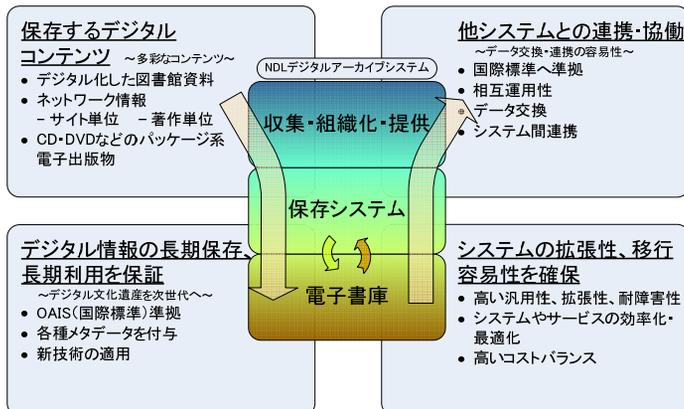
- 普及・広報活動
- システム連携・相互運用性
- 標準化、規格化

35



まとめ

NDLデジタルアーカイブシステム



36

公文書の電子的保存に関して ― 内閣府の取り組み ―

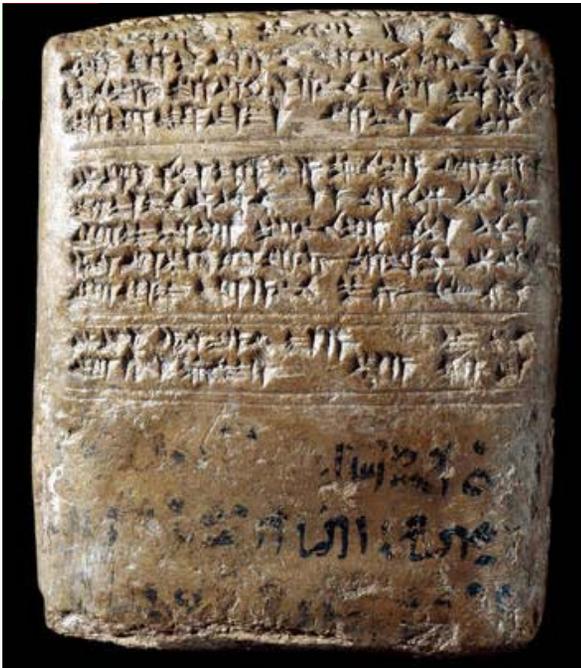
国立公文書館 牟田昌平

内容

- ・ (Letter from Tushratta to Amenhotep III 1370-1350 BC, British Museum)
- ・ (英国公文書館内展示物)
- ・ 「デジタルアーカイブ」は和製英語
- ・ 公文書館の目指す「デジタル・アーカイブ」・Born Digital のアーカイブ
- ・ (参考 総合的な文書管理システムの概念図)
- ・ 内閣府の取り組み① 歴史資料として重要な公文書等の適切な保存・利用のための研究会
- ・ 内閣府の取り組み② 公文書の適切な管理、保存及び利用に関する懇談会
- ・ 内閣府の取り組み③ 電子媒体による公文書等の管理・移管・保存に関する研究会
- ・ 「電子媒体による公文書等の管理・移管・保存のあり方」 電子公文書等の長期保存を考える上での9つの基本的視点
- ・ 「電子媒体による公文書等の管理・移管・保存のあり方」 電子公文書長期保存の4つの論点
- ・ 国立公文書館デジタルアーカイブの将来像

デジタルアーカイブシンポジウム
筑波大学知的コミュニティ基盤センター
デジタルアーカイブとその長期利用に関する研究会
公文書の電子的保存に関して
内閣府の取り組み

国立公文書館
公文書専門官
牟田昌平
平成19年2月16日



Letter from Tushratta to
Amenhotep III 1370-1350 BC
British Museum

- ・国際語のアカド文字で書かれた公的な手紙
- ・焼かれた粘土板
- ・アマルナ遺跡の「ファラオの書庫」から大量に発見
- ・エジプト文字には受け入れた年代の記述
- ・3000年以上経っても「歴史記録」としての意味を持つ基本情報(メタデータ)を含む



物理的に記録された情報は長期保存に耐える



英国公文書館内展示物
撮影:国立公文書館

・電子データは
そのままでは読
めません
・記録媒体の耐
用年数は不明
・技術進歩に
よって電子記録
媒体そのものが
時代遅れ



・Being Digital
・ユビキタス社会

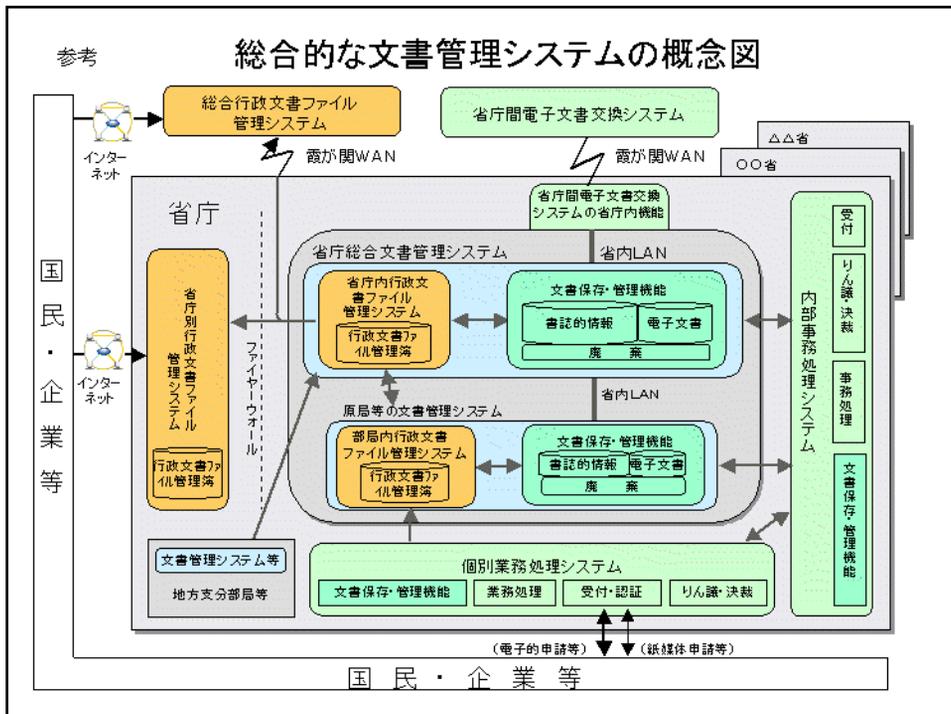
「デジタルアーカイブ」は和製英語

- 1990年代に日本で生まれた言葉
 - 1995年日本で開催された「デジタルアーカイブ国際会議 '95」
 - 主導したのは
平山郁夫 学長(東京芸術大学)
杉田繁治 教授(国立民族学博物館)
月尾嘉男 教授(東京大学)
文化資産とマルチメディアとの融合による
- デジタルアーカイブとは？(デジタルアーカイブ国際会議の定義)
「人類の文化遺産には、修復と保存のための懸命な努力にもかかわらず崩壊、消失していくものが多数あります。これを次世代に正しく継承するための手段として、マルチメディア技術を駆使してデジタル映像の形で記録保存し、さらに情報ネットワークを利用して随時閲覧でき、広く情報発信するシステムが「デジタルアーカイブ」です。」



公文書館の目指す「デジタル・アーカイブ」 Born Digitalのアーカイブ

- e-Japan構想と電子政府(e-Gov)
 - E-Govの最大の目的は情報化による効率化
 - 電子決裁→電子記録は電子で保存
 - 何を残すのか、何処まで残すのか
 - 文書館制度の要である評価選別の不在
 - 当初の総合文書管理システムに公文書館制度は含まれず
 - 文書のライフサイクルに組み込まれていない永久保存と公開
- 電子文書の長期保存の課題
 - 電子情報は目に見えない(可視性の確保)
 - メタデータが無ければ只のゴミデータ
 - 宇宙線でも影響を受ける電子記録
 - 微細になればなるほど脆弱 マルチビットエラー
 - 保存媒体と保存環境の課題





内閣府の取り組み①

歴史資料として重要な公文書等の適切な保存・利用のための研究会

- 2003年4月「歴史資料として重要な公文書等の適切な保存・利用のための研究会」が内閣府大臣官房長の研究会として発足
- 2003年7月同研究会中間報告発表
 - ITの進展とともに、文書作成が電子化され、ネットワーク環境の下におかれる中で、多くの公文書館のデジタル化対応は遅れている。…国立公文書館を広く国民が利用できる施設とするため、一層の利便性の向上を目指して、所蔵公文書等のデジタル化に早期に着手し、インターネットで提供するデジタルアーカイブ化を早急に進める必要があり、情報技術を活用した利用者サービスの改善を早期に図る必要がある。
- 2003年12月同海外実態調査報告
 - 電子記録についても、近年の電子政府推進の動向にかんがみ、従来の政府の現用記録管理システムを見直し、記録の作成段階から将来の廃棄・移管を見据えた管理を行うための新しい取り組みが見られた。



内閣府の取り組み②

公文書の適切な管理、保存及び利用に関する懇談会

- 2003年12月内閣官房長官主催の「公文書の適切な管理、保存及び利用に関する懇談会」設置
- 2004年6月「公文書等の適切な管理、保存及び利用のための体制整備について－未来に残す歴史的文書・アーカイブズの充実に向けて－」発表
 - 情報技術及び電子政府化の急激な進展への対応
 - 電子媒体の移管・保存方法等の検討
 - 各府省において電子化された文書を原本とする公文書等が急速に増加することが見込まれる。電子媒体である公文書等についても、紙である公文書等と同様に歴史資料として国立公文書館に移管・保存しなければならないのは当然のことであり、その仕組み、方法等について、検討する必要がある。



内閣府の取り組み③

電子媒体による公文書等の管理・移管・保存に関する研究会

- 2005年6月懇談会のもとに「電子媒体による公文書等の管理・移管・保存に関する研究会」が発足
- 2006年6月懇談会報告「中間段階における集中管理及び電子媒体による管理・移管・保存移管する報告書」に「電子媒体による公文書等の管理・移管・保存のあり方」を公表



「電子媒体による公文書等の管理・移管・保存のあり方」

電子公文書等の長期保存を考える上での9つの基本的視点

- ① 電子公文書等の増加及び内容と形式の多様化適切な管理・移管・保存等への積極的取り組み、公文書等の記録としての価値を維持しつつ、その多様性、利便性の将来への継承
- ② 電子公文書等は、基本的には**電子媒体のまま保存**すべき
- ③ 電子公文書等の「**エッセンス**」を**特定**して保存すべき
- ④ 長期的な安定性を重視しつつ、現時点での可能な方法で保存
- ⑤ 各府省等における作成時から歴史資料としての保存・利用段階までの**ライフサイクル全体の管理を行う。メタデータ、フォーマット等についても作成前から適切に対応**する
- ⑥ 対象電子公文書等は、保存期間満了前の可能な限り早期に、媒体を問わず同一の基準に基づいて選定する
- ⑦ 対象電子公文書等は、保存期間満了前の可能な限り早期に、一定の集中管理下で長期保存上の措置を講じる
- ⑧ 情報公開法や個人情報保護法等の法制と電子公文書等の保存期間満了前における集中管理等との関係を整理する
- ⑨ インターネット及びイントラネットのウェブ上の電子公文書移管について国会図書館とは異なる公文書館としての要件を満たす形での移管を行う

「電子媒体による公文書等の管理・移管・保存のあり方」 電子公文書長期保存の4つの論点

- ① 電子公文書等の特性を踏まえた長期保存上の措置とその技術的課題
 - ・セキュリティー等のガイドライン策定
 - ・移送後の媒体変換のルール化
 - ・作成時からのメタデータの付与と内容
 - ・移送後のフォーマット標準化と長期保存フォーマットの策定
- ② 電子公文書等の保存対象となる種類、範囲、属性とは。原本とは何か。
 - ・保存すべき「エッセンス」の明確化
 - ・移送後の原本性確保のルール化
- ③ 電子公文書等の適切な移送時期及び移送方法について
 - ・保存期間満了後のモノとしての電子媒体による移管の開始
- ④ ウェブアーカイブのあり方について
 - ・ウェブ上の公文書の選別基準
 - ・国立国会図書館との調整

国立公文書館デジタルアーカイブの将来像

(1) 公文書をめぐる環境変化への対応

- ・受入れ資料の構成の変化、多様化
- ・行政文書の電子化

(2) EAD適用の標準化

- ・階層、概念、記述内容の整理
- ・公文書への適用

(3) 機能拡張と標準化

- ・横断検索システムの拡充
- ・電子記録、マルチメディアへの対応

(1) New Task for Archives

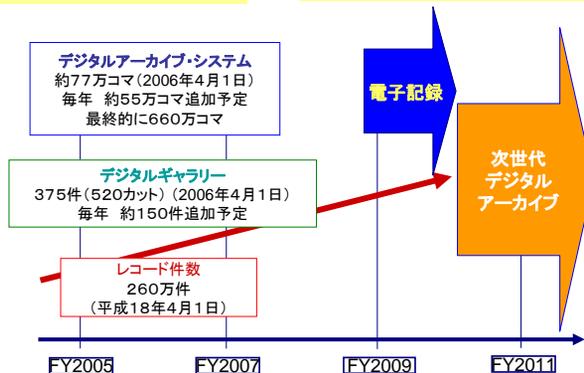
- ・diversification of material
- ・coming governmental electrical records

(2) Standardization and pervasion of EAD

- ・Continuing to research on categorize, description and tag definition, and pervasion of EAD

(3) Future Requirements

- ・expanding cross-search system (+Image data)
- ・system needs for electric record and multimedia



国立公文書館 中島康比古

内容

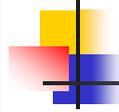
- ・ はじめに
- ・ 2005年度実施の調査研究
- ・ 電子政府化の進展状況
- ・ 電子公文書長期保存に係る技術動向と課題
 - 概要
 - メタデータ
 - 移管前の技術：フォーマット標準化、推奨
 - 移管時の技術：ファイル形式の変換
 - 移管後の技術(保管)
- ・ ウェブページの移管・保存に係る技術動向
 - 日本及び諸外国における取組み状況
 - 検討課題
- ・ 諸外国における電子公文書の管理等の取組み状況
 - アメリカ合衆国
 - カナダ
 - オーストラリア
- ・ 報告書等に関する情報



電子公文書の長期保存 — 国立公文書館の取組み —

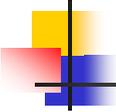
デジタルアーカイブシンポジウム
筑波大学知的コミュニティ基盤センター

2007年2月16日
国立公文書館 中島康比古



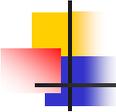
はじめに

- 中期計画(2005年度～2009年度)
電子媒体の公文書等の効率的な管理・保存に向け最適な保存媒体と管理方策等についての検討を平成17年度より行い、平成18年度を目途に結論を得る。
- 2005年度: 電子公文書の管理・移管・保存等について調査研究を実施し、報告書を作成
- 2006年度: 長期保存上の措置の対象となる電子公文書等の作成時又は作成前における選定を可能にする評価選別実施のための実証的調査研究に着手



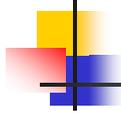
2005年度実施の調査研究

- 電子政府化の進展状況
- 電子公文書の管理・移管・保存・公開に係る技術動向と課題
- ウェブページの移管・保存に係る技術動向
- 諸外国における電子公文書の管理等の取組み状況
- 国際公文書館会議(ICA)報告書の翻訳・公開



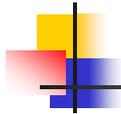
1. 電子政府化の進展状況

- 行政機関情報公開法第22条は、行政文書の適正管理、行政文書の管理に関する定め等について規定
- 行政機関は、行政文書管理規則等を定め、行政文書ファイル管理簿を作成
- 電子文書に関しては、「原本性」の確保、原本の修正、紙媒体と電子媒体の併存、電子署名等について、管理規則等で規定する省庁が存在



1. 電子政府化の進展状況

- 霞が関WAN: 政府内専用ネットワーク
電子メール、バーチャルフォーラム、電子文書交換、共通情報検索等
- 総合的な文書管理システムにより、電子文書本体と紙文書の記述情報を管理している省もあり
- 電子決裁、電子申請システムの導入・運用は、省によってバラツキあり
- 電子メールの取扱いについては、文書管理規則等に特段の規定なし



1. 電子政府化の進展状況

- 行政事務の電子化: 94.5% (2002年度)
(電子化対象57事務(「一部電子化」を含む。))
- 媒体別情報保存割合: 電子媒体保存が95.5% (2002年度)
- 電子公文書の移管・長期保存等から見る場合、ファイル管理簿の記述用語、ファイルの構成・フォーマット、紙文書との併存等に課題

2. 電子公文書長期保存に係る技術動向と課題

① 概要

- 電子公文書等の「原本性」を確保して長期保存する必要
- 「原本性」を損なうリスク要因
記録媒体の劣化、ハードウェアの陳腐化、ソフトウェアの入手困難、OS環境の喪失、改ざん、既存技術の陳腐化等
- 調査項目：メタデータ、移管前の技術、移管時の技術、移管後の技術（保管・利用）等

2. 電子公文書長期保存に係る技術動向と課題

② メタデータ

- ICA報告書16による類型化
 - 記録管理(recordkeeping)メタデータ
記録そのものの中から、または記録を作成した組織の中で生じたもの。(作成者、作成日、タイトル、機密度、キーワード等)
 - アーカイバル・メタデータ
記録が最初に作成された後にその記録を管理しやすくするように追加したもの(記録の再確認日、作成元の組織名)
 - 技術的メタデータ
記録の理解や処理に必要なもの(ファイル形式、バージョン、作成OS等)

2. 電子公文書長期保存に係る技術動向と課題

③ 移管前の技術:フォーマット標準化

- ICA報告書が推奨する保存フォーマットの要件
 - 全ての情報のまたは元の記録の中で有意義とみなされる情報相互の関係を表現できる。
 - 国際標準、国内標準、または広く一般に利用できる標準で規定されている。
 - 長期にわたる利用可能性や、幅広い適用が証明されている。
 - アクセスの目的に直接利用可能であるか、利用可能なフォーマットに変換可能である。
 - 特定のソフトウェア環境、ハードウェア環境に依存しない。
 - 元のフォーマットから保存用のフォーマットに自動的に変換できる。必要に応じて、変換時の問題やエラーを自動で検出・報告する機能がある。
 - (必要に応じて)保存用のフォーマットから元の記録作成システムや現在の記録作成システムで使用しているフォーマットへ自動的に変換できる。

2. 電子公文書長期保存に係る技術動向と課題

③ 移管前の技術:フォーマットの推奨

カナダ連邦政府の取組

種類	推薦	容認
ドキュメント	XML、HTML、SGML	.txt、.doc、PDF
オーディオ	WAV、AIFF	MPEG-1、RealAudio
画像	ITU-T、PNG、TIFF	GIF、JPEG、JFIF
動画	MPEG-2	AVI、MPEG-4、MOV

2. 電子公文書長期保存に係る技術動向と課題

④ 移管時の技術:ファイル形式の変換

- オーストラリア国立公文書館が開発したXena
 - 移管された電子記録を長期保存用のXMLフォーマットに変換するソフトウェア(ハードウェアやOSに依存しない。)
 - MicrosoftのWord・Excel・PowerPoint、OpenOffice.org Writer、RTF、CSV、Relational Database、JPEG、GIF、TIFF、PNG、BMP、HTML、Plaintext、PDF、XML等に対応
 - Xenaの利用には、JREとOpen Officeを別途導入する必要
 - Xenaは、変換対象ファイルを読み込んで、オーストラリア国立公文書館が定義した、長期保存用XMLフォーマットに変換
 - Xenaをテスト
 - Word、Excel、PowerPointのファイル(日本語・英語)は、変換できた。
 - JPEGは、変換できた。
 - PDFとPlaintextは、英語は変換できたが、日本語は変換できなかった。
 - HTMLは、PCがフリーズした(テスト環境に起因する可能性あり)。
 - ビューフは、日本語を表示できない。

2. 電子公文書長期保存に係る技術動向と課題

⑤ 移管後の技術(保管)

- 媒体寿命
 - マイクロフィルム等の非電子媒体は、電子媒体より媒体寿命が長いが、電子媒体から非電子媒体への変換により、メタデータ等が失われるおそれがある。
- バックアップ
- セキュリティ対策
- 電子認証、電子署名、タイムスタンプ等
 - 移管対象となる電子公文書は、電子認証等の有効性証明可能期間よりもはるかに長期間保存することとなることから、今後も技術動向を注視する必要がある。

3. ウェブページの移管・保存に係る技術動向

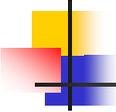
① 日本及び諸外国における取組み状況

- ウェブページ: 公開ウェブ、イントラネット上のウェブ
- 日本: 国立国会図書館インターネット資源選択的蓄積実験事業(WARP)
- アメリカ: 議会図書館MINERVA
インターネットアーカイブ
- オーストラリア: 国立図書館PANDORA
- オーストラリア国立公文書館: イン트라ネット上の記録、ウェブ上の事務処理に係る記録を管理・移管の対象に。

3. ウェブページの移管・保存に係る技術動向

② 検討課題

- 収集方法
 - 選択収集
 - バルク収集
- 収集対象
- 収集頻度
- 安定した識別子
- メタデータの標準化
- 大容量保存と利用時の利便性の両立
- 長期にわたる再現性の確保



5. 諸外国の取組み

① アメリカ合衆国

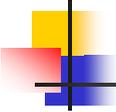
- 「電子政府戦略」(2001年)における24のイニシアチブの中の一つに、国立公文書記録管理局(NARA)が担当する「電子記録管理」がある。
- 「電子記録管理」では、全政府機関に適用可能なガイダンス等を提供するほか、電子記録を移管するためのツールを開発することとされている。
- NARAは、国防総省電子記録管理アプリケーション設計基準の全政府機関への適用を承認したほか、電子記録移管のための指針を発行
- 電子記録を永久に保存し、オンライン上で利用に供することを目的に、ERAプロジェクトに着手



5. 諸外国の取組み

② カナダ

- 「ガバメント・オンライン」計画を担当する国家予算委員会事務局CIO部及び公共事業サービス省と共に、カナダ図書館公文書館(LAC)は、政府の情報マネジメントを担当し、各種の標準等を発行
- LAC、国家予算委員会、議会図書館等が組織する連邦図書館評議会は、ウェブ情報のメタデータに関するガイドを発行
- LACは、記録情報ライフサイクル・マネジメントに関するガイドライン、連邦政府内部における電子情報の相互運用性向上を目的として、文書・画像・映像・データベース等の推奨規格等を定めたガイドライン等を発行



5. 諸外国の取組み

③ オーストラリア

- 国立公文書館(NAA)は、政府機関における記録管理のためのメタデータ標準、記録記述用語集等のほか、媒体を問わず記録を適切に関するためのDIRKSマニュアル等、ガイドライン、標準、アドバイス等を発行
- NAAは、インターネット上の情報検索を容易にするため、AGLSメタデータ標準を開発
- NAAは、電子記録を長期に安定的に保存するためのフォーマットへの変換ソフトであるXENAを開発
- NAAは、イントラネット上の文書、ウェブ上での業務記録、電子メール及び公開ウェブのスナップショットを保存することとしている。



報告書等に関する情報

- 「中間段階における集中管理及び電子媒体による管理・移管・保存に関する報告書」(2006年)
<http://www8.cao.go.jp/chosei/koubun/kondankai14/houkoku.pdf>
- 「電子媒体による公文書等の適切な移管・保存・利用に向けて－調査研究報告書－」(2006年)
http://www.archives.go.jp/hourei/hourei3_12.pdf
- 「アーカイブズの観点から見る電子記録管理ガイド」(ICA報告書8) 日本語版(2006年)[原版:1997年]
http://www.archives.go.jp/hourei/ICASTUDY8_ELECTRONIC_RECORDS_JPN.pdf
- 「電子記録:アーキビストのためのワークブック」(ICA報告書16) 日本語改訂版(2006年)[原版:2005年]
http://www.archives.go.jp/hourei/ICASTUDY16_ELECTRONIC_RECORDS_JPN.pdf

学術雑誌のデジタルアーカイブと長期利用に関して
— JST 電子アーカイブ事業とその後 —

独立行政法人科学技術振興機構・文献情報部 和田光俊

内容

- ・ JST 電子アーカイブ事業について
 - 事業の背景
 - J-STAGE とは
 - ◇ J-STAGE の利用状況
 - ◇ J-STAGE 日本語 TOP 画面
 - 公開システム Journal@rchive
 - ◇ Journal@rchive での公開
 - ◇ 前身誌・カレント分とのリンク
 - ◇ Journal@rchive と J-STAGE での横断検索
 - 電子アーカイブ事業の進め方
 - ◇ 電子アーカイブの作成方法
 - ◇ 自動スキャンロボット
 - ◇ 平成 17 年度電子アーカイブ事業の実績
 - 平成 17 年度対象誌
 - 平成 17 年度電子アーカイブ対象誌 74 誌の創刊からの年数分布
 - ◇ 平成 18 年度電子アーカイブ事業の状況
 - 平成 18 年度対象誌
 - 平成 18 年度電子アーカイブ対象誌 65 誌の創刊からの年数分布
 - ◇ Journal@rchive 掲載論文数と月間アクセス数の推移
 - Journal@rchive 掲載記事への国別アクセス状況
- ・ デジタル学術雑誌のアーカイビング 長期保存、長期利用のための課題
 - デジタル学術雑誌のアーカイビング
 - デジタル学術雑誌 コンテンツ長期保存のための課題
 - デジタル学術雑誌 長期利用のための課題
- ・ JST 電子アーカイブ事業について

学術雑誌のデジタルアーカイブと 長期利用に関して

— JST電子アーカイブ事業とその後 —



独立行政法人 科学技術振興機構
文献情報部 和田光俊

2007年 2月16日
「デジタルアーカイブとその長期利用に関する研究会」

発表内容

- JST電子アーカイブ事業について
- デジタル学術雑誌のアーカイビング
長期保存、長期利用のための課題

JST電子アーカイブ事業

- 国内学協会の学術雑誌の国際発信力の強化と重要な知的財産の保存等を目的とし、特に重要な学術雑誌について過去の紙媒体の論文に遡って電子化(電子アーカイブ)
- 平成17年度から開始 (5年間で500誌を目標)
- 日本の主要な学術雑誌を創刊号から電子化し、Journal@rchiveサイトにて全文公開
- 電子アーカイブ対象誌選定委員会(委員長:黒川清前日本学術会議会長)にて対象誌を選定

事業の背景

- JSTでは、平成10年度より国内の学協会誌の電子出版を支援するためにJ-STAGEを開発・運用
- 近年、海外では学術雑誌の創刊号からの電子化が急速に進展
- 過去に遡って電子化されていない学術雑誌は国際競争力を著しく失う恐れ
- 我が国の学協会誌の単独での電子化は困難
- 学協会からの強い要望を受け、JSTは平成17年度から予算を得て、電子アーカイブ事業を開始

J-STAGE とは

科学技術情報発信・流通総合システム

Japan Science and Technology Information Aggregator,
Electronic

JSTが運営する、学協会による電子ジャーナル出版
のための共同利用センター

- 学協会のシステム利用は無料
- 登載用データは学協会が準備
- JSTリンクセンターにより、他の電子ジャーナルサイト
や文献データベースとの相互リンクを実現
- 投稿・審査から編集、公開まで全プロセスをサポート

J-STAGEの利用状況

・J-STAGEを利用している学協会・資料タイトル数

	公開誌数	利用申請数	利用学協会数
ジャーナル	358	428	347
報告書	8	8	8
予稿集	109	161	129
合計	475	597	428

(平成19年1月末現在)

- ・登載記事数: 23.5万件
- ・月間総アクセス数: 約250万件
- ・月間PDFダウンロード数: 約60万件
- ・登載記事へのアクセスの約6割は海外から

J-STAGE 日本語TOP画面



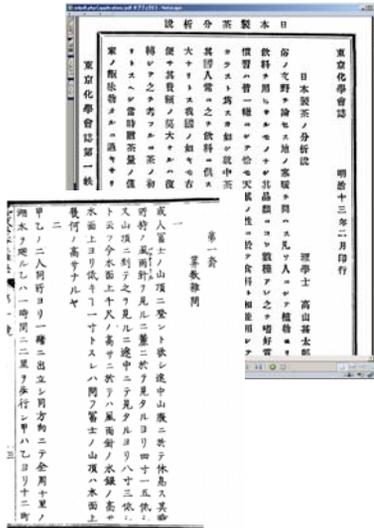
<http://www.jstage.jst.go.jp/ja/>

Journal@rchive

公開システム Journal@rchive

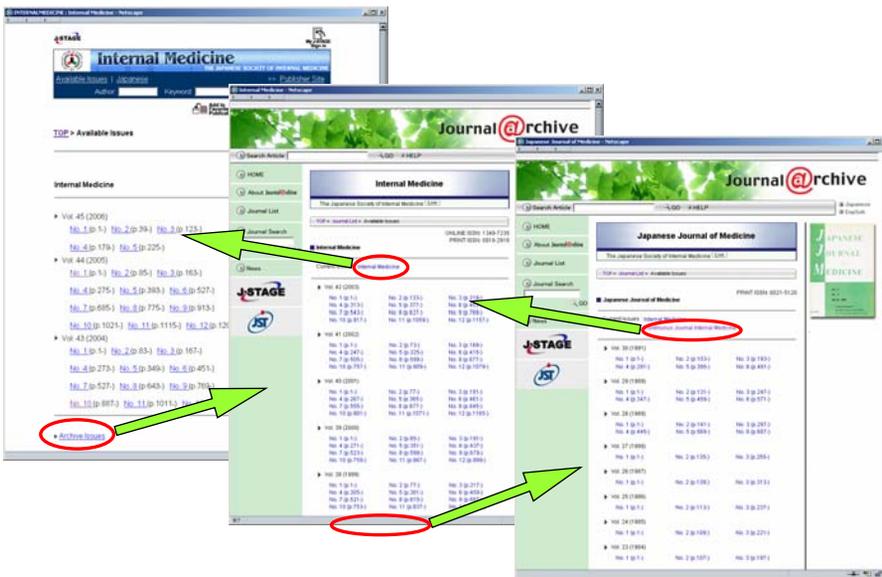
- J-STAGEのアーカイブサイト
- J-STAGEで培われた電子ジャーナル技術を電子アーカイブの大容量データに適用
- J-STAGEと同様の引用リンク、被引用リンクや全文検索を実現
- J-STAGEに収録されている最近発行された論文とも相互にリンク。横断検索も可能に

Journal@rchive での公開



<http://www.journalarchive.jst.go.jp/japanese/>

前身誌・カレント分とのリンク



Journal@rchiveとJ-STAGEでの横断検索

The image shows a screenshot of the Journal@rchive website. At the top, the search bar contains the name 'Hideo Tanihara'. Below the search bar, the search results are displayed. A red circle highlights the search bar. A red box highlights a search result entry: 'Chapter 2 Post-Newtonian Approximation to Foundation and Applications'. Green arrows point from the search bar to the result and from the result to the full article page on the left.

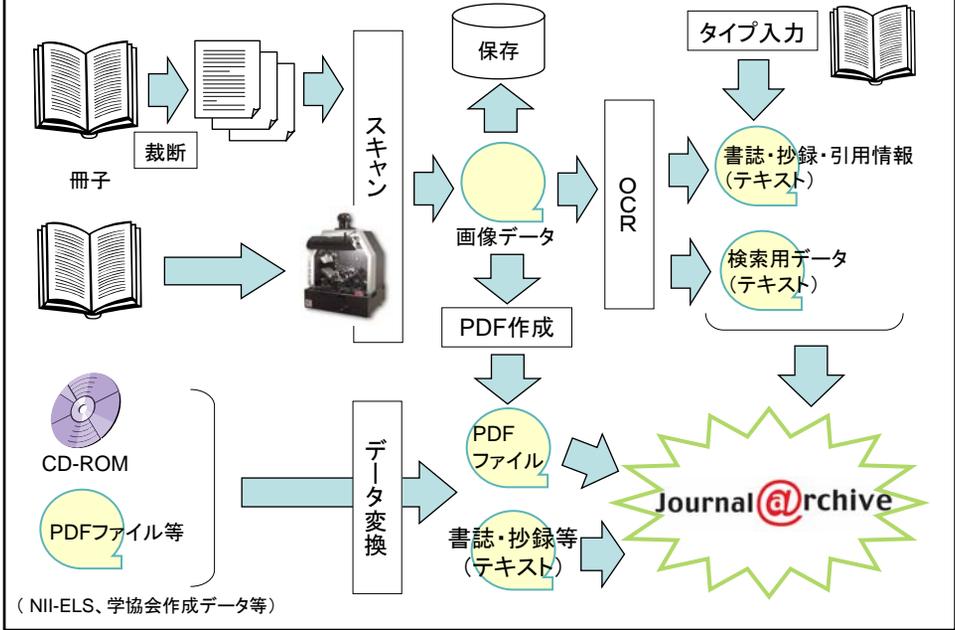
10

電子アーカイブ事業の進め方

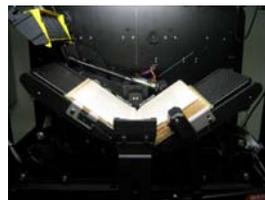
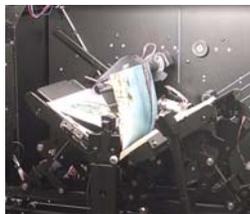
- アーカイブ化はジャーナル単位に実施。
可能な限り創刊号から。
- 元資料の準備および著作権処理については各学協会が行い、JSTは学協会から許諾を得て電子化及び公開までの作業について責任をもつ。
- 作成されたアーカイブはJournal@rchiveで公開。
- 検索や引用リンクをできるだけ可能にする。
- 作成データは、学協会の希望に応じて無償提供可。
- 既存の電子化済データがあれば、できるだけ活用。

11

電子アーカイブの作成方法



自動スキャンロボット



APT 1200 (米国Kirtas社製)

平成17年度電子アーカイブ事業の実績

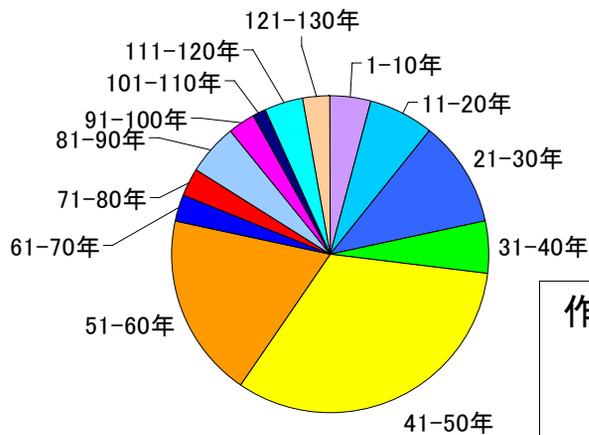
- 6月 対象候補誌基礎調査実施
延べ1,890団体に依頼
484学協会から550誌について回答
- 7～9月 対象誌選定委員会にて
平成17年度アーカイブ対象誌74誌を決定
- 8～10月 対象誌詳細調査実施
- 11月～ 順次、データ作成・登載作業を実施
- 平成18年3月27日 Journal@rchive公開開始
(当初52誌の約3万件)
データは順次追加公開 (現在53誌の約9.3万件)

平成17年度対象誌

- 全74誌
- 原則として、
カレントが英文誌のもの
(次年度以降の実験として
和文誌1誌を含む)
- J-STAGE既公開誌は48誌

人文社会	7
理学	14
生物学	8
工学	17
農学	9
医学薬学	19

平成17年度電子アーカイブ対象誌 74誌の創刊からの年数分布



作成予定データ量
延べ3,500年
37万論文
295万ページ

平成18年度電子アーカイブ事業の状況

(平成17年度電子アーカイブ対象誌)

- 4月～ 引き続き、データ作成・公開

(平成18年度電子アーカイブ対象誌)

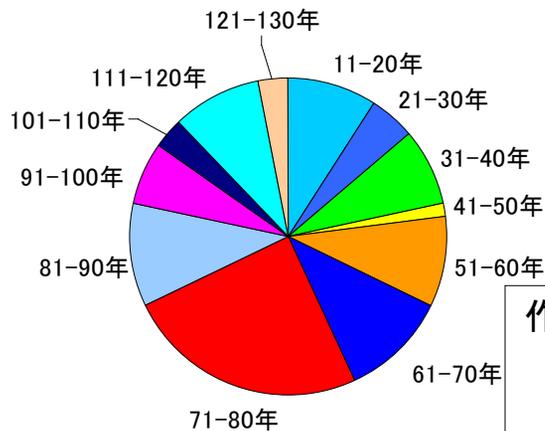
- 4～5月 電子アーカイブ対象候補誌基礎調査実施
延べ1,604団体に依頼
549学協会から612誌について回答
- 5～7月 電子アーカイブ対象誌選定委員会にて
平成18年度アーカイブ対象誌65誌を選定
- 7～9月 電子アーカイブ対象誌詳細調査実施
- 9月～ 順次、データ作成・公開

平成18年度対象誌

- 全65誌
- うち、英文誌18誌、和文誌22誌、和欧混載誌25誌
- 歴史の古いものを優先
- J-STAGE既公開誌は37誌

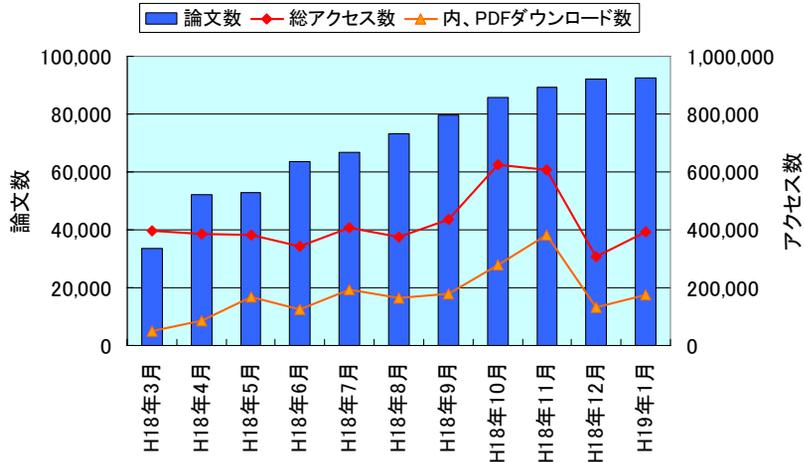
人文社会	5
理学	12
生物学	6
工学	22
農学	8
医学薬学	12

平成18年度電子アーカイブ対象誌 65誌の創刊からの年数分布



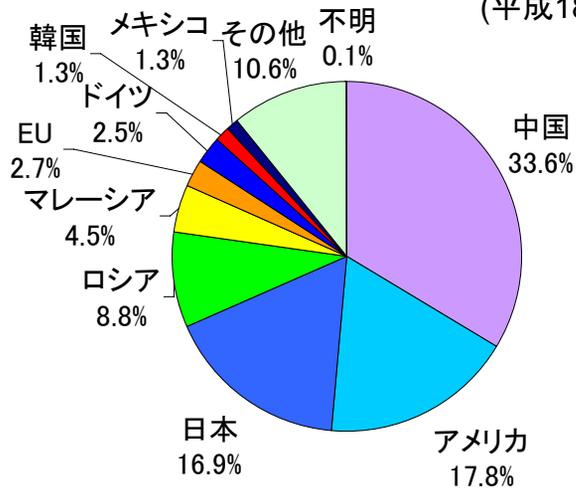
作成予定データ量
延べ4,300年
29万論文
340万ページ

Journal@rchive 掲載論文数と 月間アクセス数の推移



20

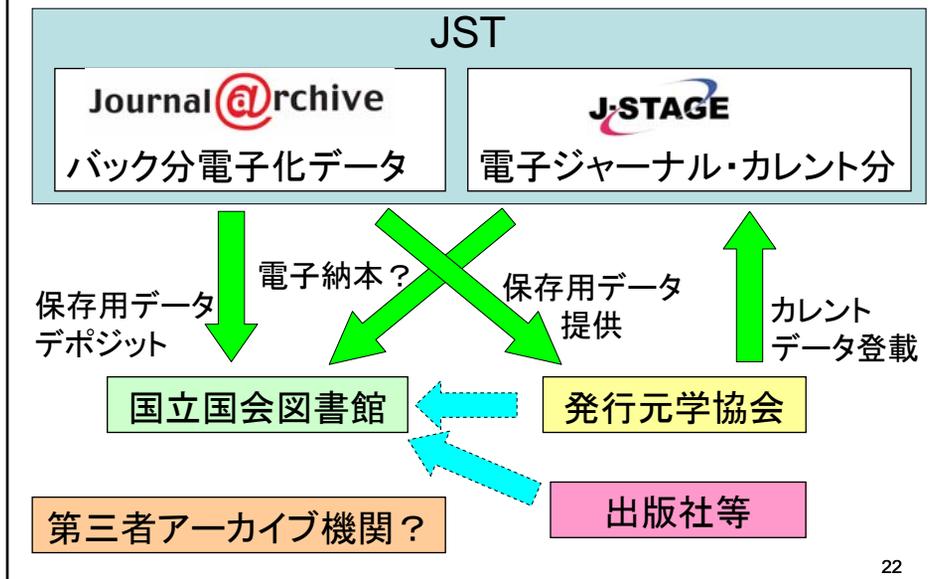
Journal@rchive 掲載記事への 国別アクセス状況 (平成18年3月～12月)



延べ141ヶ国

21

デジタル学術雑誌のアーカイビング



デジタル学術雑誌コンテンツ 長期保存のための課題

- 保存対象
 - 本文データ(画像、テキスト)、書誌データ
 - 引用リンク情報?
- 保存形式
 - 画像ファイル、テキストファイル、PDFファイル?
- 保存場所
 - 災害等を考慮、複数箇所
 - 冊子体は不要になるのか?

デジタル学術雑誌 長期利用のための課題

- 安定した組織・体制・財源
- 公開方法（購読者へのアクセス権の確保）
- 技術変化に伴うデータのコンバート
- バックアップ、ミラー・サイト

JST電子アーカイブ事業について

- 公開トップページ:
[英語]
<http://www.journalarchive.jst.go.jp/>
[日本語]
<http://www.journalarchive.jst.go.jp/japanese/>
- お問い合わせ先:
jarchive@jst.go.jp

文化財のデジタルアーカイブと長期保存に関する問題

筑波大学・図書館情報メディア研究科 西岡貞一

内容

- バーチャルリアリティの文化遺産デジタルアーカイブへの応用
 - バーチャルリアリティとデジタルアーカイブ
- **Virtual Reality**の応用
- 映画、テレビに続く第三のスクリーン
- バーチャルリアリティを実現するための道具
 - 没入感を提供するVRシアター
 - インタフェース
 - 没入感(Immersive Projection System)
- バーチャルリアリティによる文化の体験
 - 洛中洛外図の中にフライスルーし、近世京都の暮
- バーチャルリアリティのデジタルアーカイブへの応用
 - 古代マヤ文明展
 - マヤの神殿群を俯瞰し、Flythroughする
 - マルガリータ神殿の壁面レリーフ
 - 測量、復元図
 - 1200年前のマヤの色彩を再現
 - 写真撮影
 - マヤ人の復元(文化人類学)
- バーチャルリアリティによる文化の可視化(古代マヤ文明展)
 - ナスカ展
 - 大英博物館 ミイラと古代エジプト展
 - 国宝鑑真和上展
- 先導的アーカイブ映像制作支援事業(平成10年度経済産業省補正予算)
- ムーアの法則
- 高精細化 (Hi-Definition)
- 映像システムの表示性能の推移
- バーチャルリアリティのデジタルアーカイブへの応用

文化財のデジタルアーカイブと 長期保存に関する問題

2007.2.16

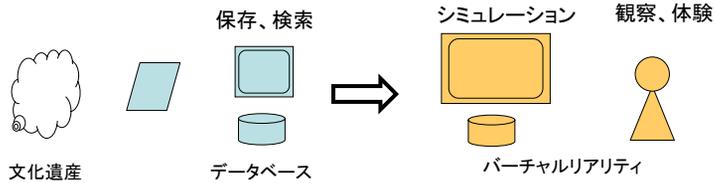
図書館情報メディア研究科
西岡 貞一

バーチャルリアリティの文化遺産デジタルアーカイブへの応用

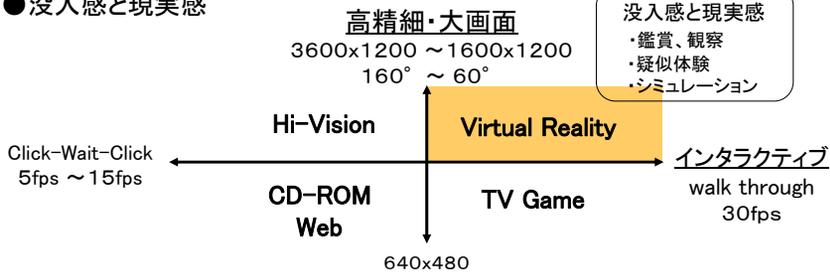
- Digital Archives
デジタル技術を活用した文化遺産の保護と公開
- 書籍やビデオでは伝えられなかった知識の伝達
- 画像データベースの検索では得られない新しい感動
- リアルタイムグラフィックスを応用した体験の共有

バーチャルリアリティとデジタルアーカイブ

●保存、検索から観察、体験へ



●没入感と現実感



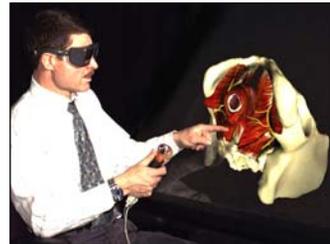
Virtual Realityの応用



- フライトシミュレータ
航空機の操縦訓練
スクリーン上に映し出された景色を見ながら空中戦の訓練を行う。



- 船舶シミュレータ
航海の訓練
海の景色を映した大型スクリーンを見ることで、仮想的な航海を経験する。



- 医療用バーチャルリアリティ

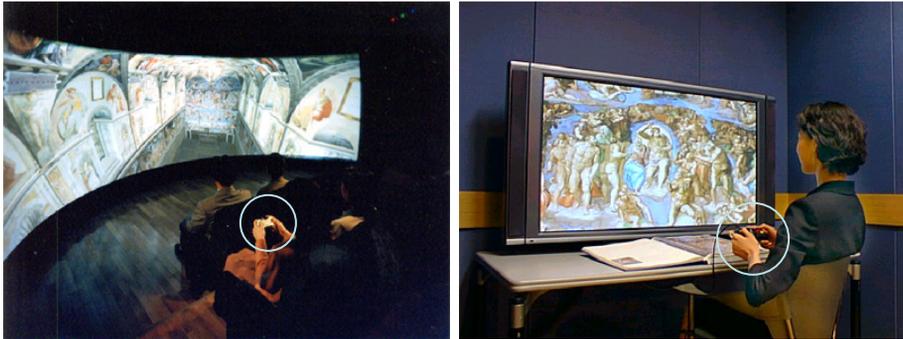


- 米国海軍で利用されるVR
パラシュート訓練機



- 映像の壁に囲まれた空間に再現された都市空間

映画、テレビに続く第三のスクリーン



- **Click & Wait**の退屈な操作とは異なり、自分が歩いている気分になれるインターフェース。
- 書物では伝えることの難しい空間的な広がり、身体的な感動を伝える。

バーチャルリアリティを実現するための道具

- ・ 鑑賞者の働きかけに映像が小気味よく応える
⇒ 高精細(1280x1024画素x3面)な画像を滑らか(毎秒30コマ)に表示するグラフィクス
- ・ 鑑賞者の意思を伝える直感的なインターフェース
⇒ 手の中に入るゲームパッド

没入感を提供するVRシアター

- Horizontal FOV: 150deg
- Vertical FOV: 40deg
- radius: 3m
- Audio System: 4ch digital



インターフェース



フライトシミュレータ

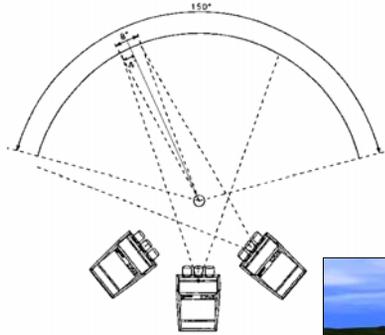


ゲームパッド

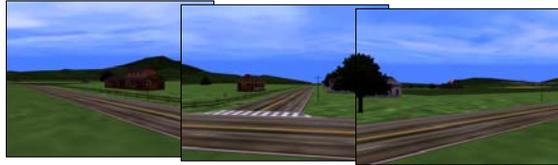


- 手になじむ
- モーションモデル

没入感 (Immersive Projection System)



- 一つのスクリーン上で三枚の映像をつなぎ合わせる
- 映像間の境界線の平滑化
- 曲面への投影によって生じる歪みの補正



バーチャルリアリティによる文化の体験



3200x1000 毎秒30コマ



4mx12m 左右150度の視域

- 映画並みの画質の映像をビデオゲームの様に自分の意思で操る。
- 没入感と対話性により自分がその空間を歩いている様な気分になる。
- 文化遺産を身体的に鑑賞する。



直感的なインターフェース

洛中洛外図の中にフライスルーし、近世京都の暮らしを体験



洛中洛外図を元に400年前の近世京都の街を復元。文章や絵では想像できない人々の暮らしを生き生きと再現し、等身大のスケールで体験する。
人物表現の実験として洛中洛外図に描かれている人物を一つ一つ切り抜きビルボードとして配した。

(国際日本文化研究センターとの共同研究、1998年)

バーチャルリアリティのデジタルアーカイブへの応用

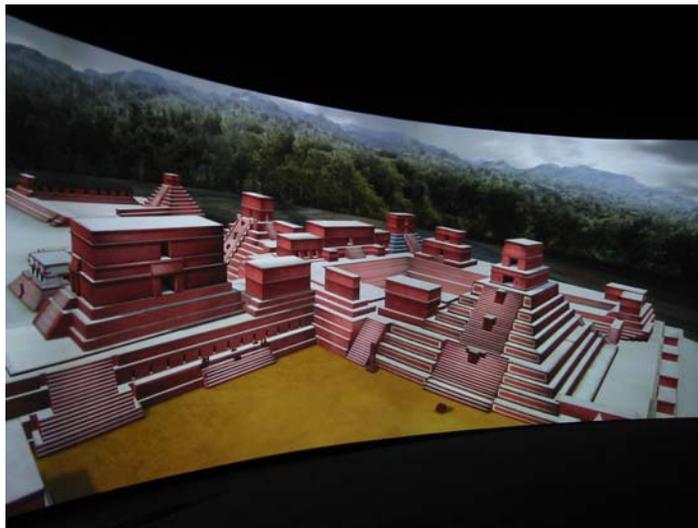
- リアルタイムに三次元空間を認識する
 - － 身体的な体験、空間的な理解
 - － 歴史的建造物の前に立ったときの感動
 - － ミュージアムに行った時の楽しさ
- インタラクティブに経験を拡大する
 - － 好きな視点から、好きなだけ鑑賞する
 - － 個別選択性
- もちろんオリジナルが一番
 - － アクセス件の民主化(非公開文化財の公開)
 - － 追体験、見たいときにいつでも

古代マヤ文明展



場所: 国立科学博物館
期間: 3月18日～5月18日
主催: 科学博物館、TBS、朝日新聞

マヤの神殿群を俯瞰し、Flythroughする



ホンジュラス人類学歴史学研究所、国立科学博物館、TBSと共同して古代マヤの世界を復元。上野の国立科学博物館にいながらにして中米ホンジュラスにワープし、1200年前の世界にタイムスリップ。
世界遺産コバン遺跡という第一級の題材を日本の自然史博物館の総本山である国立科学博物館で2ヶ月間上演。

マルガリータ神殿の壁面レリーフ



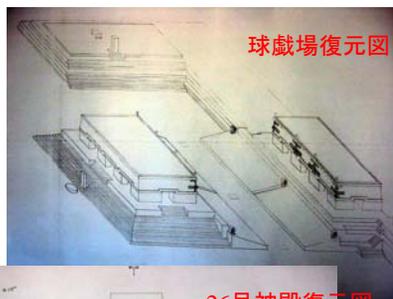
アクロポリス地下の調査風景



壁面レリーフのVR映像
文化財保護のため通常は非公開。
撮影スペースに制限があり全体を
同時に撮影するのは不可能。
30枚の分割撮影。

16号神殿の内部深くにマルガリータ神殿が埋められている。450年ごろに作られたこの神殿の入り口には宇宙を表す壮大な装飾が施されている。ケツアル鳥(右図)とコンゴウインコの紋章は初代王を表している。そのくちばしからは小さく太陽神(キニチ)が顔を出している。

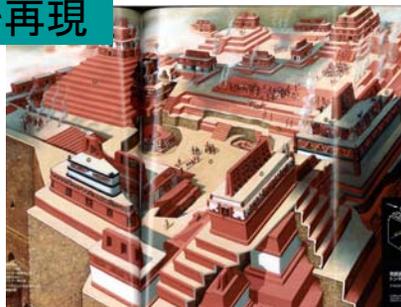
測量、復元図



1200年前のマヤの色彩を再現



アーキオ(講談社)のイラストレーション



ナショナルジオグラフィ(ハーバード大学監修)



マルガリータ棟での調査



遺跡サンプル

写真撮影

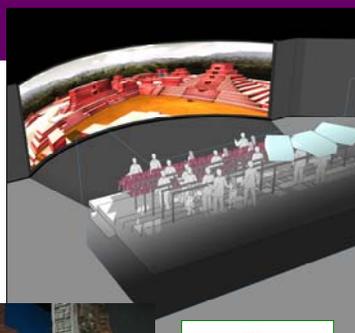


写真撮影 53枚(4x5)、440枚(6x7)、322(35mm)

マヤ人の復元(文化人類学)



バーチャルリアリティによる文化の可視化(古代マヤ文明展)



バーチャルリアリティによる文化の可視化

- 研究活動の拡大
 - コンピュータ・シミュレーション
 - 第三の研究資料
オリジナル、書物、VR
 - 空間の中での思考
検索から「ひらめき」へ
 - コンセプトの可視化
異分野の相互理解
- 公開・展示機能の拡大
 - 鑑賞者の違いに合わせた説明
子供、学生、研究者、VIP、外国人
 - 解説者の数だけストーリーがある
 - 最新の情報の提供、ソフトの更新頻度
 - 二次利用
DVD、ビデオ、バーチャルスタジオ



大英博物館
ミイラと
Mummy: the inside story
古代エジプト展



2006年10月7日～
2007年2月18日

国立科学博物館(上野)

日時指定券 一般・学生 1200
当日券 一般・学生 1500

カルトナーージュ製ミイラマスク 古代エジプトでは、生前の姿をマスクとしてミイラにつけて埋葬した。凹凸が表現できるカルトナーージュはマスクの制作に向いており、中王国時代から作られた。簡略化した顔、頭の有翼スカラベ、襟飾りの鮮やかな色違いなどから、新しい時代のものであることが分かる。



ネスベルエンネブのカルトナーージュ棺 ルクソールで1890年代に発見されたミイラ。棺に書かれたヒエログリフから、カルナクのコンス神殿に仕える神官ネスベルエンネブのものであることが判かる。カルトナーージュ棺は亜麻布とプラスチックで作られていて、木棺の内側に納められた。第26中期になると壁面に装飾を施す墓は作られなくなり、代わりに棺自体が華やかに装飾されるようになった。



古代エジプトの研究を進める上で難題の1つは、ミイラが当時を知る上で情報の宝庫であるにも関わらず、いったんその包帯を解くと取り返しがつかないという点でした。これまでもX線などによって内部の調査が行われてきましたが、十分ではなく、ミイラは多くの箇所に包まれたままでした。そこで大英博物館は、医療分野、石油・ガス探査など幅広い分野で使われてきたコンピューターによる可視化技術(物を壊さずに内部を探査する技術)に着目し、ミイラ研究に応用しました。紀元前800年のエジプトの神官・ネスベルエンネブのミイラを1mmごとにCTスキャンにかけて3D映像化し、ミイラの身体の中を探検するバーチャルツアーが実現、棺の中に隠された秘密を解き明かすことに成功しました。



国宝 鑑真和上展 2001.1.27～3.26 主催 TBS 東京都美術館 朝日新聞



先導的アーカイブ映像制作支援事業 (平成10年度経済産業省補正予算)

1. 21世紀に伝えたい日本の美術「狩野派」400年の歴史 (株)PHP研究所 [MUSE-LD](#)
2. 加賀前田家の工芸遺産 ～百工比照～ (株)北陸スタッフ [DVD-ROM](#)
3. 電子セット「幻の江戸城」(株)NHKエンタープライズ21 [DAT](#)、[UNIHI](#)
4. 日光東照宮 ～美の構造～ (株)バスプラスワン [DAT](#)、[UNIHI](#)
5. 「国宝源氏物語絵巻」高精細デジタルアーカイブの試作 (株)日立製作所 [5-MO](#)
6. 甞る日本の美 ～掛軸の修復～ (株)電通テック [UNIHI](#)
7. 西芳寺庭園 築庭当初の再現と現状 名勝庭園デジタルアーカイブ (株)電通 [MUSE-LD](#)
8. 20世紀の遺産 ～デジタル映像で見る昭和史～ (株)ナトコ [DVD-VIDEO](#)
9. 映画美術監督のスケッチによる戦後日本のまちなみアーカイブ (株)黒澤プロダクション
DVD-RAM、Photo-CD
10. 歴史・民俗資料に基づく時代観察 ～江戸時代の北国紀行～ 日本ビクター(株) [DVD-RAM](#)、[DVD-VIDEO](#)、[MO](#)
11. 東京国立博物館 法隆寺献納宝物デジタルアーカイブ (財)NHKエンジニアリングサービス
[UNIHI](#)
12. 唐招提寺 一鑑真と東山魁夷芸術ー 凸版印刷(株) [CD-ROM](#)(データ格納)
13. 萬狂言の世界 七世野村万蔵の秘曲 (株)ベンチャーユニオン [DVD-ROM](#)
14. バイオイメージング(生命への探索)アーカイブ (株)桜映画社 [MUSE-LD](#)
15. デジタルアーカイブ「西表の動物達」(株)沖縄映像センター [UNIHI](#)
16. 衰退滅亡の危機にある日本古来の人形芸術の収録保存 (株)光映 [UNIHI](#)
17. 市民が撮った震災映像アーカイブ (財)阪神・淡路大震災記念協会 [MUSE-LD](#)、[DVD-ROM](#)
18. 浮世絵等版本からのデジタル復刻の試み (株)日立製作所 [5'-MO](#)
19. 超高精細画像ブラウジングシステムのコンテンツ制作 ー平等院の国宝仏像ー (株)PFU
[DVD-RAM](#)

ムーアの法則②



Color Management System
映像(RGB)と印刷(CMY)の一致

年代	1990	2006
カメラ	ハイビジョンカメラ	デジカメ
コンピュータ	ミニコン、WS	PC
記録装置	1/2inch 磁気テープ	USB memory
画像圧縮	JPEG 専用hard	CPU
ネットワーク	ISDN64 (0.064M)	Internet 3～5M
価格	1億円	20万円

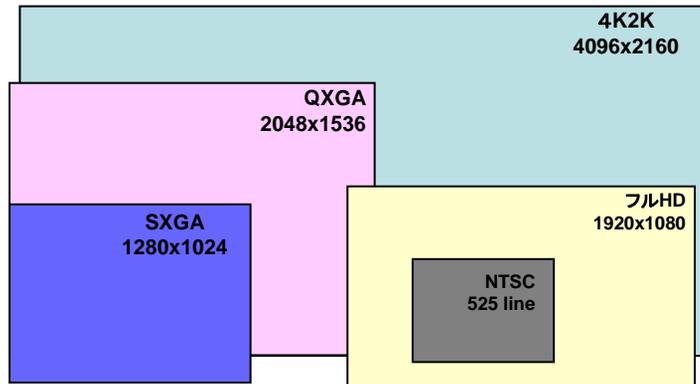


Mini-computer
& 1/2inch MT



画像伝送ユニット
JPEG圧縮、ISDNターミナル

高精細化① (Hi-Definition)



35mmオリジナルネガの画素数 3840x2070 800万画素
 上映ポジの画素数 1980x1080 約200万画素 (HD規格相当)

NTSC画面鑑賞推奨距離 6H (視力1.0で1.15mm画素を識別)
 HDTV画面鑑賞推奨距離 3H (100インチの画面を3.8mの距離から鑑賞)

映像システムの表示性能の推移

	名称	画素数	総画素数	アスペクト比	登場時期
1	VGA	640 x 480	307,200	4:3	90年代中期
2	XGA	1024 x 768	786,432	4:3	90年代後期
3	HDTV 720p	1280 x 720	921,600	16:9	2000年
4	SXGA	1280 x 1024	1,310,720	5:4	90年代後期
5	フルHD (1080i 1080p)	1920 x 1080	2,073,600	16:9	2006年
6	2K (デジタルシネマ)	2048 x 1080	2,211,840	2:1	2004年
7	4M (デジカメ)	2304 x 1728	3,981,312	4:3	2003年
8	4k2k (デジタルシネマ)	4096 x 2160	8,847,360	2:1	2006年
9	Super Hi-Vision (4k8k)	7680 x 4320	33,177,600	2:1	2015年頃

デジカメ、ビデオカメラの高精細化にともないPCの表示画素数は200万画素を超えつつある。BlueRay-DiscやHD-DVD等のフルHD対応の映像記録システムの本格普及が始まる。デジタルシネマでは2kから4kへの移行が始まる。

バーチャルリアリティのデジタルアーカイブへの応用

- リアルタイムに三次元空間を認識する
 - 身体的な体験、空間的な理解
 - 歴史的建造物の前に立ったときの感動
 - ミュージアムに行った時の楽しさ
- インタラクティブに経験を拡大する
 - 好きな視点から、好きなだけ鑑賞する
 - 個別選択性
- もちろんオリジナルが一番
 - アクセス件の民主化(非公開文化財の公開)
 - 追体験、見たいときにいつでも

デジタルアーカイブと法

筑波大学・図書館情報メディア研究科 新保史生

内容

1. 情報の保存(電子化)
 - 1.1 公文書の電子化
 - 1.1.1 公文書の電子化推進の背景
 - 1.1.2 電子政府構築の原則(8原則)
 - 1.1.3 基盤となる法制度
 - 1.1.4 検討状況
 - 1.2 文書・私文書の電子化
 - 1.2.1 文書の電子化推進の背景
 - 1.2.2 文書の電子化に関する法律
2. 知的財産の保護
 - 2.1 知的財産・知的財産権とは
 - 2.2 デジタルアーカイブと複製権
 - 2.2.1 著作権法31条の定める図書館(著作権法施行令第1条の3)
 - 2.2.2 複製の主体、対象及び態様
 - 2.2.3 図書館資料保存のための複製
3. 個人情報・プライバシーの保護
 - 3.1 個人情報保護制度
 - 3.2 プライバシーの保護
4. 時の経過
 - 4.1 現用文書→非現用文書
 - 4.2 著作権
 - 4.3 プライバシー
5. 情報のセキュリティの確保

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

デジタルアーカイブと法

筑波大学大学院 図書館情報メディア研究科 新保 史生

デジタルアーカイブの法的側面における論点

1. 情報の保存（電子化）
2. 知的財産の保護
3. 個人情報・プライバシーの保護
4. 時の経過
5. 情報のセキュリティの確保

1. 情報の保存

1.1. 公文書の電子化

1.1.1 公文書の電子化推進の背景

- 「電子政府構築計画」（平成 15 年 7 月 17 日策定、平成 16 年 6 月 14 日改定）
 - （1）年間申請件数が 10 万件以上の手続を重点に手続の簡素化・合理化の徹底、業務処理の短縮化の推進
 - （2）電子政府の総合窓口（e-Gov）を活用したワンストップサービスの推進
 - （3）IT 導入による政府全体の業務・システムの最適化の戦略的・横断的推進

1.1.2 電子政府構築の原則（8原則）

- ① 国民にとって使いやすく分かりやすい、高度な行政サービスの提供
行政機関ごとの縦割りサービスを排除し、国民が利用したい時間・場所において簡単に行政サービスが受けられる機会を確保すること。
- ② 政策に関する透明性の確保、説明義務の履行及び国民参加の拡大
電子政府の総合窓口<www.e-gov.go.jp>などを通じ、政策に関する多様な情報提供を徹底するとともに、政策立案過程における国民の意見提起の機会を最大限確保すること。
- ③ ユニバーサル・デザイン（誰もが使いやすい設計）の確保
高齢者、障害者の使いやすさにも十分に配慮されたシステム（音声による読み上げ機能に配慮した情報内容の整備等）の導入に努めること。

④ 業務効率の徹底的追求

業務や制度、システムの抜本的な見直しを行い、行政運営の簡素化、業務 効率の向上を徹底的に追求すること

⑤ 民間活力の活用

情報通信技術の専門性と変化の早さにかんがみ、業務・システムの最適化 に当たり、民間の専門家の活用や民間への委託に努めること。

⑥ 情報システムの安全性・信頼性の確保と個人情報保護

情報システムについて、常に最高水準の安全性、信頼性を確保するとともに、 I T 社会の基盤である個人情報保護法制の早急な整備と厳格な運用を図ること。

⑦ 国の行政機関以外の機関との連携及び国際連携の確保

独立行政法人、地方公共団体、国会、裁判所等国の行政機関以外の機関との連携協力により、国民の利便性・サービスの向上等を総合的・一体的に推進し、諸外国とも十分な連携を図りつつ、システム構築等にあたること。

⑧ 活力ある社会形成への配慮

電子政府を推進することによって、電子商取引をはじめとする国民生活や企業活動における I T 利用促進の触媒的機能を十分に果たすこと。

1.1.3 基盤となる法制度

○「行政手続オンライン化関係三法」

①行政手続等における情報通信の技術の利用に関する法律(平成 14 年法律第 151 号)

法令に根拠を有する国民等と行政機関との間の申請・届出等の行政手続(約 5 2, 0 0 0 手続)について、書面によることに加え、オンラインでも可能とするための法を新たに整備し、行政手続のオンライン化により、国民の利便性の向上と、行政運営の簡素化・効率化を図ることを目的とした法律。

②行政手続等における情報通信の技術の利用に関する法律の施行に伴う関係法律の整備等に関する法律(平成 14 年法律第 152 号)

行政手続オンライン化法の規定のみでは手当てが完全ではないもの、例外を定める必要があるものについて、7 1 の個別法律の改正を束ね一つの法律としてとりまとめたもの。

③電子署名に係る地方公共団体の認証業務に関する法律(平成 14 年法律第 153 号)

申請・届出等行政手続のオンライン化に資するため、第三者による情報の改ざんの防止・通信相手の確認を行う高度な個人認証サービスを全国どこに住んでいる人に対しても安い費用で提供する制度を整備するための法律。

○「情報公開法」

行政機関の保有する情報の公開に関する法律(平成 11 年 5 月 14 日法律第 42 号)

独立行政法人等の保有する情報の公開に関する法律(平成 13 年 12 月 5 日法律第 140 号)

- 「公文書館法」
公文書館法（昭和 62 年 12 月 15 日法律第 115 号）
国立公文書館法（平成 11 年 6 月 23 日法律第 79 号）

1.1.4 検討状況

- 総務省「共通課題研究会中間報告 ―電子文書の原本性確保方策を中心として―」（平成 11 年 4 月）
- (財)ニューメディア開発協会「電子文書の原本性保証ガイドライン」（平成 12 年 3 月）
- (財)ニューメディア開発協会「原本性保証に係る評価・認定制度に関する調査研究報告書」（平成 13 年 3 月）
- (財)ニューメディア開発協会「原本性保証システムガイドライン」（平成 13 年 3 月）
- 国立国会図書館「電子情報保存に係る調査研究報告書」（平成 15 年 3 月）
- 内閣府「歴史資料として重要な公文書等の適切な保存・利用のための研究会 中間取りまとめ」（平成 15 年 7 月）
- 内閣府「諸外国における公文書等の管理・保存・利用にかかる実態調査報告書」（平成 15 年 12 月）
- 国立国会図書館「電子情報の長期的保存とアクセス手段の確保のための調査報告書」（平成 16 年 3 月）
- 電子商取引推進協議会「電子文書の長期保存と見読性に関する調査報告書」（平成 16 年 3 月）
- 内閣府「公文書等の適切な管理、保存及び利用のための体制整備について―未来に残る歴史的な文書・アーカイブの充実にむけて―」（平成 16 年 6 月）
- 国立国会図書館「平成 16 年度電子情報の長期的保存とアクセス手段の確保のための調査報告書」（平成 17 年 3 月）
- 内閣府「中間段階における集中管理及び電子媒体による管理・移管・保存に関する報告書（平成 18 年 6 月）」

1.2. 文書・私文書の電子化

1.2.1 文書の電子化推進の背景

- 「e-Japan 戦略Ⅱ加速化パッケージ」（平成 16 年 2 月 6 日 IT 戦略本部決定）
 - 4. IT 規制改革の推進（D：Deregulation）
 - 1) e-文書イニシアティブ
- 法令により民間に保存が義務付けられている財務関係書類、税務関係書類等の文書・帳票のうち、電子的な保存が認められていないものについて、近年の情報技術の進展等を踏まえ、文書・帳票の内容、性格に応じた真実性・可視性等を確保しつつ、原則としてこれらの文書・帳票の電子保存が可能となるようにすること。

1.2.2 文書の電子化に関する法律

○「e-文書法」(平成17年4月1日施行)

民間事業者等が行う書面の保存等における情報通信の技術の利用に関する法律(平成16年法律第149号)

民間事業者等に対して法令で課せられている書面(紙)による保存等に代わり、電磁的記録による保存等を行うことを容認する法律

○「IT書面一括法」

書面の交付等に関する情報通信の技術の利用のための関係法律の整備に関する法律(平成12年法律第126号)

書面の交付をネットワークにおいて行うことを可能にする法律

2. 知的財産の保護

2.1. 知的財産・知的財産権とは

○「知的財産基本法」(平成14年12月4日法律第122号)

知的財産保護の基本となる事項を定めた法律

○「**知的財産**」とは、「発明、考案、植物の新品種、意匠、著作物その他の人間の創造的活動により生み出されるもの(発見又は解明がされた自然の法則又は現象であつて、産業上の利用可能性のあるものを含む。)、商標、商号その他事業活動に用いられる商品又は役務を表示するもの及び営業秘密その他の事業活動に有用な技術上又は営業上の情報をいう。」(知的財産基本法第2条1項)

○「**知的財産権**」とは、「特許権、実用新案権、育成者権、意匠権、著作権、商標権その他の知的財産に関して法令により定められた権利又は法律上保護される利益に係る権利をいう。」(知的財産基本法第2条1項)

著作権	著作権法(思想・感情の創作的表現の保護)
特許権	特許法(発明の保護)
実用新案権	実用新案法(物品の形状等に係る考案の保護)
意匠権	意匠法(工業デザインの保護)
商標権	商標法(商標・サービスマークの保護)
回路配置利用権	半導体回路配置保護法(半導体回路配置の保護)
育成者権	種苗法(種苗の品種の保護)
営業秘密	不正競争防止法(営業秘密、ドメイン名、著名標識など)

2.2. デジタルアーカイブと複製権

2.2.1 著作権法31条の定める図書館(著作権法施行令第1条の3)

- 国立国会図書館(柱書)
- 公共図書館(1号:図書館法2条1項の図書館)
- 大学や高専の図書館(2号:学校教育法1条の大学又は高等専門学校に設置された図書館及びこれに類する施設)
- 特別法に基づく高等教育機関(防衛大学校や水産大学校)の図書館(3号:学校教育法以外の法律に特別の規定があるものに設置された図書館)
- 公共施設(4号:一般公衆の利用に供する業務を主として行う施設で法令の規定によって設置されたもの)
- 研究所等が設置する施設(5号:学術の研究を目的とする研究所、試験所その他の施設で法令の規定によつて設置されたもので、資料を一般公衆の利用に供する業務を行うもの)
- 公益法人が設置する施設(6号:公益法人が設置する施設で文化庁長官が指定するもの)

指定施設(40:統合・廃止により現在36)

国立東京第二病院図書室、国立療養所東京病院図書室、(社)日本医師会医学図書館、(社)日本歯科医師会資料室、(社)日本原子力産業会議資料室、東京商工会議所商工図書館、名古屋商工会議所図書館、大阪商工会議所図書館、(社)経済団体連合会経団連図書館、(社)東京銀行協会銀行図書館、(財)社会経済生産性本部生産性研究所資料室、(社)全国市有物件災害共済会防災専門図書館、(財)原子力弘済会資料センター、金属鉱業事業団資料センター、宗教法人立正佼成会佼成図書館、(社)鋼材倶楽部資料室、(財)日本医薬情報センター附属図書館、雇用促進事業団全国勤労青少年会館図書館、(財)機械振興協会機械工業図書館、(社)日本鉄鋼連盟資料室、(社)日本化学会化学図書館・情報センター、日本貿易振興会大阪本部資料センター、(社)日本看護協会看護研修センター図書室、中小企業事業団中小企業情報センター、(財)海事産業研究所海事資料センター、(財)原子力安全研究協会資料室、高圧ガス保安協会保安情報センター、国際交流基金図書館、国立婦人教育会館婦人教育情報センター、社会福祉法人恩賜財団母子愛育会 日本子ども家庭総合研究所、(社)都市開発協会都市問題図書館、(社)土木学会附属土木図書館、科学技術振興事業団情報資料館、科学技術振興事業団筑波資料センター、日本貿易振興会図書館、(財)国際通信経済研究所資料室

2.2.2 複製の主体、対象及び態様

- 図書館その他の施設で政令で定めるもの
- 営利を目的としない事業
- 対象となる図書館資料の範囲

2.2.3 図書館資料保存のための複製

- デジタルアーカイブ
- 各種記録媒体への保存
- 技術の進歩と記録媒体の変遷への対応

3. 個人情報の保護

3.1. 個人情報保護制度

- 個人情報の保護に関する法律（平成 15 年 5 月 30 日法律第 57 号）
- 行政機関の保有する個人情報の保護に関する法律（平成 15 年 5 月 30 日法律第 58 号）
- 独立行政法人等の保有する個人情報の保護に関する法律（平成 15 年 5 月 30 日法律第 59 号）

3.2. プライバシーの保護

- ①自らに関する事柄について、外部からの干渉を受けずに自分の意思に基づいて行う「**個人の自律**」
 - ②他人から干渉を受けたり望まない侵入を受けない隔絶された状態や利益を保護するための「**私的な領域**」の保護
 - ③**他人に知られたくない個人に関する情報**の保護
- 判例（「宴のあと」事件判決（東京地判昭和 39 年 9 月 28 日判時 385 号 12 頁）「私生活をみだりに公開されないという法的保障ないし権利」
- プライバシー侵害による不法行為の成立要件
- ①公開された内容が私生活の事実またはそれらしく受けとられるおそれのある事柄であること
 - ②一般人の感受性を基準にして当該私人の立場に立った場合公開を欲しないであろうと認められること
 - ③一般の人々に未だ知られない事柄であること

4. 時の経過

4.1. 現用文書 → 非現用文書

歴史公文書等

4.2. 著作権

著作権の原則的保護期間

著作者が著作物を創作した時点から著作者の死後50年

4.3. プライバシー

時の経過によってプライバシー保護の対象になる場合

(「ノンフィクション『逆転』」事件判決(最判平成6年2月8日民集48巻2号149頁))

時の経過によってプライバシー保護の対象外になる場合

人格的利益の一身専属性

5. 情報のセキュリティの確保

- 「不正アクセス行為の禁止等に関する法律」(平成11年法律第128号)
 - IDやパスワードなどでアクセスが制御されているネットワークに接続されているパソコンに、他人の識別符を用いたり、セキュリティホールを攻撃するなどして不正にアクセスする行為を処罰する法律
- 「電子署名及び認証業務に関する法律」(平成12年法律第102号)
 - 情報内容の改ざん防止
- 「システム監査基準(平成16年10月公表)」
 - システムが適切に稼働しているか否かを監査する際の基準
- 「ソフトウェア等脆弱性関連情報取扱基準(平成16年経済産業省告示第235号)」
 - セキュリティホールの原因となるソフトウェア等の脆弱性情報の取扱いに関する基準

パネルディスカッションのまとめ

杉本重雄

以下は、パネルディスカッションの内容を、議論の記録をもとに筆者がまとめたものである。

パネリスト

大場利康（国立国会図書館関西館）
田窪直規（近畿大学短期大学部）
中井万知子（国立国会図書館書誌部）
中島康比古（国立公文書館）
司会 杉本重雄

いくつかの提起された視点・問題点

- (1) 過去に進められたデジタル化プロジェクトを振り返って考えてみることの必要性
- (2) デジタルアーカイブに関するいろいろな機関、組織の間での協調の必要性
- (3) デジタルアーカイブということばに関する理解
- (4) 紙資料の保存の方法をもう一度確認してみることに

- (1) 過去に進められたデジタル化プロジェクトを振り返って考えてみることに

90年代に行われた実験的なプロジェクトで作られたいデジタルコンテンツや文化財をデジタル化して作られたいものなどで、現在使えなくなっているものがある。デジタルアーカイブ自体はコストも考えないと長生きできないであろう。アメリカ議会図書館が推進した国家情報基盤と保存プロジェクト（National Information Infrastructure and Preservation Program, NDIIPP）では非技術的な視点を含めいくつかのトピックをあげている。こうしたプロジェクトから得た知見をもっと有効に活用すべきではないか。たとえば、技術主体で行われたプロジェクトで作られたいものの中で、現在使えなくなったものがどの程度あるのか、またその理由は何であるのか、デジタルアーカイブの構築を進めるチームはどのように構成すべきか、といったことを広く共有できるようにすべきであろう。

- (2) デジタルアーカイブに関するいろいろな機関、組織の間での協調の必要性

異なる機関・組織においてデジタルアーカイブに携わる人の中でこれまでの経験から得た知見、海外の参考にすべき例などについて情報交換する機会が求められる。また、デジタルアーカイブを構築していく際に用いるファイルフォーマット、プロトコルなど技術的情報に関する情報交換を進めることで相互運用の可能性を高めていくための努力も必要であろう。

デジタルコンテンツの長期保存に関する技術開発、標準化といったことは民間ができることではなく国レベルでの取り組みが必要である。

(3) デジタルアーカイブということばに関する理解

過去 10 年余りの間の経験で電子図書館やデジタルアーカイブといったことばに関する理解は進んだ。また、知的財産権やプライバシーなど長期利用のための社会制度的な問題についても理解が進んできた。デジタルアーカイブを構築する機関・組織による取り組みの違いについても過去の経験から理解が進んできたと思える。その一方、組織の記録（レコード）として扱われるものと、出版物として扱われるものとの間の垣根が低くなってきたといった変化もある。また、情報技術やネットワーク環境はどんどん進化し続けている。従来のような縦割りではなく、異なる機関・組織の間での協調がより強く求められるといえる。

(4) 紙資料の保存の方法をもう一度確認してみる

紙資料の保存がうまくいかないところでデジタル資料の保存がうまくいくとは考えられない。紙媒体での長い蓄積を踏まえて、収集、蓄積、利用といった総合的な観点から、電子資料のアーカイブについても進めるべきであろう。

参考

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

筑波大学・知的コミュニティ基盤研究センターでは、デジタルアーカイブの長期利用とデジタル情報資源の長期保存に関する現状と将来への課題をいろいろな観点から議論するため、下記の研究会を開催いたします。どうかふるってご参加ください。

日時： 2007年2月16日（金）

場所： 筑波大学・春日キャンパス・情報メディアユニオン3階・共同研究会議室
（つくば市春日1-2）

つくばセンター（つくばエクスプレス・つくば駅）より徒歩約10分

主催： 筑波大学・図書館情報メディア研究科／知的コミュニティ基盤研究センター

参加： 無料（どなたでも参加できます。）

プログラム

10:00-11:45

杉本重雄（筑波大学・図書館情報メディア研究科／知的コミュニティ基盤研究センター）

背景と目的、デジタルアーカイブとその長期利用に関する問題の提供

藤原誠（国立国会図書館・関西館）

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

牟田昌平，中島康比古（国立公文書館）

公文書の電子的保存に関して

<<昼休み>>

13:15-15:00

和田光俊（独立行政法人科学技術振興機構 文献情報部）

学術雑誌のデジタルアーカイブと長期利用に関して

西岡貞一（筑波大学・図書館情報メディア研究科）

文化財のデジタルアーカイブと長期保存に関する問題

新保史生（筑波大学・図書館情報メディア研究科）

デジタルアーカイブとの関係における知的財産権・プライバシーの権利の保護について

<<休憩>>

15:30-16:30

ディスカッション（パネル討論）

自治体、博物館、資料館、企業といったいろいろな視点からの意見交換

<<終了>>

問い合わせ先： 筑波大学・知的コミュニティ基盤研究センター

Tel. 029-859-1524, <http://www.kc.tsukuba.ac.jp/>

本報告書は、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の講演資料をまとめたものである。本シンポジウムの報告は知的コミュニティ基盤研究センターのホームページ(<http://www.kc.tsukuba.ac.jp/>)からも提供する予定である。