

Digital Curation and Archives, on/as/in Clouds

杉本重雄

知的コミュニティ基盤研究センター
図書館情報メディア研究科

概要

前置き:

- ここで述べることは、筆者の素朴な疑問に基づくものであり、研究報告といったものではない。自由な意見交換のタネになれば幸いと考えている。

内容:

- Digital Curation, Digital Archive (DCA)
 - 自分なりのことばの意味の確認
 - 最近の自分自身の活動とDCA
- DCAの中でのメタデータの役割
- DCAとクラウド環境

Terms - ことば

- 図書館(L), 文書館(A)と博物館・美術館(M),
 - L/A/Mの違いは？
 - Digital L/A/Mとは？
- Digital Archive, Digital Curation
 - 価値あるものを選び、集め、組織化して、保存し、アクセスできるようにする
 - Digital Archive: 電子公文書館ではない
 - Curation: 価値ある資源を選び、収集し、保存提供する
- Resource (資源、情報資源)
 - デジタル化資源とデジタル形式で作られた資源
 - インターネット上で識別し、アクセスすることができる

Digital CurationとDigital Archive

- 長期間にわたるデジタル形式での資源の保存
 - 有形無形にかかわらず、対象資源のデジタルコピーの制作、保存、利用
 - デジタル形式での保存と利用
 - いつでもどこからでもアクセスできる
- Digital Curation, Digital Archiveの対象範囲
 - Cultural Heritage(文化遺産)
 - 写真、絵画、本、文書、建物、ダンス、他
 - 記録文書、公文書
 - 科学技術データ
 - その他いろいろ

Digital CurationとDigital Archive

- 有形・無形の文化財を電子化して、蓄積・組織化したもの (a narrower view)
- 内容やフォーマットを問わず、価値ある資源を、デジタル形式で、保存し、長期にわたって提供するもの (a broader view)
- 「主体的に集め、保存する」⇒Webとの違い
- Digital CurationとDigital Archiveの違い
 - 実質的にはそんなに大きな違いはない
 - CurationとArchiveの言葉の意味の違いによる違い

DCAとメタデータ

- L/A/Mは
 - 選ぶ／集める／組織化する／提供する／保存する
(Select, Collect, Organize, Access, Preserve)
 - すべてにメタデータを必要とする
- 利用者にとっては
 - 個別のアーカイブを使う
 - 特定の領域のアーカイブを使う
 - 領域の違いを越えてアーカイブを使う
 - 適切なアーカイブを探すためにメタデータを必要とする

Internet上でのメタデータの利用



ネット上では、いろいろな利用者が、いろいろな資源を利用する。
メタデータなしに資源を探し、アクセスし、利用できるか？

メタデータ情報基盤構築事業

- 総務省・新ICT利活用サービス創出支援事業による支援 (http://www.soumu.go.jp/menu_news/s-news/01ryutsu02_01000005.html)
 - 3省書デジ懇の報告をベースにした支援事業
- メタデータ情報を流通させるための基盤としてメタデータスキーマレジストリを核とするシステム (MetaBridge) の開発
 - <http://www.meta-proj.jp/index.html> , 2010.11-2011.3
 - 筑波大学・知的セは代表として、インフォコム(株)、(株)インフォコム西日本他による提案
 - DCMI Application Profile, RDF, OWL等のSemantic Web技術の利用
 - 技術的リード: 神崎正英氏、永森光晴氏

メタデータ情報基盤構築事業

- メタデータスキーマを収集し、それをDCMIの Application Profile (Singapore Framework) の概念に基づいて整理し、RDF等の標準的な形式で表現し、蓄積し、提供する
- Application Profile
 - メタデータの語彙と構造定義を分離する
- 現在、MetaBridgeを知的セへ移行中

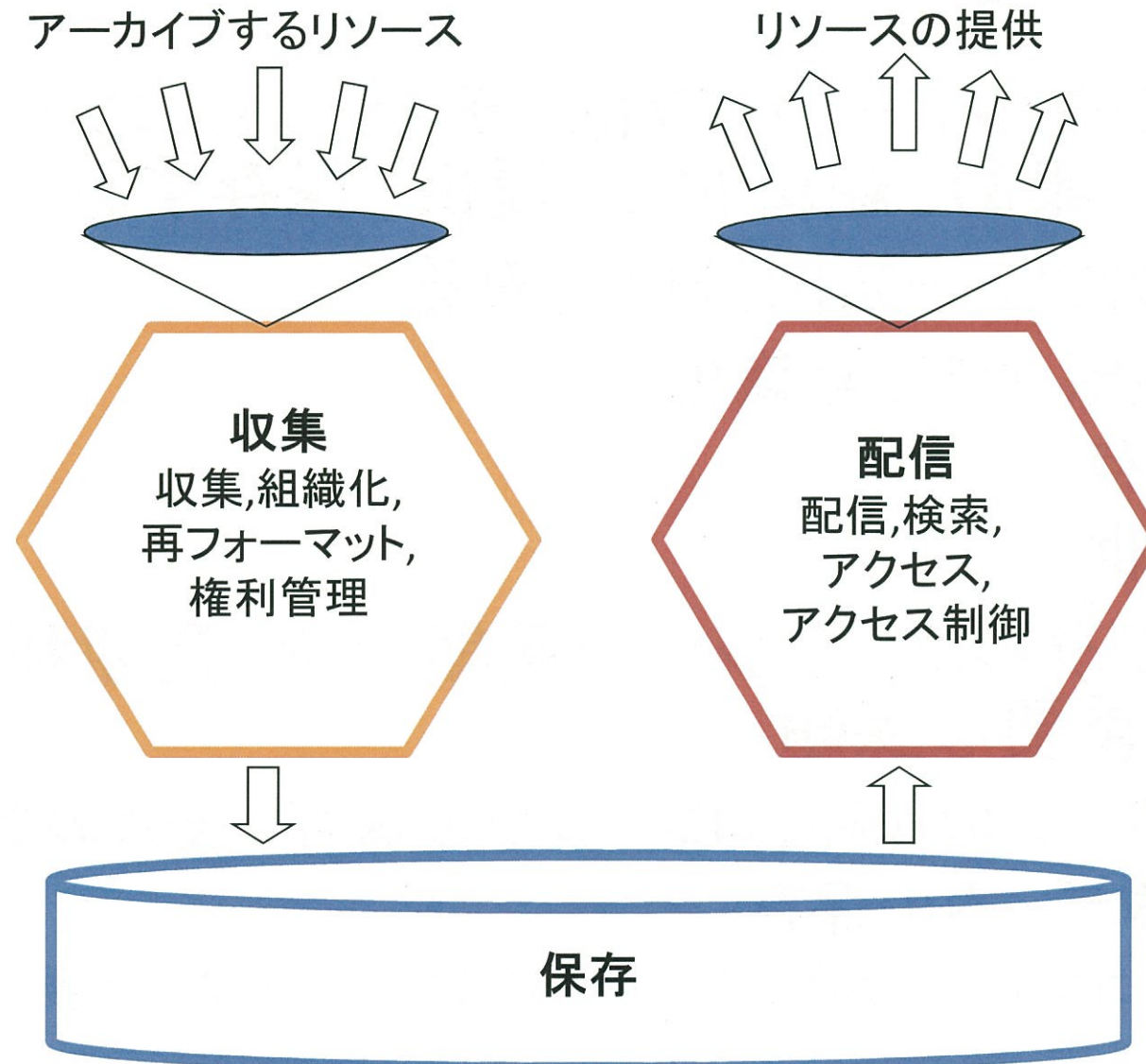
メタデータ情報基盤構築事業

- 今後、知的セをベースにした運営組織の立ち上げ
 - 国立国会図書館他、メタデータにかかわる専門知識を持つ機関との連携
 - メタデータ情報基盤研究会
 - メタデータ情報の収集と提供にかかわる産官学民を問わない連携
 - メタデータスキーマの長期利用や相互運用性にかかわる課題の検討

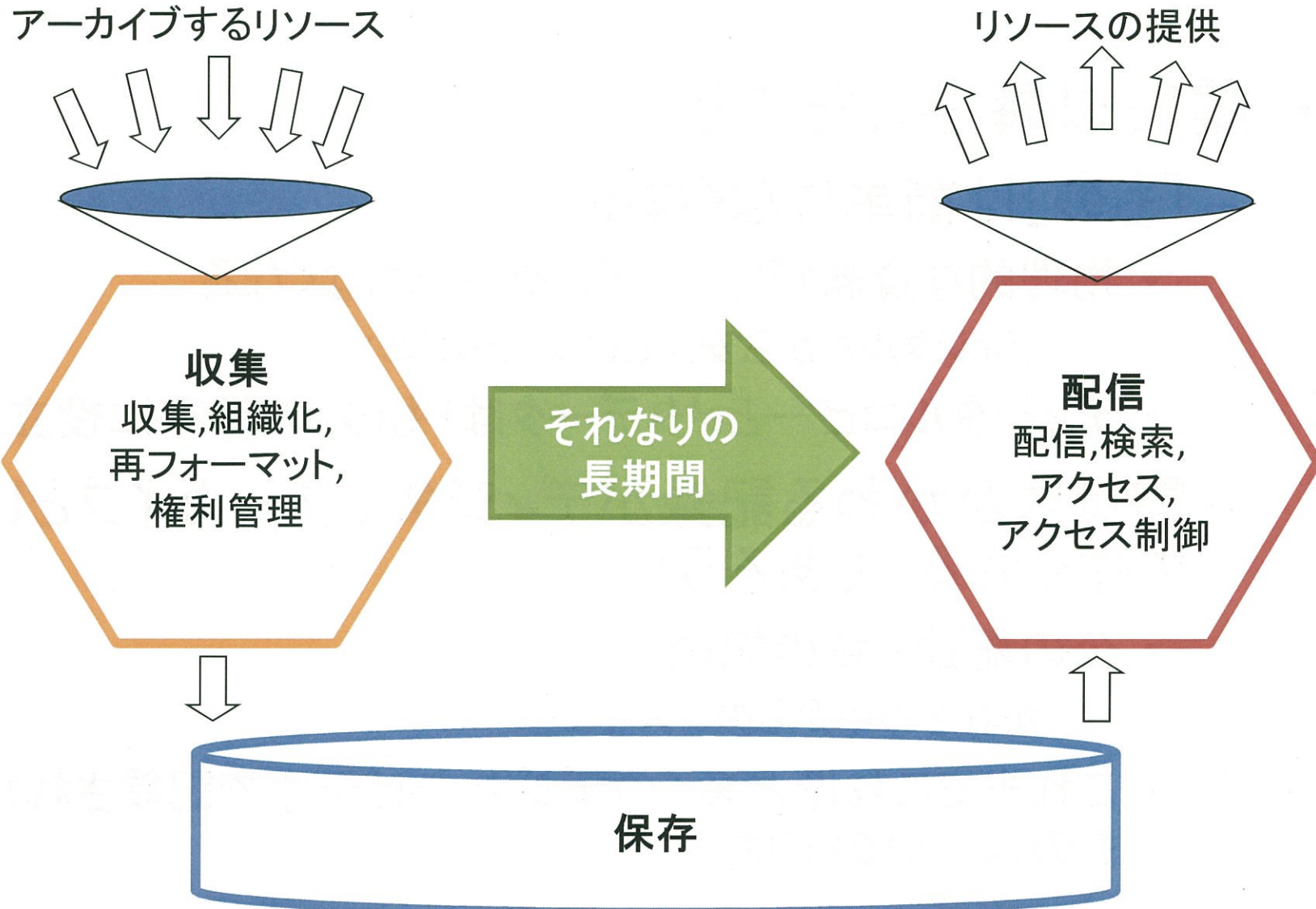
Digital Archiveのモデル

- デジタルアーカイブは、デジタルコンテンツを選び、集め、組織化し、提供、保存する
- 3つのフェーズに分けてみる
 - 選択、収集、組織化
 - 保存
 - 提供
- デジタルコンテンツの長期保存・長期利用は非常に重要な課題
 - メディアの保存、コンテンツの保存、技術の保存
 - メタデータの保存
 - コンテンツ保存にはいろいろなメタデータが必要

アーカイブのモデル – 一般化した視点



アーカイブのモデル – 一般化した視点



デジタルアーカイブについて思うこと(1/4)

- 震災以降感じたこと

- 「もの」は簡単になくなる

- 物理的な資料だけでなく、サーバもなくなる

- デジタルだから安全ということはない

- デジタルコピーとメタデータは「もの」の保存に役立つ

- 震災にかかわる記録がデジタルアーカイブとして作られる(であろう)

- 公の記録と私の記録

- がれきの中からのレスキュー

- これまでこれほど多くのデジカメ・ビデオで記録されたものはないのでは？

デジタルアーカイブについて思うこと(2/4)

- MLA連携に関して
 - デジタルだから連携したい
 - デジタルコンテンツの作り方や提供方法の違い
- 資料のデジタル化に対する取り組み
 - 海外の取り組みに比べてどうか？
 - MLAによる違い、規模による違い
- デジタル化本とBorn Digital本、電子配信
 - 組み合わせることの意味
 - デジタルアーカイブと電子出版は隣同士

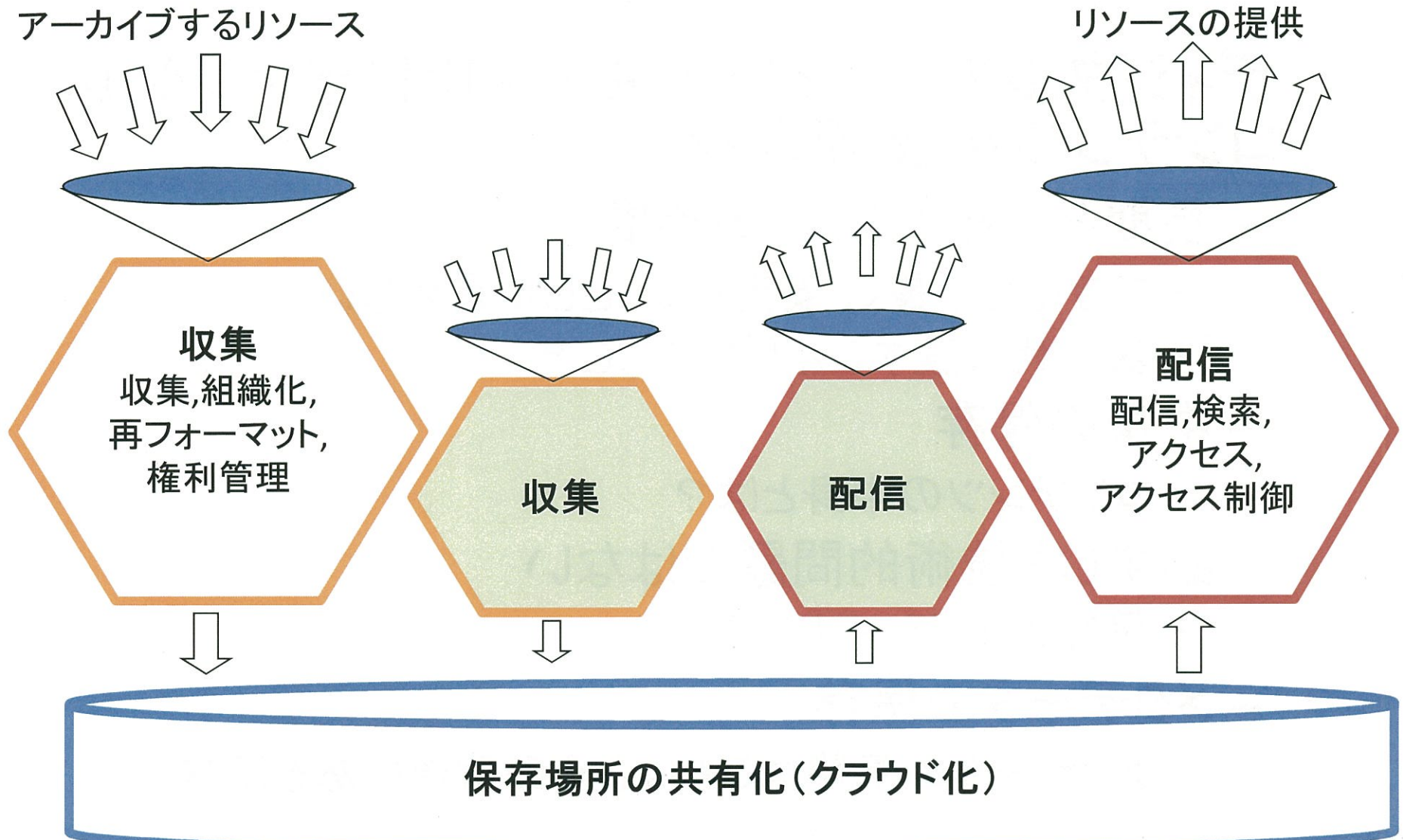
デジタルアーカイブについて思うこと(3/4)

- 電子公文書と公文書の電子化
 - 電子政府に電子公文書はつきもの
 - 業務の効率化、情報公開、セキュリティ
 - 電子公文書の長期保存とRecords Lifecycle
- メタデータに関して
 - 図書館はそれなりに標準化されているが、それ以外では？
 - MLAによる違い、規模による違い
 - 個別のニーズと共有のためのニーズのギャップ
 - L/A/M全体のパイを大きくするための工夫の必要性

デジタルアーカイブについて思うこと(4/4)

- デジタルアーカイブの長期利用、デジタルコンテンツの長期保存
 - 長期とはどれくらいの期間？
 - できるだけ長く、安定して、としか言えない
 - 次世代に渡すことができること
 - 中身の保存
 - コンテンツの中身とは？
 - 必ずしも技術的問題ではない
 - 保存のポリシーと戦略
 - メタデータの保存
 - メタデータの基盤はことばとその背後にある意味

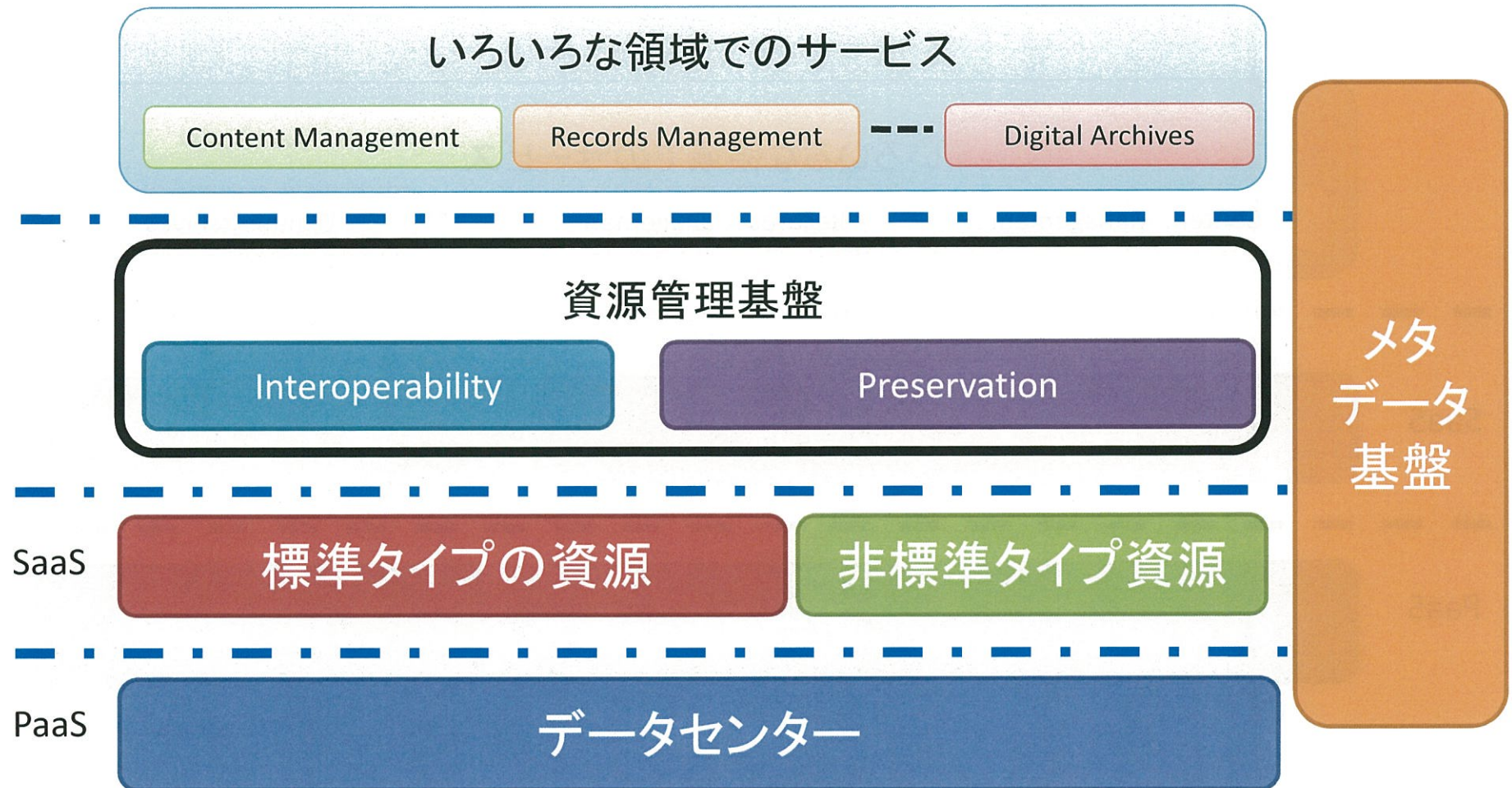
アーカイブのモデル－一般化した視点



クラウドを階層としてみる



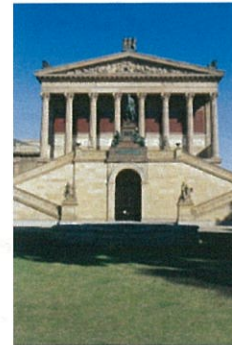
クラウドを階層としてみる —少し視点を変えて—



デジタルアーカイブについて思うこと

- デジタルアーカイブのインフラとしてクラウド
- Digital Archive on a Cloud
(クラウド上のデジタルアーカイブ)
 - デジタルアーカイブを作る基盤としてのクラウド
- Digital Archive as a Cloud
(クラウドとしてのデジタルアーカイブ)
 - デジタルアーカイブをまたぐリソースアクセス
- Digital Archive in a Cloud
(雲の中のデジタルアーカイブ)

Digital Archives on a Cloud

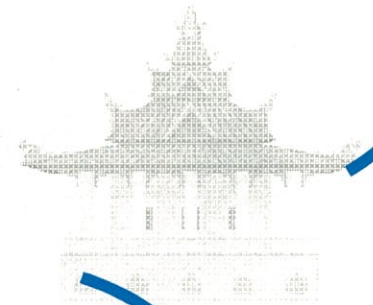
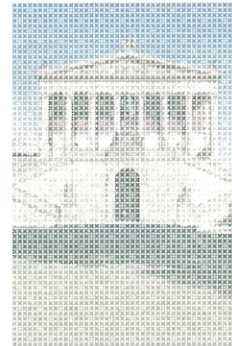
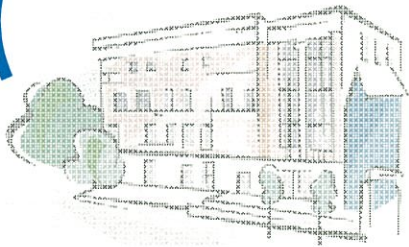


デジタルアーカイブを作る
基盤としてのクラウド

Digital Archives as a Cloud



Digital Archives in a Cloud



雲の中のデジタルアーカイブ