

文化情報資源の共有化

宇陀則彦

資源共有プロジェクト

- 総合研究大学院大学共同研究プロジェクト
 - 「文化科学研究分野における情報資源共有化のためのコラボレーション研究」
 - 2003年度-2004年度
- 人間文化研究機構 国文学研究資料館 複合領域研究系プロジェクト
 - 「文化情報資源の共有化システムに関する研究」
 - 2004年度-2006年度
- 人間文化研究機構 人間文化研究資源共有化推進事業
 - 2005年度-2007年度

プロジェクトメンバー

- 安永尚志 (国文学研究資料館)
- 原正一郎 (国文学研究資料館)
- 相田 満 (国文学研究資料館)
- 五島敏芳 (国文学研究資料館)
- 久保正敏 (国立民族学博物館)
- 山本泰則 (国立民族学博物館)
- 合庭 惇 (国際日本文化研究センター)
- 早川聞太 (国際日本文化研究センター)
- 安達文夫 (国立歴史民俗博物館)
- 鈴木卓治 (国立歴史民俗博物館)
- 関野 樹 (総合地球環境学研究所)
- 及川昭文 (総合研究大学院大学)
- 大山敬三 (国立情報学研究所)
- 神門典子 (国立情報学研究所)
- 北本朝展 (国立情報学研究所)
- 石上英一 (東京大学資料編纂所)
- 横山伊徳 (東京大学資料編纂所)
- 柴山 守 (京都大学東南アジア研究センター)
- 森 洋久 (大阪市立大学)
- 宇陀則彦 (筑波大学)
- 松村 敦 (筑波大学)

プロジェクトの目的

- 多くの研究機関でデータベースが公開されているが、現在では個々の仕様、事情、制限に従ってその都度データベースを切り替えなければならない。
- 同一概念や主題などから、関連するデータベースを同時的に渡り歩くのは不可能に近い。
- そこで本プロジェクトは、情報資源を切り替えて利用するのではなく、関連する情報を横断的に検索し、利用する仕組みを構築することを目的とする。

人文科学の学際的広がり

- 地震史料のデータベース化
 - 神戸大学 都市安全研究センター 石橋克彦
- 地震史料:地震が起きたときの伝聞、事例、伝承などの記録・・・歴史地震学、史料地震学
- 歴史地震のカタログを作ること
- 『大日本地震史料』田山実 1904(明治37年):日本書紀以来の地震の記録を集めたもの
- 1923年(大正12年)関東大震災
- 武者金吉 『日本地震史料』

人文科学の学際的広がり

- 日本最古の歴史地震:416年8月23日
 - 允恭天皇の時代科学的にはかなり疑わしい
- 確実な歴史地震:599年5月28日
 - 推古天皇の時代 聖徳太子もコメント
- Fake Earthquake 捏造地震
 - 1233年3月24日 天福南海地震
 - 『日高郡誌』から『蓮専寺記』を調べると、蓮専寺が開基されたのは1485年であることが判明
 - その他、同時代の記録には記述なし
- 地理情報システムと史料データベースのリンク

人文科学の国際的広がり

- 明治学院大学 Michael Watson
 - 「残念ながら、アメリカの大学とかイギリスの大学と比べると、日本の大学は研究者のメーリングリストやリストサーブの重要性を理解してくれない」
- ヴェネチア大学 アルド・トリーニ
 - 「日本人は外国の研究者が書いたものともっと読んだほうがよい。外からみた日本文学はどういうものかを知ると、新しい世界が開けるはずだ」

資源共有システムの必要性

- 「伊能忠敬」Googleでも調べられる。
- しかし、伊能忠敬に関する論文がどこにあるか、伊能家はどういう家か、「日本緯経度実測」はどういう本で、どこで見ることができて、全文データベースはあるのか、手に入れられるのか、地図データはあるのか、関連主題である「幕末」「シーボルト」などはどう関連するか、等々、学術的に正確な情報を入手するのは困難
- 専門組織のデータベースを(も)調べる必要がある。
- 関連するデータベースを渡り歩くのか？ (非効率)

資源共有システムの必要性

- 現実的な課題
 - データベースごとに検索手法が異なる
 - 類似の事項、用例が同時に検索できない
 - 関連情報を適切に検索できない
- 解決策(案)
 - 全てのデータベースを同じ作りにする。
 - 利用者が全ての検索方法をマスターする。
 - システムがサポートする。

資源共有システムの難しさ

- データベースは全て異なる
 - 独自性、専門性(同一内容であっても異なる)
 - 情報資源の多様性
 - 非移植性
 - 非互換性
- 研究計画
 - 第1段階: 研究機関の相互接続
 - 第2段階: メタデータDBによる同一操作
 - 第3段階: 標準検索プロトコルによる同一操作

プロジェクトメンバーのコメント

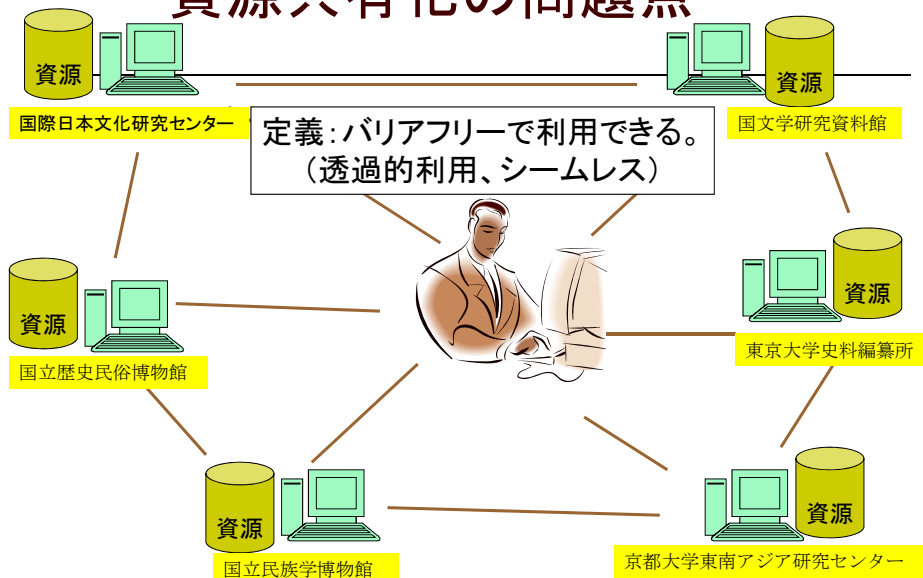
- 石上英一(東京大学史料編纂所)
 - 「機関間の資源共有化の前提には、各機関における学術的使命の実現のために、資源の公開と相互の利用環境構築が不可欠という内発的な意思が存在しなければならない」
- 久保正敏(国立民族学博物館)
 - 「自分でデータを使わない人は、データを作ってもあまり意味がないことがはっきりした。データを作るとしても、ちゃんとしたアーキビストやカタログャーがいないとどうにもならないこともはっきりした」

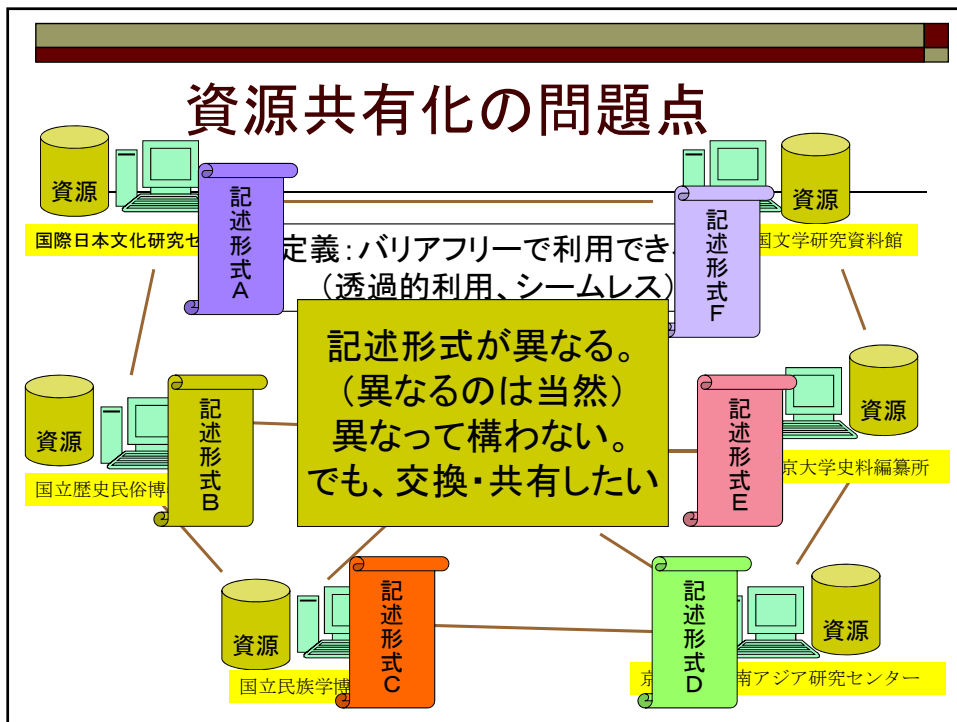
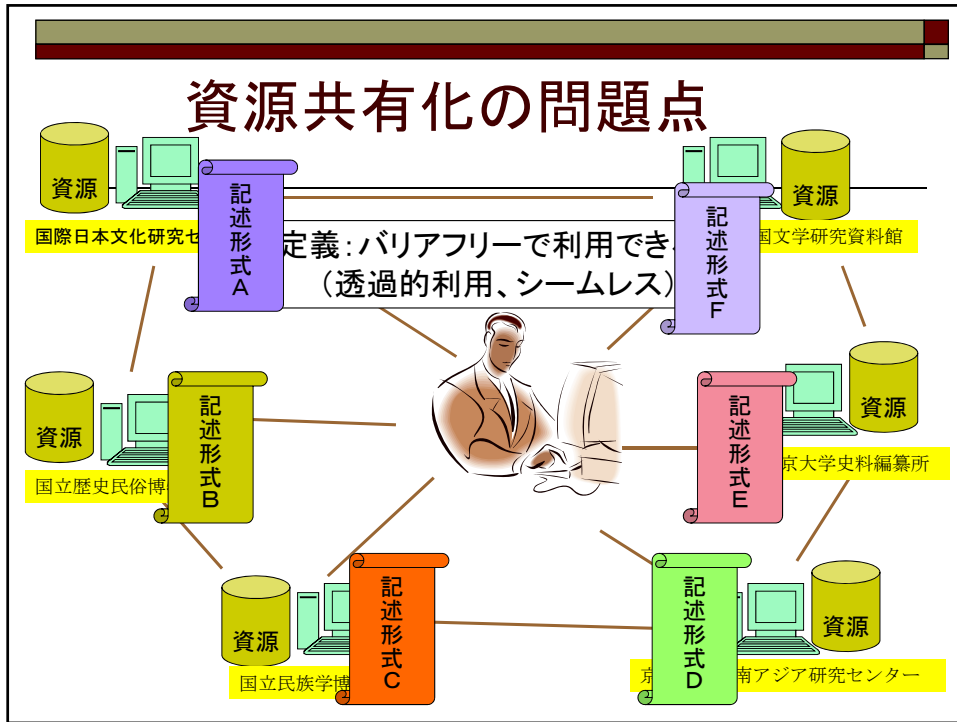
プロジェクトメンバーのコメント

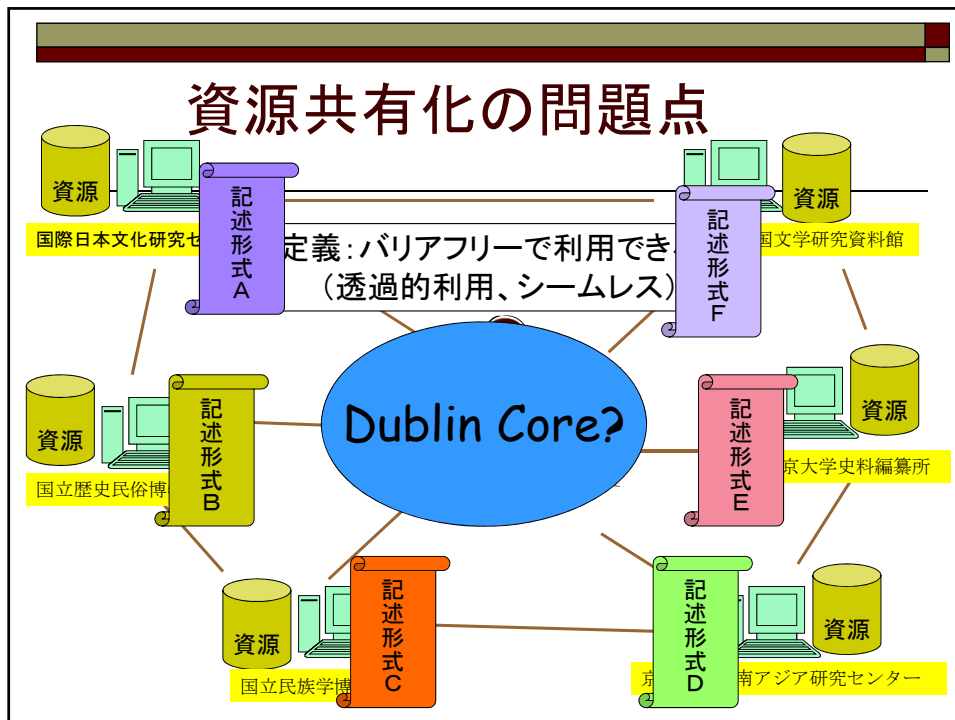
- 横山伊徳(東京大学史料編纂所)
 - 「機能としては、図書と史料を統合したデータベースが欲しいが、図書館学系の目録と記録管理学系の目録はしばしば発想が逆方向に向いているので難しい。フィールド構成は図書室の実際の記述項目からとり、階層構成は記録管理学の発想からとった。要は折衷案をとったということだ」

メタデータ関連のチュートリアル

資源共有化の問題点







Dublin Core (DCMES)

- Dublin Core Metadata Element Set
- Resource Discovery (資源発見)のため
- Interoperability (相互運用性)を最重視
- Element Setの最大公約数的位置づけ
- Coreとなる15項目

Dublin Core Metadata Element Set

- Title(タイトル)
- Creator(責任者)
- Subject(主題)
- Description(記述)
- Publisher(出版社)
- Contributor(寄与者)
- Date(日付)
- Type(ジャンル)
- Format(物理的形式)
- Identifier(識別子)
- Source(出所)
- Language(言語)
- Relation(他との関連)
- Coverage
(地理的、時間的特性)
- Rights(権利、条件)

Dublin Core (DCMES)の考え方

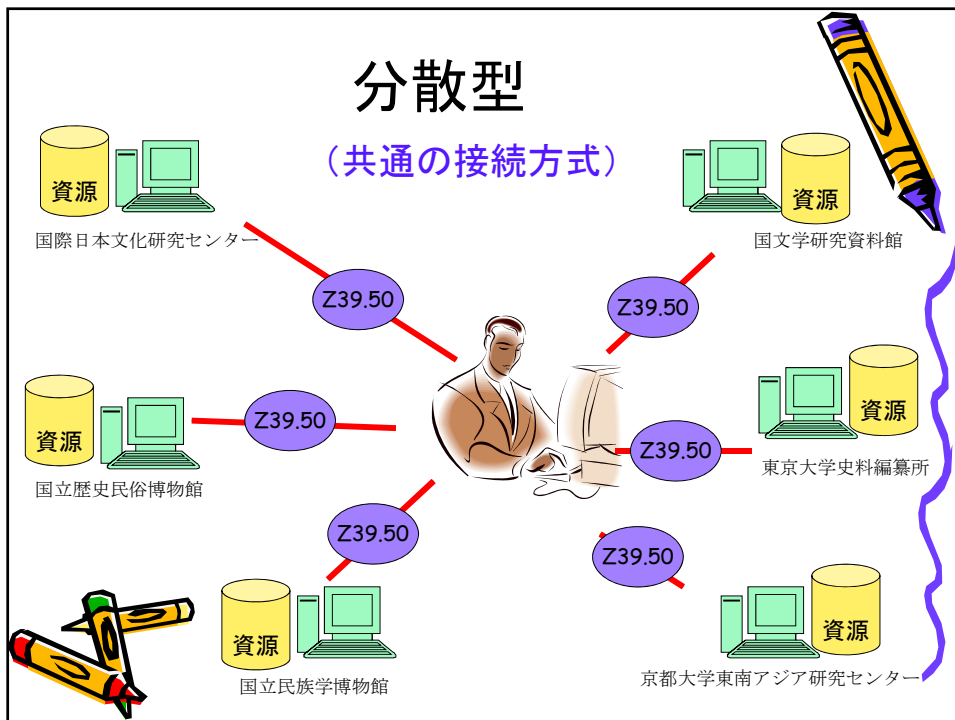
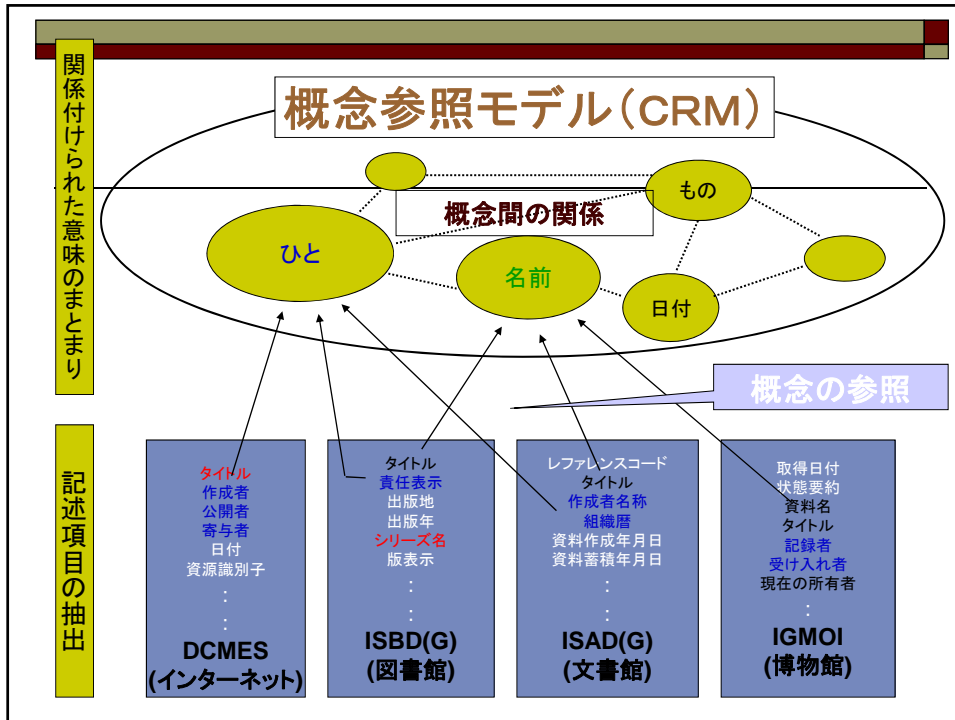
- Dublin Core Metadata Element Set
- 草の根的活動から始まった。
- ISBDのような強制力はない。
- 全てのElementが省略可能で繰り返し可能
- 独自のElementを拡張可能
- <http://dublincore.org/>

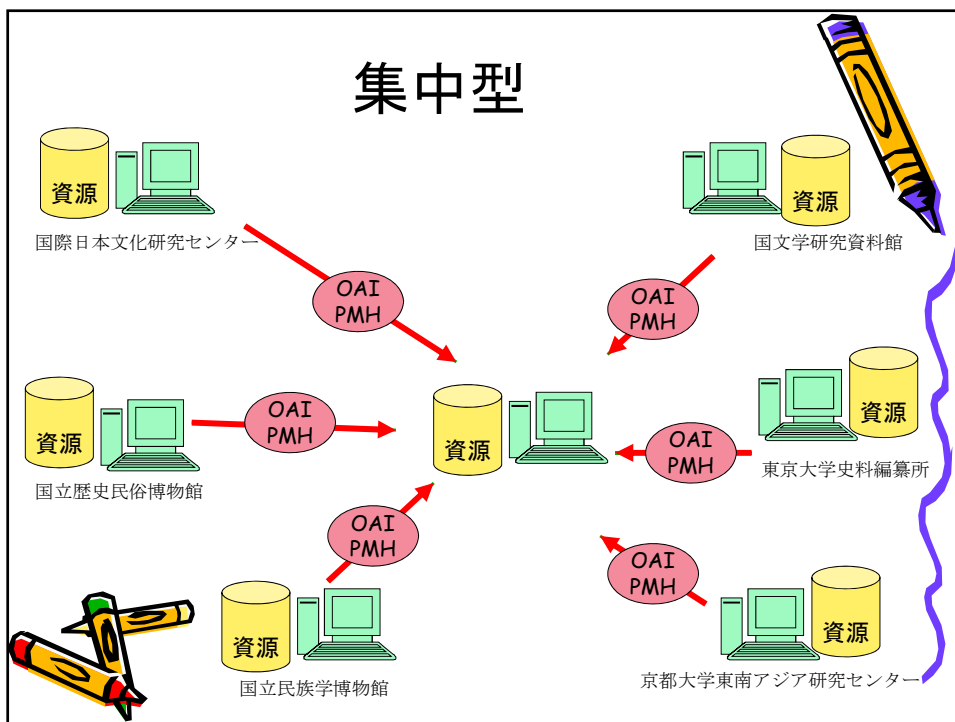
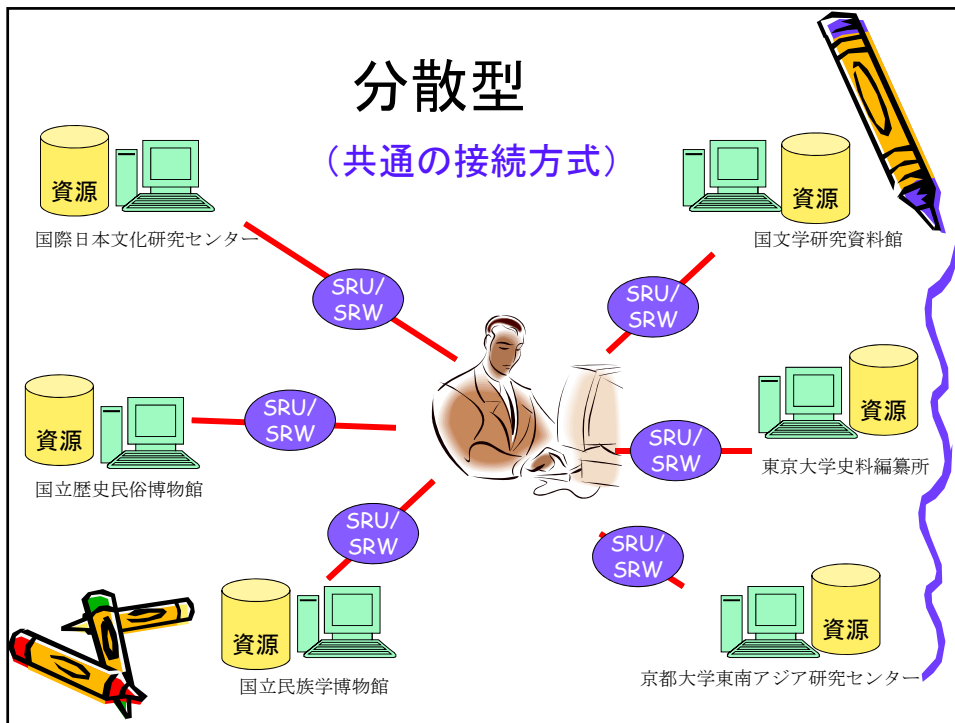
多様なメタデータセット

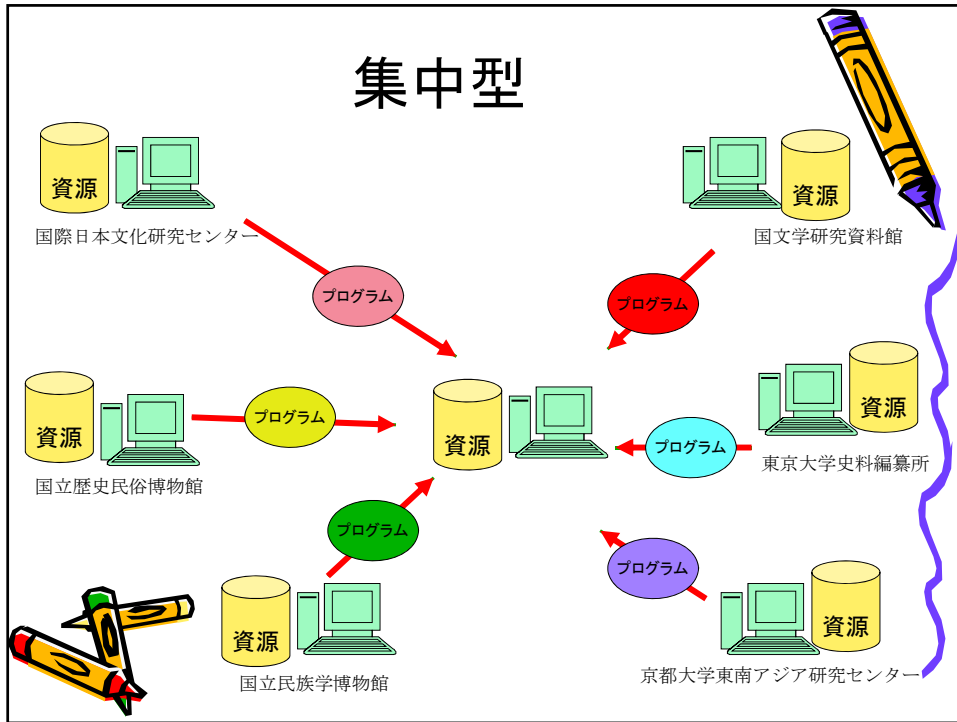
- CHIN(カナダ文化遺産情報ネットワーク)
 - <http://www.chin.gc.ca/English/Standards/>
- AMICO(美術館画像コンソーシアム)
 - <http://www.amico.org/AMICOLibrary/dataspec.html>
- CDWA(美術品記述のためのカテゴリー)
 - http://www.getty.edu/research/conducting_research/standards/cdwa/
- CIMI(米国文化遺産コンソーシアム)
 - <http://www.cimi.org/>
- 参考ページ
 - <http://www.ndl.go.jp/jp/library/current/no222/doc0005.htm>

メタデータ関連仕様

- DCMES: Dublin Core Metadata Element Set
- ISBD: International Standard Bibliographic Description
- ISAD: International Standard Archival Description
- IGMOI: International Guidelines for Museum Object Information
- MODS: Metadata Object Description Schema
- EAD: Encoded Archival Description
- RDF: Resource Description Framework
- CRM: Conceptual Reference Model
- ABC: ABC Model
- FRBR: Functional Requirements for Bibliographic Records
- Topic Maps:
- OWL: Web Ontology Language







留意点

分散型

- ・ それぞれのデータベースにZ39.50あるいはSRU/SRWのサーバプログラムが必要
- ・ プロトコルに合わせてデータベースの改造が必要かもしれない。
- ・ データの重複(更新頻度はどのくらいか)
- ・ データエレメントのマッピングが必要
- ・ OAI-PMHでなければ、最悪、入力プログラムを別々に作る必要あり

集中型

メタデータ関連のチュートリアル

- 「アプリケーションプロファイル設計のための枠組みや方法論はあるのか？」
- 「現在のところない」
- 「OAI-PMHのモデルは無理ではないか。データをくださいと言っても多分くれない。あげるのはごみのようなデータになりかねない。あげるとしても、作った側の思い入れがあり、それがどのように使われるかはかなり不安に思うだろう」
- 「無理によこせというのは当然無理。けれども、コミュニティで合意がとれれば便利なプロトコル」

プロジェクトメンバーのコメント

- 及川昭文(総合研究大学院大学)
- 「広く解釈すると、共有化は実現できている」
 - 書誌情報を通じた論文の共有
 - データベースの公開
- さらに何ができるか
 - 客観的な議論が可能になる
 - 共有した情報に対して複数の視点から議論できる
 - 議論で得られた知見が共通の情報となる
 - 若い研究者の育成に役に立つ

プロジェクトメンバーのコメント

- 及川昭文(総合研究大学院大学)
- メタデータとシステムに対する懸念
 - 定型化することによる自由な発想の阻害
 - メタデータによってある種の限界が生じる
 - 情報の蓄積や探索に多くの時間を割くようになる
 - 道具作りは不可欠だが、そればかりになる。
 - おもしろいからやめられない
 - 情報と道具への依存