

世界のデジタル化プロジェクト

土屋俊
(千葉大学)

デジタル化(電子化、digitization)

- 既存媒体資料の電子的な保存と利用
 - 紙媒体(印刷物、手稿)資料の画像・テキスト
 - 音楽・音声・画像・映像資料
 - ネットワーク化の進展と微妙な関係
- 1990年代以前(1990年代初頭まで)
 - 古典的テキストの電子化:トマス・アキナス、ウィトゲンシュタイン、初期アメリカ文献、古英語・中英語、岩波古典体系、日蓮遺文などなどなどなど
 - 図書館スペース問題解消プロジェクト(JSTOR)
 - 資金源(NEH・メロン財団・JISC・科研費データベース等および、営利ないし準営利⇒図書館の購入)

1990年代前半の問題

- 記録形式の標準化
 - マークアップ言語による方式(たとえば、Text Encoding Initiative)が主流(日本ではほとんど普及せず)
 - 音声資料、画像資料などについては独自の標準化
- メタデータの標準化
 - ファイル記述方式とMARCとの関係
- NSF/JISC型Digital libraryプロジェクトへ
- ネットワーク経由の利用の普及と新たな問題
- 生産時から電子的である(born digital)資料の問題

1990年代後半における 日本の電子図書館プロジェクト

- 紙媒体資料からの(画像)電子化が中心
 - 奈良先端科学技術大学院大学、京都大学、筑波大学、東工大、神戸大学、(図書館情報大学)、そしてその他の大学(独自資金、科研費等)の各図書館のプロジェクト(他の情報関連施設は関心なし)
- 貴重資料中心の電子化(インターネット発信)、ただし
 - 新刊雑誌、新刊本も対象とした奈良先端(ただし、スペース節約ではなかった)
 - 学内生産物にも展開した筑波大学の構想
- 現在では、「散在」
 - **メタデータの整備によってこれからの大学機関リポジトリに吸収、再編が可能⇒緊急の課題**

2000年代初頭

- 電子化・デジタル化そのものは前景から退き、
- 電子ジャーナルの急速な普及
- 電子資源管理(ERM)の技術的、商業的展開
 - FEDORAなどのNSFプロジェクト
 - Innovative, Endeavor, Ex Librisなどなど
 - 機関リポジトリソフトウェア(DSpace, Eprint等)
 - 日本のシステムベンダーは追従せず(なぜ?)
- メタデータ・ハーベスティング標準の普及
- 日本はどうも蚊帳の外？

2005: 突然デジタル化の再燃

- 2004年12月: Google Print の発表
 - Harvard, Stanford, the University of Michigan, and Oxford University, the New York Public Library (NYPL)所蔵の3千万巻をスキャン
- 2004年夏から秋: やや休止
 - 著作権者からの意見多数
 - 9月にはAuthors Guildからの提訴(デジタル化するかどうかを決めるのは、権利者だ)

ヨーロッパの反発

- 2005年4月 : Jean-Noel Jeanneney (フランス国立図書館) の発言とシラク大統領の支持発言
- 引き続き : ヨーロッパ各国の支持 (ドイツ、ハンガリー、イタリア、スペイン) ⇒ EU資金を求める
- しかし、出版界はここでも抵抗。

さらに別のプロジェクト

- 2005年10月 : Open Content Alliance発表。
Yahoo, the Internet Archive,などがパブリック・ドメインの印刷物をdigitize。HPとAdobeが技術を提供。
 - the University of California system, The University of Toronto, the European Archive, the National Archives in the United Kingdom, O'Reilly Media and Prelinger Archives がコンテンツを提供。
- Googleがなんでも無許諾でdigitizeして、利用で制限に対して、無許諾でよいものだけをやる

しかし、技術は進む

- 1台で1日10000ページをスキャンする機械を数台導入。過去半年のパイロット段階で2000巻を入力済み(トロント大学)
- たとえば、Kirtas社のAPT BookScan 1200は、1時間に1200ページをスキャンして、1台120,000USD
- 大量電子化時代への移行の兆候
 - 日本語文献については止まっている。すくなくとも見えない。
 - 英語などについても脆弱。

JSTORの日本語文献電子化

- パイロットプロジェクト
- ハーバード大学図書館が印刷媒体を提供
- 著作権が切れているもの
 - Garakuta bunko 我楽多文庫, No. 1-16 (1888-1889)
 - Toyo gakuho 東洋学報, Vol. 1- 14 (1911-1924)
 - Rekishi chiri 歴史地理, Vol. 1-48 (1899-1926)
 - Shigaku zasshi 史学雑誌, Vol. 1-34 (1889-1926)
 - Shirin 史林, Vol. 1-10 (1916-1925)

図書館とデジタル化

- 図書館資料がデジタルに利用できない状態は、国際的には、当然ではない
- デジタル化のコストは、極度に下がっている
- 当面、機関リポジトリは紙資料のデジタル化を行うだろう。図書館資料の電子化については著作権の問題は残るだろう
- それらのメタデータ、保存については、各機関の連携が必要であろう。