

20 Oct 2017

Asian Library Café 002

古典籍 on flickr !

～漢籍・法帖を写真サイトでオープンしてみると～

はじめに: なぜU-PARL は flickrを選んだのか?

富澤かな



U-PARL

東京大学附属図書館アジア研究図書館上席倫理財団寄付研究部門
UEHIRO PROJECT FOR THE ASIAN RESEARCH LIBRARY
THE UNIVERSITY OF TOKYO LIBRARY SYSTEM

本日目指すところは...

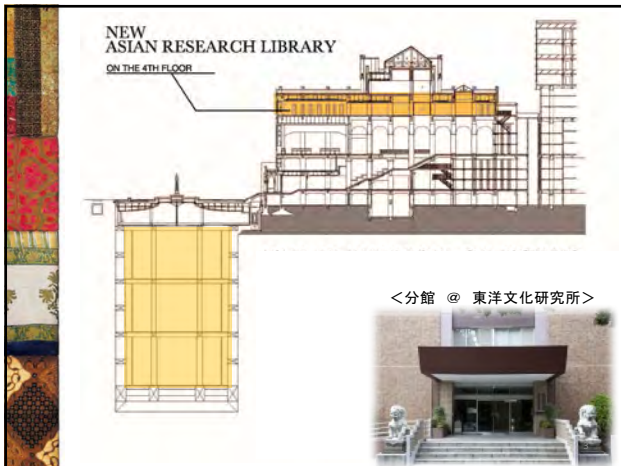
- U-PARLが進めている **flickr** を利用した試みを紹介

デジタル資料公開の ↓
「小さいモデル」「裾野のモデル」を具体的・前向きに考える

高度な・大規模なデジタルアーカイブ構築の動きの一方に、
それにつながる、
簡単で持続性ある資料デジタル化のモデルも求められるのでは？

- * 「アグリゲート」される側の知恵、ハウツーの必用
- * デジタル化の「スペクトラム」的な広がりイメージ

2



さて、どうしよう？

大きな 資金 and/or 人力(知恵・技術 + 労働) の必要性
例・ 2,000点の撮影およびデータベース構築を業者にまとめて依頼して約500万円
複数コレクションを束ねるシステムをつくらせてイニシャル1000万・アニュアル1000万 etc.

- 撮影と画像 △
... 原本アクセス度合、テキストとOCRの問題
- メタデータ ?
...「日本版ヨーロッパアナ」? 標準化? マッピング地獄?
参考: Metadata Standards Crosswalk, Getty Research Institute
http://www.getty.edu/research/publications/electronic_publications/intrometadata/crosswalks.html
- サーバーとパーマリンク ×
... 自前で用意? 維持できるの? レンタルで解決する?
- データベース・ビューアー・公開インターフェイス ×
... お金で解決? 維持は? セキュリティは? どこで検索? DL?

メタデータをめぐる試行錯誤の中で...

「XMLかJSONか云々...の部分はあとからどうにでもなる」
「スプレッドシートでよい」「とにかく区切ってあればよい」
「メタデータの統一基準をつくるのは難しい」
「とりあえず見つければよい」
「その「もの」をちゃんと記述できることが大切」

↓

- アジア研究者の集団として、資料種に即した記述をまず考える
- そのすべてを拾うアンブレラができる・必用だとは思えない
- とりあえず見つけてもらう上で必要なものを落とさないようにしつつシンプルに eg. 漢籍・法帖のメタデータに英語はどこまで必要か?
- 画像とメタデータを自由に使えれば、使用者自ら加工・拡張可

“極論すれば、固有番号と写真があれば、
あとから・多くの人の力で、情報整備は可能”

6

希望と条件はさまざまあれど...

- 安上り・簡単 ↔ 便利・安定
- 量 ↔ 質
- できることから ↔ 恒久性
(パーマリンク、識別子の安定性、全般的持続性...)
- できることから ↔ 拡張性
(テキスト化、TEI、タグ機能、アノテーション機能...)

Q: このどのあたりに「U-PARLのデジタル資料公開」を位置付けるべきなのか？
→ 「小さいモデル」「裾野のモデル」も有意義では？
→ では最低限の要件は何か？

7

一番大事なことは何か？

- 「見つかる」
検索にかかる
- 「使える」
フルサイズJPGをDL可 + メタデータ + ライセンス
→ それぞれの手元で好きに使ってもらえる
- 「持続する」
サーバーやシステムは構築も維持も大変。JPGはしばらく大丈夫
(多分)



8

flickrは...

- 「見つかる」
Googleやflickrで検索可 * BLのタグプロジェクト
- 「使える」
フルサイズで1TBまで無料
メタデータ内容・タグ情報・ライセンス表示可能、ダウンロード可能。
- 「持続する」
データベースやサーバーの管理不要。東大全体の基盤が完成しないうちから稼働し、その後も必要に応じて残せる。拡張可能性もあり。

とはいえ問題も...

- 「見つからない？」 ... flickrの検索機能問題
- 「使えない？」 ... メタデータはテキストで記述
DLするとファイル名変化 ライセンス2.0
- 「持続しない？」 ... 民間のクラウドサービス類にも不安

不満も不安もあり万能策ではないですが、でも...

9

