

Research on Federated Search for Ukiyo-e Prints

「浮世絵の横断検索に関する研究」

Report of Research Results on JSPS
International Training Program (ITP) 2011
2013.02.11

バトジャルガル ビルゲサイハン
Batjargal Biligsaikhan



Goals and Objectives

- ❑ A computer system for accessing multiple Ukiyo-e Databases in the World simultaneously
 - ❑ for **finding information** of **Ukiyo-e prints**
 - ❑ regardless of **format, channel, language** or **location**

❑ ITP Program

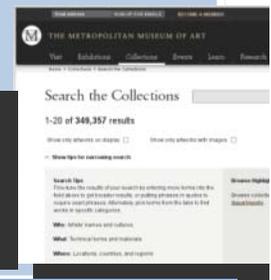
- ❑ **Partner institution** : School of Oriental and African Studies (SOAS), University of London
- ❑ **Duration**: Dec 3, 2011 – Feb 4, 2012 (64days)



Scope and challenges

Simultaneous access to multiple web databases

- **Remote Ukiyo-e databases**
 - Searching and viewing databases in parallel
 - Using a single user-interface
 - Single query input for all databases



Searching databases not in her/his native language

- Finding information in languages that are not familiar with
- Query input in other languages

Existing trends

Gathering multiple databases in advance

- ①
 - Centralized indices
 - Data availability

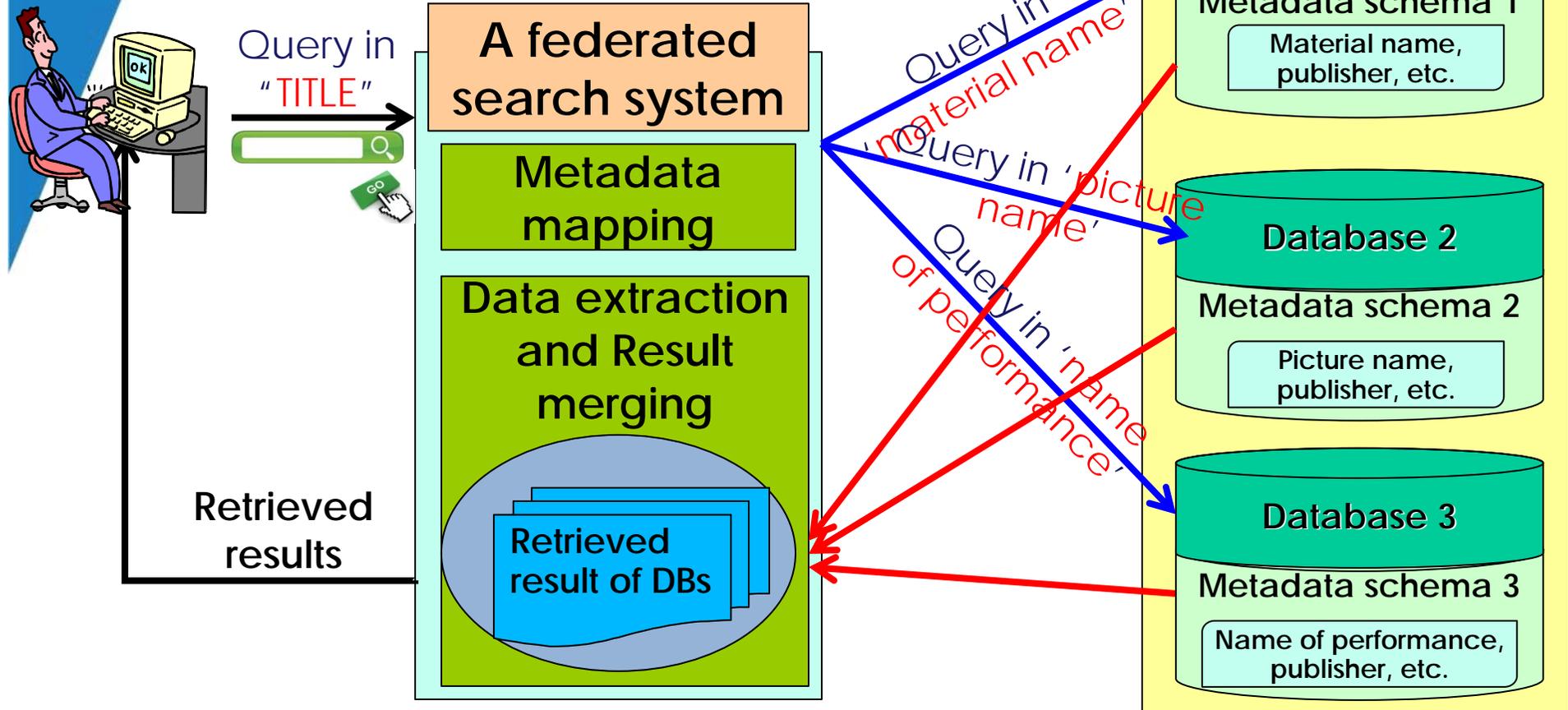
② Accessing in real time

- Federated search
 - **neither** search engine (*web crawler*)
nor *harvester*
 - deep-web, hidden web, dynamic, unlinked content
 - repositories, licensed sources, subscriptions, library data, and password protected databases

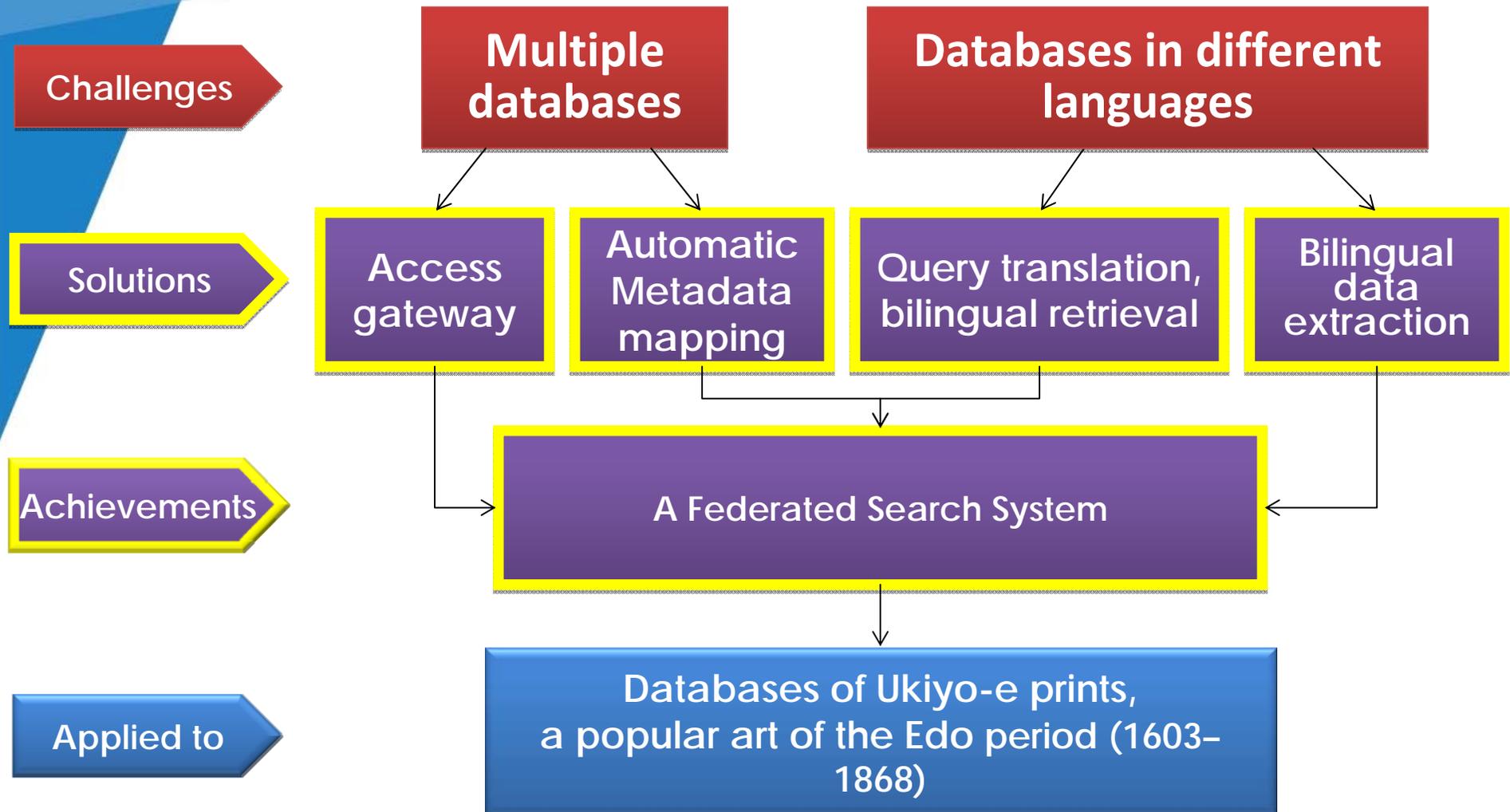
Federated Search concept

Federated search (aka **distributed information retrieval**)

- Accessing or searching across multiple resources simultaneously in one search box in real-time



Research results



① Extracting data

- Cleaning irrelevant HTML tags, JavaScript codes
- Assigning element names

② Identifying & assigning the descriptive metadata

- "title", "author", "subject", "date", "description", "publisher", "type", "identifier", "source", "coverage", "rights"

The Metropolitan Museum of Art The Boston Museum of Fine Arts

作品番号 : arcUP1759	画題等 「忠臣蔵 大尾」
co-setarcUP1748(1) all-setarcUP1748(1)	上演: 嘉永02 . . 江戸 .
絵師: 広重<1> 改印: 馬込、浜	演目:
板元: 佐野屋 喜兵衛	配役:

The British Museum

番号	JP1023
作品名/資料名	No 114, Bikunibashi setchu; Me...
作家/制作者	一遊斎広重(歌川広重)/画
制作年	文政期
所蔵館	江戸東京博物館

The ARC of INSOULINGAL UNIVERSITY

作品番号 : arcUP1759	画題等 : 「忠臣蔵 大尾」
co-setarcUP1748(1) all-setarcUP1748(1)	上演: 嘉永02 . . 江戸 .
絵師: 広重<1> 改印: 馬込、浜	演目:
板元: 佐野屋 喜兵衛	配役:

Chrysa

番号	JP1023
作品名/資料名	No 114, Bikunibashi setchu; Me...
作家/制作者	一遊斎広重(歌川広重)/画
制作年	文政期
所蔵館	江戸東京博物館

Y-O-Ec

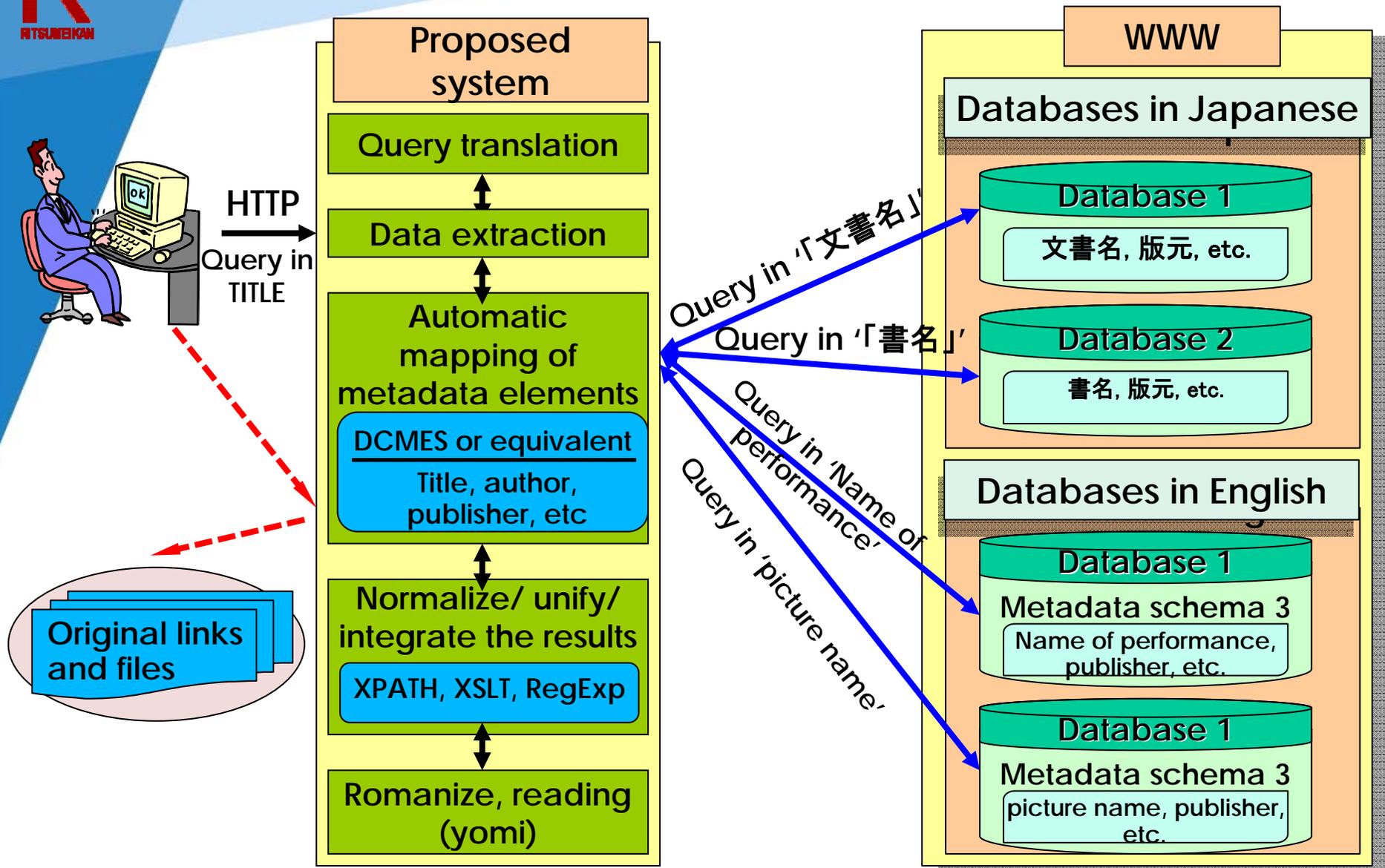
番号	JP1023
作品名/資料名	No 114, Bikunibashi setchu; Me...
作家/制作者	一遊斎広重(歌川広重)/画
制作年	文政期
所蔵館	江戸東京博物館

<title>?

<author>?

絵師, 編著者等, 原所蔵者, 彫師等, 担当者
画題等, 画題2, 役名, 外題, 所作題, 細目題, 主外題, 系統分類題, 演目(統合), 画題統合, 資料名, 解題
版元文字, 異版, 版印1No, 版元1No, 版元1, 版印2No, 版元2No, 版元2, 版元備考, 地域版, 版元統合
西暦, 和暦, 年月日備考, 月日一計算, 西暦版, 和暦版, 月日版, 年月日備考版

The proposed system



Linked Databases

- National Diet Library Digital Archive Portal**
 - Mainly in Japanese, few in English, approx. 10,000 Ukiyo-e print
- The Tokyo-Edo Museum**
 - in Japanese, approx. 6,100 Ukiyo-e prints
- Ukiyo-e Database of Art Research Center, Rits**
 - only in Japanese, approx. 12,500 Ukiyo-e prints

Diverse data
in Japanese

- The Museum of Fine Arts, Boston**
 - Mainly in English, few in Japanese, approx. 57,000 Ukiyo-e prints
- The Victoria and Albert Museum**
 - Mainly in English, few in Japanese, approx. 38,000 Ukiyo-e prints
- The British Museum**
 - In English and Japanese, approx. 15,000 Ukiyo-e prints
- The Metropolitan Museum of Art, New York**
 - In English, approx. 4,000 Ukiyo-e prints
- The Boston Museum of Fine Arts**
 - Mainly in English, few in Japanese, approx. 57,000 Ukiyo-e prints
- The Museum of Art and Archaeology, University of Oxford**
 - Only in English, approx. 640 Ukiyo-e prints
- The Art Institute of Chicago**
 - Only in English, approx. 2,600 Ukiyo-e prints
- Fine Prints: Japanese, pre-1915, Library of Congress**
 - Mainly in English, few in Japanese. Approx. 2,740 Ukiyo-e prints

Diverse data
in English

DEMONSTRATION

- ❑ Please visit a demo website at the <http://ukiyoe.media.ritsumei.ac.jp/>

Federated search for Ukiyo-e prints

- ❑ **Useful to find information from multiple resources**
 - ❑ searching and viewing diverse Ukiyo-e databases
 - ❑ a single query input in a single user-interface
 - ❑ saves users' access time to many databases
- ❑ **Automatic metadata mapping**
 - ❑ diverse data into an unified metadata
 - ❑ providing universal access to multiple databases
- ❑ **Query translation and Data extraction**
 - ❑ Bilingual data



Thank you for your attention !

Questions or Comments?