Case study

3 DH Projects and the Data Modeling of Shanghai Library

Xia Cuijuan (夏翠娟)
Senior Engineer of Shanghai Library's R&D Department
Summary: Architecture of Shanghai Library's DH Knowledge bases

**Basic knowledge bases**
- Name Authority Database
- Geographic Names
- Chinese History Chronology
- Events Knowledge base

**Special Collection knowledge bases**
- Genealogy
- Sheng Xuanhai's Archives
- Ancient Books
- Old Movie
- Manuscripts and Archives for Celebrities

**Rest APIs**
- service for people
- Rest API for machine

Every entity is connected with the others by the Abstract Model.
Case 1: Genealogy (Special Collection)

- One-stop Chinese family search
  60000+ genealogy documents of 597 organizations in the world, about 628 Chinese family names.

- Crowd Sourcing
  - Upload
  - Transcription
  - Create a family tree or a genealogy document

- Social network of librarians, users and researchers
Case 2: Chinese Ancient Books (Special Collection)

- Catalog of 600,000+ existing books held by 1400+ organizations in the world.
- 4 organizations provide scanned images and have their own customized website. The manager of an organization can modify or add their own metadata records and configure their own website.

**An Union Catalog of Chinese Ancient Books**
Case 2: Chinese Ancient Books (Special Collection)

- extract structural data of non-existing ancient books from dozens of catalog books
- mashup every item of the ancient catalog books with the related item of existing organizations' collections
- provide evidence for some specific ancient book about its different versions, the historical annotations, and the metadata records of organizations' collections
Case 2: Knowledge base and tools

Seal knowledge base

machine learning for image recognition
Case 2: Data Modeling

Data Model

ontology and vocab: http://gj.library.sh.cn/ontolgy
Case 3: Name Authority Database
(General and basic knowledge base)

- Amount: 900,000+

Every person has an http URI as the universal identifier and location.

Purpose:

- Name authority control on the web
- Connect different library collections
- Connect resources and entities on the web
- Data source
  - Metadata: authors and contributors or some important people
  - Authority files of libraries
  - Dictionary of names
  - Web Encyclopedias
Case 3: Data Modeling for Name Authority

http://names.library.sh.cn/ontology
Case 3: ongoing——Integrate Data of CBDB