

Collaboration for Digital Archiving and Open Government Data

Mei-Xue Wang

Academia Sinica Research Center for Information Technology Innovation
miseru@iis.sinica.edu.tw

Abstract

Transforming archiving webpages to open data or combining with government data will make information more useful. In this paper, we make the process about how we transform webpages to open data.

1. The National Digital Archiving Project

The National Digital Archiving Project are build for more than ten years. Most digitization resources are came from academia institutes, museums and national profiles. They are produced into text , images, audio or video file type. There are many different websites build for helping users searching and using these resources. One of the portal is : Online Public Access Catalogue for the Digital Resource of the Institute of History and Philology (<http://ndbuilding.iis.sinica.edu.tw/IHPCatalog/>) , which archiving historic books and utensils. The other famous portal websites are : Digital Taiwan - Culture & Nature (<http://culture.teldap.tw/culture/>) 、Academia Sinica Digital Resources (<http://digiarch.sinica.edu.tw/>) ◦

In this process, many technologies have been developed with archiving, for examples : resources categorization 、ancient writing digital identifying 、cross-domain semantic searching 、historical photos(video) searching 、copyright and privacy protection technology ◦

The Website : Online Public Access Catalogue for the Digital Resource of the Institute of History and Philology , contains the following topics :

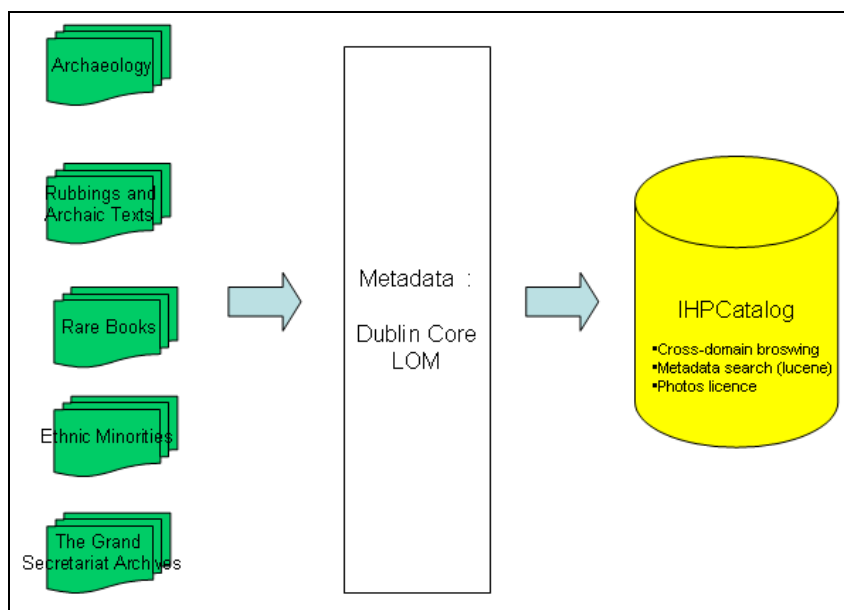
- Archaeology
 1. Relics (Artifact)
 2. Remains
 3. Ruins
- Rubbings and Archaic Texts
 1. Bronze Rubbings collected
 2. Rubbings of Han Dynasty Stone Relief

3. Digital Arhive of the Oracle Bones Rubbing
 4. Han Wooden Slips collected
 5. Digital Archives Project For The Liao-Chin-Yuan Rubbings
- Rare Books
 - Ethnic Minorities
 - The Grand Secretariat Archives



IHPCatalog

Online Public Access Catalogue for the Digital Resource of the Institute of History and Philology



Architecture of IHPCatalog

2. Open Data

In these years , because the simplicity and convenience, many cities build open data providers to let more people using these data easily. Peopel develop mobile APP and website programes to make these data more useful.

The formate of oepn data :

- API
- JSON
- AtomPub
- ATOM
- RSS

Open Data in Taiwan

Taipei City Govermanet Open Data platform (<http://data.taipei.gov.tw/>) collections data from each department. Most of the data are about communication and department statistics profile. If there're any apps about theose data, then that would be put on the platform : <http://apps.taipei.gov.tw/> .

The Government

In the second half 2012 , Government will provide people's livelihood information by the nine major categories: food, medicine, housing, transportation lines, recreation, employment, culture, economic, and life, etc. For example : food security, agricultural products prices, bus dynamic data, real estate prices and information, tourism information, and the financial tax information.

3. Process of archiving websites to open data

- It's not difficult transferring website pages to open data , the processes are :
 1. Key purposes extraction
 2. Metadata (column) mapping
 3. Copyright definition (Open Comments , CC)
- Related API aggregation:

1. Timeline API
2. Geo API
3. Social API
4. News API
5. Photo / Image API
6. Video API

Future Developing

For a archiving project , making data useful is a continue working. How to make historical data and news becoming alive and useful , open data will be a good bridge to work on .