

Providing Value-added Services for Digital Musical Images

Database of Chinese Musical Relics (DCMR) based on CADAL's Resource

Hongjie Sun
Shanghai Normal University
Shanghai, China
wongersan@shnu.edu.cn

Chen Huang
Zhejiang University
Hangzhou, China
beside.huang@gmail.com

Donna Kurtz
The University of Oxford
Oxford, United Kingdom
donna.kurtz@oerc.ox.ac.uk

Joshua Seufert
The University of Oxford
Oxford, United Kingdom
joshua.seufert@bodleian.ox.ac.uk

Dominic Oldman
British Museum
London, United Kingdom
doint@oldman.me.ox.ac.uk

ABSTRACT

This paper^{*} presents a pilot project of CADAL that uses the latest digital technologies of the next generation semantic web to open, link and share data for research and public benefit. The pilot collaborates with the University of Oxford's semantic web pilots and the Mellon Foundation's ResearchSpace project in the British Museum.

KEYWORDS

China, archeology, music, semantic web, Linked Open Data

1 BACKGROUND

This article is about a pilot project of CADAL, the China Academic Digital Associative Library, which was launched by the Chinese government in 2001. Today it has 2.76 million digital records; 2.42 million are online. It is one of the largest digital library projects in China with a range of subjects that includes tangible and intangible cultural heritage. [1] Since the collections of CADAL are continuously increasing in both number and type, there are many digital images of music and dance from ancient China from different periods in history and different sources. Being in books which could not be reached easily outside certain libraries in China, they were largely unnoticed by international scholars and students who are interested in Chinese studies. The situation has not been much improved even after the digitalization of those books, because few people are aware of their existence. The increasing interest in digital cultural heritage coincides with the aim of CADAL to make more special resources related to Chinese culture available to everyone. For that reason, the program of building a database for digital images of Chinese musical relics has been on the agenda since 2015. The database aspires to collect and arrange various kinds of musical images from Chinese history and provide value-added services for them through data optimization and data association.

2 PLANS AND PROGRESS

To build a full-fledged database of Chinese musical relics is a long-term cooperative project; a pilot will show the basic concept and the advantages through the following four stages.

2.1 Pre-process: Resource classification and metadata framework design

The mass of digital musical images from diverse raw resources have been divided into five categories in terms of their subjects and applications: 1) musical instruments, 2) musical scores, 3) musical books, 4) the properties of theatrical performance, and 5) the scenes of musical performance depicted on paintings, murals, cliff paintings, brocades, wares, pictorial bricks and stones, coffins and dancing figurines. [2] Different categories call for different levels of detail in metadata creation. For example, musical instruments need information about shape, structure, ornamentation and craft; musical painting, however, needs size, material, content. Likewise, an object from excavation needs informations about its discovery. Such differences have been considered in building the metadata framework.

2.2 Metadata annotation for sample data

About 800 images of 350 items were chosen as samples for metadata annotation, including 120 musical instruments, 6 musical scores, 4 musical books, 20 performing properties, and 200 performing scenes. Existing descriptions from original sources have been referenced and re-edited according to the new framework of metadata. English translation was given during the process, so as to benefit the international scholars and students. Musical paintings with high resolution have been in process of data optimizing, for instance, to describe significant details in a piece - for example, persons, clothing, instruments, plants, calligraphy and seals in a given musical painting - using additional metadata lists. With this operation, the given record has been given more links to connect to others, so the possibilities of being retrieved has been increased, and the potential academic value of the data has been reinforced as well.

2.3 Data association

^{*}© Hongjie Sun, Chen Huang, Donna Kurtz, Joshua Seufert, Dominic Oldman. This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

The dataset of musical images could benefit from relating its data with that of other collections which have been recorded in CADAL's database, for example, books, photographs, paintings, audio files and so on. In addition, the project wishes to develop collaborations with others internationally to make its resources stronger, more accessible and to be able to link to public resources on the web, for example Wikipedia (DBpedia) and in research institutions, for example the Chinese Biographical Database (CBDB at Harvard and Peking University), OXLOD (Oxford University) and the [ResearchSpace](#) (British Museum), [Chinese Text Project](#), etc.

2.3.1 Collaborations with the University of Oxford. One of the international collaborations has been with Joshua Seufert, the HD Chung Chinese Studies Librarian, Bodleian Library, University of Oxford, who has worked with CADAL for many years. He has been to Zhejiang and CADAL's Huang Chen, Deputy Director of Zhejiang University Libraries and Deputy of the Administrative Centre for CADAL, has been to Oxford. Huang Chen and Dr Hongjie Sun took part in the Professor Kurtz's "*Digital Cultural Heritage China*" workshop in October 2015 [1, 3, 4]. CADAL sent Dr Sun back to Oxford for one year from December 2015 to work with Professors David de Roure, Director of the University of Oxford's e-Research Centre, and Professor Kurtz, Senior Research Fellow of the e-Research Centre. [5, 6] Professor Kurtz and Joshua Seufert have also collaborated on the project Digital Cultural Heritage China and on a Linked Data pilot project in the University of Oxford (OXLOD- Oxford Linked Open Data) - <http://www.oerc.ox.ac.uk/news/oxford-linked-open-data>. Initially OXLOD will map sample data from the University's libraries, museums and selected academic research. International databases were linked in CLAROS (<http://clarosnet.org>) between 2008 and 2011. [3, 5].

2.3.2 Collaborations with the British Museum. Sample data of musical paintings have been linked to ResearchSpace, a digital discovery project platform in the British Museum. The objective of the RS is to build a research environment in which data from different sources and with different perspectives (influenced by different factors including disciplinary, geographical, cultural, etc.) can be integrated without a common fixed central data schema. Since RS itself is an international platform (whose development has been funded by the Mellon Foundation), it is dedicated to integrating various cultural heritage databases and can help CADAL to make its resource better known to the world. The CADAL test case in progress used the following process: 1) Dr Hongjie Sun constructed a small database of manuscripts using Excel spreadsheet. 2) This was converted to XML format and mapped to CIDOC CRM using an online mapping tool, converting the data from a flat field to a semantic hierarchical knowledge graph. 3) The mapped data were transformed into Resource Description Framework (RDF of Linked Data). 4) Additional RDF statements were inferred from the data that abstracted a more generalized framework for use in a more accessible user interface. Entities were abstracted to six fundamental categories: Things, Place, Actor, Time, Event and Concept. Properties were also abstracted to fundamental relationships. 5) The data were uploaded to the ResearchSpace

system. A report referring to the progress of collaborations between CADAL, the University of Oxford's e-Research Centre and the British Museum's ResearchSpace project has been drafted by Dr Hongjie Sun and his colleagues: Donna Kurtz of Oxford e-Research Centre, Dominic Oldman of British Museum, and Joshua Seufert of Oxford University's HD Chung Chinese Studies Librarian, Bodleian Library. [6]

2.4 Open for expertise

The database will be open to the public. Subject experts will be able to contribute their knowledge to the platform. The public will be able to deliver their comments, corrections or suggestions.

3 VISIONS

Providing value-added services for digital musical images will make data more accessible to musicologists, archaeologists and others who want to know more about Chinese music and its history. The following work will be carried out in the future for that purpose: 1) More examples of musical relics will be added to the pilot project, a special section focused on the images of the Guqin (古琴) - a traditional Chinese musical instrument, will be a highlight for demonstration. 2) Further collaboration with ResearchSpace project and with the University of Oxford's e-Research Centre. Building on three years of working together with OERC, CADAL has an opportunity to take part in OXLOD, Oxford Linked Open Data. The pilot project is dedicated to connecting sample data from the University of Oxford's Museums and Collections. CADAL is ready to share sample data of Chinese musical images to OXLOD, and to explore the possibilities of linking data between digital Chinese collections.

REFERENCES

- [1] Chen Huang and Hongjie Sun. October, 2015. The Work of CADAL and Digital Libraries in China. In *Digital Cultural Heritage China Workshop*. University of Oxford, Oxford, UK.
- [2] Hongjie Sun. June, 2017. A Case for Providing Value-added Services for Digital Art Collections: The Collaborations Related to CADAL's Resource Between China and the UK. In *The International Conference on Digital Resources Sharing and Application (DRSA 2017)*. Zhejiang University, Hangzhou, China.
- [3] Donna Kurtz. October, 2015. CLAROS: Linked Data Pilot. In *Digital Cultural Heritage China Workshop*. University of Oxford, Oxford, UK.
- [4] Dominic Oldman. October, 2015. Opening, Linking and Sharing Heritage Data: Museums and Libraries. In *Digital Cultural Heritage China Workshop*. University of Oxford, Oxford, UK.
- [5] Donna Kurtz, Dominic Oldman, Joshua Seufert, Hongjie Sun. July, 2016. Linked Data: An Institutional Approach. In *Digital Humanities at Oxford Summer School 2016 (DHOxSS 2016)*. University of Oxford, Oxford, UK.
- [6] Donna Kurtz, Hongjie Sun, Dominic Oldman, and Joshua Seufert. 2016. The China Academic Digital Associative Library (CADAL) and the University of Oxford: A Semantic Web Representation of Classical Chinese Paintings (CADAL). 12 pages.