

ChinaDigitalMuseumProject

Contents

- 1 [The China Digital Museum Project](#)
 - 1.1 [Approach](#)
 - 1.2 [Status](#)
 - 1.3 [Other tools](#)
 - 1.4 [Contact](#)

The China Digital Museum Project

The Chinese Digital Museum Project is a collaborative project between the Chinese Ministry of Education, Hewlett-Packard Company and several Chinese universities, including Beijing Normal University and Beihang University.

[Presentation at second DSpace user group meeting](#)

Many universities in China have one or more museums. In order to improve access of research communities, students/teachers and the general public to the artefacts in these museums, these universities are undertaking the digitisation of those artefacts. The principal aim of the project is to enable these universities to provide infrastructure based on DSpace to store, manage, preserve and disseminate the digitised versions of the artefacts. In the final phase of the project, there will be around 100 university museums with digital artefacts stored in federated DSpace installations.

In fulfilling this requirement, we will need to address many problems associated with managing distributed digital asset management, including persistent identifiers, metadata standardisation, deployment processes and management, and content and metadata replication.

Another challenge that this project faces is how, given the vast diversity of these digital museums across China, to consolidate the digital assets that they each manage so that users do not have to know where the assets are being stored in which website, but rather find and search them on by subject, such as "Chinese Calligraphy Paintings". Hence we aim to create "Virtual Museums", formed by arrangements of digital assets by subject, regardless of physical location.

Approach

The target solution is a large-scale, federated deployment of DSpace to serve the Chinese universities' digital museums and end users:

- Each university museum runs a local instance of DSpace to hold the digitised content from its local holdings, and associated metadata.
- Two "data centres" running large-scale DSpace instances will each hold a replicated copy of all the digitised content and metadata from all of the museums for long-term preservation purposes, and to provide the foundation for virtual museums.

In order to federate the metadata and content, METS packages will be harvested by the data centres from the university museums via the OAI-PMH protocol. The METS packages will contain references to the constituent content files, enabling the data centre to retrieve and replicate the full content of the university museums.

In addition, a central registry tool is being developed to:

- Keep track of DSpaces in the federation
- Allow those DSpaces to discover other DSpaces in the federation
- Provide a 'select mirror' tool when accessing objects

Status

Single university museum DSpace and data centre up and running and sharing content.

Development of enhancements progress shown above.

- Chinese language translation of user interface. Almost complete.
- (/) Search of Chinese metadata and text. Complete, [patch available](#) also allowing plug-in of different Lucene Analysers
- OAI-PMH harvester/client - Build on [existing client from Old Dominion University and Los Alamos National Laboratories](#).
- Supporting extended metadata - "Pretend" its Dublin Core to start; build on [metadata support extension](#) proposed by HP Labs, developed by Martin Hald at Eduworks (patch available, currently in beta)
- METS Importer tool - in prototype
- (/) Create Handles using remote Handle Server - [patch available](#) – fixing a scaling problem – Handles need to be created in batches
- (/) Configurable item display page. complete, in DSpace 1.3.2

Other tools

For testing DSpace at large scale, a large data set is needed. A quick way to fake one is to import the same data set several times over. However, you really need at least the titles to be different for each item.

The following modified (basic, non-METS) item importer slightly morphs metadata on the way in, so that each time it's run, you'll get a slightly altered set of items. Right now, it just adds increasing numbers to the titles, but the technique could be adapted for more sophisticated morphing if needed.

[\[ItemImport.java\]](#)

Contact

[James Rutherford](#)