

CADAL项目： 面向数字人文的 金石拓片系统平台构建

聂华 北京大学图书馆

基于CADAL资源的特藏建设：
蒋介石资料数据库建设学术研讨会
2019.02.16 @杭州

背景.现状.需求

- 1 北大馆馆藏资源
- 2 北大馆平台建设
- 3 标准规范
- 4 问题需求



北京大学金石拓片馆藏资源情况

- 北京大学图书馆收藏有近代两位著名藏书家繆荃孙艺风堂、张仁蠡柳风堂的全部藏拓，加上本馆自主采访金石拓片资料等（以下简称馆藏号拓片），北大馆藏金石拓片资料（含拓轴、拓本等）共40000余种、80000余份，藏量仅次于国家图书馆和上海图书馆，在全国范围的图书馆、博物馆等机构的金石拓片收藏中占有重要的一席之地。

北大图书馆金石组(胡海帆、汤燕等)编纂：

- 《1996-2012北京大学图书馆新藏金石拓本菁华》（北京大学出版社），包括金器类、砖陶类、碑志类和法帖类四类拓本，公布了242种（281品）新入藏的金石拓本资料。
- 《中国古代砖刻铭文集》、《北京大学图书馆藏历代金石拓本菁华》



北京大学金石拓片数字资源建设情况

- 2003年至今，已完成扫描：
 - 墓志拓片：15000余份、20000余拍，
 - 北京地区拓片：2889种、4000余拍，
 - 艺风堂：历代拓片8618种、11000余拍
 - 其他专藏零种拓片：500余种、5000余拍；
- 尚未完成扫描的有：1152种艺风堂拓片、10816种柳风堂拓片、11596种馆藏号拓片以及拓本若干。
- 剩余部分拓片扫描计划：
 - 2018年底艺风堂藏拓全部扫描完毕
 - 2020年底之前柳风堂藏拓全部扫描完毕，
 - 2021年开始扫描馆藏号拓片。另，每年新购入约500种左右拓片资料应于当年年底之前扫描完毕。



金石拓片数据建设及相关标准规范

- 金石碑拓数字化和数据建设已建立一系列的数据标准规范体系，目前进行数字化处理建设主要有国家图书馆、北京大学图书馆、浙江大学图书馆、台湾中研院以及哈佛大学、加州大学伯克利分校等机构。
- 北京大学图书馆从2001年开始进行拓片元数据规范研究并形成科技部相关规范标准。



北京大学金石拓片数字资源建设情况

- 从2001年北京大学图书馆开始进行拓片元数据规范研究并形成科技部相关**规范标准**的同时，**数字化工作**持续开展并取得很大进展。在CALIS项目支持下**建成秘籍琳琅拓片数据库**
- <http://rbd.calis.edu.cn/aopac/pages/Browse.htm>
- 然而，由于包括技术原因在内的各种因素限制，现有资源平台大部分只能提供传统的浏览、检索和有限精度的图像查看或图标预览功能，不能满足数字人文研究背景下学者对资源利用的进一步需求。



问题和需求

- 数字化的数量不够：各大机构数字化加工和酝酿很多年了，但是由于多种因素加工总量还远远不够，数字化工作仍需投入大量资源继续推进；
- 数据标准不统一：目前主要的元数据标准是国家图书馆标准、北京大学标准、台湾中研院标准。因为各自考量的因素不一，虽然核心元数据大致一致，但是描述各有侧重不同；
- 无法满足数字人文研究环境下的用户需求：无法支持数字人文研究环境下学者进行资源检索和获取、图像高精度转换查看、历史时间和空间层次检索、知识挖掘、关联网络分析等进一步需求。

系统平台建设规划

- 1 年度建设目标
- 2 平台建设内容
- 3 项目实施计划
- 4

1 本年度建设目标

- 构建金石拓片系统平台，推动北大图书馆拓片资源的数字化加工和整合；
- 基于国家图书馆标准、北京大学标准等设计平台通用元数据标准，以北大元数据标准为例实现转换互通，完成数据整理、清洗、加工完善和导入；
- 以满足数字人文需求为导向开发面向用户的服务功能，包括资源检索和获取、图像高精度转换查看、历史时间和空间层次检索、外部数据关联融合等。
- **推动金石拓片资源的保护，共享，开放。**

2 项目建设内容

(1) 完成平台元数据方案

(2) 完成数据整理、清洗、加工完善和导入

(3) 完成后台数据管理功能开发

(4) 完成面向用户的服务功能开发

- 资源检索与获取基本功能：资源浏览、全文检索
- 资源检索与获取增强功能：高清图像查看和互动，基于GIS技术的时空检索功能
- 外部数据整理加工、外部数据融合功能：通过接口引入外部数据库（例如CBDB等），对拓片内容进行标记和链接

3 项目实施规划Roadmap

2019年

4月

8月

12月

(1) 完成样本范围选定、数据导出、平台元数据方案

2019年4月底之前

(2) 完成数据整理、清洗、加工完善和导入

2019年8月底之前

(3) 提出后台数据管理功能需求、完成功能开发

2019年8月底之前

(4) 完成面向用户的服务功能开发

2019年9月底之前

- 资源检索与获取基本功能：资源浏览、全文检索
- 资源检索与获取增强功能：高清图像查看和互动，基于GIS技术的时空检索功能
- 外部数据整理加工、融合功能：通过接口引入外部数据库（例如CBDB等），对拓片内容进行标记和链接

(5) 测试修改、项目结题报告

2019年11月底之前

4 讨论

