



广西壮族自治区
桂林图书馆



唯一标识符注册与维护项目

——广西桂林图书馆2014年度建设情况



广西桂林图书馆 白薇

2015.12 北京

唯一标识符注册与维护项目建设情况



项目完成情况

▶元数据注册9158条

▶<http://202.103.241.77/cdoi/PB/DA0101InitAction.do>

The screenshot displays the '元数据注册' (Metadata Registration) interface of the '广西桂林图书馆唯一标识符系统' (Guangxi Guilin Library Unique Identifier System). The page shows a search form at the top and a table of registered metadata records below. The table has columns for '唯一标识符' (Unique Identifier), '题名(来源名称)' (Title (Source Name)), '文件格式' (File Format), '资源类型' (Resource Type), and '基本操作' (Basic Operations). The records list various historical and organizational entities, such as '中共梧州市委员会(1928.4...)', '中共广西善工委(1936.11...)', and '中共广西特委(1929.1-1...)', among others. The interface also includes a sidebar with navigation options like '元数据注册', '元数据维护', and '元数据删除'.

唯一标识符	题名(来源名称)	文件格式	资源类型	基本操作
108.ndic.33.4503049031010001/78F35.20120813153622m1	中共梧州市委员会(1928.4...	XML	网页采集	[Edit] [Delete]
108.ndic.33.4503049031010001/78F35.20120809160743m1	中共广西善工委(1936.11...	XML	网页采集	[Edit] [Delete]
108.ndic.33.4503049031010001/78F35.20120809160310m1	中共广西特委(1929.1-1...	XML	网页采集	[Edit] [Delete]
108.ndic.33.4503049031010001/78F35.20120809155217m1	中共广西临时特委(1928.9...	XML	网页采集	[Edit] [Delete]
108.ndic.33.4503049031010001/78F35.20120809154656m1	中共广西特委(1928.1-1...	XML	网页采集	[Edit] [Delete]
108.ndic.33.4503049031010001/78F35.20120809154243m1	中共广西地委(1927.5-1...	XML	网页采集	[Edit] [Delete]
108.ndic.33.4503049031010001/78F35.20120813154448m1	中共田南临时特别委员会(192...	XML	网页采集	[Edit] [Delete]
108.ndic.33.4503049031010001/78F35.20120813155949m1	中共滇黔边区委员会(1936...	XML	网页采集	[Edit] [Delete]
108.ndic.33.4503049031010001/78F35.20120813160334m1	中共桂西区特别委员会(1937...	XML	网页采集	[Edit] [Delete]
108.ndic.33.4503049031010001/78F35.20120813160846m1	中共郁江区特别委员会(1937...	XML	网页采集	[Edit] [Delete]
108.ndic.33.4503049031010001/78F35.20120813161206m1	中共浔江区特别委员会(1937...	XML	网页采集	[Edit] [Delete]
108.ndic.33.4503049031010001/78F35.20120813162200m1	中共桂西南特别委员会(1941...	XML	网页采集	[Edit] [Delete]
108.ndic.33.4503049031010001/78F35.20120813162453m1	中共柳州市委员会(1941.3...	XML	网页采集	[Edit] [Delete]
108.ndic.33.4503049031010001/78F35.20120813162714m1	中共柳州特别委员会(1940...	XML	网页采集	[Edit] [Delete]
108.ndic.33.4503049031010001/78F35.20120813163219m1	中共桂北临时联合工作委员会(1...	XML	网页采集	[Edit] [Delete]
108.ndic.33.4503049031010001/78F35.20120813163440m1	中共靖镇边区工作委员会(194...	XML	网页采集	[Edit] [Delete]
108.ndic.33.4503049031010001/78F35.20120813163913m1	中共钦廉西属工作委员会(194...	XML	网页采集	[Edit] [Delete]
108.ndic.33.4503049031010001/78F35.20120813164310m1	中共十万山地方委员会(1949...	XML	网页采集	[Edit] [Delete]





- ▶ 文件解读，明确目标
- ▶ 拟定工作方案
- ▶ 培训，掌握要点
- ▶ 实施
- ▶ 质检
- ▶ 完成，提交成果

唯一标识符注册与维护项目建设情况



文件解读，明确目标（做什么，怎么做）

《关于开展数字图书馆推广工程2014年度数字资源联合建设工作的通知》

《推广工程数字资源联合建设唯一标识符注册与维护项目建设方案》

《唯一标识符注册数据著录规则》

国家图书馆

国图函〔2014〕56号

关于开展数字图书馆推广工程 2014年度数字资源联合建设工作的通知

图书馆：

为进一步加大数字图书馆推广工程（以下简称“推广工程”）在资源建设方面的实施力度，集全国公共图书馆之力共建优秀数字资源，2014年推广工程将继续开展数字资源联合建设项目。根据《财政部关于下达2014年公共文化建设专项资金的通知》（以下简称《通知》），国家向省级图书馆每馆下达50万元专项资金，部分市级图书馆每馆下达20万元专项资金，用于各馆开展推广工程数字资源联合建设工作。按文化部要求，国家图书馆负责数字资源联合建设工作的牵头管理与组织实施。经文化部批准，具体通知如下：

一、建设目标

以“合作共建”和“合理共享”为原则，将各地区分散、异构的资源有效集中与整合，有重点地共建一批主题明确、特色鲜明、类型丰富、组织有序的优秀数字资源库，使我国数字图书馆资源建设和保障的整体水平得到跨越式提升。

- 1 -

附件3

推广工程数字资源联合建设 唯一标识符注册与维护项目建设方案（2014）

为加强我国公共图书馆数字资源的管理能力，促进全国公共图书馆数字资源的共建共享，推广工程联合全国各省、市级公共图书馆共同开展全国数字资源唯一标识符注册与维护工作，共同建设公共图书馆数字资源唯一标识符体系。

一、建设内容

公共图书馆数字资源唯一标识符注册与维护项目主要围绕全国公共图书馆已发布的馆藏自建资源进行注册。资源类型包括但不限于文本资料、数据集、图书、照片、连续出版物、音频资源、视频资源、软件、抽象作品、艺术作品等。本项目将其归纳为八类，分别是图书、期刊、论文、报纸、图片、音频、视频，以及网页采集数据。

二、工作流程

1. 数据准备

各图书馆应按附件“唯一标识符注册数据著录规则”提交 excel、xml 或 access 格式的唯一标识符注册数据。注册数据模板可以在唯一标识符系统【元数据注册（批量）】页面下载。

2. 唯一标识符注册

28

附录

唯一标识符注册数据著录规则

一、数据质量总体要求

1. 唯一标识符注册数据应完整、准确、命名规范，不得出现错别字，数据组织合理；
2. 无IP地址的唯一标识符数据需要在本项目规定时间内补充完整；
3. 进行全部数据的检查，错误率不得超过万分之三。

二、唯一标识符注册数据标准

1. 注册数据

唯一标识符的标识对象包括多种形式，如：单册图书和一套丛书；图书的章节；期刊；期刊中的一期；一期中的目次、一篇论文；一篇论文的几页、单页；单页中的一幅图片、公式；多媒体资源集合；一首乐曲；乐曲中的一段音乐等。本项目中，资源的标识粒度以“种、卷（册/首/幅）”为原则。以下注册此事将按照此原则进行说明。

本规范包含唯一标识符注册数据的23个元素，其中元数据注册字段19项，URL地址注册字段4项，分别描述如下。

2. 元数据注册字段

注册文件模板请参见唯一标识符系统【元数据注册】菜单中【批量元数据注册】字表的“模板下载”项。元数据注册字段共有19项，具体描述如下：

表1 元数据注册字段说明表

序号	数据元素	注册文件表 名称 对应字段	字段注释	必 备 性	可 重 复 性
1	编目系统系 统号	系统号	按照资源对应的编目系统号 加工流水号，若无，则使用 资源来源系统流水号。	必备	不可重复
2	MARC 记 录号	MARC 记 录号 001 标识号	MARC 记录 001 字段	有 必 备	不可重复

29



广西壮族自治区
桂林图书馆



明确几个要点：

馆藏自建资源

已发布（互联网访问）

资源类型：

图书、期刊、论文、报纸、图片、音频、视频、网页





对我馆自建数字资源进行梳理

- 我馆数字资源建设始于**20世纪90年代初**
- **1992年**，开始书目数据库建设，
- **1997年**，开始制作全文数据库
- **2002年**，引进**CDI CM**内容管理系统，开始专题资源库建设
- 已建成的数字资源包括：

书目数据库（题录数据库、目录数据库、索引与文摘数据库）

全文数据库

集文字、图片、音频、视频为一体的多媒体数据库



唯一标识符注册与维护项目建设情况



- 广西壮族自洽区地方志述评
- 省志
- 河池地区
- 南宁市
- 桂林市
- 玉林地区
- 南宁地区
- 桂林地区
- 钦州地区
- 柳州市
- 梧州市
- 百色地区
- 柳州地区
- 梧州地区



序号	题名	作者	出处
1	教育行政计划重点及实施要领 (桂图)	黄朴心	广西教育, 1942年第1卷第1期
2	中学英语教学研究 (桂图)	李业武	广西教育研究, 1941年第2卷第1期
3	社会的不安与教育的不安 (桂图)	康德尔	广西教育, 1942年第1卷第1期
4	新子的教育思想 (桂图)	杨保国	广西教育, 1942年第1卷第1期
5	中学生学习指导 (桂图)	姚应华	广西教育, 1942年第1卷第1期
6	中学学校学生组织活动的领导问题 (桂图)	李福麟	广西教育, 1942年第1卷第1期
7	中学学校导师制之理论与实际 (桂图)	环德善	广西教育, 1942年第1卷第1期
8	广西省中等学校学生身体素质检查的分析研究 (桂图)	李行健工作室	广西教育, 1942年第1卷第1期
9	关于广省中等学校教师的调查问题 (桂图)	汤有耀	广西教育, 1942年第1卷第1期
10	中学学校语文教学的一个实例 (桂图)	莫奇	广西教育, 1942年第1卷第1期
11	精英团 (桂图)	云从龙	广西教育, 1942年第1卷第1期
12	广西省教育后之教育概况 (桂图)		广西教育, 1942年第1卷第1期
13	三十五年度第一学期中等教育概况 (桂图)		广西教育, 1942年第1卷第1期
14	三十五年度广西省教育文化费占总经费百分比 (表)		广西教育, 1942年第1卷第1期
15	三十五年度广西省教育文化费之分配 (表)		广西教育, 1942年第1卷第1期

桂林旧影

社会政治 建筑遗址 人物春秋 风景名胜

【资源类型】 社会政治、建筑遗址、人物春秋、风景名胜
地址：广西桂林市安铁北路1号 电话：0773-3841387、3844047

山水中的古老文化 桂林石刻

桂林石刻简介

“唐宋题名之渊藪，以桂林为甲。”
这是一百多年前，清代著名金石学家叶昌炽对桂林石刻的评价。桂林石刻是国内外具有重要影响的摩崖石刻，它以历史悠久、数量众多、内容丰富、精品纷呈而闻名天下。桂林石刻分布于奇山秀水之间，二千多件摩崖石刻构成了桂林山水一道亮丽的人文风景线，成为这座古者的历史文化名城的重要组成部分。

返回首页 | 普陀山月牙山石刻 | 南溪山石刻 | 象鼻山石刻 | 虞山石刻 | 伏波山石刻 | 叠彩山石刻 | 西山隐山石刻 | 独秀峰石刻 | 其它
地址：广西桂林市安铁北路1号 电话：0773-3841387、3844047 < 桂ICP备0800626号 >



广西师范大学
桂林图书馆

唯一标识符注册与维护项目建设情况





拟定工作方案

确定实施部门：馆领导挂帅，信息技术部具体负责

确定工作量：1万条

确定工作流程：

唯一标识符系统部署——培训——选取数字资源——提交样例数据——唯一标识符注册——对象数据URL地址管理——质检——提交数据

确定建设周期：3个月完成



唯一标识符注册与维护项目建设情况



培训，掌握要点

2014年12月初，国家图书馆老师到馆部署系统

2014年12月10-11日，参加数字图书馆推广工程2014年度数字资源联合建设培训班

2015年1月，培训相关工作人员





实施

成立建设小组：4人

1名副主任：总体负责，质检

1名计算机技术人员：参数设置、系统维护

2名资源加工组人员：元数据注册

注册方式选择：

两种注册方式：手工单条注册、批量导入注册





手工单条注册页面

系统菜单 << 您的当前位置：元数据注册>单笔元数据注册

元数据注册 >>

>元数据注册

>元数据注册(批量)

单笔元数据注册

资源类型	<input type="text"/>	*	数据来源系统	<input type="text"/>	*
文件格式	<input type="text"/>	*	系统号	<input type="text"/>	*
MARC记录001标识号	<input type="text"/>		题名(资源名称)	<input type="text"/>	*
作者	<input type="text"/>		ISBN	<input type="text"/>	
ISSN	<input type="text"/>		语种	<input type="text"/>	
出版者	<input type="text"/>		出版时间	<input type="text"/>	
关联	<input type="text"/>				
来源	<input type="text"/>				
描述信息	<input type="text"/>				
URL	<input type="text"/>				

确定

元数据维护 <<

标识符解析 <<

系统管理 <<



唯一标识符注册与维护项目建设情况



批量导入注册excel模板

1	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
2	资源种类编号	格式编号	系统号	颗粒度K1	K1值	颗粒度K2	K2值	颗粒度K3	K3值	MARC记录00	题名(资源名称作者)	ISBN	ISSN	语种	出版社	出版时间	关联	来源	描述信息	扩展字段1	扩展字段2	扩展字段3	
3																							
4																							
5																							
6																							
7																							
8																							
9																							
10																							
11																							
12																							
13																							
14																							
15																							
16																							
17																							
18																							
19																							
20																							
21																							
22																							
23																							
24																							
25																							
26																							
27																							
28																							
29																							
30																							
31																							
32																							
33																							
34																							
35																							
36																							
37																							
38																							
39																							
40																							
41																							
42																							
43																							
44																							



唯一标识符注册与维护项目建设情况



最初，采用手工单条注册，2周注册了大约**2000**条数据

The left screenshot shows the 'Single Record Registration' (单笔元数据注册) form in the 'Guangxi Library Unique Identifier System'. The form includes fields for Resource Type, File Format, MARC Record 001 Number, Author, ISSN, Publisher, Associated, Source, and Description. A 'Confirm' (确定) button is at the bottom.

The right screenshot shows the details for the '1st Guangxi Provincial Party Congress' (中共广西省第一次代表大会). The details are as follows:

题名:	中共广西省第一次代表大会
作者:	广西大百科全书编纂委员会编
出版者:	中国大百科全书出版社
出版地:	北京
出版日期:	2008.11
主题关键词:	广西红色历史文化 建党沿革 重要会议 中共广西省第一次代表大会
文献出处:	广西大百科全书: 历史 (下)

Below the details is a paragraph of text describing the congress: '中共广西省第一次代表大会, 1929年9月10~14日在南宁津头村召开。出席大会代表19人, 代表全省420名党员。中共广贺昌出席大会并作政治报告, 广西特委书记何晋达作工作报告。大会通过《中国共产党广西省第一次代表大会政治任务决议、宣传、工运旅运、军事、土地革命、妇女、共青团8个工作草案大纲》。《决议案》提出今后工作的总路线是利用一切反动领导广大群众的日常斗争, 逐步深入开展土地革命、职工运动、士兵运动、农民武装暴动, 推翻国民党的统治, 创建红军。斗争策略是深入进行土地革命宣传, 争取广大群众, 扩大党的政治影响, 发展党和群众组织, 注重以贫雇农为主要同盟者。强调'目前农村中的主要路线是领导农民做日常的斗争, 从斗争中转变到深入土地的斗争, 尤其扩大农民的武装到游击战争, 使土地斗争更加深入。'大会选举产生新的省特委, 何晋达仍任书记。大会强调要把全省工作重点放在南宁、三大城市及西江流域城镇。9月下旬, 俞作柏、李明瑞仓促举兵反蒋, 中共中央代表邓小平果断地把广西特委的主要干部冯、严敏、何建南、麦锦汉等调到左右江开展创建红军和革命根据地工作, 在实际行动上纠正了省'一大'的城市中心论偏。'

At the bottom of the right screenshot, there is a navigation bar with links: '返回首页 / 建党沿革 / 革命武装 / 历史事件 / 农民运动 / 陵园墓碑 / 纪念馆 / 文物遗址 / 抗战文化 / 八桂英烈 / 红色追忆 / 红色旅游' and a copyright notice: '版权所有: 广西桂林图书馆 <桂ICP备10002506号> 联系电话: 0773-3844047'.





单条注册的优点：

- 1、可以多人同时操作
- 2、在注册的同时能够发现已建资源出现的问题，可及时进行修改

问题：

- 1、人工加工容易发生错误
- 2、效率较低



解决技术难点，批量导入

我馆数字资源加工使用**CDI CM**资源加工和发布系统

使用**XML+XSL**展示数据

用 **C#**制作一个网站元数据抓取程序，把准备注册数据的元数据自动抓取至**EXCEL**模板里

该程序可以根据数据类型的不同，自行判断转化为导入时所用值

例如：视频资源格式为**WMV**，自动抓取到**EXCEL**模板时，格式编号自动转为**F39**



唯一标识符注册与维护项目建设情况



全国文化信息资源共享工程

广西桂林分中心资源检索

检索结果

元数据检索

检索词	检索点	检索途径	匹配关系	所在 单位	关系词
广西红色历史文化	全部	全部	前方一致		与
	全部	全部	前方一致		检索



CDI内容抓取

网址 <http://202.103.241.70:4237/UserCenter/usercenter?querytype=2&querypolicyid=20090326130948&type=get> 快速选择 民国报刊

获取底层内容

抓取当页 抓取页面至末尾

开始抓取

抓取到Excel

抓取到MDB

数据抓取



广西壮族自治区
桂林图书馆

唯一标识符注册与维护项目建设情况



2天注册了7000余条数据，加快了工作效率，减少了加工过程中由于人工产生的错误。

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
1	资源种类编号	格式编号	系统号	颗粒度K1	K1值	颗粒度K2	K2值	颗粒度K3	K3值	MARC记录00	题名(资源名称)	作者	ISBN	ISSN	语种	出版者	出版时间	关联	来源	描述信息	扩展字段1	扩展字段2	扩展字段3	扩展字段4	扩展字段5
2	T8	F35	20120813153622	T8K1V1	1						中共梧州市委	广西大百科全书编纂委员会编			中文(简体)	2008-11			广西桂林图书馆	20120813153622					广西红色历史文化
3	T8	F35	20120809160743	T8K1V1	1						中共广西省工	广西大百科全书编纂委员会编			中文(简体)	2008-11			广西桂林图书馆	20120809160743					广西红色历史文化
4	T8	F35	20120809160310	T8K1V1	1						中共广西省委	广西大百科全书编纂委员会编			中文(简体)	2008-11			广西桂林图书馆	20120809160310					广西红色历史文化
5	T8	F35	20120809155217	T8K1V1	1						中共广西临时	广西大百科全书编纂委员会编			中文(简体)	2008-11			广西桂林图书馆	20120809155217					广西红色历史文化
6	T8	F35	20120809154656	T8K1V1	1						中共广西特委	广西大百科全书编纂委员会编			中文(简体)	2008-11			广西桂林图书馆	20120809154656					广西红色历史文化
7	T8	F35	20120809154243	T8K1V1	1						中共广西地委	广西大百科全书编纂委员会编			中文(简体)	2008-11			广西桂林图书馆	20120809154243					广西红色历史文化
8	T8	F35	20120813154448	T8K1V1	1						中共田南临时	广西大百科全书编纂委员会编			中文(简体)	2008-11			广西桂林图书馆	20120813154448					广西红色历史文化
9	T8	F35	20120813155949	T8K1V1	1						中共滇黔桂边	广西大百科全书编纂委员会编			中文(简体)	2008-11			广西桂林图书馆	20120813155949					广西红色历史文化
10	T8	F35	20120813160334	T8K1V1	1						中共桂西特	广西大百科全书编纂委员会编			中文(简体)	2008-11			广西桂林图书馆	20120813160334					广西红色历史文化

唯一标识符导入截图





批量注册需注意的问题：

- 1、元数据批量注册与**WEB**页面上注册有所不同，需要先对元数据批量导入注册唯一标识符后，再导出元数据的唯一标识符，随后才进行**URL**的批量增加。
- 2、出版时间应设置为文本格式，否则导入后会自动转化成数字。



URL批量导入时遇到的问题：

导入URL时有可能遇到同名而资源类型不同的情况，例如：

元数据信息一览					
		总计2条记录			
<input type="checkbox"/>	唯一标识符	题名(资源名称)	文件格式	资源类型	基本操作
<input type="checkbox"/>	108.ndlc.33.4503049031010001/T7F39.20130206153031m1a2	桂林文化城	WMV	视频	  
<input type="checkbox"/>	108.ndlc.33.4503049031010001/T8F35.20121022110339m1	桂林文化城	XML	网页采集	  

两条数据题名同为“桂林文化城”，而一个为WMV视频格式，一个为XML数据。导入时由于排序问题可能致使两条资源的URL地址互换，这就需人工进行干预。



系统方面问题：

唯一标识符系统批量处理通常能在几分钟能完成，但有时也会发生程序卡死的情况。文件上传成功后，批量处理结果长时间未有完成，需要重启唯一标识符系统或者服务器后才能继续完成导入工作。

元数据批量注册	已处理	注册成功数量为{0}注册失败数量为{1}	2015-04-01 17:02:17	2015-04-02 08:41:46	 
元数据批量注册	已处理	注册成功数量为{174}注册失败数量为{0}	2015-04-01 09:09:52	2015-04-02 08:38:53	

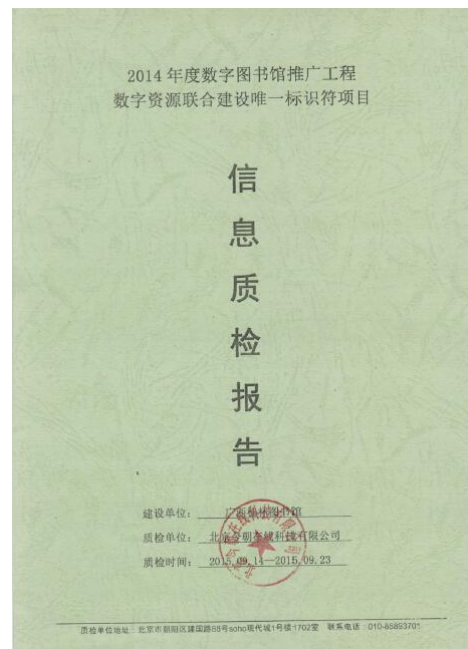


质检

建设小组成员互检——主任审核——第三方质检

质检过程中发现的问题：

部分信息作者字段著录不规范，
多著录了“题、撰”等责任方式；
少数信息出版时间错误；
少数信息出版社错误；





几点体会：

- 1、馆领导重视
- 2、成立工作小组，明确目标与责任
- 3、遇到问题，及时与国图老师沟通
- 4、数据质量是关键（元数据完整、准确、规范）
- 5、批量导入优于手工操作





谢谢大家！

敬请指正！

