

数字图书馆推广工程

唯一标识符系统二期平台培训

国家图书馆信息技术部

童忠勇

2017年6月

- 系统建设背景
- 系统建设内容概述
- 系统核心功能详解
- 系统数据迁移

一、系统建设背景

一期系统概述：

国家图书馆数字资源唯一标识符系统于2012年11月完成建设，在近五年的运行中，完成全国48个省、市级图书馆唯一标识符系统的部署，完成220万条唯一标识符的注册，基本构建了公共图书馆数字资源唯一标识符体系。

系统应用情况：

- 参与推广工程资源联建项目：实现了地方图书馆30万条唯一标识符的注册。
- 与国图Aleph系统对接：实现了约50万册纸本图书与电子图书的数据关联展示。
- 与国图发布与服务系统对接：实现了数字资源的唯一标识和统一管理；
- 外购资源整合：完成35万条方正电子书唯一标识符的注册，并在国图发布与服务系统中实现了与自建资源的整合及可视化展示。

[登录](#)[检索首页](#)[我的图书馆](#)[参数设置](#)[最近检索](#)[帮助](#)[ENGLISH](#) [少儿版](#)

正题名

《红楼梦》中的建筑与园林

中文文献库

[搜索](#)[二次搜索](#)[返回结果列表](#) | [添加到电子书架](#) | [保存/邮寄](#) | [重新查询](#)第 2 条记录(共 2 条) < [上一条记录](#) [下一条记录](#) >[标准格式](#)[网摘目次](#)[文献索取](#)

头标区 -----nam0-22----- 450-

ID 号 003854691

通用数据 20080804d2008 em y0chiy50 ea

题名与责任 ● 《红楼梦》中的建筑与园林 [专著] / 关华山著

出版项 ●天津：百花文艺出版社，2008

载体形态项 254页：图，折图；23cm

语言 chi

内容提要 本书意在经由小说，抽析其中反映之生活与建筑空间的契合关系，选取《红楼梦》中的荣、宁二府与大观园为研究对象，加以分析、考订、组合，重建荣、宁二府与大观园的建筑风貌，绘出其配置图。全书配48幅精美的彩色插图，100余幅黑白建筑示意图。

主题 ● 《红楼梦》--建筑艺术--研究

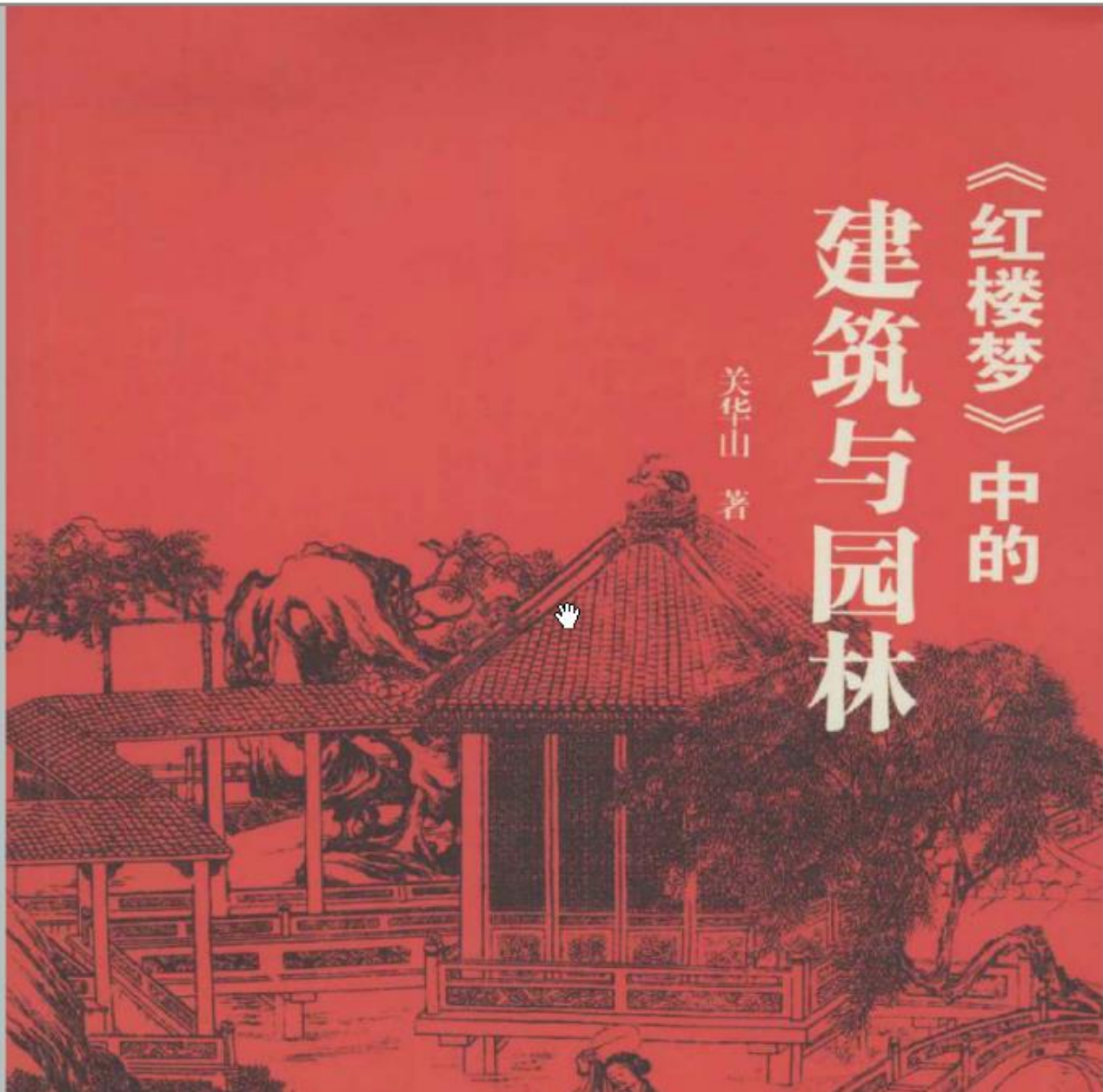
● 《红楼梦》--古典园林--研究

中图分类号 ●I207.411

●TU-092.49

著者 ●关华山著

在线阅读 [全文](#)并列 [访问电子文献](#)



二、项目建设内容概述

提升系统核心性能

包括注册及解析性能，数据同步性能，聚类性能等。

优化完善系统功能

改善现有唯一标识符注册及命名规则，增加注册数据核心字段，增加唯一标识符生成规则。

丰富系统平台接口

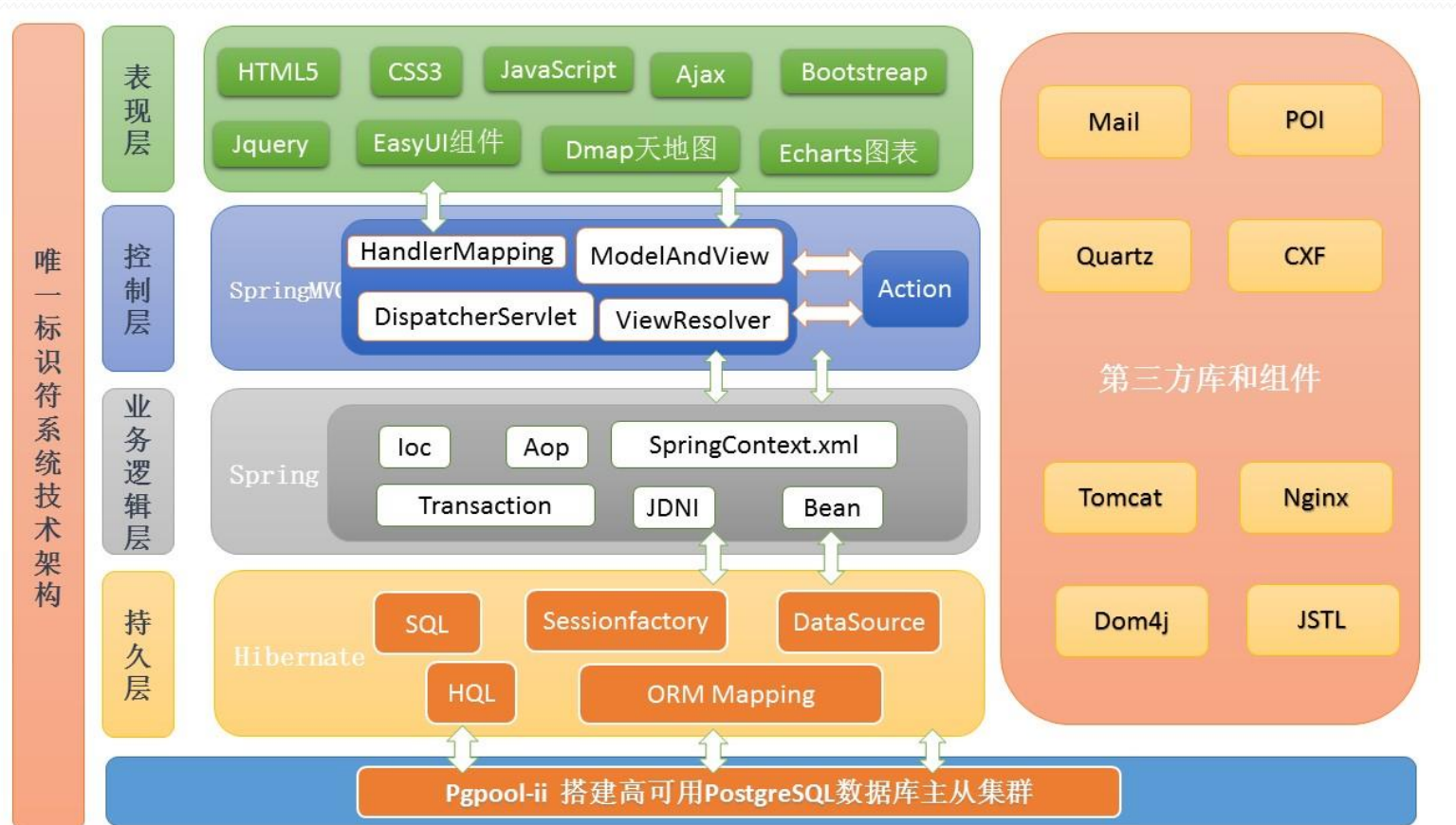
适应地方馆的个性化服务需求，为各级图书馆基于资源加工、保存、检索和服务等系统平台提供数据调用服务。

新增数据聚类功能

对全国唯一标识符数据进行统一规范整理，设计数据关联规则，对现有唯一标识符数据进行聚类处理，以解决唯一标识符数据的离散化带来的服务不完整性。

1. 提升系统核心性能

(1) 摒弃传统的Spring+Struts+Hibernate开发框架，采用目前主流的SpringMVC框架设计，避免Struts2框架本身的高危漏洞。



1. 提升系统核心性能

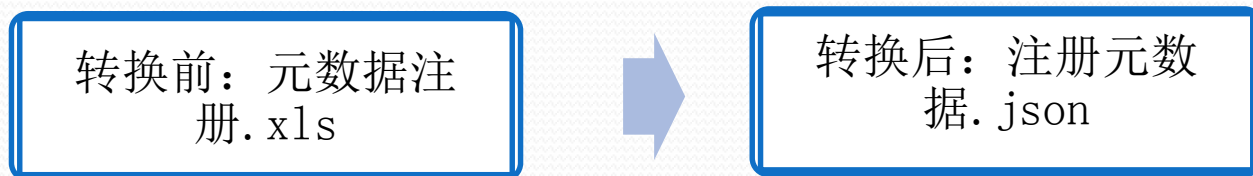
(2) 系统性能

- 注册数据量在500万数据量以内情况下：支持500用户并发正向解析，响应时间不超过200ms；支持200用户并发逆向解析，响应时间不超过1s；
- 注册数据量在500~1000万数据量情况下：支持500用户并发正向解析，响应时间不超过300ms；支持200用户并发逆向解析，响应时间不超过2s。
- 支持1000万以上数据量的唯一标识符解析能力，可通过追加服务器设备，有效提升唯一标识符的解析性能。

2. 优化完善系统功能

(1) 注册功能优化

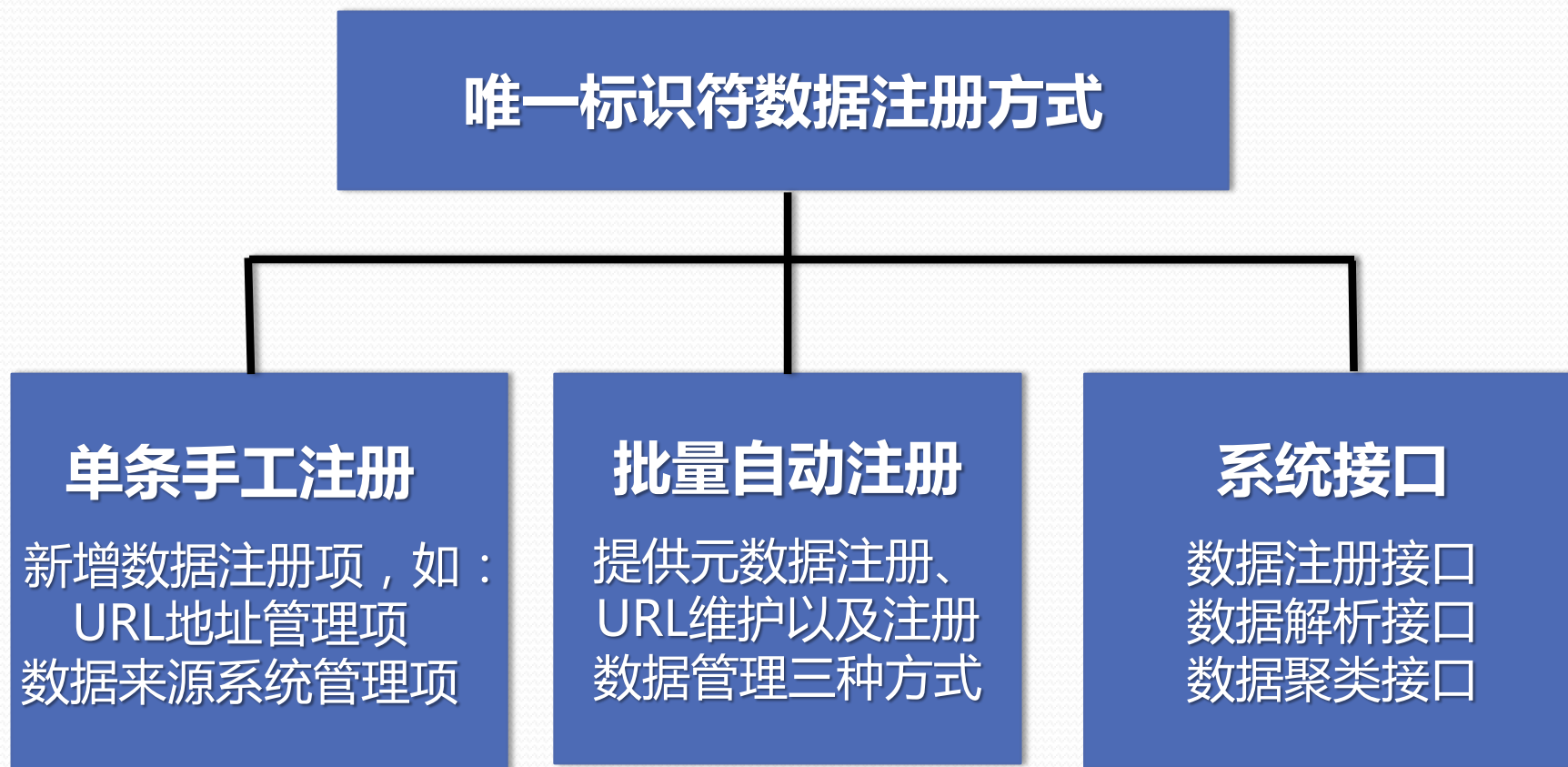
- 优化注册数据转换工具，高效压缩元数据文件大小，减轻服务器压力。



- 新增主题词、标准号值、所属数据库等数据信息，丰富了注册数据的描述信息，以便更方便的与其他应用系统进行对接；
- 完善了标识符生成规则，参照DOI的著录方式，将原有的CDOI生成规则拓展为三种：“默认规则”、“随机字符串”、“指定标准号”。
- 完善了URL地址著录规则，新增URL地址管理元数据，如状态位管理项，以区分不同用途（如多终端、内外网访问等）的唯一标识符解析。
- 新增数据审核机制，增强了唯一标识符数据服务的安全性和可靠性；

2. 优化完善系统功能

(2) 数据注册方式优化



2. 优化完善系统功能

(3) 完善数据导出功能：支持元数据按任务方式批次导出或一次性全部导出

S 系统菜单
ystem menu

- 元数据注册
- 元数据维护
 - 元数据维护
 - 元数据维护 (批量)
 - URL维护 (批量)
 - 批量处理结果
 - 元数据导出列表
- 标识符解析
- 元数据审核
- 系统管理

创建时间： 至

检索字段： 唯一标识符 精确

查询结果

批量导出 全部导出 ~~批量删除~~

<input type="checkbox"/>	唯一标识符	题名 (资源名称)	文件格式	资源类型	审核状态	基本操作
1	<input type="checkbox"/> 108.ndlc.2.1100009031010006/2017031017592	到白马涧龙池微旅行	TXT	报纸	已通过	
2	<input type="checkbox"/> 108.ndlc.2.1100009031010006/T4F32.16010170	苏州市非遗馆开馆展	TXT	报纸	已通过	
3	<input type="checkbox"/> 108.ndlc.2.1100009031010006/F32.1601017011	苏州同里将建文遗数	TXT	报纸	已通过	

10 第 1 共 1 页

显示1到3,共3记录

2. 优化完善系统功能

(4) 优化数据同步功能

数据同步可视化，可查看全国各图书馆同步数量，同步时间，同步状态，可对全国图书馆指定时间内数据，进行重新同步。



唯一标识符系统

(中国国家数字图书馆)

帮助说明 修改密码 退出系统

【nb测试账号 / 系统管理员】您好！

S 系统菜单
system menu

元数据导出

元数据审核

手动聚类

系统管理

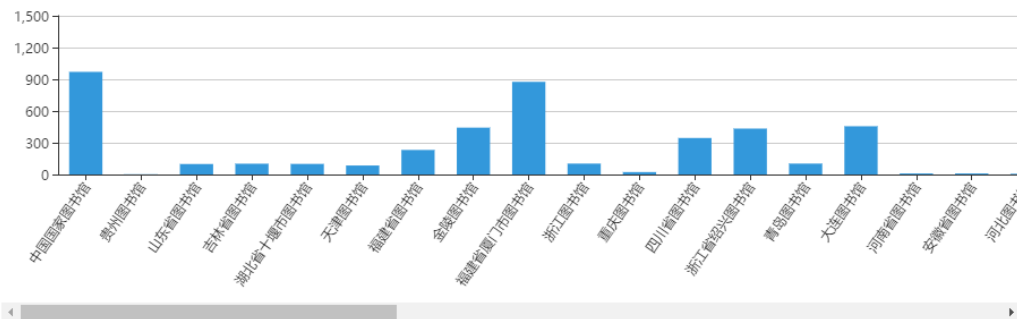
节点名称: 中国国家数字图书馆

重新同步

搜索

开始时间:

结束时间:



	节点名称	同步数量	同步时间	同步类型	同步状态	同步详情
1	中国国家图书馆	7	2017-02-23 14:34:53.427	手动同步	同步成功	元数据同步成功;
2	中国国家图书馆	1	2017-02-23 14:14:52.65	手动同步	同步成功	元数据同步成功;
3	浙江省舟山图书馆	645	2017-02-21 17:01:22.148	手动同步	同步成功	元数据同步成功;
4	云南省图书馆	458	2017-02-21 17:01:22.148	手动同步	同步成功	元数据同步成功;
5	湖南图书馆	450	2017-02-21 17:01:22.148	手动同步	同步成功	元数据同步成功;
6	南京图书馆	455	2017-02-21 17:01:22.148	手动同步	同步成功	元数据同步成功;
7	中国国家图书馆	20	2017-02-21 17:01:22.148	手动同步	同步成功	元数据同步成功;
8	哈尔滨市图书馆	455	2017-02-21 16:08:10.479	自动同步	同步成功	元数据同步成功;
9	首都图书馆	454	2017-02-21 16:08:10.479	自动同步	同步成功	元数据同步成功;
10	内蒙古图书馆	454	2017-02-21 16:08:10.479	自动同步	同步成功	元数据同步成功;

10 | 1 | 共6页 | 显示1到10共60记录



唯一标识符系统

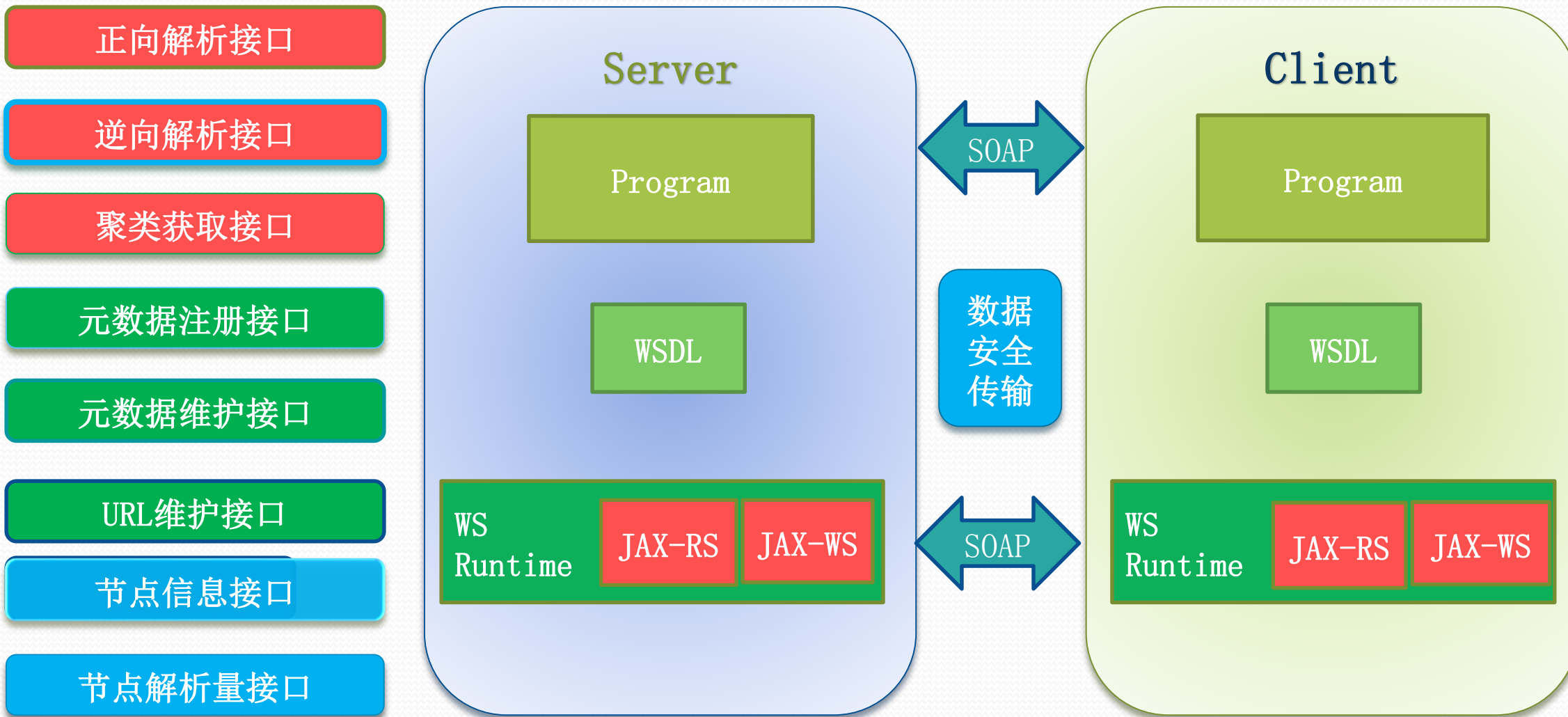
(中国国家图书馆)

电子邮件: chinadoi@163.com

京ICP备05014420

国家图书馆版权所有©

3. 丰富系统平台接口



正向解析接口

逆向解析接口

聚类获取接口

元数据注册接口

元数据维护接口

URL维护接口

节点信息接口

节点解析量接口

Server

Program

WSDL

WS
Runtime

JAX-RS

JAX-WS

SOAP

数据
安全
传输

SOAP

Client

Program

WSDL

WS
Runtime

JAX-RS

JAX-WS

4. 新增唯一标识符聚类功能

➤ 聚类种类

- ① 同一资源
- ② 多版本资源
- ③ 多卷册资源
- ④ 多语种资源
- ⑤ 跨文献资源

➤ 聚类方式

- ① 自动聚类
- ② 手工聚类

**唯一标识符系统**
(中国国家图书馆)

唯一标识符: 108.ndlc.2.2200009031010001/T5F14.1605004057

资源类型: T5

文件格式: JPG

MARC记录001标识号:

作者:

出版者:

关联:

描述信息: 谭竹青, 吉林省长春市人, 汉族, 1948年参加革命工作, 1975年6月加入中国共产党。他情系老百姓, 一心为民的精神赢得了群众的赞誉, 被人们亲切地称为“小巷总理” 图: 谭竹青同志先进事迹图片展

创建人:

数据来源系统:

系统号: 2

题名(资源名称): 谭竹青同志先进事迹图片展

语种:

出版时间:

来源: 吉林省图书馆

创建时间: 2017-02-28 00:00:00

同一资源

唯一标识符	题名	作者	出版社	文件格式	资源类型	详细信息
1 108.ndlc.2.2200009031010001/T5F14.1605004054	谭竹青同志先进事迹图片展			JPG	图片	?
2 108.ndlc.2.2200009031010001/T5F14.1605004055	谭竹青同志先进事迹图片展			JPG	图片	?
3 108.ndlc.2.2200009031010001/T5F14.1605004058	谭竹青同志先进事迹图片展			JPG	图片	?
4 108.ndlc.2.2200009031010001/T5F14.1605004053	谭竹青同志先进事迹图片展			JPG	图片	?
5 108.ndlc.2.2200009031010001/T5F14.1605004052	谭竹青同志先进事迹图片展			JPG	图片	?
6 108.ndlc.2.2200009031010001/T5F14.1605004050	谭竹青同志先进事迹图片展			JPG	图片	?
7 108.ndlc.2.2200009031010001/T5F14.1605004051	谭竹青同志先进事迹图片展			JPG	图片	?

10 | 第 1 页 共 1 页 | 显示 1 到 7 共 7 记录

- 多版本资源
- 多卷册资源
- 跨语种资源
- 跨文献资源
- 规范人名库资源

**唯一标识符系统**
(中国国家图书馆)

电子邮件: chinadai@163.com
京ICP备05014420
国家图书馆版权所有©

通过聚类分析，可以为第三方系统提供数据支撑，进而较好的实现资源推荐，便于读者使用。

聚类类型: 操作类型:

资源类型: 状态:

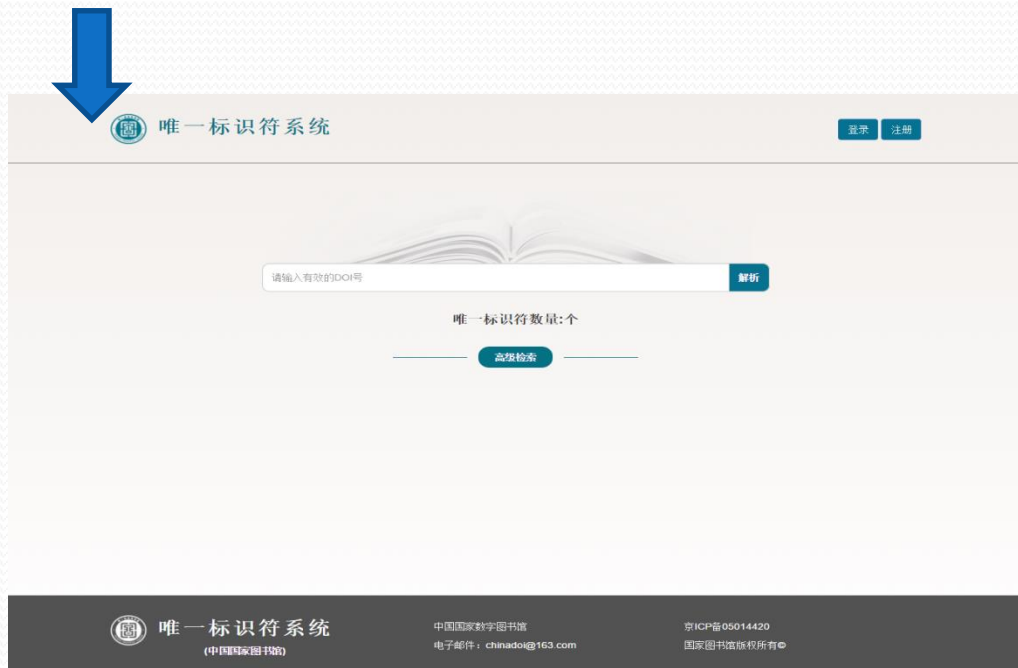
聚类审核信息一览

<input type="checkbox"/>	聚类	唯一标识符	资源名(题名)	出版者	来源	
<input type="checkbox"/>	40031	1383857	108.ndlc.2.1100009031010001/T1F24.411999010937m7	隋堂摘稿	李金黄中等	国家图书馆
<input type="checkbox"/>	40032	1383857	108.ndlc.2.1100009031010001/T1F24.411999010937m5	隋堂摘稿	李金黄中等	国家图书馆
<input type="checkbox"/>	40033	1383857	108.ndlc.2.1100009031010001/T1F24.411999010937m4	隋堂摘稿	李金黄中等	国家图书馆
<input type="checkbox"/>	40034	1383857	108.ndlc.2.1100009031010001/T1F24.411999010937m3	隋堂摘稿	李金黄中等	国家图书馆
<input type="checkbox"/>	40035	1383857	108.ndlc.2.1100009031010001/T1F24.411999010937m2	隋堂摘稿	李金黄中等	国家图书馆
<input type="checkbox"/>	40036	1383856	108.ndlc.2.1100009031010001/T1F24.411999010931m9	安陽集	張士隆	国家图书馆
<input type="checkbox"/>	40037	1383856	108.ndlc.2.1100009031010001/T1F24.411999010931m10	安陽集	張士隆	国家图书馆
<input type="checkbox"/>	40038	1383856	108.ndlc.2.1100009031010001/T1F24.411999010931m3	安陽集	張士隆	国家图书馆
<input type="checkbox"/>	40039	1383856	108.ndlc.2.1100009031010001/T1F24.411999010931m4	安陽集	張士隆	国家图书馆
<input type="checkbox"/>	40040	1383856	108.ndlc.2.1100009031010001/T1F24.411999010931m5	安陽集	張士隆	国家图书馆

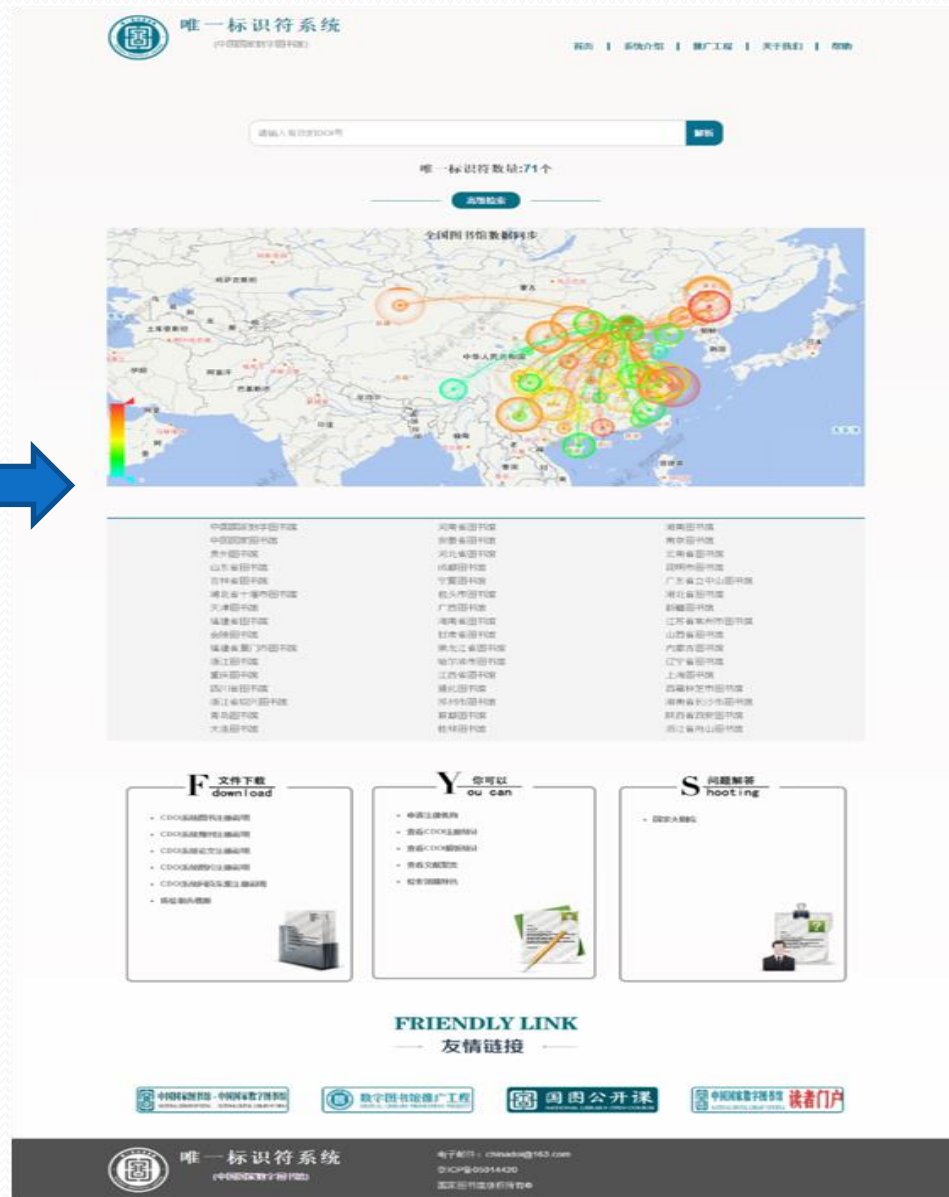
5. 系统首页面可视化设计

系统采用HTML5技术，支持PC端、移动端等多终端访问。

二级节点首页面



中心节点首页面



系统首页面功能：向互联网用户提供高级搜索功能。

请输入有效的DOI号 解析

唯一标识符数量:个

高级检索

书名: 作者: 出版社: 查询

唯一标识符	题名	作者	出版社	文件格式	资源类型	详细信息
1 108.ndlc.2.1100009031010001/T1F24.0191025607	实用税收指南	杨振谟主编	新华出版社	PDF	图书	
2 108.ndlc.2.1100009031010001/T1F24.002379293	郭店楚简老子校释	廖名春著	清华大学出版社	PDF	图书	
3 108.ndlc.2.1100009031010001/T1F24.004700289	2011年公务员考试精讲	高守国编著	清华大学出版社	PDF	图书	
4 108.ndlc.2.1100009031010001/T1F24.003852664	C#3.0完全自学宝典	杨明羽编著	清华大学出版社	PDF	图书	
5 108.ndlc.2.1100009031010001/T1F24.003843163	化学课堂教学技能训练	袁孝凤主编	华东师范大学出版社	PDF	图书	
6 108.ndlc.2.1100009031010001/T1F24.003842605	小学综合实践活动	郭元祥, 沈旒主编	华东师范大学出版社	PDF	图书	
7 108.ndlc.2.1100009031010001/T1F24.003237012	业务流程与信息技术	(美)Ulric J. Gelinas,	清华大学出版社	PDF	图书	
8 108.ndlc.2.1100009031010001/T1F24.002182960	管理学原理自学辅导	关键, 杜强编著	华中科技大学出版社	PDF	图书	
9 108.ndlc.2.1100009031010001/T1F24.002191115	建国以来中国共产党科技政策研	崔禄春著	华夏出版社	PDF	图书	
10 108.ndlc.2.1100009031010001/T1F24.004045420	让小学生受益一生的100个美德	陈雄, 黄棋主编	华东师范大学出版社	PDF	图书	

10 第 1 共 4449 页

显示 1 到 10, 共 44483 记录

系统首页面功能：各省、市图书馆唯一标识符的同步及解析可视化展示



系统首页面功能：提供帮助文件下载，操作指导和问题解答

F 文件下载 download

- 唯一标识符系统培训(2013-06)
- 唯一标识符注册与维护项目(2014-12)



U 用户手册 user manual

- 如何申请注册机构
- 如何进行唯一标识符解析
- 如何进行注册元数据
- 如何进行元数据维护
- 如何进行元数据审核
- 如何添加/删除/搜索用户
- 如何进行数据来源管理



S 问题解答 shooting

- 什么是注册机构
- 什么是注册机构操作员、注册机构管理员



三、系统核心功能详解

1. 注册机构申请

步骤一、输入网址<http://doi.nlc.cn>，进入首页；

步骤二、点击 **机构注册** 按钮，在协议内容下方勾选“同意该协议”，点击“我要注册”；

步骤三、进入注册界面并完成信息填写。

注意：省级/副省级图书馆申请注册机构；市级/区级图书馆由省馆分配机构操作员帐号，不需要申请注册机构。

步骤四、用户登录邮箱进行验证（**注意：**邮箱必须是申请人可用的邮箱，否则不能认证）。

步骤五、系统管理员（国家图书馆）审核，请联系系统管理员审批注册申请，审批通过后，就可以登录系统。

2. 用户登录

点击右上角“用户登录”，在弹出的登录窗口中输入用户名和密码，点击“登录”，进入系统操作页面。



3. 忘记密码

输入用户名以及注册该用户时填写的邮箱，点击“立即找回”，系统提示“密码已发送到注册时的邮箱，请注意查收”。



关闭

 唯一标识符系统
(中国国家图书馆)

用户名 

密码 

记住密码

找回密码

登录



用户名: 用户名

邮箱: 注册时使用的邮箱

立即找回

4. CDOI注册

元数据注册分为两大部分：元数据单条注册和元数据批量注册。

(1) 元数据单条注册

S 系统菜单
system menu

- 元数据注册
- 元数据注册**
- 元数据注册（批量）
- 注册数据颗粒度对照表
- 颗粒度输入规则对照表
- 元数据维护
- 标识符解析
- 元数据审核
- 系统管理

标识符生成规则: 默认规则 *
资源类型: 请选择 *
文件格式: 请选择 *
MARC001标识号:
标准号: 请选择 *
主题词:
出版者:
关联信息:
原始资源库标识号:
描述信息:

标题(资源名称):
系统号: *
作者:
版本信息:
标准号值:
语种: 请选择 *
出版时间:
所属数据库:
馆藏信息:

URL: * 状态位: 请选择 *
数据来源系统: 请选择 * 版权信息:

➤ **标识符生成规则：**必填项，即标识符后缀的生成方法，分为三种：默认规则（默认选中）、标准号值、唯一字符串。

默认规则：分为两段，第一段是综合信息编号，第二段是系统内部唯一号，两个结合作为元数据的后缀。

标准号值：将注册元数据的标准号值作为标识符的后缀。（注：如果选择标准号值作为生成规则，则标准号值字段为必填项）

唯一字符串：将随机字符串作为标识符的后缀。

- **标题（资源名称）：**必填项，根据要生成标识符的元数据真实信息填写。
- **资源类型：**必填项。
- **文件格式：**必填项。
- **系统号：**必填项。根据资源的真实系统号填写，不能填写汉字。

标识符生成规则：默认规则 *
资源类型：请选择 *
文件格式： *
MARC001标识号：
标准号：
关键词：
出版者：
关联信息：
原始资源库标识号：



The screenshot shows a form with several fields. A red circle highlights the '文件格式' (File Format) dropdown menu, which is currently open and showing a list of options: 图书 (Books), 期刊 (Journals), 论文 (Papers), 报纸 (Newspapers), 图片 (Images), 音频 (Audio), 视频 (Video), and 网页采集 (Web Scraping). The '视频' (Video) option is highlighted in grey.

标识符生成规则：默认规则 *
资源类型：请选择 *
文件格式： *
MARC001标识号：
标准号：
关键词：
出版者：
关联信息：
原始资源库标识号：



The screenshot shows a form with several fields. A red circle highlights the '文件格式' (File Format) dropdown menu, which is currently open and showing a list of file formats: ARC, ARJ, ASP, AVI, BMP, CAB, DBF, DOC, FLV, and GIF. The 'GIF' option is highlighted in grey.

➤ URL管理

注册元数据：如果要填写URL，则URL中的必填项也要填写。

URL字段：必填项，查询唯一标识符时，链接到该资源对应的URL地址。

状态位字段：必填项，即URL的显示状态。

数据来源系统：必填项，URL的数据来源。数据来源系统后台可维护。



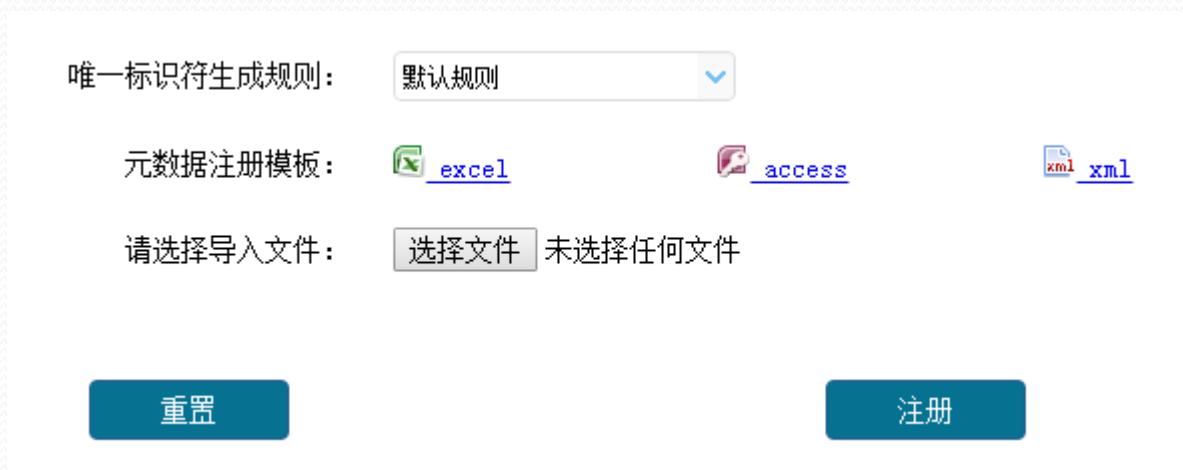
The screenshot shows a form for URL management. On the left, the label "URL:" is displayed. The form contains several fields: a text input for "url:" with a red asterisk, a dropdown for "数据来源系统:" with "方正电子书" selected, a dropdown for "状态位:" with "请选择" selected and a red asterisk, and a text input for "版权信息:". A red rectangular box highlights the "状态位:" dropdown menu, which lists "手机" and "电脑" as options. To the right of the form is a green plus icon. Below the form is a blue button labeled "注册".

(2) 元数据批量注册

步骤一：导出模板，用户根据自己的需求选择不同格式的注册模板；

步骤二：在模版中填写元数据属性值；

步骤三：导入填好的元数据模板，选择生成规则并注册。



The screenshot shows a web interface for metadata batch registration. It features three main sections: a dropdown menu for '唯一标识符生成规则' (Unique Identifier Generation Rule) set to '默认规则' (Default Rule); a section for '元数据注册模板' (Metadata Registration Template) with three options: 'excel', 'access', and 'xml', each with a corresponding file icon; and a '请选择导入文件' (Please select import file) section with a '选择文件' (Select file) button and the text '未选择任何文件' (No file selected). At the bottom, there are two buttons: '重置' (Reset) and '注册' (Register).

唯一标识符生成规则: 默认规则

元数据注册模板: excel access xml

请选择导入文件: 选择文件 未选择任何文件

重置 注册

元数据注册方式

注册方式一：注册元数据、回写URL同步进行；



注册方式二：先注册元数据，再回写URL；



元数据注册元素

序号	元数据注册模板	字段注释	必备性
1	资源类型	图书 (T1)、期刊 (T2)、论文 (T3)、报纸 (T4)、图片 (T5)、音频 (T6)、视频 (T7)、网页采集 (T8)	必备
2	文件格式	如：ARC (F1)、AVI (F4)、PDF (F24)、TXT (F32)、HTML (F40) 等	必备
3	系统号	使用编目系统的资源加工流水号；若无，则使用资源来源系统流水号。建议使用10位数字从0000000001开始。	必备
4	颗粒度K1	标识对象的一级粒度，如种、卷 (册/首/幅)	有则必备
5	颗粒度K1描述	对一级粒度的详细说明，如第2册，第5卷等	有则必备
6	颗粒度K1值	描述一级粒度的值，如第2册K1值为2,第5卷K1值为5等	有则必备

颗粒度描述规则

资源类型	颗粒度1	颗粒度1(Key)	颗粒度1(可输入值)
图书(T1)	单卷/册	无输入	无输入
	多卷/册	T1K1V2	输入1
期刊(T2)	全年	T2K1V1	输入2
	期刊	T2K1V2	输入3
	增刊	T2K1V3	输入4
	合订本	T2K1V4	输入2
论文(T3)	无输入	无输入	无输入
报纸(T4)	全年	T4K1V1	输入2
	期刊	T4K1V2	输入3
图片(T5)	图号	T5K1V1	输入1
音频(T6)	单曲号	无输入	无输入
	多曲号	T6K1V2	输入1
视频(T7)	单场次	无输入	无输入
	多场次	T7K1V2	输入1

颗粒度输入规则

输入名称	输入数据	数据类型	数据解释
输入1	a	数字	第a卷/册/页
	b-c	起始数字	第b-c卷/册/页
	a,b-c	数字、起始数字任意组合(逗号分割)	
输入2	YYYY	年份(4位年)	
	YYYY-YYYY	起始年份-结束年份	某年至某年
	YYYY,YYYY-YYYY	年份、起始年份-结束年份任意组合(逗号分割)	
输入3	YYYYim	起始年份(4位年)+字母i+m期(最少1位数字)	某年的m期
	YYYYia-YYYYib	起始年份(4位年)+字母i+a期(最少1位数字)-结束年份(4位年)+字母i+b期(最少1位数字)	某年的a期至某年的b期
输入4	YYYYsm	起始年份(4位年)+字母s+第m增刊(最少1位数字)	某年的第m增刊
	YYYYsa-YYYYsb	起始年份(4位年)+字母s+第a增刊(最少1位数字)-结束年份(4位年)+字母s+第b增刊(最少1位数字)	某年的第a增刊至某年的第b增刊

元数据注册元素

序号	元数据注册模板	字段注释	必备性
7	MARC记录001标识号	MARC记录001字段	有则必备
8	题名(资源名称)	标识对象最通常被使用的名字，如题名	必备
9	作者	创建资源内容的主要责任者或者对资源内容做出贡献的其他实体。	有则必备
10	标准号	国际标准书号（ISBN）、国际标准连续出版物编号（ISSN）、音像制品的国际标准编码（ISRC）	有则必备
11	标准号值	与标准号配对使用	有则必备

元数据注册元素

序号	元数据注册模板	字段注释	必备性
12	语种	已在系统中规范化处理，可多选，多选用“，”分割。	有则必备
13	出版者	使资源成为可以获得并可用的责任者，包括个体，组织，或服务。	有则必备
14	出版时间	用于注明电子资源出版或颁布的日期。	有则必备
15	关联	描述该资源之间的关联关系。	有则必备
16	描述信息	资源内容的说明。描述可以包括但不限于以下内容：文摘、目录、对以图形来揭示内容的资源而言的文字说明、或者一个有关资源内容的自由文本描述。	有则必备

元数据注册元素

序号	元数据注册模板	字段注释	必备性
17	主题词	在标引和检索中用以表达文献主题的规范化的词或词组。	有则必备
18	版本信息	如：第一版第一次印刷	有则必备
19	所属数据库	著录资源内容的类型，如：抗日战争图片库、徽州家谱数据库、长白山动植菌物数据库等。	有则必备
20	原始资源库标识号	原始资源库流水号	有则必备
21	馆藏信息	描述文献资源与编目系统中元数据的挂接关系	有则必备

元数据注册元素

序号	元数据注册模板	字段注释	必备性
22	URL地址	资源的访问地址。	有则必备
23	数据来源系统	资源的提供系统，本系统已做规范化处理，分为自建、商购、其他等，可建二级子系统。	有则必备
24	状态位	如：手机端、电脑端。	有则必备
25	版权信息	资源的版权描述	有则必备

(2) 元数据批量注册

注册成功：实时显示数据注册进度；

注册失败：如果有失败的数据，会显示出来，包括失败的条数以及失败原因，操作员可以根据失败原因修改注册文件相应项后，再次提交失败的数据。



唯一标识符生成规则： 默认规则

元数据注册模板：[excel](#) [access](#) [xml](#)

请选择导入文件： REGDATA.xls

注册失败！

- 第1条数据：标题名(资源名称)未填写；

5. 元数据维护

分为五部分：元数据维护，元数据维护（批量），URL维护（批量），批量处理结果和元数据导出列表。

(1) 元数据维护

查询结果						
 批量导出  全部导出  批量删除						
<input type="checkbox"/>	唯一标识符	题名（资源名称）	文件格式	资源类型	审核状态	基本操作
1	<input type="checkbox"/> 108.ndlc.2.1100009031010001/F9.041071508010120m1	美国因素对中日关系的影响	FLV	视频	待审核	 
2	<input type="checkbox"/> 108.ndlc.2.1100009031010001/F9.041071508010120m2	美国因素对中日关系的影响	FLV	视频	待审核	 
3	<input type="checkbox"/> 108.ndlc.2.1100009031010001/F9.041071506010117m1	非遗保护——苏州读本	FLV	视频	待审核	 
4	<input type="checkbox"/> 108.ndlc.2.1100009031010001/F9.041071406020070m1	周易谈人生——地道做人，厚德载物	FLV	视频	待审核	 
5	<input type="checkbox"/> 108.ndlc.2.1100009031010001/F9.041071406020070m2	周易谈人生——地道做人，厚德载物	FLV	视频	待审核	 
6	<input type="checkbox"/> 108.ndlc.2.1100009031010001/F9.041071406020069m1	辛亥革命与近代中国的地方主义	FLV	视频	待审核	 
7	<input type="checkbox"/> 108.ndlc.2.1100009031010001/F9.041071406020069m2	辛亥革命与近代中国的地方主义	FLV	视频	待审核	 
8	<input type="checkbox"/> 108.ndlc.2.1100009031010001/F9.041071406020068m1	辛亥革命：“革故”与“鼎新”	FLV	视频	待审核	 
9	<input type="checkbox"/> 108.ndlc.2.1100009031010001/F9.041071406020068m2	辛亥革命：“革故”与“鼎新”	FLV	视频	待审核	 

(2) 元数据导出

元数据维护页面有批量导出和全部导出功能，导出的数据需要在“元数据导出列表”下查看。



The screenshot displays the metadata management interface. On the left is a navigation menu with the following items:

- 元数据注册
- 元数据维护
- 元数据维护
- 元数据维护 (批量)
- URL维护 (批量)
- 批量处理结果
- 元数据导出列表

The main content area shows a table titled "查询结果" (Query Results). Above the table are three action buttons: "批量导出" (Batch Export), "全部导出" (All Export), and "批量删除" (Batch Delete). These buttons are circled in red. Below the buttons is a checkbox labeled "唯一标识符" (Unique Identifier). The table contains six rows of data:

	<input type="checkbox"/>	唯一标识符
1	<input type="checkbox"/>	108.ndlc.2.1100009031010001/F9.041071506010120m1
2	<input type="checkbox"/>	108.ndlc.2.1100009031010001/F9.041071506010120m2
3	<input type="checkbox"/>	108.ndlc.2.1100009031010001/F9.041071506010117m1
4	<input type="checkbox"/>	108.ndlc.2.1100009031010001/F9.041071406020070m1
5	<input type="checkbox"/>	108.ndlc.2.1100009031010001/F9.041071406020070m2
6	<input type="checkbox"/>	108.ndlc.2.1100009031010001/F9.041071406020069m1

(3) 元数据批量维护

步骤一： 导出元数据维护模板，用户根据自己的需求导出需要的维护模板并填写维护信息。

注意： 模版中必填项是“唯一标识符”字段，其他字段如果需要修改，则填写修改的内容，如果不需要修改，则保持空值。

步骤二： 文件导入，点击“元数据维护（批量）”，进入界面，在界面中导入准备好数据的元数据维护模板，点击“维护”。

(4) URL维护

同一条唯一标识符数据，可能会包含多条URL地址，操作员可能在注册数据后需要追加或修改URL地址。

步骤一：导出模板，用户根据自己的需求选择不同格式的URL维护模板；

步骤二：导入填写好的URL维护模板，并点击维护。

(5) 批量处理结果

可对数据进行处理类型、上传时间、处理结果的查询。

S 系统菜单
system menu

元数据注册

元数据维护

元数据维护

元数据维护(批量)

URL维护(批量)

批量处理结果

处理类型: 查询所有

上传时间:

查询

查询结果

	批次	上传用户	处理类型	处理结果描述	上传时间	基本操作
1	20170531171936086	童忠勇	元数据批量注册	1条错误数据!	2017-05-31 17:19:36.086	
2	20170527165055960	张宝文	元数据批量注册	注册成功数据7507条; 保存URL数据7507条;	2017-05-27 16:50:55.96	
3	20170527162515061	张宝文	元数据批量注册	0条错误数据!	2017-05-27 16:25:15.061	
4	20170527162312313	张宝文	元数据批量注册	1条错误数据!	2017-05-27 16:23:12.313	
5	20170527161720121	张宝文	元数据批量注册	7507条错误数据!	2017-05-27 16:17:20.121	
6	20170527161347738	张宝文	元数据批量注册	7507条错误数据!	2017-05-27 16:13:47.738	
7	20170505101027	孙梦晖	元数据批量维护	更新元数据1条;	2017-05-05 10:10:27.151	
8	20170505100723	孙梦晖	元数据批量维护	更新元数据1条;	2017-05-05 10:07:23.016	

5. 唯一标识符解析

正向解析

- URL1 <http://mylib.nlc.cn/system/application/search/displayClient/metaDataObjectDisplayRedirectPage.jsp?autologin=yes&metaData.id=316450&metaData.lld=320945&ldLib=40283415347ed8bd013483174ef60002> 机构：国家图书馆
系统：数字资源与发布系统 版权： 状态：电脑
- URL2 <http://mylib.nlc.cn/system/application/search/displayClient/metaDataObjectDisplayRedirectPage.jsp?autologin=yes&metaData.id=306909&metaData.lld=311404&ldLib=40283415347ed8bd013483174ef60002> 机构：国家图书馆
系统：数字资源与发布系统 版权： 状态：电脑
- URL3 <http://mylib.nlc.cn/system/application/search/displayClient/metaDataObjectDisplayRedirectPage.jsp?autologin=yes&metaData.id=316563&metaData.lld=321058&ldLib=40283415347ed8bd013483174ef60002> 机构：国家图书馆
系统：数字资源与发布系统 版权： 状态：电脑
- URL4 www.test.com 机构：国家图书馆 系统：方正电子书 版权： 状态：电脑
- URL5 www.test1.com 机构：国家图书馆 系统：方正电子书 版权： 状态：电脑

正向解析：当元数据对应多个URL地址时，提示每个URL的来源和系统、版权、状态位等信息，供用户参考选择；

逆向解析

简单检索 高级检索

服务节点： 国家图书馆【本节点】 请选择

检索词： 赏析

查询 二次查询

查询结果

	唯一标识符	题名(资源名称)	文件格式	资源类型	来源	基本操作
1	108.ndlc.2.1100009031010001/T1F24.fd2015331885	红楼梦诗词赏析	PDF	图书	国家图书馆	
2	108.ndlc.2.1100009031010001/T1F24.006770702	《红楼梦》诗词赏析	PDF	图书	国家图书馆	
3	108.ndlc.2.1100009031010001/T7F9.WJVAW0605121308m2	《红楼梦》人物赏析——3	FLV	视频	国家图书馆	
4	108.ndlc.2.1100009031010001/T7F9.WJVAW0605121308m1	《红楼梦》人物赏析——3	FLV	视频	国家图书馆	
5	108.ndlc.2.1100009031010001/T1F42.00016766	红楼梦诗词赏析	CEBX	图书	方正电子书	
6	108.ndlc.2.1100009031010001/T1F42.00000473	红楼梦诗词赏析	CEBX	图书	方正电子书	
7	108.ndlc.2.1100009031010001/T7F9.1186577	中国古典四大名著赏析——	FLV	视频	国家图书馆	
8	108.ndlc.2.1100009031010001/T1F24.003228966	红楼梦诗词赏析	PDF	图书	国家图书馆	
9	108.ndlc.2.1100009031010001/T1F24.0188005161	《红楼梦》赏析	PDF	图书	国家图书馆	
10	108.ndlc.2.1100009031010001/T1F24.003751117	语用学与《红楼梦》赏析	PDF	图书	国家图书馆	

逆向解析：有简单查询和高级查询两种解析方式，简单查询可以在查询结果基础上二次查询；高级查询可以添加多种条件组合查询数据；

6. 元数据审核

元数据注册成功后，需要机构管理员进行元数据审核，审核通过方可提供解析服务。

创建时间: 至 批次号:

唯一标识符: 状态: 全部

元数据审核信息

通过 未通过 撤销

批次号: 批次通过 批次撤销

	批次号	唯一标识符	资源名(题名)	出版者	来源	审核状态	基本操作
1		108.ndlc.2.1100009031010001/T1F24.003252976	亡友鲁迅印象记	上海文化出版社	国家图书馆	已通过	
2		108.ndlc.2.1100009031010001/T1F24.004531807	卡米不想洗澡	希望出版社	国家图书馆	已通过	
3		108.ndlc.2.1100009031010001/T1F24.003208285	导弹动力系统故障机理分	西北工业大学出版社	国家图书馆	已通过	
4		108.ndlc.2.1100009031010001/T1F24.004308812	独船	中国少年儿童出版社	国家图书馆	已通过	
5		108.ndlc.2.1100009031010001/T1F24.004364480	恐龙世界	科学普及出版社	国家图书馆	已通过	
6		108.ndlc.2.1100009031010001/T1F24.003208288	企业社会资本与战略管理	人民出版社	国家图书馆	已通过	
7		108.ndlc.2.1100009031010001/T1F24.004547969	樱桃雨	江苏少年儿童出版社	国家图书馆	已通过	
8		108.ndlc.2.1100009031010001/T1F24.004300666	做一回上帝	江西高校出版社	国家图书馆	已通过	
9		108.ndlc.2.1100009031010001/T1F24.003209867	积极的未来	江西人民出版社	国家图书馆	已通过	
10		108.ndlc.2.1100009031010001/T1F24.005182222	新编常用集成电路及元器	机械工业出版社	国家图书馆	已通过	

7. 添加/删除用户

点击“系统管理”-“用户管理”，进入“用户管理”界面

用户名: 姓名:

用户信息

 增加  删除

	用户名	姓名	所属机构	角色	所属单位	基本操作
1			国家图书馆	机构操作员		 
2			国家图书馆	机构操作员		 
3			国家图书馆	机构操作员		 
4			国家图书馆	机构操作员		 

8. 数据来源管理

细分为：一级来源系统、二级来源系统。

数据来源信息添加

一级来源系统名称： *

二级来源系统名称：

系统主页URL地址：

商购资源

数据来源系统信息					
 增加  删除					
	<input type="checkbox"/>	机构名称	系统名称	系统主页URL地址	基本操作
1	<input type="checkbox"/>	国家图书馆	数字资源与发布系统	http://mylib.nlc.cn	
2	<input type="checkbox"/>	国家图书馆	方正电子书	http://10.100.2.1/Usp/apabi_usp/pub.mvc?pid=book.detail&me	

9. 唯一标识符聚类

聚类方式:

自动聚类分析

手工聚类分析

聚类类型:

对所有同步至国数图中心节点的元数据进行分类整理。唯一标识符聚类可以是**同一资源**的聚类；也可以是**跨文献、多语种、多版本、多卷册**的数据聚类。

聚类配置

系统菜单
ystem menu

元数据导出

元数据审核

手动聚类

系统管理

用户管理

权限管理

行业管理

文件格式管理

资源类型管理

语种管理

聚类规范管理

日志统计

节点数据统计

聚类规范管理信息

添加

1

聚类规则修改

聚类类型：同一资源

规则名称：同一资源聚类规则

版本信息

主题词

资源类型

文件格式

系统号

颗粒度K1描述

标准号

标准号值

扩展字段1

扩展字段2

扩展字段3

颗粒度K3值

颗粒度K1

颗粒度K1值

颗粒度K2

聚类相等项

聚类不相等项

ID	聚类项字段名称	聚类项条件	相似度 (1~100%)
1	颗粒度K3值	相等	100
2	颗粒度K1	相等	100
3	颗粒度K1值	相等	100
4	颗粒度K2	相等	100
5	颗粒度K2值	相等	100

确定

取消

聚类配置：中心节点用户可以在后台对不同聚类方式进行规则配置，如相等项，相似度，不相等项字段的配置；

聚类信息

聚类类型：多卷册资源 108.ndlc.2.1100009031010001/T1F24.4008000m4 查询

聚类数据信息

+ 新增聚类 - 删除

<input type="checkbox"/>	聚类		资源名(题名)	出版者	来源	基本操作
<input type="checkbox"/>	1334717	108.ndlc.2.1100009031010001/T1F24.4008000m53	空同先生集六十三卷 嘉靖		国家图书馆	
<input type="checkbox"/>	1334717	108.ndlc.2.1100009031010001/T1F24.4008000m63	空同先生集六十三卷 嘉靖		国家图书馆	
<input type="checkbox"/>	1334717	108.ndlc.2.1100009031010001/T1F24.4008000m62	空同先生集六十三卷 嘉靖		国家图书馆	
<input type="checkbox"/>	1334717	108.ndlc.2.1100009031010001/T1F24.4008000m35	空同先生集六十三卷 嘉靖		国家图书馆	
<input type="checkbox"/>	1334717	108.ndlc.2.1100009031010001/T1F24.4008000m29	空同先生集六十三卷 嘉靖		国家图书馆	
<input type="checkbox"/>	1334717	108.ndlc.2.1100009031010001/T1F24.4008000m24	空同先生集六十三卷 嘉靖		国家图书馆	
<input type="checkbox"/>	1334717	108.ndlc.2.1100009031010001/T1F24.4008000m33	空同先生集六十三卷 嘉靖		国家图书馆	
<input type="checkbox"/>	1334717	108.ndlc.2.1100009031010001/T1F24.4008000m26	空同先生集六十三卷 嘉靖		国家图书馆	
<input type="checkbox"/>	1334717	108.ndlc.2.1100009031010001/T1F24.4008000m31	空同先生集六十三卷 嘉靖		国家图书馆	
<input type="checkbox"/>	1334717	108.ndlc.2.1100009031010001/T1F24.4008000m25	空同先生集六十三卷 嘉靖		国家图书馆	

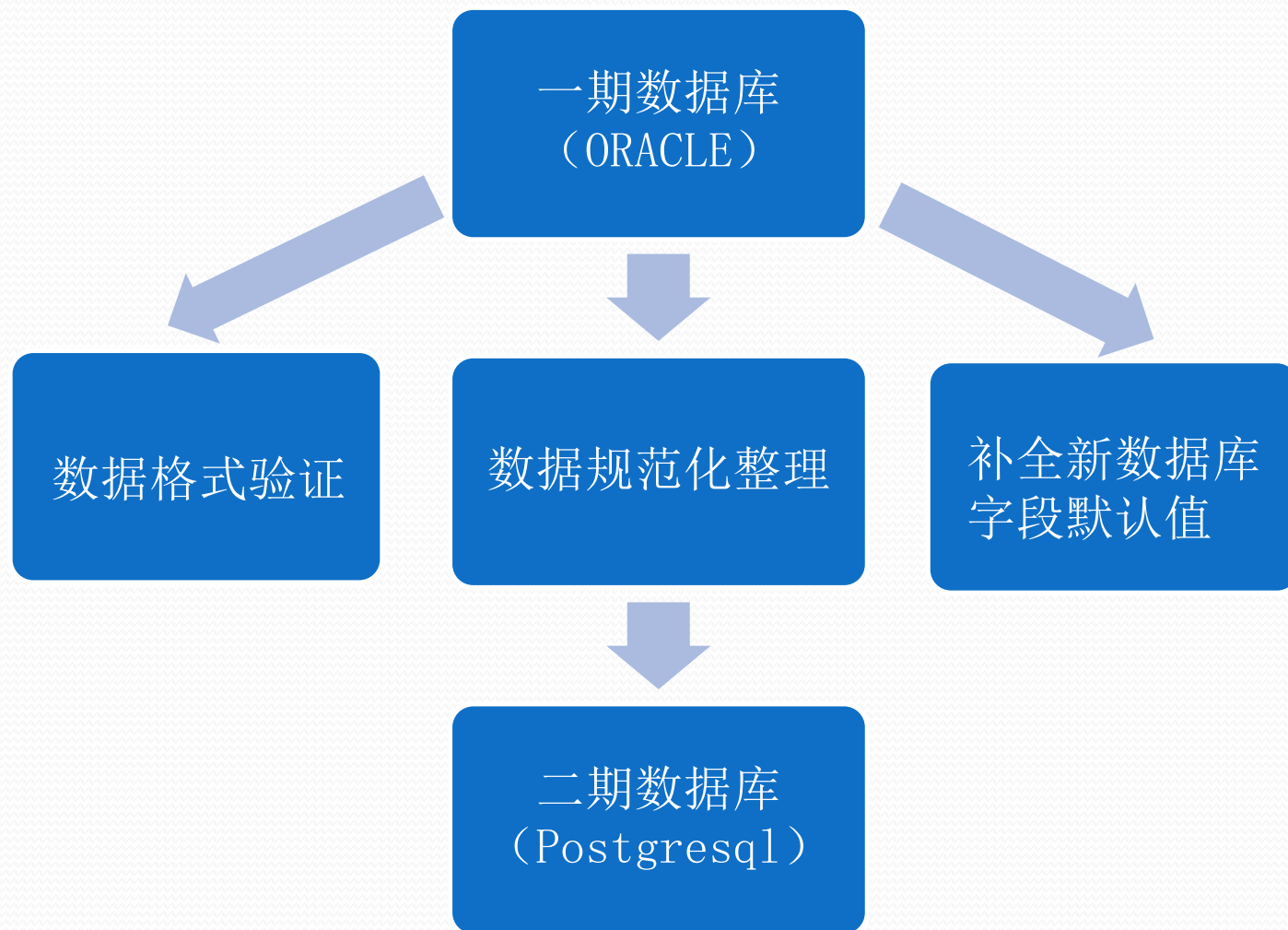
10 ⏪ ⏩ 第 1 共 7 页 🔄 显示1到10,共63记录

聚类信息：在中心节点聚类页面，可以选择不同的聚类方式，查看数据的聚类情况；

四、系统部署规划

1. 数据迁移

一期数据的迁移工作由数据迁移程序完成，数据迁移程序从一期数据库里提取数据，对数据进行**数据格式验证**；**数据规范化整理**，例如：一期元数据语言字段数据，有chi, 中文, 汉语。整理后统一为CHI，对应语言表中文，可以后台管理；**补全新库字段默认值**后插入到二期数据库，完成数据迁移功能。



2. 迁移策略

(1) 原有服务节点数据迁移

CDOI数据：数据库拷贝，需要各图书馆配合国家图书馆完成；

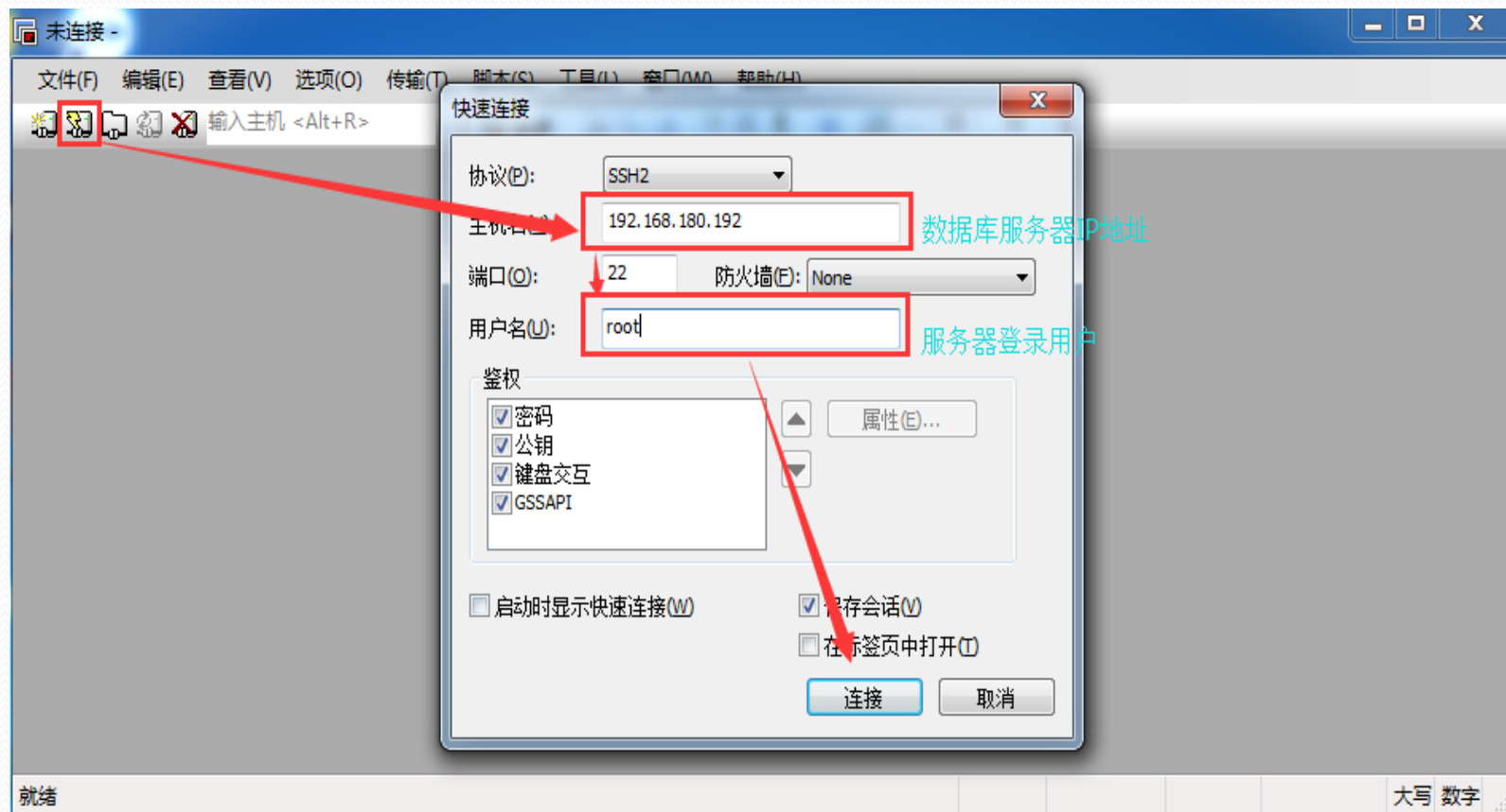
用户数据：原有省馆系统管理员/操作员取消；

原有省馆注册机构（前台）管理员保留，市馆注册机构（前台）管理员调整为省馆注册机构操作员；

(2) 原有注册机构数据迁移

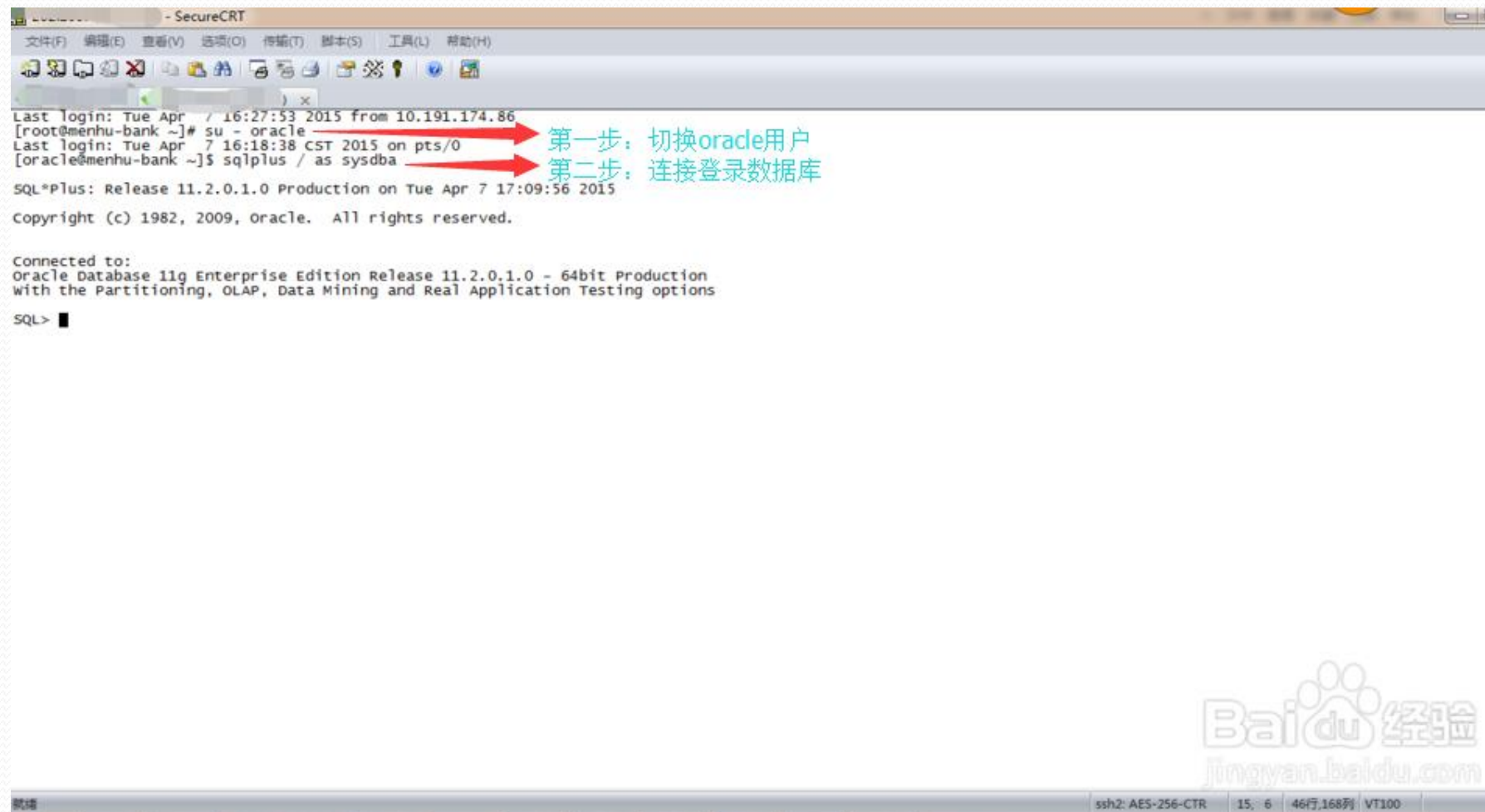
2. 数据迁移

通过root用户登录到Linux服务器



root用户下首先切换到oracle用户 命令:

su - oracle



导出命令:

```
exp user/pwd@SID file=path/app.dmp full=y log=path/app.log
```

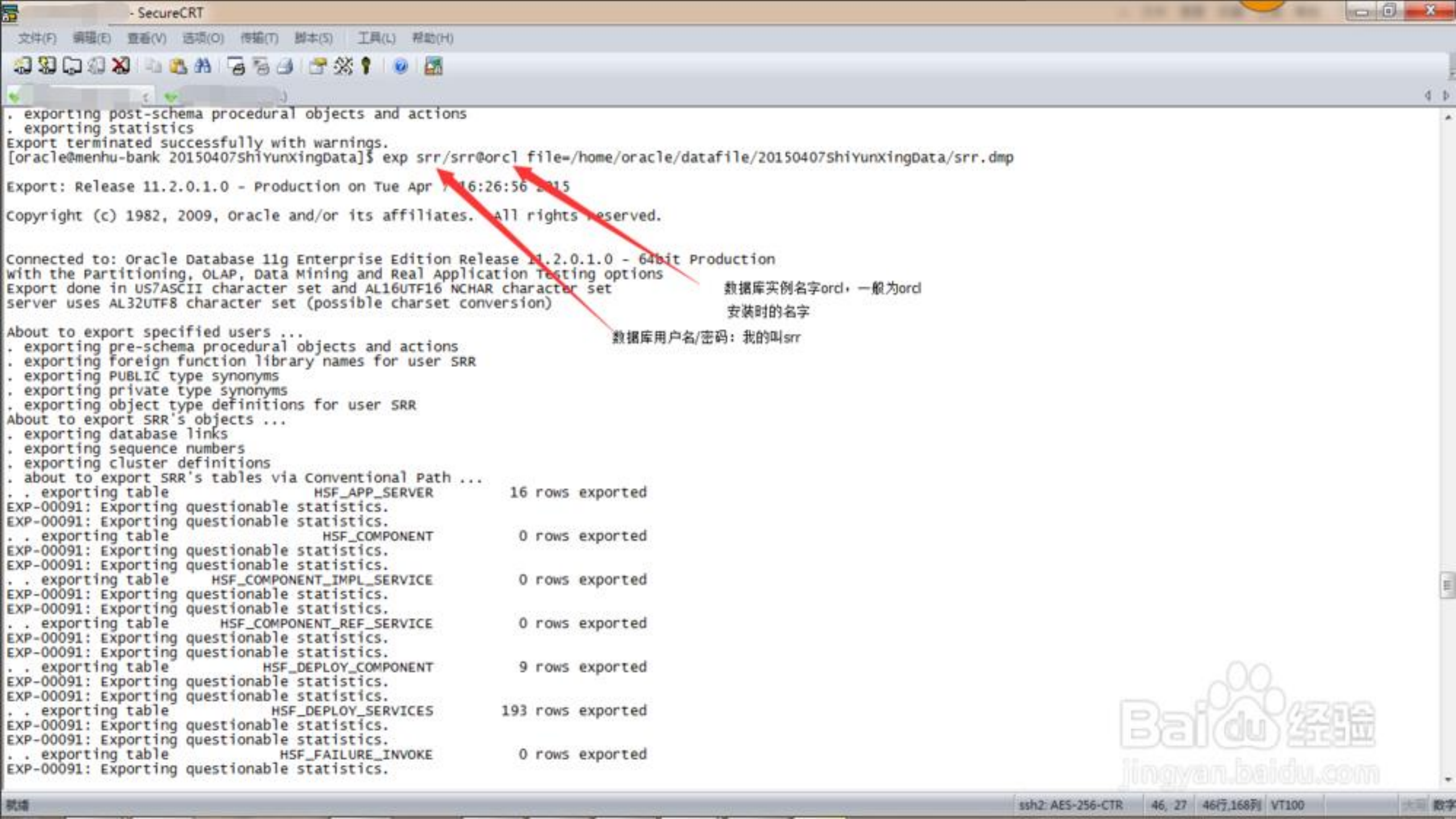
其中user/pwd是你要导出的表空间用户名与密码

SID为oracle启动实例

path是你导出的路径

app.dmp是你导出的数据库文件名

以上都是必须参数一样都不能少



```

. exporting post-schema procedural objects and actions
. exporting statistics
Export terminated successfully with warnings.
[oracle@menhu-bank 20150407ShiYunXingData]$ exp srr/srr@orc1 file=/home/oracle/datafile/20150407ShiYunXingData/srr.dmp
Export: Release 11.2.0.1.0 - Production on Tue Apr 7 16:26:56 2015
Copyright (c) 1982, 2009, Oracle and/or its affiliates. All rights reserved.

```

数据库实例名字orc1, 一般为orc1
安装时的名字

数据库用户名/密码: 我的叫srr

```

Connected to: Oracle Database 11g Enterprise Edition Release 11.2.0.1.0 - 64bit Production
With the Partitioning, OLAP, Data Mining and Real Application Testing options
Export done in US7ASCII character set and AL16UTF16 NCHAR character set
server uses AL32UTF8 character set (possible charset conversion)

About to export specified users ...
. exporting pre-schema procedural objects and actions
. exporting foreign function library names for user SRR
. exporting PUBLIC type synonyms
. exporting private type synonyms
. exporting object type definitions for user SRR
About to export SRR's objects ...
. exporting database links
. exporting sequence numbers
. exporting cluster definitions
. about to export SRR's tables via Conventional Path ...
. exporting table          HSF_APP_SERVER          16 rows exported
EXP-00091: Exporting questionable statistics.
EXP-00091: Exporting questionable statistics.
. exporting table          HSF_COMPONENT            0 rows exported
EXP-00091: Exporting questionable statistics.
EXP-00091: Exporting questionable statistics.
. exporting table          HSF_COMPONENT_IMPL_SERVICE 0 rows exported
EXP-00091: Exporting questionable statistics.
EXP-00091: Exporting questionable statistics.
. exporting table          HSF_COMPONENT_REF_SERVICE 0 rows exported
EXP-00091: Exporting questionable statistics.
EXP-00091: Exporting questionable statistics.
. exporting table          HSF_DEPLOY_COMPONENT      9 rows exported
EXP-00091: Exporting questionable statistics.
EXP-00091: Exporting questionable statistics.
. exporting table          HSF_DEPLOY_SERVICES      193 rows exported
EXP-00091: Exporting questionable statistics.
EXP-00091: Exporting questionable statistics.
. exporting table          HSF_FAILURE_INVOKE       0 rows exported
EXP-00091: Exporting questionable statistics.

```



谢谢！

欢迎加入唯一标识符QQ群：**378362263**

工作人员：孙梦晖，谭焯 88545316

童忠勇 88545655