

文本框可根据需求改变颜色、移动位置；文字可编辑

33.2.55.33

Manufacturing
Supply chain
Product
Cargo
Customer
Delivery
Inventory
Management
Freight

50.36.2

Innovation
Branding
Solution
Marketing
Analysis
Ideas
Success
Management

60.50.3.1



Innovation
Branding
Solution
Marketing
Analysis
Ideas
Success
Management

Branding
Solution
Marketing
Analysis
Ideas
Success
Management

Manufacturing
Supply chain
Product
Cargo
Customer
Delivery
Inventory
Management
Freight

22.3.3.55

国家图书馆 2018.12

推广工程网络资源保存与服务系统

目录

系统建设

系统功能

系统操作

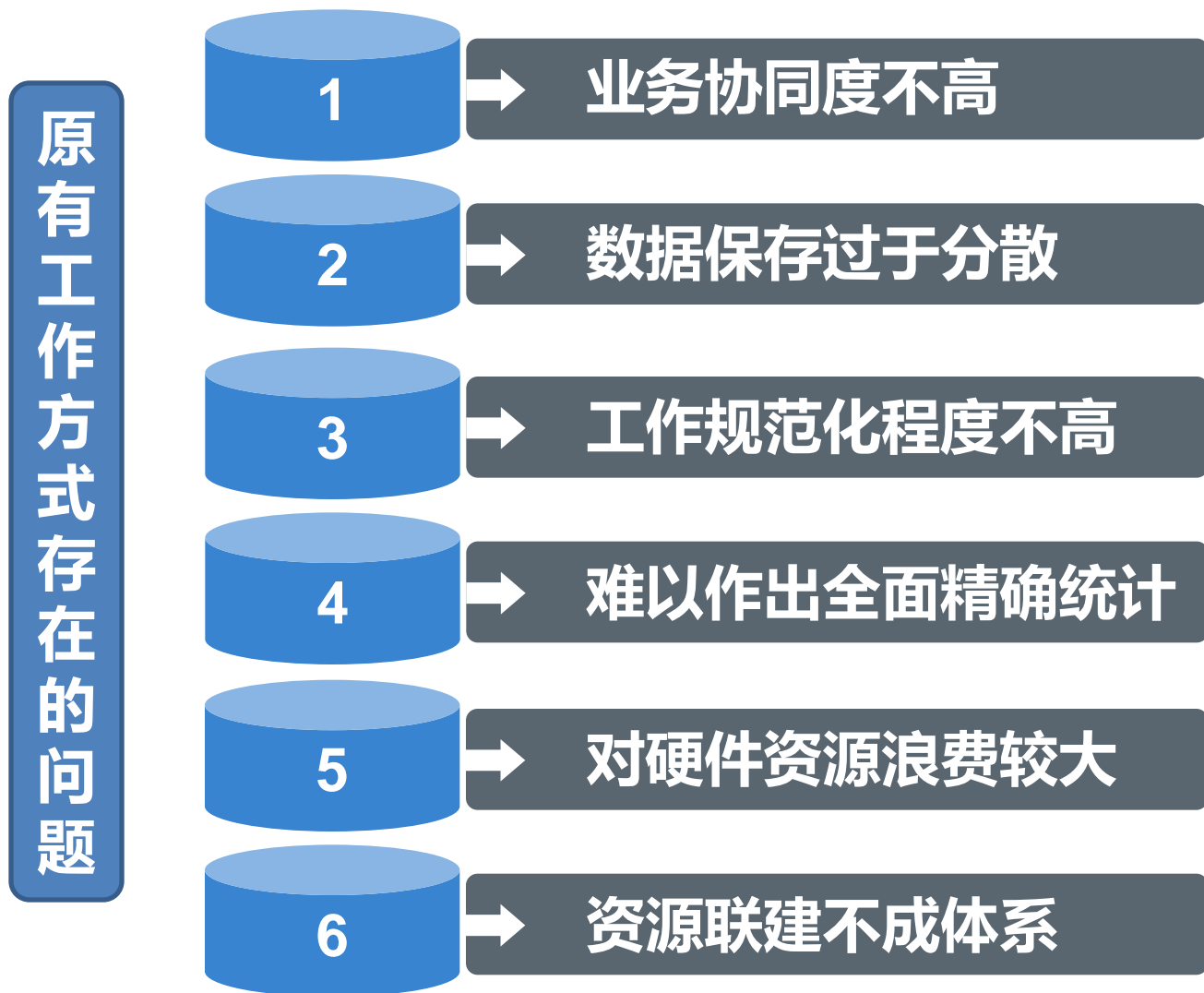


第一部分

系统建设



第一部分：系统建设



第一部分：系统建设

建设目标

以国际通用 **WARC** 采集文档格式为基础，通过系统开发和功能定制，实现网络资源保存与服务全流程业务的规范化管理和可视化操作。

1

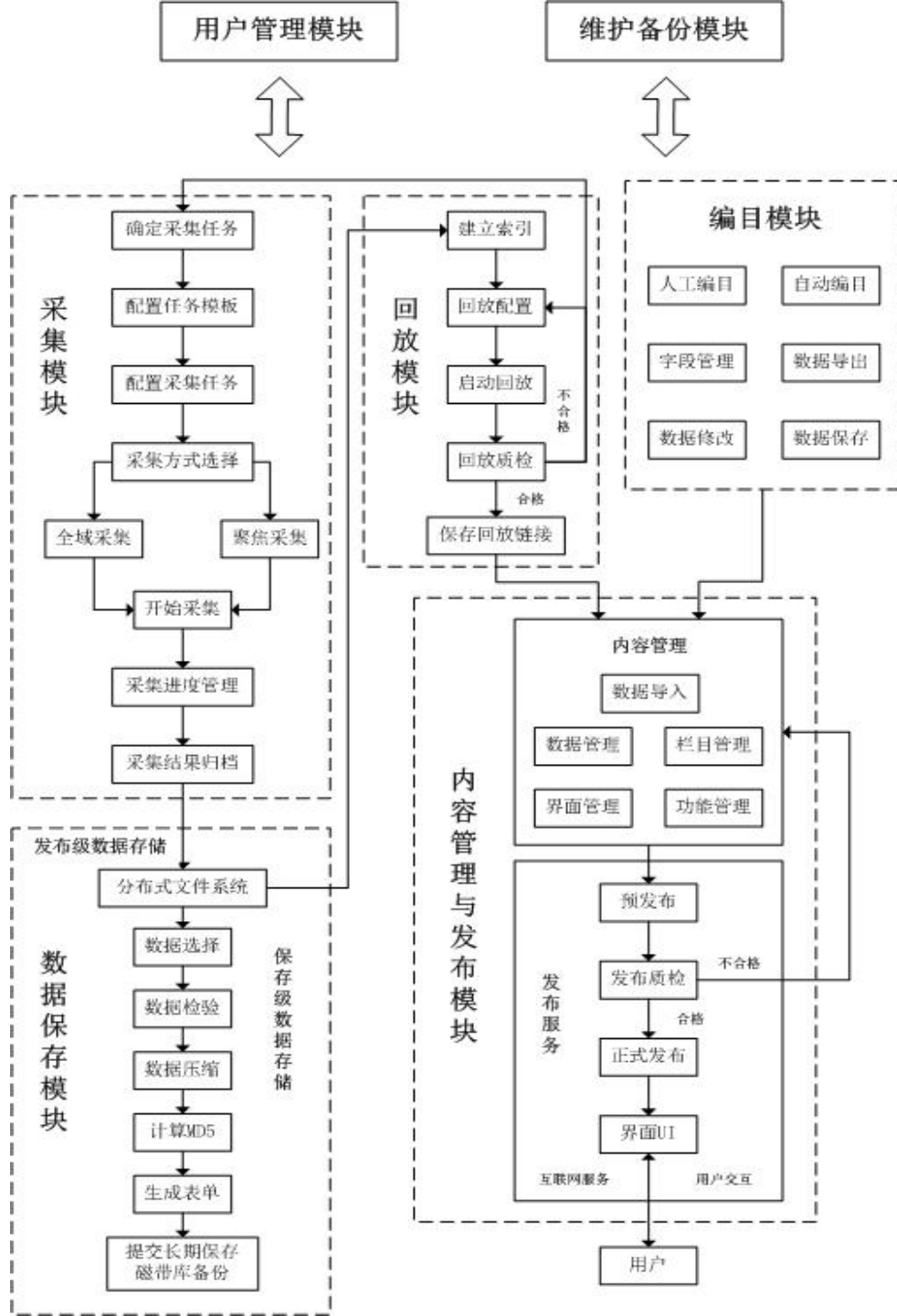
系统能够在现有软硬件环境下，支持大批量（TB级别）数据的采集、加工、管理和发布，并提供互联网展示服务和检索。

2

操作和使用过程便捷，具备 **可视化** 的界面操作，支持与其它模块之间的无缝衔接，同时还应具备 **可扩展性** 和 **可复用性**，能够较好满足地方馆开展网络资源保存与服务需求，并适应未来业务和用户需求的改变。



第一部分：



第一部分：系统建设

系统架构——B/S架构

1

服务器端

- 操作系统：Linux操作系统，内核版本2.6或以上。Windows Server 2012操作系统用于网站发布
- 支持环境：Java、Tomcat、Hadoop
- 数据库：My SQL

2

客户端

- 操作系统：Microsoft windows7以上
- 支持环境：Internet Explorer9.0以上



第一部分：系统建设

系统健壮性——安全性

1

在本系统与其他业务系统的接口能相互衔接并安全运作的情况下，保证本系统的独立性和完整性，防止非本系统的用户对本系统进行任何操作。

2

系统软件本身提供可靠的备份和恢复机制，在硬件、服务器软件、数据库软件等出现严重故障时，保留故障点的信息并能够彻底恢复到故障前的正常运行环境。



第一部分：系统建设

系统健壮性——安全性

3

系统具备7×24小时不间断服务，系统本身稳定、健壮、安全、可靠，能够完全满足当前技术环境下对于海量网络资源发布与服务需求。

4

系统需具备严格的权限控制和管理机制，严格控制系统后台各项操作和管理功能的使用和访问权限。



第一部分：系统建设

系统性能特点

- (1) 数据量增加时，能够方便对存储容量进行扩展。
- (2) 用户并发数增加时，能够对硬件平台支撑能力进行扩充。
- (3) 系统有良好的扩展性，能够方便地增加新的子系统或功能。
- (4) 对系统的参数、配置等进行修改可以简单易行。
- (5) 系统提供方便、友好的操作界面，操作简单、易懂。
- (6) 支持资源联建和推广。地方馆部署后可独立开展全流程业务，并与国图系统进行关联，可实现数据的统一管理 with 个性化发布。



第一部分：系统建设

建设原则



第一部分：系统建设

系统部署

硬件准备

- 服务器需有1TB以上（建议）可用的数据存储空间
- 支持以太网接口和TCP/IP协议
- 以Raid5划分磁盘

准备工作

软件准备

- 操作系统：Centos7.* 64位及以上、分布式操作系统Hadoop（有需求时）
- 磁盘分区：大容量且为一个区（分布式系统不用）

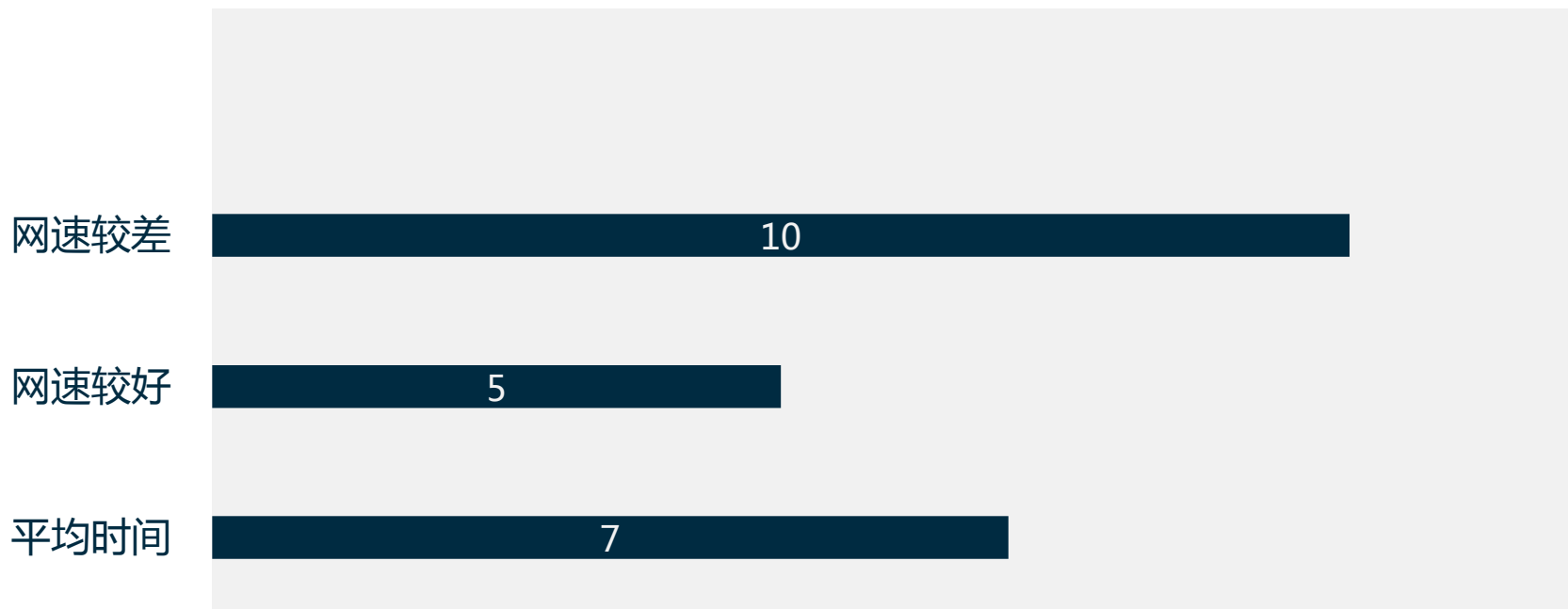
网络准备

- 开通IP的互联网访问权限
- 开通端口保证数据的传输



第一部分：系统建设

部署时长（天）



第一部分：系统建设



国图主节点

- 完成系统部署、分布式系统部署，建立虚拟机满足多人同时使用。
- 完成超级管理员、国图管理员、国图用户的权限分配。
- 开始实际运行。
- 国图中心服务器支持100个并发请求，同时在线用户最多2000人。



第一部分：系统建设



地方馆分节点

- 完成在广东、广西、云南省馆的部署，与国图主系统实现互联。
- 完成地方馆管理员、地方馆用户的权限分配。
- 完成系统的调试，并开始试运行。

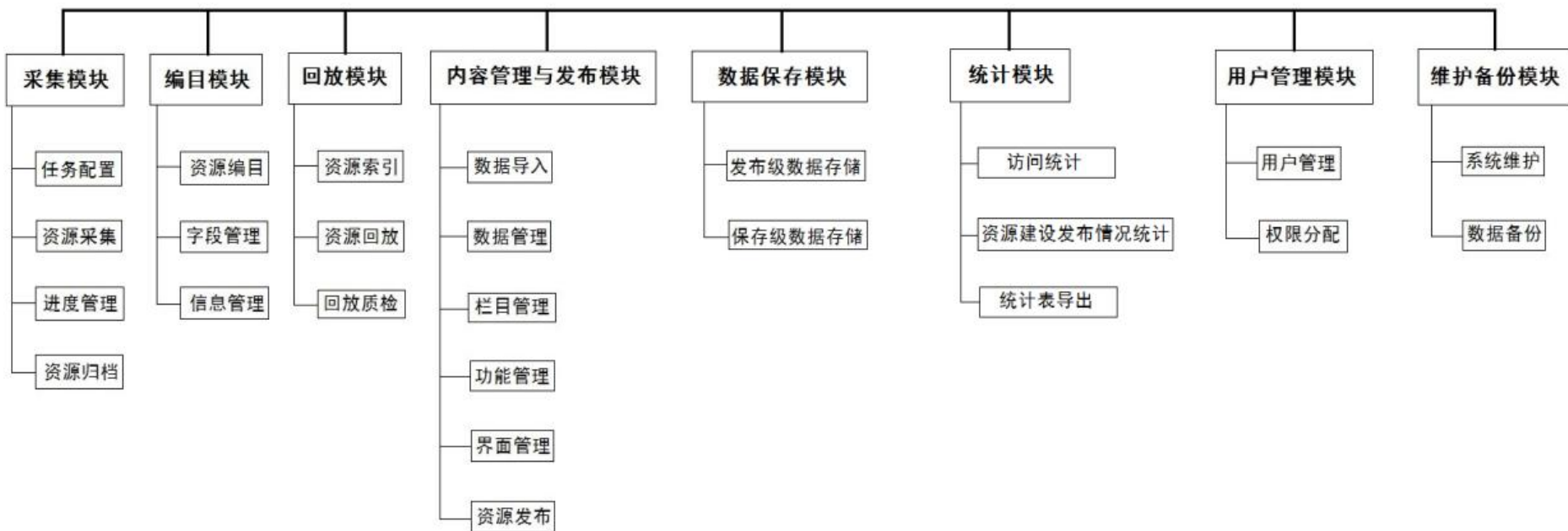
第二部分

系统功能

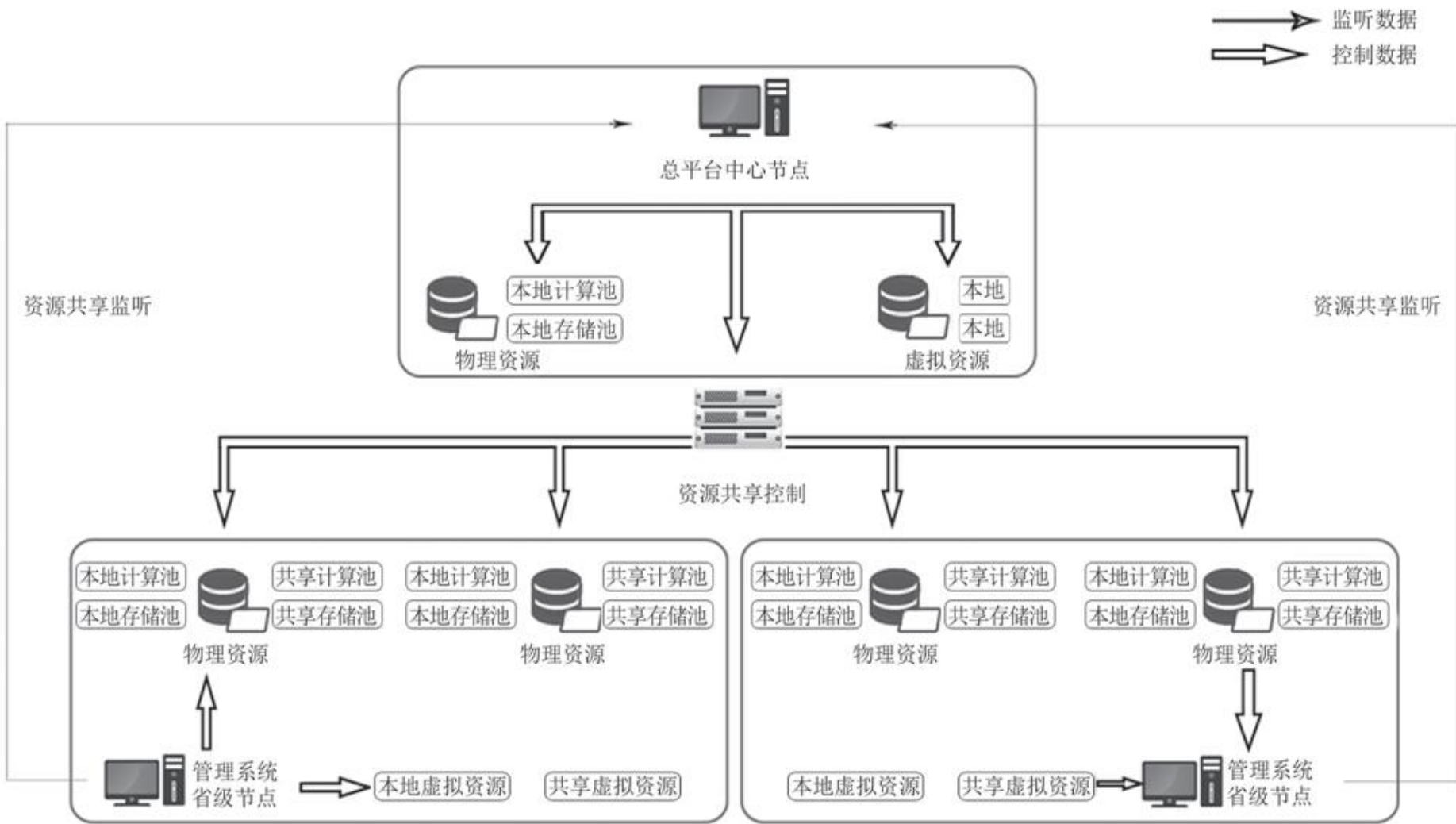


第二部分：系统功能

网络信息发布和服务



第二部分：系统功能

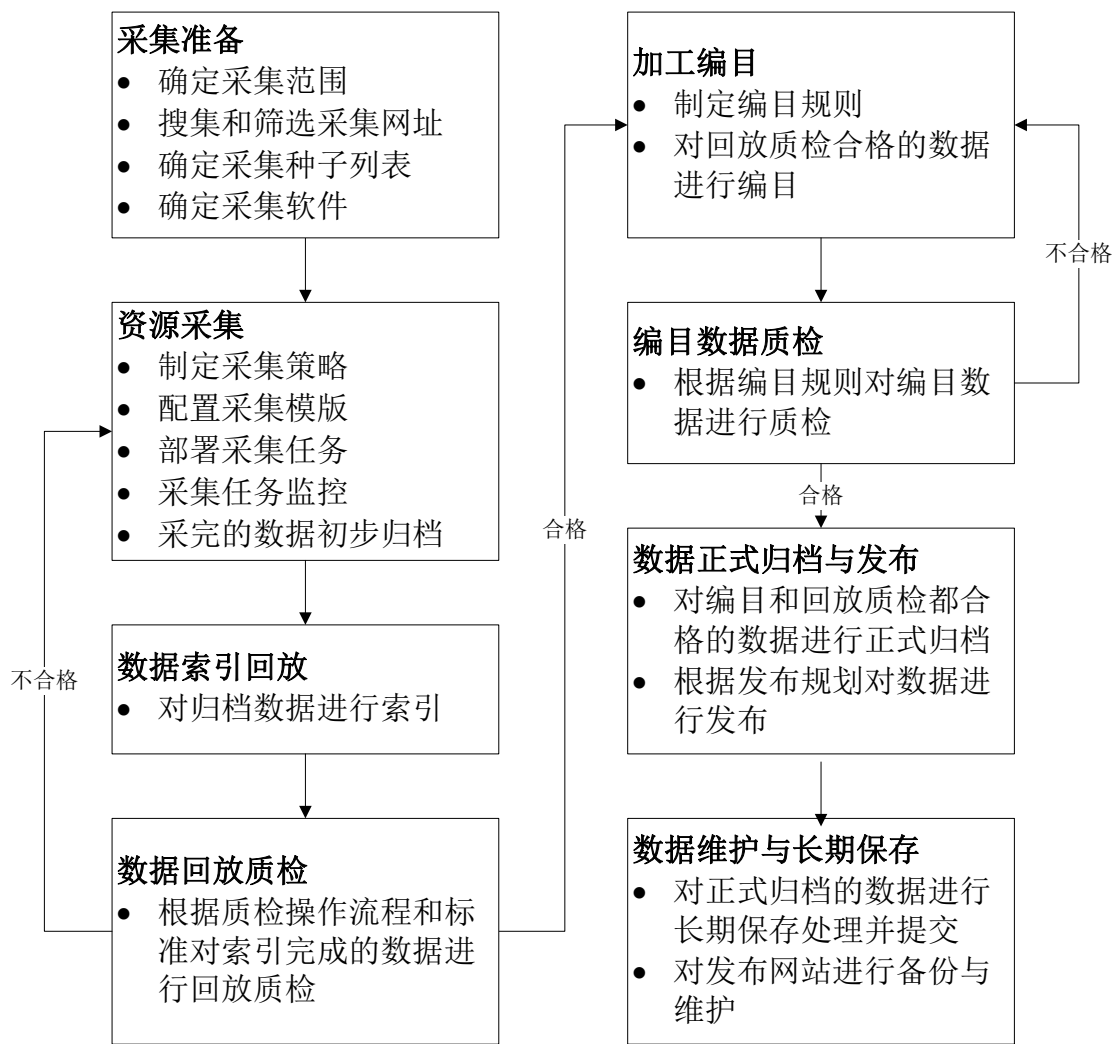


第三部分

系统操作



第三部分：系统操作



网络资源保存工作流程



第三部分：系统操作



采集准备

资源采集开始前的所有工作，主要包括：

- (1) 根据采选原则，对网络资源进行梳理，确定采集范围；
- (2) 完成采集目标种子网址（ URL地址 ）的搜集和筛选，整理成网络采集的资源种子网址列表。



第三部分：系统操作



资源采集

使用采集软件，对种子网址列表进行采集，具体过程包括：

- (1) 根据业务需求，制定采集策略。
- (2) 根据采集策略，在采集软件中对各类种子网址（网页、专题、网站）进行采集模版的配置。
- (3) 选择相应的采集模版对各种子网址进行采集任务部署，开始采集。
- (4) 对采集过程监控并进行相应处理，直至采集进程完成。
- (5) 对采集完成的任务数据进行初步归档整理。



第三部分：系统操作



数据索引回放

使用索引软件对采集完成的数据进行索引处理。

- 本质是对采集的各类对象数据进行组织后生成统一的相对路径（回放地址）；
- 用户可用浏览器通过回放地址对采集资源进行回放显示。



第三部分：系统操作



资源编目

- 制定网络资源元数据著录规范，根据规范对采集完成并且回放质检合格的数据进行资源的元数据著录。
- 元数据规范符合DC标准，并有适当的扩展和修改。



第三部分：系统操作



数据质检

- 回放质检
根据质检操作流程和标准，对完成索引的数据进行回放质检。合格数据进行正式归档，不合格数据根据情况进行处理。
- 编目质检
根据编目规则对编目数据进行质检。不合格数据进行修改并二次质检，直到全部质检合格。



第三部分：系统操作



发布服务

- 建设内容发布网站，对采集的网络资源进行有序组织和整理，按照分类和专题进行整合发布。
- 网站建设可充分利用信息技术创新，增加可视化、全文检索、人机交互等功能，提升服务效果。



第三部分：系统操作



数据保存

- 对正式归档数据进行长期保存处理，定期提交长期保存管理部门进行长期保存。
- 对发布网站数据库进行更新、备份与维护。



第三部分：系统操作

推广工程网络资源保存与服务系统

账号登录

用户名

请输入用户名！

密码

验证码

g A Q W

登录

客户端操作系统为Windows，需保证稳定地连接互联网；
为保证显示和功能正常，推荐使用谷歌浏览器；
使用系统前需登录，按照相应权限进行操作。



第三部分：系统操作

系统整体布局

推广工程网络资源保存与服务系统 2018/11/7 上午9:36:18

报警信息 通知 读者来信 用户名 退出

系统名称和logo 时间

欢迎使用“推广工程网络资源保存与服务系统”

采集中的任务(个)	25	累计采集完成(个)	377	活跃线程(个)	27	当前排队任务(个)	51
累计采集时间	2590天18小时12分51秒	平均采集速度(KB/s)	16.6052	运行服务器(个)	9	存储状态(%)	55

一级菜单栏目

功能区 展示区 页头区



第三部分：系统操作

首页——概况

推广工程网络资源保存与服务系统 2018/12/5 上午9:06:34

4133 33 0 gdzadmin

首页

概况

异常报警消息 ^N

通知 ³³

用户来信

采集模块

编目模块

回放模块

欢迎使用“推广工程网络资源保存与服务系统”

采集中的任务(个)	0	累计采集完成(个)	22	活跃线程(个)	0	当前排队任务(个)	0
累计采集时间	122天18小时42分1秒	平均采集速度(KB/s)	8.9677	运行服务器(个)	3	存储状态(%)	6

首页——异常报警信息

推广工程网络资源保存与服务系统 2018/12/5 上午9:11:11

4133 33 0 gdzadmin

首页

概况

异常报警消息 ^N

通知 ³³

用户来信

采集模块

编目模块

回放模块

内容管理和发布模块

数据保存模块

服务器报警 ^N | 索引报警 ²⁰

服务器名称: 全部 状态: 全部 报警时间: 任务ID: 任务名称: 服务器线程: 全部 搜索

批量删除

<input type="checkbox"/>	序号	报警时间	服务器	线程	采集任务名称	任务ID	报警原因	操作
<input type="checkbox"/>	1	2018-10-02 20:11:34	广东省立中山图书馆-1	Heritrix	-	-	close() ros[ToeThread #29: http://news.ifeng.com/a/20180611/F_Widget...	删除 查看详情
<input type="checkbox"/>	2	2018-10-02 19:56:30	广东省立中山图书馆-1	Heritrix	-	-	close() ros[ToeThread #29: http://news.ifeng.com/a/20180611/F_Widget...	删除 查看详情



第三部分：系统操作

首页——通知

推广工程网络资源保存与服务系统 2018/12/5 上午9:17:08

4133 83 0 gdzadmin

首页

概况

异常报警消息 N

通知 33

用户来信

采集模块

编目模块

回放模块

采集任务完成通知 21 | 索引完成通知 12

任务名称 查看状态 全部 创建时间 服务器 全部 服务器线程 全部 搜索

批量删除

<input type="checkbox"/>	序号	时间	采集任务名称	任务ID	服务器	线程名	消息内容
<input type="checkbox"/>	1	2018-10-02 20:28:02	shanghe-d-10	190020180921110844aa011	广东省立中山图书馆-1	_Heritrix1	此任务已经采集完成, 点击查看详情
<input type="checkbox"/>	2	2018-10-01 23:29:29	21yidaiyilu2017-B2-D-431	190020180912112524ab009	广东省立中山图书馆-1	_Heritrix1	此任务已经采集完成, 点击查看详情

首页——用户来信

推广工程网络资源保存与服务系统 2018/12/5 上午9:17:59

4133 83 0 gdzadmin

首页

概况

异常报警消息 N

通知 33

用户来信

采集模块

编目模块

用户来信

姓名 联系电话 处理状态 全部 创建时间 搜索

邮箱

批量删除

<input type="checkbox"/>	序号	时间	姓名	邮箱	联系电话	处理状态	来自	操作
没有找到匹配的记录								



第三部分：系统操作

采集模块——采集任务部署管理——界面布局

推广工程网络资源保存与服务系统 2018/12/5 上午9:39:39

4193 33 0 gdzadmin

采集任务列表 | 模板管理 | 质检不合格任务列表 — 列表管理

条件检索

任务名称 状态 已完成 服务器 全部

任务 ID 线程 _Heritrix4 机构 全部 搜索

创建时间

创建单个采集项目 批量导入任务 批量导出任务 — 资源采集

批量删除 批量开始 批量暂停 批量继续 自动刷新

序号	采集任务名称	任务ID	所属机构	状态	已采时间	服务器名称	服务器IP	线程	当前速度	平均速度	已采文档数	剩余文档数	已采容量	进度	任务创建时间	操作
1	shanghe-d-4	190020180921101635aa004	广东省立中山图书馆	已完成	1d4h24m53s	广东省立中山图书馆-1	10.10.123.198	_Heritrix4	0 KB/s	0 KB/s	1438	4	30.12 M	100%	2018-09-21 10:16:35	操作
2	21yidailyu2017-B2-D-423	190020180912112522ab007	广东省立中山图书馆	已完成	20h48m39s	广东省立中山图书馆-1	10.10.123.198	_Heritrix4	0 KB/s	2 KB/s	2200	0	148.74 M	100%	2018-09-12 11:25:22	查看日志 编辑 删除 检查点 查看详情 归档路径

显示第 1 到第 2 条记录，总共 2 条记录

资源列表及操作



第三部分：系统操作

采集模块——采集任务部署管理——新建单个采集任务

创建采集任务

url

任务id

任务名称

任务描述

选择服务器

线程

选择模板

备注

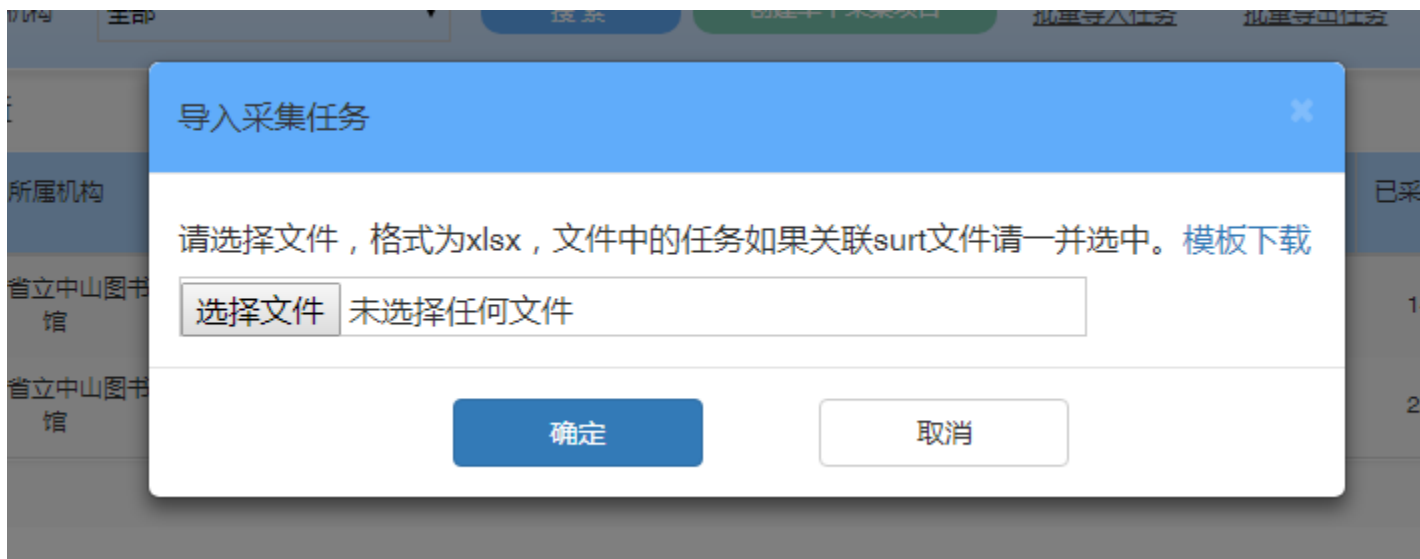
未选择任何文件

任务创建完成后立即开始采集 采集完成后自动开始索引



第三部分：系统操作

采集模块——采集任务部署管理——批量导入采集任务



第三部分：系统操作

采集模块——采集任务部署管理——模板管理

推广工程网络资源保存与服务系统 2018/12/5 上午10:18:40

4133 33 0

gdzsadmin

采集任务列表 | 模板管理 | 质检不合格任务列表

参数模板名称 搜索

创建参数模板

批量删除 导入参数模板 导出参数模板

<input type="checkbox"/>	序号	模板名称	操作
<input type="checkbox"/>	1	default	导出 编辑 删除
<input type="checkbox"/>	2	domainCrawl	导出 编辑 删除
<input type="checkbox"/>	3	domainCrawl-h6	导出 编辑 删除
<input type="checkbox"/>	4	domainCrawl-noh	导出 编辑 删除
<input type="checkbox"/>	5	events	导出 编辑 删除
<input type="checkbox"/>	6	events-d	导出 编辑 删除
<input type="checkbox"/>	7	events-d-gov	导出 编辑 删除

显示第 1 到第 7 条记录，总共 7 条记录



第三部分：系统操作

采集模块——采集任务部署管理——质检不合格列表

推广工程网络资源保存与服务系统 2018/12/5 上午10:24:58

4199 33 0 gdzadmin

采集任务列表 | 模板管理 | 质检不合格任务列表

任务名称 采集状态 全部 任务创建时间

任务 ID 服务器 全部 质检不合格原因 全部 搜索

批量删除 批量重采

序号	采集任务名称	任务ID	状态	已采时间	服务器名称	服务器IP	姓名	平均速度	剩余文档数	已采文档数	已采容量	进度	任务创建时间	不合格原因	操作
没有找到匹配的记录															



第三部分：系统操作

采集模块——采集任务调度管理

推广工程网络资源保存与服务系统 2018/12/5 上午10:27:42

4133 33 0 gdzadmin

首页 采集模块 采集任务部署管理 采集任务调度管理 编目模块 回放模块 内容管理和发布模块 数据保存模块 统计模块 用户管理模块

服务器监控

服务器名称 服务器IP 状态 全部 搜索

添加服务器

批量删除 heritrix重启说明

序号	服务器名称	服务器IP	所属机构	状态	已用线程	空闲线程	操作
1	广东省立中山图书馆-1	10.10.123.198	广东省立中山图书馆	正常	0	8	编辑 增减空闲线程 删除 查看详情 关闭Heritrix
2	广东省立中山图书馆-2	10.10.123.198	广东省立中山图书馆	正常	0	8	编辑 增减空闲线程 删除 查看详情 关闭Heritrix
3	广东省立中山图书馆-3	10.10.123.198	广东省立中山图书馆	正常	0	6	编辑 增减空闲线程 删除 查看详情 关闭Heritrix

创建服务器

服务器IP 输入可包含汉字,字母,数字,连接符和下划线

ip输入错误!

JMX端口 8849

服务器名称 任意字符

请输入服务器名称,至少4位,不能超过30位

线程数

取消 添加

服务器监控 | 广东省立中山图书馆-1详情 X

采集任务名称 任务ID 线程状态 全部 搜索

增减空闲线程

批量删除

序号	线程	状态	采集任务名称	任务ID	任务状态	采集速度	采集进度	对任务的操作	操作
1	_Heritrix	空闲	-	-	-	-	-	-	删除线程
2	_Heritrix1	空闲	-	-	-	-	-	-	删除线程
3	_Heritrix2	空闲	-	-	-	-	-	-	删除线程
4	_Heritrix3	空闲	-	-	-	-	-	-	删除线程
5	_Heritrix4	空闲	-	-	-	-	-	-	删除线程
6	_Heritrix5	空闲	-	-	-	-	-	-	删除线程
7	_Heritrix6	空闲	-	-	-	-	-	-	删除线程
8	_Heritrix7	空闲	-	-	-	-	-	-	删除线程

当前使用中的线程为 0 个,共 8 个线程

空闲线程数 8

显示第 1 到第 8 条记录, 总共 8 条记录

取消 确定



第三部分：系统操作

编目模块——编目管理——界面布局

推广工程网络资源保存与服务系统 2018/12/5 下午1:28:46

4133 33 0 gdzadmin

首页 采集模块 编目模块 编目管理 编目模板 编目专题 资源类型 回放模块 内容管理和发布模块 数据保存模块 统计模块 用户管理模块 维护备份模块

编目URL列表 | 编目不合格URL列表

任务名称 编目状态 未审核 格式 全部 内容分类 国别 全部

采集地址 质检结果 全部 编目模板 全部 摘要 语种

名称 任务状态 全部 数据提交单位 全部 CDOI 创建时间

姓名 全部 创建者 类型 全部 关键词 服务器 全部

任务ID 版权拥有 编目专题 全部 其他名称

专题ID 访问地址 快照日期 采集日期

条件检索

搜索 新建编目 批量审核

导出数据 导入数据 批量删除 下载编目说明 导入导出说明 批量操作

单个操作

CDDOI	语种	所属专题	专题ID	时间范围	空间范围	版权所有者	使用权限	姓名	数据提交单位	创建时间	编目状态	编目不合格原因	编目质检操作人	任务采集完成情况	回放质检结果	发布状态	操作
36	chi	一带一路	ZT-2018-010	42086		北京商报网	局域网访问	中山用户3	广东省立中山图书馆	2018-09-12 15:15:19	未审核	-	-	已完成	未质检	未发布	编辑 删除 编目审核
34	chi	一带一路	ZT-2018-010	42086		新浪网	局域网访问	中山用户3	广东省立中山图书馆	2018-09-12 15:15:14	未审核	-	-	已完成	未质检	未发布	编辑 删除 编目审核
33	chi	一带一路	ZT-2018-010	42086		中国经济网	局域网访问	中山用户3	广东省立中山图书馆	2018-09-12 15:15:12	未审核	-	-	采集中异常	未质检	未发布	编辑 删除 编目审核
31	chi	一带一路	ZT-2018-010	42086		中青在线	局域网访问	中山用户3	广东省立中山图书馆	2018-09-12 15:15:10	未审核	-	-	已完成	未质检	未发布	编辑 删除 编目审核

编目信息及状态显示

显示第 1 到第 4 条记录，总共 4 条记录



第三部分：系统操作

编目模块——编目管理——新建单个编目

新建编目

编目模板

注意：请谨慎选择编目模板——编目记录创建后只能修改该条编目记录的字段内容，若更换编目模板则需删除该条编目记录后重新创建！

名称

其它题名

创建者

国别

其他责任者

关键词

内容分类

摘要

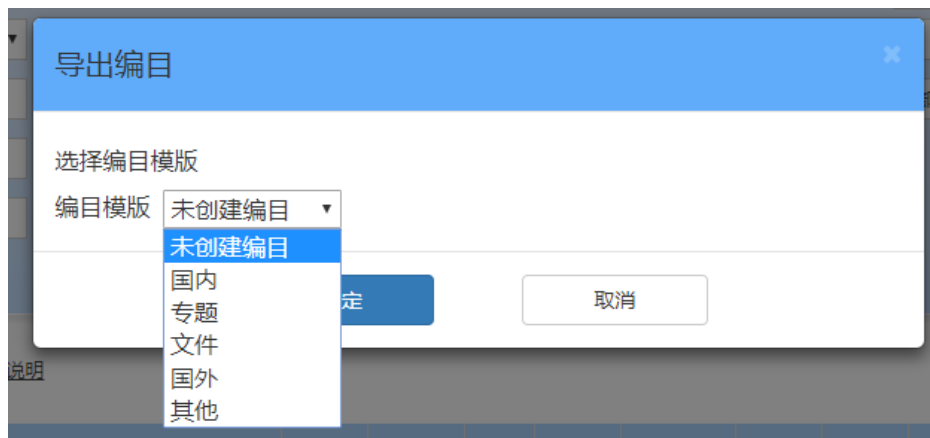
出处附注



第三部分：系统操作

编目模块——编目管理——编目信息的批量导入导出

导入导出说明



第三部分：系统操作

编目模块——编目管理——编目信息的审核

编目审核

编目审核是否合格

姓名	数据提交单位	创建时间	编目状态	编目不合格原因	编目质检
中山用户3	广东省立中山图书馆	2018-09-12 15:15:19	未审核	-	
中山用户3	广东省立中山图书馆	2018-09-12 15:15:14	未审核	-	
中山用户3	广东省立中山图书馆	2018-09-12 15:15:10	未审核	-	

编目不合格原因

- 网站名称不规范
- 摘要内容不完整/不成句/信息不足/语法错误
- 摘要翻译内容不完整/不成句/信息不足/语法错误
- 网站语种不完整
- 其他



第三部分：系统操作

编目模块——编目管理——编目模板

推广工程网络资源保存与服务系统 2018/12/5 下午2:21:05

4133 83 0 gdzadmin

首页 < 采集模块 < 编目模块 > 编目管理 编目模板 编目专题 资源类型 回放模块 < 内容管理和发布模块 < 数据保存模块 < 统计模块 <

编目模板 状态 全部

<input type="checkbox"/>	序号	编目模板	描述	状态	排序	操作
<input type="checkbox"/>	1	国内网站		开启	<input type="button" value="上"/> <input type="button" value="下"/> <input type="button" value="顶"/> <input type="button" value="底"/>	<input type="button" value="编辑"/> <input type="button" value="暂停"/> <input type="button" value="导出新增模板"/> <input type="button" value="导出更新模板"/> <input type="button" value="删除"/>
<input type="checkbox"/>	2	专题资源		开启	<input type="button" value="上"/> <input type="button" value="下"/> <input type="button" value="顶"/> <input type="button" value="底"/>	<input type="button" value="编辑"/> <input type="button" value="暂停"/> <input type="button" value="导出新增模板"/> <input type="button" value="导出更新模板"/> <input type="button" value="删除"/>
<input type="checkbox"/>	3	国外网站		开启	<input type="button" value="上"/> <input type="button" value="下"/> <input type="button" value="顶"/> <input type="button" value="底"/>	<input type="button" value="编辑"/> <input type="button" value="暂停"/> <input type="button" value="导出新增模板"/> <input type="button" value="导出更新模板"/> <input type="button" value="删除"/>
<input type="checkbox"/>	4	规范字段	默认模板	开启	<input type="button" value="上"/> <input type="button" value="下"/> <input type="button" value="顶"/> <input type="button" value="底"/>	<input type="button" value="编辑"/> <input type="button" value="暂停"/> <input type="button" value="导出新增模板"/> <input type="button" value="导出更新模板"/> <input type="button" value="删除"/>



第三部分：系统操作

编目模块——编目管理——编目专题

专题ID 专题名称 状态 **全部** 专题时间 搜索

新建专题

批量删除 批量暂停 批量开启

序号	专题名称	专题ID	专题时间	专题子目	创建时间	描述	状态	排序	操作
1	一带一路	ZT-2018-010	2017-09-08		2018-08-30 16:47:11		开启	上 下 顶 底	编辑 暂停 删除

编辑详情

专题名称

专题ID

专题时间

专题子目

描述

确定

确定 取消



第三部分：系统操作

编目模块——编目管理——资源类型

推广工程网络资源保存与服务系统 2018/12/5 下午2:35:05

4133 33 0 gdzadmin

首页 采集模块 编目模块 编目管理 编目模板 编目专题 资源类型 回放模块 内容管理和发布模块 数据保存模块 统计模块 用户管理模块 维护备份模块

资源类型 状态 全部 搜索

新建资源类型

批量删除 批量暂停 批量开启

<input type="checkbox"/>	序号	资源类型	描述	状态	排序	操作
<input type="checkbox"/>	1	整站	整站	开启	上 下 顶 底	编辑 暂停 删除
<input type="checkbox"/>	2	频道	频道	开启	上 下 顶 底	编辑 暂停 删除
<input type="checkbox"/>	3	网页	网页	开启	上 下 顶 底	编辑 暂停 删除
<input type="checkbox"/>	4	资源	资源	开启	上 下 顶 底	编辑 暂停 删除
<input type="checkbox"/>	5	其他	其他	开启	上 下 顶 底	编辑 暂停 删除

显示第 1 到第 5 条记录，总共 5 条记录



第三部分：系统操作

回放模块——索引管理

推广工程网络资源保存与服务系统 2018/12/5 下午2:41:20

4133 33 0 gdzadmin

索引任务列表

任务名称 URL索引状态 全部 采集完成时间 姓名 全部

任务ID 全文检索状态 全部 质检情况 全部 搜索

[批量删除索引](#) [批量索引](#)

<input type="checkbox"/>	任务名称	任务ID	姓名	采集完成终止时间	抓取文档数量	抓取文档容量	URL索引	全文索引	质检情况	操作
<input type="checkbox"/>	shanghe-d-10	19002018092111084aa011	中山机构管理员	2018-10-02 20:28:02	1271	333542653	<input type="button" value="建立索引"/>	<input type="button" value="建立索引"/>	未质检	<input type="button" value="重建索引"/>
<input type="checkbox"/>	shanghe-d-5	190020180921101636aa005	中山机构管理员	2018-09-21 12:40:27	386	6104029	http://183.63.187.57:8099/openwayback/...	索引已完成	质检合格	<input type="button" value="重建索引"/>
<input type="checkbox"/>	shanghe-d-4	190020180921101635aa004	中山机构管理员	2018-09-22 14:42:03	1438	31586380	http://183.63.187.57:8099/openwayback/...	索引已完成	质检合格	<input type="button" value="重建索引"/>



第三部分：系统操作

回放模块——索引管理

推广工程网络资源保存与服务系统 2018/12/5 下午2:46:21

4138 33 0 gdzadmin

首页 采集模块 编目模块 回放模块 索引管理 质检管理 内容管理和发布模块 数据保存模块 统计模块 用户管理模块 维护备份模块

质检任务列表 | 质检不合格任务列表

任务名称 质检情况 全部 采集完成/终止时间 搜索

任务 ID

批量删除索引 批量重建索引 批量重采任务 批量删除任务

<input type="checkbox"/>	任务名称	任务ID	姓名	采集完成/终止时间	抓取文档数量	抓取文档容量	URL索引	质检情况	操作
<input type="checkbox"/>	shanghe-d-5	190020180921101636aa005	中山机构管理员	2018-09-21 12:40:27	386	6104029	http://183.63.187.57:8009/openwayback/...	质检合格	操作
<input type="checkbox"/>	shanghe-d-4	190020180921101635aa004	中山机构管理员	2018-09-22 14:42:03	1438	31586380	http://183.63.187.57:8009/openwayback/...	质检合格	质检审核 重建索引 重采任务 删除任务 查看任务详情
<input type="checkbox"/>	shanghe-d-3	190020180921101634aa003	中山机构管理员	2018-09-21 19:56:14	4431	1833962010	http://183.63.187.57:8009/openwayback/...	质检合格	查看任务详情
<input type="checkbox"/>	shanghe-d-2	190020180921101634aa002	中山机构管理员	2018-09-21 22:13:10	4414	1833879609	http://183.63.187.57:8009/openwayback/...	未质检	
<input type="checkbox"/>	shanghe-d-1	190020180921101633aa001	中山机构管理员	2018-09-25 17:22:41	4617	1893971434	http://183.63.187.57:8009/openwayback/...	未质检	
<input type="checkbox"/>	21yidaiyilu2017-B2-D-165	190020180912151515ad005	中山用户3	2018-09-12 15:27:48	223	8439609	http://183.63.187.57:8009/openwayback/...	未质检	操作

显示第 1 到第 0 条记录，总共 0 条记录



第三部分：系统操作

内容管理和发布模块——内容管理

推广工程网络资源保存与服务系统 2018/12/5 下午2:53:05

4133 33 0 gdzadmin

首页 采集模块 编目模块 回放模块 内容管理和发布模块 内容管理 网站管理 界面管理 发布管理 数据保存模块 统计模块 用户管理模块

编目质检合格任务列表

任务名称: 国别: 格式: 数据提交单位: 采集时间:

采集地址: 语种: 编目模板: 内容分类: 快照日期:

任务ID: 发布状态: 专题ID: CDOI:

名称: 创建者: 类型: 关键词:

姓名: 版权信息: 编目专题: 其他名称:

格式:

序号	名称	其他名称	创建者	国别	其他责任者	关键词	内容分类	摘要	摘要翻译	出处附注	其他附注	快照日期	采集日期	类型	格式	馆藏原生数字资源记录标识号	任务ID
1	【图解天下】第262期：一图看懂...			中国		上海合作组织	门户专题			国际在线		2018-09-25 17:50:28	2018-09-21 10:21:37	网页	WARC		19002018092110163
2	亲亲上合 政协有为			中国		上海合作组织	门户专题			东方网		2018-09-25 16:54:34	2018-09-21 10:16:38	网页	WARC		19002018092110163



第三部分：系统操作

内容管理和发布模块——网站管理

推广工程网络资源保存与服务系统 2018/12/5 下午2:58:26

4133 33 0 gdzadmin

首页 采集模块 编目模块 回放模块 内容管理和发布模块 内容管理 网站管理 界面管理 发布管理 数据保存模块 统计模块 用户管理模块

网站管理

网站名称 姓名 创建时间 更新时间 更新人

<input type="checkbox"/>	网站名称	创建时间	姓名	最后更新时间	更新人	状态	操作
<input type="checkbox"/>	测试026	2018-09-26 15:15:57	中山机构管理员	2018-09-26 15:48:03	中山机构管理员	开启	操作
<input type="checkbox"/>	一带一路	2018-09-12 16:59:20	中山用户1	2018-09-12 16:59:20	-	开启	编辑 栏目管理 停用 删除 发布

显示第 1 到第 2 条记录，总共 2 条记录

编辑

网站名称

第三部分：系统操作

内容管理和发布模块——网站管理——栏目管理

推广工程网络资源保存与服务系统 2018/12/5 下午3:20:51

4133 33 0

gdzadmin

- 首页
- 采集模块
- 编目模块
- 回放模块
- 内容管理和发布模块
 - 内容管理
 - 网站管理
 - 界面管理
 - 发布管理
- 数据保存模块
- 统计模块
- 用户管理模块
- 维护备份模块

网站管理 | 栏目管理(测试926) X

栏目名称 栏目类型 全部 模板 全部 状态 全部 搜索

添加栏目

批量删除 批量使用

栏目名称	栏目类型	模板	状态	移动	操作
首页	首页	template_1	开启		操作
专题保存	专题列表页	template_1	开启	↑ ↓ →	操作
2018上海合作组织青岛峰会	专题资源列表页	template_1	开启	↑ ↓ ← →	
改革开放	专题资源列表页	template_1	开启	↑ ↓ ← →	
网站保存	分类列表页	template_1	开启	↑ ↓ →	
国内网站保存	文件资源列表页	template_1	开启	↑ ↓ ← →	

- 添加子栏目
- 编辑
- 分类配置
- 停用
- 删除

新建栏目

栏目名称

栏目类型

上级栏目

备注

栏目名称不能为空

确定 取消

添加资源

任务名称: 名称: 数据提交单位: 全部 采集日期:

编目模板: 全部 内容分类: 专题创建时间: 资源专题: 全部 搜索 批量添加

序号	编目模板	任务名称	名称	国别	语种	内容分类	所属专题	专题ID	快照日期	采集日期	类型	数据提交单位	创建时间	操作
1	专题资源	shanghe-d-4	亲亲上合 政协有力	中国	chi	门户专题	2018上海合作组...	ZT-2018-011	2018-09-25 16:54:34	2018-09-21 10:16:38	网页	广东省立中山图...	2018-09-21 10:16:36	添加
2	专题资源	shanghe-d-3	【图解天下】第252期: 一图看懂...	中国	chi	门户专题	2018上海合作组...	ZT-2018-011	2018-09-25 17:50:28	2018-09-21 10:21:37	网页	广东省立中山图...	2018-09-25 15:35:28	添加



第三部分：系统操作

内容管理和发布模块——界面管理

推广工程网络资源保存与服务系统 2018/12/5 下午3:25:49

4133 33 0

gdzsadmin

- 首页
- 采集模块
- 编目模块
- 回放模块
- 内容管理和发布模块
 - 内容管理
 - 网站管理
 - 界面管理
 - 发布管理
- 数据保存模块
- 统计模块
- 用户管理模块
- 维护备份模块

首页模板

首页模板1

首页模板2

首页模板3

The screenshot displays the '界面管理' (Interface Management) section of the system. It features a sidebar on the left with a menu where '界面管理' is selected. The main content area is titled '首页模板' (Homepage Templates) and shows three preview cards for different homepage designs: '首页模板1' (Homepage Template 1), '首页模板2' (Homepage Template 2), and '首页模板3' (Homepage Template 3). Each template preview shows a complex web layout with various banners, text blocks, and images. The system's top navigation bar includes the title '推广工程网络资源保存与服务系统', the current date and time '2018/12/5 下午3:25:49', and notification icons for messages (4133), voice (33), and email (0). The user 'gdzsadmin' is logged in.

第三部分：系统操作

内容管理和发布模块——发布管理

推广工程网络资源保存与服务系统 2018/12/5 下午3:27:27

4133 83 0 gdzadmin

发布历史

发布名称 发布地址 发布状态 全部 发布时间 搜索

<input type="checkbox"/>	网站名称	发布名称	发布时间	发布地址	状态
<input type="checkbox"/>	测试026	test026	2018-09-26 15:48:03	http://183.63.187.57:81/webz/web_30/首页.html	成功

显示第 1 到第 1 条记录，总共 1 条记录

内容管理
网站管理
界面管理
发布管理



第三部分：系统操作

数据保存模块

推广工程网络资源保存与服务系统 2018/12/5 下午3:30:43

4133 33 0 gdzadmin

任务数据保存

数据提交单位 全部 类型 全部 采集时间

搜索

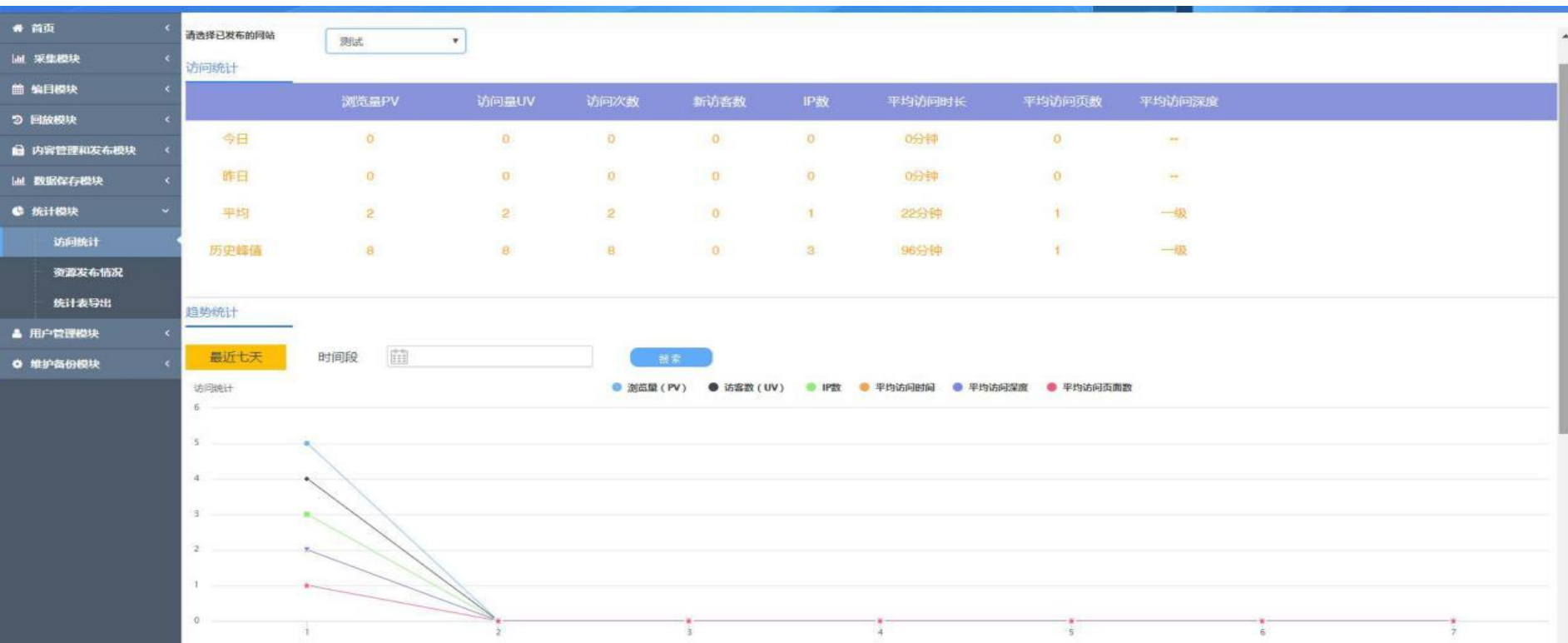
批量打包并生成报表 批量删除

<input type="checkbox"/>	序号	名称	任务名称	任务ID	采集时间	所属专题	专题ID	类型	数据提交单位	相对路径	打包状态	打包进度	报表	查看日志
<input type="checkbox"/>	1	【图解天下】第252期：一图看懂青岛峰会重要成果	shanghe-d-3	190020180921101634aa003	2018-09-25 15:35:28	2018上海合作组织青岛峰会	ZT-2018-011	专题资源	广东省立中山图书馆		未打包	0%	无报表	查看
<input type="checkbox"/>	2	亲亲上台 政协有为	shanghe-d-4	190020180921101635aa004	2018-09-21 10:16:36	2018上海合作组织青岛峰会	ZT-2018-011	专题资源	广东省立中山图书馆		未打包	0%	无报表	查看



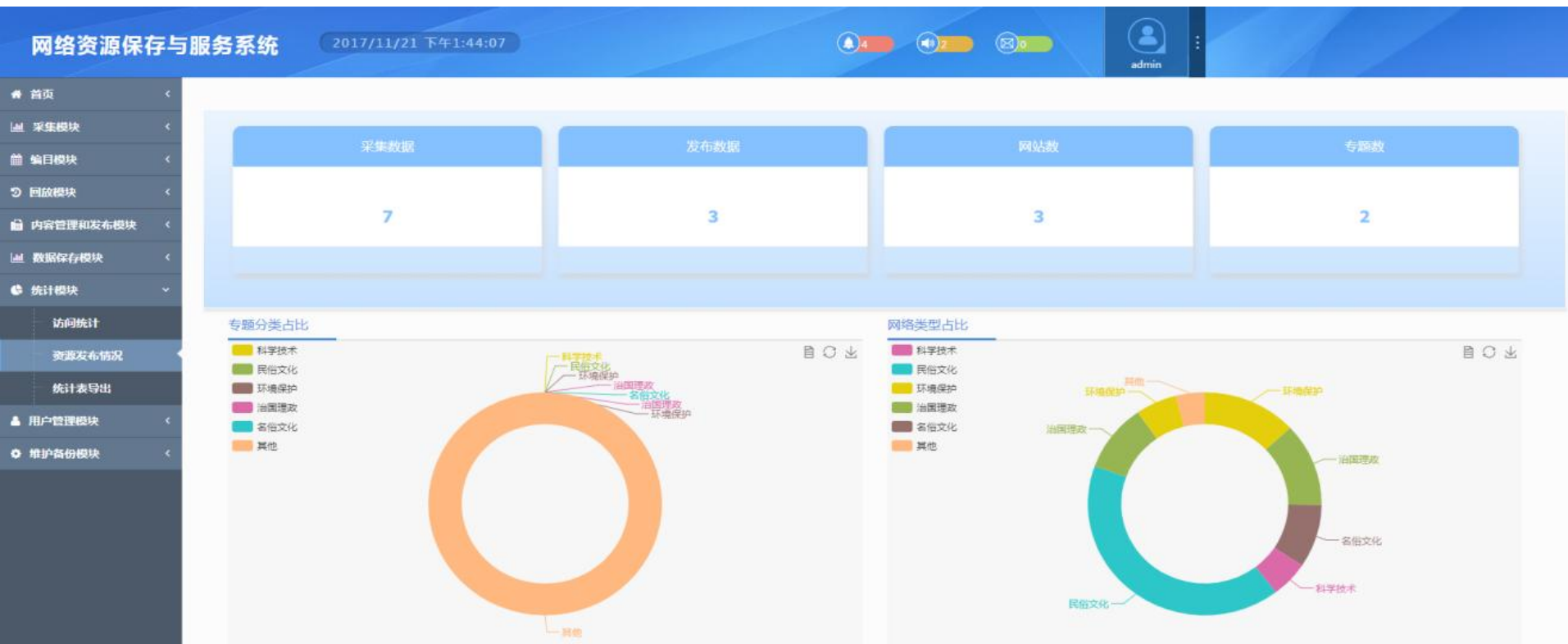
第三部分：系统操作

统计模块——访问统计



第三部分：系统操作

统计模块——资源发布情况



第三部分：系统操作

统计模块——统计表导出

[增量统计表](#) | [汇总统计表](#) | [分类明细简表](#) | [分类明细详表](#)

时间段

查询

增量统计表

国内网站 专题资源 国外网站 规范字段

2018年08月17日 - 2018年12月05日 网络信息资源增量统计表 [下载增量统计表](#)

资源种类	资源格式	建设情况						发布情况						备注
		建设数量 (个)			建设容量 (G)			发布数量 (个)			发布容量 (G)			
		月增量	年增量	总量	月增量	年增量	总量	月增量	年增量	总量	月增量	年增量	总量	
专题资源	WARC	0	2	2	0.00000	0.00000	0.00000	0	2	2	0.00000	0.00000	0.00000	填写

制表单位：广东省立中山图书馆 制表时间：2018年12月05日 统计人：中山机构管理员

[增量统计表](#) | [汇总统计表](#) | [分类明细简表](#) | [分类明细详表](#)

时间段

查询

汇总统计表

国内网站 专题资源 国外网站 规范字段

2018年08月17日 - 2018年12月05日 网络信息资源汇总统计表 [下载汇总统计表](#)

资源种类	资源格式	建设情况				发布情况		备注
		URL数量 (个)	采集次数 (次)	压缩前大小 (TB)	压缩后大小 (TB)	种子数 (个)	发布次数 (次)	
专题资源	WARC	2	2	0.00170	0.00000	2	2	填写

制表单位：广东省立中山图书馆 制表时间：2018年12月05日 统计人：中山机构管理员



第三部分：系统操作

统计模块——统计表导出

增量统计表 | 汇总统计表 | 分类明细简表 | 分类明细详表

时间段

分类明细简表

查询

国内网站 专题资源 国外网站 规范字段

网络信息资源分类明细简表 [下载分类明细简表](#)

资源种类	序号	任务ID	名称	采集情况				存储情况	发布情况		备注
				URL数量(个)	采集次数(次)	压缩前大小(GB)	压缩后大小(GB)	存储目录	种子数(个)	发布次数(次)	
专题资源	1	190020180921101635aa004	shanghe-d-4	1	1	0.02942	0	/mnt/glusterfs/nraas/jobs/10.10.123.198-8848/shanghe-d-4-20180921021636010	1	1	填写
	2	190020180921101634aa003	shanghe-d-3	1	1	1.70801	0	/mnt/glusterfs/nraas/jobs/10.10.123.198-8848/shanghe-d-3-20180921021635266	1	1	填写
合计				2	2	1.73743	0.00000	-	2	2	填写

制表单位：广东省立中山图书馆 制表时间：2018年12月05日 统计人：中山机构管理员

增量统计表 | 汇总统计表 | 分类明细简表 | 分类明细详表

时间段

分类明细详表

查询

国内网站 专题资源 国外网站 规范字段

网络信息分类明细详表 [下载分类明细详表](#)

资源序号	任务ID	名称	采集情况										存储情况		质检情况						
			任务年份	URL数量(个)	采集网址	采集日期	采集次数(次)	运行时间	压缩前大小(字节)	压缩后大小(字节)	压缩前大小(GB)	压缩后大小(GB)	采集服务器IP	采集人	存储目录	存储人	编码目录	数据目录	是否发布	种子数(个)	发布次数(次)
1	190020180921101635aa004	shanghe-2018-d-4	2018-09-21	1	http://news.eastday.com/eastday/13news/auto/news/society/20180612/u7ai7804307.html	2018-09-21	1	1d4h24m53s	31586380	0	0.02942	0	10.10.123.198-8848	中山机构管理员	/mnt/glusterfs/nraas/jobs/10.10.123.198-8848/shanghe-d-4-20180921021636010	中山机构管理员	中山机构管理员	是	1	1	http://183.63.187.5
2	190020180921101634aa003	shanghe-2018-d-3	2018-09-21	1	http://news.cri.cn/20180611/7cbc2a3b-3a56-da54-7697-f32dfdc836d9.html	2018-09-21	1	9h39m10s	1833962010	0	1.70801	0	10.10.123.198-8848	中山机构管理员	/mnt/glusterfs/nraas/jobs/10.10.123.198-8848/shanghe-d-3-20180921021635266	中山机构管理员	中山机构管理员	是	1	1	http://183.63.187.5
合计			-	2	-	-	2	1d14h4m3s	1865548390	0	1.73743	0	-	-	-	-	-	-	2	2	-
合计			-	2	-	-	2	1d14h4m3s	1865548390	0	1.73743	0	-	-	-	-	-	-	2	2	-

制表单位：广东省立中山图书馆 制表时间：2018年12月05日 统计人：中山机构管理员

第三部分：系统操作

用户管理模块——用户管理

推广工程网络资源保存与服务系统 2018/12/5 下午3:55:38

4133 83 0

gdzsadmin

用户管理

用户名 机构 全部 用户代码 姓名 邮箱

状态 全部 创建时间 角色 全部 查询

新建用户

批量启用 批量停用 批量删除 导入数据 导出数据

<input type="checkbox"/>	序号	用户名	机构	用户代码	姓名	邮箱	角色	账号状态	创建时间	操作
<input type="checkbox"/>	1	gdzsadmin	广东省立中山图书馆	aa	中山机构管理员	wlb@zslib.com.cn	机构管理员	启用	2018-08-30 17:59:18	停用 编辑用户 删除
<input type="checkbox"/>	2	gdzsuser1	广东省立中山图书馆	ab	中山用户1	gdzs@user1.com	维护备份, 回放管理, 采集管理, 内容管理, 首页管理, 统计管理, 编目管理, 数据保存管理	启用	2018-09-11 16:40:26	停用 编辑用户 删除
<input type="checkbox"/>	3	gdzsuser2	广东省立中山图书馆	ac	中山用户2	gdzs@user2.com	维护备份, 采集管理, 回放管理, 内容管理, 首页管理, 统计管理, 编目管理, 数据保存管理	启用	2018-09-11 16:50:14	停用 编辑用户 删除
<input type="checkbox"/>	4	gdzsuser3	广东省立中山图书馆	ad	中山用户3	gdzs@user3.com	维护备份, 回放管理, 采集管理, 内容管理, 首页管理, 统计管理, 编目管理, 数据保存管理	启用	2018-09-11 16:50:52	停用 编辑用户 删除



第三部分：系统操作

用户管理模块——角色管理

推广工程网络资源保存与服务系统 2018/12/5 下午3:56:35

4133 33 0 gdzadmin

角色管理

角色 状态 全部

<input type="checkbox"/>	序号	角色	状态	用户	操作
<input type="checkbox"/>	1	首页管理	启用	10	<input type="button" value="停用"/> <input type="button" value="编辑"/> <input type="button" value="删除"/> 查看角色下用户
<input type="checkbox"/>	2	采集管理	启用	10	<input type="button" value="停用"/> <input type="button" value="编辑"/> <input type="button" value="删除"/> 查看角色下用户
<input type="checkbox"/>	3	编目管理	启用	10	<input type="button" value="停用"/> <input type="button" value="编辑"/> <input type="button" value="删除"/> 查看角色下用户
<input type="checkbox"/>	4	回放管理	启用	10	<input type="button" value="停用"/> <input type="button" value="编辑"/> <input type="button" value="删除"/> 查看角色下用户
<input type="checkbox"/>	5	内容管理	启用	10	<input type="button" value="停用"/> <input type="button" value="编辑"/> <input type="button" value="删除"/> 查看角色下用户
<input type="checkbox"/>	6	数据保存管理	启用	10	<input type="button" value="停用"/> <input type="button" value="编辑"/> <input type="button" value="删除"/> 查看角色下用户
<input type="checkbox"/>	7	统计管理	启用	18	<input type="button" value="停用"/> <input type="button" value="编辑"/> <input type="button" value="删除"/> 查看角色下用户
<input type="checkbox"/>	8	用户管理	启用	1	<input type="button" value="停用"/> <input type="button" value="编辑"/> <input type="button" value="删除"/> 查看角色下用户
<input type="checkbox"/>	9	维护备份	启用	10	<input type="button" value="停用"/> <input type="button" value="编辑"/> <input type="button" value="删除"/> 查看角色下用户

显示第 1 到第 0 条记录，总共 0 条记录



第三部分：系统操作

用户管理模块——机构管理

推广工程网络资源保存与服务系统 2018/11/15 下午2:58:57

648108 1067 19 admin

机构管理

机构名称 机构代码 状态 全部 查询

新建机构

[批量启用](#) [批量停用](#) [批量删除](#) [导入数据](#) [导出数据](#)

序号	机构名称	机构代码	联系人	用户数	后台地址	状态	操作
1	国家图书馆	0000	11	12	http://202.106.125.203/	启用	编辑 查看机构下用户
2	广东省立中山图书馆	1000	颜虎	4	http://183.63.167.61/login.html	启用	停用 编辑 删除 查看机构下用户
3	广西壮族自治区图书馆	2100	莫晓明	4	http://116.252.35.139/8888/login.html	启用	停用 编辑 删除 查看机构下用户
4	云南省图书馆	2600	金航	4	http://220.163.11.121/login.html	启用	停用 编辑 删除 查看机构下用户

显示第 1 到第 4 条记录，总共 4 条记录



第三部分：系统操作

维护备份模块——服务器监控

The screenshot displays a web-based server monitoring interface. On the left is a dark sidebar with a menu containing items like '首页', '采集模块', '编目模块', '回放模块', '内容管理和发布模块', '数据保存模块', '统计模块', '用户管理模块', and '维护备份模块'. The '维护备份模块' is expanded to show '服务器监控' and '备份'. The main content area is titled '服务器列表' and features search filters for '服务器名称' and '服务器IP', a '状态' dropdown menu set to '全部', and a '搜索' button. Below the filters is a table with 10 rows of server data. The table columns are: 序号, 服务器名称, 服务器IP, 已用线程, 空闲线程, 状态, and 操作. The '操作' column for the first row contains a '查看详情' link, which is highlighted with a red box. At the bottom of the table, there is a pagination control showing '显示第 1 到第 10 条记录, 总共 18 条记录 每页显示 10 条记录' and a '跳转' button with page numbers 1 and 2.

序号	服务器名称	服务器IP	已用线程	空闲线程	状态	操作
1	202.106.125.203-1	202.106.125.203	2	7	正常	查看详情
2	202.106.125.203-2	202.106.125.203	6	3	正常	查看详情
3	202.106.125.203-3	202.106.125.203	1	7	正常	查看详情
4	192.168.30.16-1	192.168.30.16	0	6	正常	查看详情
5	192.168.30.16-2	192.168.30.16	4	6	正常	查看详情
6	192.168.30.16-3	192.168.30.16	5	4	正常	查看详情
7	192.168.30.17-1	192.168.30.17	2	4	正常	查看详情
8	192.168.30.17-2	192.168.30.17	6	1	正常	查看详情
9	192.168.30.17-3	192.168.30.17	0	8	正常	查看详情
10	广东省立中山图书馆-1	10.10.125.198	0	8	正常	查看详情



第三部分：系统操作

维护备份模块——服务器监控

The screenshot displays a web-based server monitoring interface. On the left is a dark sidebar with navigation options: 首页, 采集模块, 轴目模块, 回放模块, 内容管理和发布模块, 数据保存模块, 统计模块, 用户管理模块, 维护备份模块, 服务器监控, and 备份. The main content area is titled '服务器列表 | 202.106.125.203-1详情 X'. It features search filters for '采集任务名称', '任务ID', and '线程状态', along with a '搜索' button and a '增加空闲线程' button. Below the filters is a table with columns: 序号, 线程, 状态, 采集任务名称, 任务ID, 任务状态, 采集速度, 采集进度, 对任务的操作, and 操作. The table lists 9 threads, with threads 2 and 4 in '采集中' (collecting) and '暂停' (paused) states respectively. Each row has a '删除线程' (delete thread) button. At the bottom, it shows '显示第 1 到第 0 条记录, 总共 0 条记录'.

序号	线程	状态	采集任务名称	任务ID	任务状态	采集速度	采集进度	对任务的操作	操作
1	Hertrix	空闲	-	-	-	-	-	-	删除线程
2	_Hertrix1	采集中	35e2zgjefh1-d-031	000020180912093526ab003	采集中	2	97.17%	操作	删除线程
3	_Hertrix2	空闲	-	-	-	-	-	-	删除线程
4	_Hertrix3	采集中	35e2zgjefh1-d-038	000020180914093324ab002	暂停	0	61.50%	操作	删除线程
5	_Hertrix4	空闲	-	-	-	-	-	-	删除线程
6	_Hertrix5	空闲	-	-	-	-	-	-	删除线程
7	_Hertrix6	空闲	-	-	-	-	-	-	删除线程
8	_Hertrix7	空闲	-	-	-	-	-	-	删除线程
9	_Hertrix8	空闲	-	-	-	-	-	-	删除线程



第三部分：系统操作

维护备份模块——备份

The screenshot displays the 'Network Resource Preservation and Service System' (网络资源保存与服务系统) interface. The top navigation bar includes the system name, a timestamp of '2018/7/6 下午2:16:10', and notification icons for '11296', '51', and '0'. A user profile for 'admin' is visible in the top right.

The left sidebar contains a menu with the following items: 首页 (Home), 采集模块 (Collection Module), 编目模块 (Cataloging Module), 回放模块 (Playback Module), 内容管理和发布模块 (Content Management and Release Module), 数据保存模块 (Data Preservation Module), 统计模块 (Statistics Module), 用户管理模块 (User Management Module), and 维护备份模块 (Maintenance Backup Module). The '维护备份模块' is expanded to show '服务器监控' (Server Monitoring) and '备份' (Backup), with '备份' selected.

The main content area features a search bar with '操作人' (Operator) and '时间' (Time) filters, and a '搜索' (Search) button. Below the search bar are links for '数据库导入' (Database Import) and '数据库导出' (Database Export). A table with the following columns is shown: 序号 (Serial Number), 时间 (Time), 导入/导出 (Import/Export), and 操作人 (Operator). The table currently displays the message '没有找到匹配的记录' (No matching records found).



部署申请方式



联系电话：010-88545587

联系邮箱：szzyzh@nlc.cn



× 推广工程网络资源保存与服务平台 ...



请填写以下问卷并点击提交，如有疑问请联系国家图书馆

联系电话：010-88545587
联系邮箱：szzyzh@nlc.cn

Q1. 申请平台部署的单位名称

Q2. 申请时间

Q3. 申请人姓名

Q4. 申请人联系方式（请填写移动手机和座机电话）

Q5. 申请人邮箱

提交





谢谢!

