

# 数字资源唯一标识符应用系统完善建设 开放试验方案

数字资源唯一标识符应用系统完善建设（简称唯标二期）是在一期数字图书馆应用系统中有效使用唯一标识符的基础上，依托国家科技文献平台，探讨跨越数字科研、数字教育、数字出版和数字媒体等领域利用唯一标识符系统实现互操作机制，支持跨领域的数字对象发现、确认、管理与调用，支持数字信息集成服务。本次开放试验的目的是，通过选取国内 4 家有代表性的中文学位论文服务系统，利用唯一标识符应用系统对异构数据库实行整合，试验中文学位论文服务系统的统一命名和全文提供机制，探索并完善跨系统的唯一标识符管理机制。

## 一、试验目标

在保持现有各学位论文服务系统基本架构和服务功能的基础上，尝试利用唯一标识符系统将多个异构学位论文服务系统整合为一个“虚拟中文学位论文服务系统”。利用该系统提供涵盖 CALIS、国家图书馆、中国科学院国家科学图书馆、中国科学技术信息研究所等国内四家主要信息机构的博、硕士学位论文的查询及全文提供服务。

具体试验目标为：

1. 在论文查重的基础上，利用唯一标识符系统实现各系统中中文学位论文的唯一标识、解析和全文定位。
2. 通过唯一标识符，从参与试验的任意一个学位论文服务系统实现查询所有试验系统中学位论文文摘的能力，并可提供全文获取服务。
3. 验证唯标一期开放的唯一标识符解析系统，试验中心服务器

GHR 和本地服务器 LHS，在现实数字图书馆跨系统信息服务系统中的可行性。

4. 为了进一步拓展试验目标，希望将通过唯一标识符定位学位论文全文的功能，整合到国家科学图书馆新开发的桌面信息工具“E划通”之中。

## 二、试验框架

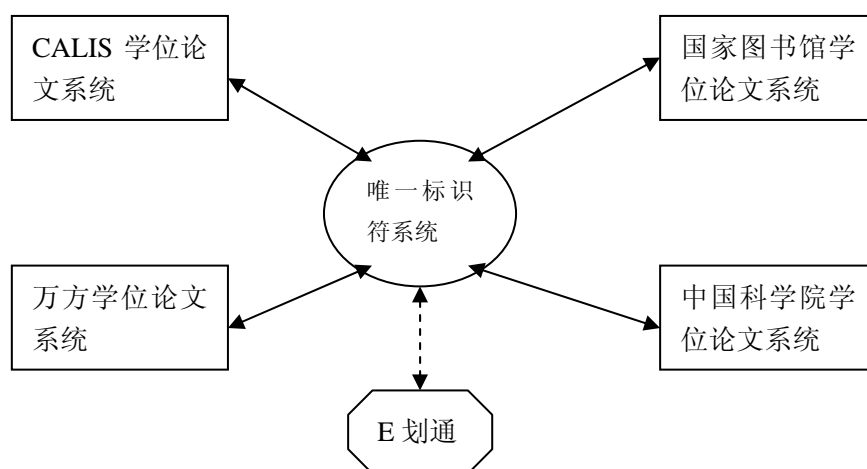


图 1: 唯标二期开放试验结构示意图

## 三、试验设计

1. 实验数据准备。选择国图、CALIS、万方公司和中科院四家学位论文数据库中 2004-2005 年所有硕士和博士论文的元数据作为实验对象。由各家分别负责将数据转换为项目组指定的中间格式（以逗号分隔的.txt 或.csv 格式），由项目组统一对数据进行查重和分配唯一标识符，并返还给各家。

试验中间格式定义如下：

本地记录控制号（RecordID）

中文题名（TitleCN）

英文题名（TitleEN）

作者中文名 ( AuthorCN )

作者拼音 ( AuthorPY )

学位授予年 ( Year ): 格式 ( YYYY )

导师姓名 ( TutorName )

学位授予单位 ( Organization )

学位级别 ( Level ): 博士或硕士

学位授予时间: YYYY-MM-DD

页码 ( Pages ) (注: 页码可以是整型或不包含其它字母、标点, 可转换为整形的字符型)

专业 ( Major ) (注: 按原系统取值, 不必作规范)

论文中文文摘 ( AbstractCN )

论文中文关键词 ( KeywordCN ) 多个关键词用%分隔

论文英文文摘 ( AbstractEN )

论文英文关键词 ( KeywordEN ) 多个关键词用%分隔

论文题录信息链接 ( AbstractLink ) (必备字段)

论文前 16 页或 24 页链接 ( PartTextLink )

论文全文链接 ( FullTextLink )

唯一标识符 ( URI ): 唯一标识符, 由 GHR 统一分配

2. 实验数据导入。项目组统一将分配了唯一标识符的学位论文元数据导入 GHR, LHS 部署单位将数据导入 LHS;

3. 实验环境搭建。上述 4 家单位在本地搭建以项目组返回的实验数据为基础的学位论文原型系统, 通过改造后(或直接改造原系统)为学位论文记录增加唯一标识符链接。

4. 实验工作实施。项目组模拟用户从唯一标识符途径对 4 家单位的学位论文数据库发出全文请求, 进行请求解析和定位等实验工作。

#### 四、试验进度和分工

2007 年 2 月-4 月, 确定开放试验方案, 各单位完成数据准备工

作包括：数据导出（各单位），数据查重（国家科学图书馆），数据分配唯一标识符（国家科学图书馆）。

2007年4月-6月，各单位完成本单位的学位论文系统的数据导入和系统改造工作，增加唯一标识符链接（各单位）。试验多个单位唯一标识符分配和管理机制（CALIS和国家科学图书馆）。

2007年6月-8月，各单位模拟用户通过唯一标识符获取学位论文全文数据，进行请求解析和定位等试验（各单位）。

如有可能，将唯一标识符系统与“E划通”链接试验（国家科学图书馆），其他项目参建单位也可进行相应的试验。

2007年9月，开放试验初步完成，形成试验报告。

## 五、学位论文数据库参考网址

### 1. 国家科学图书馆学位论文数据库

<http://dpaper.csdli.ac.cn/xwlv/index.jsp>

前台 JSP 编程，后台使用 TRS 数据库和 SQL server 数据库

### 2. CALIS 学位论文数据库

<http://etd.calis.edu.cn/>

### 3. 国家图书馆学位论文数据库

<http://res1.nlc.gov.cn:9080/doctor/index.htm>

前台 JSP 编程

### 4. 万方学位论文数据库

站点 1: [http://www.nstl.gov.cn/htm/wxjs/ywjsdg.jsp?db\\_type=ENPS5](http://www.nstl.gov.cn/htm/wxjs/ywjsdg.jsp?db_type=ENPS5)

前台利用 JSP 编程，后台为 TRIP 数据库

站点 2: <http://www.wanfangdata.com.cn/>

前台利用 Asp 编程，后台为 iBASE 数据库

5. 国家科学图书馆“E划通”

<http://desktool.csdl.ac.cn/>

本试验方案上报数字图书馆标准规范总项目组，并在总项目组的指导下实施。

数字资源唯一标识符应用系统完善建设项目组

2007年4月8日