

清华大学图书馆
期刊数据支撑学位申报服务的应用案例

报告人：武小楠

2015-11



提纲

- 问题的提出
- 清华图书馆的数据基础
- 期刊规范数据池的建设
- 学位申报服务接口的建设
- 研究生院学位申报系统应用事例
- 研究生院学位申报系统应用效果
- 结语



提纲

➤ 问题的提出

- 清华图书馆的数据基础
- 期刊规范数据池的建设
- 学位申报服务接口的建设
- 研究生院学位申报系统应用事例
- 研究生院学位申报系统应用效果
- 结语



问题的提出

国务院学位办要求研究生在学位申报时需提交在读期间发表的论文，同时应标注收录引用情况。

研究生个人提交的数据准确性极差，主要表现在：期刊名称不准确，收录引用信息混乱；合作发表文章时多作者提交数据不一致；缺乏必要的元数据信息。因此，由学生自提交的数据是不足以支撑分类统计和深度分析工作进行的。

清华大学研究生院希望图书馆在数据规范方面提供帮助。



问题的提出

人工自存缴或者自保存的论文列表，因元数据不完整和描述失范而无法支撑对论文的科学管理和深度分析。

失范数据

频繁错误

Title: Arbitrary phase rotation of the marked state cannot be used for Grover's quantum search algorithm.

频繁遗漏

Authors: Long, G1 ; Zhang, W1 ; Li, Ys ; Niu, L

全部失范

Journal: Communications In Theoretical Physics

基本正确

Year: 1999

基本正确

Vol: 32

频繁遗漏

Issue: 3

频繁遗漏

StartPage: 335

频繁遗漏

EndPage: 338

丢失

ISSN: 0253-6102



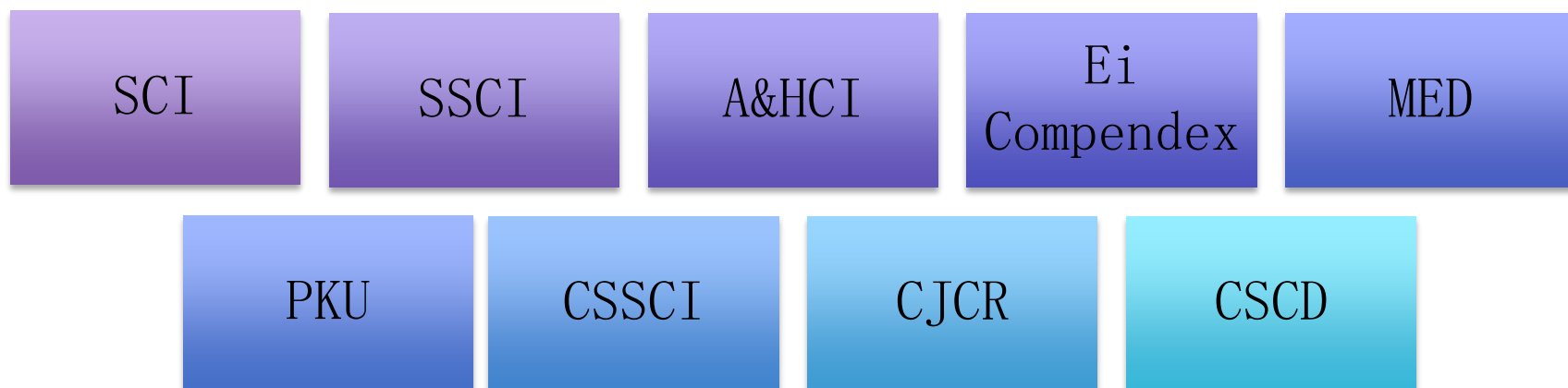
提纲

- 问题的提出
- 清华图书馆的数据基础
- 期刊规范数据池的建设
- 学位申报服务接口的建设
- 研究生院学位申报系统应用事例
- 研究生院学位申报系统应用效果
- 结语



清华图书馆的数据基础

学术期刊是高校图书馆文献资源的主要组成部分，具有重要的信息参考价值。而期刊信息，特别是期刊收录情况，能够反映出该种刊物，乃至其收录论文的学术水平。自2006年起，清华大学图书馆依托电子期刊管理系统在期刊基础数据规范化管理、期刊收录引用数据的收集、期刊出版数据的融汇方面积累了大量的经验。



提纲

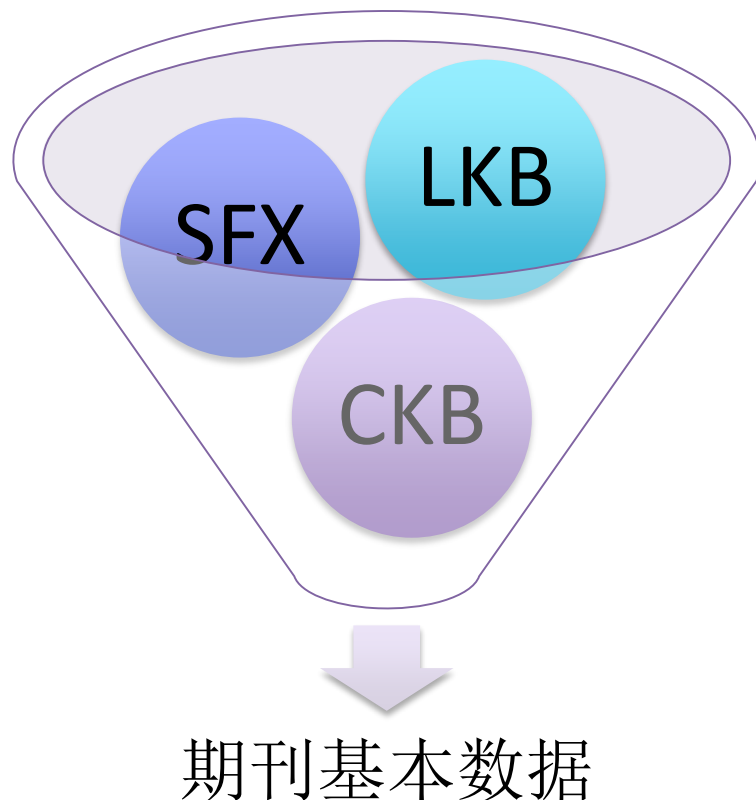
- 问题的提出
- 清华图书馆的数据基础
- 期刊规范数据池的建设
- 学位申报服务接口的建设
- 研究生院学位申报系统应用事例
- 研究生院学位申报系统应用效果
- 结语



期刊规范数据资源池的建设

期刊基本数据整理

清华大学图书馆借助现有平台(SFX)和数据(SFX LKB CKB)优势,定期收集全球近10万种期刊列表及基本信息,包括:期刊正题名、缩略题名、ISSN、出版周期、出版机构、历史沿革等内容。其中多个期刊题名与ISSN的对应关系直接作用于未来数据的分析和关联处理。期刊基本数据采用MySQL管理,支持XML数据导出,本项目根据实际需要解析XML文件获取期刊基本信息数据。



期刊规范数据资源池的建设

期刊收录引用数据整理

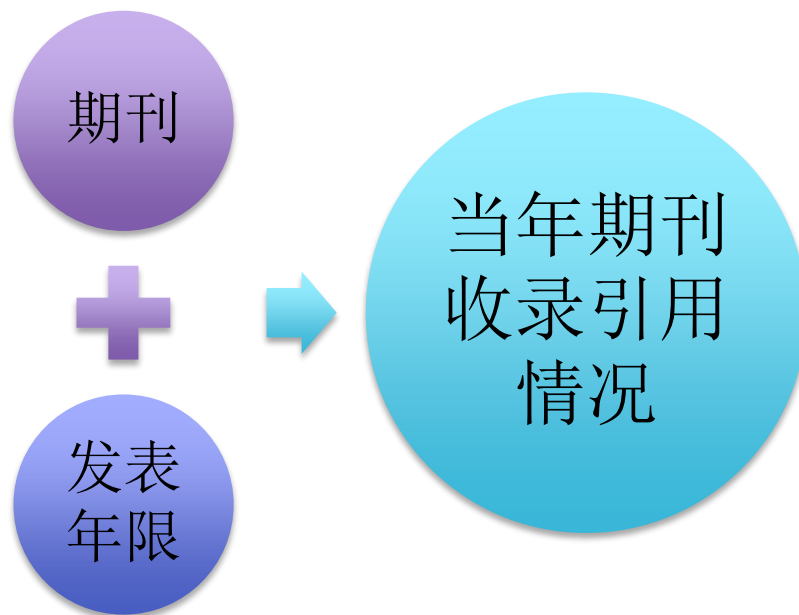
期刊评价体系多种多样，清华图书馆采用现行的期刊收录工具收集整理多评价体系中期刊收录情况，与期刊信息在数据层面做深度整合。这些期刊评价信息源主要有：SCI（科学引文索引）、SSCI（社会科学引文索引）、A&HCI（艺术与人文科学引文索引）、Ei Compendex（工程索引）、MED（MEDLINE数据库）、PKU（中文核心期刊要目总览）、CSSCI（中文社会科学引文索引）、CJCR（中国科技期刊引证报告）、CSCD（中国科学引文数据库）等，上述收录引用工具在学科领域上各有侧重，更新时间不尽相同，本项目整理多评价体系的期刊收录列表、更新周期。如SCI每年更新，CSCD、CSSCI2年更新、PKU自2011年起3年更新一次。本案例，清华图书馆提供自2011年以来的所有收录引用指标。



期刊规范数据资源池的建设

期刊基本数据与期刊收录引用数据的深度融合

针对一本期刊，可根据研究生发表文献的年限自动关联该期刊的收录引用情况。



提纲

- 问题的提出
- 清华图书馆的数据基础
- 期刊规范数据池的建设
- 学位申报服务接口的建设
- 研究生院学位申报系统应用事例
- 研究生院学位申报系统应用效果
- 结语



学位申报服务接口的建设

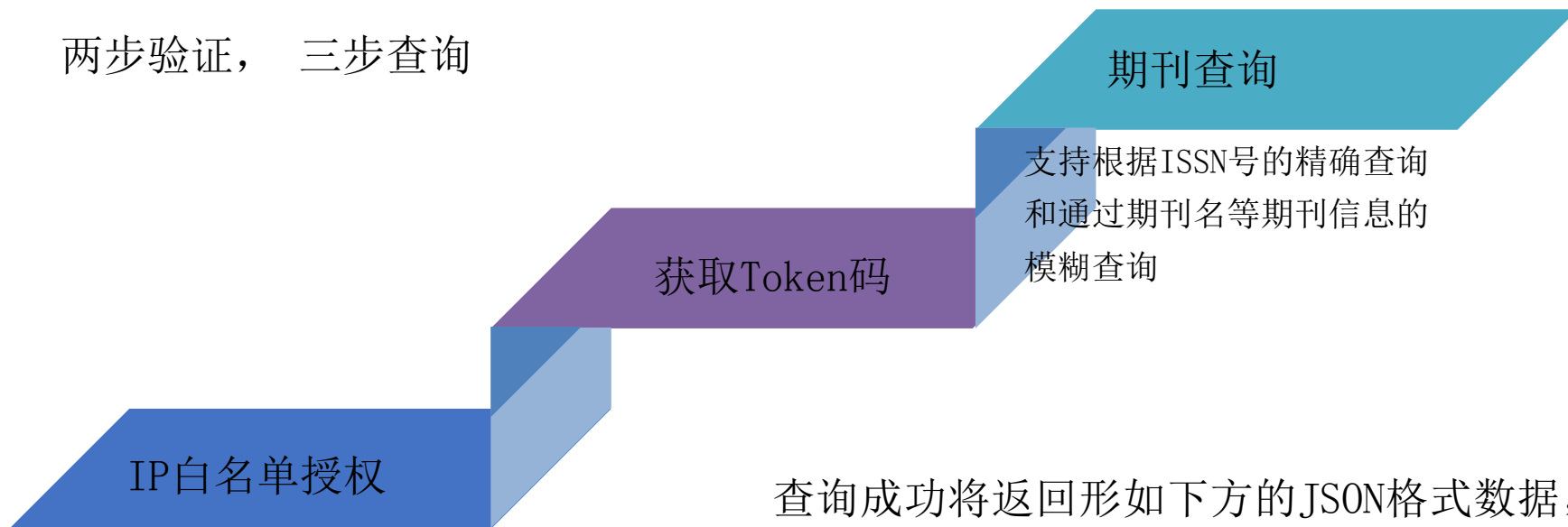
清华大学图书馆基于Open API服务理念，着手为学位申报系统开发期刊应用接口建设。考虑到研究生学位申报系统的重要性的稳定性，图书馆在接口建设过程考虑了如下几个核心要素，包括：**接口授权、接口稳定性、接口数据请求及返回规范**等内容。

接口开发本着“两步验证，三步发起查询请求”的原则，具体流程如下：首先，验证环节包括IP白名单授权和获得Token。期刊服务接口程序在后台设置了严格的白名单限制，只有通过授权，才能进行接口调用。而接口调用之初，需要发出形如：`http://服务IP/getToken`的HTTP请求，获得Token码。然后，才能进行实际的期刊查询请求。本接口支持通过ISSN号的精确查询和通过期刊名等期刊信息的模糊查询。



学位申报服务接口的建设

两步验证， 三步查询



支持根据ISSN号的精确查询
和通过期刊名等期刊信息的
模糊查询

查询成功将返回形如下方的JSON格式数据，
便于第三方系统调用。

```
{"result":true,"data": [[{"jtitle":"AAPG bulletin","ISSN":"0149-1423","EISSN":"1558-9153","Source":"Available from 2000 volume: 84 issue: 1. ","alternative":["AAPG BULL AM ASSOC PETR GEOL","AAPG BULL","AAPG BULL AMER ASSN PETROL G","AAPG BULLETIN AMERICAN ASSOCIATION OF PETROLEUM GEOLOGISTS","AM ASSOC PET GEOL BULL" ]}] ]}
```



提纲

- 问题的提出
- 清华图书馆的数据基础
- 期刊规范数据池的建设
- 学位申报服务接口的建设
- 研究生院学位申报系统应用事例
- 研究生院学位申报系统应用效果
- 结语



研究生院学位申报系统应用示例

清华大学
Tsinghua University

学位申报系统

学位申报

- 填报提示
- 硕士学位论文
- 发表论文 (3)
- 出版专著 (0)
- 科研成果 (4)
- 就业单位
- 对外交流
- 硕士学位论文电子版
- 毕业信息核对
- 最终确认

• 论文收录、检索、查询

发表论文

论文信息	
* 论文题目	<input type="text"/> (不超过200字符)
* 第几作者	<input type="text" value="1"/> 请填写阿拉伯数字
* 导师是否第一作者	否
* 论文发表情况	已发表
* 发表类别	国内期刊
* 发表年月	<input type="text" value="201306"/> (日期格式: YYYYMM)
刊物信息	
ISSN	<input type="text"/>
* 刊物名称	<input type="text"/> (不超过72字符) natu Nature Bay nature Our Nature Actanature Natuur. blad Natura 2000 Natural life Human nature Nature Methods naturalSCIENCE
* 刊物级别	核心期刊
* 刊物卷期号	<input type="text"/>
检索信息	
* 是否EI检索	否
* 是否SCI检索	否
* 是否SSCI检索	否
* 是否AMCI检索	否
* 是否CSSCI检索	否
* 是否CSCD检索	否

在学位申报系统中嵌入图书馆期刊服务接口之后，学生填报发表论文的期刊名称时，系统会调用接口，通过自动完成（Autocomplete）功能，根据输入的信息，实时查询期刊信息，予以搜索热点提示。学生只需选择确认的期刊名称，这样保证了入库信息的准确性。



研究生院学位申报系统应用示例

清华大学
Tsinghua University

学位申报系统

发表论文

学位申报

- 填报提示
- 硕士学位论文
- 发表论文 (3)
- 出版专著 (0)
- 科研成果 (4)
- 就业单位
- 对外交流
- 硕士学位论文电子版
- 毕业信息核对
- 最终确认

论文收录、检索、查询

论文信息

- 论文题目: (不超过200字符)
- 第几作者: 1 (请填写阿拉伯数字)
- 导师是否第一作者: 否
- 论文发表情况: 已发表
- 发表类别: 国内期刊
- 发表年月: 201306 (日期格式: YYYYMM)

刊物信息

- ISSN: 0028-0836
- 刊物名称: Nature (不超过72字符)
- 刊物级别: 核心刊物
- 刊物卷期号:

检索信息

- 是否EI检索: 否
- 是否SCI检索: 是
- 是否SSCI检索: 否
- 是否AMCI检索: 否
- 是否CSCSI检索: 否
- 是否CSCD检索: 否

保存

清华大学计算机与信息管理中心
Copyright 2008 All rights reserved.

当学生填入论文发表年月，以及期刊ISSN或者期刊名称后，系统将再次进行服务接口的调用，在如图所示的检索信息中，同步显示出该种期刊当年是否被各大收录体系所收录。

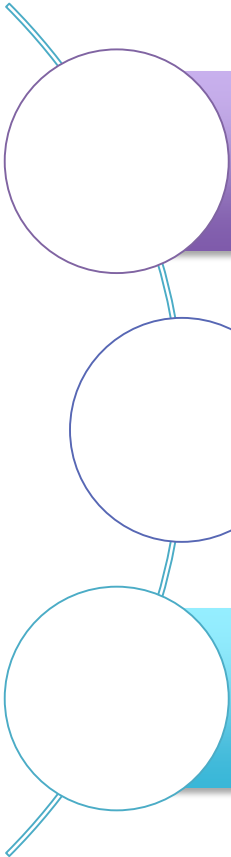


提纲

- 问题的提出
- 清华图书馆的数据基础
- 期刊规范数据池的建设
- 学位申报服务接口的建设
- 研究生院学位申报系统应用事例
- 研究生院学位申报系统应用效果
- 结语



研究生院学位申报系统应用效果



2015年5月在研究生院学位申报工作中首次使用，使用体验平滑流畅。

近5000名硕士生和博士生进行了论文情况申报。

图书馆因此得到研究生院不菲的经费支持。



提纲

- 问题的提出
- 清华图书馆的数据基础
- 期刊规范数据池的建设
- 学位申报服务接口的建设
- 研究生院学位申报系统应用事例
- 研究生院学位申报系统应用效果
- 结语



结语

图书馆利用大量学术资源数据，对期刊基本信息和收录信息进行追踪、梳理创建期刊规范数据资源池。期刊规范数据资源池的建设可支撑各种与期刊相关的上层应用，图书馆也可快速方便地融入学校的教学科研管理流程，满足学校对学术产出规范化管理和深层分析的需求。该案例的成功应用也为图书馆利用数据优势支撑学校的教学科研及人才培养工作奠定了扎实的基础。

开发环境（编程环境，开源软件的利用）：

Python、Apache、MySQL

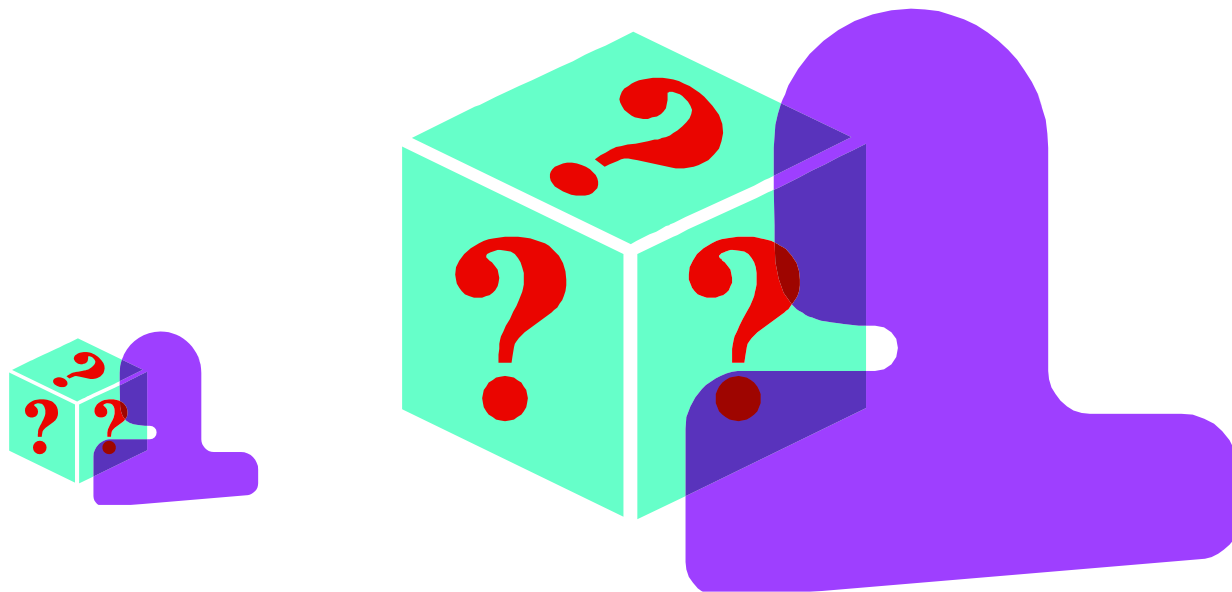
应用环境（运行环境，应用范围）：

WEB

开发人员：

张蓓、窦天芳、陈信欢、钟文锋、武小楠





谢谢各位

