



融入学术成果生态系统 —机构知识库的应用 扩展与开发实践

北京大学图书馆
信息化与数据加工中心

目录

Contents

- PKU IR 3.0
- 科研信息管理系统
- 国家自然科学基金委开放知识库
- 地球与空间学院 & 国家地质调查局
- 总结

01

Part One

PKU IR 3.0

1-1 What's New ?



新的数据



新的机遇

新的设计



新的思想



1-2 门面最重要



热门
POPULAR

全部馆藏 教师学者 分类导航 使用指南 中文 | English 登录

理学部	3D Haar-Like Features for Pedestrian Detection 2007 IEEE International Conference, 1992 Xinyi Cui 6703 阅	中国科学院机构知识库增值服务体系建设与发展 人民黄河, 1992 祝步明, 刘巍 6366 阅
信息与工程学部	对黄河三角洲开发的认识 人民黄河, 1992 胡兆量 6703 阅	CALIS三期机构知识库建设及推广项目平台建设总结与演示 人民黄河, 1992 胡兆量 6366 阅
人文学部	大城市高层知识分子迁移特征——北京教授籍贯地分析 人民黄河, 1992 胡兆量 6326 阅	地理学的不倦探索者——悼念陈传康教授 人民黄河, 1992 蔡运龙 6366 阅
社会科学部	3D Haar-Like Features for Pedestrian Detection 2007 IEEE International Conference, 1992 Xinyi Cui 6703 阅	中国科学院机构知识库增值服务体系建设与发展 人民黄河, 1992 祝步明, 刘巍 6366 阅
医学部	台湾加工出口区规划与建设经验 人民黄河, 1992 胡兆量 6326 阅	《北京通史》学术顾问、特约编审发言(摘要) 人民黄河, 1992 蔡运龙 6366 阅
跨学科类	边疆优势论与边境口岸建设 城市问题, 1993 胡兆量 5669 阅	《北京通史》学术顾问、特约编审发言(摘要) 人民黄河, 1992 蔡运龙 6366 阅

面向学者、突出学者

- 学者是学术成果的创造者
- 学者关心个人学术产出的关注度
- 领导在乎机构学术产出的影响力
- 用户喜欢追求热门
- 图书馆希望系统更有生命力

1-3 数据是王道(1)

去重后元数据量为**412,796**篇

去重前是**784,326**条

全文**294,092**篇



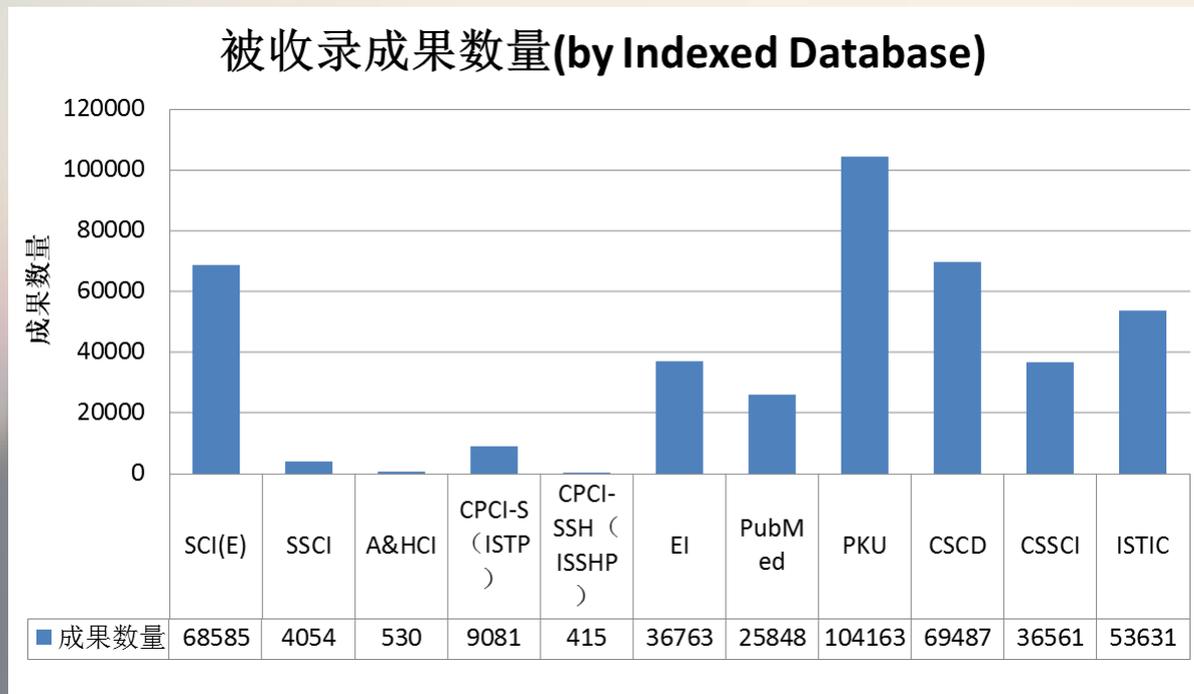
感谢刘丹同学在内容建设方面细致的工作

1-3 数据是王道(2)

元数据主要包括以下19个属性：

- 成果类型
- 语种
- 标题
- 作者
- 英文作者
- 作者排序
- 作者机构
- 刊名/会议名称/出版社
- 年卷期/会议时间/出版日期
- 页码
- 关键词
- 英文关键词
- 摘要
- 英文摘要
- 受资助情况
- 来源数据库
- 全文链接
- 收录索引情况：SCI(E)、SSCI、A&HCI、CPCI-S (ISTP)、CPCI-SSH (ISSHP)、EI、PubMed、中文核心期刊要目总览 (PKU)、中国科学引文数据库 (CSCD)、中文社会科学引文索引 (CSSCI)、中国科技论文统计源期刊 (中国科技核心期刊)
- DOI：44%的文章有DOI
- 被引情况：被引次数，引证文献

1-3 数据是王道(3)



感谢刘丹同学在内容建设方面细致的工作

02

Part Two

科研信息管理系统

2-1 PKU CRIS

实现科研信息统计
真实化、自动化

实现科研管理全流
程自动化和信息化

辅助学校发展决策
科学化



2-2 图书馆做什么？



2-3 PKU IR ↔ CRIS



调研与借鉴香港大学

DSpace-cris



IR学术成果提交嵌入项目管理过程



保持一致性，无缝衔接，
成为CRIS的有机组成部分



学校出台一定的强制与奖励政策

03

Part Three

国家自然科学基金委开放知识库

3-1 项目由来



政策声明

2014年5月19日，国家自然科学基金委员会发布了《关于受资助项目科研论文实行开放获取的政策声明》。



全球研究理事会

全球研究理事会（Global Research Council, GRC）2014年全体大会于2014年5月26日至28日在北京举行。第四届全球研究理事会会议（GRC）于2015年5月26-28日在日本东京召开。



成果共享网

国家自然科学基金委现有的成果共享网主要是针对项目，缺乏成果全文，网站结构不太适合开放获取的需求



机缘巧合

双方前期探讨；项目急迫，不到一个学期；北大图书馆在平台、内容、政策方面的建设有一定的基础。

3-2 团队成员

聂华



Project
Leader

崔海媛



Team
Leader

韦成府



System
Develop

刘丹



Content
Develop

朱玲



Policy
Research

罗鹏程



System
Develop

吴越



UI & Policy
Research

张乃帅



Operation
Maintenance

周义刚



System
Test

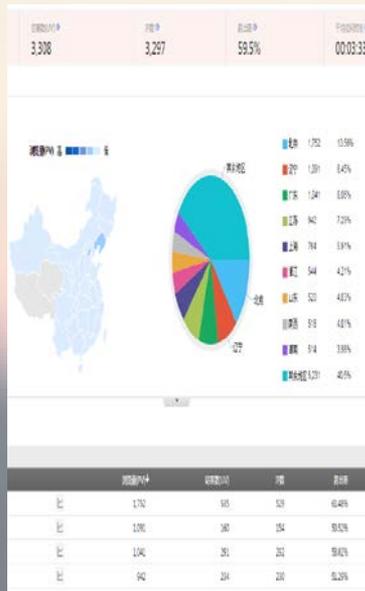
3-3 NSFC OR



首页



政策



访问



内容

3-4 内容建设

原始数据

309,303条

- 项目一一对应数据：269,463条。
- 全文PDF一一对应：214,174条。

CALIS期刊论文库中对比，得到142,643条。

- 去重，得到135688条。
- 入库135100条。

135688条数据的所有字段全部清洗，按字段的重要性程度分别进行不同程度的清洗：

- 标题、期刊名、ISSN、年、作者列表、语种：对这6个字段进行了数据完整性、准确性的检查，保证这6个字段的数据是完整的，而且是准确的；
- 卷、期、页码：保证数据项尽量完整；
- 关键词、摘要、作者单位：保证数据格式的正确性；

入库135100条

数据按标准格式打包，导入到NSFC-OR系统中。

04

Part Four

地球与空间学院 & 地质调查局

4-1 项目背景



全国地质资料馆作为各高校、科研机构研究人员、生产人员的资源提供者，更需要紧跟时代的发展，不断加快科研资源的数字化和信息化进程；在进行资源数字化的同时，更好地为科研人员、生产人员提供信息服务

建设学科知识库成为解决上述问题的
首选方法

4-2 项目需求

1

用户、用户
组与授权、
认证

2

workflow管理

3

关联数据发布

4

检索和浏览

5

知识关联

4-3 一种双赢



国土资源部“十二五”重点项目：国家地质数据库一体化管理系统建设



来自教师科研的需求，对于图书馆来说，责无旁贷



无形中，IR建设已经被深入宣传，而且也将得到进一步推广

05

Part Five

总结

5-1 学术成果生态化



学者为主导

嵌入科研

提供研究支持

与出版方、数据库商合作

5-2 随想



应用扩展

NSFC OR——机构与学科相结合的知识库

NGA IR——学科机构知识库

数字昆曲——特色资源知识库

CRIS——学术成果管理系统

.....

开发实践

增值服务——ORCID、Altmetrics、学术轨迹、数据挖掘

外部合作——WOS、Google Scholar、百度学术、CSSCI、CSCD、

CNKI、万方、.....





Thank You

北京大学图书馆