



清华大学图书馆
Tsinghua University Library

全网域资源的发现与服务

Web-scale Discovery Services at Tsinghua University

邓景康 DENG Jingkang

清华大学图书馆 Tsinghua University Library

第四届中美高校图书馆合作发展论坛 多伦多 2017.03

- **资源发现服务现状**
The status of discovery service
信息资源环境
- **从资源发现到资源获取**
From discovery to delivery
电子资源管理 电子资源揭示
- **从资源发现到知识发现**
From resource discovery to knowledge discovery
发现服务新动向



- 资源发现服务现状
The status of discovery service
信息资源环境

- 电子资源的蓬勃发展给传统图书馆服务带来冲击和挑战
- 全球学术图书馆在电子资源上的投入逐年增加
- 在线访问是获取信息资源的主要手段

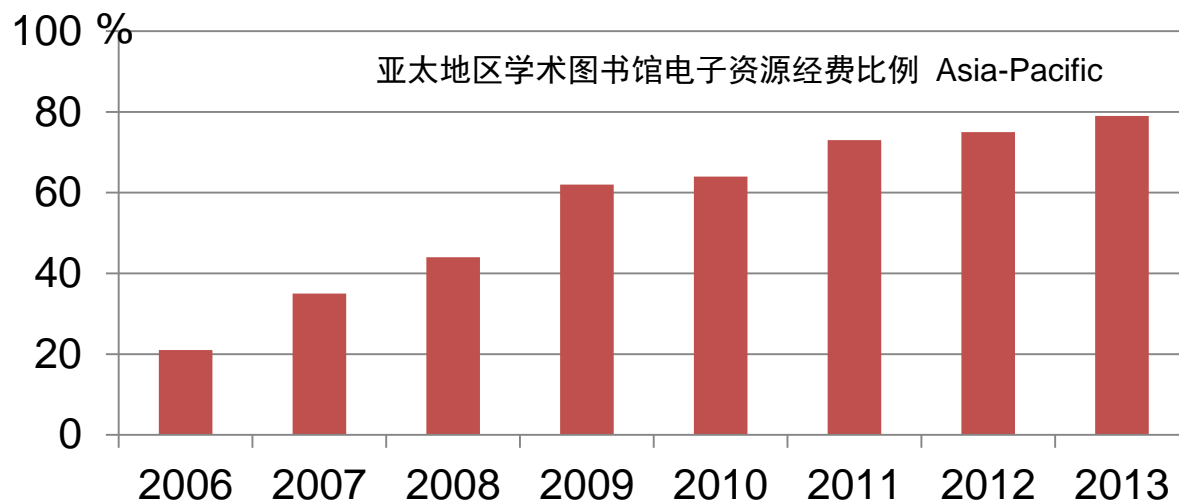
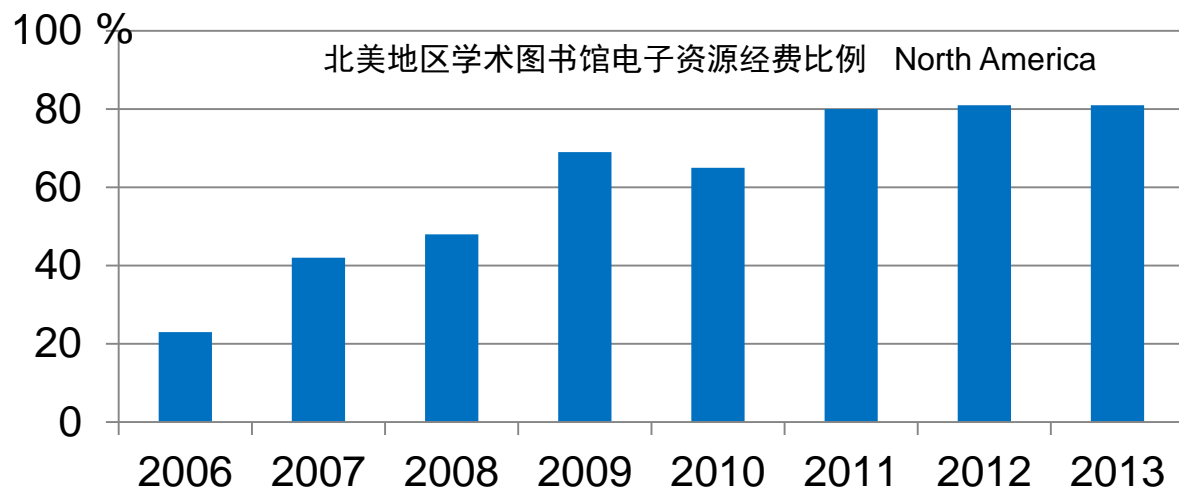
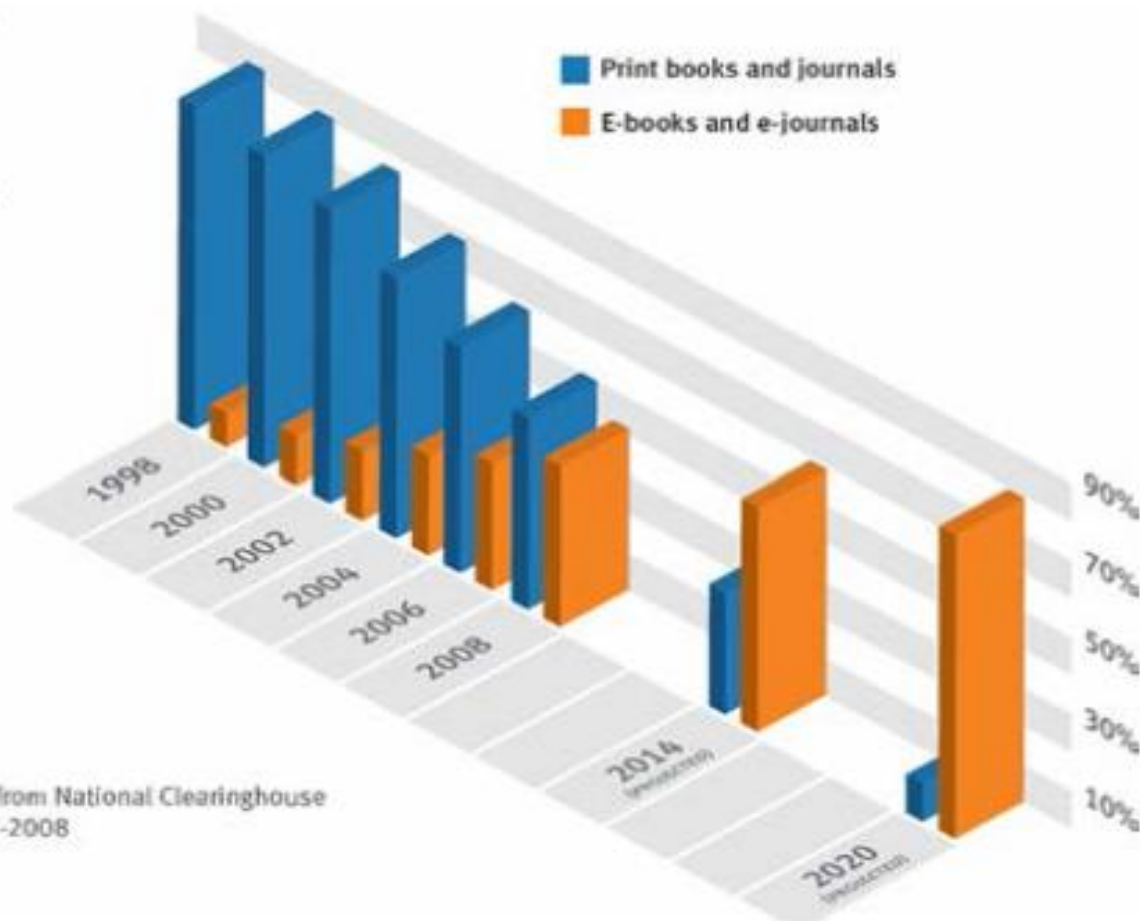


Figure 1. *Academic libraries expenditures since 1998*

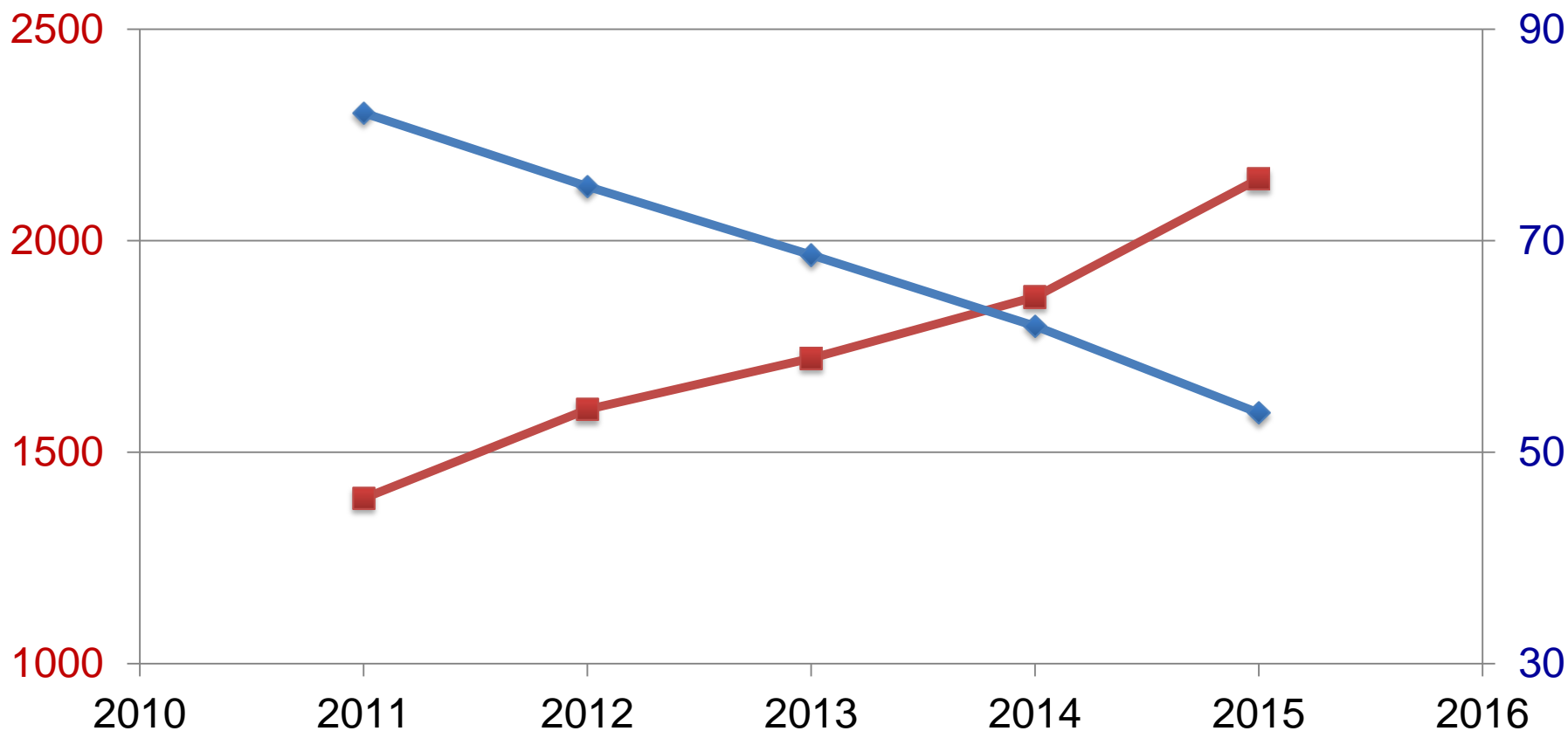
Academic Library
Expenditures on
Purchased and
Licensed Content



OCLC Research, 2010. Derived from National Clearinghouse on Educational Statistics, 1998-2008

Source: OCLC (2010)

电子资源全文下载量以年均15%的速度增长
纸质图书借阅量以年均8.1%的速度下降



实体馆藏

- 440 万种，503.6万册/件
- 形成多学科、多类型、多载体的综合性馆藏体系

电子资源

- 数据库 586 个
- 全文电子期刊 8.9 万种（含OA刊2.3万种）
- 电子图书 855.3 万册
- 电子版学位论文 392.5 万篇

信息环境

● 传统意义上的馆藏

- ✓ 多类型
- ✓ 多来源

● 数字化资源

- ✓ 分布广泛
- ✓ 内容丰富
- ✓ 数量巨大
- ✓ 获取灵活

Google

搜索引擎
培养的用户



- ✓ 有用的 (useful)
- ✓ 快速的 (fast)
- ✓ 简单的 (simple)
- ✓ 有吸引力的 (engaging)
- ✓ 创新的 (innovative)
- ✓ 适合大众的 (universal)
- ✓ 有用的 (profitable)
- ✓ 漂亮的 (beautiful)
- ✓ 值得信赖的 (trustworthy)
- ✓ 个性化的 (personable)

- ✓ 清华收集OA期刊2.3万种，占全文电子期刊总量的1/3
- ✓ Springer有562种OA期刊，占Springer期刊总量的18.8%
- ✓ Elsevier有458种OA期刊，占Elsevier期刊总量的10.5%
- ✓ JCR（2014）有1614种OA期刊，占JCR期刊数量的18.6%

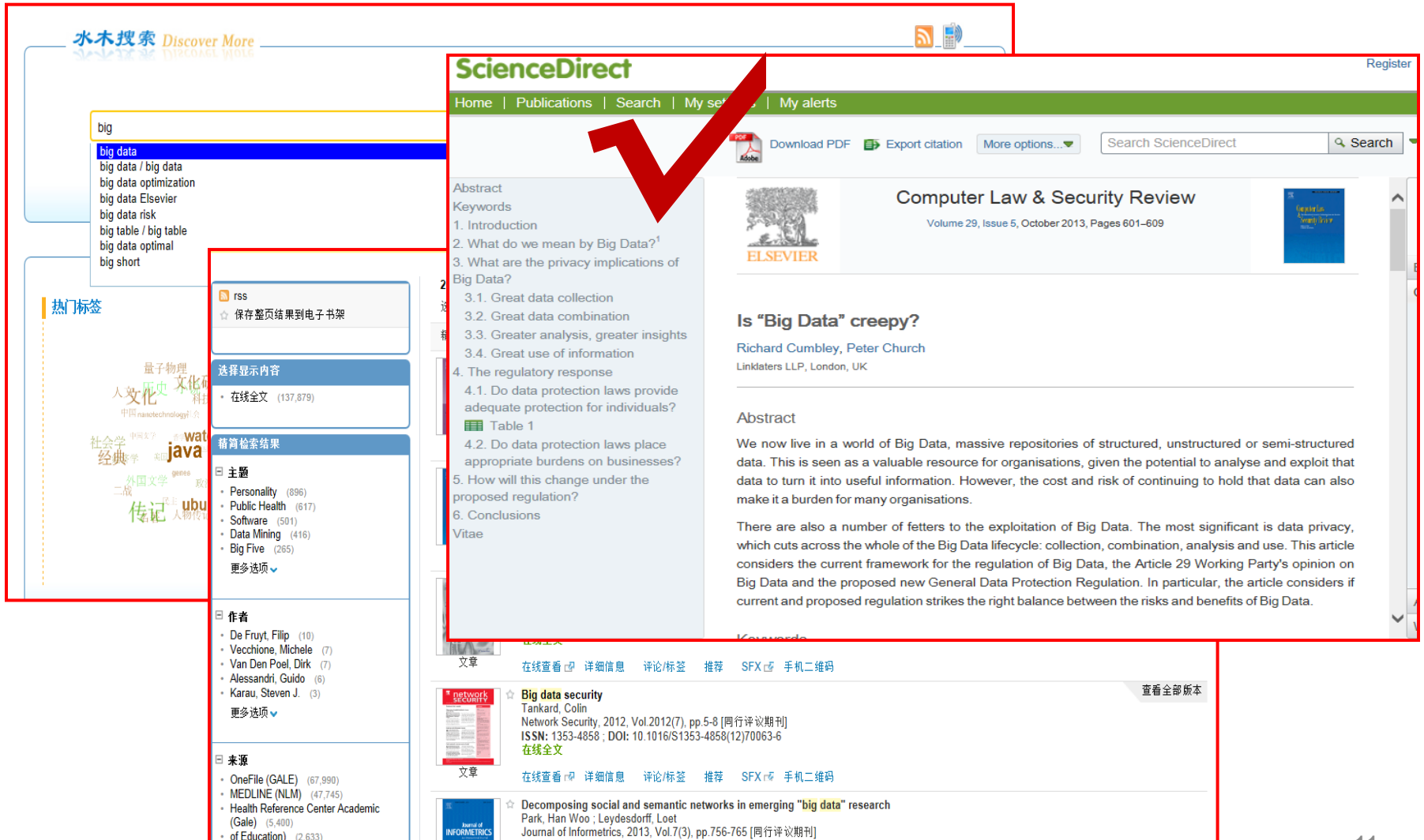
HKIR, Plos, DOAJ, Library of Congress的元数据都免费开放，图书馆可追踪这些高品质的开放资源，引导读者不要错过学术价值较高的OA资源。

从出版商、大学、公开的网站收集学术文献信息，形成中心知识库，通过预索引的方式，为用户提供快速、简单、易用的资源发现及获取服务。

标志性事件：

2009年1月的ALA(American Library Association)年会上，Series Solution公司宣布即将发布全球第一个网络级资源发现系统Summon。

资源发现系统诞生



The screenshot displays a search interface with the following components:

- Water Search (水木搜索):** Includes the text "Discover More" and a search bar.
- Search Results:** A dropdown menu for the search term "big" lists suggestions such as "big data", "big data / big data", "big data optimization", "big data Elsevier", "big data risk", "big table / big table", "big data optimal", and "big short".
- Hot Tags (热门标签):** A word cloud featuring terms like "量子物理", "历史文化", "社会", "经典", "java", "ub", "传记", "人物", "ub", "人物", "传记", "ub", "人物", "传记".
- Navigation and Actions:** Includes an RSS icon, a button to "保存整页结果到电子书架", and a "选择显示内容" section with a "在线全文 (137,879)" link.
- Refined Search Results (精简检索结果):** Lists subjects (主题) such as "Personality (806)", "Public Health (617)", "Software (501)", "Data Mining (416)", and "Big Five (265)".
- Authors (作者):** Lists authors like "De Fruyt, Filip (10)", "Vecchione, Michele (7)", "Van Den Poel, Dirk (7)", "Alessandri, Guido (6)", and "Karau, Steven J. (3)".
- Source (来源):** Lists sources like "OneFile (GALE) (67,990)", "MEDLINE (NLM) (47,745)", "Health Reference Center Academic (Gale) (5,400)", and "of Education (2,633)".
- ScienceDirect Article:** A large red checkmark is placed over the article details for "Computer Law & Security Review". The article title is "Is 'Big Data' creepy?". Authors are "Richard Cumbley, Peter Church". The abstract discusses the challenges of Big Data, particularly data privacy, and the need for regulation.
- Related Articles:** Below the main article, two other articles are listed: "Big data security" by Tankard, Colin and "Decomposing social and semantic networks in emerging 'big data' research" by Park, Han Woo.

国内资源发现系统应用现状

表1 “985” 高校资源发现系统使用情况

序号	名称	资源发现名称	系统类型
1	清华大学	学术信息发现平台（水木搜索）	Primo
2	北京大学	未名学术搜索	Summon
3	厦门大学	中文发现	指针搜索
		学术搜索（xsearch）	dspace
4	南京大学	Find + 外文知识发现平台	Find +
		超星中文知识发现平台	超星
		乐致安 Find + 中文发现平台	Find +
5	天津大学	超星经纬学术发现系统	超星
6	浙江大学	求是学术搜索	Summon
7	西安交通大学	中文学术资源发现系统	超星
		学术资源发现系统	Summon
		http://lib.xjtu.edu.cn/ （西交利物浦大学）	EDS
8	东南大学	成贤搜索	Summon
9	武汉大学	图书馆电子资源门户 Primo Central	Primo Central
10	上海交通大学	思源探索	Primo

国内资源发现系统应用现状

平台名称	功能特色	第三方资源混搭 (词条、书封、 书评、目次)	参 考 咨 询	虚 拟 书 架	计 量 工 具	社 群 服 务	手机二 维 码	按学科定制 检 索 结 果
水木搜索 (清华大学图书馆)		√	√	√			√	√
思源探索 (上海交通大学图书馆)		√	√	√		√	√	√
学术资源发现平台 (中国社会科学院图书馆)		√	√	√		√	√	
学术资源发现平台 (香港中文大学图书馆)			√			√		√
资源探索系统(台湾新竹清 华大学图书馆)		√	√			√		
BU Libraries (波士顿大学图书馆)		√	√					
Catalogue(SOLO) (牛津大学Bodleian图书馆)		√	√	√	√			
Search Library (范德堡大学图书馆)		√	√			√		
Everything (普渡大学图书馆)		√	√	√				
Explore Further (大英图书馆)		√	√			√		

国内资源发现系统应用现状

RSS
☆ 保存整页结果到电子书架

选择显示内容

- 在线全文 (35)
- 本馆收藏 (41)

精简检索结果

▣ **主题**

- Computer science (16)
- 单片微型计算机 (8)
- Engineering, Civil (5)
- Arduino (Programmable controller) (4)
- Microcontrollers (4)

更多选项 ▾

▣ **作者**

- SpringerLink (Online service) (15)
- 程晨 (5)
- Arduino, P (5)
- 唐乐 (4)
- University of Washington (4)

更多选项 ▾

▣ **来源**

- SpringerLink (15)
- Safari (5)
- Morgan Claypool (3)
- World ebook Library (2)
- ehron (1)

arduino

Arduino, 是一个基于开放原始码的软硬件平台, 构建于开放原始码 simple I/O 介面版, 并且具有使用类似Java, C 语言的Processing/Wiring开发环境。

显示热门文章 ▾

Results 1 - 10 of 72 : **馆藏图书** 排序: 相关性 ▾ 1 2 3 4

选择显示内容 在线全文 (35) | 本馆收藏 (41)

☆ **Arduino从基础到实践 = Beginning Arduino** 有 2 个版本或分

麦克罗伯特, M. (McRoberts, Michael) 杨继志; 郭敬 2013
国际标准书号 978-7-121-19201-2

预约/续借 馆藏地 详细信息 评论/标签 SFX 虚拟书架 手机二维码 试读信息 **网摘目录** 操作

图书



TP368.1 M283



TP368.1 M353



TP368.1 M607

« Previous

☆ **动手玩转Arduino = Arduino workshop**

博克索尔, J. (Boxall, John), 翁恺 2014
国际标准书号 978-7-115-33596-8

预约/续借 馆藏地 详细信息 评论/标签 SFX 虚拟书架 手机二维码

目录

- 第1章 Hilbert 空间初步
 - 1 Hilbert 空间
 - 2 Sobolev 空间
 - 3 线性算子
 - 4 紧算子与特征展开
- 参考文献
- 第2章 边值问题的变分形式
 - 1 抽象变分形式
 - 2 二次泛函的临界点
 - 3 二阶椭圆边值问题
 - 4 弹性理论的变分原理
 - 6 四阶椭圆方程的边值问题
- 参考文献
- 第3章 里兹-迦辽金(Ritz-Galerkin)方法
 - 1 极小化序列、Ritz 法
 - 2 紧算子方程的 Galerkin 解法

国内资源发现系统应用现状

 RSS
☆ 保存整页结果到电子书架

选择显示内容

- 在线全文 (35)
- 本馆收藏 (41)

精简检索结果

主题

- Computer science (16)
- 单片微型计算机 (8)
- Engineering, Civil (5)
- Arduino (Programmable controller) (4)
- Microcontrollers (4)

更多选项 ▾

作者

- SpringerLink (Online service) (15)
- 程晨 (5)
- Arduino, P (5)
- 唐乐 (4)
- University of Washington (4)

更多选项 ▾

来源

- SpringerLink (15)
- Safari (5)
- Morgan Claypool (3)
- World ebook Library (2)
- ehron (1)

arduino

Arduino, 是一个基于开放原始码的软硬体平台, 构建于开放原始码 simple I/O 介面版, 并且具有使用类似Java, C 语言的Processing/Wiring开发环境。

显示热门文章 ▾

Results 1 - 10 of 72 : 馆藏图书 排序: 相关性 ▾ 1 2 3 4

选择显示内容 在线全文 (35) | 本馆收藏 (41)

☆ **Arduino从基础到实践 = Beginning Arduino** 有 2 个版本或分

麦克罗伯茨, M. (McRoberts, Michael) 杨继志; 郭敬 2013
国际标准书号 978-7-121-19201-2

预约/续借 馆藏地 详细信息 评论/标签 SFX  虚拟书架 **手机二维码** 试读信息 网摘目次

操作 ▾ 

图书



TP368.1 M283

« Previous

动手玩转Arduino = Arduino works

博克索尔, J. (Boxall, John), 翁恺 2013
国际标准书号 978-7-115-33596-8

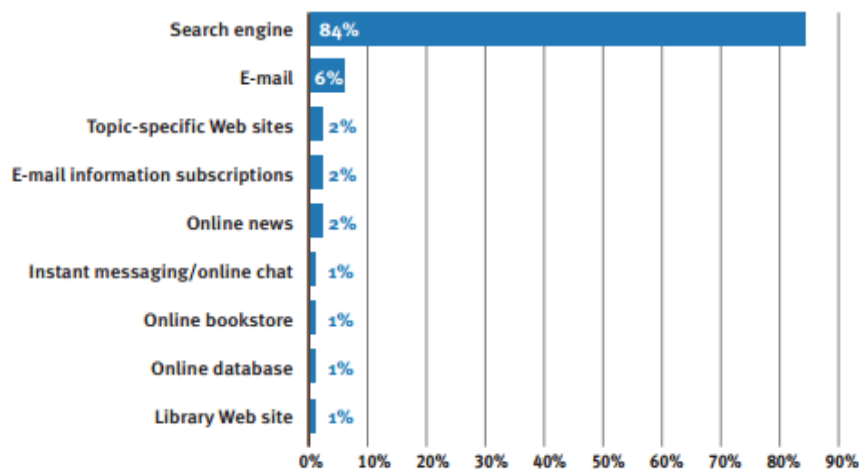
预约/续借 馆藏地 详细信息



资源发现服务给图书馆带来的变化

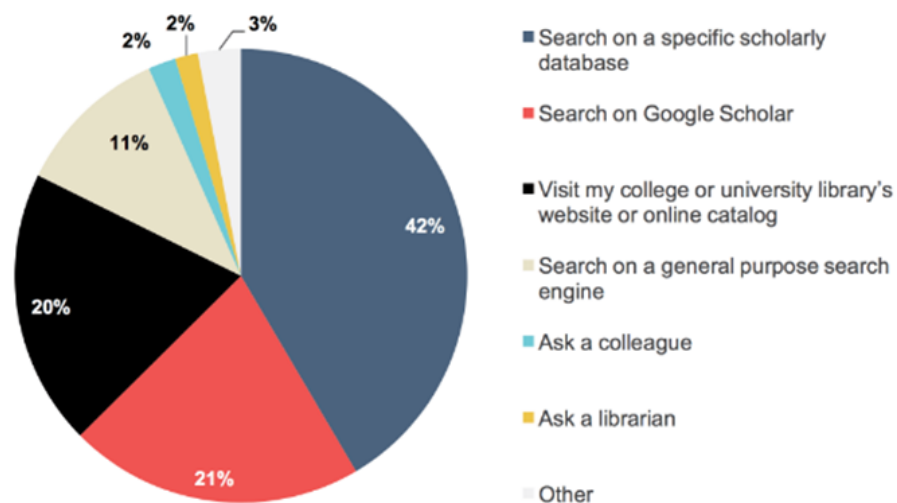
Where Electronic Information Searches Begin— by Total Respondents

Where do you typically begin your search for information on a particular topic?



Source: *Perceptions of Libraries and Information Resources*, OCLC, 2005, question 520.

Figure 3: When you explore the scholarly literature to find new journal articles and monographs relevant to your research interests, how do you most often begin your process? *Percent of respondents who indicated that each option is the starting point for their exploration.*





DNA-guided genome editing using the *Natronobacterium gregoryi* Argonaute

Feng Gao, Xiao Z Shen, Feng Jiang, Yongqiang Wu & Chunyu Han

[Affiliations](#) | [Contributions](#) | [Corresponding author](#)

Nature Biotechnology **34**, 768–
Received 03 June 2015 | Accepted

[Citation](#) [Reprints](#) [Search](#)



Abstract

[Abstract](#) • [Accession codes](#) • [References](#) • [Author information](#) • [Supplementary information](#)

The RNA-guided endonuclease Cas9 has made genome editing a widely accessible technique. Similar to Cas9, endonucleases from the Argonaute protein family also use oligonucleotides as guides to degrade invasive genomes. Here we report that the *Natronobacterium gregoryi*

对科研数据的发现



Accurate and robust genomic prediction of celiac disease using statistical learning
Abraham, Gad ; Tye-Din, Jason A. ; Bhalala, Oneil G. ; Kowalczyk, Adam ; Zobel, Justin ; Inouye, Michael
Arxiv ID: 1301.5948
● 在线全文

在线查看 详细信息 评论/标签 SFX 手机二维码

PLOS Early Career Travel Award Program
Application Deadline: May 31, 2016
APPLY NOW

plos.org create account sign in

PLOS ONE Publish About Browse Search advanced search

OPEN ACCESS PEER-REVIEWED
RESEARCH ARTICLE

Accurate and Robust Prediction of Genetic Relationship from Whole-Genome Sequences
Hong Li, Gustavo Glusman, Chad Huff, Juan Caballero, Jared C. Roach
Published: February 28, 2014 • <http://dx.doi.org/10.1371/journal.pone.0085437>

24 Save	3 Citation
2,523 View	9 Share

Download PDF Print Share

CrossMark

Article Authors Metrics Comments Related Content

Abstract Introduction Abstract

在线查看 详细信息 评论/标签 SFX 手机二维码

Accurate and robust genomic prediction of celiac disease using statistical learning
Gad Abraham 2013
DOI: 10.6084/M9.FIGSHARE.154193.V3



文章

Accurate
Abraham,
Arxiv ID:
● 在线全

在线查看



文章

Accurate
Abraham,
2014, Vol
ISSN: 151
● 在线全

在线查看



研究数据

Accurate
Gad Abra
DOI: 10.6
● 在线全

在线查看



研究数据

Accurate
Gad Abra
DOI: 10.6
● 在线全

在线查看



研究数据

Accurate and robust genomic prediction of celiac disease using statistical learning
Gad Abraham 2013
DOI: 10.6084/M9.FIGSHARE.154193
● 在线全文

在线查看 详细信息 评论/标签 SFX 手机二维码

Accurate and robust genomic prediction of celiac disease using statistical learning
Gad Abraham 2013
DOI: 10.6084/M9.FIGSHARE.154193.V3



DATASET

grs.csv (4.79 kB)



1. The Generic Risk Score

The SNM model that produces the generic risk score (Abraham et al., 2012) is given in Table 1. Also provided in the intercept β_0 . The intercept is not required if we are only interested in ranking patients according to their risk, and do not wish to compare the results with the results in the meta-analysis.

$$\hat{\beta}_0 = \beta_0 + \sum_{j=1}^M \alpha_j \beta_j \quad (1)$$

where α_j is the genotype for the j th SNP in the risk sample.

The easiest way to produce a risk score for a dataset is using PLINK, as it will take care to use the correct minor allele. Assuming that the data are called MAF as in BED/BIM/P3M format:

```
plink --bim bim --score grs_scores.txt --maf MAF
```

which will produce a file named DATA.genetic

The public file can be read into R, and written back to the file grs.txt:

```
R <- read.table("DATA.genetic", header=FALSE)
```

```
intercept <- -0.797229
```

```
grs <- @GRS@ + @SNP@ + @INTERCEPT
```

```
write.table(grs, "grs.txt", "w")
```

supplementary.pdf (183.33 kB)

Download all (188.12 kB)

Share

Cite

Embed

+ Collect (you need to log in first)

Accurate and robust genomic prediction of celiac disease using statistical learning

Version 1 (old) 25.01.2013, 08:58 (GMT) by Gad Abraham

Practical application of genomic--based risk stratification to clinical diagnosis is appealing yet performance varies widely depending on the disease and genomic risk score (GRS) method. Celiac disease (CD), a common immune--mediated illness, is strongly genetically determined and requires specific HLA haplotypes. HLA testing can exclude diagnosis but has low specificity, providing little information suitable for clinical risk stratification. Using six European CD cohorts, we provide a proof-- of--concept that statistical learning approaches which simultaneously model all SNPs can generate robust and highly accurate predictive models based on genome--wide SNP profiles. The high predictive capacity replicated both in cross--validation within each cohort (AUC of

278
views

50
downloads



CATEGORIES

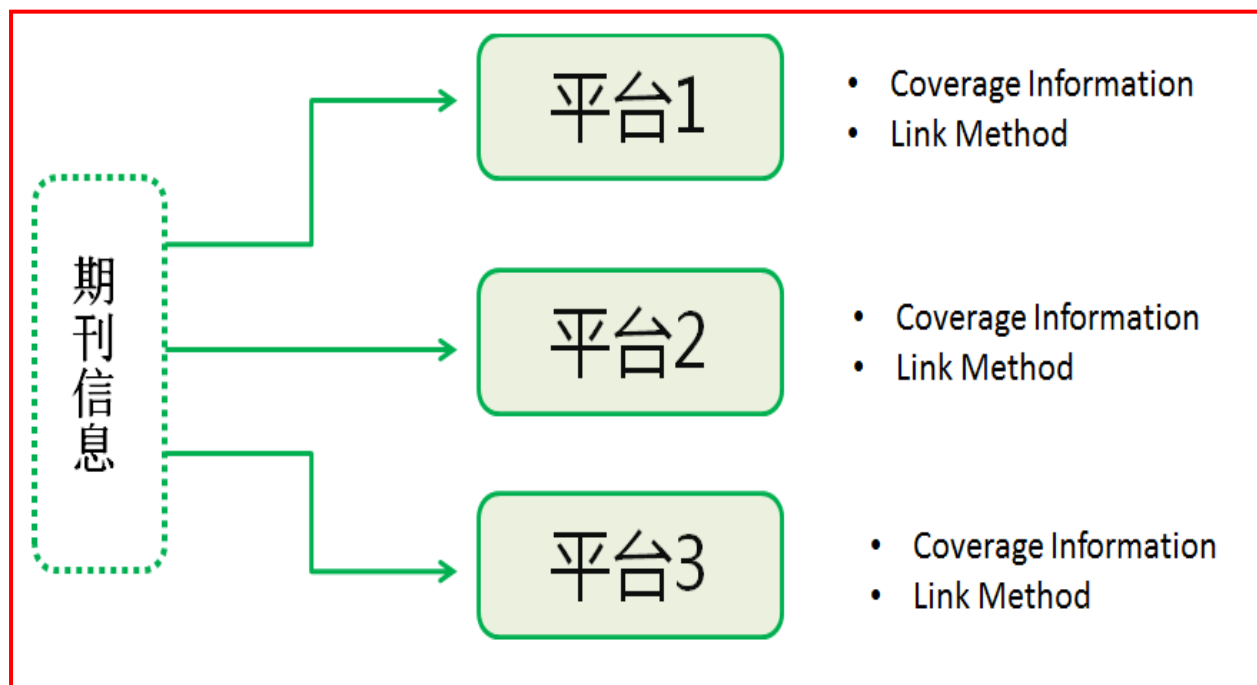
- Bioinformatics
- Genetics

TAGS

- 从资源发现到资源获取
From discovery to delivery
电子资源管理 电子资源揭示

从资源发现到资源获取

用户更关注资源获取，图书馆应利用外部工具做好电子资源的揭示工作，保障资源从发现到获取的渠道畅通。



从资源发现到资源获取



[期刊] CaMnO₃热电薄膜的制备及其性能表征
作者: 詹斌, 兰金功, **林元华** (清华大学 新型陶瓷与精细陶瓷国家重点实验室)
出处: 稀有金属材料与工程 2013 第S1期 P54-56
ISSN: 1002-185X
关键词: 热电 薄膜 CaMnO₃ 溶胶-凝胶
摘要: 采用溶胶-凝胶旋转涂覆法制备CaMnO₃热电薄膜。用PPMS对其热电性能进行了测试。XRD、SEM及Raman光谱表征了薄膜的晶体结构及形貌。随着热处理温度...
基金: 国家“973”计划支持项目(2011CB932506)
获得途径: [电子全文](#) [邮箱](#)

[学位论文] 薄膜流涎机模块化参数化设计技术
作者: **林元华** (南京理工大学)
学位名称: 硕士
出处: 南京理工大学 2012
关键词: 薄膜流涎机 模块化设计 参数化设计 Visual-Basic S
摘要: 薄膜流涎机是生产包装薄膜的主要生产设备。随着国民经济的发展,对薄膜流涎机的要求也越来越高,这就促使薄膜流涎机生产...
获得途径: [电子全文](#) [邮箱](#)



旧版入口 | 手机版 | English | 网站地图 | 帮助中心

cnki中国知网 cnki.net

文献 期刊 博硕士 会议 报纸 外文文献 年鉴 百科 词典 统计数据 专利 标准 更多>>

出版物检索 学位论文 高级检索

文献全文分类 全文

期刊 博硕 会议 报纸 图片 年鉴 统计数据 作者投稿 采编 评价 引文 企业 医院 农业 政府 商务杂志 文艺 科普 手机知网 知识管理软件 下载 专利 成果 标准 法规 古籍 外文 工具书 学术不端 独家 优先出版 项目 法律 教育 城建 书报超市 文化 政报 手机报 数字出版平台 更多

资源总库 海量中外文献, 一站式检索

国际文献总库 14,000多家国外出版社的文献

行业知识服务平台 面向行业网络的知识发现及共享平台

个人/机构数字图书馆 个性化定制及服务的全新资源应用方式

知网动态 更多...
高校图书馆知识服务与创新应用高级研修班成功举办
感恩节, 为导师整理成果!
“学习党的历史 践行群众路线”学习、竞赛活动启动
CNKI E-Learning上线两周年回馈活动开始啦!
《智爱助教辅导平台》免费试用!
《CNKI合同文本库》全新上线, 即查即用!

中国党政系列专业期刊 更多...
· 公共管理 · 经济管理 · 法治政府 · 政党建设
· 马克思主义 · 政治学 · 社会管理 · 文化建设

中国高校系列专业期刊 更多...

糟糕! 谷歌浏览器无法找到
www.sl.shulixhxx.cn
您是不是要找: www.shulixhxx.cn

系统 译助手 形搜索

与资源出版商/平台商携手推动数字资源开放标准、链接标准、建立可行的数字资源管理机制, 提升数字资源的发现及获取能力仍然是图书馆的重要职责。

从资源发现到资源获取

Google

biochemical correlates of calling



学术搜索

找到约 29,700 条结果 (用时0.07秒)

文章

小提示：只搜索中文(简体)结果，可在 学术搜索设置 指定搜索语言

我的图书馆

Biochemical correlates of calling activity in neotropical frogs

Get via Tsinghua Library

Physiological Zoology > Vol. 68, No. 6, Nov. - Dec., 1995 > Biochemical Correlat...



Biochemical Correlates of Calling Activity in Neotropical Frogs

Catherine Robb Bevier
Physiological Zoology
Vol. 68, No. 6 (Nov. - Dec., 1995), pp. 1118-1142

Published by: [The University of Chicago Press](#). Sponsored by the [Division of Comparative Physiology and Biochemistry, Society for Integrative and Comparative Biology](#)
Stable URL: <http://www.jstor.org/stable/30163796>
Page Count: 25

Download PDF
Cite This Item
Journal Info

« Previous Item | Next Item »

Viewing page 1118 of pages 1118-1142

JSTOR [PDF]
Get via Tsinghua Library

Get via Tsinghua Library

4大出版集团和75家大学出版社开放期刊文章的元数据

	主流发现工具	Google	Doi/Metadata
Springer	√	√	√
Elsevier	√	√	√
Wiley	√	√	√
Taylor	√	√	√

- 从资源发现到知识发现
From resource discovery to knowledge discovery
发现服务新动向

从资源发现到知识发现

新工具：Yewno “Knowledge discovery, contextualized”



Hyper Knowledge

We unlock the ability for people to explore and digest information based on seemingly unrelated concepts at scale.



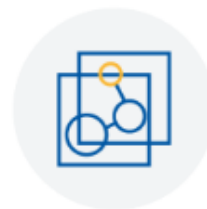
High Quality Data

All data sources are vetted, high quality content providers, creating a level of unparalleled accuracy and sophistication.



Accuracy From Depth

Yewno continuously evolves as it intakes data, expanding its algorithms and improving the depth of its contextualized knowledge.



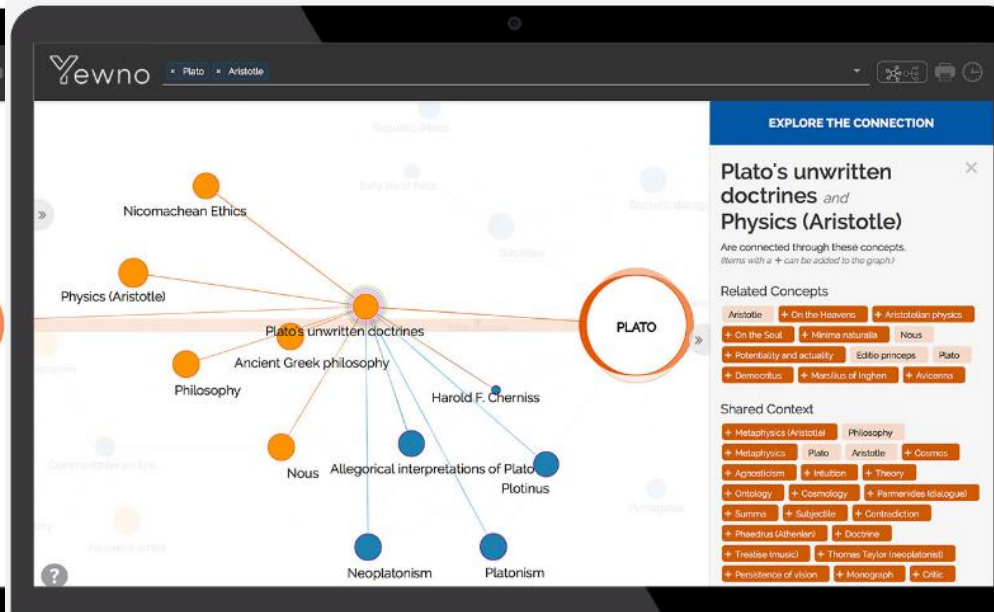
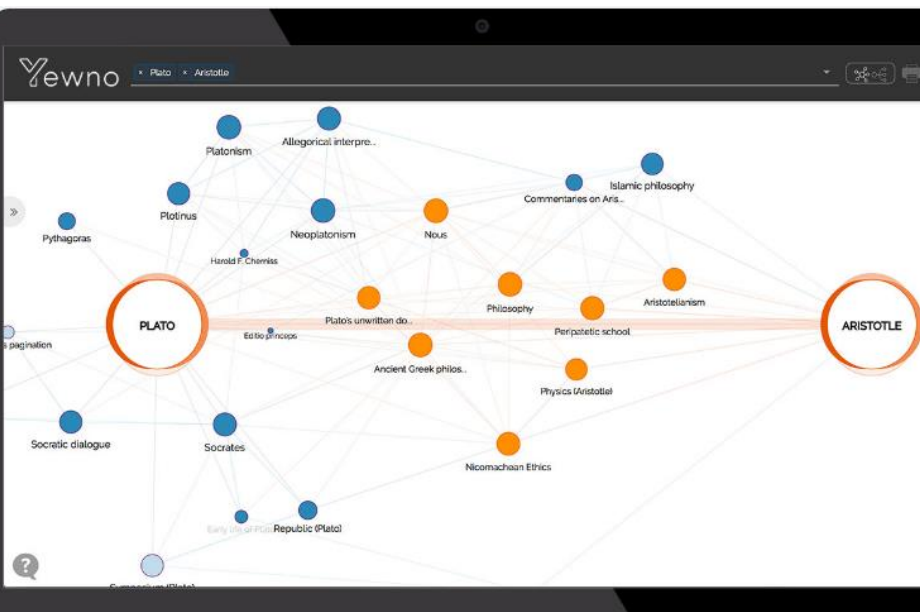
Multi-Dimensional Connections

Yewno analyzes data in a non-linear environment, drawing tangible connections across many concepts.

从资源发现到知识发现

新工具：Yewno “Knowledge discovery, contextualized”

Yewno for Finance



目前，Cornell、MIT等部分图书馆正在试用Yewno。

新方法：discovery + linked data

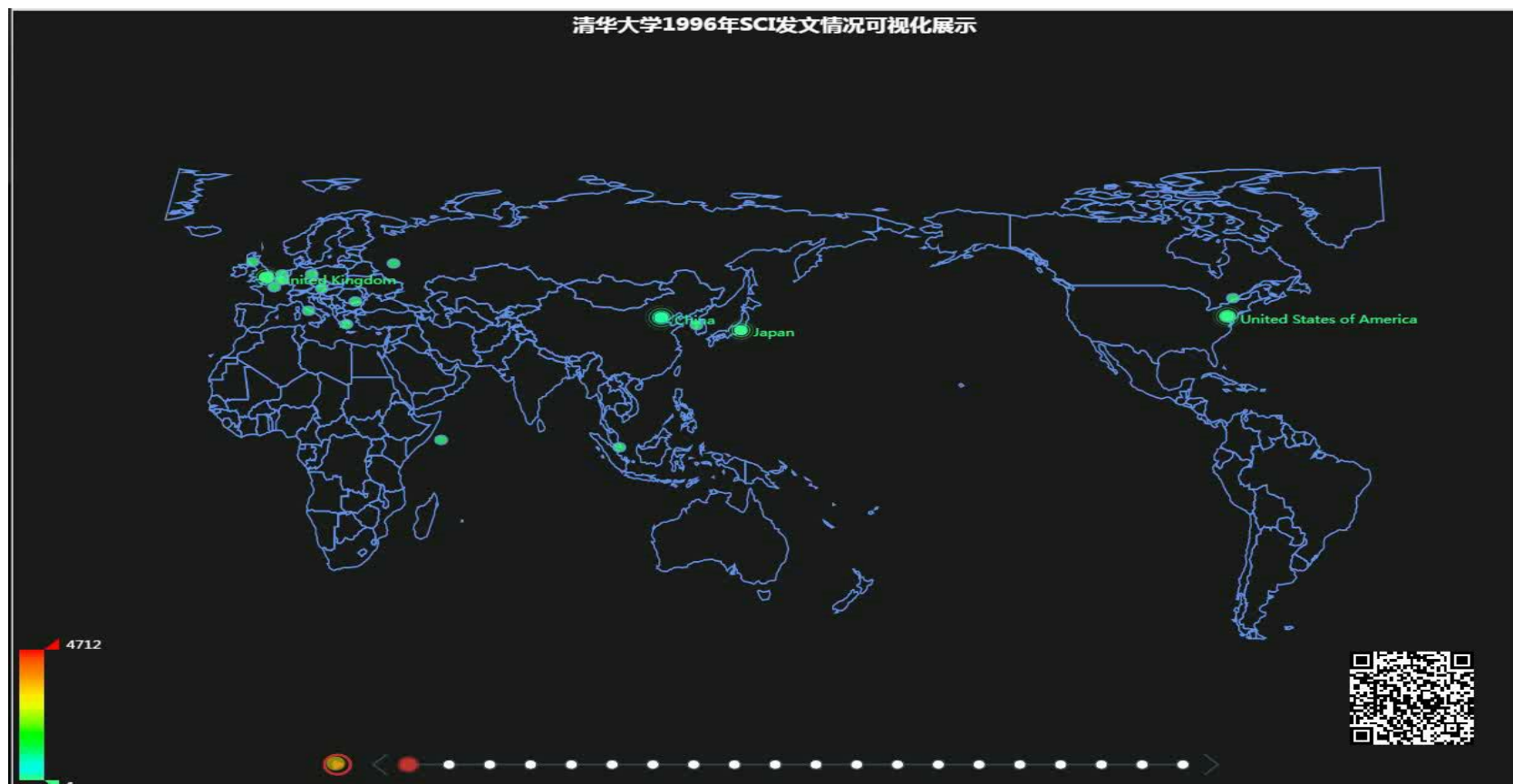
LD4L Linked Data for Discovery:

“ Implement a demonstration SRSIS Search discovery and access system across the three SRSIS instances (Jon Corson-Rikert, Brian Lowe, Cornell Hydra developer, Cornell SRSIS developer) ”

图书馆的数据关联化，支持关联数据的发现服务。

从资源发现到知识发现

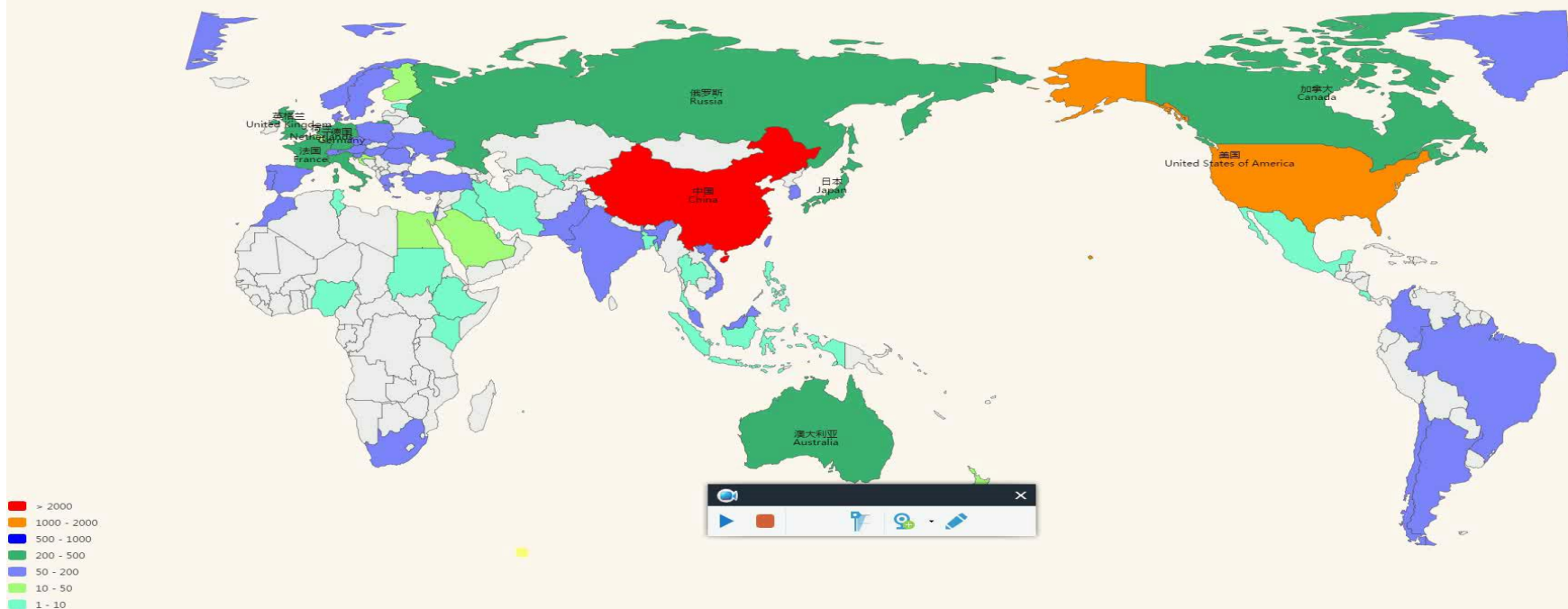
新展示：可视化手段展示数据促进知识发现



从资源发现到知识发现

新展示：可视化手段展示数据促进知识发现

清华大学2015年SCI发文情况可视化展示





Thank you!

