

# **Polyinfo** **恒通博联**

基于大数据的电子资源解决方案  
助力高校双一流建设

北京恒通博联科技有限公司

史保屹  
13501082353/18610271523



# Agenda



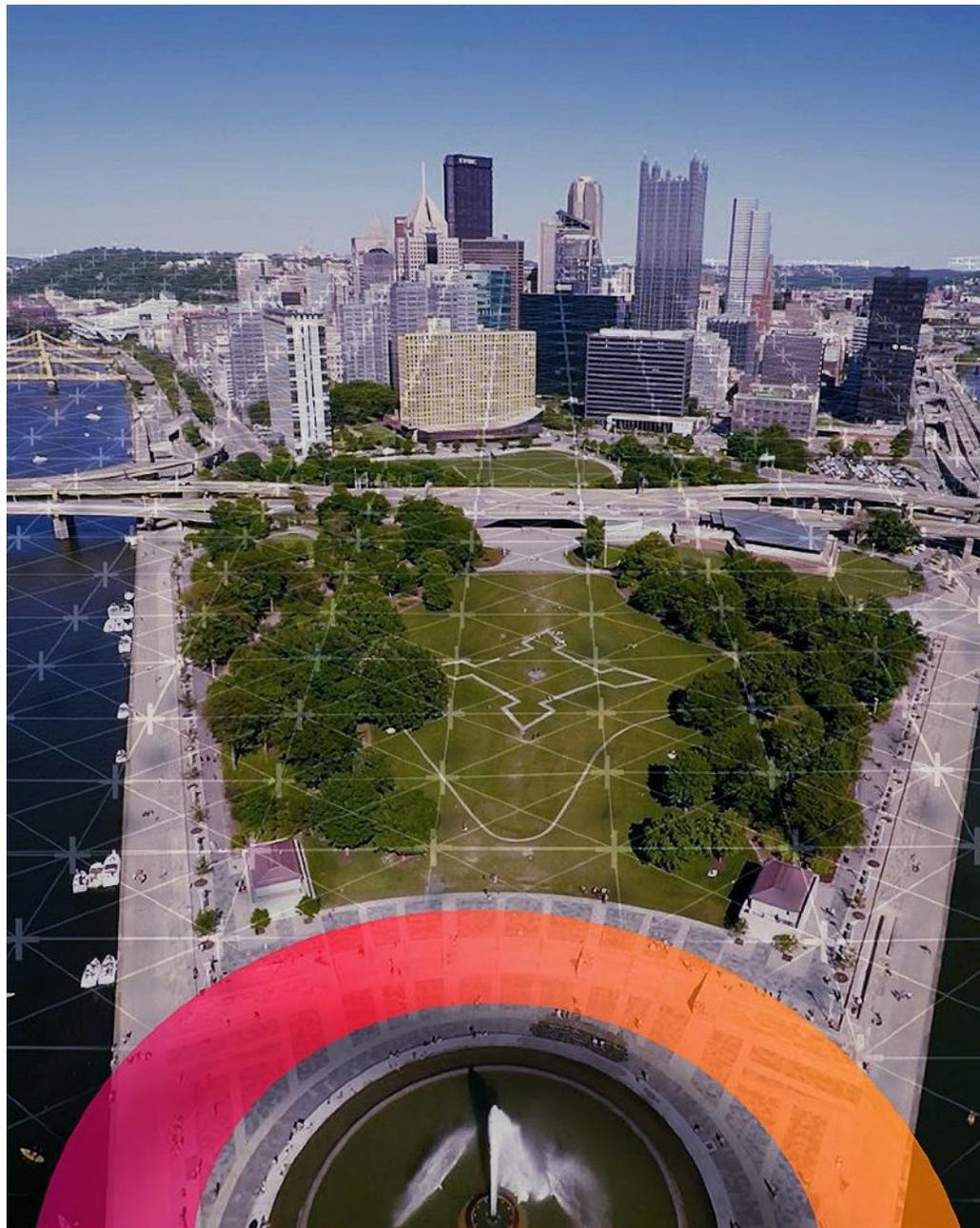
1. 公司介绍
2. 大数据双一流
3. PolyInfo图书馆电子资源解决方案大数据云平台
4. 展望未来



# 公司介绍

---

- 北京恒通博联科技有限公司是一家高技术企业，拥有领先的互联网数据采集、统计和挖掘三大技术，
- 公司的业务主要集中在网络流量分析、网络行为审计、图书馆电子资源分析、大数据、数据挖掘。
- 公司致力于成为一家大数据、机器学习、数据挖掘、人工智能的专业研究Data的公司



# 数字化走向数据化

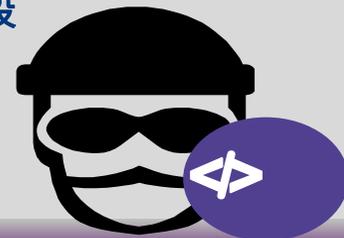
## 大数据时代到来

- 掌握的数据越来越多，越来越活，数据的价值也日益凸显
- 数据的挖掘、分析与整合
- 如今，每个人在数据空间中都会留下痕迹
- 折射其兴趣爱好、需求意愿、研究方向
- 电子资源访问行为数据档案



## 数据挖掘价值

- 利用数据进行科学决策
- 在校领导、管理者及师生中培养与建立“拿数据说话”习惯已是势在必行
- 绘制出读者的“阅读轨迹”
- 分析相关电子资源访问数据便可以洞悉和预判现实中师生员工的专业行为，准确定位师生的需求，从而实现精细化管理，支撑高校双一流建设
- 从资源对学校发文引文，教学科研
- 的支撑



# 双一流建设给大学带来的机遇

- “双一流”不再将高校分成三六九等
- 不管是211、985，不论是部属高校还是地方院校，只要有学科优势和特色,都有机会获得政府财政支持
- 学科建设成为重中之重



# PolyInfo 图书馆电子资源分析大数据云系统组成

## 电子资源用户行为分析

- 详细记录每一个用户访问电子资源的搜索，浏览，下载行为，精确到时间、用户、平台、库、刊、作者、单位、学科
- 提供大量针对用户、资源等多维度统计分析报表

## PolyInfo 图书馆电子资源分析优化云系统

## 电子资源绩效评估

- 馆藏核心期刊的保障率
- 学科分析，重点学科保障，学科跟踪
- **JCR 1区 2区** 本校发文引用期刊的馆藏保障情况
- 基于学科的核心资源保障情况
- 各种库、刊、学科的使用报告

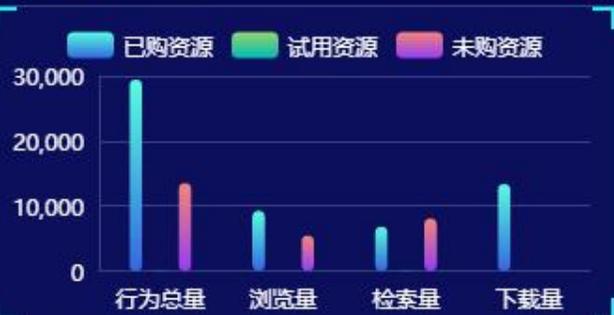
## Polyinfo 图书馆电子资源绩效评估云系统

# 基于大数据的图书馆电子资源优化分析云系统

## 图书馆电子资源分析优化云系统

进入系统

### 资源概述 (近一月)



### 资源使用行为趋势

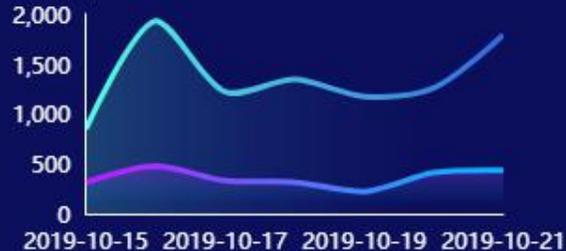


### 资源使用占比 (近一月)

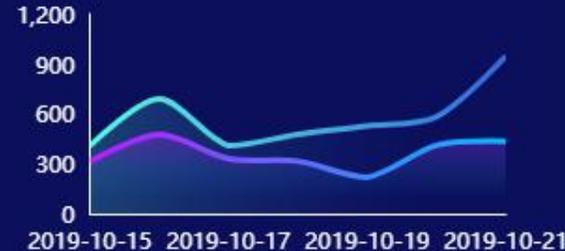


### 资源使用趋势对比

#### 平台行为总量



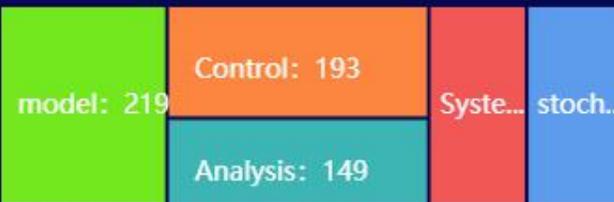
#### 数据库行为总量



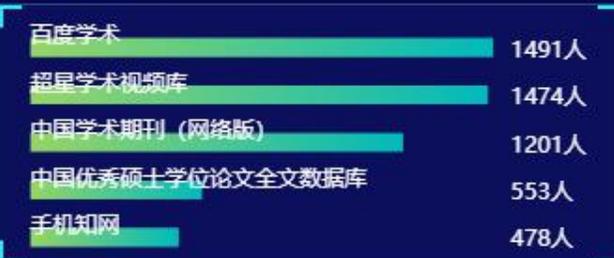
### 热门学科 (近一月)



### 检索关键词 (近一月)



### 最热资源库 (近一月)



# 基于大数据的图书馆电子资源绩效评估云系统

学科分析 数据库分析 期刊分析 发文引文

核心资源分析 分级汇总 馆藏分析 授权管理

## 图书馆电子资源绩效评估云系统

### 馆藏分析



### 资源保障率



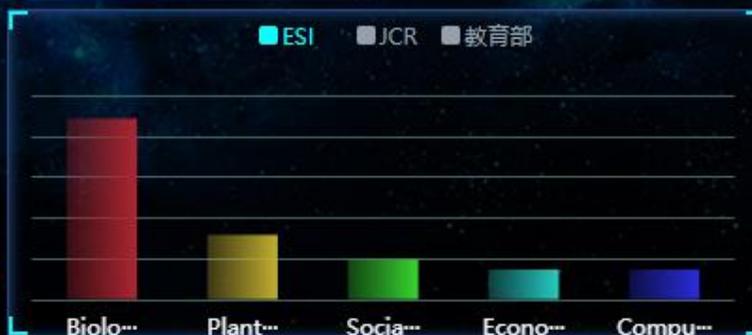
### 热门期刊



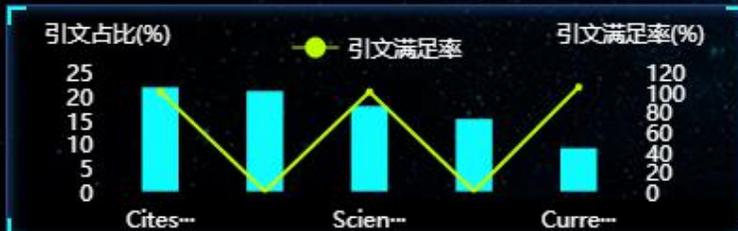
### 资源分级汇总



### 热门学科



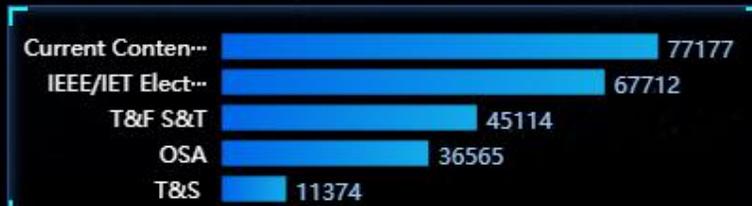
### 引文数据库分析



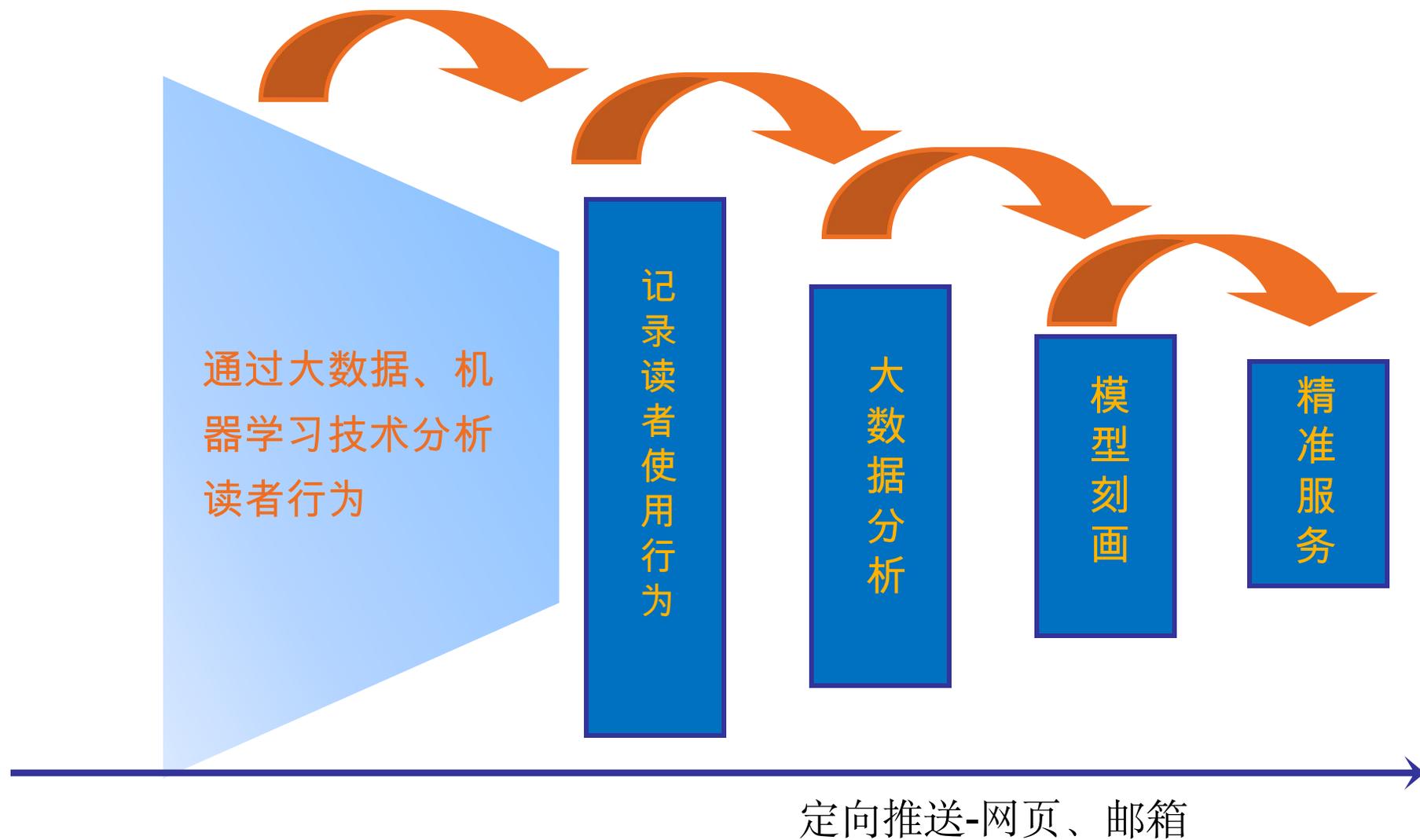
### 热门学科保障率趋势



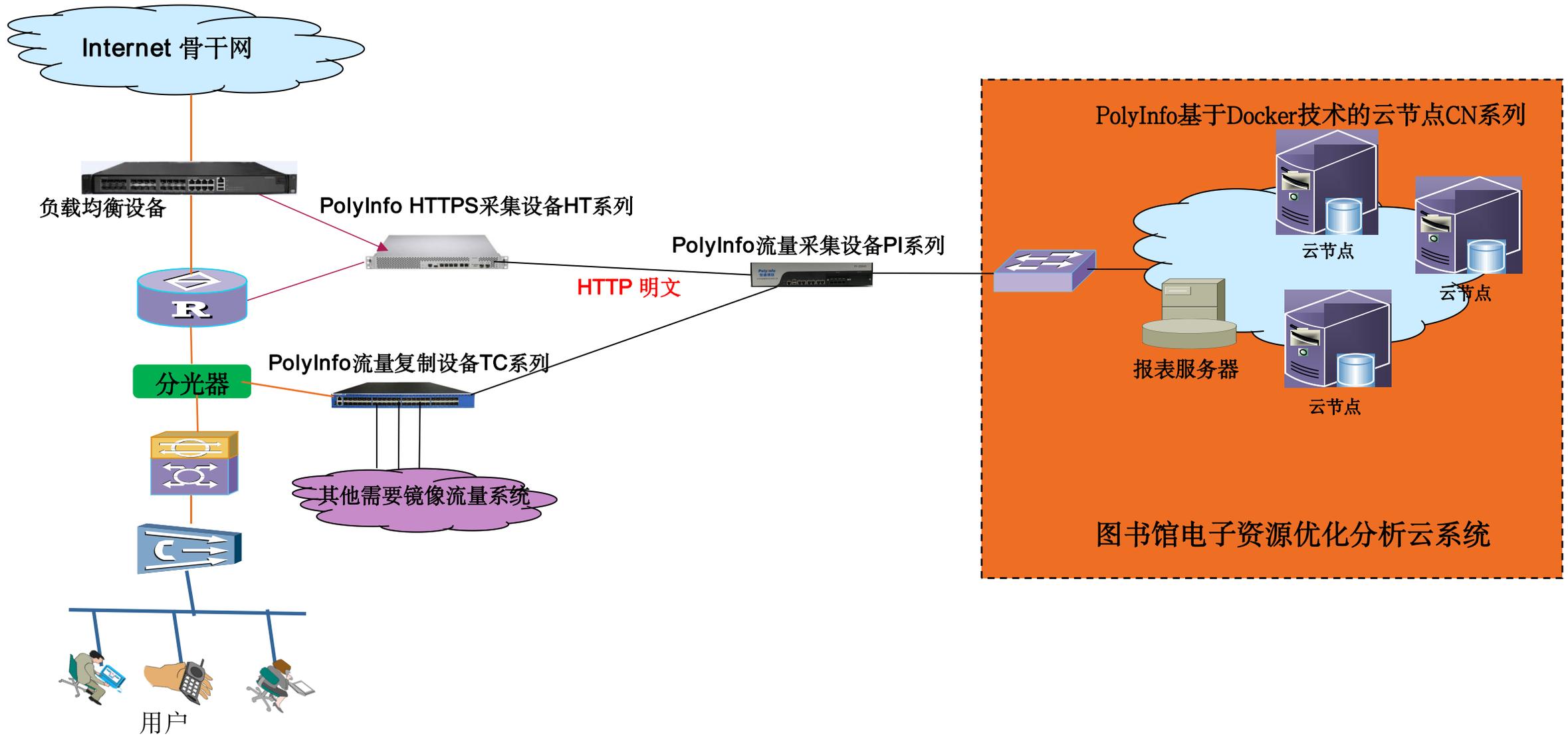
### 热门数据库



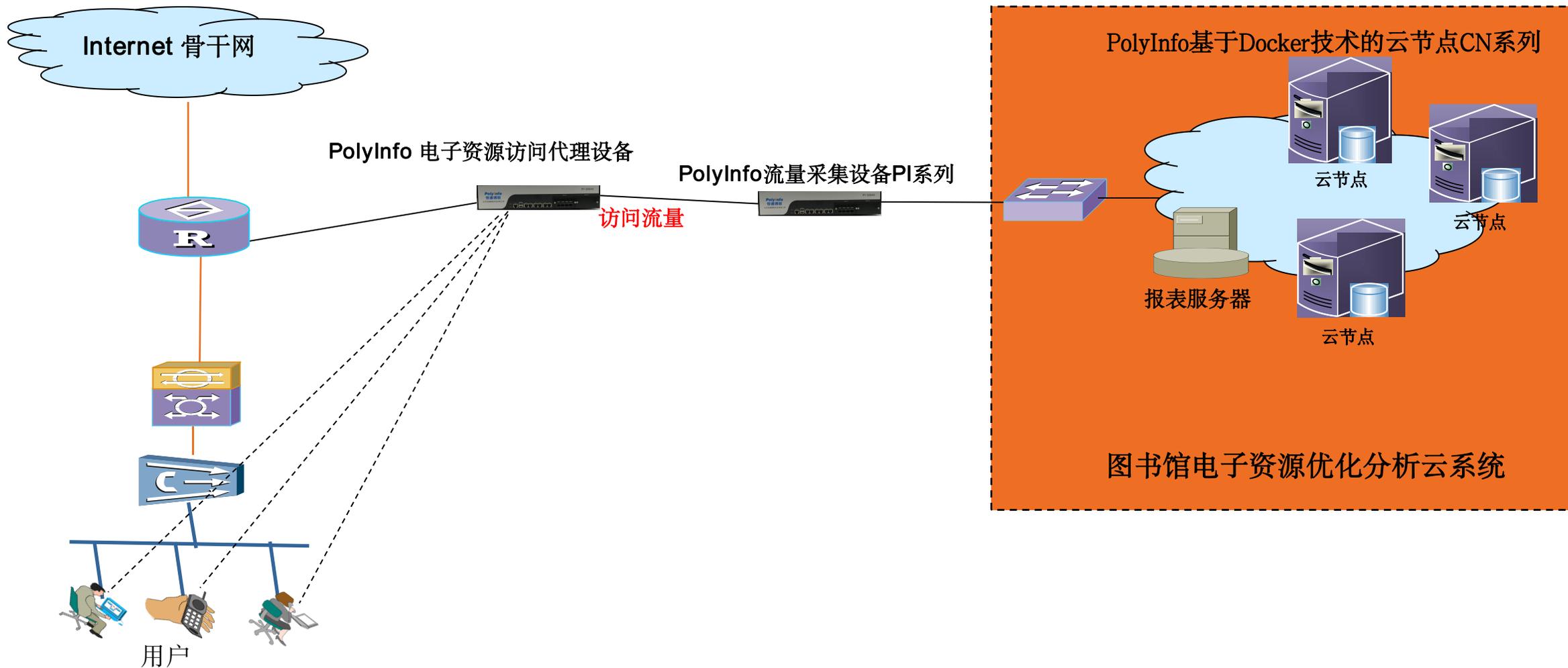
# 大数据驱动精准服务



# PolyInfo 图书馆电子资源优化分析云系统（用户行为）解决方案1



# PolyInfo 图书馆电子资源优化分析云系统（用户行为）解决方案2



# HTTPS 带来的挑战

## 免费方案

基于国际标准  
**SUSHI**协议进  
行数据传送

**HTTPS**资源方  
返回**XML**格式文  
件

## 收费方案 1

**HTTPS** 引流到  
专业解密  
**HTTPS**设备

得到相应的  
**HTTPS**资源的  
**HTTP**访问记录

## 收费方案2

基于远程访问代  
理设备导航页面

得到相应的  
**HTTPS**资源的  
**HTTP**访问记录

率先支持**Counter R5** 报告收割

# HTTPS 采集数据展示

The screenshot displays the PolyInfo library system interface, which is used for analyzing and optimizing electronic resources. The interface is in Chinese and features a navigation menu on the left, a main content area with various analysis tools, and a data table at the bottom.

**Navigation Menu (Left):**

- 主页 (Home)
- HTTP CDR
- 综合统计 (Overall Statistics)
- 读者行为 (Reader Behavior) - Expanded
  - 行为查询 (Behavior Query)
  - 库预览行为 (Library Preview Behavior)
  - 库搜索行为 (Library Search Behavior)
  - 库下载行为 (Library Download Behavior)
  - 库行为分析 (Library Behavior Analysis)
  - 平台预览行为 (Platform Preview Behavior)
  - 平台搜索行为 (Platform Search Behavior)

**Main Content Area:**

The main content area shows the system title "PolyInfo 图书馆电子资源分析优化云系统" and the user "您好 ht". It includes a navigation bar with options like "实时URL", "行为查询", "资源统计", "URL查询", "地域分布", "预览达人", "来源分析", "资源库总排行榜", "健康查询", and "库学科统计".

**Data Table:**

时间	用户名	源IP	标题	归属库	归属网站	归属平台	学校资源组	文献种类	学科种类	库的种类	操作
11-29 13:49:23	201812544	101.76.214.218	Phys. Rev. Lett. 121, 220001 (2018) - Essay: Half a Century of the Standard Model	PHYSICAL REVIEW JOURNALS	APS Journals	APS Journals		期刊	综合	外文	
11-29 13:50:54	201812544	101.76.214.218	Phys. Rev. Lett. 121, 190503 (2018) - Disentanglement Cost of Quantum States	PHYSICAL REVIEW JOURNALS	APS Journals	APS Journals		期刊	综合	外文	
11-29 14:24:27	201712303	172.25.181.39	Phys. Rev. E 70, 016608 (2004) - Robust method to retrieve the constitutive effective parameters of metamaterials	PHYSICAL REVIEW JOURNALS	APS Journals	APS Journals		期刊	综合	外文	

# 资源库支持多达上千种

- 支持数千种电子资源库，可以自定义分类和编辑。

电子资源平台 电子资源网站 电子资源库 学校资源配置 自定义用户组

电子资源平台 电子资源网站 电子资源库 学校资源配置 自定义用户组

请输入组名称 查询

+新增 编辑 删除

+	序号	+	序号	组名	创建时间	修改时间	操作
+	1	+	1	开放资源	2018-03-16 19:44:29	2018-03-16 19:44:29	+
+	2	+	2	试用资源	2018-03-16 18:48:54	2018-03-16 18:48:54	+
+	3	+	3	外文资源	2018-03-16 18:04:42	2018-03-16 18:04:42	+
-		-	4	中文资源	2018-03-16 17:26:42	2018-03-16 17:26:42	+

学校资源组配置

库名称	URL	操作
	http://www.cnki.net/	

# 记录用户访问每一个电子资源的动作

---

- 预览行为
- 搜索行为
- 下载行为
- 统计报表



# 用户访问电子资源行为查询

- 发生时间
- 用户名
- 电子资源类别
- 资源网站名
- 场景-包括预览、下载、搜索
- 还原原始页面

发生时间	源IP	用户名	归属库	归属网站	归属平台	学校资源组	文献种类	学科种类	库的种类	文件名称
11-07 08:02:16	10.13.4 0.223	10.13.4 0.223	中国知网文献检索	中国知网	中国知网	中文资源	检索平台	综合	中文	山东半岛城市群城市就业结构变动特征分析_曹佩.caj
11-07 08:02:16	10.13.1 69.199	10.13.1 69.199	中国学术期刊(网络版)	中国知网	中国知网	中文资源	期刊	综合	中文	甘肃北山Eoredlichia的发现及下寒武统划分对比_范国琳.pdf
11-07 08:02:16	10.13.2 16.134	10.13.2 16.134	中国学术期刊(网络版)	中国知网	中国知网	中文资源	期刊	综合	中文	地形分析中栅格中间层构建的尺度适应性_张立朝.pdf
11-07 08:02:16	10.13.4 0.223	10.13.4 0.223	中国知网文献检索	中国知网	中国知网	中文资源	检索平台	综合	中文	青岛市海洋第三产业与整体经济发展关系实证研究_穆丽娟.caj
11-07 08:02:16	10.13.1 69.199	10.13.1 69.199	中国知网文献检索	中国知网	中国知网	中文资源	检索平台	综合	中文	我国商品房价格影响因素分析_基于区域差异视角的实证检验_葛红玲.caj

能够为用户提供:

- 用户对所有的图书馆电子资源的各种行为的访问日志记录
- 提供基于 (源IP、目的IP、用户名、URL、网站名、类型、子类型、电子资源类别) 的查询功能

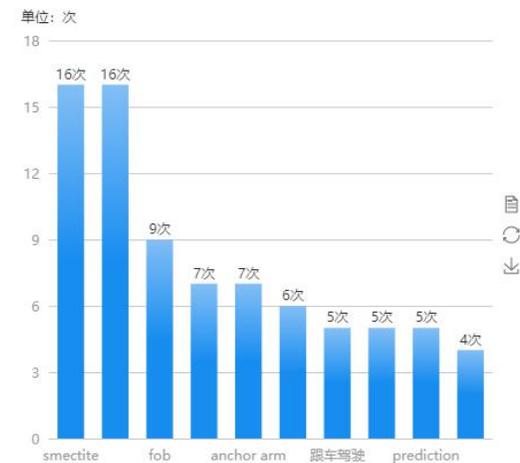


# 库搜索关键词统计

## 读者兴趣热点

— 关键词风云榜 —

关键词	搜索次数
1 smectite	16↑
2 IEEE Access	16↑
3 fob	9↑
4 journal of cleaner produ...	7↑
5 anchor arm	7↑
6 受付身	6↑
7 跟车驾驶	5↑
8 环境公益诉讼	5↑
9 prediction	5↑
10 金属矿山	4↑



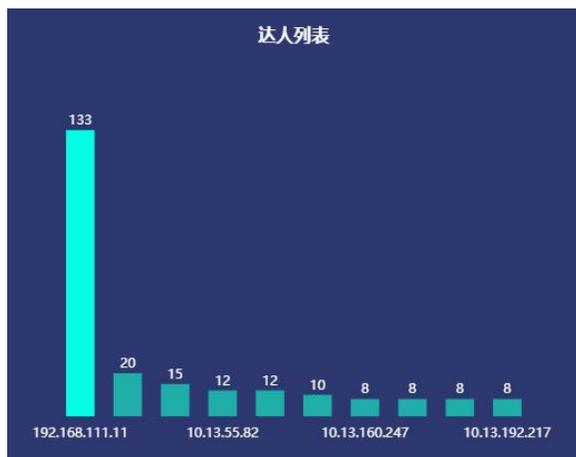
库关键字统计列表					
序号	日期	关键词	次数	独立ip/账号	操作
1	2018-05-18_2018-05-18	data	48	15	
2	2018-05-18_2018-05-18	Mining	42	13	
3	2018-05-18_2018-05-18	电子商务	42	23	
4	2018-05-18_2018-05-18	e-commerce	32	15	
5	2018-05-18_2018-05-18	based	25	9	

显示第 1 到第 5 条记录, 总共 4572 条记录 每页显示 5 条记录

1 2 3 4 5 ... 915

# 活跃读者

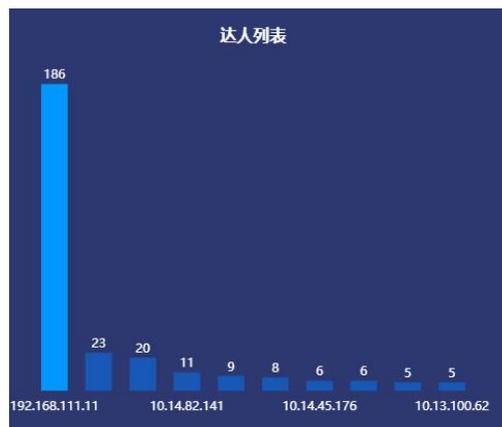
## 预览达人



192.168.111.11浏览标题

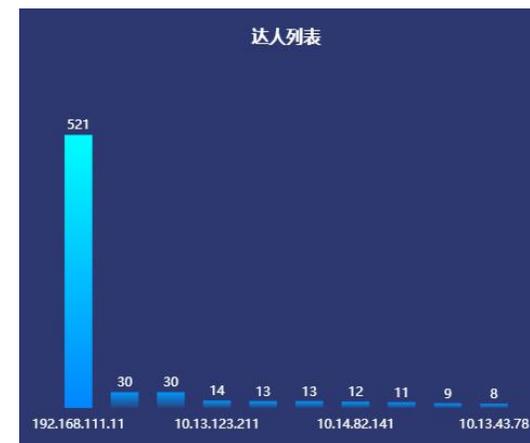
标题	浏览次数
JOURNALOFCHEMICALTHERMODYNAMI--	3
ACTACHIMICASINICA	2
CONTINUUMMECHANICSANDTHERMODY--	2
IETControlTheoryandApplicatio--	2
JournalofEnergyChemistry	2
MultistepSchemesforForwardBac--	2
创新视角下知识搜索的研究回顾与展望	2
抓住MOOC之机促进计算机与软件工程专业教学改革	2
深度学习及其在图像分类识别中的研究综述	2
股权融资还是债务融资——绩效反馈与融资决策的关系研究	2

## 搜索达人



时间	用户名	源IP	检索内容	扫描站	归属网站	归属平台	学校资源库	修改类型
10-22-2 3:52:09	192.168.111.11	192.168.111.11	measurement & control	Web of Science核心合集	Web of Science	Web of Science	中文资源	Windows 10
10-22-2 3:51:30	192.168.111.11	192.168.111.11	measurement & control	Web of Science核心合集	Web of Science	Web of Science	中文资源	Windows 10
10-22-2 3:51:14	192.168.111.11	192.168.111.11	xue jianliang	Web of Science核心合集	Web of Science	Web of Science	中文资源	Windows 7
10-22-2 3:50:46	192.168.111.11	192.168.111.11	measurement & control	Web of Science核心合集	Web of Science	Web of Science	中文资源	Windows 10
10-22-2 3:50:37	192.168.111.11	192.168.111.11	xue jianliang	Web of Science核心合集	Web of Science	Web of Science	中文资源	Windows 7
10-22-2 3:47:58	192.168.111.11	192.168.111.11	zhou guangzhu	Web of Science核心合集	Web of Science	Web of Science	中文资源	Windows 7

## 下载达人



192.168.111.11下载文件

文件名	下载次数
_沪蚁_生存现状调查研究_周选.pdf	3
CMC粉尘火灾爆炸风险分析及预防措施研究_何勇.caj	2
N <sub>2</sub> 喷射对瓦斯爆炸抑制作用的初步实验研究_陈秋敏.caj	2
_蚁族_弱势群体在租来空间中的呐_省略_广州天河区_蚁族...	2
_蚁族_的困惑_看当前大学生就业难题_赵红艳.pdf	2
_蚁族_聚居群落生态环境研究_高永良.pdf	2
中国生态文明建设评价指标体系构建与发展策略研究_彭一然.c...	2
原油管道油气爆炸抑制机理研究_袁健威.caj	2
含NaCl荷电细水雾对甲烷爆炸抑制特性的研究_吴丽洁.caj	2
基于我国资源环境问题区域差异的生态文明评价指标体系研究_李...	2

# 库综合统计-平台-库-基于counter报告

年度报告		详细查询		时间: 2019 2018 2017 2016 2015						分类: 资源库 平台				
1	中国知网文献检索	年度报告	详细查询	2019	2018	2017	2016	2015	资源库	平台				
2	中国学术期刊 (网络版)													
3	中文社会科学引文索引 (CSSCI)													
4	百度学术													
5	中国优秀硕士学位论文全文数据库													
6	中国学术期刊 (光盘版)													
7	新思维													
8	Web of Science核心合集													
序号	资源库	序号	归属平台	行为总量	浏览次数	检索次数	下载次数	1月总量	2月总量	3月总量	4月总量	5月总量	6月总量	7月总量
1	中国知网文献检索	1	中国知网	433030	167518	178446	87066	31505	12746	47847	99631	95557	38622	40819
2	中国学术期刊 (网络版)	2	中国社会科学引文索引	75620	612	74953	55	346	207	287	270	152	21	73645
3	中文社会科学引文索引 (CSSCI)	3	百度学术	63650	20762	42888	0	0	0	2444	15240	14500	8397	6690
4	百度学术	4	读秀	20763	4979	14583	1201	1673	3095	3007	2920	2718	736	1639
5	中国优秀硕士学位论文全文数据库	5	新思维	19060	0	19060	0	0	0	1297	8099	7754	1910	0
6	中国学术期刊 (光盘版)	6	Web of Science	18516	1868	15909	739	1742	588	1479	1689	1213	407	1578
7	新思维	7	超星	14315	5105	3892	5318	1149	513	887	1179	1128	1170	1648
8	Web of Science核心合集	8	百链学术搜索	8003	2067	5842	94	826	225	753	1001	699	350	1324
		9	万方数据知识服务平台	5592	4349	1232	11	360	91	588	1155	1160	516	493

# 库排行榜

**升降榜:** 库访问排名前十的名次升降变动统计。

**热度榜:** 库访问排名前十的访问指数变动统计。

**积分榜:** 库积分总额排名前十的积分变动统计。

**【注】:** 本榜单统计的是上个月到上个月的排名明细，到本月初进行数据刷新，重排。

排名	库名称	名次变动
1	中国知网文献检索	0-
2	中国学术期刊(网络版)	0-
3	中国优秀硕士学位论文全文数据库	1↑
4	中文发现	6↑
5	中国优秀博硕士学位论文全文数据库	0-
6	百链_期刊数据库	13↑
7	读秀_期刊数据库	1↓
8	万方中国学术期刊数据库	1↑

排名	库名称	访问指数变动
1	中国知网文献检索	52334↑
2	中国学术期刊(网络版)	15769↑
3	中国优秀硕士学位论文全文数据库	2112↑
4	中文发现	3952↑
5	中国优秀博硕士学位论文全文数据库	2509↑
6	百链_期刊数据库	4441↑

# 库健康审查

## 健康查询、统计

连接正常 连接超时 连接失败 未到检测时间

健康查询表

时间	库名称	URL	8am	12am	2pm	6pm	9pm
2018-05-19	Lexis Advance全球法律信息数据库	http://advance.lexis.com/?identityprofileid=S9JNZW52643	🟢	🕒	🕒	🕒	🕒
2018-05-19	美国生物科学研究所电子期刊	http://www.ingentaconnect.com/content/aibs	🟢	🕒	🕒	🕒	🕒
2018-05-19	美国芝加哥大学出版社电子期刊	http://www.jstor.org/publisher/ucpress	⚠️	🕒	🕒	🕒	🕒
2018-05-19	Scopus	https://www.scopus.com/	🟢	🕒	🕒	🕒	🕒
2018-05-19	INFORMS美国运筹学和管理学研究协会期刊	https://pubsonline.informs.org/	🔴	🕒	🕒	🕒	🕒
2018-05-19	Science online科学	http://www.sciencemag.org/subscriptions/accessinfo.dtl	⚠️	🕒	🕒	🕒	🕒

健康统计表

时间	库名称	URL	连接正常	连接失败	连接超时	总检查次数
2018-05-19~2018-05-19	地球科学世界出版社期刊	https://pubs.geoscienceworld.org/#D	1	0	0	1
2018-05-19~2018-05-19	英国玻璃技术学会电子期刊	http://www.ingentaconnect.com/content/sgt	1	0	0	1
2018-05-19~2018-05-19	国家图书馆Emerald回溯内容全国在线	http://www.emeraldinsight.com	1	0	0	1
2018-05-19~2018-05-19	美国生物科学研究所电子期刊	http://www.ingentaconnect.com/content/aibs	1	0	0	1
2018-05-19~2018-05-19	IOS Press出版社电子期刊	https://metapress.com	0	1	0	1

## 告警配置

+新增 编辑 删除

邮箱配置表

序号	发件人	收件人	操作时间
1	15225404398@163.com	13898957@qq.com	2018-05-19 11:35:43

通知条件配置

名称查询 查询

+新增 编辑 删除

通知条件配置表

序号	名称	异常类型	异常次数	周期类型	周期天数	创建时间	修改时间
1	测试	失败	3	每天	0	2017-12-27 10:52:39	2017-12-27 10:52:39

# 恶意下载-行为封堵



## 行为封堵-保障电子资源有序下载

### 分析

- 通过大数据分析发现异常下载用户
- 掌握异常下载用户在具体电子资源下载量

### 控制

- 设定相应的下载阈值，超过阈值的用户进行限制
- 制定策略，对具体用户到某个电子资源访问控制

### 日志

- 统计每日被挡用户
- 过滤用户的排名

## 封堵策略配置

添加

资源库： 请选择

用户组：

撤销

»

>

<

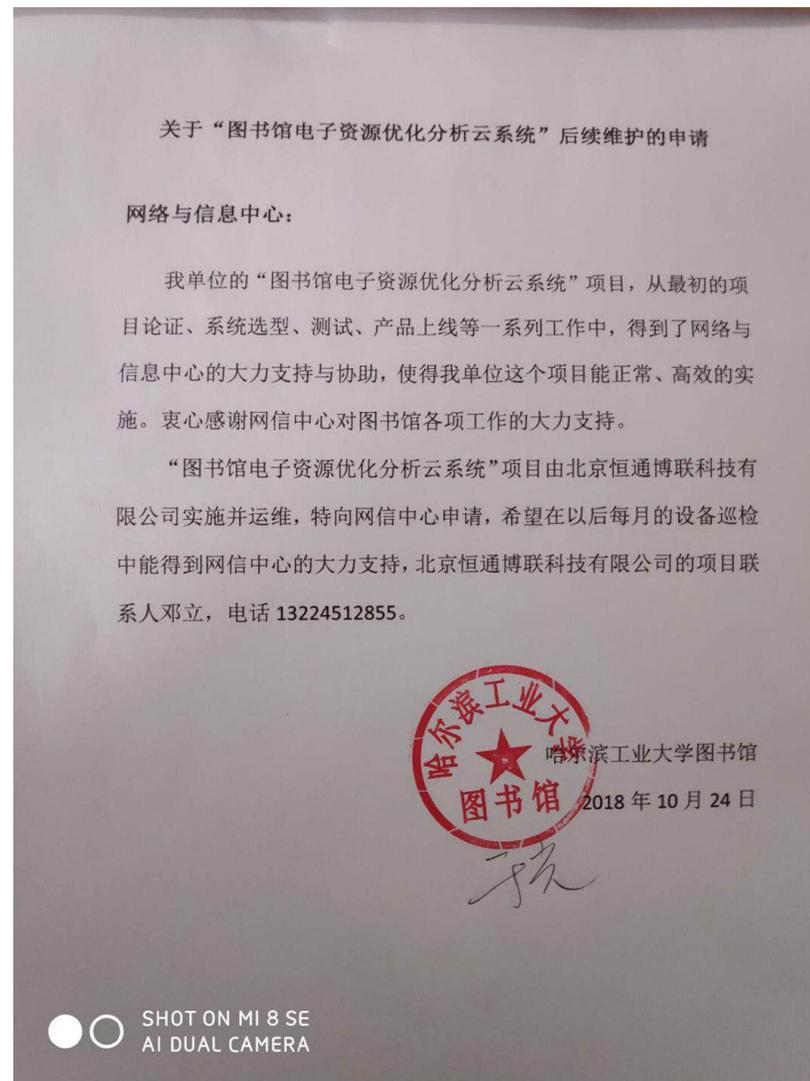
«

天下载次数：

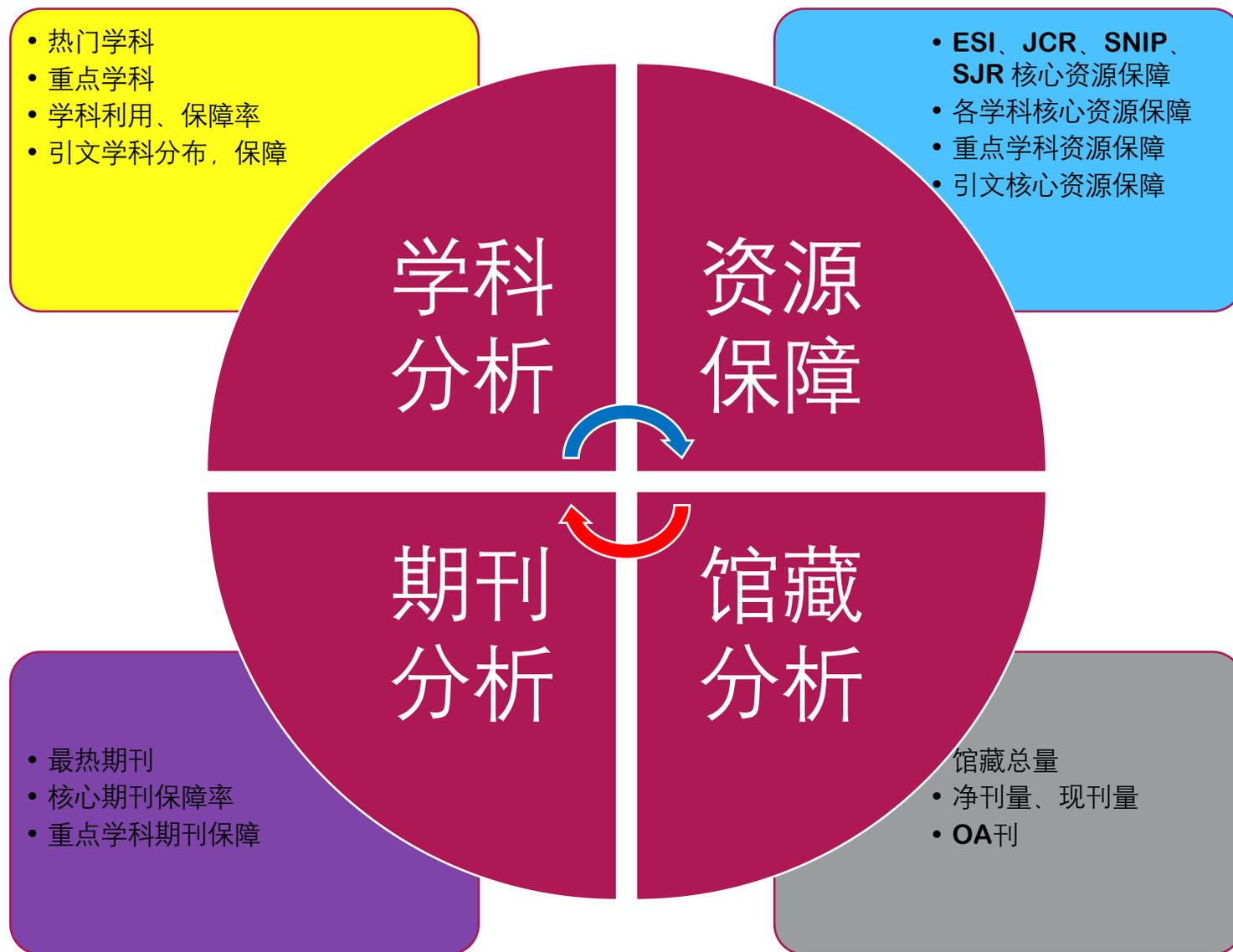
月下载次数：

取消 保存

# 相关资质

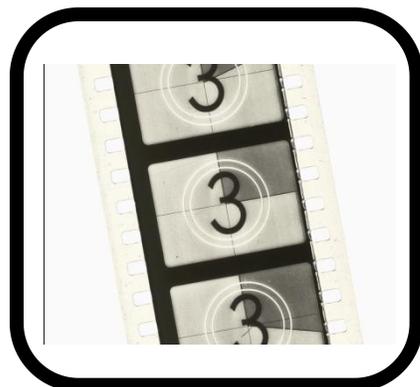
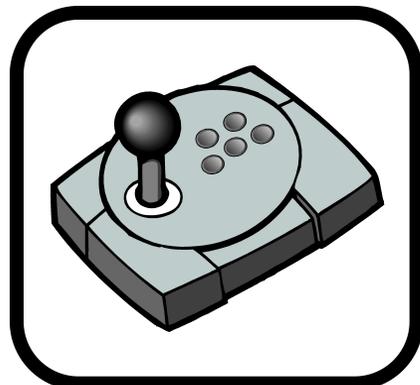
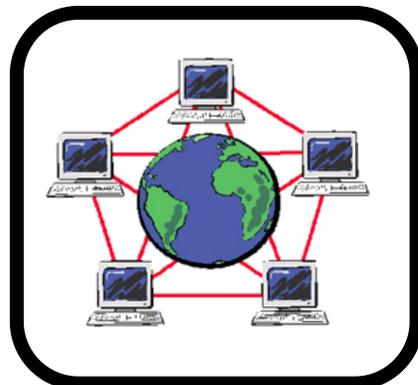


# 大数据分析电子资源绩效助力双一流建设



# 电子资源绩效评估的必要性

- 高校电子资源经费投入逐年增加
- 核心资源对高校学术研究、核心期刊发文，高校排名有很大帮助
- 保障教学、科研所需的文献信息资源
- 图书馆改进服务手段，确保服务质量的重要手段。



# 图书馆电子资源绩效分析云系统（1）



## 馆藏分析

- 馆藏总量、现刊量、净刊量、OA刊量、数据库数量
- 馆藏核心期刊匹配率（ESI、JCR、SNIP、SJR）



## 学科分析

- 热门学科
- 学科保障率
- 学科有下载期刊量
- 学科利用率
- 学科保障量
- 重点学科



## 数据库分析

- 热门数据库
- 数据库核心期刊匹配（JCR、中科院JCR）
- 数据库期刊统计（期刊数量、独有刊、重复刊、现刊、OA）

# 图书馆电子资源绩效分析云系统（2）



## 期刊分析

- 热门期刊
- 期刊概述
- 详细期刊信息（下载量、学科、期刊收录等）



## 核心资源分析

- 资源保障量
- 资源保障率
- 核心资源分析列表



## 分级汇总

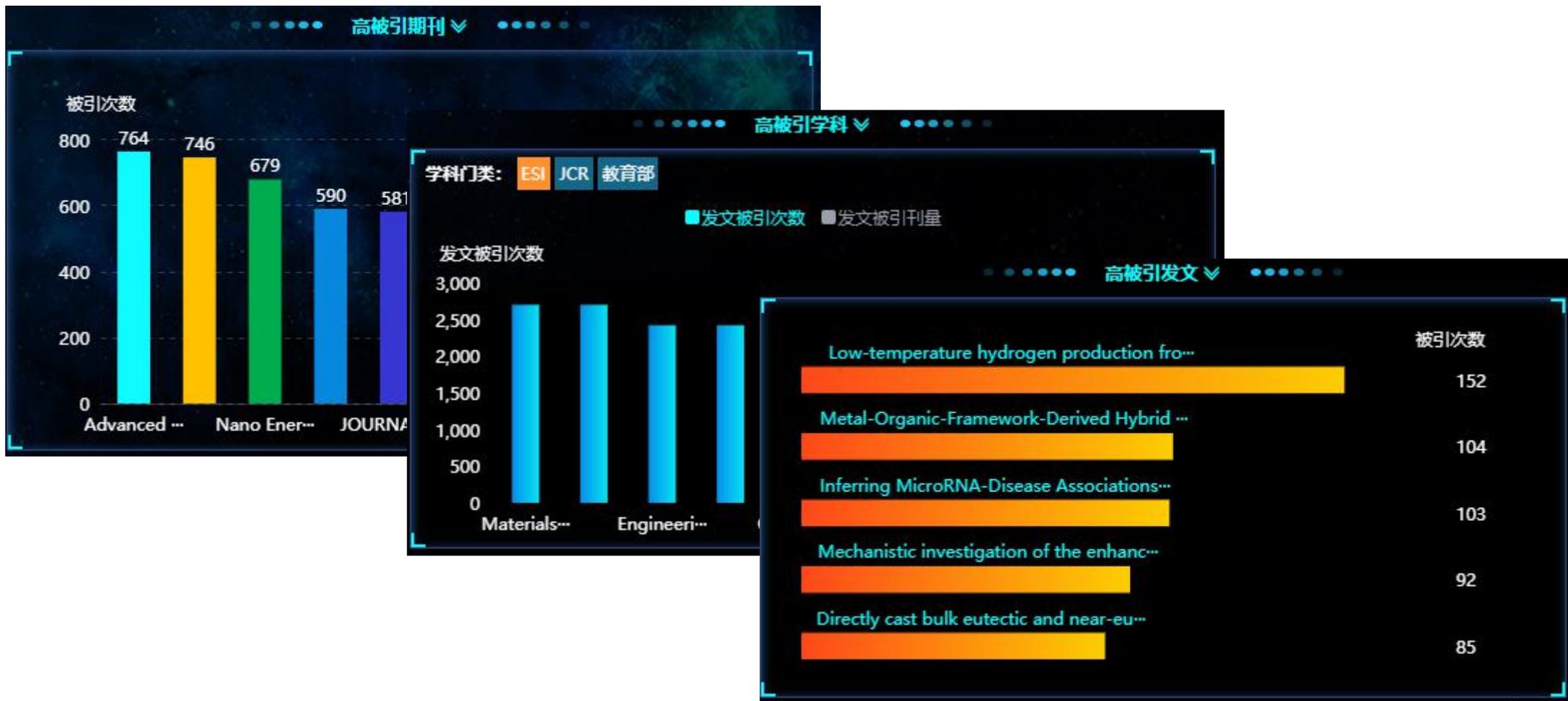
- 按下载量统计
- 按SNIP指数统计
- 下载分级列表

# 国家自然-社会基金馆藏资源保障

2018年最新国家自然科学基金项目



# ESI-发文引文保障



# 学科分析-学科概述



# 学科分析-学科列表

外文期刊

- ESI 、 JCR
- 教育部

中文期刊

- 中文社会科学引文索引
- 中文核心期刊要目总览

The screenshot shows a web interface for '学科列表' (Subject List). On the left, there are search filters: '学科选择' (Subject Selection) with '类型' (Type) set to '外文期刊' (Foreign Journals) and '时间' (Time) set to '2018'. Below this, there are checkboxes for 'ESI', 'JCR', and '教育部' (Ministry of Education), all of which are checked. A '查询' (Query) button is present. On the right, a table titled '学科列表- 外文期刊- 2018' (Subject List - Foreign Journals - 2018) displays data for various disciplines. The table has three columns: the discipline name, '已保障资源(10967)' (Secured Resources (10967)), and '未保障资源(0)' (Unsecured Resources (0)).

学科列表- 外文期刊- 2018	已保障资源(10967)	未保障资源(0)
- ESI		
Agricultural Sciences	336	0
Biology & Biochemistry	431	0
Chemistry	527	2
Clinical Medicine	1873	6
Computer Science	396	0
Economics & Business	422	1
Engineering	832	3

# 学科分析-学科利用率 (1)

学科利用率

年份  
(  
2018.2017.)

学科门类  
(ESI、JCR、  
教育部)

保障维度  
(已保障、未  
保障、全部)

学科下载量

年份  
(  
2018.2017.)

学科门类  
(ESI、JCR、  
教育部)

保障维度 (已  
保障、未保障、  
全部)

# 学科分析-学科利用率 (2)



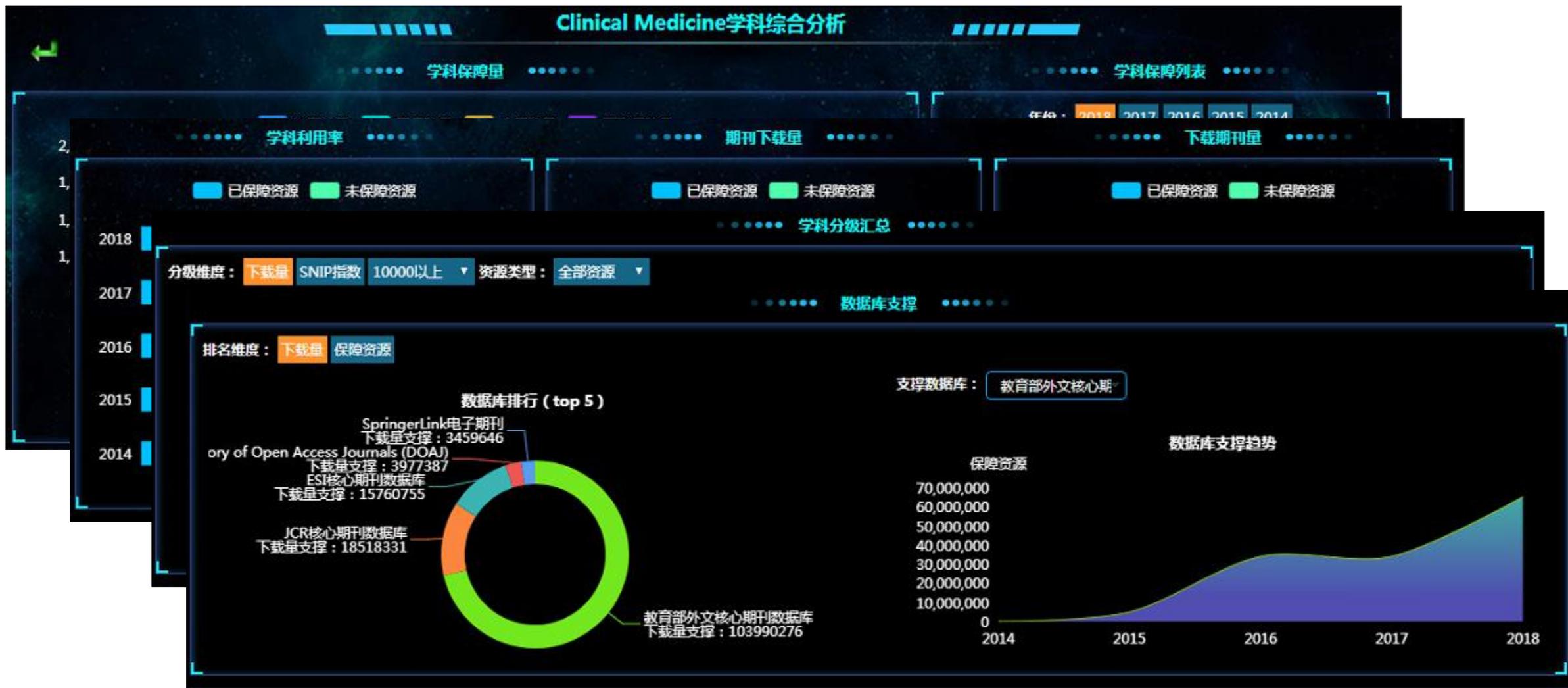
# 学科分析-学科保障统计



# 学科分析-学科分级统计



# 学科分析-具体某学科综合分析



# 学科分析-重点学科

配置重点学科

重点学科

类型: 外文期刊 时间: 2018 配置重点学科

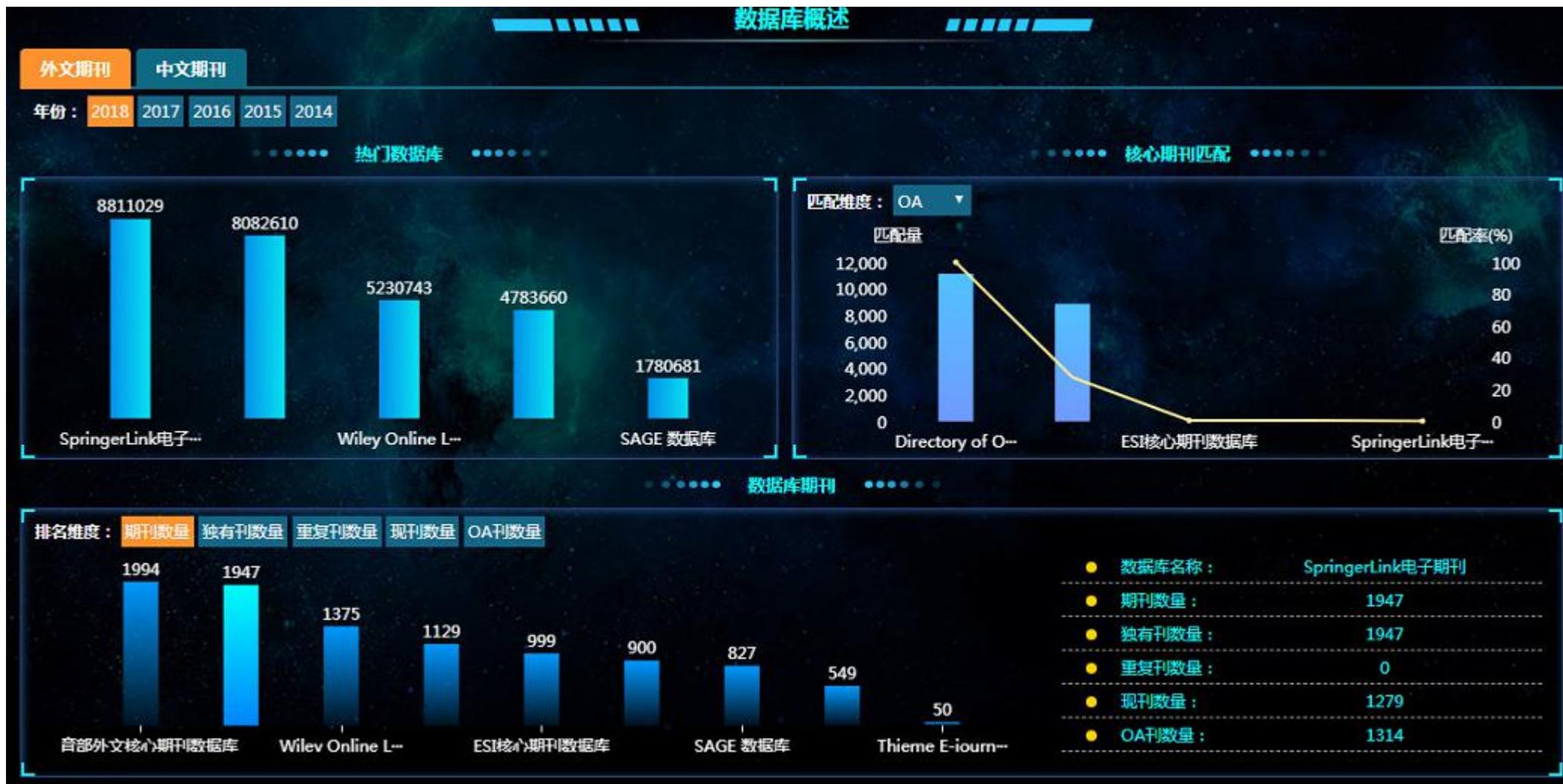
	已保障资源	未保障资源	保障率	下载量	利用率
— 教育部					
— 经济学					
— 理论经济学					
政治经济学	1	0	0%	37079	100%
经济思想史	116	0	0%	10965631	100%
经济史	113	0	0%	7284595	100%
西方经济学	2	0	0%	37445	100%
世界经济	0	0	0%	0	0%
人口、资源与环境经济学	0	0	0%	0	0%
— 应用经济学					
国民经济学	1	0	0%	37079	100%



# 学科分析-数据库支撑



# 数据库分析-数据库概述



# 数据库分析-热门数据库



# 数据库统计-数据库期刊统计



# 期刊分析-核心期刊匹配



# 期刊分析



# 核心资源分析



# 分级汇总



# 馆藏分析



# Future-用户学术画像及推荐系统大数据系统

## 用户画像

- 对用户多个维度进行画像
- 单位、专业、年龄、电子资源的使用行为

## 机构画像

- 对用户多个维度进行画像
- 单位、专业、年龄、电子资源的使用行为

## 学科画像

- 对单位学科设置，学科预测，学科发展趋势
- 根据电子资源学科行为，热点，**SCI** 发文，引用等



THANK  
YOU

[sby@polyinfo.net.cn](mailto:sby@polyinfo.net.cn)