

智能时代的图书馆 知识服务

2020年7月6日 杨新涯

图书馆

CHONGQING UNIVERSITY
LIBRARY

第一单元 啥子是知识服务？

信息-知识-情报



信息
知识
情报

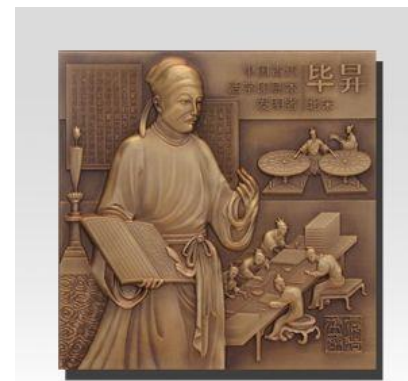
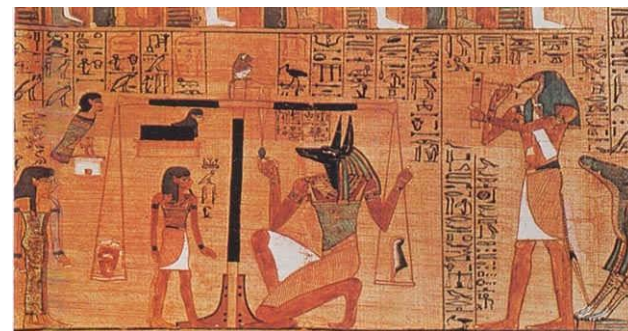
文献

图书馆

- 信息是物质的基本属性，是客观世界中物质的状态和运动变化的反映。知识是被人类认知到的信息。情报是知识传递过程中，使受众的知识结构发生变化的那一部分知识。
- 文献：记录有知识的一切载体。
- 图书馆是收集、整理文献（**知识**）并提供利用的社会机构，是人类社会最重要的社会制度之一。

文献发展的4个关键点

- 文字的出现（公元前3400年）
- 纸张（公元前3000年莎草纸）
- 活字印刷术（1000年毕昇，1440年谷腾堡铅活字）
- 计算机（1946年ENIAC）



猝不及防的变化：信息不对称

实体图书馆时代

知识
不足

数字图书馆时代

知识
过剩

有什么不同?



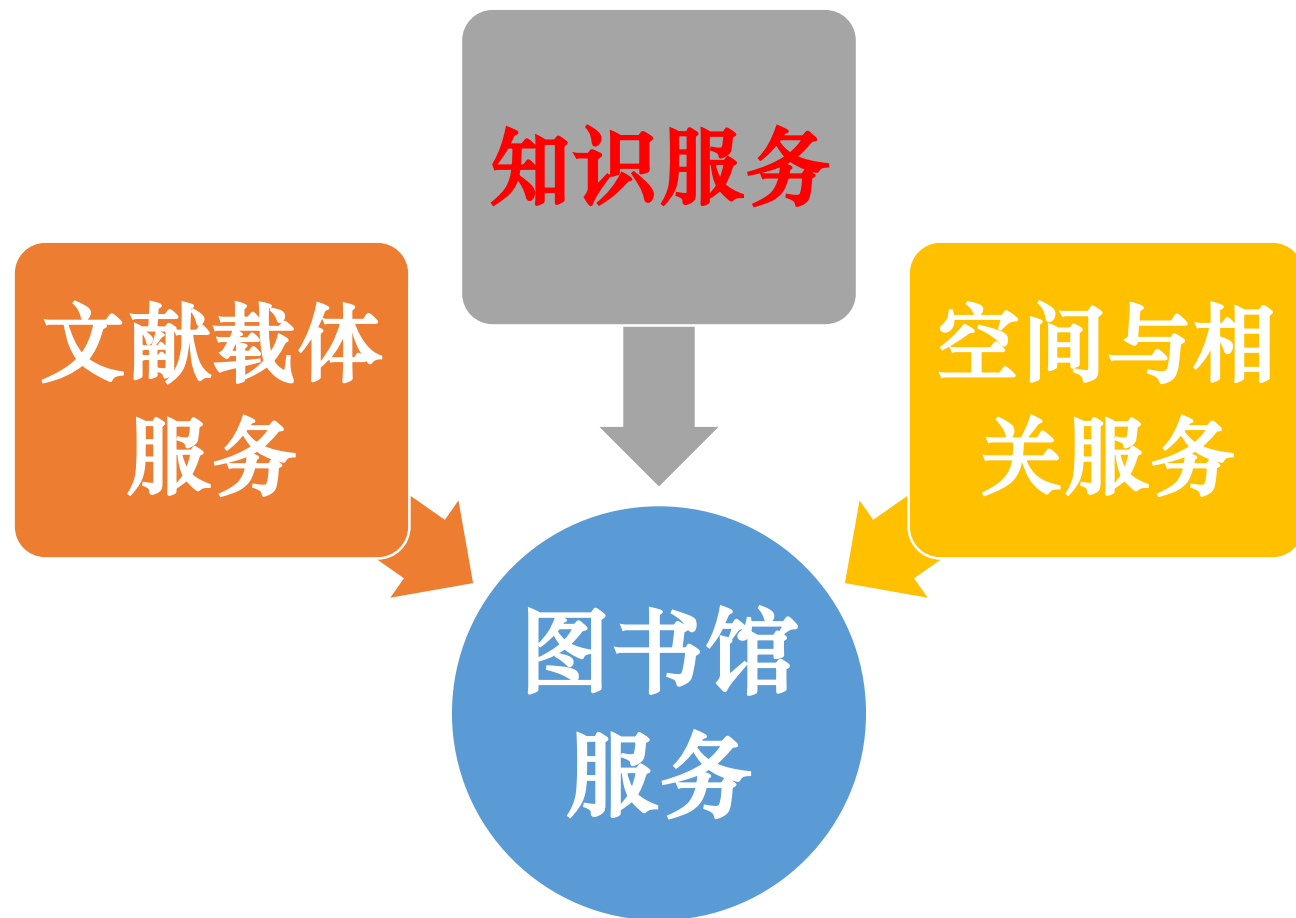
知识
不足

- 目录
- OPAC
- 联合目录
- 馆际互借
- 文献传递

知识
过剩

- 科技查新
- 纸电融合
- 知识发现
- 个性服务
- 精准推送
- 学术评价
- 研究引导
- 知识产权
-

图书馆服务的三方面



需要的图书馆学、情报学方法

- ✓ 核心期刊的二次文献
 - ✓ 主题词标引
 - ✓ 归一化处理
 - ✓ 引文分析
 - ✓ 相关评价指标体系的确定
-
- ✓ 用途：管理与决策的智库、课题选题、高被引方向

欢迎商榷

- 对象：一次文献内容
- 方法：纸电融合的、标准化的二次文献
- 手段：全数据化的
- 目标：解决知识过剩
- 服务：在线的、精准的、个性化的

第二单元 人工智能

无人工智能，无智慧图书馆

- 2016年1月，中国工程院启动了《中国人工智能2.0发展战略研究》重大咨询项目。
- 2017年7月，国务院印发《新一代人工智能发展规划》，提出了五种人工智能的技术形态，即从**数据到知识到决策**的大数据智能、从处理单一类型媒体数据到不同模态（视觉、听觉和自然语言等）综合利用的跨媒体智能、从“个体智能”研究到聚焦群智涌现的群体智能、从追求“机器智能”到迈向人机混合的增强智能[8]、从机器人到智能自主系统。
- 科技部启动建设了十五家国家新一代人工智能开放创新平台，分别是自动驾驶（百度）、城市大脑（阿里云）、医疗影像（腾讯）、智能语音（科大讯飞公司）、智能视觉（商汤集团）、视觉计算（上海依图）、营销智能（明略科技）、基础软硬件（华为）、普惠金融（中国平安）、视频感知（海康威视）、智能供应链（京东），图像感知（旷视）、安全大脑（360）、智慧教育（好未来）、智能家居（小米）。
- 科技部批准了十一个地方政府为国家新一代人工智能创新发展试验区，分别是北京、上海、合肥、杭州、深圳、天津、德清（唯一一个县域地方政府）、重庆、成都、西安、济南。
- 2018年10月和2020年1月，科技部发布“科技创新2030-新一代人工智能重大科技项目”指南，指南内容涵盖基础理论（深度学习、因果推理、博弈决策、群智涌现、混合增强智能、类脑智能等）、技术（知识计算、跨媒体分析、自适应感知等）以及应用。
- 国家自然科学基金委员会于2018年在信息学部设立了人工智能新的学科代码，包括人工智能基础、复杂理论与系统、机器学习、知识表示与处理、机器视觉、模式识别、自然语言处理、人工智能芯片与软硬件、智能系统与应用、新型和交叉人工智能、仿生智能（类脑机制）和人工智能安全等。
- 我国的重点领域：**互联网消费、自动驾驶、智能医疗和智能物联**



在图书馆的一些应用



南京大学：盘点机器人



清华大学：通借通还送书车



深圳宝安图书馆：智能分拣机器人1个小时可以分拣2000册书籍



宁波大学：学术头条

猜你喜欢

换一批

期刊文献 培养孩子爱上阅读方法初探

袁媛 儿童大世界 2期

图书 别等了，快做自己该做的事

周乔蒙编著 哈尔滨出版社 2019

期刊文献 雷蒙德·卡弗小说《大教堂》主题研究

孔艳艳 兰州教育学院学报 1期

学位论文 从接受美学角度解读莫里斯·葛雷兹曼的《往

李明珠 安徽大学 2017

图书 中华散文 萧红散文 珍藏版

萧红著 人民文学出版社 2016

期刊文献 有一种阅读叫严肃阅读

曾于里 视野 8期

序号	题名	作者	出版社	年份	价格
1	培养孩子爱上阅读方法初探	袁媛	儿童大世界	2019	2.00
2	别等了，快做自己该做的事	周乔蒙	哈尔滨出版社	2019	15.00
3	雷蒙德·卡弗小说《大教堂》主题研究	孔艳艳	兰州教育学院学报	2019	1.00
4	从接受美学角度解读莫里斯·葛雷兹曼的《往	李明珠	安徽大学	2017	1.00
5	中华散文 萧红散文 珍藏版	萧红	人民文学出版社	2016	1.00
6	有一种阅读叫严肃阅读	曾于里	视野	2018	1.00

- 万方的咨询机器人
- 在线咨询系统
- 人脸门禁

重大園的点滴应用



人工智能的**应用原则**

- 节约人工
- 服务提升
- 成本合理

第三单元：人工智能+知识服务

大数据-超强算力

一次文献+二次文献

Q

图书馆关心的问题？

- 1、数据馆藏，该如何建设馆藏数据资产？
- 2、门户安全，图书馆如何提高信息门户安全？
- 3、混合检索，图书馆如何实现不同文献载体的纸电合一检索？
- 4、智慧门户，如何提升信息门户的服务可见度？
- 5、学院图书馆，怎样给每个学院建立单独的虚拟学院图书馆？
- 6、成果管理，怎么管理学校的学术成果？
- 7、成果加工，该怎么自己加工学校的学术成果？
- 8、资源管理，该怎么管理数据库等电子资源？
- 9、智能报表，怎么提供智能生成馆员日常的统计数据报表？
- 10、读者运营，怎么生成读者的阅读分析报告？
- 11、数据发布，怎么大屏等多终端进行图书馆运行数据发布展示？
- 12、API标准接口，怎么杜绝不合理的接口费，建立数据馆藏API标准接口？
- 13、用户管理，怎么对图书馆用户进行分级分层管理？
- 14、统一认证，怎么对网络服务进行统一认证？
- 15、馆外访问，怎么支持网络服务的馆外访问？
- 16、数据汇编，怎么对数据馆藏进行二次组织汇编？
- 17、移动图书馆，如何为读者提供移动端的服务？
- 18、知识发现，如何实现浩如烟海的知识发现？
- 19、原文保障，怎么提升资源原文保障水平？
- 20、信息素养，怎么开展信息素养服务？

- 21、课程图书馆，怎么开展嵌入式教学服务？
- 22、个人图书馆，怎么提供一站式的读者服务？
- 23、读者运营，如何实现对用户的积分运营？
- 24、采集加工，怎么对自有资源进行数据采集加工？
- 25、统计分析，怎么对数据馆藏实施统计和分析？
- 26、修订屏蔽，怎么对错误、敏感数据进行修订和屏蔽？
- 27、访问权限管理，如何设置分级访问权限？
- 28、资源查重，在资源订购评估时，如何进行数据查重？
- 29、打破信息孤岛，如何开展数据集成管理？
- 30、智能PDA，如何开展资源的智能荐购？
- 31、自主离校，如何简化离校审批流程？
- 32、查收查引，如何开展智能查收查引？
- 33、查新服务，可以在线优化查新流程？
- 34、自助借还，怎么减少人工服务，支持自助借还？
- 35、身份识别，怎么快速甄别用户身份？
- 36、馆藏建设，如何优化馆藏评价、调节馆藏资源？
- 37、智能采访，如何优化采购决策和采购流程？
- 38、三级导航，如何实现库级、刊级、偏级的资源导航。
- 39、流程再造，分散的单个流程应用，能不能重构和集中整合？
- 40、系统更换，原来的自动化系统不好用，是否方便更换？

基于机器学习的高校图书馆用户偏好检索系统研究

📖 笔记

沈敏 杨新涯 王楷

重庆大学图书馆 重庆大学自动化学院

📄 导出/参考文献 🔗 分享 🔔 创建引文跟踪 ⭐ 收藏 🖨 打印 🖨 印刷版 ▼

摘要: [目的/意义]针对大数据环境下高校图书馆检索系统的信息超载问题,提出一种机器学习方法,为用户提供个性化的偏好检索服务。[方法/过程]通过提取用户检索行为大数据中的用户特征,利用监督机器学习方法,在线学习可随用户偏好同步变化的自适应检索模型,预测用户对文献的选择概率,优化检索结果的排序。[结果/结论]设计出用户偏好检索原型系统,介绍用户偏好检索系统工作流程,对比分析系统效果,并对系统进行客观评价。

基金: 国家社会科学基金项目“智慧图书馆理论与系统实践研究”(项目编号:13XTQ009)研究成果之一;

关键词: 机器学习; 高校图书馆; 用户偏好; 偏好检索;

DOI: 10.13266/j.issn.0252-3116.2015.11.020

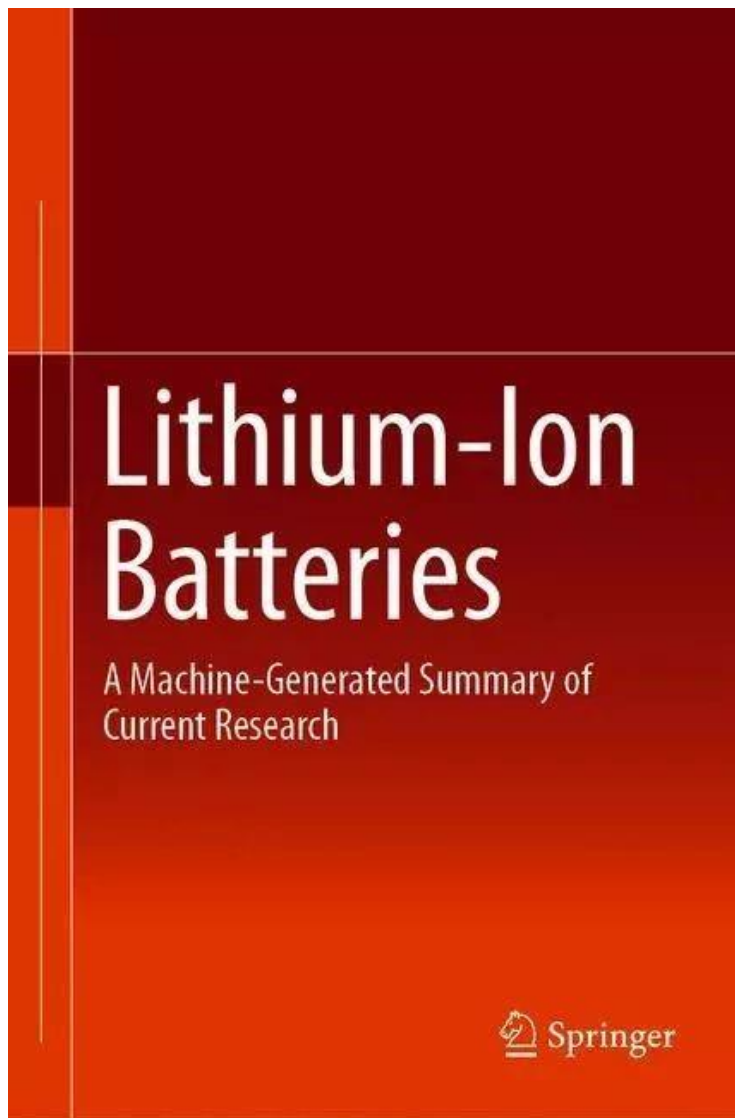
分类号: G258.6;G252.7



图书情报工作

Library and Information
Service

2015年11期

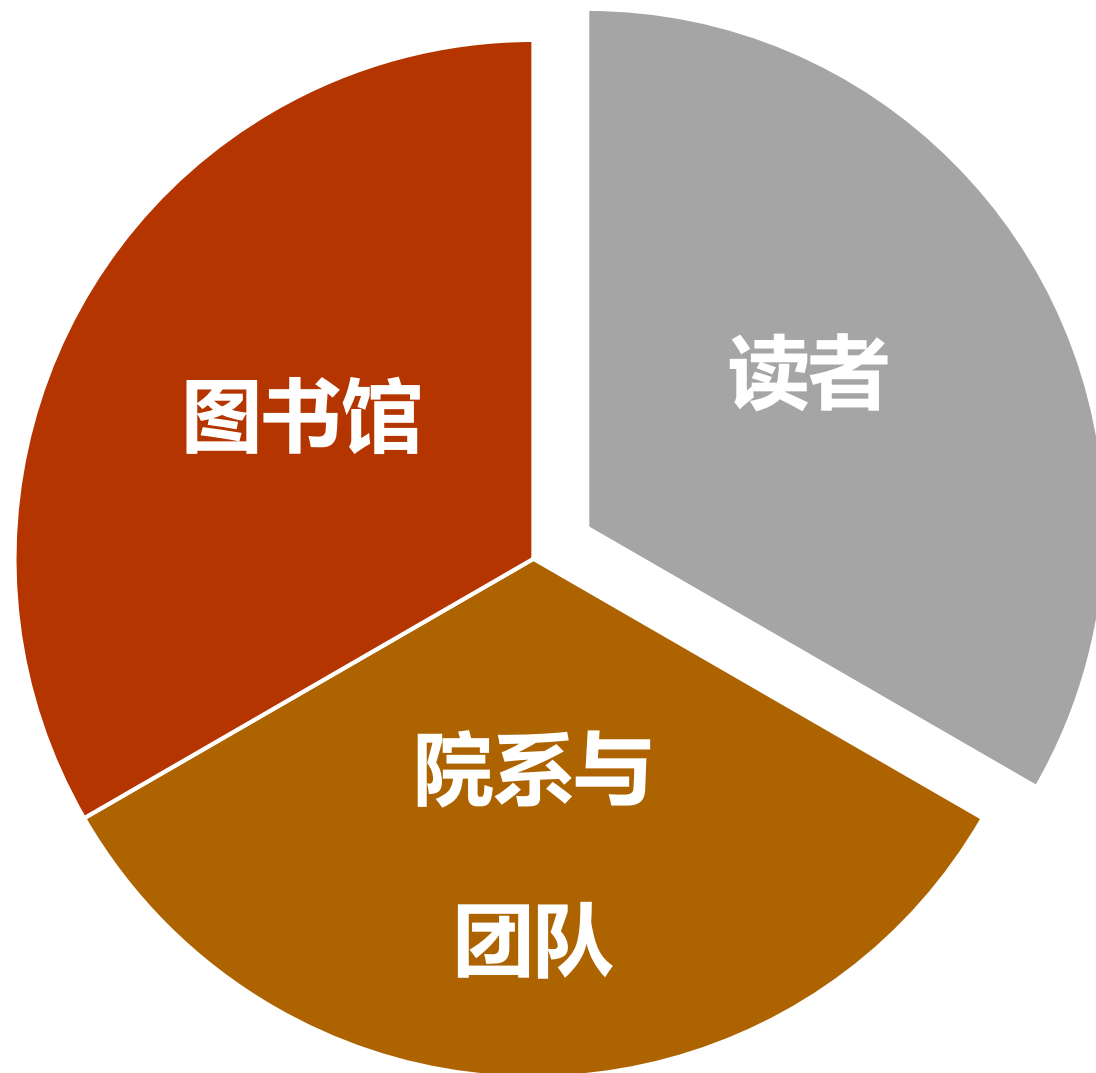


《锂离子电池：机器生成的当前研究摘要》

由机器学习算法总结了2016年到2018年的53000多篇相关论文，找出150多篇权威并将其摘要进行集合，帮助读者进一步阅读原始的文章。

用户驱动的知识服务体系

- 全文数据化
- 数据处理：自动编目、自动标引
- 学科文献支撑
-



读者

研究型读者

- 文献综述
- 项目选题
- 科研动态推送
- 投稿智能引导
- 成果推荐

学习型读者

- 用户偏好的搜索
- 精准和个性化的文献资源导读

院系与团队

院系决策者

- 研究团队发展态势与影响力分析
- 全球研究趋势分析

学科与团队负责人

- 学术热点/冷点分析
- 学科前沿分析
- 知识产权导航与分析

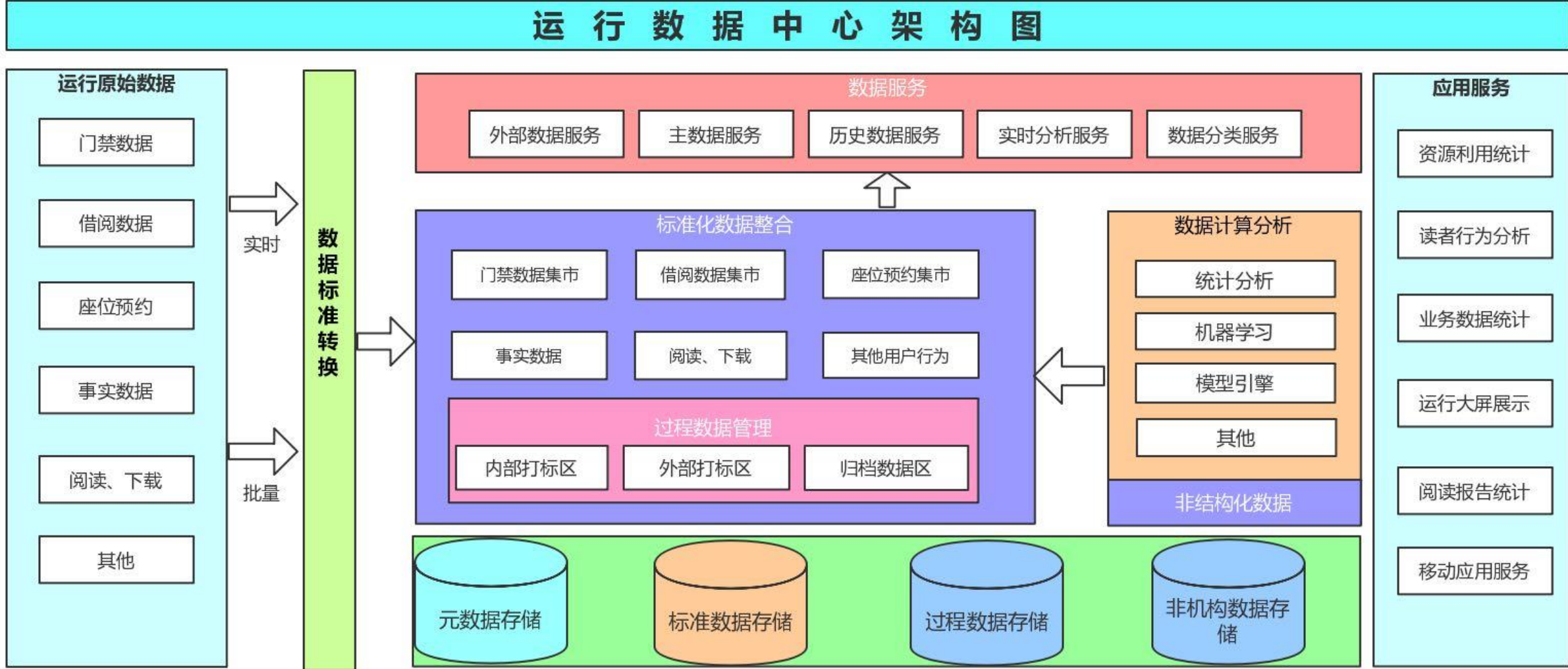
第四单元：三种手段

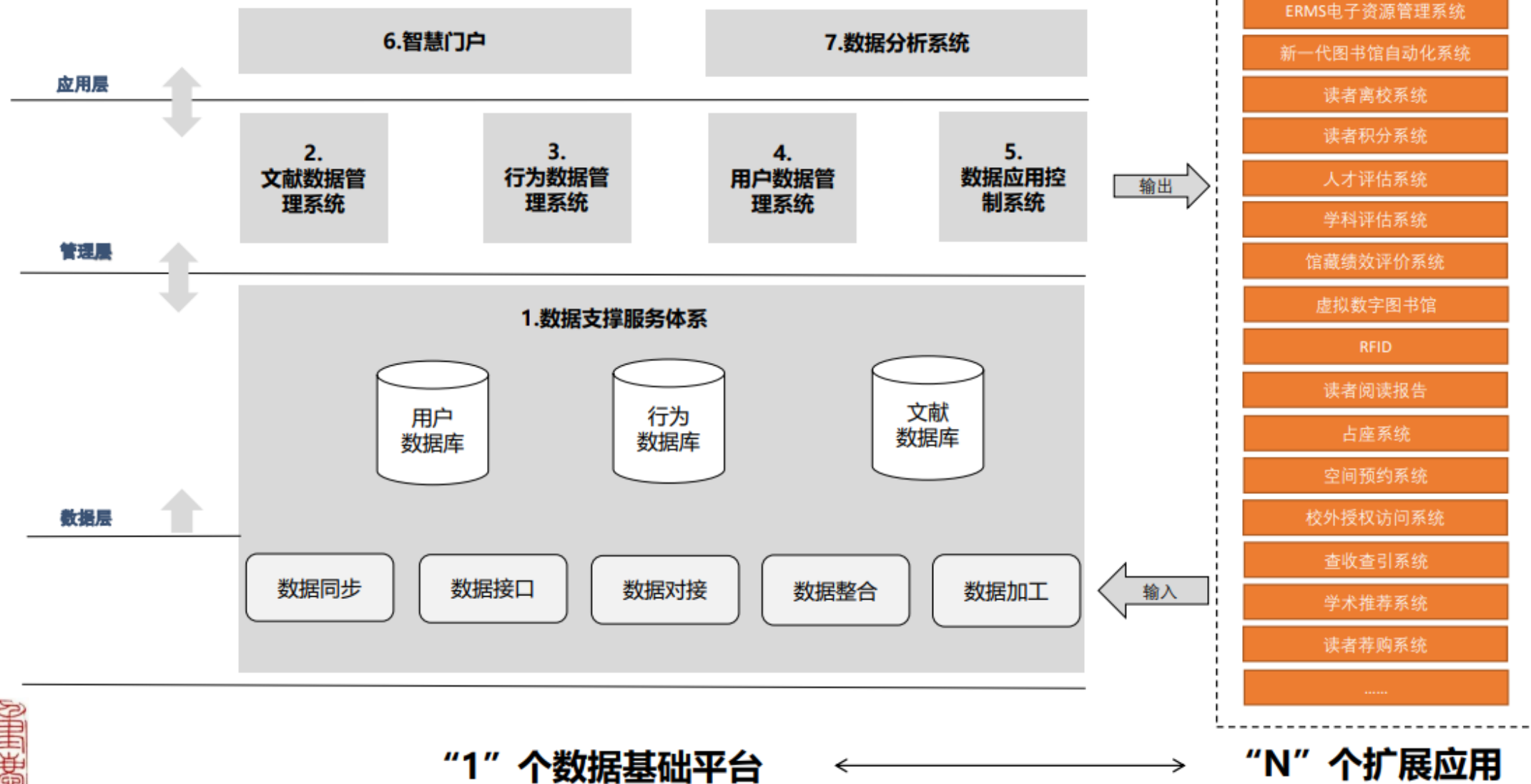
数据平台-机构库-馆员

1 数据及其平台

- 图书馆数据很多，但没有大数据
- 数据属于图书馆
- 图书馆对任何数据都要**极度**重视，以优化读者服务和业务管理水平

像管理图书一样管理数据！





策略一：用数据中心升级资源中心，建设“数据生产体系”。

Data

事实数据

- 图书馆事实数据、分类、大学排名（QS、TIMES、武苏连...）、分类（中图、教育部学科分类、EIS分类、WOS分类...）、高校名录、双一流高校名录、自然科学基金、社会科学基金...

要素：复杂性、多样性

用户行为数据

- 纸本借阅、电子资源使用（访问、检索、浏览、下载）行为、系统使用记录
- 按时间、按类型、按院系、按地区...

要素：复杂性、完整性

学术数据

- 电子资源、纸本资源、自建资源、OA资源
- 数据库厂商名录、数据库名录、数据库收录清单、中外文期刊总表、IF、JCR/SJR/SNIP、SCI/CSCD/CSSCI

要素：获取、时效、完整、机构化、体系化

用户数据

- 学生身份信息、教研人员身份信息、决策管理人员身份信息、学科带头人、两院院士、人文社科资深教授、杰青基金、长江学者、千人计划、青年海外高层次人才引进计划、新世纪优秀人才支持计划、高等学校教学指导委员会、优秀青年教师支助计划...

要素：来源多样性、数据完整性欠缺

业务数据

- 机构知识库、教务系统、科研系统、人事系统、门禁系统、空间预约系统、大屏展示系统

要素：复杂性、整体结构化艰难

去重后的统计结果 >

175,307,812 (篇)

93,444(种)期刊资产总量

去重后的统计结果 >

8,252,877 (篇)

学位论文资产总量

去重后的统计结果 >

8,173,324 (种)

图书资产总量

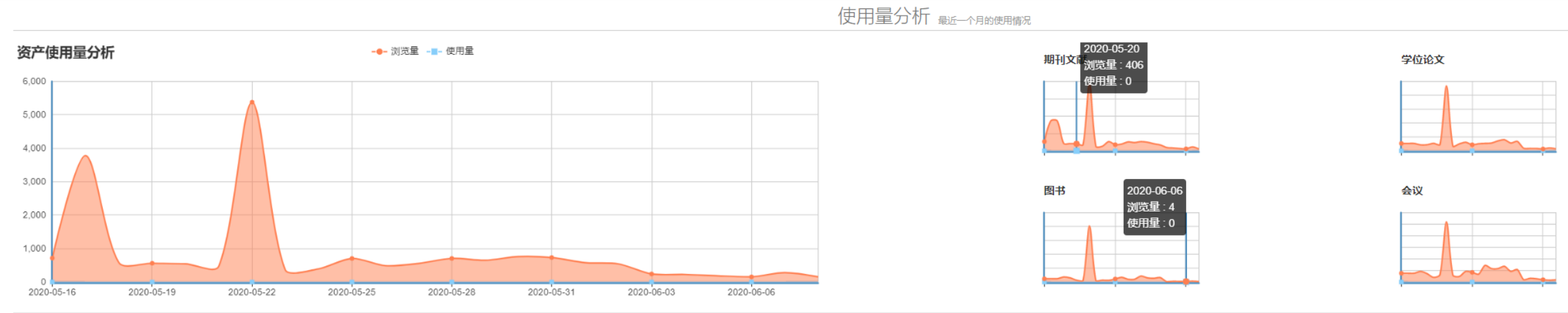
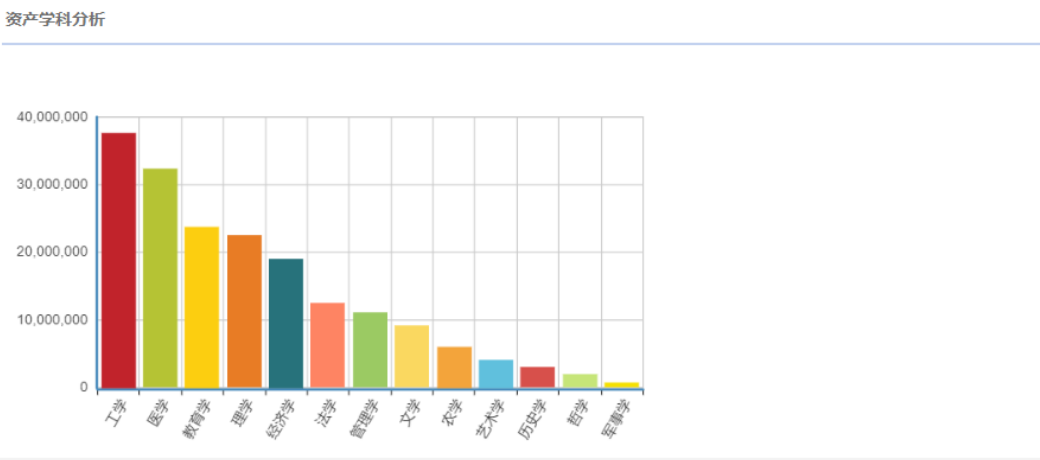
去重后的统计结果 >

890,104 (篇)

标准资产总量

最新数据批次

类型	供应商产品	批号	数据量	添加日期
论文	DDS学位论文	2020061501	528348	2020-06-15
期刊文献	T&F期刊	2020061501	7126	2020-06-15
期刊文献	美国化学学会 (ACS) 期刊	2020061501	92	2020-06-15
期刊文献	Emerald全文期刊库	2020061501	1200	2020-06-15
期刊文献	IEEE期刊	2020061501	678	2020-06-15
期刊文献	万方期刊数据库	2020061501	49847	2020-06-15
期刊文献	台湾华艺期刊	2020061501	3103	2020-06-15

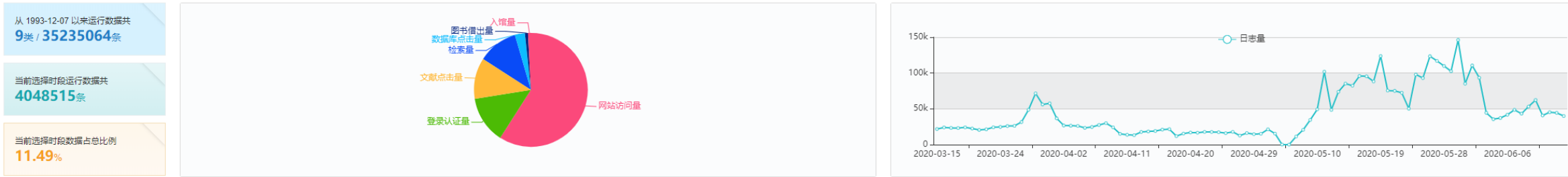


最近30天

最近3个月

今年

自定义时间



导出

按天

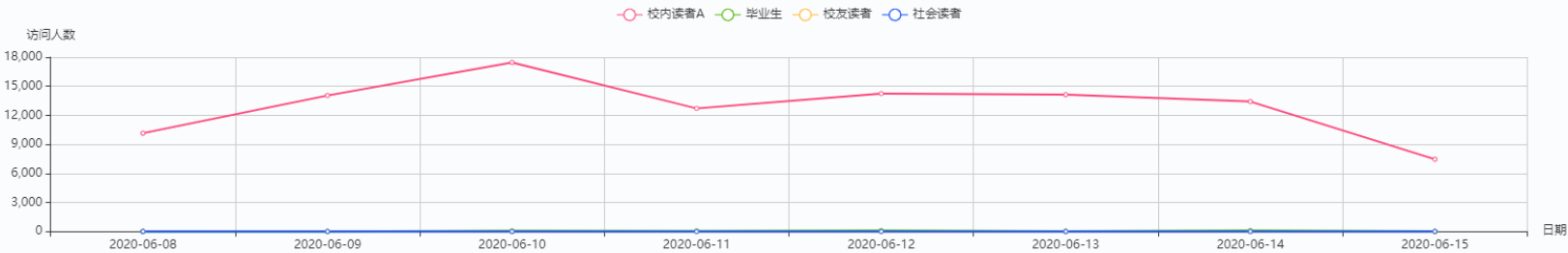
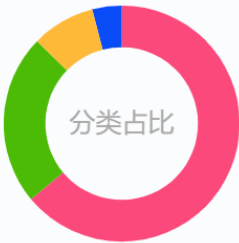
按月

按年

日期	图书借出量	文献点击量	数据库点击量	检索量	入馆量	登录认证量	网站访问量	总量
2020-06-14	0	1941	553	4643	0	5967	26860	39964
2020-06-13	0	2374	653	7208	0	6049	28096	44380
2020-06-12	0	2035	620	9858	0	6033	26741	45287
2020-06-11	0	1747	379	7111	0	5596	25892	40725
2020-06-10	0	1620	424	12664	0	7790	39752	62250
2020-06-09	0	1073	422	17446	0	6209	28007	53157
2020-06-08	376	1062	211	14790	0	4524	22302	43265
2020-06-07	0	1999	350	17058	0	4991	24174	48572
2020-06-06	0	1434	334	13056	0	4842	22098	41764
2020-06-05	0	1181	317	7675	0	5306	22728	37207
2020-06-04	0	1188	338	7300	0	4698	22035	35559
2020-06-03	0	1480	632	7648	0	6972	27715	44447

读者分类统计

● 校内读者A: 50,506
● 毕业生: 18,469
● 校友读者: 6,886
● 社会读者: 3,153



7天活跃用户数

0

占总数的百分比: 0.00%(79573)

30天活跃用户数

13660

占总数的百分比: 17.17%(79573)

90天活跃用户数

20322

占总数的百分比: 25.54%(79573)

365天活跃用户数

48550

占总数的百分比: 61.01%(79573)

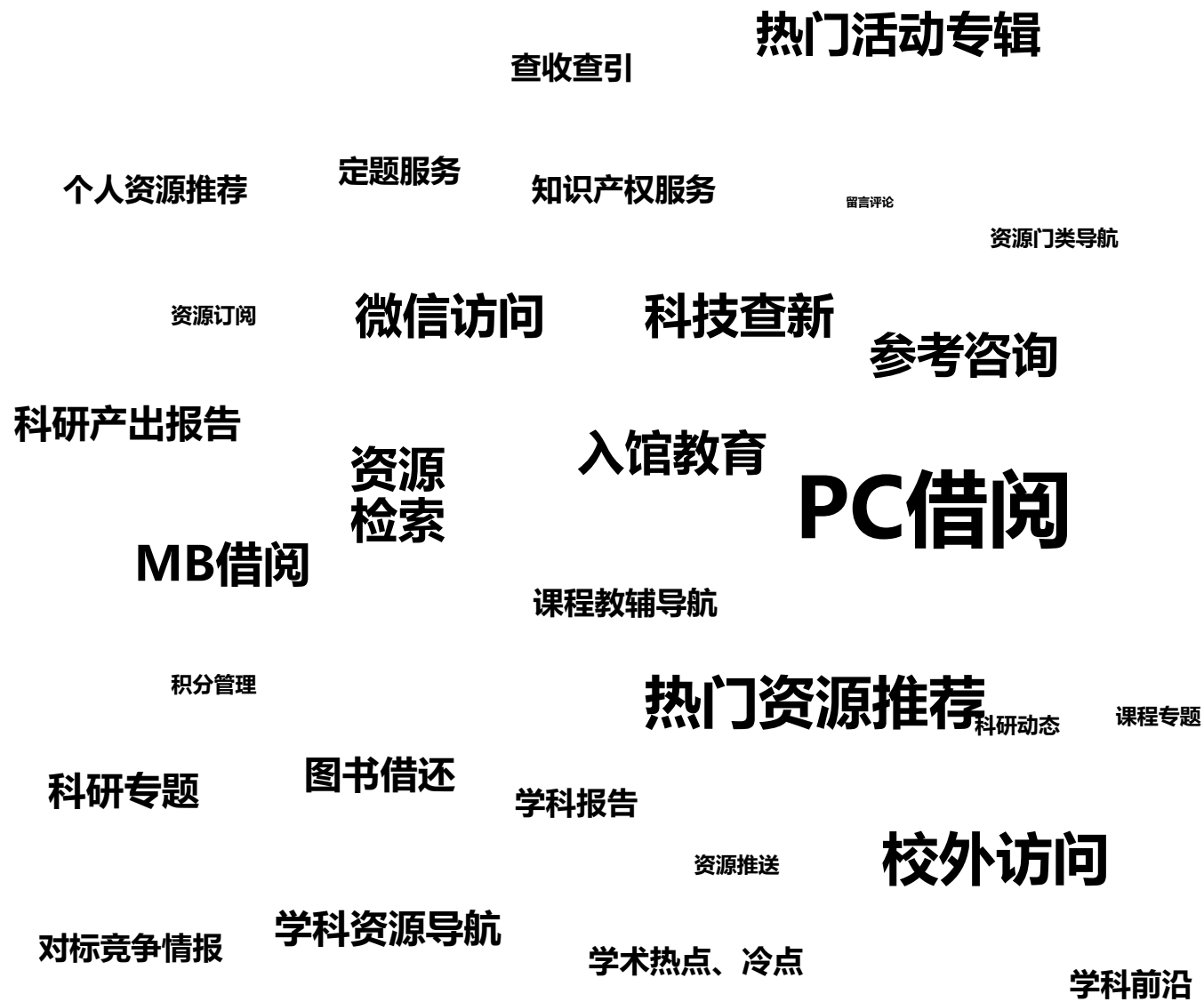
最近7天活跃用户排名

用户名	访问次数	最后登录时间
夏明	300	2020-06-15 11:15:43
黄应平	140	2020-06-15 07:24:23
税艳玲	118	2020-06-15 09:28:24
马建树	111	2020-06-14 07:24:34
王睿	95	2020-06-13 11:49:05

最近更新读者

用户名	部门/院系	更新时间
冉波	图书馆	2020-06-15 11:10:29
孙晓冬	公共管理学院	2020-06-15 10:24:56
高天勇	化学化工学院	2020-06-15 09:51:52
李俊萌	材料科学与工程学院	2020-06-15 09:40:49
宋凯胜	化学化工学院	2020-06-15 08:58:05

策略二：场景化



策略三：用主动运营替代被动服务，建设“用户运营体系”。

身份画像

专科生、本科生、研究生、博士生、教师、行政人员、科研人员

用户画像

正式读者、社会读者、特殊读者、校友读者

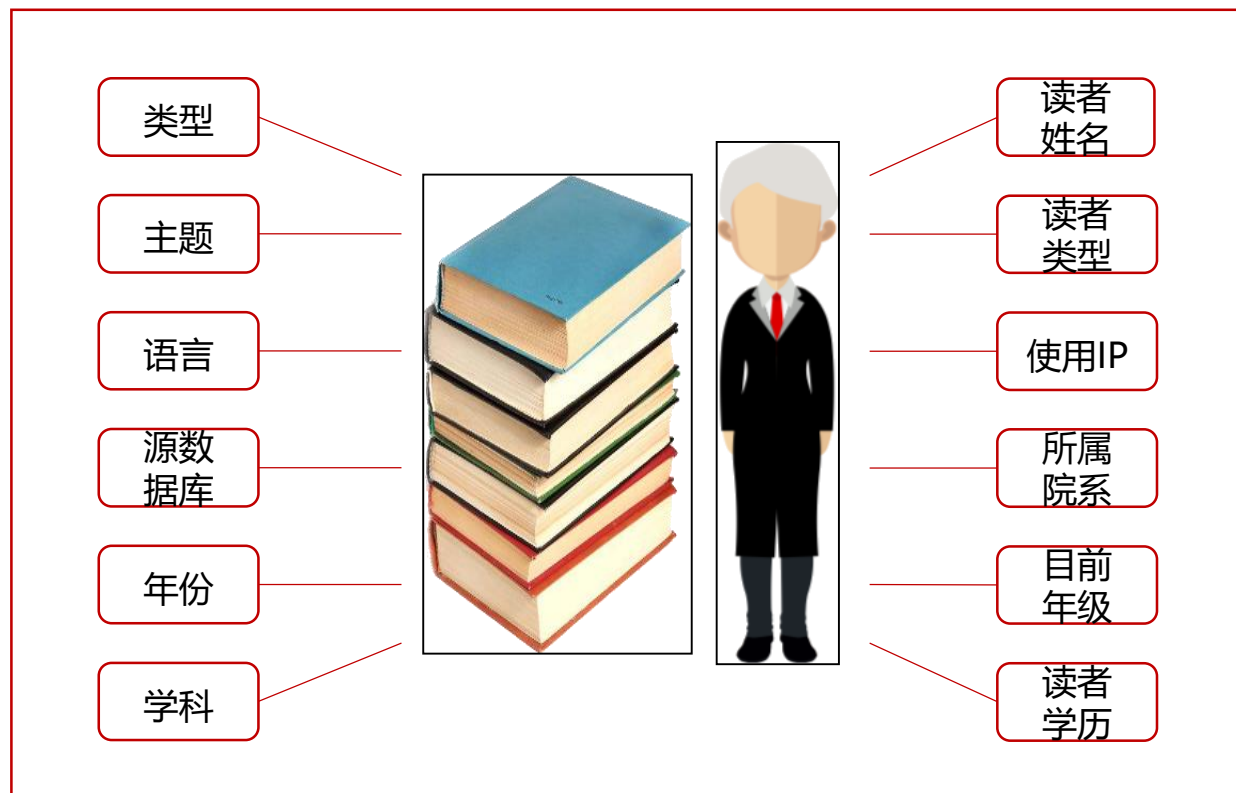
职称画像

助教、讲师、副教授、教授、实习研究员、助理研究员、副研究员、研究员

职务画像

校长、副校长、学科带头人、院长、副院长、处长（主任、部长）、副处长（副主任、副处长）

部门画像、需求画像.....



2 馆员-知识服务能力

科研评价

- 学科发展评价
- 科研成果认定
- 重点人才评价
- 人才引进前期评估
- 人才引进后期评价

科研规划与学科规划

- 学科发展态势
- 学科研究热点和前沿
- 学科学术影响力
- 学科发展趋势
- 潜在重点学科

- 信息技术能力
- 需求设计能力
-

研究能力

专业研究能力

智慧图书馆协同创新中心

图情研究所

学术评价与分析中心

开放学术论坛

科研奖励

基本业务科研经费

馆内管理与服务创新项目

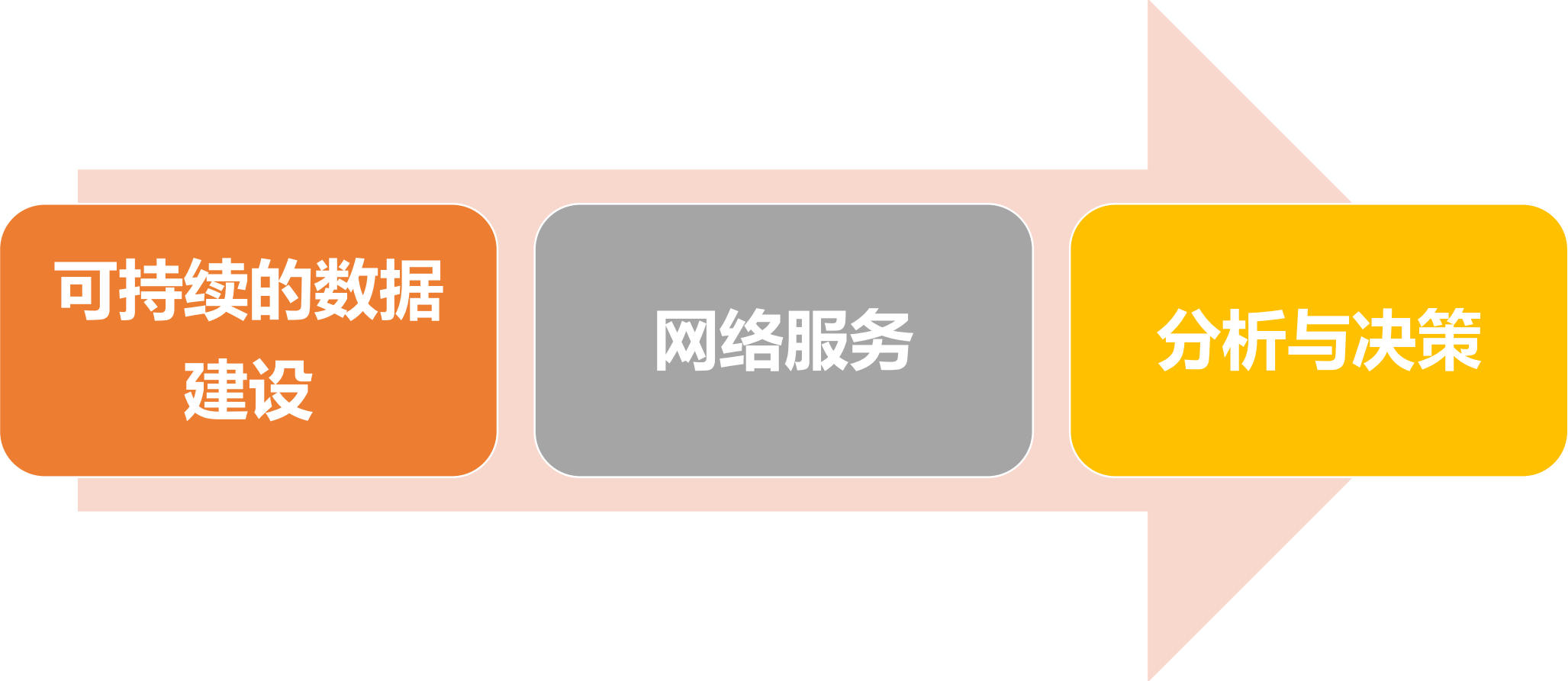
知识产权信息服务中心

- 研究馆员
- 学术型机构



3 机构知识库与相关方法论

主要的关注点



**可持续的数据
建设**

网络服务

分析与决策





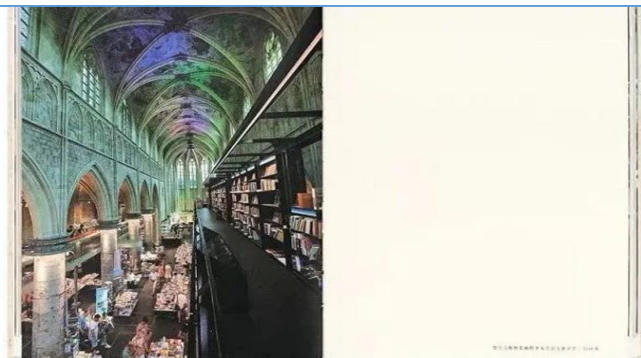
微信号: bookinages



因书而在 Notebook



手机天猫扫一扫



微信号: bookinages



有书而美 Notebook



重庆大学

谢谢!

QQ: 40018777

E-mail: xinya@cqu.edu.cn