

2019'高校图书馆发展论坛

从数据驱动到知识导引 -新一代图书馆探索实践

刘柏嵩@大连
宁波大学，2019.6.6

汇报提纲

- 一、进展与挑战
- 二、探索与实践
- 三、展望与未来

一.进展与挑战

未来已至：人类已经进入智能时代

- 大数据的日益积累、计算能力的快速增长为人类进入智能时代奠定了基础；
- 大数据为智能技术的发展带来了前所未有的数据红利；
- 机器计算智能、感知智能达到甚至超越人类。

2012年，在**图像识别**的国际大赛ILSVRC(大型视觉辨识挑战竞赛)中，加拿大多伦多大学的研究团队基于深度卷积神经网络的模型[1]夺冠，把TOP5错误率降到15.3，领先第二名超过十个百分比，震惊学术圈。

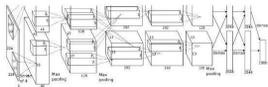


Figure 2. An illustration of the architecture of a CNN, depicting the definition of receptive fields between the two GPUs. One GPU runs the layer maps (a) in the top of the figure while the other runs the layer maps (b) in the bottom. The GPU communication is performed across the network. The network inputs is 224x224x3, and the number of outputs in the network's remaining layers is given by 251,440-184,320, 61,496-44,800, 43,264-4096, 4096-1000.

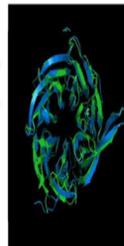


2016年，Google全资收购的DeepMind推出名为AlphaGo 的围棋程序[2]，以4:1的总比分击败世界顶级职业围棋选手李世石，让全世界开始关注人工智能技术巨大的应用前景。

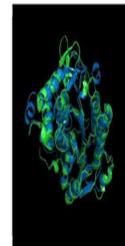


2018年“阿尔法折叠”，2018年年底，谷歌在科学杂志上发布了“阿尔法折叠”的研究成果，它可以根据DNA的序列测出一个蛋白质的三维结构。

T0954 / 6CVZ



T0965 / 6D2V

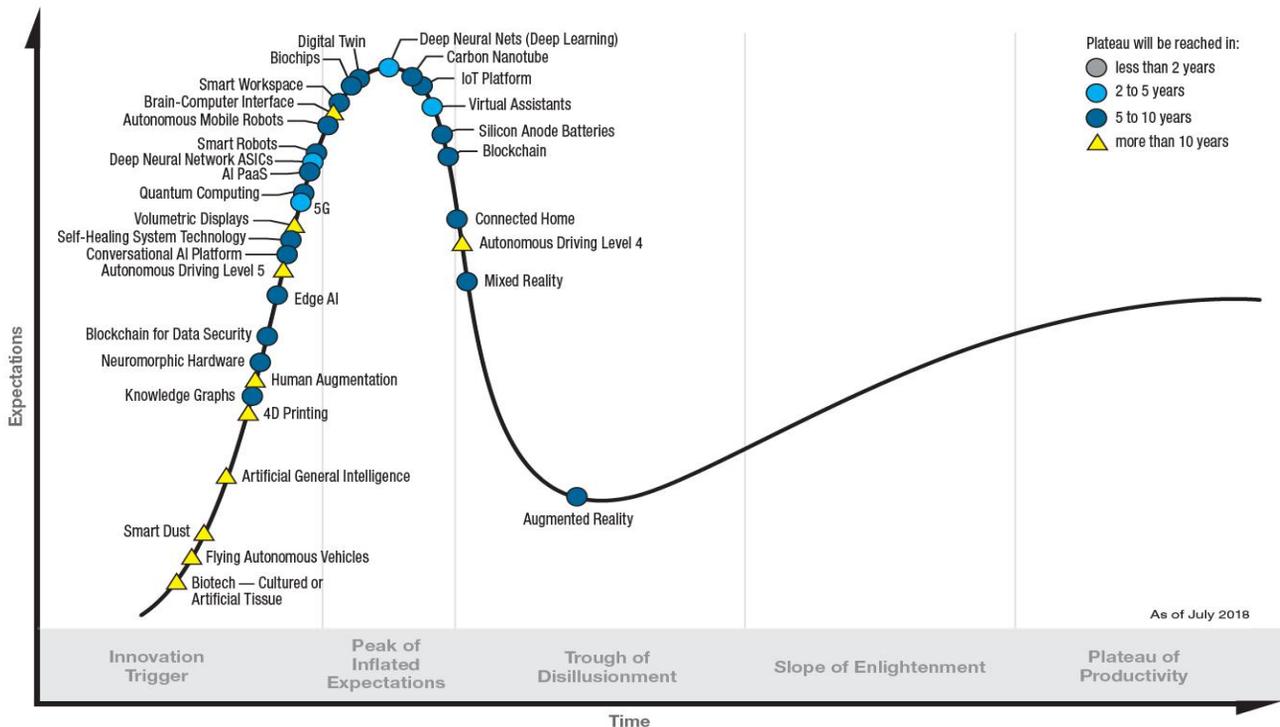


T0955 / 5W9F



Structures:
Ground truth (green)
Predicted (blue)

Hype Cycle for Emerging Technologies, 2018



处在期望膨胀期：深度学习、智能机器人、智能工作空间、虚拟助手、自动移动机器人

Gartner®

AI四个重要时间节点

- 第一代（1956-1970）模仿人
- 第二代（1971-2000）数据驱动
- 第三代（2001-2016）大数据+深度学习
- 第四代（2016-至今）超级人工智能
 - IA：智能增强，通过数据来增强人的创造和智能
 - II：智能基础设施，象网约车和餐馆推荐系统
 - 2018年11月美国商务部工业与安全局制定出台《特定新兴技术管制审视》政策，人工智能（AI）和机器学习技术，也列其中

吴军，《超级智能时代的两个关键因素》CNCC2018

带来智能化升级与转型

- 智能化升级与转型已经成为各行各业的普遍诉求
- 从信息化走向智能化是必然趋势
- AI+成为AI赋能传统行业的基本模式
- 战略意义
 - ▣ 全方位、深度渗透到各行各业、各个环节
 - ▣ 颠覆性影响，重塑行业形态，甚至社会形态



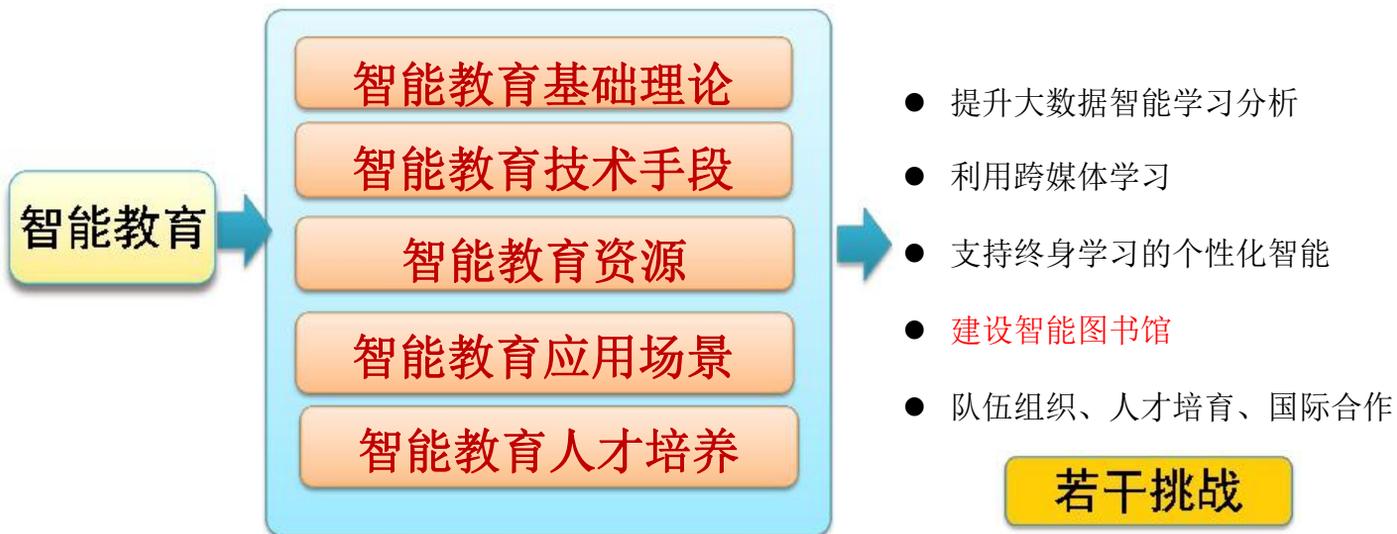
《中国教育现代化2035》

🏛️ 2019年2月国务院发布，要加快信息化时代教育变革。

建设**智能化校园**，统筹建设一体化智能化教学、管理与服务平台。

🏛️ 利用现代技术加快推动人才培养模式改革，实现规模化教育与个性化培养的有机结合。创新教育服务业态，建立数字教育资源共建共享机制，完善利益分配机制、知识产权保护制度和新型教育服务监管制度

智能图书馆是智能教育的重要组成



智慧图书馆的提出



——刘兹恒，北京大学信息管理系

对照信息新时代的几个标准

🏠信息融入一切

🏠网络互连一切

🏠数据表征一切

🏠智能灵化一切

我们首先需要问自己4个问题

- 我的图书馆能否最大限度地实现**网络化**？
- 我的图书馆能否尽可能地引入机器的**学习效应**？
- 我的图书馆能不能在服务师生与资源建设过程中，尽可能地利用**机器决策**取代人工决策？
- 最后一个问题，我的图书馆能否全面收集各类**数据**与更多不同类型的数据产生交换？

目前还是概念化阶段

🏠 概念很美好

🏠 现实很骨感

➤ 业务互联程度不高：

- 数据商应用商百花齐放，数据标准并不统一

➤ 管理与服务效率不高

➤ 用户体验并不便利

从当前人工智能三要素来分析



A 算法

B 数据

C 算力



智慧化路径的可能选择

- ▣ 大数据涌现和计算能力大幅提升的趋势
- ▣ 图书馆如何顺应人工智能带来的红利？

赋能路径：从数字化-》数据化-》
智能化-》智慧化



二、探索与实践



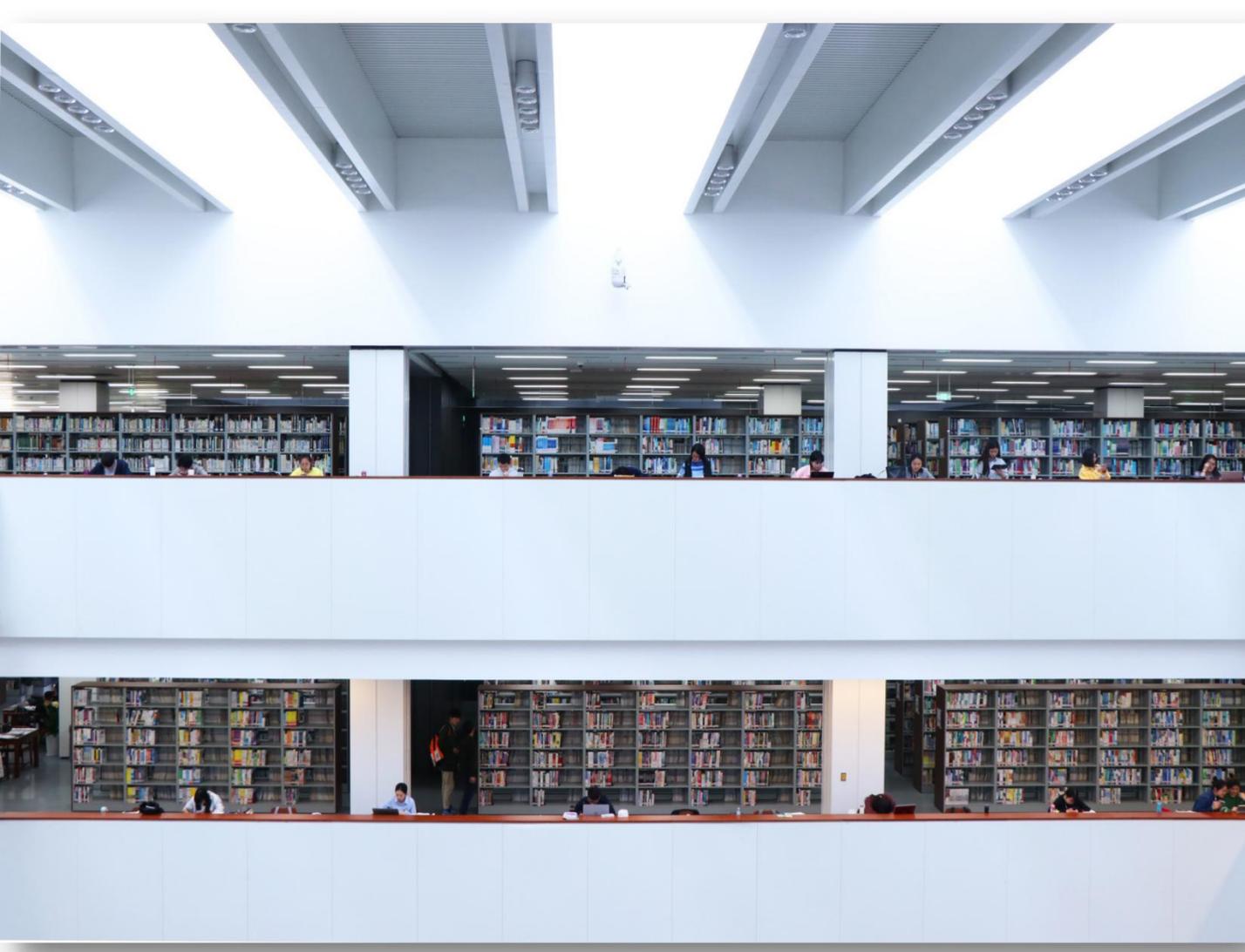
真诚图书馆

沈从文



梅山分馆





当前是数据驱动的AI

🏛️在DT时代，图书馆普遍缺乏话语权，能否在AI时代崛起？

🏛️图书馆的目标是什么？

🏛️智慧图书馆：数据智能+网络协同

🏛️如何表现？特征是什么？

🏛️馆员能有哪些作为？

数据从四个维度产生作用

- ▣ **决策**：通过数据让我们更好理性决策，不是基于经验，而是通过数据（校级、部门级）；
- ▣ **优化**：通过数据分析，进行服务的优化；
- ▣ **推广**：基于读者的个性化需求，进行精准推荐与阅读推广（读者画像）；
- ▣ **创新**：通过数据的方式，形成新的**服务**，新的业务增长点（对学习、科研和管理方面）。

来自三类主要服务

🏠 流程服务

- 借阅、采购、预约

🏠 协作服务

- 入馆教育、行为数据分析、科技查新

🏠 咨询服务

- 空间服务、智能问答、个性化推荐

增加与读者接触点

📖 记录每一位读者的行为数据

📖 到馆的服务

✓好的借阅环境只是基础

✓读者感知、互动

📖 不到馆的服务

✓接触点分布在网站和移动APP中

➤ 支付宝：公众号+余额宝+快车

✓如何更多更方便地收集到读者的兴趣

✓如何让读者参与服务中

加强对读者的洞察

📖 读者的交互数据（线上线下）都会产生很多的数字足迹

- 他们什么时候开始搜索论文、图书关键词
- 他们的兴趣爱好是什么
- 他们喜欢电子还是纸质图书
- 他们喜欢坐在哪里阅读
- 他们经常预约研讨间否
- 除了他自己的专业资源外，还爱看哪些资源

📖 将各类数据进行关联

没有数据体系支撑的都不是新一代系统；都不是智慧图书馆

智慧型图书馆建设核心是图书馆**数据能力**的构建，**掌握数据**才能具备有资源、懂读者、知运行的服务能力。



初步形成图书馆智能应用体系

📖 图书馆应用体系:

- 📖 流程服务（借还书及预约）
- 📖 协作服务（学科查新）
- 📖 咨询服务（馆内、校内百事通）

📖 数据赋能中台:

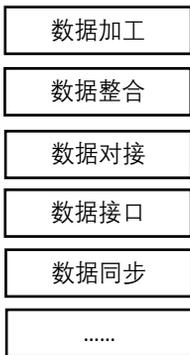
- 📖 本校相关学科学术态势感知
- 📖 资源采购与动态分析平台
- 📖 基于学生学习成长的数据支撑平台:数据模型分析, 数据标签处理

2.1 业务管理决策数据的多元、立体化

建立图书馆数据支撑服务体系

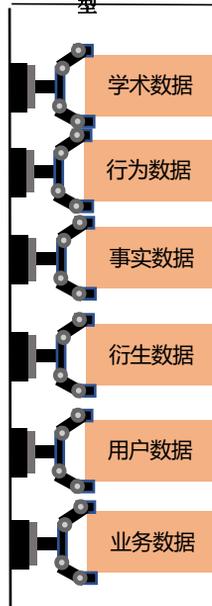
以“数据”为对象，提供基于数据的“藏用一体”的服务体系。

数据支撑服务体系

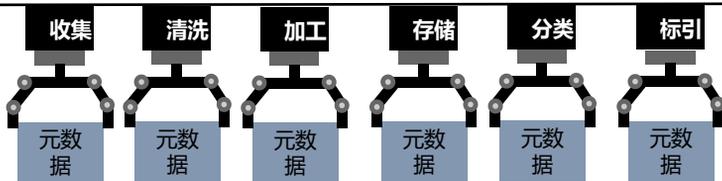


藏用

数据加工类型



数据加工流程



元数据

元数据处理流水线

摸清资源家底

数字资产管理 » 主面板



去重后的统计结果 >

162,177,632
(篇)

61,173(种)期刊资产总量



去重后的统计结果 >

4,450,729 (篇)

学位论文资产总量



去重后的统计结果 >

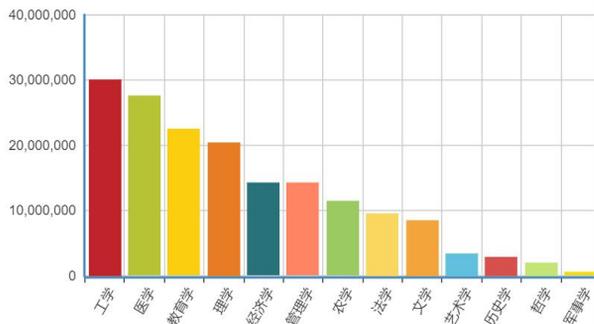
2,370,512 (种)

图书资产总量

最新数据批次

类型	供应商产品	批号	数据量	添加日期
期刊文献	维普中文期刊服务平台	2018111601	72484	2018-11-16
期刊文献	JSTOR西文过刊全文库	2018111601	530104	2018-11-16
期刊文献	IOP英国物理学会	2018111601	5087	2018-11-16
期刊文献	Elsevier ScienceDirect期刊	ELS2018...	4600	2018-11-16
期刊文献	CNKI期刊学位会议报纸	CNKI201...	51118	2018-11-16
期刊文献	ASCE土木工程	2018111301	16	2018-11-13
期刊文献	APS美国物理学会(NSTL资助)	2018111301	385	2018-11-13

资产学科分析



文献数据管理系统



文献类型覆盖图书、期刊、学位论文、专利、标准、样本、音视频、资讯等15大类型



数据来源，学科分类，更新批次等多维度的资源查看和管理，最小管理粒度可以到篇级。



提供查重、核对、权限等辅助工具，在智能算法技术的帮助下，轻松管理上亿规模资源。

期刊详细分析

期刊详细分析

期刊分析

学科分析

序号	刊名	ISSN	出版类型	出版单位
1	红专专	1009-8569	半月刊	南京出版社
2	初中生英语	1674-375X	旬刊	湖北楚天教育文化新闻出版
3	红专专	1671-3484	月刊	浙江经济和信息化委员会
4	江苏教育·教育管理	1005-6009	月刊	江苏教育期刊社
5	红专专·教育	1673-9442	月刊	黑龙江普通文字期刊社
6	红专专·教育	1671-3141	旬刊	世界图书出版公司·中国期刊协会
7	红专专	2095-2627	周刊	西藏高原区报刊出版中心
8	红专专·教育	2096-4110	旬刊	黑龙江文化产业促进会
9	红专专	1671-220X	半月刊	宁波日报报业集团
10	红专专	1674-1900	月刊	北京教育行业协会
11	红专专·教育		月刊	杭州市委办公厅
12	红专专·教育	1002-6932	旬刊	贵州教育期刊社有限公司
13	红专专·教育	1002-0330	月刊	江苏少年儿童出版社
14	红专专	2096-1154	月刊	光明日报出版社
15	红专专·教育	2096-1995	旬刊	山东教育期刊社
16	红专专·教育	1004-2377	旬刊	中国和平出版社有限责任公司
17	红专专	1005-9806	双月刊	中国农业出版社有限责任公司
18	红专专·教育	2095-3968	旬刊	山东人民出版社
19	红专专	1003-5484	月刊	河南科学技术协会
20	红专专·教育	1009-4091	月刊	现代教育行 编辑部

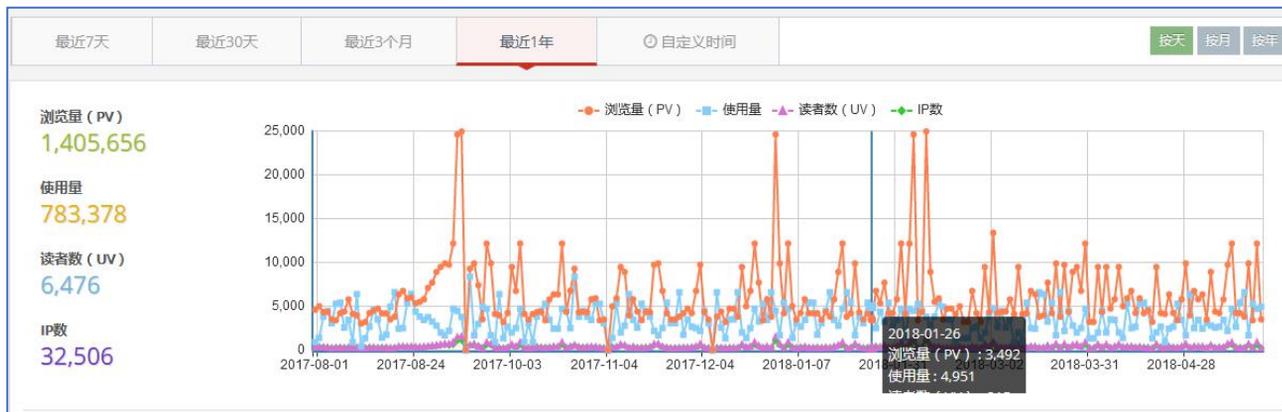
共2,614条 共135页

★ 期刊详细信息

封面

刊名 江苏教育：教育管理
出版类型 月刊
主办单位 江苏教育期刊社
ISSN 1005-6009
统一刊号 32-1410
语言 中文

实现资源的管理决策



读者：谁在用？都用了些什么资源？

资源：都被谁用了？哪些用得好，哪些用得不好？

情报决策服务



馆藏评价

按照不同学科纸+电资源，根据不同纬度和指标分析学科资源支撑情况。



根据重点学科资源需求，建立科资源导航门户，平台数据驱动资源导服务，自动更新，提升学科数据服务的生态性。



学科决策

系统梳理137个双一流学校学术成果情况，学科评估报告生成平台、学科评估数据处理平台。



大数据中心作为数据基础，提供机构词表、学者词表工具，通过大数据中心提取结构化机构成果数据，同时提供数据优化处理平台。



参考咨询

面对教学、科研、学科服务的数据资源汇编，提升核心资源展示度和使用率。



基于海量数据平台，提供一站式检索。综合展示纸电资源，分发服务，提供多纬聚类，提升便捷性。



不仅能为用户提供投稿推荐，还可以为高校提供学科建设引导性投稿推荐



通过大数据平台智能获取收录情况及引用情况，实时通知，报告生成……



学院图书馆

学院专业资源聚类、学院馆藏资源分析、学院查查索引服……

数据驱动采访与决策

- 建立图书馆“数据集市”
 - 实时看到整体运行状况和用户需求
- 馆藏政策：纸质书借阅率，刊的保障率
 - 自动统计分析本校的专业和学科
 - 自动匹配、生成纸质与数字资源采访列表
- 决策分析
 - 📊 空间布局：读者的学习习惯，粉丝的需求
 - 📊 与学业成绩的关联；与科研产出的关联
 - 对学科（团队）的数据分析

管理数据——馆藏资源评价分析



期刊整体数量

期刊、图书等各种资源类型的总体数量。



期刊整体质量

JCR、ESI、SNIP、SJR来源期刊馆藏保障率，JCR Q1、Q2、Q3、Q4各分区馆藏保障率，学校导向期刊保障率。



期刊整体引用情况

高引用量期刊馆藏保障情况。



数据库收录分析

数据库资源收录情况。



学科期刊保障

ESI学科期刊保障率、JCR学科保障率、教育部学科保障率



SCI收录期刊馆藏保障情况

每年SCI收录期刊馆藏纸电保障情况



类似机构资源保障率对标

对类似高校期刊建设情况对标



本校学术成果引用馆藏分析

分析学术产出者使用资源建设情况

馆藏资源评价分析

- 资源评价从理论走向实践

缺少数据获取途径

- 从单一指标走向多元指标

点击量、下载量——保障率、影响力、作者权重

图书馆运行数据的管理

从单一的日志数据分析，扩展到图书馆全面运行数据管理。整合多种类型、来源的运行数据集成。

运行数据管理系统

数据来源管理

运行标准管理

事实数据管理

行为数据统计

.....

运行元数据：统一集成&编辑



门户/移动



流通借还



厂商日志



门禁数据



空间设备

专业的运行元数据：规范&管理

多维度：借阅行为、下载行为、浏览行为、收藏行为、评价行为、到馆行为.....

全方位：人事数据、空间数据、设备数据.....

多标准：页面记录 Open API SUSHI COUNTER

ORM

多角度：门禁数据、厂商日志、流通借还、页面记录、空间设备.....

图书馆运行数据分析



2.2 服务方式自动化与个性化

图书馆用户数据的管理

对用户数据的分级管理与控制，
可实现全面了解读者信息，打
通与读者之前的互动沟通桥梁。

用户 数据 管理 系统

读者数据管理

智能消息管理

读者兴趣画像

统一认证授权

社交帐号绑定

读者标准管理

● 用户数据类别



基础数据



身份数
据



属性数
据

● 用户身份认证 用户分组设定 用户单点登录 用户权限

设置 用户数据管理 接 ILS 读者数据开放

API

同步一卡
通



提供单点登
陆

CAS

关联社交帐
号

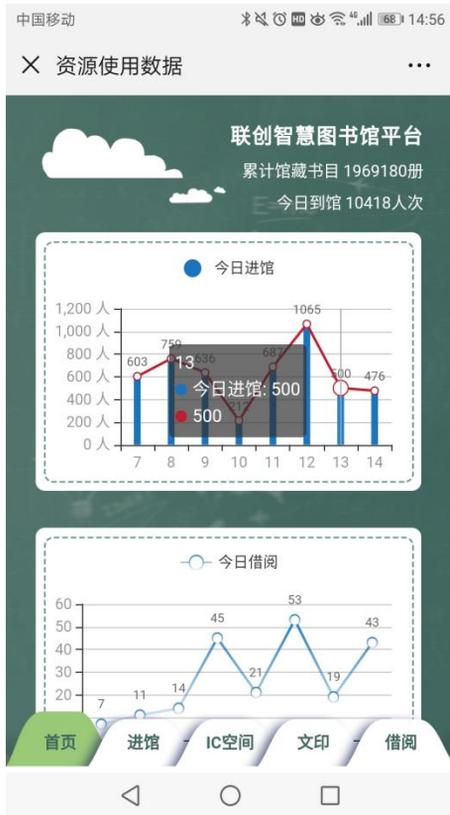


第三方登
录

Open ID

用户数据分析自动化





阅读报告的个性化



用户行为全面记录（最近7天）

	读者	类型	院系	浏览
1	刘燕燕	高级	未登录访问	88
2	陈德锋	一般	阳明学院	117
3	王一峰	一般	阳明学院	15
4	刘晓杰	研究生	机械工程与力学学院	166
5	罗佳峰	研究生	海洋学院	22
6	杨璐	一般	未登录访问	48
7	潘彬	高级	未登录访问	107
8	王贤斌	高级	未登录访问	129
9	郝值值	研究生	未登录访问	35
10	刘泽圭	一般	阳明学院	91
11	钱露	高级	未登录访问	125

生成读者画像

▣做到“千人千面+每时每刻”

▣基于在物理空间和数字空间的数据收集

▣引入互联网数据

▣行为分析

✓用户建模

✓社交关系分析：同学、兴趣组、师兄弟

✓关联分析

洞察用户行为与个性化服务

新书提醒、资源推荐、课堂教学

	B	C	D	E
	读者名称	所属院系	浏览内容	访问类型
0	陈德锋	阳明学院	王阳明的生活世界	图书
1	陈德锋	阳明学院	王阳明的生活世界	图书
9	陈德锋	阳明学院	基于循证实践方法的老年人口健康干预研究	图书
3	陈德锋	阳明学院	代际关系对老年人死亡风险影响研究	图书
4	陈德锋	阳明学院	老年社会工作:从理论到实践.第2版	图书
9	陈德锋	阳明学院	老年学概论	图书
1	陈德锋	阳明学院	老年学理论与实践	图书
4	陈德锋	阳明学院	中老年学电脑操作与网络应用	图书
1	陈德锋	阳明学院	中国老年人长期照护筹资制度研究	图书

洞察用户行为与个性化服务

B	C	D	E
读者名称	所属院系	浏览内容	访问类型
刘燕燕		法国汉学史	图书
刘燕燕		法国汉学研究史概述(下)	期刊文献
刘燕燕		法国汉学研究史概述(中)	期刊文献
刘燕燕		法国汉学研究史概述	期刊文献
刘燕燕		耶稣会士罗明坚《大学》拉丁文译本初探	期刊文献
刘燕燕		非英语专业学生英语词典使用情况调查研究——以三江学院日语系新生为例	期刊文献
刘燕燕		非英语专业学生英语词典使用情况调查研究——以三江学院日语系新生为例	期刊文献
刘燕燕		英语专业学生英语词典使用情况调查	期刊文献
刘燕燕		20世纪德国的汉学研究	期刊文献
刘燕燕		讲座-外文期刊投稿简介	新闻
刘燕燕		晚清来华西方人汉语学习与研究	学位论文
刘燕燕		西方传教士汉语闽方言教育史略——从天主教徒早期南洋闽南话著述说起	期刊文献
刘燕燕		西方的文学方言翻译策略研究述评	期刊文献

实现图书馆的智慧空间

在线化

 让你的图书馆永不打烊

数据化

 让你的图书馆生产数据；识别读者及其与空间的联系

交互化

 让你的图书馆更懂读者

知化

 让你的图书馆更聪明

用户近期感兴趣的点——图书订购

检索主题词	检索量
中国知网	3,047
知网	1,005
python	532
音乐美学	458
海洋渔业	327
matlab	265
日本声优	244
东野圭吾	241

B	C	D
检索词	读者名称	读者院系
python	张静	教师教育学院
python	邹毓颖	建筑工程与环境学院
python	刘政	海运学院
python	丁晨航	阳明学院
Python	陈华	信息科学与工程学院
python	项建军	信息科学与工程学院
python	项建军	信息科学与工程学院
python	张心远	机械工程与力学学院
python	阮班超	机械工程与力学学院
python	祝鹏飞	阳明学院
python	邹毓颖	建筑工程与环境学院
python	邵熠羽	机械工程与力学学院
python	李飞云	阳明学院
python	李飞云	阳明学院
python	周诗丹	阳明学院
python	周诗丹	阳明学院
python	付福音	教师教育学院
python	刘晓杰	机械工程与力学学院
python	谢普初	机械工程与力学学院
python	谢普初	机械工程与力学学院
python	谢普初	机械工程与力学学院

医学、农学、理学科用户检索效率高，

目标更明确——懂图书馆的用户、重点用户——高端信息利用

学科分类	阅读未下载	下载使用量	读者数	IP数	未下载比
医学	8852	1403009	786	732692	0.6%
农学	4405	441995	409	209493	1.0%
军事学	254	18651	53	10047	1.4%
理学	12982	705903	1270	408289	1.8%
工学	35214	1359892	2105	660231	2.6%
历史学	3052	109234	434	57989	2.8%
艺术学	3387	99300	373	53277	3.4%
经济学	14884	339353	1192	184119	4.4%
管理学	16768	368577	1321	187289	4.5%
法学	10769	223780	908	120147	4.8%
教育学	25629	528729	1797	294764	4.8%
文学	15131	282262	1048	151519	5.4%
哲学	2821	45604	424	23514	6.2%

从人找资源—>为资源找人

- **图书推荐**
 - 基于借阅历史、学科、内容、社会关系、热门搜索
 - 线下阅览习惯、其他书（网）店的热门书
- **空间资源：座位推荐**
 - 基于个人习惯
 - 馆藏与学科背景相似度、活动与个人兴趣匹配
- **学术资源：学术头条**
 - 从搜索引擎——向推荐引擎转变

学术头条



- 以用户为中心, 让**资源去找人**;
- 海量学术资源的动态推荐;
- 基于大数据的深度学习, 对用户信息、行为动态等方面进行捕捉, 多维度制作精准的用户**画像**, 进而形成个性化、智能化、精准化的推荐;
- 资源类型包括: 期刊、图书、论文、会议、专利、新闻、博客等;
- 人工智能: 自动收割、学科分类、标签生成。

实现功能

- 主动推荐：从“搜”到“推”
- 个性化定制：基于用户身份+用户行为
- **跨类型**服务：多种类型学术资源混合推荐
 - 不同的资源类型
 - 跨语种
 - 跨格式

学术服务个性化

学术
资源

- 学术交流的工具不断变革和创新
- 被动服务转向主动服务
- 大众服务转向个性化服务
- 传统文献库转向开放型学术资源

学术
头条

AI产品化: 让算法利用足够的数据（学术资源），使得产品（学术头条）运行起来，然后通过产品来获取用户，用户在提供更多的数据……周而复始。

数据更新自动化

	网站源	数量
学术会议 持续更新中...	Ourglocal、WikiFPC、丁香园、中国化学会、中国学术会议在线、中国教育网、医学全在线 医学生物会展网、学术会议网 新浪财经一会展、新浪财经网 科学网、艾会网.....	两万余条 实时抓取，日更新量50余条

	网站源	数量
学术新闻 持续更新中...	科学网 http://www.sciencenet.cn/	10万余条 实时抓取，日更新量1000余条
	中新网 http://www.chinanews.com/	
	Nature中文 http://www.natureasia.com/zh-cn/	
	生物帮 http://www.bio1000.com/news/	
	中国社会科学网 http://www.cssn.cn/	
	中国科技网 http://www.stdaily.com/	
.....		

	年份	数量 (篇)
学术论文 持续更新中...	2014	278.8001万
	2015	245.3132万
	2016	235.6011万
	2017	129.4645万
	2018	208.1072万

	类型	数量 (篇)
2019年 持续更新中...	中文会议论文	110746
	英文会议论文	64963
	SCI英文文献	618955
	学位论文	291419
	图书	290625
	专利	33万余条

推荐策略

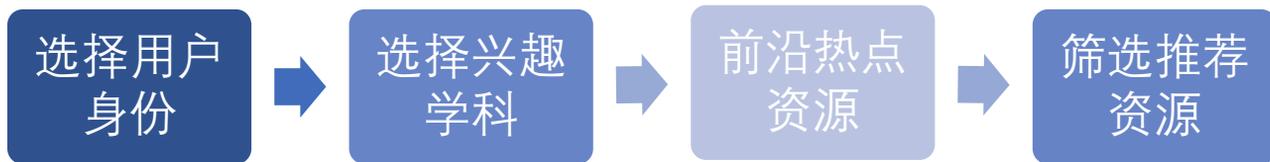
冷启动推荐

- 用户身份
- 用户兴趣学科
- 热点研究内容
- 期刊类型选择
(是否核心)

推荐技术

- 基于资源内容
- 协同过滤
- 混合推荐
- 基于知识图谱

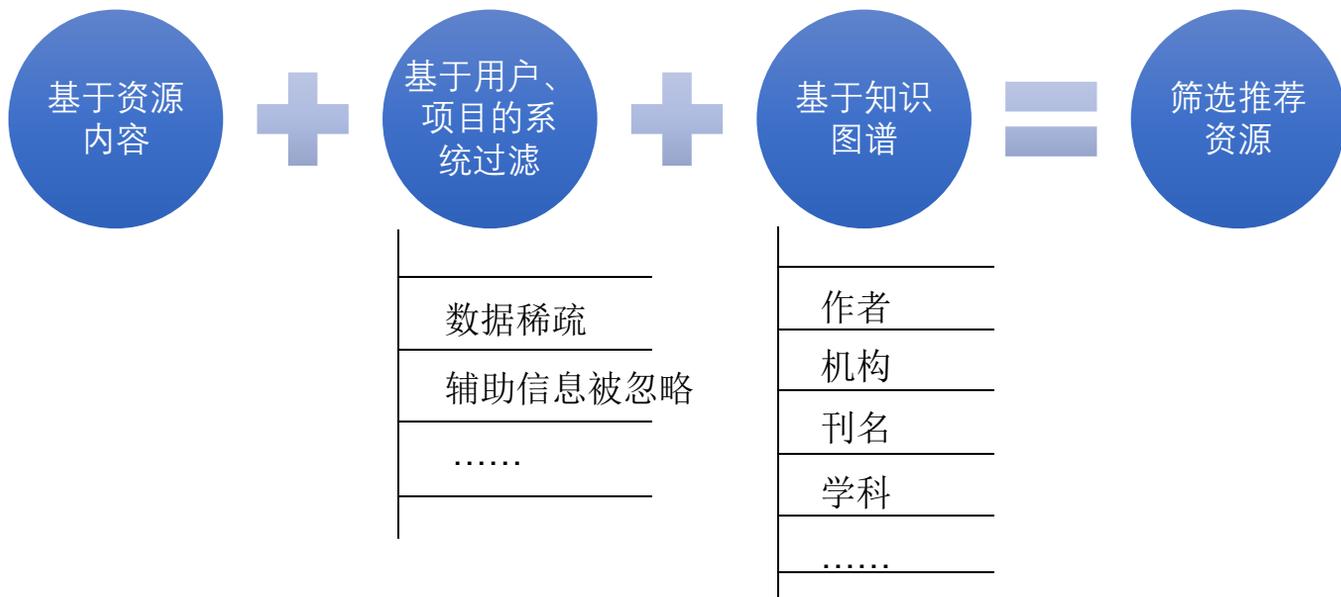
冷启动推荐阶段



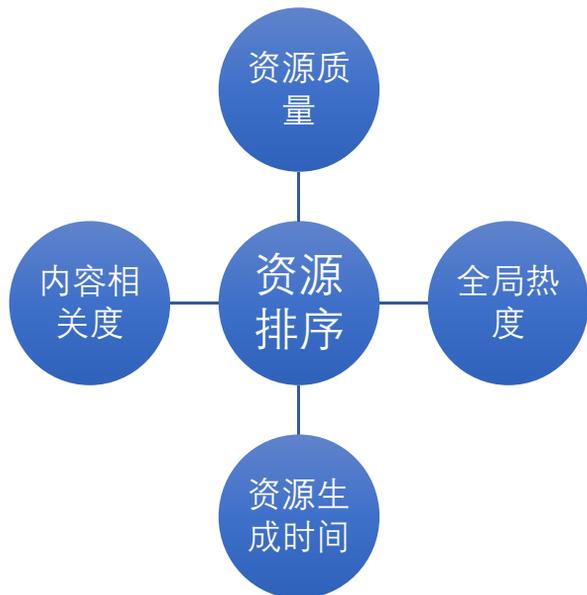
角色、领域、兴趣词、期刊级别



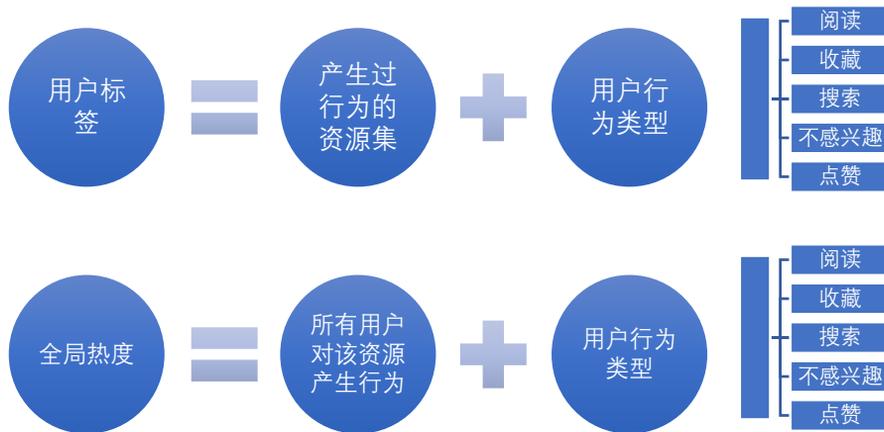
常规推荐阶段



排序策略



推荐模型动态更新



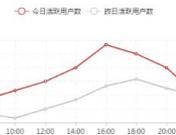
实时更新，
实时影响推
荐资源顺序

两分钟更新
所有资源的
全局热度

机构数据统计

2018年08月14日 星期一 18:06:56

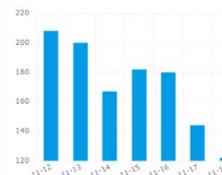
今日活跃用户数: 663 昨日活跃用户数: 753



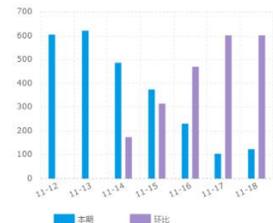
访客趋势



新游趋势



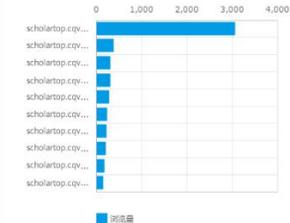
访客活跃天数分布



平均访问时长



页内点击量前10名



重点关键词或排行:

- 军事学 311
- 法学 299
- 政治学 258
- 社会学 230
- 民族学 223
- 马克思主义理论 223
- 历史学 218
- 理学 190
- 天文学 144
- 地球科学 138

IP数趋势



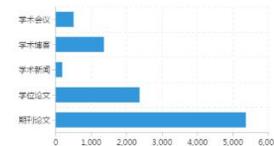
推荐榜单排行:

- 中国影视剧作品走进西视电视频 311
- 科学家创建全球首个自动化火山预警系统 299
- 5省份持续领跑老年政策进步指数 258
- 我国计划生育家庭奖励扶助标准及其测算研究 230
- 晋世界备灾如何预防生育 223
- 中国农村家庭信任度、人力资本与金融资产配置... 223
- 中国重婚家庭成员身份关系的法律定位研究 218
- 美国行政解释及其司法审查标准初探 190
- 司法治理的不平衡不充分发展及其对策研究 144
- 客观归责论与过失犯的认识 138

获取资源排行:

- 我国计划生育家庭奖励扶助标准及其测算研究 311
- 科学家创建全球首个自动化火山预警系统 299
- 晋世界备灾如何预防生育 258
- 我国计划生育家庭奖励扶助标准及其测算研究 230
- 晋世界备灾如何预防生育 223
- 法所保留农村人民陪审制度存在的困境及出路 223
- 中国重婚家庭成员身份关系的法律定位研究 218
- 美国行政解释及其司法审查标准初探 190
- 司法治理的不平衡不充分发展及其对策研究 144
- 客观归责论与过失犯的认识 138

文章类型关注度:



阅读排行:

- 透视重整程序中的债转股 311
- 科学家创建全球首个自动化火山预警系统 299
- 论网络消费权益的法律保护 258
- 商品化权及保护模式研究 230
- 晋世界备灾如何预防生育 223
- 忽略智能科学的到来——晋国中国法律大... 223
- 中国重婚家庭成员身份关系的法律定位研究 218
- 法所保留农村人民陪审制度存在的困境及出路 190
- 司法治理的不平衡不充分发展及其对策研究 144
- 公司行政处置与破产程序的冲突与协调探讨 138

用户数据统计

全部数据 用户群 是否关注 访客类型

91个人
点击查看群体画像

时间排序

请输入ip,用户名,备注,省份或者城市

	111.204.125.163	11-19 15:40:05	新访客
	36.149.11.118	11-19 15:37:08	新访客
	1860118****	11-19 15:36:36	2次回访
	124.114.148.30	11-19 15:30:27	新访客
	106.12.145.38	11-19 15:28:22	新访客
	192.102.204.38	11-19 15:25:33	新访客

王丹
XXX 机构

用户资料

用户id: 8910

年龄: 35

浏览器: qqbrowser

回访: 0次

访问时长: 00:43:52

最后访问: 2018-11-15

email: ****uanwangc

realname: 王丹

用户标签

VIP 在线超过5分钟用户

最近30天行为统计

访问页面	19个	注册成功	1次
注册试用	2次	30天免费体验	1次
注册完成	1次	用户登陆	1次
登录按钮	1次		

访问轨迹

2018-11-19

按类型 按页面

访问页面

2018-11-19 16:59:01	“全国二胎政策”背后的家庭社会学研究	1
2018-11-19 16:59:35	AA制户外运动事故的法律争议及归责	1
2018-11-19 17:07:29	“全国二胎政策”背后的家庭社会学研究	1
2018-11-19 17:10:21	AA制户外运动事故的法律争议及归责	1
2018-11-19 17:10:50	“全国二胎政策”背后的家庭社会学研究	1
2018-11-19 17:12:28	AA制户外运动事故的法律争议及归责	1

访问结束

用户画像

全部数据 用户群 是否关注 访客类型 搜索

91个人
点击查看群体画像

时间排序
请输入ip,用户名,备注,省份或城市

	111.204.125.163	11-19 15:40:05	新访客
	36.149.11.118	11-19 15:37:08	新访客
	1860118****	11-19 15:36:36	2次回访
	124.114.148.30	11-19 15:30:27	新访客
	106.12.145.38	11-19 15:28:22	新访客
	192.102.204.38	11-19 15:25:33	新访客

点击加载更多

王丹
XXX 机构

用户资料

用户id: 8910
年龄: 35
浏览器: qqbrowser
回访: 0次
访问时长: 00:43:52
最后访问: 2018-11-19 16:59
email: ****uanwangdan@1
realname: 王丹

学术兴趣词
平台 智慧 计算机 双高分析

关注学科
军事学 法学 政治学 社会学

常访问学科
历史学 数学 物理学 化学

用户标签

VIP 在线超过5分钟用户 在线超过15个页面用户 在线5分钟+5个页面用户 PV大于1

常访来源类型

期刊论文	23次	学位论文	12次
学术新闻	6次	学术博客	4次
学术会议	2次		

最近30天行为统计

访问页面	19个	注册成功	1次
注册试用	2次	30天免费体验	1次
注册完成	1次	用户登陆	1次
登录按钮	1次		

访问轨迹

2018-11-19

按类型 按页面

访问页面

2018-11-19 16:59:01	“金十二预政策” 首届的家规社会学研究	🔍
2018-11-19 16:59:35	AA制户外运动事故的法律争议及归责	🔍
2018-11-19 17:07:29	“金十二预政策” 首届的家规社会学研究	🔍
2018-11-19 17:10:21	AA制户外运动事故的法律争议及归责	🔍
2018-11-19 17:10:50	“金十二预政策” 首届的家规社会学研究	🔍
2018-11-19 17:12:28	AA制户外运动事故的法律争议及归责	🔍

访问结束

用户画像（手机端）



王丹
XXX机构

导出

访问来源: 微信公众号 EMAIL: 10755566@163.COM

用户资料

OPENID: 25880 昵称: 丹儿
回访: 0次 真实姓名: 王丹
访问时长: 20分钟 访问次数: 3次
首次访问: 2018-10-10 浏览量: 1000kb
最后访问: 2018-10-10

学术兴趣词

平台 计算机 数据分析 区块链 云计算
智慧 图书情报

关注学科

军事学 法学 政治学 社会学 民族学 数学
历史学 化学 天文学 物理学 地球科学
海洋科学 马克思主义理论

常访问学科

历史学 数学 物理学 化学 天文学
地球科学 海洋科学

用户标签

VIP 在线超过15分钟用户 在线5分钟+5个页面用户
pv大于3

常访资源类型

期刊论文: 23次 会议论文: 28次
学术新闻: 6次 学术博客: 45次
学术会议: 2次

最近30天行为统计

访问页面: 19个 注册成功: 1次
注册试用: 2次 30天免费体验: 1次
注册完成: 1次 用户登录: 1次
登录按钮: 1次

访问轨迹

2018-9-18 2018-9-21 2018-9-22 2018-9-26 2018-9-28

按类型 按页面

- 2018-9-21 16:10:21 全面二孩政策, 背学研究社会... ❏ ❶
- 2018-9-21 16:10:21 全面二孩政策, 背学研究社会... ❏ ❷
- 2018-9-21 16:10:21 全面二孩政策, 背学研究社会... ❏ ❸
- 2018-9-21 16:10:21 全面二孩政策, 背学研究社会... ❏ ❹
- 2018-9-21 16:10:21 全面二孩政策, 背学研究社会... ❏ ❺
- 2018-9-21 16:10:21 全面二孩政策, 背学研究社会... ❏ ❻
- 2018-9-21 16:10:21 全面二孩政策, 背学研究社会... ❏ ❼
- 2018-9-21 16:10:21 全面二孩政策, 背学研究社会... ❏ ❽

学者名片

× 机构知识库 ...

学者名片

 **潘复生**
材料科学与工程学院
发文量:1120 被引量:8040 H指数:38

主要领域

工学 理学 管理学 经济学 教育学 医学

研究主题

镁合金 力学性能 显微组织 组织 腐蚀
AZ31镁合金 微观组织 铝合金 性能

合著作者

潘复生 Pan, Fusheng Jiang, Bin
Pan, Fu-Sheng Pan, Fusheng 汤震涛
王敬丰 杨明波 Tang, Aitao 彭建

成果概览

EI:715 WOS:558 SCI-EXPANDED:491
CSCD:343 BDHX:249 SCOPUS:214
CPCI-S:82

× 机构知识库 ...

学者名片

详细介绍

潘复生, 1962年7月生, 浙江兰溪人, 博士, 重庆大学教授、博士生导师, 中国工程院院士, 重庆市科学技术协会主席, 重庆市科学技术研究院院长, 国家铁合金材料工程技术研究中心主任。德国斯图加特大学和日本千叶大学客座科学家, 英国玛丽皇后学院访问教授, 澳大利亚昆士兰大学荣誉教授, 俄罗斯矿业科学院院士, 亚太材料科学院院士。现兼任国际标准化组织 (ISO) “镁及镁合金技术委员会”主席、Elsevier出版社《Journal of Magnesium and Alloys》国际刊物主编、科学出版社《大学科普丛书》编委会主任、中国材料研究会副理事长等。

学者成果

Effects of phase composition on the mechanical properties and damping capacities of as-extruded Mg-Zn-Y-Zr alloys | 被引量:66

Wang, Jingfeng
SCI-EXPANDED WOS EI

镁及镁合金的资源、应用及其发展现状
被引量:93

张丁非
CSCD

Synergetic effect of graphene nanoplatelets (GNPs) and multi-walled carbon nanotube (MW-CNTs) on mechanical properties of pure magnesium | 被引量:80

Rashad, Muham

× 机构知识库 ...

学者名片

学者成果

Effects of phase composition on the mechanical properties and damping capacities of as-extruded Mg-Zn-Y-Zr alloys | 被引量:66

Wang, Jingfeng
SCI-EXPANDED WOS EI

镁及镁合金的资源、应用及其发展现状
被引量:93

张丁非
CSCD

Synergetic effect of graphene nanoplatelets (GNPs) and multi-walled carbon nanotube (MW-CNTs) on mechanical properties of pure magnesium | 被引量:80

Rashad, Muhammad
SCI-EXPANDED WOS EI

碳化硅颗粒增强铝基复合材料开发与应用的研现状
被引量:92

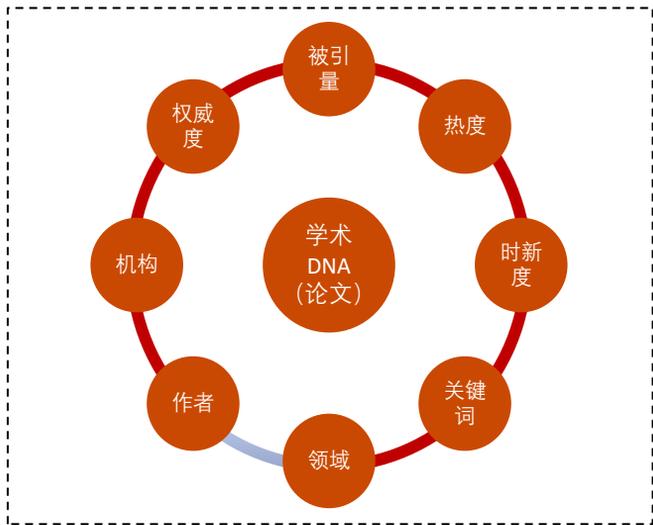
王文明
CSCD

Current research progress in grain refinement of cast magnesium alloys: A review article
被引量:94

Ali, Yahia
SCI-EXPANDED WOS EI

A one-step, cost-effective green method to in

后续：学术基因提取

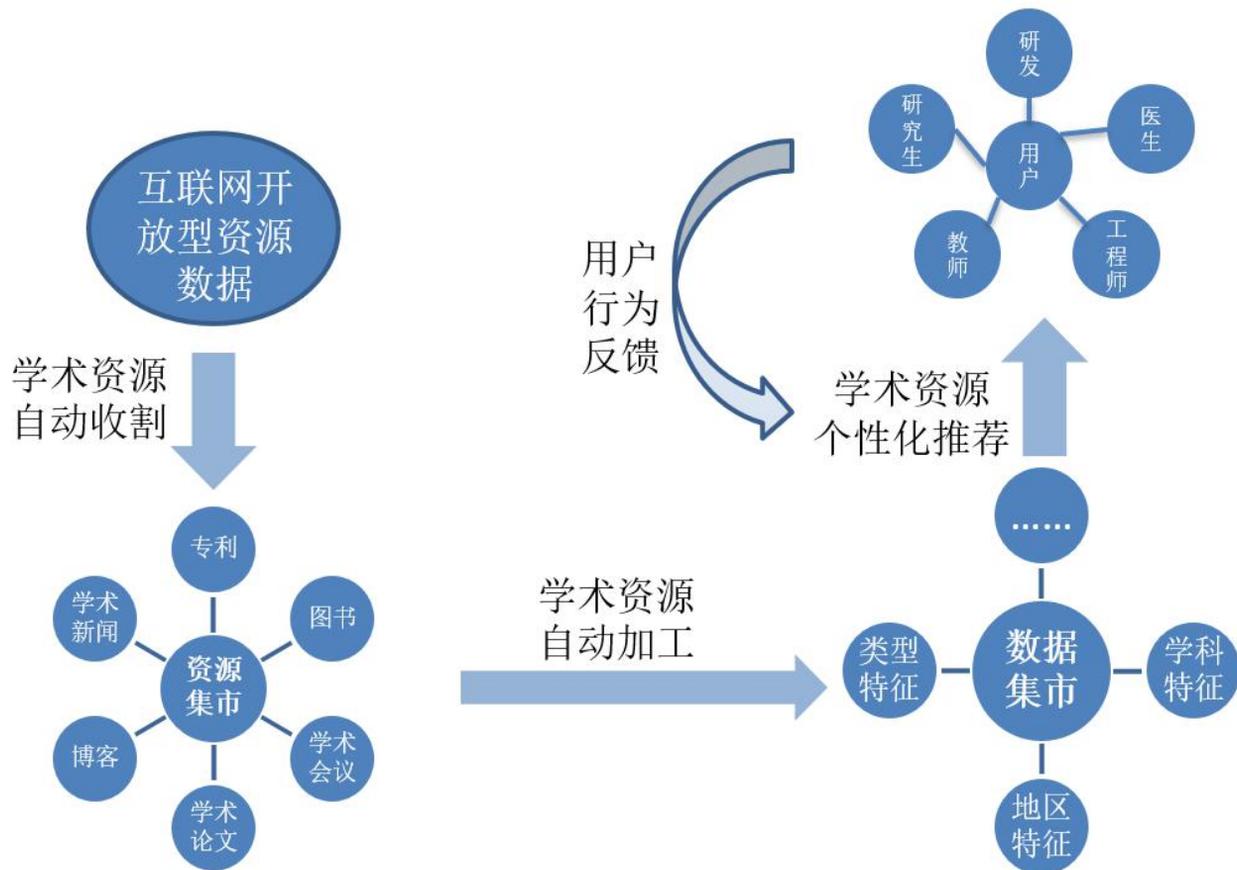


- 学术资源的多标签
 - 学术特征库的建立数十个特征， n 个标签
 - 关键词：大数据、深度学习、超级计算机
 - 题材分类，分类虚构故事的体裁
- 📖 方法：决策树、SVM、神经网络等监督学习模型，用来预测已标记数据

学术进展机器人

- 当前某领域的学术进展如何？
- 某领域的学术热点有哪些？
- 有哪些新锐学者？
- 领域的高被引、高下载的论文有哪些？
- 某机构（作者）有哪些新学术成果？
- 某领域即将要召开的学术会议有哪些？
- 某疾病的相关学术论文和专利有新成果否？

构建学术资源的数据闭环



2.3 服务内容的知识化

数据分析——学术资源

- 全文全媒体的知识点抽取
 - 实体识别
 - 人物；事件；关联关系
 - 数据抽取
 - 时间、GDP
- 图表抽取
- 引用关系及引用态度

如，唐代人物资料的数据化

- 史料的电子化-》“数据化”
 - 45,000多人的姓名、时间、地址、职官、入仕途径、著作、社会区分、亲属关系、社会关系、财产、事件
- 训练计算机判断古籍文本中人物之间的关系
- 哈佛大学、北京大学、“中央研究院”历史语言研究所共同主持

基于全文的知识服务



书是输入



分解：知识元

Basic information:

full name	Wen Jiabao
date of birth	September 1942 (aged 69 years)
place of birth	Tianjin, Tianjin

Images:

official position:

official position	Premier
country	China
start date	2003-03-16

Timeline:

Wen Jiabao

Timeline axis: 1900, 1200, 1400, 1600, 1800, 2000, 2010

The screenshot shows a knowledge graph interface for Wen Jiabao. It includes a 'Basic information' table with fields for full name, date of birth, and place of birth. Below this is an 'Images' section with a portrait of Wen Jiabao. Further down is an 'official position' table with fields for official position, country, and start date. At the bottom, there is a 'Timeline' section with a horizontal axis labeled 'Wen Jiabao' and a scale from 1900 to 2010.

基于需求组合

智慧情报

🏠 语义搜索

- 数据平台：学术数据集市

🏠 智能咨询

- 在线客服
- 个人助理

- 扎克伯格定律（每隔一年互联网内容翻番）

🏠 知识服务

- 无连接-》弱连接-》强连接（信用建立）-》知识社群（圈子）

语义搜索——评价科学家的影响力

- 基于4000万篇文献进行语义分析，且持续增加；
- 识别论文中的不同部分（引言、方法、结果、图表），识别作者对所引用文献的态度，能在一定程度上理解文献内容的人工智能搜索引擎；
- 可对科学家或科研组织对后续研究的影响力评估。



FAQ Contact Sign in

Cut through the clutter

Find peer-reviewed research from the world's most trusted sources

Search over 40 million papers from ArXiv, PubMed, and more...

Try: [Oliver W. Press](#) [Radioimmunotherapy](#) [Deep Learning](#)

How it works



We've pulled over 40 million scientific papers from sources like PubMed, Nature, and ArXiv.



Our AI analyzes research papers and pulls out authors, references, figures, and topics.



We link all of this information together into a comprehensive picture of cutting-edge research.

精准分析

🏠 精准化数据分析

- 舆情分析
- 热点统计
- 医学情报分析
- 商业情报分析

[深扒王宝强离婚内幕 最大祸根源于谁 百山探索](#)
[深度解析宝宝离婚闹剧事件 细说婚姻幸福真谛! 央广网](#)
[宝强离婚最新动态,DNA结果公布马蓉原形毕露 新闻频道_中华网](#)

.....宝宝不知道宝宝的宝宝是不是宝宝亲生的宝宝，宝宝现在担心的是宝宝的宝宝不是宝宝的宝宝如果宝宝的宝宝真的不是宝宝的宝宝那就吓死宝宝了宝宝的宝宝为什么要这样对待宝宝，宝宝很难过，如果宝宝和宝宝的宝宝因为宝宝的宝宝打起来了，你们到底支持宝宝还是宝宝的宝宝！【宝宝心里苦，但是宝宝不说】

🏠 精细化数据分析

- 评论抽取
- 个性化定制

[军民融合南海掀波 陆渔船舰队近逼菲中业岛](#) → 菲律宾 相关
[意大利华人捐古版中国地图 证明钓鱼岛为中国领土](#) → 日本 相关

■ 大数据的精准、精细分析需要智能化技术支撑

到“精准情报”

📖 精准医学→精准情报

📖 理解读者：个性化、精准化

📖 在掌握使用数据的基础上

📖 精准服务;个性化服务

📖 理解内容：数据化、语义化

📖 从元数据级到全文级

📖 自动分类;自动摘要生成

📖 学术标签生成

读者可实现智慧搜索

🏠 精准搜索意图理解

- 精准分类、语义理解、个性化

🏠 复杂多元对象搜索

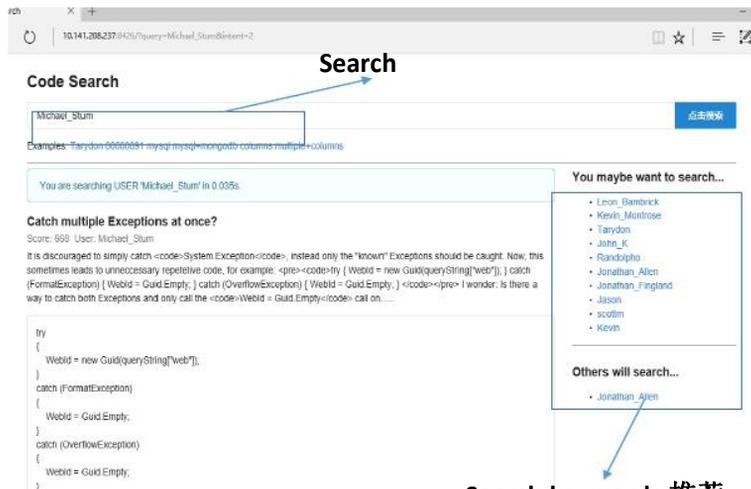
- 表格、文本、图片、视频
- 文案、素材、代码、专家

🏠 多粒度搜索

- 篇章级、段落级、语句级

🏠 跨媒体搜索

- 不同媒体数据联合完成搜索任务



Search keywords 推荐

一切皆可搜索，搜索必达

其他的智慧情报应用场景

人工智能助教：jill watson

-  在线课程，400个学生，1万个问题；
-  2016春季上岗，3个月下来，未被学生发现；
-  基于三年来4万多个问题-答案进行学习。

个人学术助理

-  学者投稿智能引导；
-  自动情报服务；
-  学术热点追踪。

智能参考咨询

-  自动简报；
 -  专题特色库的自动化智能化
-  机器文献综述、自动文摘；
-  自动编目、自动标引。

三、展望与未来

人工智能的三个阶段

🏠 弱智能（ANI）：感知智能

🦉 从基本的计算智能进化而来：利用大量算力

🦉 认知智能（智能1.5）

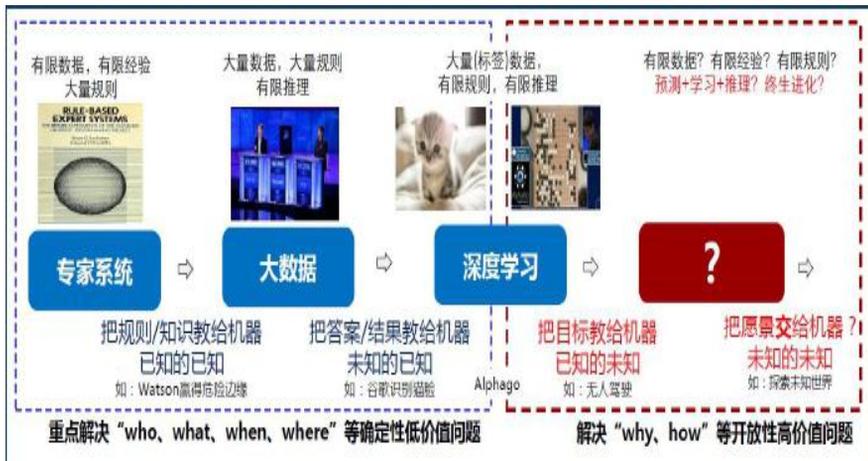
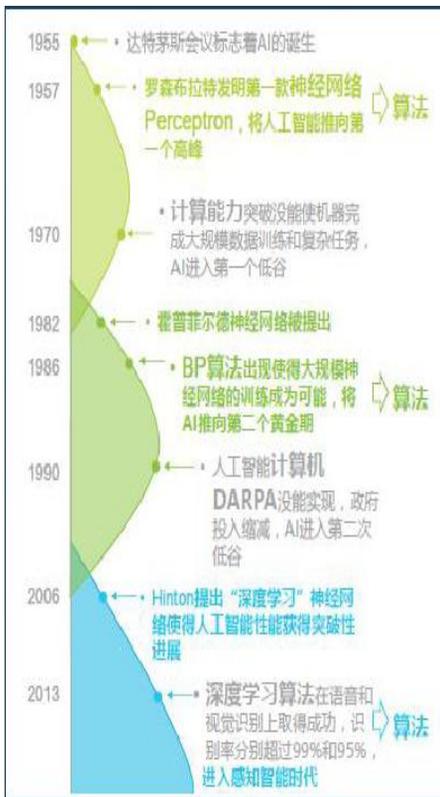
✓ 能够理解、思考、交流：语义解析、对话交互

🏠 强智能（AGI）：人类级别的智能

🦉 让机器有自主智能是最终理想：强大、通用

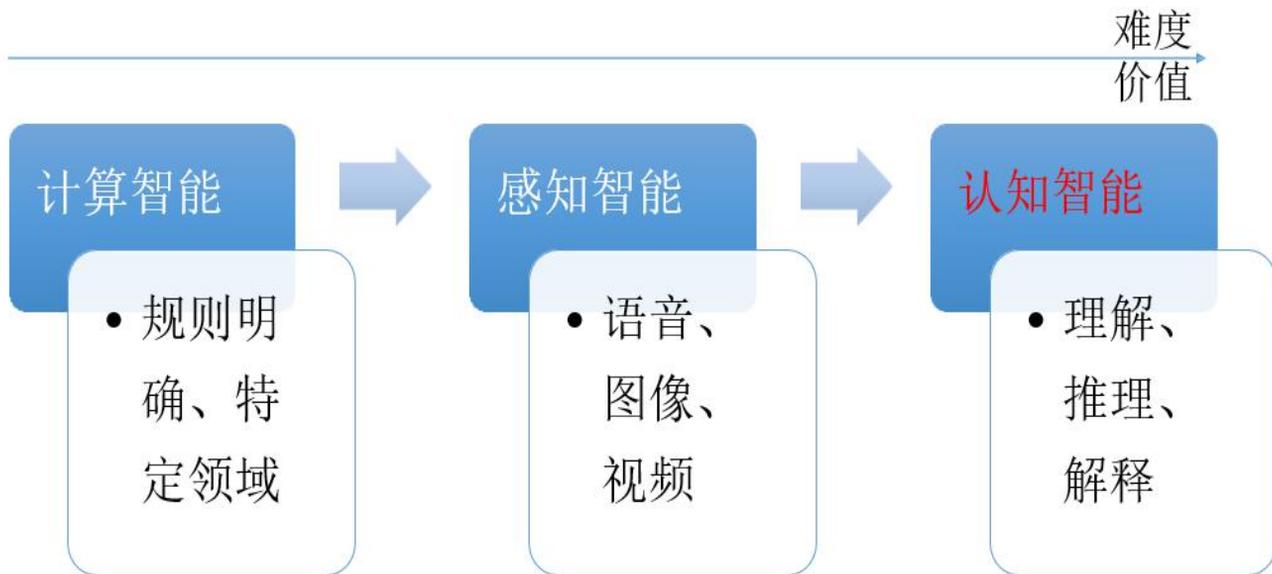
🏠 超人工智能（ASI）

深度学习之后是什么，怎样才能实现通用智能



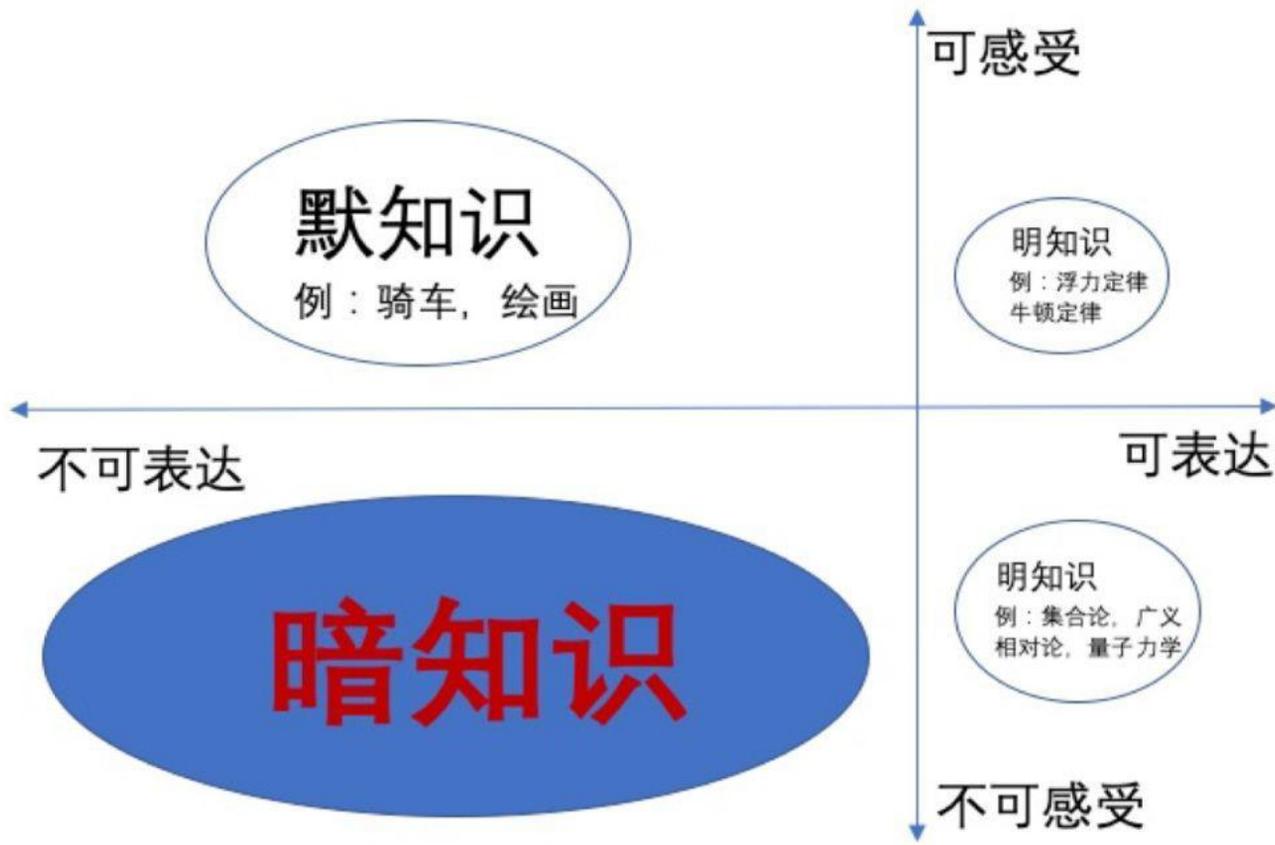
- 解决了关联，还没有解决因果；
- 解决了识别，还没有解决理解；
- 解决了判断，还没有解决推理；
- 解决了胜负，还没有解决意义；

智能化需要机器智能，特别是认知智能



- 随着数据红利消耗殆尽，以深度学习为代表的感知智能遇到天花板
- 认知智能将是未来一段时期内AI发展的焦点，是进一步释放AI产能的关键

人类知识分类



明知识

默知识

暗知识

暗知识



智能
时代

催生



智能
应用

需要



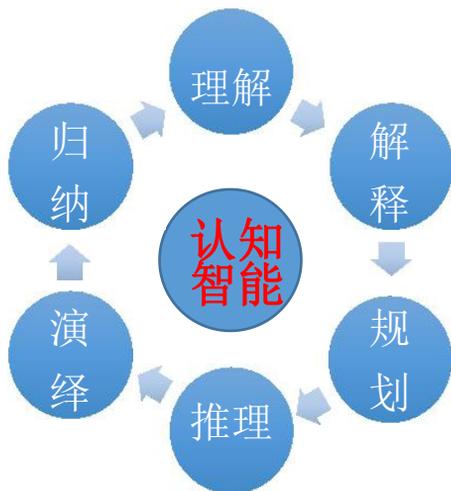
认知
智能

依赖



知识
图谱

认知智能是今后智能化的关键



Can machine *think like humans*?

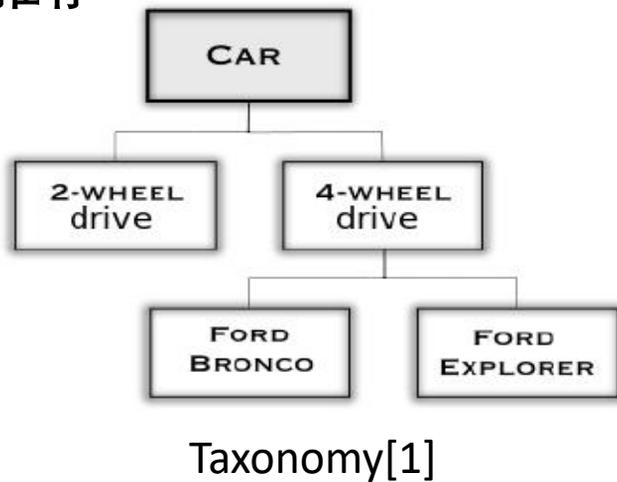
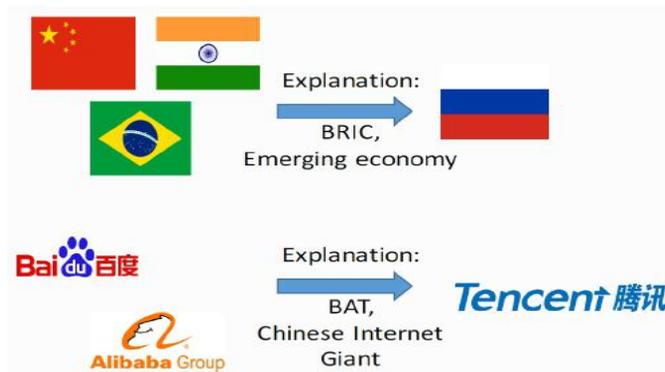


■ 理解与解释是后深度学习时代人工智能的核心使命之一

**知识图谱是实现认知智能的关键技术，
是实现机器认知智能的使能器（Enabler）**

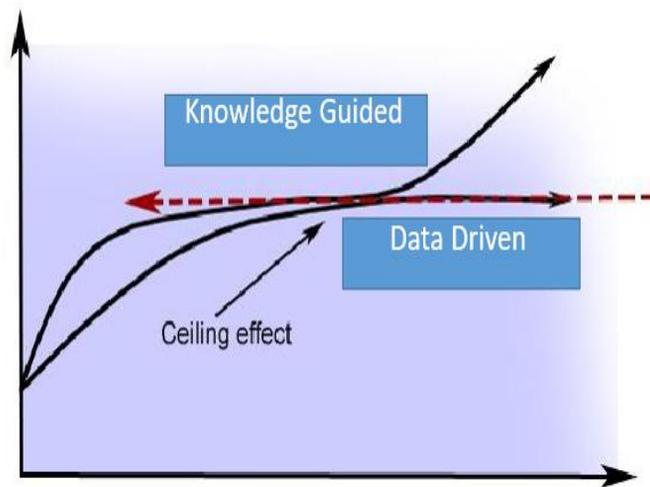
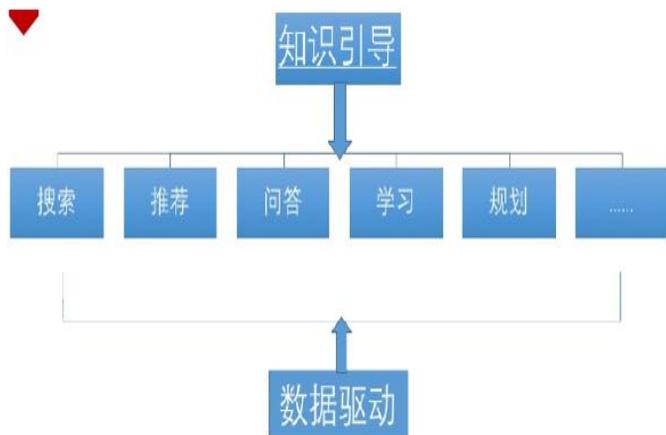
可解释性推荐

基于类型的可解释性实体推荐



- 问题：给定一个实体集合，我们能否理解它们的概念并推荐一些相关实体
- 应用：电子商务推荐场景下，如果用户正在搜索samsung s6和iPhone 6，那么我们应该推荐什么，为什么？

知识引导将成为解决问题的主要方式



- “数据驱动”利用统计模式解决问题
- 单纯依赖统计模式难以有效解决很多实际问题

张三把李四打了，他进医院了

张三把李四打了，他进监狱了

知识将成为比数据更为重要的资产

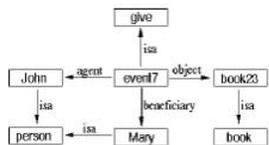
- 大数据时代是得“数据者”得天下
- 人工智能时代是得“知识者”得天下
- 经过知识沉淀的机器智能使得知识工作自动化成为可能
- 数据是石油，知识就是石油的萃取物



知识加工与石油萃取

“Knowledge is power in AI”， Edward Feigenbaum

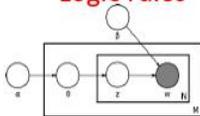
从“小”知识到“大”知识



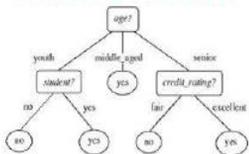
Ontology, Frame

$$\forall x \forall y (P(f(x)) \rightarrow \neg(P(x) \rightarrow Q(f(y), x, z)))$$

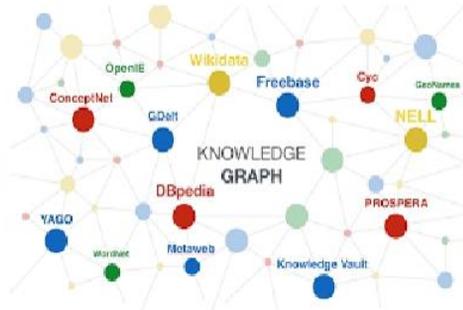
Logic rules



Bayesian network



Decision tree



Big Knowledge



Small knowledge+ big data= big knowledge

知识图谱引领知识工程复兴

新一代图书馆范式

🏠 倡导一个共生协同平台和生态，实现计算-数据-智能融合，用户只需要聚焦业务和模式创新

- 🌟 一个痛点是目前市场上的产品和服务往往是以云计算、大数据和人工智能工具包这三个维度提供的，需要图书馆自行集成组合。这让很多中小型馆苦不堪言，也不利于业务和模式的创新。

🏠 从通用化到场景化

- 🌟 LSP不能包打天下

🏠 从模式化到智能化

- 🌟 可因用户而变：图书馆的C2B
- 🌟 Small knowledge + Bigdata=Big knowledge

🏠 大知识+大协同=新一代图书馆

图书馆的C2B

- 图书馆和读者能够实时互动
 - 5G的推广普及带来万物互联+人物互联;
- 数据等于意见
 - 快速有效地收集读者的需求信息;
 - 形成各类读者社区
- 服务的快速迭代
 - 机器学习具有自我优化的能力，实时持续地对资源和服务做出相应的优化;
 - 根据需求打造个性化产品，高校图书馆更多服务于90后00后的年轻人群体，更偏向于个性化需求。

进一步提升图书馆数据服务体系

	资源标准	数据标准	流程
过去	纸	单数据标准 MARC	纸的流程规范 上一代自动化系统
未来	纸 + 电 文献/行为/用户	多数据标准 MARC/DC/RDA/...	纸+电的流程规范 新一代永远在路上

S2b2c: 系统S通过给每个馆b赋能，更好服务于读者c

智能图书馆OS：情报互联网



图书馆APP (BOOKS、LIB、DISC, etc.)

新一代系统: libOS

图书馆APP
商店

?

资源管理

空间服务

DCS

CRM

LSP

是边缘层
还是核心层?

图书馆智能的进程表

🏠 感知图书馆：弱智能-2020年

- 让图书馆能更懂读者；
- 利用大数据+AI听说读看能力

🏠 认知图书馆：中智能-2025年

- 图书馆能够理解读者、思考交流
- 全新知识服务：深度语义解析、对话式交互

🏠 自主图书馆：强智能-2030年

- 有一定自主智能、行动能力；灵活、强大而通用

数据是基础
知识是关键
协同是推力

智慧图书馆体系

单点智能
协同智能
超级智能

总结：走向智慧图书馆

📦 大数据：大量海量实时更新的数据和知识引擎

- 本地中央库B：全方位收集各类数据
- 网上云平台S：知识引擎提升对师生的服务能力
- 本地化APP：有效地将每一服务、每一份资源推送给用户

📦 借力：三个大趋势的汇聚

- 海量**数据**（Big Data）《= 互联网、物联网
- 线上线下融合**OMO**：引进各种**技术**积木，为本馆所用
- 互联网大平台+本地化数据=》本地化**服务**

谢谢！
请批评指正！
lbs@nbu.edu.cn

