



领域知识图谱构建与应用

杜一

中国科学院计算机网络信息中心

2020年4月

领域知识图谱

- 知识图谱（Knowledge Graph）
 - 以提高人工智能可解释性为主要目标的智能搜索引擎，其本质上是一种语义网络，以实体语义为核心，能够提供从关系的角度分析问题的能力。
- 领域知识图谱（Domain Knowledge Graph）
 - 针对通用知识图谱存在的精准度较差、专业性较低、领域时效性较弱的问题，面向特定领域构建知识网络，能够将知识网络赋能特定领域。

通用知识图谱

Freebase

LINKINGOPENDATA
W3C SWEO Community Project

WordNet
A lexical database for English

ZHISHI.me

PKUBASE

NELL

schema.org

...the new SEO?

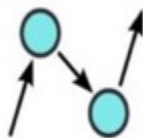
DBpedia

XLORE

WIKIDATA

WEB CHILD

CN-DBpedia



ConceptNet

An open, multilingual knowledge graph

Herbnet

yago
select knowledge

LinkedGeoData.org

WEBKB

linked life data

通用图谱成功的应用

中国是哪一年成立

All News Images Videos More Settings Tools

About 703,000,000 results (0.90 seconds)

China / Founded

October 1, 1949

People also search for

- Japan February 11, 660 BC
- Russia December 25, 1991
- United States July 4, 1776

Feedback

Google

中国人民解放军

Remove

- 中国人民解放军
- 中国人民解放军海军
- 中国人民解放军军歌
- 中国人民解放军国防大学
- 中国人民解放军进行曲
- 中国人民解放军空军
- 中国人民解放军陆军工程大学
- 中国人民解放军信息工程大学
- 中国人民解放军战略支援部队
- 中国人民解放军火箭军

Google Search I'm Feeling Lucky

Report inappropriate predictions

People's Liberation Army

Armed force



The Chinese People's Liberation Army is the armed forces of the People's Republic of China and its founding and ruling political party, the Communist Party of China. The PLA consists of five professional service branches: the Ground Force, Navy, Air Force, Rocket Force, and the Strategic Support Force. [Wikipedia](#)

Budget: 177.6 billion USD (ranked 2nd, 2019)

Founded: August 1, 1927, [China](#)

Commander in chief: [Xi Jinping](#) (General Secretary and President)

Founders: [Mao Zedong](#), [Zhu De](#)

Headquarters: [Central Military Commission, Beijing, China](#)

Did you know: People's Liberation Army is the world's second-largest employer by number of employees (2.52 million). [wikipedia.org](#)

People also search for

View 15+ more



People's
Liberation
Army Navy



People's
Armed
Police



United
States
Armed Fo...



People's
Liberation
Army Air...



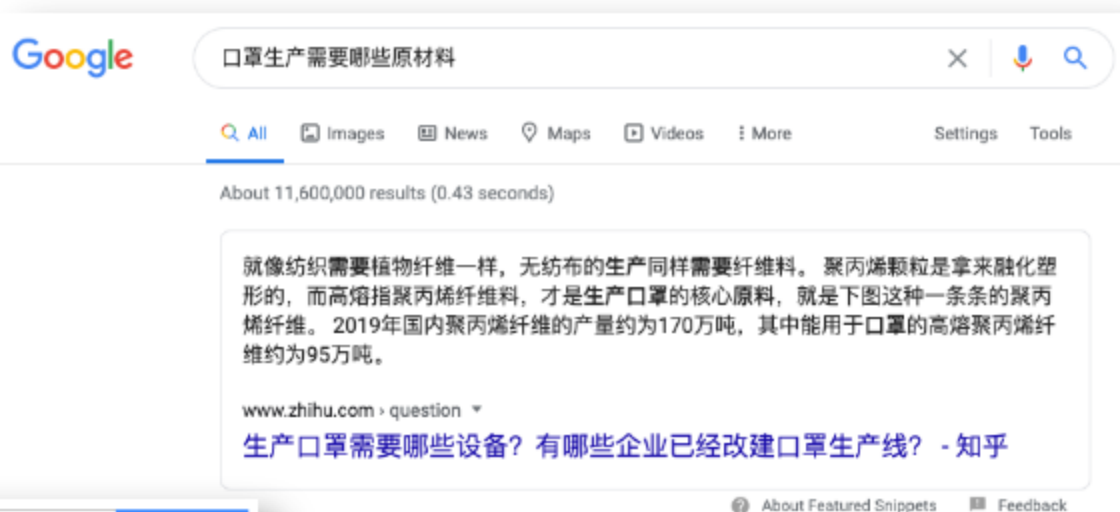
Russian
Armed
Forces

Claim this knowledge panel

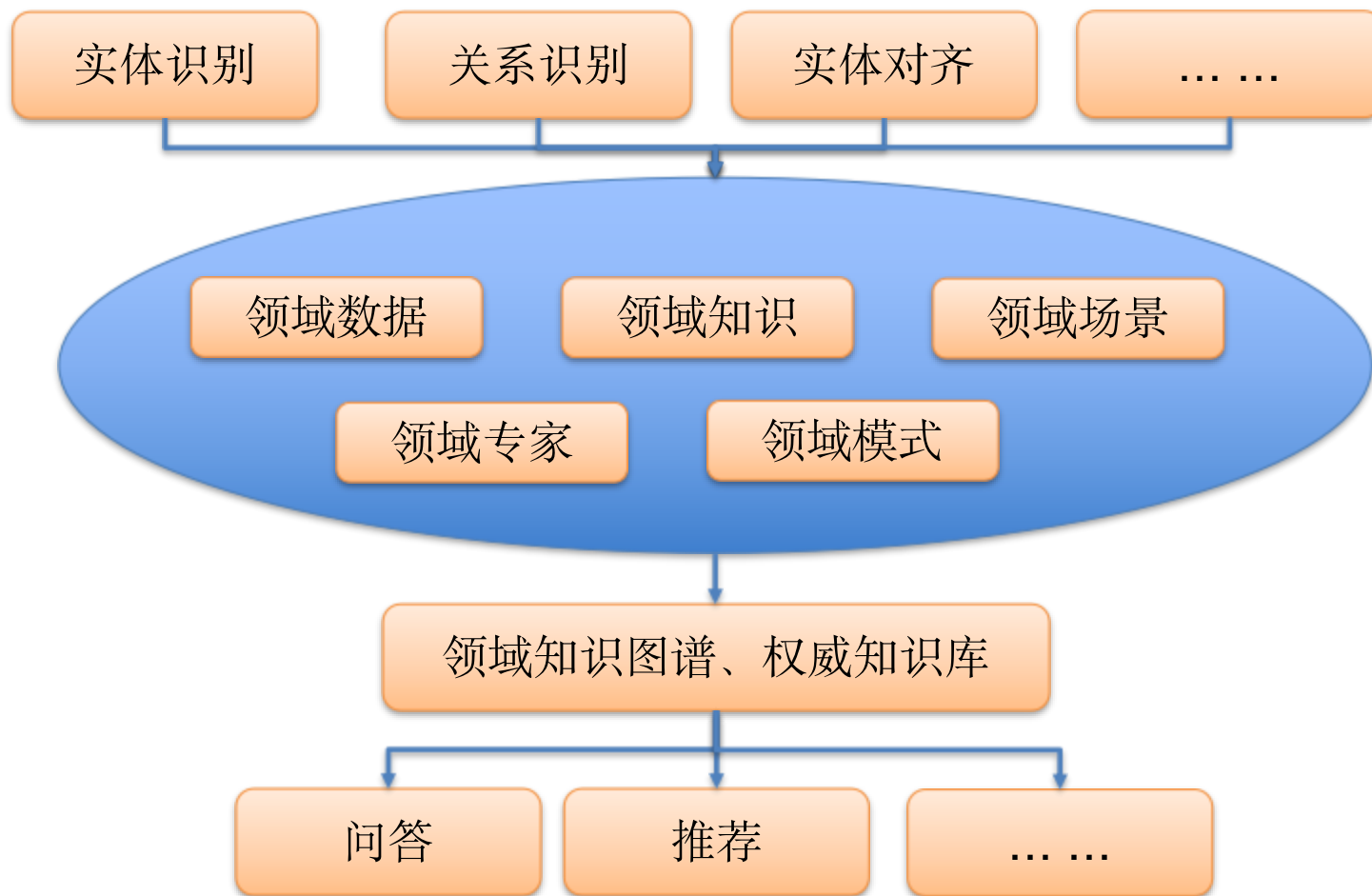
Feedback

但在领域应用时依然有很大瓶颈

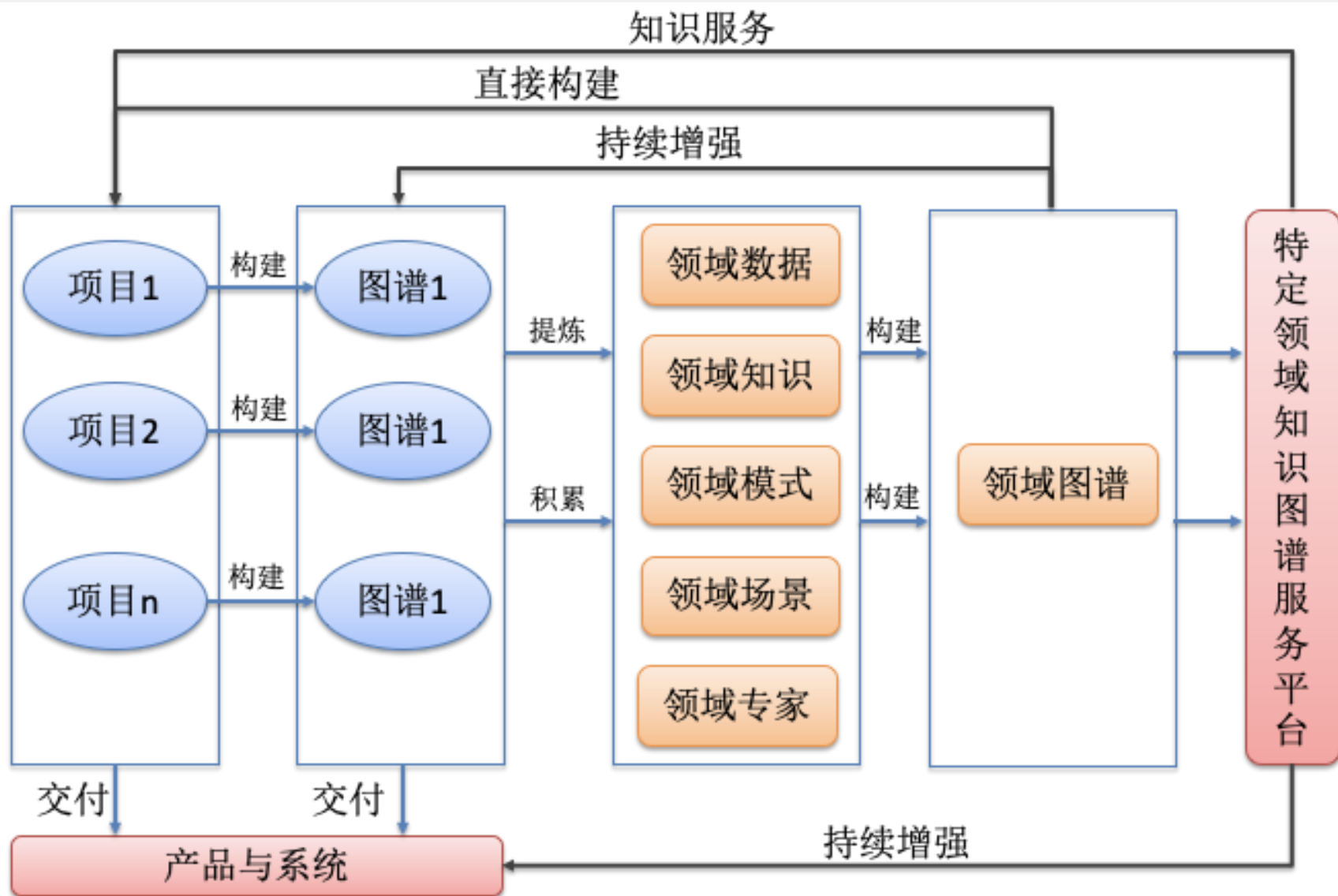
专业性
精准性
实时性



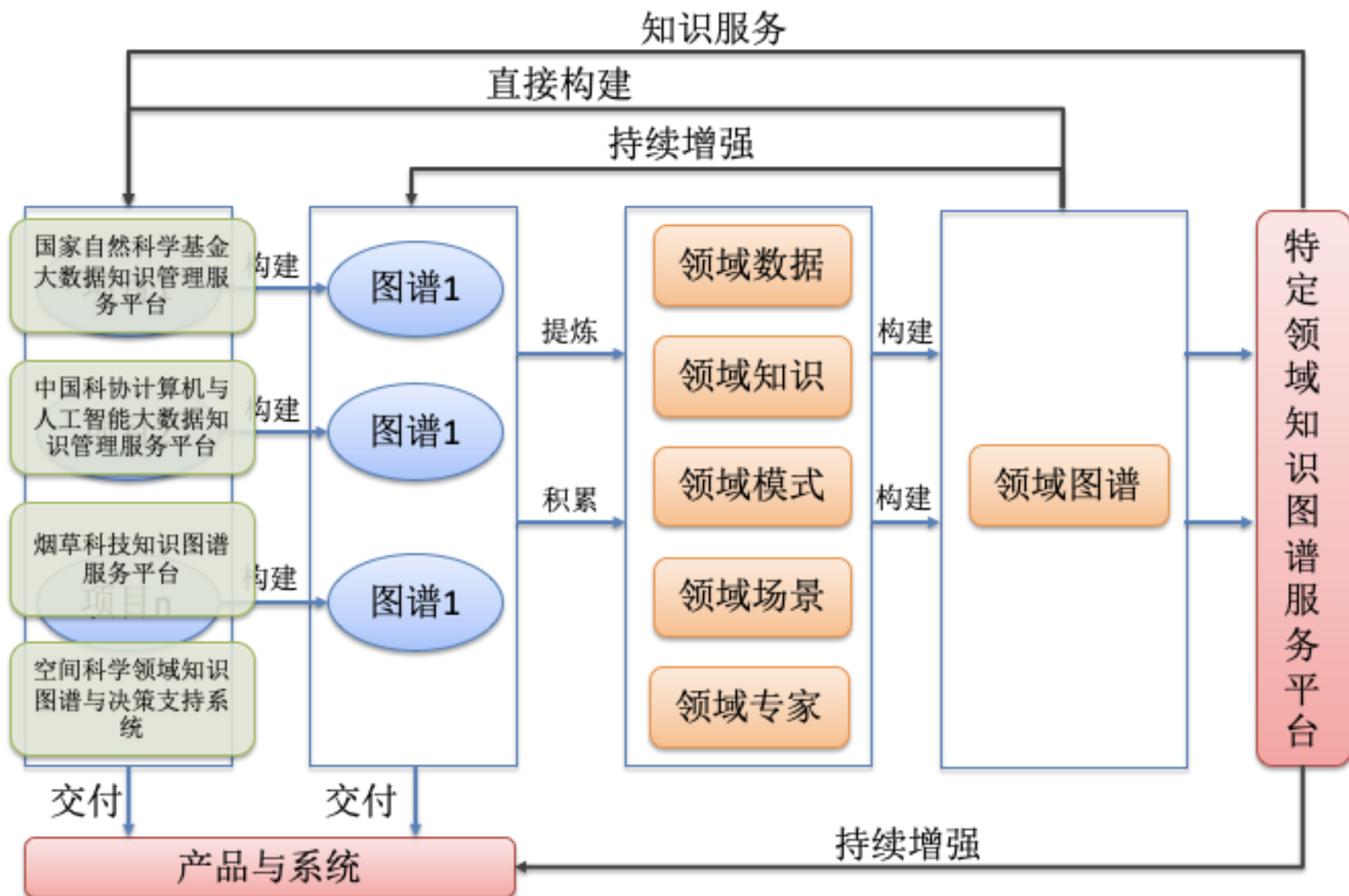
领域图谱构建思路



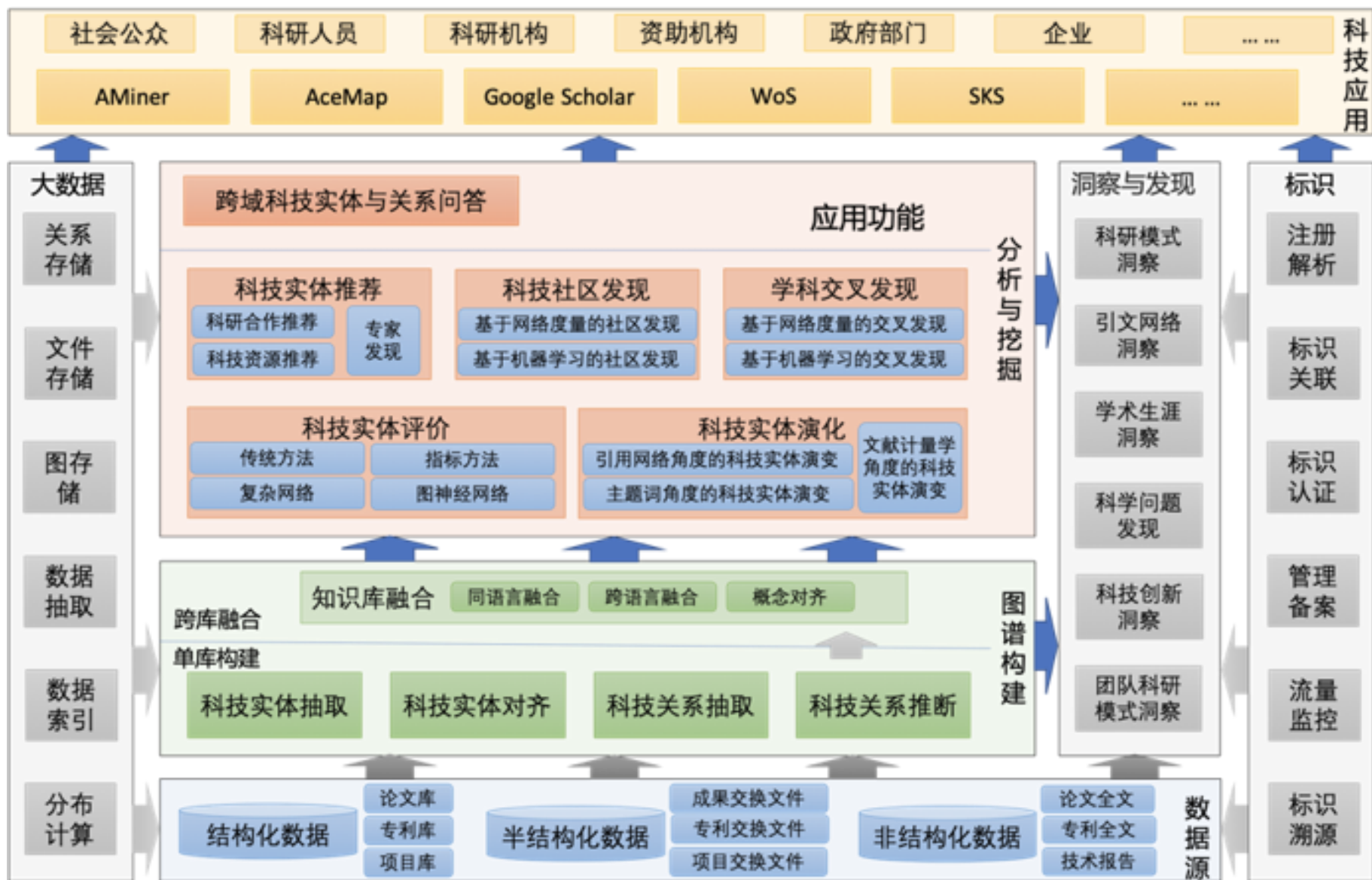
我们的建设路径



科技领域知识图谱

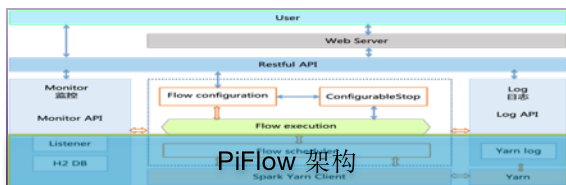


科技领域知识图谱：研究架构

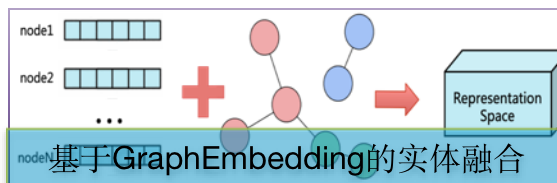
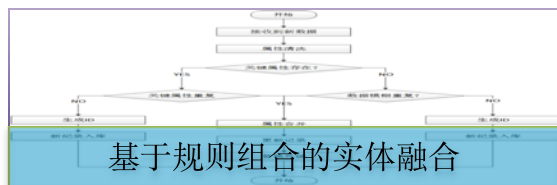


科技领域知识图谱：技术积累

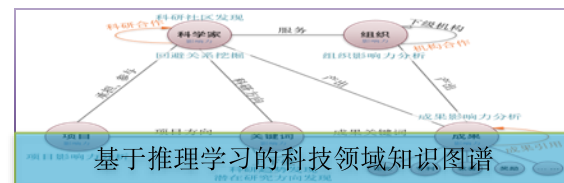
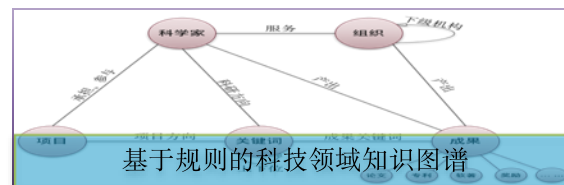
自动及半自动的数据采集、处理流水线机制



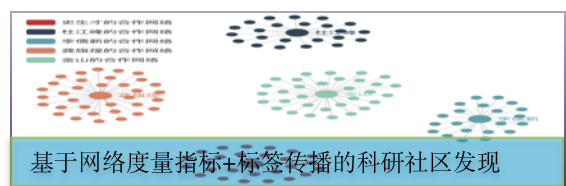
基于统计规则及深度学习方法的实体融合方法



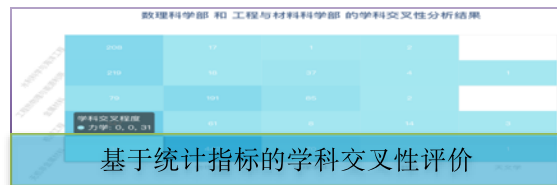
基于属性图的科技领域知识图谱构建方法



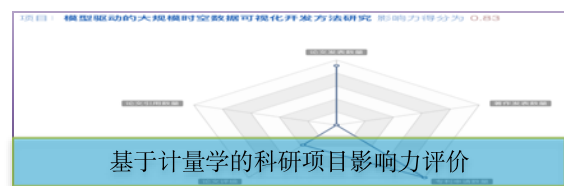
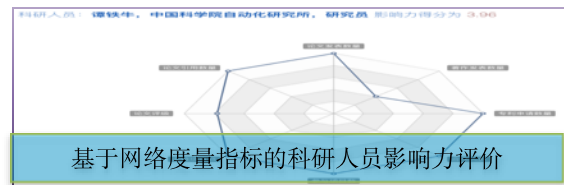
基于学科的科研社区发现算法



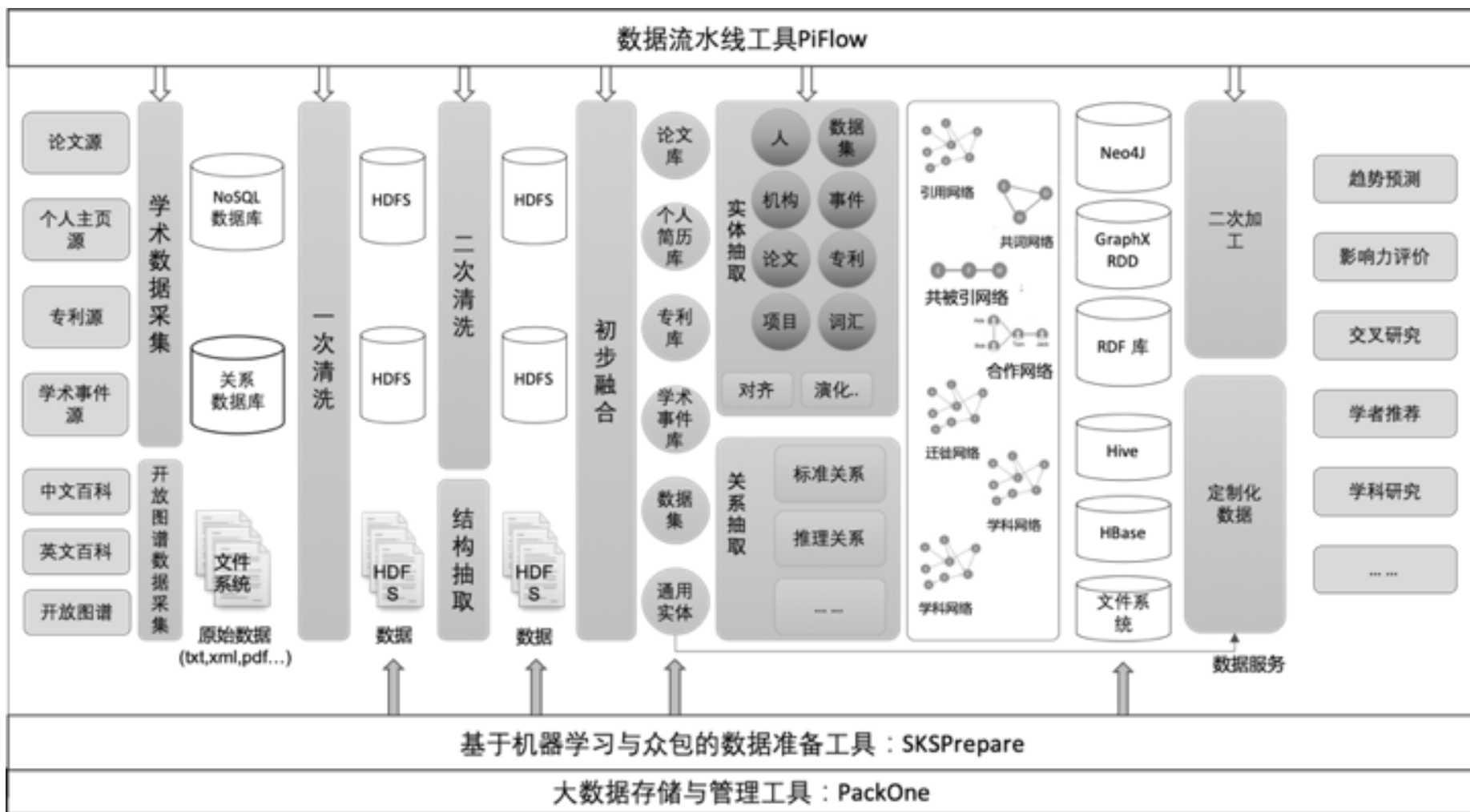
学科交叉性评估及学科趋势预测方法



学术影响力评价方法

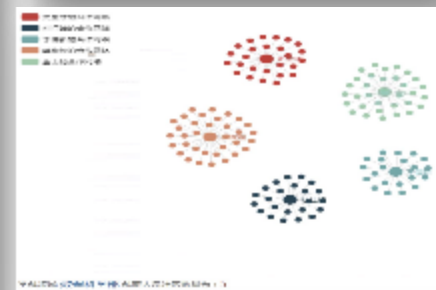
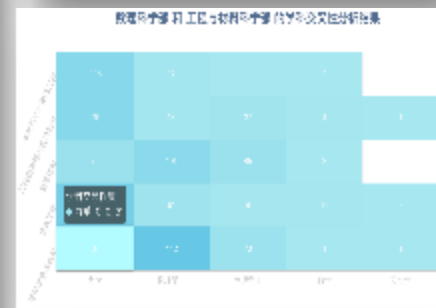
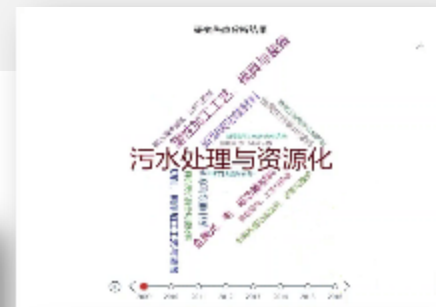


科技领域知识图谱：图谱构建框架



科技领域知识图谱：产品与系统

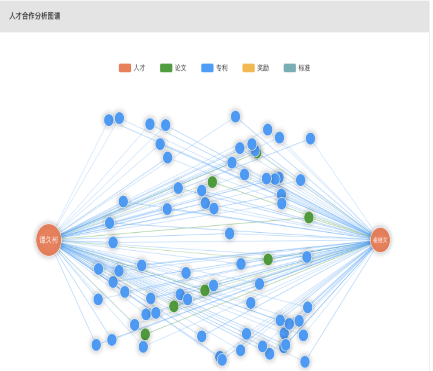
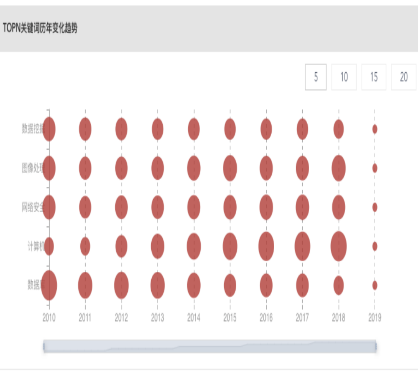
- 国家自然科学基金大数据知识管理服务门户



科技领域知识图谱：产品与系统

中国科协计算机与人工智能大数据知识管理服务平台

关键词	2016-2017	2017-2016	2016-2015	2015-2014	2014-2013	2013-2012	2012-2011
数据库	83.31%	72.76%	108.2%	79.81%	87.55%	95.66%	104.7%
计算机	106.6%	100.8%	117.8%	107.3%	115.4%	129.5%	136%
网络安全	106.1%	83.47%	107.2%	102.4%	107.5%	97.32%	112.6%
数据治理	124.2%	94.46%	86.95%	110.5%	116%	102.4%	99.63%
数据智能	104.2%	74.91%	99.11%	93.36%	104.2%	90.83%	92.79%
应用	112.8%	88.87%	118.4%	95.69%	123.6%	139.4%	122.8%
数据算法	100.5%	81.48%	102.5%	88.32%	96%	88.18%	82.16%
数据分割	103.2%	82.08%	99.2%	92.51%	99.6%	91.36%	97.38%
设计	80.82%	69.05%	124.4%	94.08%	111.2%	109.1%	128.5%
云计算	98.22%	76.75%	95.73%	93.22%	118.4%	123.2%	142.8%



智慧科协态势感知系统

16,531,090

国际、人才、团队、机构、论文、热点、专利、奖项

中国发明专利

2008-2016 有效/失效

科研专利

2008-2016 有效/失效

获奖情况

2017 国家科学技术进步奖 220, 国家自然科学奖 54, 国家技术发明奖 63

Geography-中外人才对比

2010-2019 数量/占比

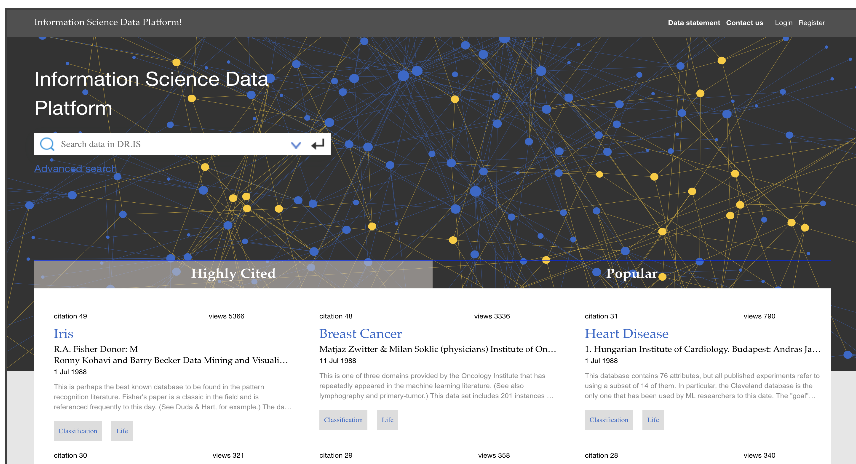
智慧科协知识服务平台主界面，包含用户信息、学科热度、人才合作网络、领域研究、主要研究项目、获奖情况、科研专利、Geography-中外人才对比等模块。

科技领域知识图谱：产品与系统

烟草科技知识图谱服务平台



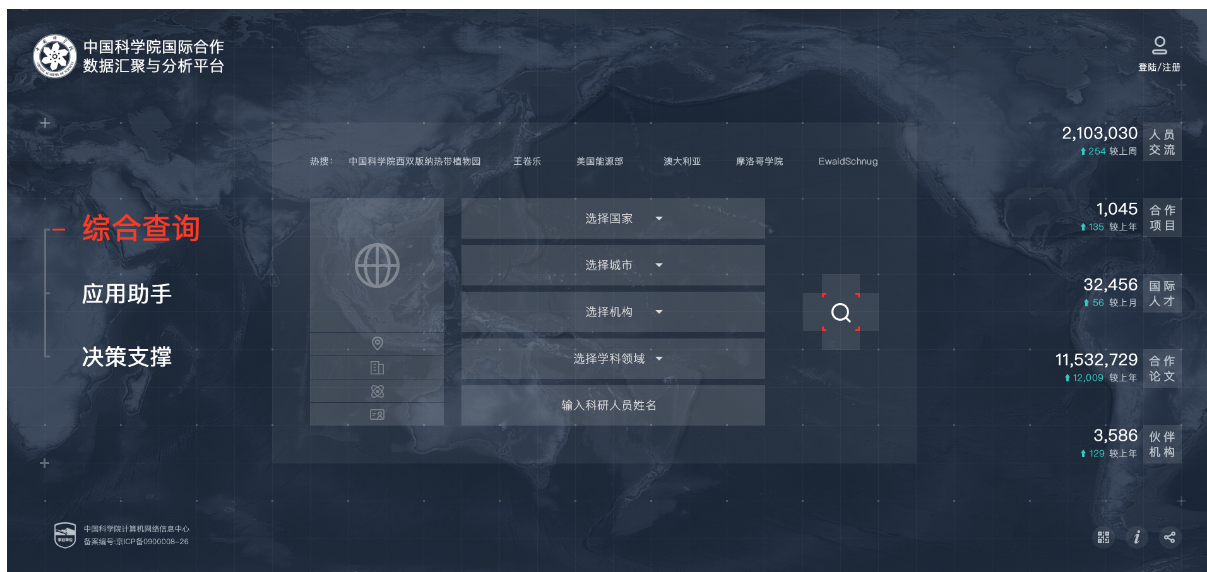
科技领域知识图谱：产品与系统



Information Science Data Platform



基于知识图谱的科技合作智能化管理服务平台



中科院国际合作数据汇聚与分析平台

还有一个现实问题

领域数字化、信息化程度



一个案例

——国家自然科学基金大数据知识管理服务平台

背景



网络信息系统



基础研究知识库



共享服务网



成果之友

数据量增大

数据孤岛

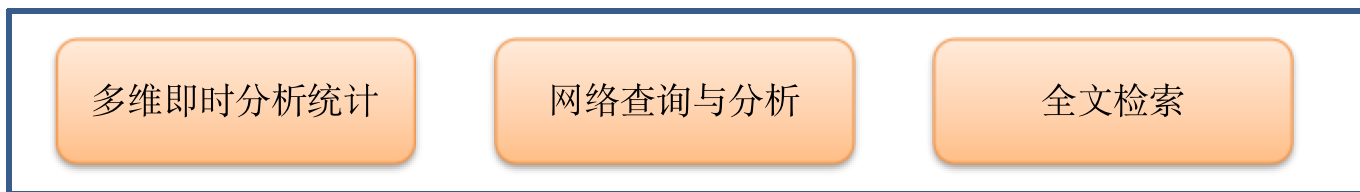
深入的分析与挖掘
(知识图谱)

系统架构

应用系统



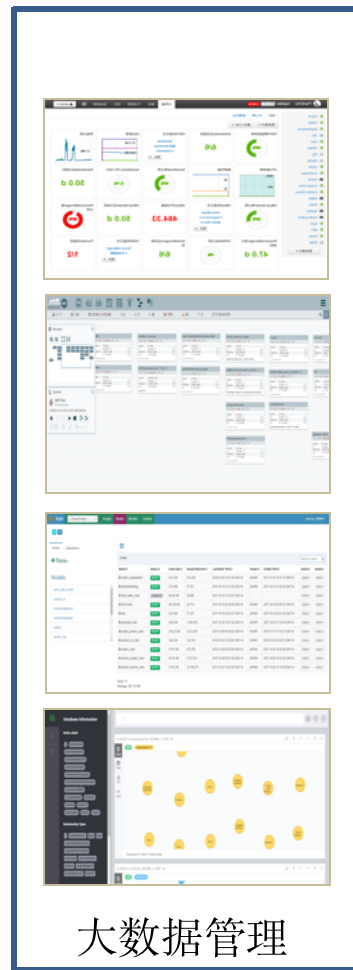
数据服务



数据模型



数据源



应用系统



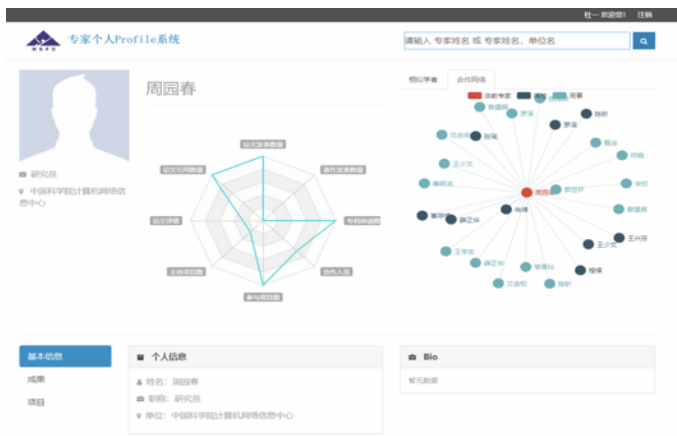
大数据知识管理服务门户



科技成果信息服务系统



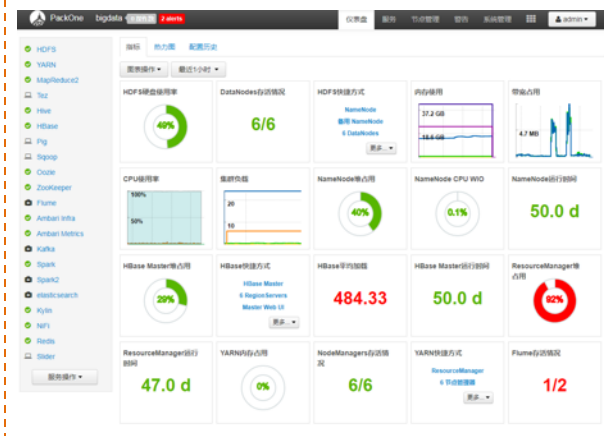
基础研究知识库服务系统



专家个人Profile系统



掌上服务APP



大数据平台集成管理



我们还在做哪些领域知识图谱？

其它领域知识图谱

工业制造

医疗健康

烟草

航空

2019年工业互联网
创新发展工程——
标识解析公共服务
支撑平台项目

国家重点研发计划
食源性疾病监测、
溯源与预警技术研
究(基于多源数据的
食源性疾病实时预
警技术体系研究)

国家重点研发计划
云计算和大数据专
项：制造业跨域异
构数据智能治理技
术研究(主数据语义
融合与分布伸缩可
配置存储计算技术)

卷烟销售行为
领域知识图谱
构建

民航用户行为
领域知识图谱
构建

基于科技领域知识图谱的成果评估方法研究; 国家自然科学基金成果开放共享政策研究与平台架构设计; 基于群智理论的创新方法新系统研究与应用示范; 面向领域大数据的知识图谱构建... ..



发个广告

招聘

- 我们发顶级论文，但更能落地
- 岗位：
 - 知识图谱**算法**研发
 - 知识图谱**平台**研发
 - 知识图谱**应用**研发
- 要求：
 - 硕士、博士，优秀本科
 - 对数据敏感、逻辑清晰、动手能力强

请批评指正！