

# 去往认知海洋的一艘船

中国知识图谱行业研究报告

2019年



“

认知智能是探寻人类学习、理解、产生决策的生物能力，并将之赋予机器的技术统称，而实现机器认知，需要突破感知泛在、意识建立、低维到高维的主观概念形成、推理决策能力唤醒，以及多模态知识持续学习等能力瓶颈，其产生的价值也将使人工智能更加贴近人类的思维方式，而知识图谱作为人类意识和概念的载体是现阶段认知智能主要的发展方向之一。

2019年是知识图谱相关技术飞速发展的一年，世界顶级NLP大会ACL一次性收录了超30篇高质量知识图谱类论文，其中对于关系向量和图神经网络的论述将对知识图谱中关系补全、推理决策和认知计算带来长足发展。在国内建设应用方面，工信部发文明确指出，2020年将围绕工业大数据融合应用、民生大数据创新应用、大数据关键技术先导应用、大数据管理能力提升4大类7个细分方向着重发展，而知识图谱作为集大数据和人工智能于一身的综合技术，也将成为重点关注领域。

”

——艾瑞咨询研究院

1

人工智能本质是解决生产力升级的问题，人类生产力可以归类为知识生产力和劳动生产力，人工智能走入产业后，可以分为感知智能、认知智能和行为智能，后两者更与生产力相对应，NLP和知识图谱是发展认知智能的基础。

2

原始数据通过知识抽取或数据整合的方式转换为三元组形式，然后三元组数据再经过实体对齐，加入数据模型，形成标准的知识表示，过程中如产生新的关系组合，通过知识推理形成新的知识形态，与原有知识共同经过质量评估，完成知识融合，最终形成完整形态上的知识图谱。

3

在面对数据多样、复杂，孤岛化，且单一数据价值不高的应用场景时，存在关系深度搜索、规范业务流程、规则和经验性预测等需求，使用知识图谱解决方案将带来最佳的应用价值。

4

2019年涵盖大数据分析预测、领域知识图谱及NLP应用的大数据智能市场规模约为106.6亿元，预计2023年将突破300亿元，年复合增长率为30.8%，其中2019年市场中以金融领域和公安领域应用份额占比最大。

5

随着整体市场数据基础的完善和需求唤醒，大数据智能领域规模持续走高，但在行业可落地性和理性建设的限制下，预计市场增速将呈现下降趋势，期间咨询性需求将会大量出现，从整体发展来看增速处于良性区间，对真正有价值的公司和产品有正向意义。

知识图谱技术概述

1

中国知识图谱市场概述

2

中国知识图谱细分市场分析

3

中国数据智能代表企业案例展示

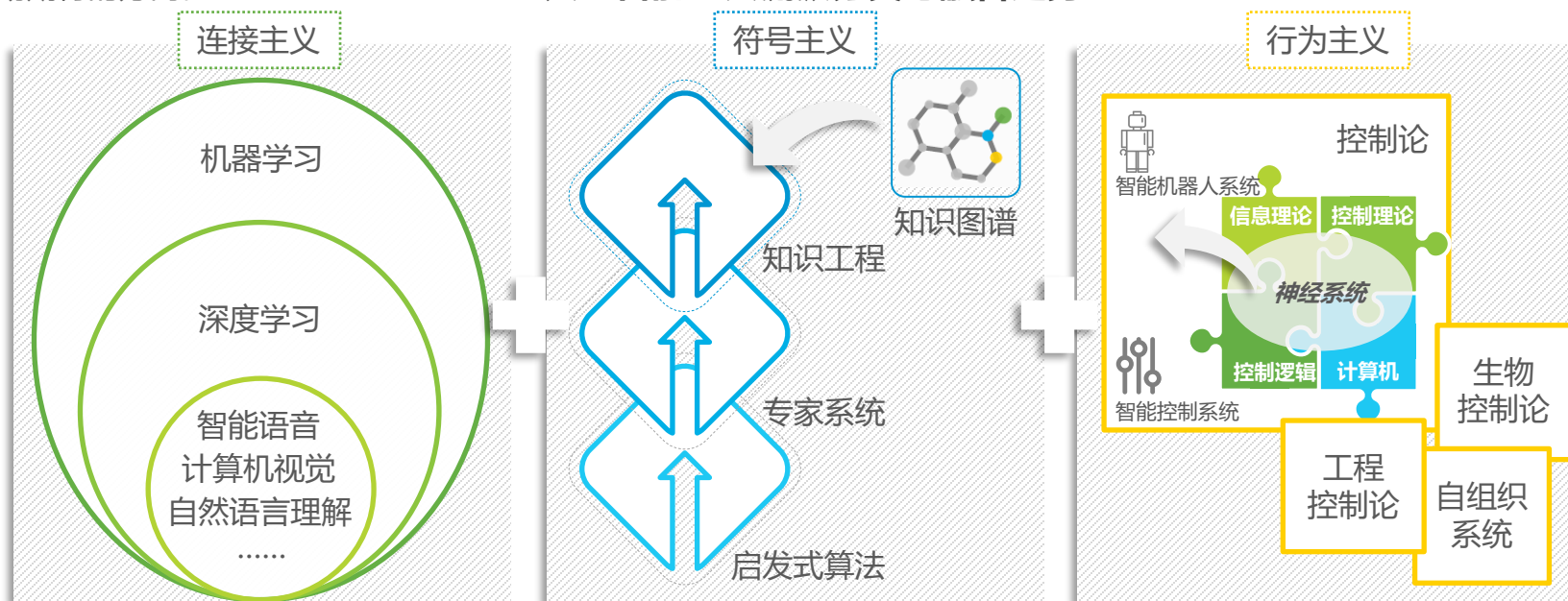
4

# 人工智能技术分类和趋势

## 三种流派的融合应用，使人工智能向想象更进一步

人工智能是对一类能够实现机器模拟智慧生命某些特征的技术统称，从学术上可以分为，对人类已有知识进行组织编辑的符号主义、通过数学理论公式推导聚类 and 预测问题的连接主义，以及利用机器模仿生物活体行为的行为主义三个流派，分别以知识工程、机器学习和仿生机器人为时代代表，而知识图谱就是新一代知识工程的具体体现。2012年，深度学习在计算机视觉和智能语音上产生重大突破，打开了人工智能商业化的大门，使得连接主义一度成为人工智能的代名词，但随着应用落地成为主旋律，缺位行业逻辑和理论概念的连接主义，往往找不到最佳的应用场景而止步于浅层尝试，在此背景下，人工智能技术应当走向融合，符号主义需要连接主义提供强大的计算支撑，连接主义需要符号主义的逻辑指导，二者又共同作用于行为主义，充当机器人的大脑和“记忆宫殿”，在多种技术综合利用下的垂直领域智能解决方案才是当今最符合市场期待的方向。

### 人工智能三大流派分类与融合趋势



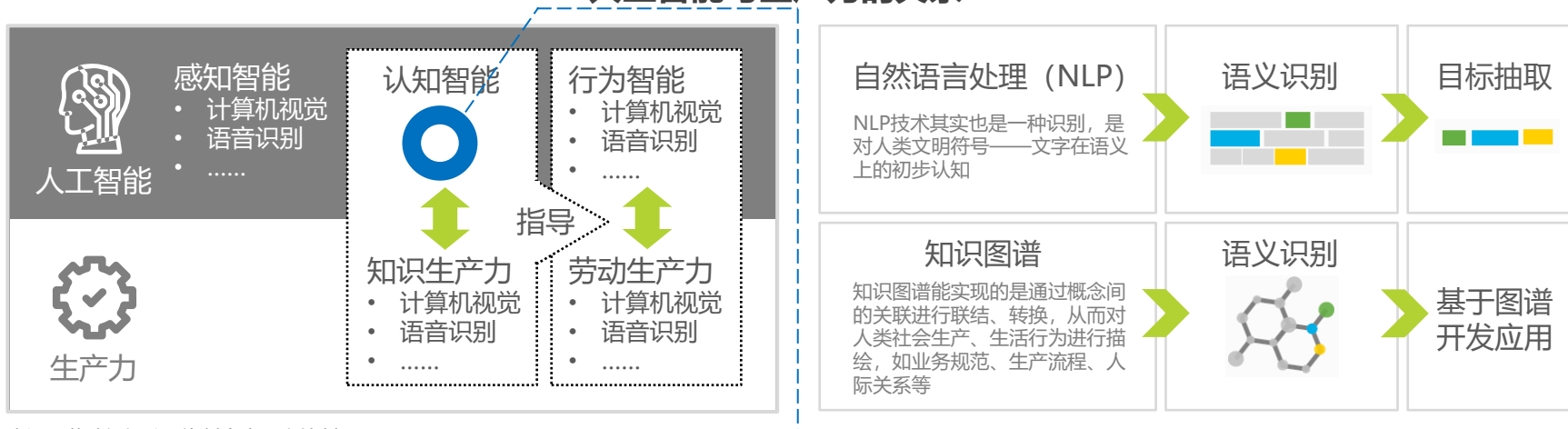
来源：艾瑞根据公开资料自主研究绘制。

# 人工智能技术应用本质

## 认知智能使AI触及生产核心，知识图谱决定认知智能的起点

人工智能与互联网或5G等技术本质上存在差别，后者解决的是信息的发出、传递、接收与反馈的闭环问题，而人工智能的本质是进行生产力升级，因此评判人工智能使用的是否有价值，要看其技术应用是否贴近生产核心。人类生产力可以归类为知识生产力和劳动生产力，人工智能走入产业后，可以分为感知智能、认知智能和行为智能，后两者更与生产力相对应。以计算机视觉、语音识别为代表的感知智能应用深度学习，在算力与数据的支撑下突破了工业红线，实现了机器对于自然界具象事物的判断与识别，但仅仅如此并没有触及核心生产环节，所以也就限制了其商业应用半径。当人们使用机器能识别更多事物的时候，自然而然的引发了，对事物背后的事理，理解、分析和决策的深层次需求，认知智能呼之欲出。认知智能核心解决的问题是对人类文明抽象概念的识别与联想，通过自然语言处理（NLP）技术对文字内容在语义上进行初步认知和自动抓取，经由知识图谱对概念间的关系属性进行联结、转换，从而对人类社会生产、生活行为进行描绘，实现业务规范梳理、生产流程可视化、人际关系挖掘等代表应用，这与注重经验、逻辑和方法论累积的知识生产力产生了直接对应，而与劳动生产力相对应的行为智能同样需要知识的指导，所以认知智能的发展才是人工智能回归本质的表征，而搭建知识图谱是认知智能可以参与生产的基础锚点。

### 人工智能与生产力的关系



来源：艾瑞根据公开资料自主研究绘制。

# 知识图谱概念定义

## 是一种描绘实体间关系的语义网络，是认知智能的必要基础

知识图谱是一种描绘实体之间关系的语义网络，是人工智能重要研究领域——知识工程的主要表现形式之一。知识图谱通过RDF（三元组），既“实体 x 关系 x 另一实体”或“实体 x 属性 x 属性值”集合的形式，以人类对世界认知的角度，阐述世间万物之间的关系，通过NLP技术、图计算、知识表示学习等手段，将非线性世界中的知识信息结构化，以便机器计算、存储和查询，起到赋予机器人类认知的效果，是人工智能技术走向认知的必要基础。

### 知识图谱以RDF形式阐述万物关系



来源：艾瑞根据公开资料自主研究绘制。

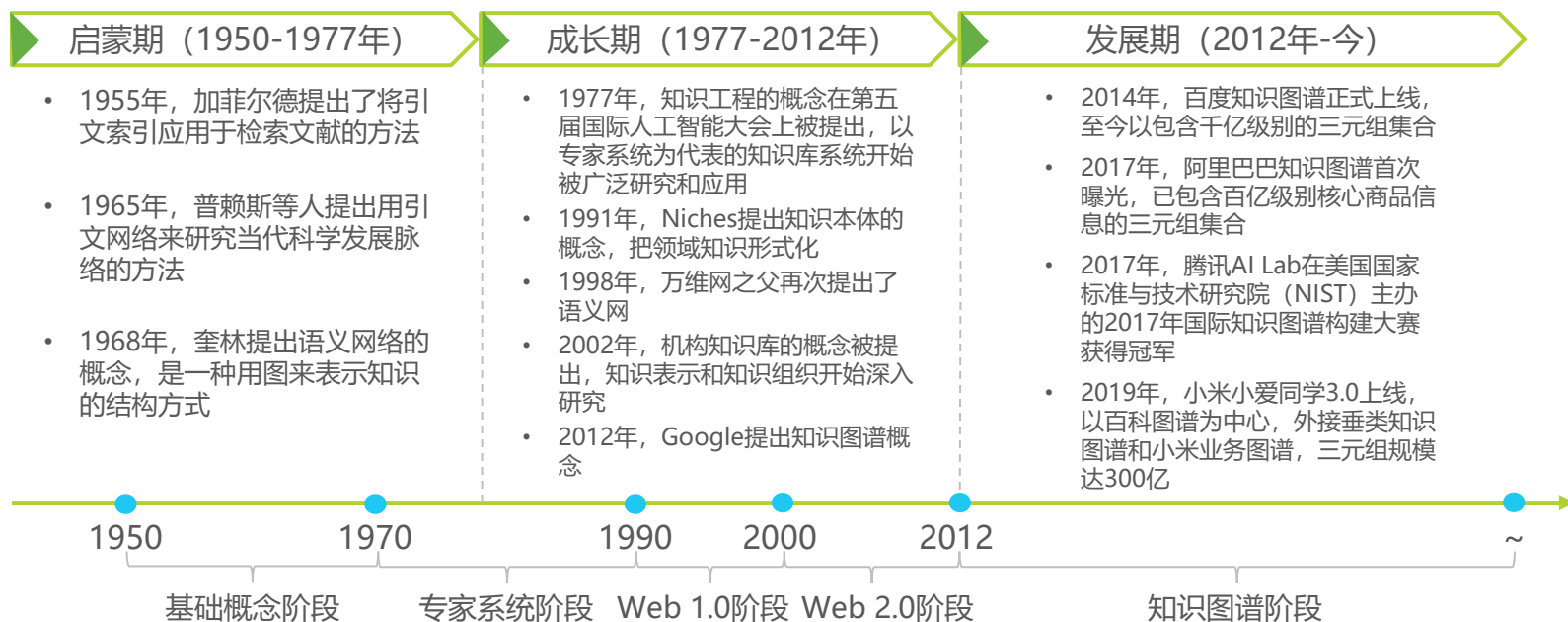


# 知识图谱发展历程

## 三个时期五个阶段奠定了现代知识图谱的雏形

知识图谱发展至今可以总结为三个时期和五个阶段。1950-1977年是启蒙期，包含了基础概念阶段和专家系统阶段的开端，这一时期文献索引的符号逻辑被提出并且应用；1977-2012年是知识图谱不断演变的成长期，包含了大部分专家系统阶段和Web 1.0和2.0阶段，在此期间出现了很多如WordNet、Cyc、Hownet等大规模的人工知识库，知识工程成为了人工智能重要的研究领域，2012年，Google正式提出的知识图谱概念，开启了现代知识图谱的序章；2012年至今是知识图谱的发展期，中国企业开始入局，以BAT为代表的科技公司依托自身业务，在搜索引擎、电商、医疗等领域开始应用知识图谱技术，解决办法服务商们也从大数据赛道中脱颖而出，将知识图谱技术拓展到安防、金融、教育等更多领域，让AI跳出感知智能的商业局限，向解决各产业生产环节中的核心痛点更进一步。

### 知识图谱的发展历程



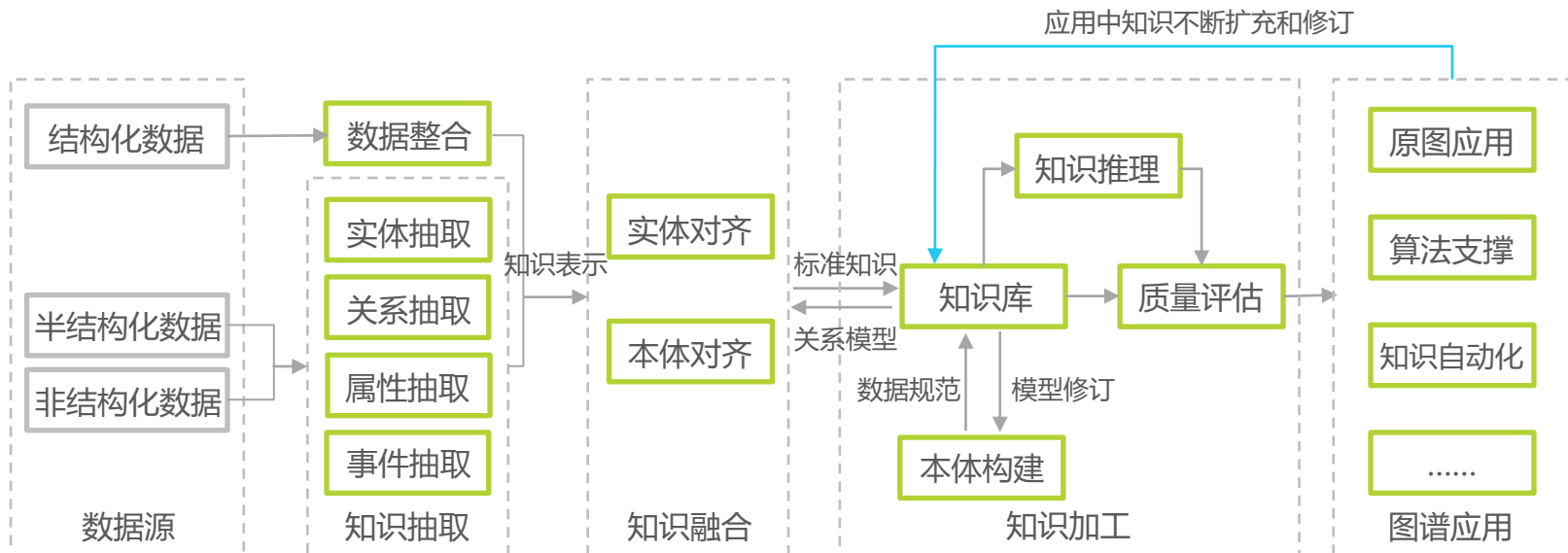
来源：艾瑞根据公开资料自主研究绘制。



## 建设本体模型和实体数据库是知识图谱结构的核心

知识图谱从逻辑上可以分为概念层和数据层，数据层指以三元组为表现形式的客观事实集合，而概念层是它的“上层建筑”，是经过积累沉淀的知识集合。建设中以本体模型和实体数据库为核心，根据二者的建设顺序又分为先定义本体和数据规范，再抽取数据的“自顶向下型”和先抽取实体数据，再逐层构建本体的“自底向上型”两种模式，前者适用于场景较为固定，存在可量化行业逻辑的领域，如金融、医疗、法律等；后者适用于新拓展的，有大量数据积累，行业逻辑难以直接展现的领域。总体而言，搭建知识图谱从数据源开始，经历了知识抽取、知识融合、知识加工等步骤。原始的数据通过知识抽取或数据整合的方式转换为三元组形式，然后三元组数据再经过实体对齐，加入数据模型，形成标准的知识表示，过程中如产生新的关系组合，通过知识推理形成新的知识形态，与原有知识共同经过质量评估，完成知识融合，最终形成完整形态上的知识图谱。

知识图谱结构拓扑图



来源：艾瑞根据公开资料自主研究绘制。

# 多源异构的数据来源

## 结构化和半结构化数据的抽取是搭建基础，非结构化数据抽取的发展将拓宽知识图谱的应用边际

知识抽取是搭建知识图谱过程中至关重要的一环，也是目前评判知识图谱优劣性的标准之一。知识抽取所处理的信息往往是明确的事实性信息，信息源通常分为已有关联数据库的结构化数据、如表格和列表类的半结构化数据，以及如TXT格式的文本类非结构化数据，前二者可以通过D2R（DRF格式转换器）、包装器（格式解析工具）等工具完成处理，而文本类非结构化数据需要通过自然语言处理的相关技术进行知识抽取，所有数据最终都要通过处理转化为标准数据供图谱使用。随着生物识别和物体识别等AI技术的应用，指纹库、人脸库、车辆库等数据库逐步建立，对于非结构化数据的知识抽取将不再局限于文本，所搭建的知识图谱维度将会更丰富，应用场景更加垂直下沉。

### 对信息源数据的抽取方法



# 大规模应用落地的保障——知识抽取

## 实体和关系抽取是核心，事件抽取是知识更新的重要手段

对信息源的知识抽取主要按照自然语义中的主谓宾语法进行抽取，分为实体抽取、关系抽取、属性抽取和事件抽取等。实体抽取指在信息源中识别出特定的元素标签，并与实体库中的标签相链接，是信息抽取中最基础的部分；关系抽取意在找到信息源中实体间的关系，可分为全局抽取和局部抽取，全局抽取是通过语料库对信息源中的所有关系对进行抽取，而局部抽取则是判断一句话中实体的关系类型，目前可以通过特征标注的有监督学习和借助外部知识库进行标注的远程监督学习实现，后者相比前者节省人工标注成本，但准确率略低；属性抽取可以看作实体和其属性值之间的关系，属于关系抽取的一种类型；事件抽取是将信息源中指定的事件信息抽取，并结构化地表现出来，包括事件的时间、地点、人物、原因、结果等，通常使用将事件划分多个分类阶段的pipeline方法和利用神经网络的深度学习方法，事件抽取拥有时间维度，可以与时俱进地迭代学习，是知识图谱知识更新的重要手段。

### 知识抽取核心内容



#### 实体抽取

1

- 从信息源中识别并提取实体
- 是信息抽取中最为基础和关键的部分
- 分为基于规则与词典、基于统计机器学习和面向开放域三种抽取方法



#### 关系抽取

2

- 从信息源中抽取实体间的关系
- 解决实体间语义链接的问题
- 分为有监督学习抽取和远程监督学习抽取



#### 属性抽取

3

- 对信息源中实体的特征和性质进行抽取
- 由于可以将实体的属性视为实体与属性值之间的一种名词性关系，因此也可以将属性抽取问题视为关系抽取问题



#### 事件抽取

4

- 从信息源中识别并抽取出事件信息，包括发生的时间、地点、参与角色以及与之相关的动作或者状态的变化
- 动态化地丰富和更新实体关系，是现有知识资源的重要补充

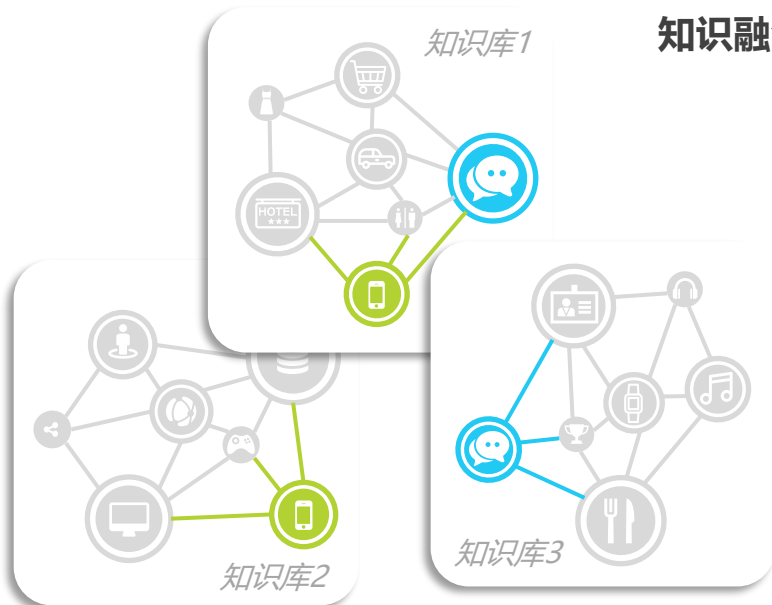
来源：艾瑞根据公开资料自主研究绘制。

# 快速工程化搭建的利器——知识融合

通过本体对齐和实体对齐实现知识库联结合并，形成更加统一、稠密的新型知识图谱

知识融合是指，从概念层和数据层两方面，通过知识库的对齐、关联、合并等方式，将多个知识图谱或信息源中的本体与实体进行链接，形成一个更加统一、稠密的新型知识图谱，是实现知识共享的重要方法。概念层的知识融合主要表现为本体对齐，是指确定概念、关系、属性等本体之间映射关系的过程，一般通过机器学习算法对本体间的相似度进行计算来实现，根据自然语言类型，可分为单语言对齐和跨语言对齐，其中跨语言对齐是实现知识国际交流的重要方式。数据层的知识融合主要表现为共指消解和实体对齐，前者意在将同一信息源中同一实体的不同标签统一，实现消歧的目的；后者是将不同信息源中同一实体进行统一，使信息源之间产生联结。知识融合的使用能够大量应用人类已有知识储备，节省成本，是快速搭建知识图谱的必要手段，也是现代知识图谱应用中重要的研究领域。

## 知识融合的核心内容



通过本体对齐和实体对齐形成更加统一、稠密的新型知识库



来源：艾瑞根据公开资料自主研究绘制。

# 由数据向知识转化的秘诀——知识加工

iResearch

艾瑞咨询

## 本体构建、知识推理和质量评估使RDF数据向“知识”跃迁

经过知识抽取和知识融合，实体和本体从信息源中被识别、抽取，并且消歧、统一，此时得到的关联数据是对客观事实的基本表达，但客观事实还不是知识图谱需要的知识体系，想要获得结构化的知识网络，还需要经过本体构建、知识推理和质量评估等知识加工过程。本体构建是知识图谱内实体连通的语义基础，以“点线面”组成的网状结构为表现形式，“点”代表不同实体，“线”代表实体间的关系，“面”既是知识网络。本体可以通过人工总结专家经验进行手动编程，也可以由机器学习驱动进行自动构建，本体构建的模型深度和广度，决定了知识图谱的应用价值；知识推理是通过对已有实体间关系的计算，找到新关联，从而丰富新知识的过程，也是知识图谱更新的重要手段；质量评估是知识加工最后的“质检”环节，确保经本体构建和知识推理得到的知识是合理，且符合知识图谱应用目的，根据所建设知识图谱的类型和具体用途，质量评估的标注有所不同。

### 知识加工核心内容

#### 本体构建

- 本体是知识图谱内实体连通的语义基础，主要以“点线面”组成的网状结构呈现
- 本体可通过人工编辑的方式手动构建，也可通过数据驱动自动构建，然后再经质量评估方法与人工审核相结合的方式加以修正与确认

#### 知识推理

- 通过对三元组的语义分析，获取新知识或结论，包含公理性推理和判断性推理
- 知识推理是知识图谱更新的重要手段
- 知识推理的主要方法包括基于Tableaux运算、基于逻辑编程改写、基于一阶查询重写、基于产生式规则等方法

#### 质量评估

- 对已产生的知识数据进行评估，将符合标准的数据导入知识图谱中
- 质量评估是确保知识图谱内容正确可用的关键步骤
- 根据所建设知识图谱的类型和具体用途，质量评估的标注有所不同

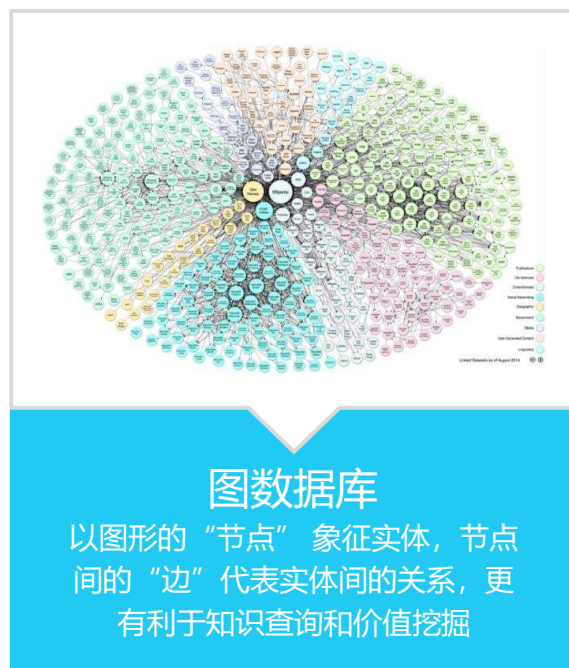
# 知识图谱的载体——知识存储

## 关系型数据库和图存储是现代知识图谱较为常见的存储方式

知识存储是将知识图谱的概念层和数据层以计算机可识别的数据格式进行保存的物理形式，是知识的现实容器。知识存储按照存储结构可以分为表格形式和图形式，常见的表格存储有三元组表、类型表和关系型数据库，前二者难于大规模建设，且查询、维护、删改等操作成本较高，不适用于现代知识图谱应用。关系数据库通过属性记录实体，能很好的和RDF完成映射，利用SQL语言也能方便快捷的进行知识检索，是目前应用较为广泛的知识存储方式；图存储是以图形结构对数据进行存储的方式，并非图片存储。以图形的“节点”象征实体，节点间的“边”代表实体间的关系，这种结构与知识图谱内本体结构非常贴合，更有利于知识查询和价值挖掘，常被应用于本体层的存储，但在某些场景中因其格式死板问题，并不适合全面应用。结合图谱使用目的，灵活的选择存储搭配方式，是符合当下知识图谱落地的最佳方案。



### 知识加工核心内容





# 知识图谱的主要算法——图计算

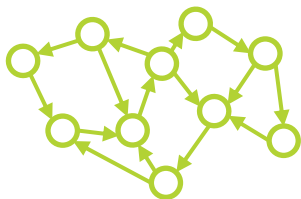
## 用于图结构化数据间的关联性推理运算，善于挖掘隐藏关系

图计算属于大数据挖掘技术的一种，专用于图结构化数据之间关联性的推理运算，基础数据格式与图存储相对应，由代表实体的“点”、代表本体的“边”和边上的权重组成，这里的权重多指一些包含时序性信息的具体数据。图计算算法主要包括全盘访问每一个节点的遍历算法、用于计算社交网络中人际关系的社区发现、源自搜索引擎，用于网页链接排序的PageRank，以及解决图结构中距离问题的最短路径算法，在知识图谱中主要应用遍历算法进行知识推理，以发现实体间隐藏的关系。为了更高效的进行图计算，在应用中会使用图计算系统和计算模型，图计算系统分为单机系统和分布式多机系统，对于大规模图结构存储的计算后者更具有优势。而图计算模型被用于解决图计算因频繁迭代带来的数据读取等待和成本高的问题、对相邻节点和边的计算依赖问题，以及难以在分布不均的区块上并行计算的问题。

### 图算法核心内容

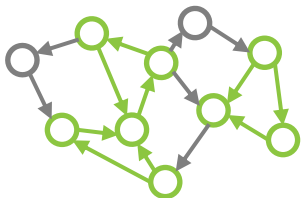
#### 遍历算法

沿着某条搜索路线，依次对树（或图）中每个节点均做一次访问，试图找到新的关联



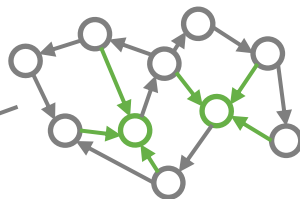
#### 社区发现

社区发现算法可以用来发现社交网络中三角形的个数（圈子），可以分析出哪些圈子更稳固，关系更紧密



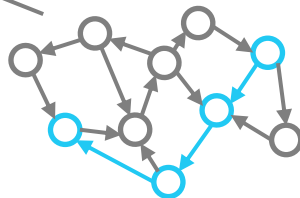
#### PageRank

如一个网页被多个网页链接，则其PageRank值较高，该算法源自搜索引擎中的网页排序



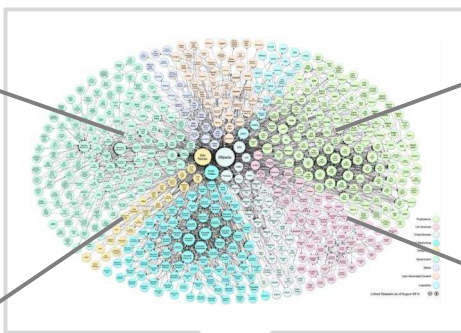
#### 最短路径

用于计算一个节点到其他节点间最短的途径，以目标节点为中心，向边缘扩散



### 图数据库

以图形的“节点”象征实体，节点间的“边”代表实体间的关系，更有利于知识查询和价值挖掘

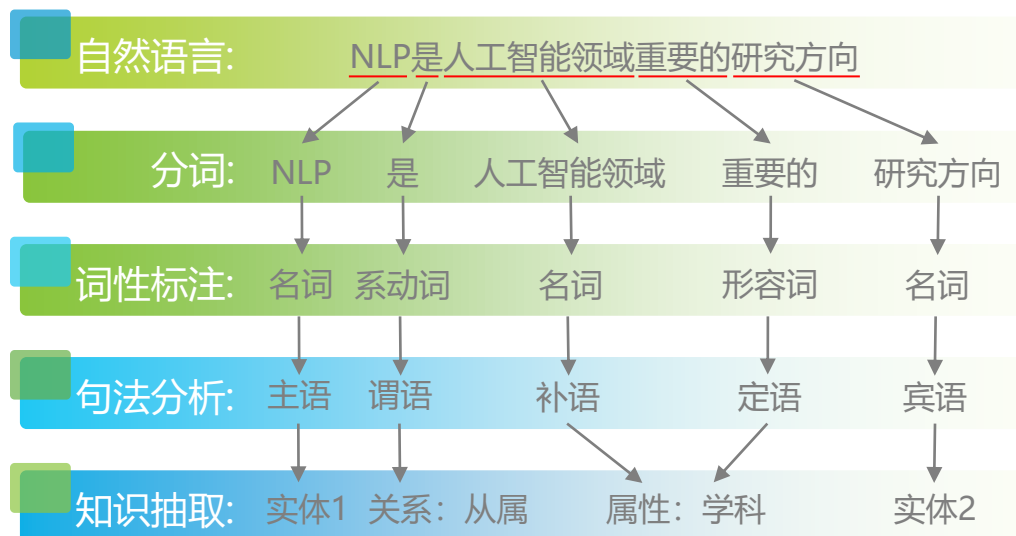


# 知识图谱的关键技术——自然语言处理 iResearch 艾瑞咨询

## 通过分词、词性标注和句法分析等方式使语言与程序相耦合

自然语言处理（NLP）是一门涉及语言学、数学和计算机科学的综合学科，主要研究人与计算机之间使用人类自然语言进行有效沟通的理论和方法。NLP技术包含非常广泛，最常应用于知识图谱中的能力是对于自然语言中信息的抽取，基本能力包括分词、词性标注和句法分析，通过词典法或统计法将文章中的词语进行分割，由最大熵、HMM或CRF等算法训练模型，对名词、动词、形容词、标点符号等词性类别进行识别和标注，再根据规则法或统计法对语句的主谓宾等句法结构进行分析，得到完整句式，最终产出一个计算机可识别的语料库，供后续需求调取。NLP技术是人工智能领域最具有挑战性的任务之一，2018年10月，谷歌公布的BERT模型在11项NLP任务中表现卓越，将下游具体NLP任务近一步推向预训练产生词向量环节，增强了NLP的泛化能力和自动化能力，为业界带来了新的思考方向，随着训练模型的不断探索与开源，类似深度神经网络于识别类任务的里程碑型训练算法终将出现，届时NLP训练成本将大大降低，从而真正走向产业化，开启认知智能的大门。

### 通过自然语言处理实现知识抽取的简单示意



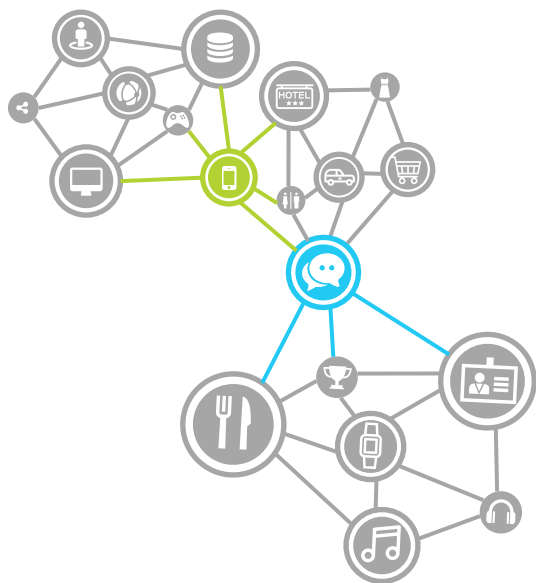
来源：艾瑞根据公开资料自主研究绘制。

# 业界前沿技术——知识表示学习

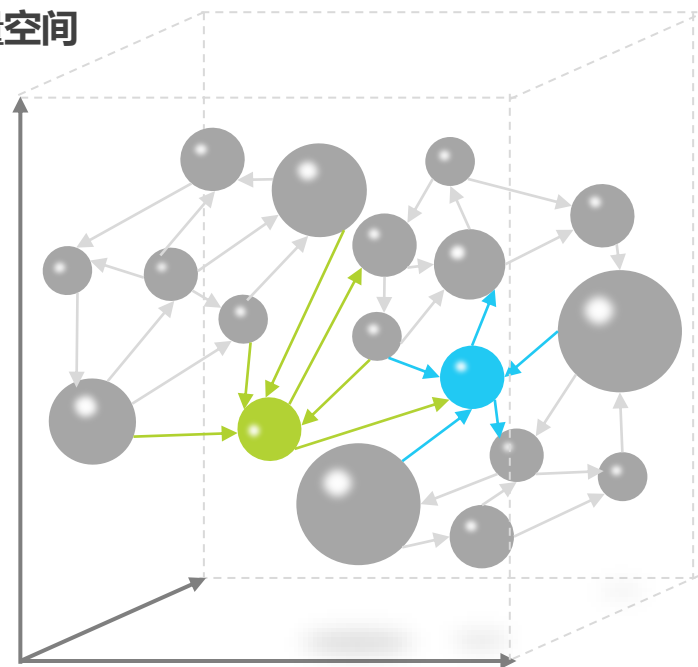
## 将知识映射到向量空间参与计算，用深度学习挖掘隐藏关系

表示学习是深度学习的核心环节，其本质是将研究对象的语义信息转化为稠密低维的向量，利用数学模型表示实例，并参与机器计算。知识表示学习是图神经网络（GNN）重要的应用表现，是面向知识库中概念和实体进行的表示学习，它的价值在于能够量化语义信息，进而计算概念间和实体间的相似度，实现关系抽取、实体对齐和知识推理的效果。知识表示学习的应用能够有效提升计算效率、降低数据稀疏带来的学习难度、解决多元异质数据的耦合问题等，是知识图谱技术的重点研究方向。目前知识表示学习的前沿研究仍以模型试验为主，以提高模型的预测准确性、可解释性、泛化能力为研究重心，使知识表示能够区分概念和实体，增强逻辑性是研究方向，随着知识表示学习的发展成熟，认知智能将更快得以实现。

### 将知识映射到低维向量空间



通过知识表示学习，将复杂的语义信息降维成表示向量，映射到向量空间中，参与计算



来源：艾瑞根据公开资料自主研究绘制。

知识图谱技术概述

1

中国知识图谱市场概述

2

中国知识图谱细分市场分析

3

中国数据智能代表企业案例展示

4

# 知识图谱的应用价值

## 对多源异构数据和多维复杂关系的处理与可视化展示是现阶段的主要价值，图神经网络下的深度学习算法应用将带来更广阔的认识市场

知识图谱是人工智能符号学派中知识工程的代表应用，其核心价值在于对多源异构数据和多维复杂关系的处理与可视化展示，其底层逻辑是将人类社会生活与生产活动中难以用数学模型直接表示的关联属性，利用语义网络和专业领域知识进行组织存储，形成一张以关系为纽带的数据网络，通过对关系的挖掘与分析，能够找到隐藏在行为之下的利益链条和价值链条，并进行直观的图例展示。在面对数据多样、复杂，孤岛化，且单一数据价值不高的应用场景时，存在关系深度搜索、规范业务流程、规则和经验性预测等需求，使用知识图谱解决方案将带来最佳的应用价值。连接主义中的深度学习算法几乎代表了当代整个人工智能技术，但深度学习需要具有明确因果关系的数据对训练，且存在尚缺乏解释性的黑箱问题，在掺杂众多非线性问题的复杂场景中应用价值有限，通过与知识图谱的配合使用，在一定程度上可以解决此类问题，随着关系向量法深入研究，图神经网络将走向产业应用，届时依托于行业知识与经验的深度学习将产生更多贴近产业核心的认知智能应用，人工智能技术也将跟进一步实现生产力升级的终极目标。

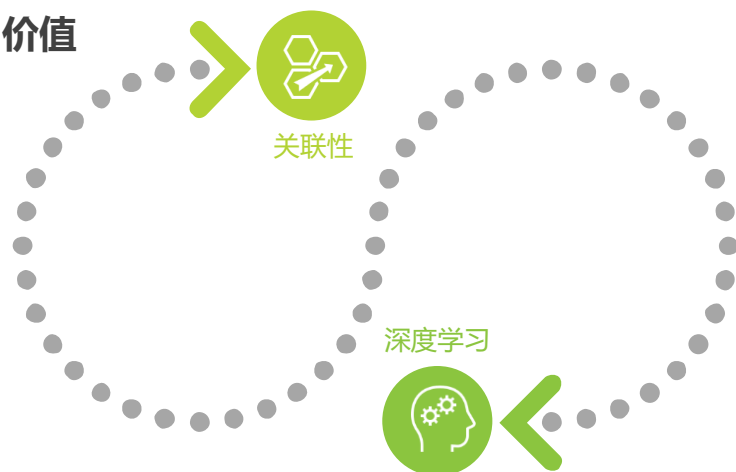
### 知识图谱的应用价值

#### 对多源异构数据和多维复杂关系的处理与可视化展示

将人类社会生活与生产活动中难以用数学模型直接表示的关联属性，利用语义网络和专业领域知识进行组织存储，形成一张以关系为纽带的数据网络，通过对关系的挖掘与分析，能够找到隐藏在行为之下的利益链条和价值链条，并进行直观的图例展示。

#### 图神经网络支撑深度学习算法应用

随着关系向量法深入研究，图神经网络将走向产业应用，届时依托于行业知识与经验的深度学习将产生更多贴近产业核心的认知智能应用，人工智能技术将跟进一步实现解放生产力的终极目标。



来源：艾瑞根据公开资料自主研究绘制。

# 知识图谱市场商业逻辑分析

## 原图应用解决查询类问题，算法支撑造就智能化解决方案

知识图谱作为一种更加形象的数据表现方式，其产品形式可以分为原图应用和算法支撑两类。原图应用是指直接通过图谱产生价值的服务形式，图谱根据概念层和数据层的区别可以分为通用知识图谱和行业知识图谱，通用知识图谱信息一般来自开放的互联网，三元组多为具有普适性的常识知识，知识覆盖的广泛性越强，价值越凸显，其特性更适用于如谷歌、百度、搜狗等百科型搜索引擎，被视为下一代搜索引擎的核心技术。而行业知识图谱则更看重具体场景中的认知深度，以及与行业know-how的结合程度，在此基础上实现的知识检索、隐藏关系挖掘和缺失数据补足，能很好地满足垂直领域知识类查询的需求，例如企业业务流程查询、司法领域案例查询、警务领域嫌疑人关系查询等；算法支撑是指通过知识图谱对于信息源的生数据进行处理，将产出的结构化关联数据用于算法模型训练和应用，得到能解决具体场景问题的研判建议，从而形成解决办法产生价值的服务形式。结合垂直行业B端市场的需求特点，由知识图谱作为算法支撑的智能解决办法具有更凸显的市场价值和更广阔的想象空间，用于投研投顾、辅助断案、业绩预测等偏向于认知智能的服务，是以知识图谱和算法能力见长的数据智能服务公司应有的核心竞争力。

### 知识图谱价值体现





来源：艾瑞根据公开资料自主研究绘制。



# 知识图谱在各领域中的应用概览

数据繁杂、单一价值有限、问题抽象需要可视化展现、五层关联维度以上的应用场景更加适合搭建知识图谱

知识图谱在各领域中的应用概览

	 行业知识库	 关联搜索	 预警应用	 研判应用	 推荐应用	 数据中台
金融领域	✓	✓	✓	✓	✓	✓
公安领域	✓	✓	✓	✓		
医疗领域	✓	✓	✓	✓	✓	
教育领域	✓	✓		✓	✓	✓
能源领域	✓	✓	✓	✓		✓
工业领域	✓	✓	✓	✓		✓
网络搜索领域	✓	✓			✓	
零售电商领域	✓	✓	✓		✓	✓
政务领域	✓	✓		✓	✓	✓
客服领域	✓	✓			✓	
营销领域	✓	✓		✓	✓	
媒体舆情领域	✓	✓	✓	✓		
企服领域	✓	✓			✓	✓
司法领域	✓	✓		✓	✓	

来源：艾瑞根据公开资料自主研究绘制。

# 中国知识图谱市场产业链战略分析

中游解决方案提供方进行全流程服务，是产业链核心环节，中国知识图谱市场产业链上游主要以各类数据源企业或机构为主，因所处理数据多具有敏感性和专属性，所以往往数据源也是服务的需求方；中上游主要包括数据采集服务商、数据库服务商、云服务商等，为产业链做支撑服务，但涉及保密数据采集加工的业务更多交由解决方案提供方直接处理；中游主要以提供知识图谱解决方案的大数据智能公司、互联网公司和AI公司为主，该部分是产业链的核心环节。其中大数据智能公司多脱身于大数据公司，凭借多年数据服务积累，在具体领域中具有先发优势和工程落地经验，随着市场对于认知智能需求的觉醒，近些年越来越多的大数据公司开始向知识图谱做技术拓展。互联网公司多以巨头企业为主，因其综合服务能力强、具有生态优势等因素，在资源调动和承接整体解决方案方面具有优势。AI公司主要以机器学习能力、NLP能力切入知识图谱市场，在技术创新和算法应用开发方面具有优势；下游为基于知识图谱而开发应用的各类具体场景应用，知识图谱应用在中国刚刚起步，主要聚焦于图中展示领域，随着市场推进将辐射更多行业。

中国知识图谱市场企业图谱



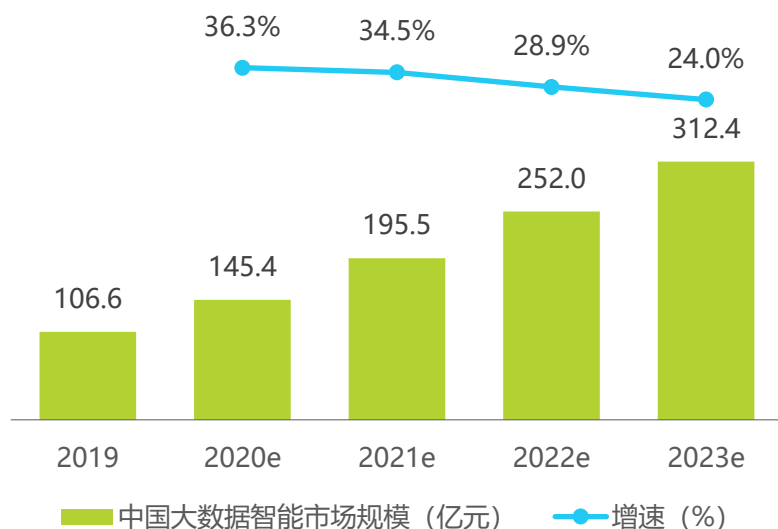
来源：艾瑞根据公开资料自主研究绘制。

# 中国大数据智能行业市场规模和结构

## 2019年市场规模约为106.6亿元，金融和公安领域占比最大

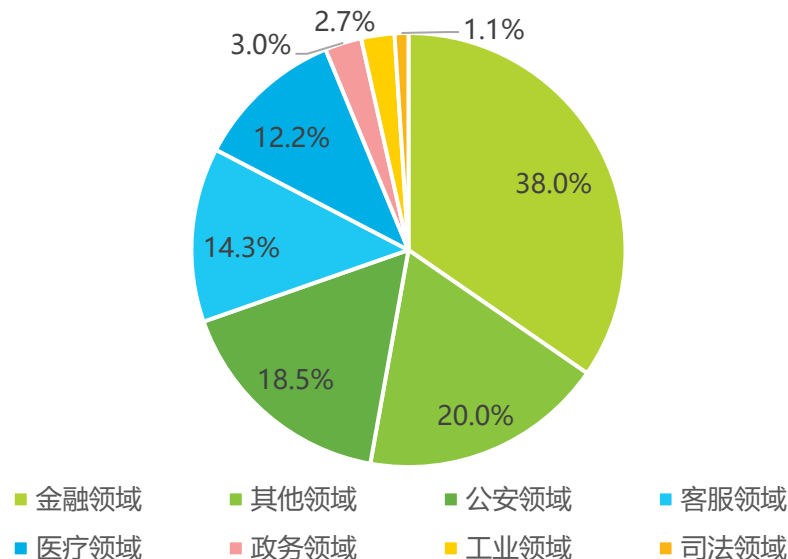
据艾瑞咨询统计推算，2019年涵盖大数据分析预测、领域知识图谱及NLP应用的大数据智能市场规模约为106.6亿元，预计2023年将突破300亿元，年复合增长率为30.8%，其中2019年市场中以金融领域和公安领域应用份额占比最大，金融领域因标准化数据积累丰富，行业认知与直接需求最为明显等因素，成为数据智能最早落地并产生价值的行业；公安领域经过三年的感知智能基础设施建设，已经初步形成人工智能应用环境，在党中央和公安部的号召指导下，公安大数据建设将成为下一阶段的主题，这一趋势也在2019年相关招标项目中得以体现。随着整体市场数据基础的完善和需求唤醒，大数据智能领域规模持续走高，但在行业可落地性和理性建设的限制下，预计市场增速将呈现下降趋势，期间咨询性需求将会大量出现，从整体发展来看增速处于良性区间，对真正有价值的公司和产品有正向意义。

### 2019-2023年中国大数据智能市场规模



来源：艾瑞根据专家访谈、招投标项目统计推算而得  
注释：文中大数据智能多指To B领域大数据分析预测、领域知识图谱、NLP应用等；C端应用和企业自研自用规模难以统计，暂未列入计算范围

### 2019年中国大数据智能市场规模细分结构



来源：艾瑞根据专家访谈、招投标项目统计推算而得  
注释：其他领域包括营销领域、零售领域、教育领域、学术领域等

知识图谱技术概述

1

中国知识图谱市场概述

2

中国知识图谱细分市场分析

3

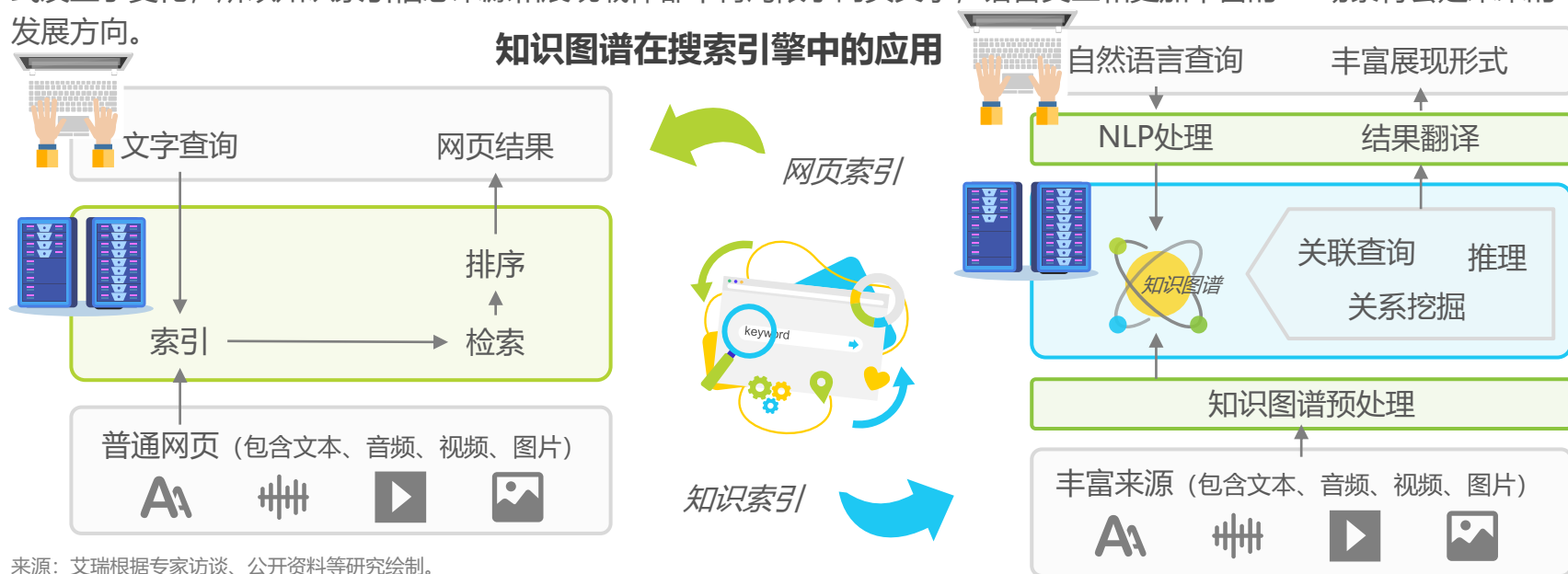
中国数据智能代表企业案例展示

4

# 知识图谱应用——搜索引擎领域

## 知识索引可将有价值的信息以自然语言的形式直接展现

搜索引擎是互联网发展中最具代表性的应用之一，它解决了人们如何方便获取信息的问题，一度成为大部分流量的入口。但随着网络中信息沉积越来越多，人们真正需要的，隐藏在大量无用信息中有价值的部分，我们称之为知识的内容，往往得不到直观展现，这成为了搜索引擎领域重点要解决的问题。2012年5月，谷歌率先提出知识图谱概念，用以更好的描述现实世界中实体关联性问题，进而提高信息搜索中的知识获取效率，随后搜狗、微软、百度纷纷推出相关概念架构，知识图谱被视为下一代搜索引擎的核心技术。传统搜索引擎一般采用网页索引，按照“网页 -- (预处理) --> 临时库 -- (索引) --> 索引好的库 -- (由用户行为触发检索) --> 为用户展示网页结果”的流程执行，信息来自网页，展现内容也是网页链接。而基于知识图谱的搜索作为和网页索引同级、并行的另一套知识索引，更注重信息间的关联性和自然语言理解，通过图存储的形式从新组织互联网中的信息，再以人类的语言习惯进行查询和展现，从而提高搜索体验，因为知识的存储形式发生了变化，所以知识索引信息来源和展现载体都不再局限于网页文字，语音交互和更加丰富的IoT场景将会是未来的发展方向。



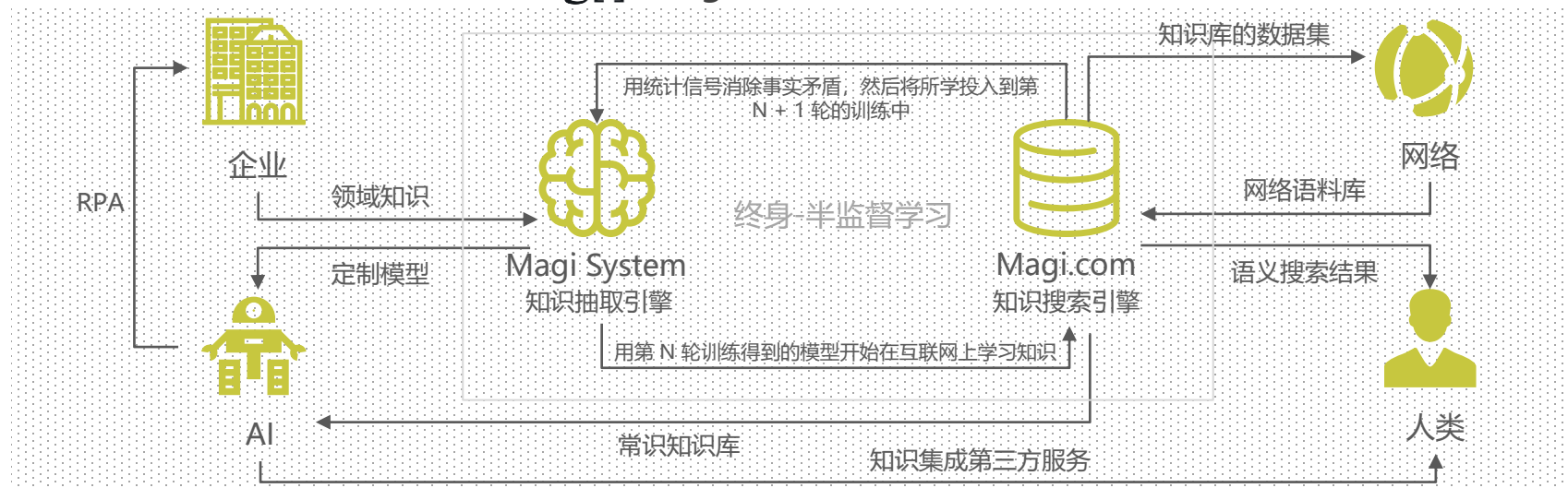
来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 知识图谱应用——搜索引擎领域

## 应用NLP的知识抽取将成为知识搜索的重要破局点

目前，互联网上只有极少数知识被人工整理成了机器可以解析的格式，如各类百科栏目和垂直领域数据库，这些资源支撑了当下绝大部分通用知识图谱的搭建，但这些信息也仅仅是沧海一粟，无论是覆盖范围、更新频率、可靠程度都无法满足日益增长的自动化和智能化需求，更严峻的是新一代互联网用户逐渐丧失了参与编辑的精神，甚至不知道可以编辑百科内容，在此趋势下基于百科的知识图谱将会变成一种静态数据留档，其支撑的知识搜索也将会大量依赖于权威组织编辑，成本和效率将会成为致命伤，因此应用准确且自动化的NLP技术，在更广泛的网页文本信息中，替代人工进行知识抽取，成为了知识搜索领域重要的破局点。知识搜索引擎Magi提供了一种NLP+评价系统的半监督学习方式，通过互联网对有监督学习的原始NLP模型进行“放养式”训练，再由“一个知识点在不同表达中被印证，则证明其可信度较高”而形成的评价系统去收敛训练成果，从而完成较低成本，可持续的广域下NLP模型训练，经过搜索引擎Magi.com的公测体验，以及B端知识抽取的定制服务，Magi逐渐形成了独立NLP产品的服务模式，为知识工程轻需求用户，和知识图谱搭建者提供了完善的RPA服务。

Mag[i] Magi引擎服务框架



来源：艾瑞根据专家访谈、公开资料等研究绘制。

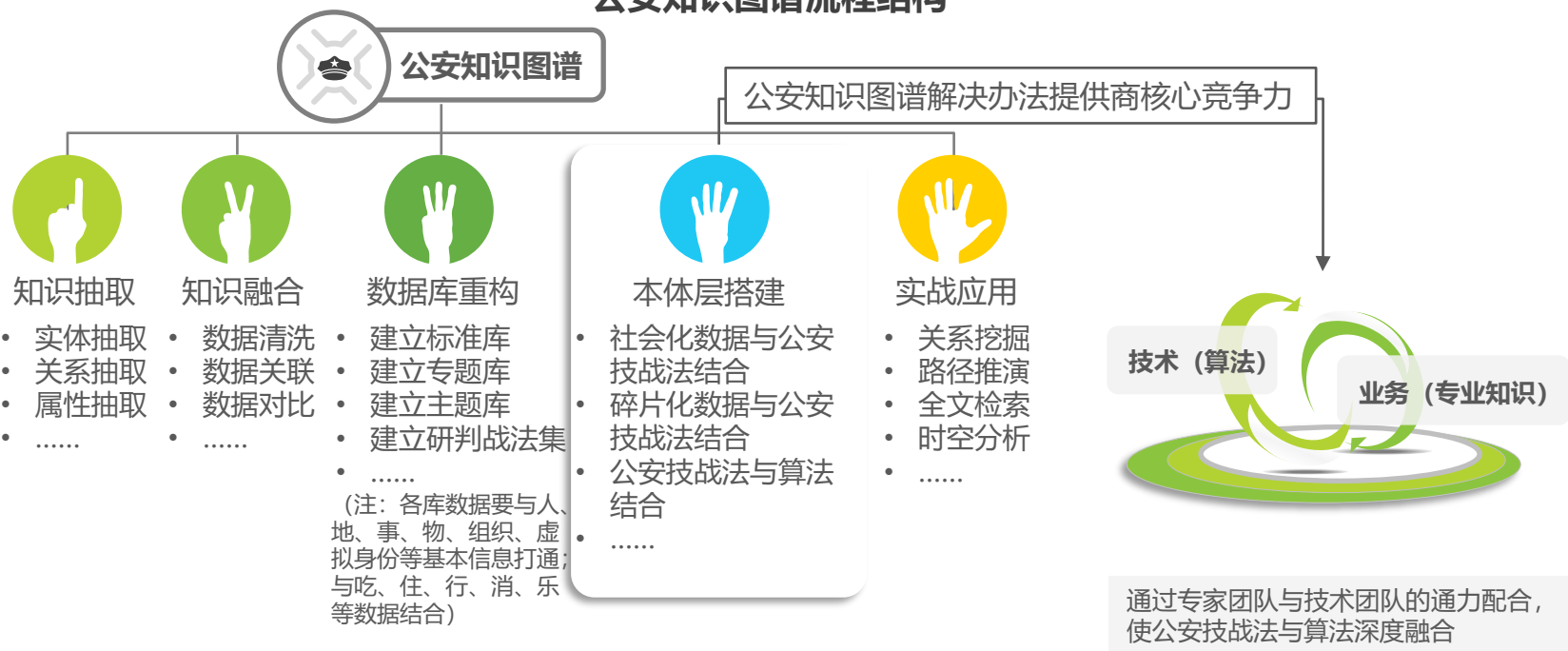


# 知识图谱应用——公共安全领域

## 公安知识图谱重点解决数据关联性和数据价值挖掘问题

公安大数据是全面助推公安工作质量变革、效率变革、动力变革的重要力量，受到党中央和公安部的高度重视，而作为大数据和人工智能双重技术的应用表现，知识图谱通过数据分析、文本语义分析等，抽取出人、物、地、组织机构、服务标识等实体，并根据实体的属性联系、空间联系、语义联系、特征联系等建立相互关联，构建一张具有公安特性的多维多层的实体与实体、实体与事件的关系网络，在解决公安大数据发展中面临的数据缺乏关联性问题时起到了重要作用。建设公安知识图谱仍然遵循知识图谱搭建逻辑，其中知识抽取、本体层建设和实战应用开发，需要将公安部门多年积累的实战经验与技术算法相互转换，重点考验公安知识图谱解决办法提供商对公安业务的理解能力和专业积累，是该行业竞争中重要的壁垒。

### 公安知识图谱流程结构



来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 知识图谱应用——公共安全领域

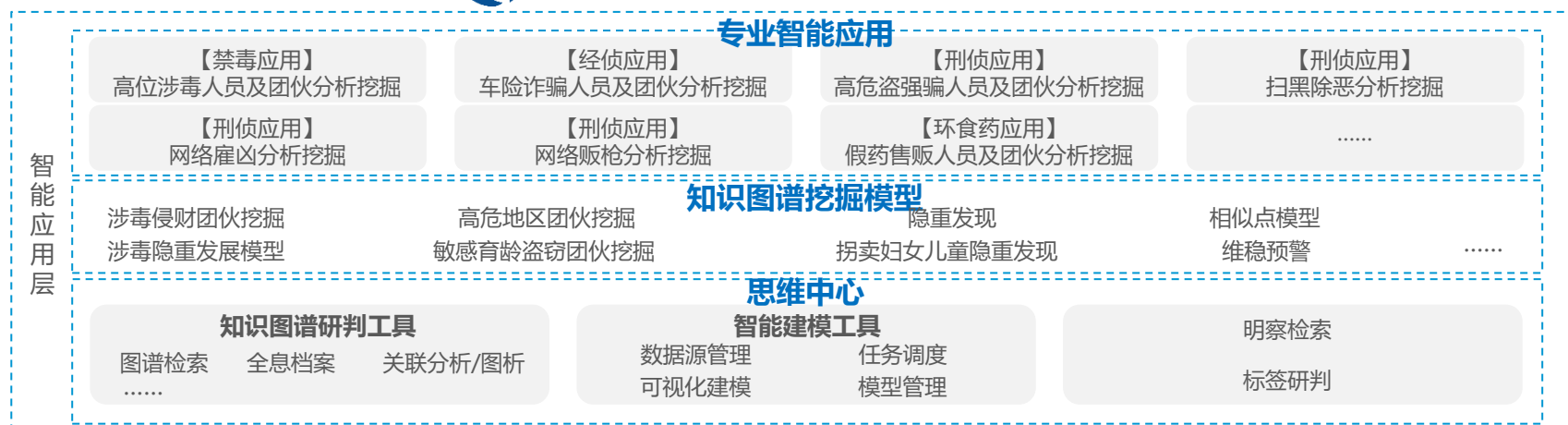
## 泛在感知加多模态数据融合应用将成为智慧警务的建设核心

近些年，在政策指导下公安大数据整体建设发展迅速，2019年感知智能硬件设备市场规模预计超350亿元，以人脸识别数据和车辆识别数据为基础的公安视图库加快建设和丰富，感知智能的成功应用不仅拓宽了公安领域结构化数据的边界，也唤起了认知智能应用的市场需求。随着公安部自上而下的推动，公安体系中各垂直系统的纵向业务开始拉通、各单位横向数据也开始分层解耦，进行初步融合，此举奠定了公安业务数据规范化和标准化的基础，也为感知智能与认知智能联动做好了准备，在十四五新一阶段的任务中，打通公安视频网和公安信息网的泛在感知加多模态数据融合应用必将成为建设重心。在此背景下，明略科技将海量感知数据纳入知识图谱的构建和积累中，从平台服务、数据服务到应用服务构建了由感知到认知再到行动的完整闭环，其智慧公安解决方案主要包括警务大脑平台、公安知识图谱平台和泛在感知应用实战平台等，提供公安知识图谱建设、知识图谱研判工具、知识图谱挖掘模型、警种专用智能应用、知识图谱服务接口等服务，目前已经部署到30多个省、市级公安局，据估计，明略科技公安知识图谱是30多个大类公安数据来源中的6529张表和1538亿条数据的结晶，沉淀了80%以上不同种类公安数据的处理经验，从常规业务到特殊警种办案，形成了基于AI技术的警务支撑体系和安全保障体系。



明略科技™  
MININGLAMP Technology

### 明略科技公安知识图谱解决方案



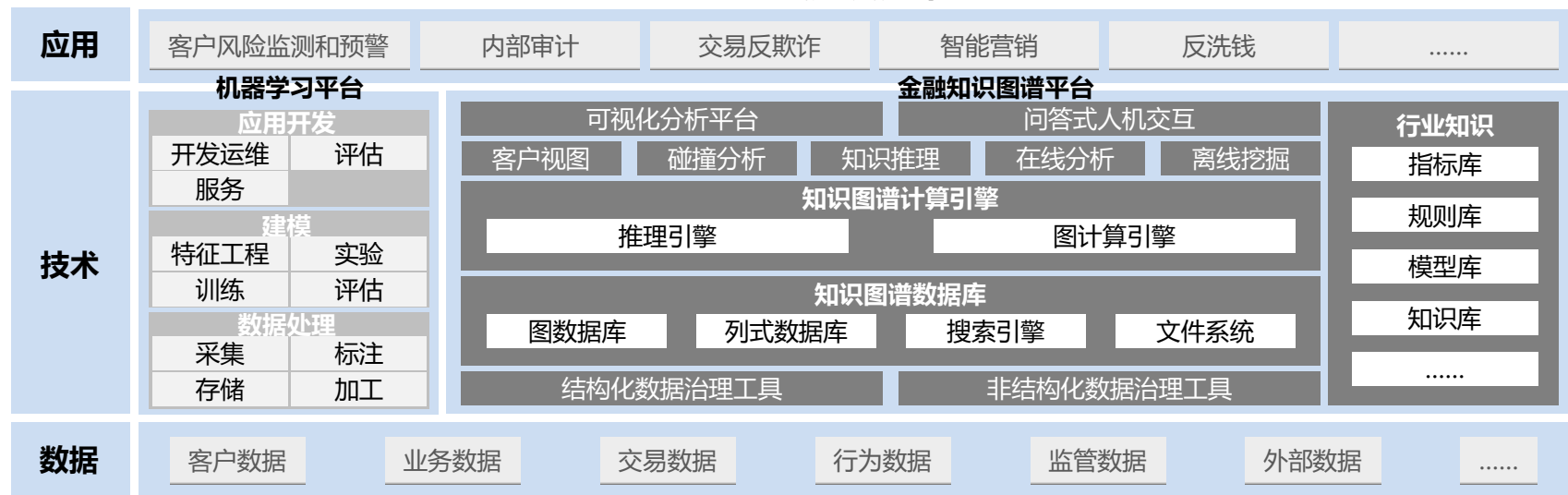
来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 知识图谱应用——金融领域

## 知识图谱与机器学习相结合的智能风控方案是主流趋势

在金融领域中无论是传统金融或是互联网金融，信用评估、反欺诈和风险控制都是最为关键的环节，随着近些年金融数据的爆发式增长，传统风控系统逐渐力有不逮，而应用机器学习算法和知识图谱的智能风控系统在风险识别能力和大规模运算方面具有突出优势，逐渐成为金融领域风控反欺诈的主要手段。机器学习和知识图谱相结合是目前主流的解决方案，其中机器学习算法通过概率计算的方式，以数学运算特征反应风险情况，形成易于机器计算的风控模型；而知识图谱通过权威经验和规则创建本体模型和抽取实体的范围，根据实体间关系形成关联数据网的图谱形式，描画囊括个人基础信息、金融行为、社交网络行为等用户综合画像，根据画像情况和模型对应，形成具有金融业务特性的风控体系，在解决方案的决策环节结合规则和概率的综合评价，给出最终的风险评估，整个过程能够实现秒级响应。知识图谱的应用不仅能够为缺乏可解释性的机器学习算法带来必要的参考系，还可以串联金融业务中产生的大量多源异构数据形成数据中台，挖掘数据深层价值，为实现精准营销、投资关系梳理、产业链风险预警、智能催收等上层应用打下基础。

### 智能金融解决方案



来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 知识图谱应用——金融领域

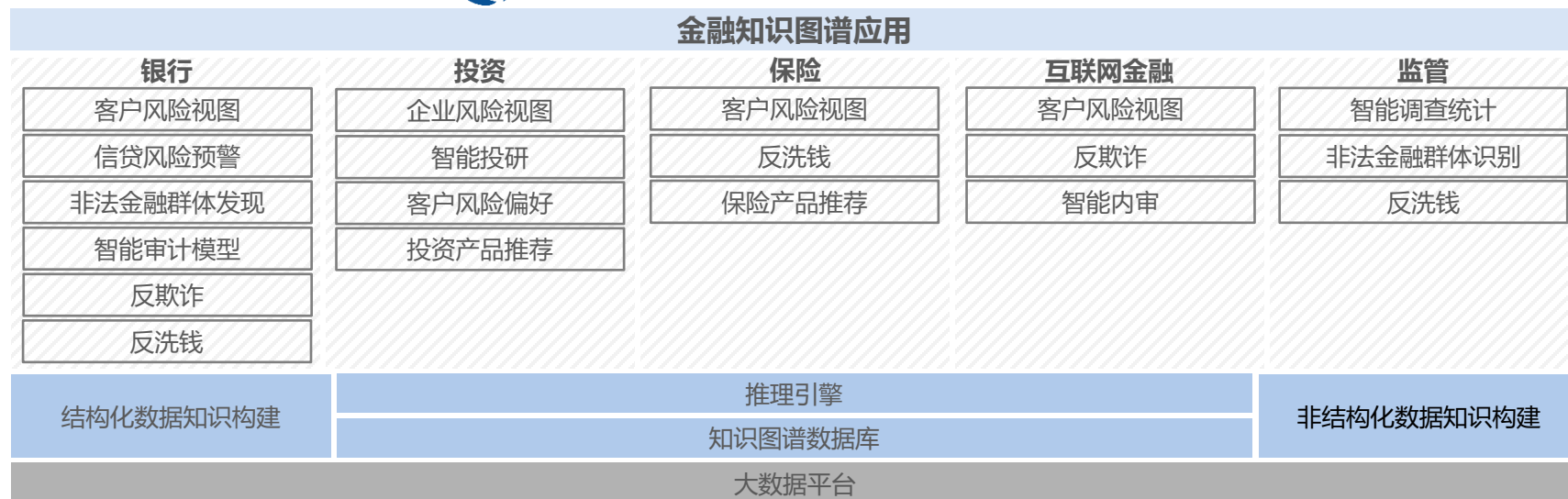
## 通过对大量多源异构金融数据的知识化整合形成中台式支撑

知识图谱在金融领域主要解决的问题是对多源异构数据的知识化整合。金融行业拥有海量包含各行业的数据信息，这些信息又以文字、表格、图形等形式存储在大量文档中，格式非标准统一且呈碎片化存在，而可用于风控反欺诈、信用评估、营销推荐、产业链分析等应用服务的数据又往往隐藏在多层关联下的细微处，因此采用集自然语言理解技术，处理非标数据，和多维多层次关系挖掘技术，展现数据关联性于一体的知识图谱应用，成为了金融领域较好的中台支撑形式。以场景最丰富、量级最大的银行业务为例，针对传统技术和手段难以实现的需求，明略科技为某全国股份制银行全行近十年的全量数据构建了包括“企业、个人、机构、账户、交易以及行为数据”在内，规模达十亿节点百亿边的知识图谱数据库，通过知识图谱平台建设来帮助该银行风控体系建立了完整的客户关系网及资金流转全貌，支持了该行非现场审计、系统性风险管控、精准营销等多项应用的研发和实施。



明略科技™  
MININGLAMP Technology

### 明略科技金融知识图谱应用场景



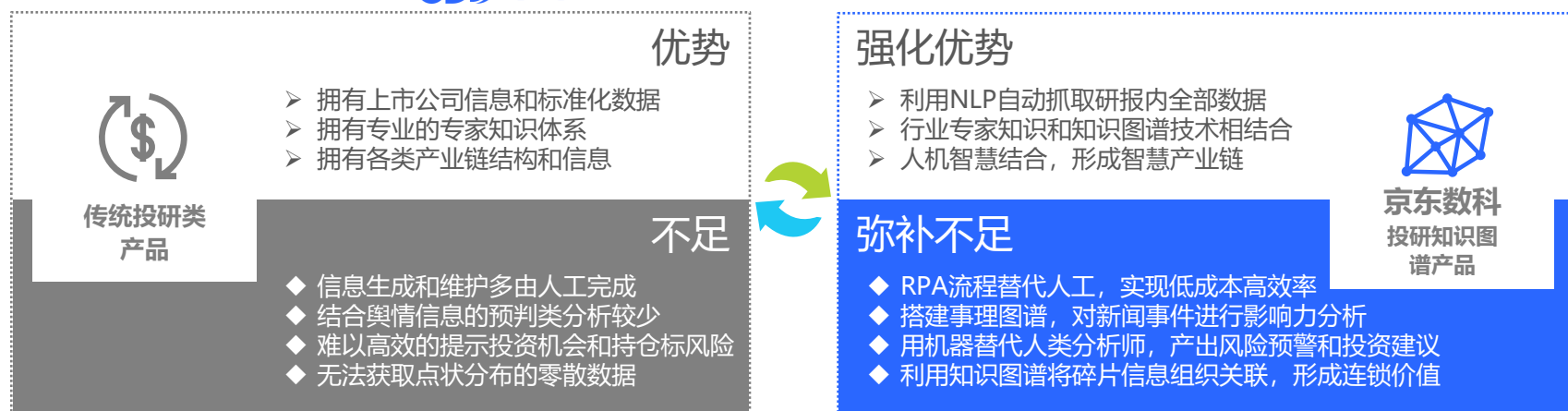
来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 知识图谱应用——投研领域

## 知识图谱通过自动抓取和产业链关系推理解决投研领域痛点

对一级市场或二级市场的投资研究，一直是泛金融领域重要的课题，上市公司或一些重要的标的公司在公开网络中披露了众多如财务数据、定期公告、公司研报等有价值的信息，为投资者行为提供了充分的参考依据。传统投研工作需要分析师通过各种渠道去搜集和判别信息，凭借个人经验对零散的数据进行组织建模，以报告的形式产出趋势观点和数据分析，大部分的物料和时间成本都花费在信息和数据的搜集上，而且成果可控性不高，纵使头部金融数据机构提供了相应的软件产品，但数据的颗粒度和产业链关联性仍难以满足多元需求，这成为了该行业长久存在的痛点。京东数科以公司信息、产业链关联为切入点，利用NLP技术自动抓取关键信息，搭建投研领域知识图谱，将各个行业的发展变化抽象导入数字层面，为知识查询和应用开发提供实现基础。据统计，券商研报中80%的数据指标在传统软件产品中无法查询到，分析师在进入一个新领域时要耗费一周左右的时间搜集类似数据，而利用京东数科投研产品可以将其必要劳动时间缩短到一分钟，大幅提高投研效率。除静态领域图谱外，京东数科还利用时间序列搭建了对网络报道、新闻事件进行抓取的事理图谱，两相结合，从行业固有逻辑和实时信息双管齐下，推导事情的发展脉络和趋势走向，为投研机构和投资者清晰的梳理关联脉络，为后续判断投资机会和持仓股票风险等研判类应用提供数据支撑。

### 京东数科 京东数科投研知识图谱产品特性



来源：艾瑞根据公开资料自主研究绘制。



# 知识图谱应用——司法领域

## 知识图谱是“智慧司法”建设的基础工程

司法领域是以公检法等国家司法机关及司法人员依照法定职权和法定程序，运用法律处理案件的专业领域。近年，司法领域积极运用大数据、云计算、人工智能等先进技术，深入业务场景解决痛点问题，有效提升办案质效、辅助司法管理、服务群众诉讼，加速推进司法智慧化、数字化、现代化转型升级。知识图谱的构建是实现智慧司法不可或缺的基础工程。司法知识图谱将法律领域中的实体、属性和关系进行体系化梳理，并建立逻辑关联，通过知识图谱和大数据技术进行数据挖掘，辅助决策，洞察知识领域动态发展规律。基于司法知识图谱，通过技术手段可实现司法业务场景的智能应用，解决“案多人少”“同案不同判”等现实问题。目前，司法知识图谱已广泛运用于法律知识检索和推送、文书自动生成、裁判结果预测预警、知识智能问答、数据可视化等方面，为司法人员办案提供高效参考和科学依据，全新定义司法数据应用和司法智能化，凝练司法智慧，服务法治建设。

### 司法知识图谱相关应用场景



#### 面向司法人员

- **法律知识检索**  
提供法律法条、典型案例、类案文书、专家观点、图书期刊等信息的精准检索，以及主动式智能推荐，为办案提供智能参考
- **文书自动生成**  
根据电子文书自动分析文书制作需求，选择模板自动生成法律文书，还能进行文书自动纠错，提高文书写作质量与效率
- **定罪量刑辅助**  
自动识别并归类法定、酌定情节，自动生成具体量刑建议，并对判案结果进行类案判决的偏离度预警，防止“同案不同判”发生



#### 面向人民群众

- **法律知识问答-强化普法教育**  
提供法律问题的专业知识查询和智能问答服务，自动匹配问题与答案，提供高效、便捷、权威的法律咨询，提升普法效果和力度
- **案情智能分析-提升调解成功率**  
提供案件纠纷相关的裁判规则、法律法规、类似案件判决结果对比，了解相似案件判决情况，帮助建立诉讼预期，从而服判息诉，提升纠纷调解成功率

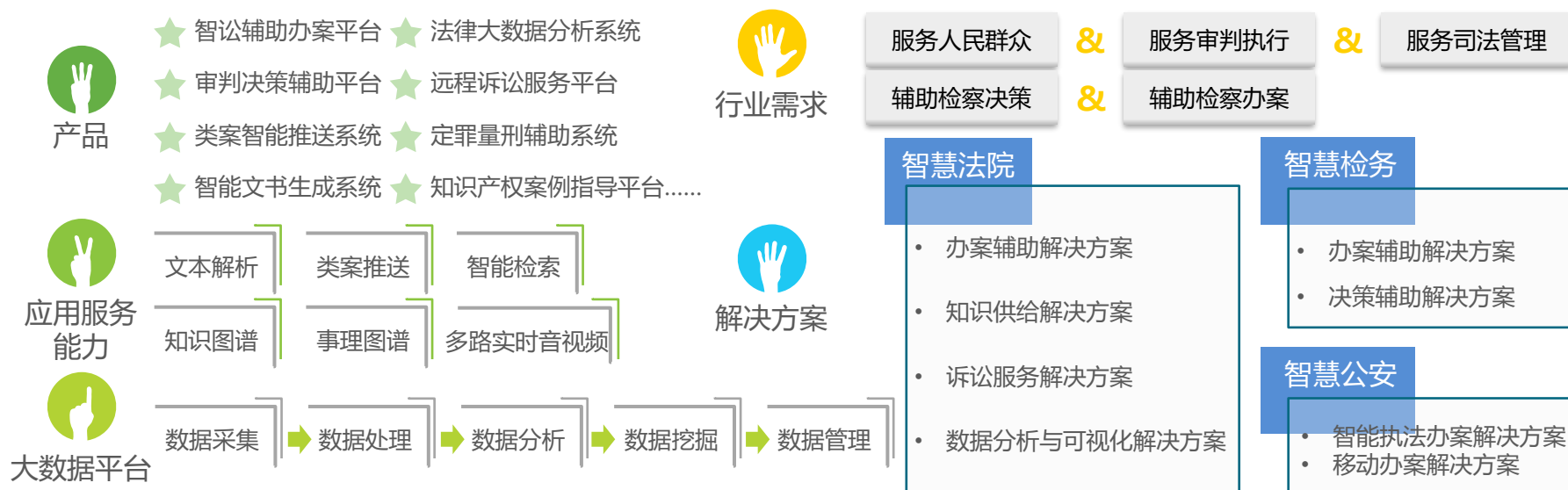


# 知识图谱应用——司法领域

## 司法领域数据治理以知识性为中心，适于知识图谱应用落地

面对大量且复杂的司法大数据，应用人工智能需要提炼共性规则，也就是依据不同的司法场景，提取具有统一性和相似性的领域规则并对抽取的数据进行数据标注，供机器学习以形成类似于人类的信息提取、逻辑分析能力。因此，与其说司法人工智能以数据为中心，不如说其以知识为中心，构建司法知识图谱是人工智能司法应用的先决性问题。目前司法知识图谱应用最广泛的是由人民法院出版集团、中国司法大数据研究院和国双共同研发建设的法律知识和案例大数据平台——法信，该平台主要解决法律人查找需求和海量知识数据供给的匹配问题，目前已覆盖全国30个省，3200家法院，为90%的法官提供全方位审判知识服务。此外国双还以大数据和人工智能技术为核心，以远程审判、语音办案、文书生成、类案推送、类案预警、量刑建议、审判规范化、决策支持、舆情监测等各项应用需求为“点”，以服务法院、检察院为“线”，以服务人民群众、审判执行、司法管理、助力司法研究、司法改革、司法创新为“面”，构建“点线面体”四维模式立体化的解决方案。

### GRIDSUM 国双 国双“智慧司法”一体化解决方案



来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 知识图谱应用——油气领域

## 知识图谱应用为油气勘探开发和安全环保生产提供决策辅助

目前，我国境内已探知的石油资源质量不高且产能有限，以至70%原油需求有赖于进口，加之国际原油价格持续走低，开采成本又很难摊薄，我国油气行业面临着生产效率与成本带来的巨大压力和挑战。在油气领域中，生产环节主要包括产业链上游的油气勘探、开发生产和中游的油气储运、炼制与加工，存在的痛点是上游如何将钻探数据、地质数据等“地下信息”与地面工程执行高效结合，从而准确探明更有开发价值的油井；中游如何利用大量自动化设备和传感器产生的数据进行高效的安全生产，以及节能环保问题。知识图谱凭借对多源异构数据关联性挖掘和知识体系信息化搭建等能力，在数字化程度较高、数据类型复杂的油气领域搭建认知网络，将领域知识与实时数据有机结合，为油气勘探、开发生产、综合研究、生产管理提供智能化分析手段，帮助决策者从海量的数据中洞悉规律，来提升效率和管理水平。

### 油气勘探开发数据及知识成果共享应用总体技术架构



来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 知识图谱应用——油气领域

## 知识专业性强来源众多且复杂，是知识图谱绝佳的应用场景

油气领域各专业是生产数据的来源，也是知识图谱构成的基础，因此，油气大数据分析与智能化应用的方向也应该从专业领域中找寻。在物探专业中，地震数据最为关键，但单一数据价值有限，需要结合多年采集的留存数据以及相关知识积累才能挖掘数据背后的规律，进行地震资料解释、地质反演与储层预测等工作；在测井专业中，测井解释需要将各类如电阻率、自然电位等物理参数经过关联性数据处理描述为地质信息，供后续勘探开发做必要参考，是寻找油气的核心环节；在钻录井专业中，随着物联网的普及，实时随钻录井数据成为生产监测重要的数据维度；在油气生产领域，根据各类实时数据采集，可以进行关联性分析与建模，产出实时动液面数据、实时产量数据和实时功图数据的关系模型；在油气藏模拟领域，对油田生产井历史数据进行关联性建模，是油田开发和油气生产动态分析的常用手段。以上均为知识图谱在油气领域中重点应用和研究方向，也是油气智能化发展的核心。

### GRIDSUM 国双 国双勘探开发数据及知识成果管理和应用

勘探应用		开发应用		生产应用		工程应用	
勘探区带前景预测		开发战略规划		油气生产优化		非常规井压裂酸化设计	
全球油气资源优选		重点气区开发决策支持		注采系统优化		油气井诊断与优化设计	
勘探部署规划		复杂油藏提高采收率		海外增产措施优选		重点工程配套技术研发	

面向应用的统一数据与知识共享服务中间件							
---------------------	--	--	--	--	--	--	--

知识服务中台DataS

知识应用	以业务/自然对象+索引标签为框架的知识应用																									
	知识检索		知识问答		知识类比		智能分析		知识聚合		可视化展示															
知识管理	勘探开发知识管理（知识分类体系、知识成果加载、知识智能标注、知识自动抽取、知识图谱构建、数据知识融合于同步、自然语言处理、文本识别...）																									
	盆地/区带/圈闭/油气藏		勘探部署		开发方案		提高采收率		开发生产优化		非常规油气		深层油气勘探		海相油气											
多源异构数据	勘探开发大数据管理平台（平台管理、数据接入、数据接入、数据处理、数据治理、数据安全、多维分析、智能分析、检索查询、图形识别、自然语言处理工具...）																									
	勘探库		开发库		生产库		实验数据		矿权储量		石油地质研究报告		资源评价研究报告		开发方案		非常规油气研究		归档资料		文献资料		专利		.....	

计算与存储平台IaaS	科学研究云				虚拟化/容器化				流计算Spark Streaming				分布式文件系统HDFS				分布式检索Elasticsearch			
-------------	-------	--	--	--	---------	--	--	--	--------------------	--	--	--	-------------	--	--	--	--------------------	--	--	--

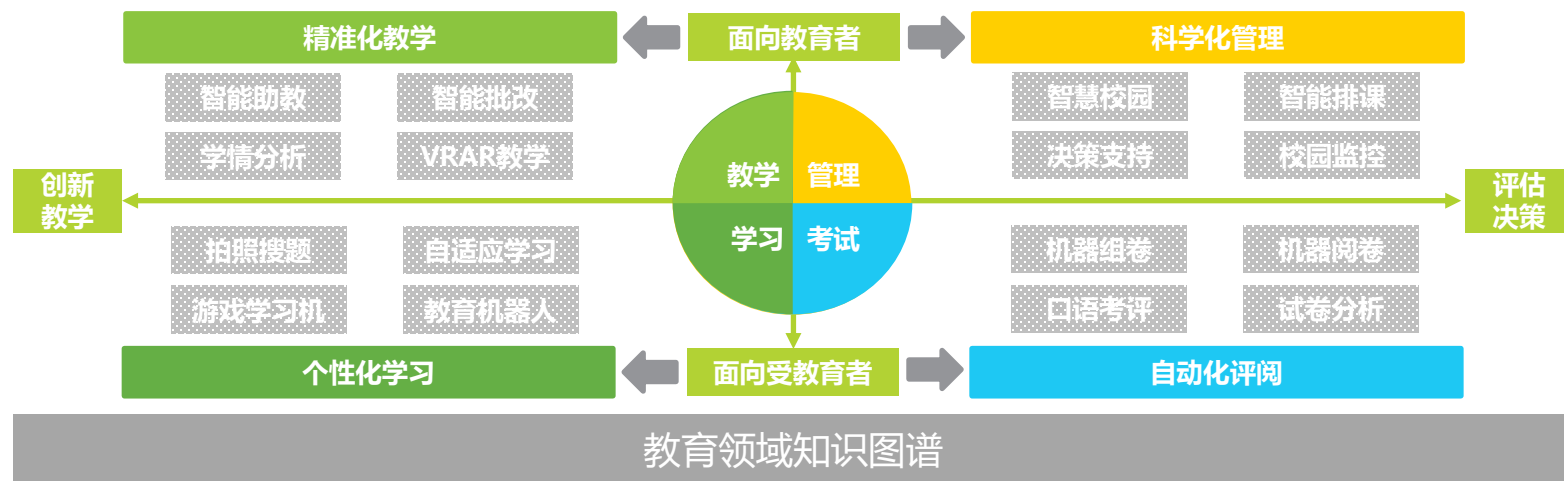
来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 知识图谱应用——教育领域

## 搭建知识图谱是AI技术落地教育领域核心环节的关键前提

当前的教育场景可以划分为教、管、学、考，主要围绕教育者和受教育者进行授课、答疑、阅卷和学习、练习、考试等活动，本质上是通过系统化的知识传授与强化练习，使受教育者掌握知识点的过程。传统教育模式以教师集中授课为主要方式，存在着特级教师资源少、优质教学资源分配不佳、教学内容和教学节奏制式化，以及作业练习布置缺乏个性化等问题。在教育信息化和线上教育发展普及之后，AI+教育的概念随之产生，人工智能公司和教育机构希望通过利用AI技术部分解决上述痛点，以达到对教育参与者减负增效的目的，但因为教育领域参与个体众多，采集到的数据驳杂且零碎，难以直接有效应用，所以目前AI技术更多应用在如拍照搜题、口语评测、课堂监控等外围需求的工具上面，并未能有效深入到教学场景中。而类似课题推送等应用更多是将教学资源再分配，重视知识的点状强化练习，却不深究知识的掌握程度，以至实际使用者反馈平平。人工智能技术真正产生生产价值，一定要建立在充分且必要的数据库基础上，搭建贯穿教材知识体系、教学资源管理和受教育者学习轨迹的知识图谱，将教与学的全过程进行可视化展现，使静态知识点数据与动态教学活动（如考试、作业等）数据产生关联，为算法应用提供支撑环境，是AI+教育能落地到产业核心的关键前提。

### 人工智能在教育领域中的应用场景



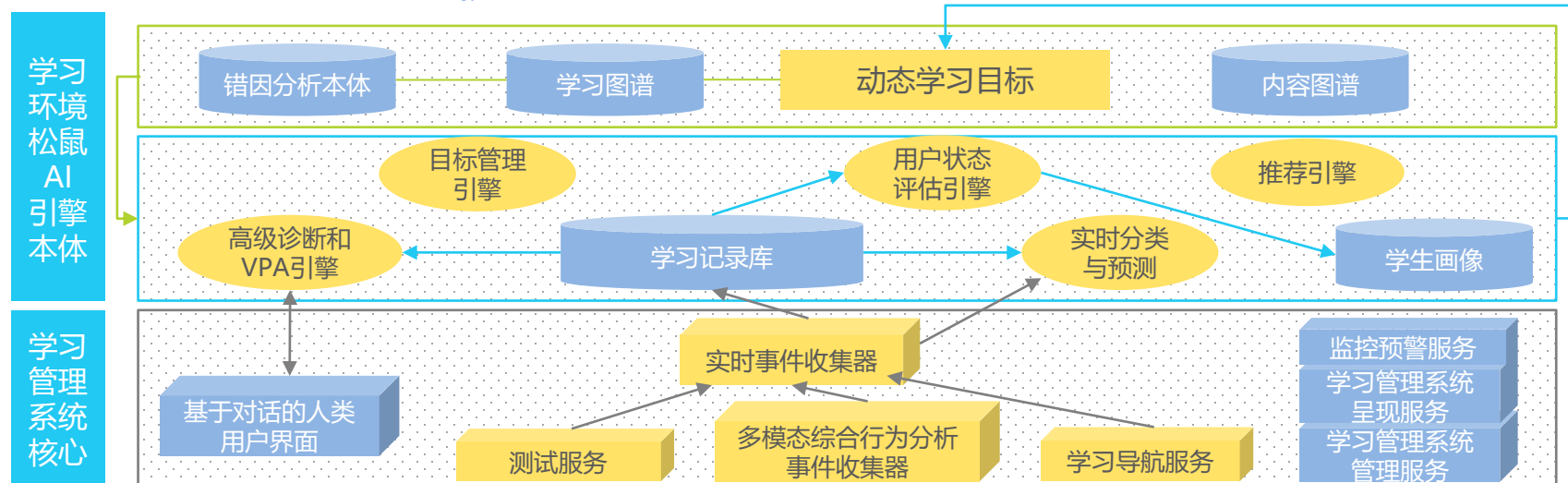
来源：艾瑞根据公开资料自主研究绘制。

# 知识图谱应用——教育领域

## 教育知识图谱为机器学习算法提供充分且必要的数据支撑

知识图谱在教育领域主要有以下几种应用场景：一是将学科教材知识进行本体建模，形成可关联性查询的知识网络；二是以图结构将教学资源以及关系进行语义化组织，以便合理调用；三是在知识图谱的基础上，应用大数据、AI等技术形成面向学习目标的个性化学习路径，实现千人千面的教学方案；四是面对受教育者搭建个人知识图谱，通过对其知识点学习进度和考试反馈数据的实时关联，形成知识掌握状态的可视化个人画像，以至于习题推送和老师一对一教学有的放矢；五是教育领域碎片化多源异构数据进行处理，形成标准化的关联数据集，为机器学习算法训练提供充要条件。通过以上五点应用，勾勒出基于知识图谱的数字中台形式，最大限度的对教育领域数据进行资源整合，为上层智能化应用提供支撑，改变了“传统教育披上人工智能外衣”的状况，用技术起底教育逻辑，形成数据指导下自适应学习的价值闭环。在自适应教学中，松鼠AI专注于学生画像与个性化学习路径的探索和应用，运用贝叶斯网络、深度学习等AI算法，通过对学生学习过程中多维数据的全方位评估，给予有针对性的学习内容、学习路径和学习节奏的推荐与把控，是国内具有代表性的智能教育机构之一。

松鼠AI 1对1 松鼠AI自适应教育技术框架



来源：艾瑞根据公开资料自主研究绘制。

知识图谱技术概述

1

中国知识图谱市场概述

2

中国知识图谱细分市场分析

3

中国数据智能代表企业案例展示

4



# 中国数据智能代表企业案例展示

明略科技致力于探索新一代人工智能技术在知识和管理复杂度较高的行业中进行更有效落地的方式方法。打通感知与认知智能，通过多模态人工智能和大数据技术，连接人、机器、组织的智慧，最终实现具有分析决策能力的高阶人工智能应用，让组织内部高效运转，让更多的人 and 资源投入到创新的工作中去。目前，明略科技已为各行各业垂直领域的2000多个组织，提供完整的人工智能闭环解决方案。基于数据中台的能力，连接行业和客户，形成智能化的应用，为公共安全、工业、金融、数字城市、智慧应急、数字卫生、智慧园区等领域的政企客户构建行业知识图谱，并逐步在诸多行业应用场景中实现了感知、认知和行动的连接。

## 明略科技多领域智能化解决方案框架



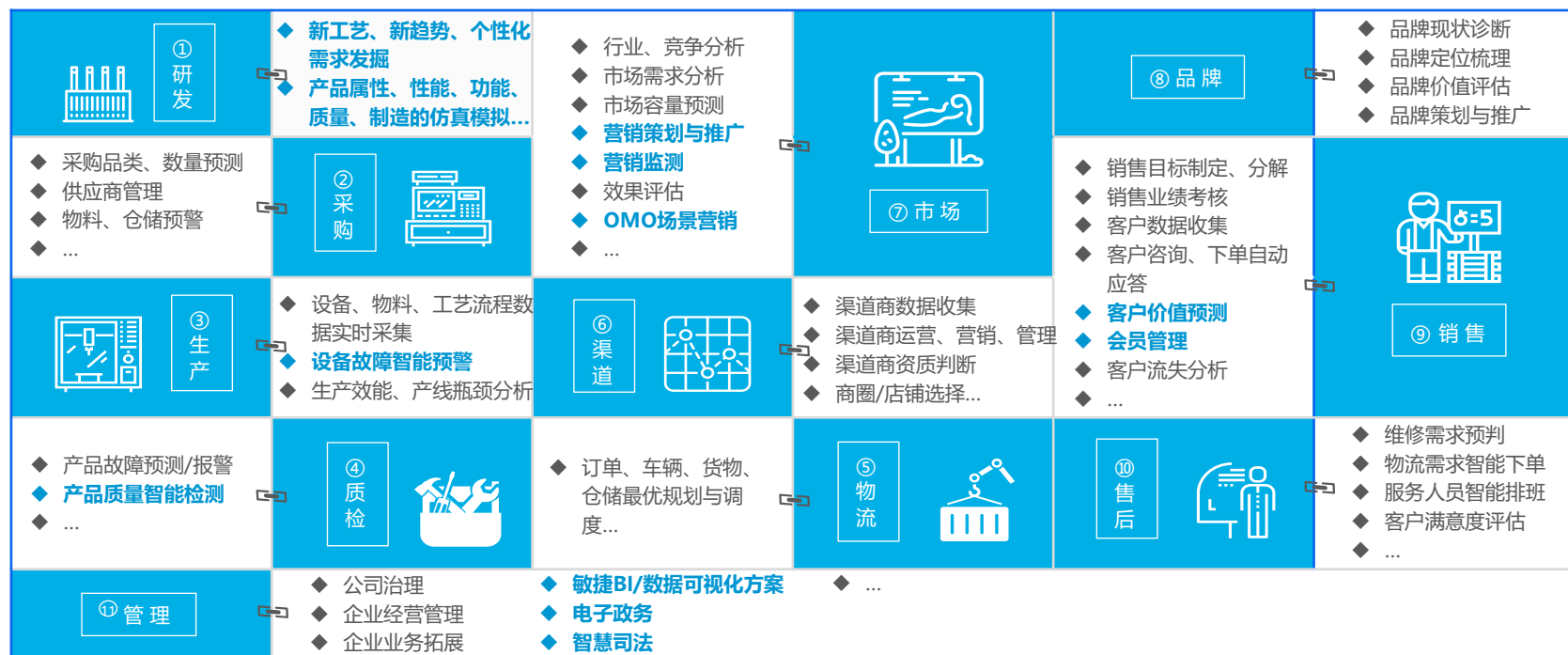
来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 中国数据智能代表企业案例展示

## 国双 GRIDSUM 国双

国双成立于2005年，于2016年9月23日，成功登陆纳斯达克，成为中国首家赴美上市的大数据企业，是一家企业级大数据和人工智能解决方案提供商，公司致力于实现“技术改变产业，促进社会发展”的愿景。国双数据科学团队成员共计30多人，涉及领域包括数据科学家、机器学习建模专家、计算机视觉专家、自然语言处理专家、语音工程专家等，目前数据科学应用主要服务能源、汽车、金融、快消、政府、司法等行业。基于大数据平台分布式数据架构和实时、多维度关联性分析技术，国双的解决方案能够高效采集并分析海量数据，使客户充分洞悉数据间的复杂关系，获得全新的商业洞察，以更好地作出商业决策。

### 国双企业数字化智能转型解决方案



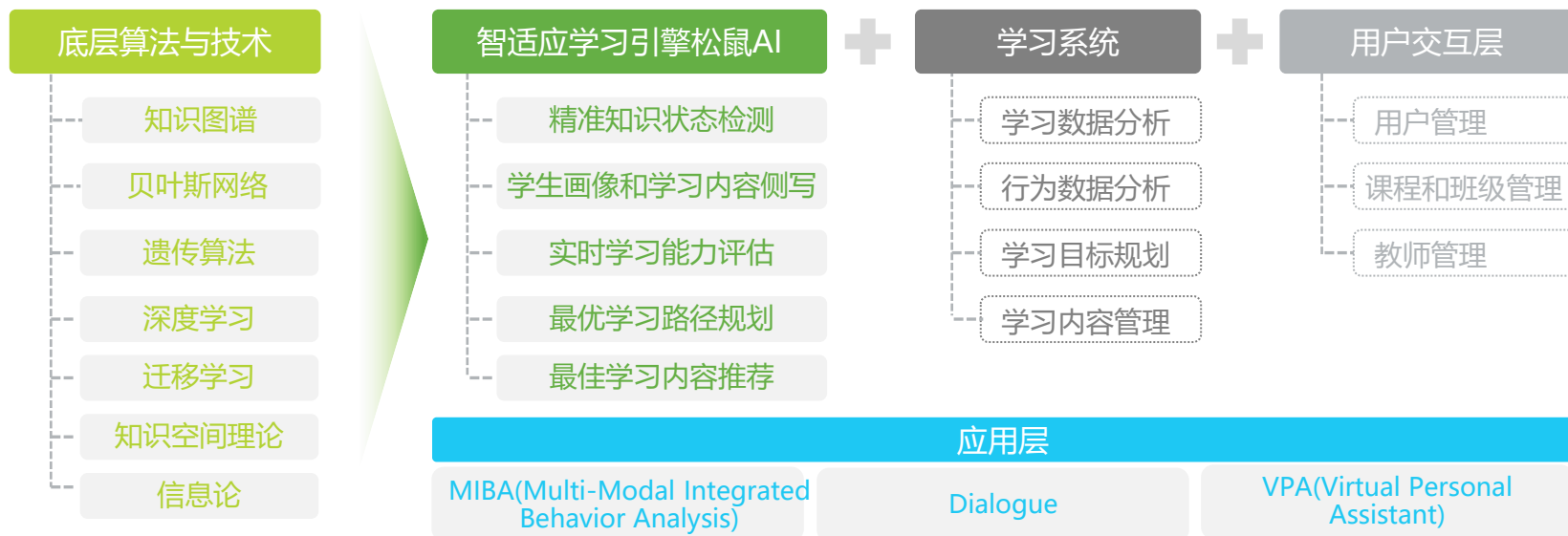
来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 中国数据智能代表企业案例展示

## 松鼠AI 松鼠AI 1对1

AI+教育的三个阶段包括第一层“信息智能+”、第二层“感知智能+”和第三层“认知智能+”，逐层递进尝试接近教育的终极理想——因材施教。在认知智能的努力上，专注自适应学习的义学教育是业内先行者之一：公司成立于2014年，是国内较早规模化针对K12学生提供人工智能自适应学习方案和服务的机构，结合自适应学习的理念以及人工智能和大数据技术，成功开发了拥有完整自主知识产权、以算法为核心的智适应学习引擎“松鼠AI”。松鼠AI目前注册用户数近200万，限定场景的教学效果测评结果可以比肩高级教师和特级教师，2019年松鼠AI旗下“慧眼”系统登台CCTV-1，该基于MCM（思想、能力、方法）的系统能够分析出每个孩子思维模式、学习能力和学习方法，通过人工智能来塑造和补足，发扬优势，补足短板，塑造更全面发展的自己。

### 松鼠AI智适应学习产品整体框架



来源：艾瑞根据公开资料自主研究绘制。

# 中国数据智能代表企业案例展示

## 京东数科 JDD 京东数科

京东数字科技集团是一家致力于以AI驱动产业数字化的新型科技公司，创立于京东集团内部，于2013年10月开始独立运营。京东数科以AI、数据技术、物联网、区块链等时代前沿数字科技为基础，建立并发展起核心的数字化风险管理能力、用户运营能力、产业理解能力和B2B2C模式的企业服务能力。截至目前，公司完成了在智能城市、数字农牧、金融科技、资管科技、数字营销、数字乡村、智能机器人等领域的布局，服务客户纵贯个人端、企业端、政府端。截至2019年6月，公司已累计服务涵盖4亿个人用户、800万线上线下小微企业、700多家各类金融机构、17000家创业创新公司、30余座城市的政府及其他公共服务机构。京东数科知识图谱技术目前支持了智慧产业链的落地，目前已经正式在JT<sup>2</sup>平台上线。另外，知识图谱还服务于京东金融App，支持各个业务线的销售对话机器人，包括白条，金条，小金库，基金等业务，为智能问答系统提供了推理问答，逻辑判断问答，查询问答等功能。

### 京东数科整体服务框架

#### 城市 x 数字科技

推出“城市操作系统”，以城市计算平台解决智能城市建设中数据结构化、标注化问题；不同场景需要定制开发应用问题；技术兼容性问题和隐私保护问题

#### 营销 x 数字科技

利用数字科技能力打造数字化场景的线下营销生态

#### 农牧 x 数字科技

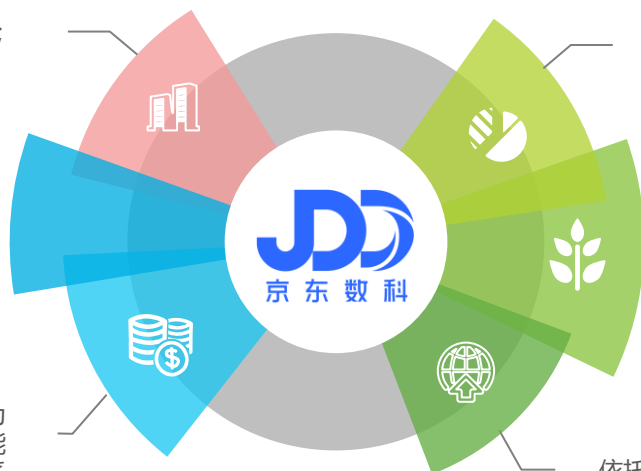
基于AI、物联网等前沿科技，为农业中的各细分行业提供智能解决方案，帮助农业提高数字化、互联网化和智能化水平

#### 金融 x 数字科技

利用关系网络、知识图谱、神经网络、行为序列、生物识别等技术构建智能风控、智能理财、智能保险科技、智能支付等服务体系

#### 创新 x 数字科技

依托于数字经济和产业数字化大潮，以创新手段不断突破边界、向各行业进行业务拓展



来源：艾瑞根据公开资料自主研究绘制。

# 关于艾瑞

在艾瑞 我们相信数据的力量，专注驱动大数据洞察为企业赋能。

在艾瑞 我们提供专业的数据、信息和咨询服务，让您更容易、更快捷的洞察市场、预见未来。

在艾瑞 我们重视人才培养，Keep Learning，坚信只有专业的团队，才能更好地为您服务。

在艾瑞 我们专注创新和变革，打破行业边界，探索更多可能。

在艾瑞 我们秉承汇聚智慧、成就价值理念为您赋能。

● 我们是艾瑞，我们致敬匠心 始终坚信“工匠精神，持之以恒”，致力于成为您专属的商业决策智囊。



扫描二维码  
读懂全行业

海量的数据 专业的报告



400-026-2099



ask@iresearch.com.cn

## 版权声明

本报告为艾瑞咨询制作，报告中所有的文字、图片、表格均受有关商标和著作权的法律保护，部分文字和数据采集于公开信息，所有权为原著者所有。没有经过本公司书面许可，任何组织和个人不得以任何形式复制或传递。任何未经授权使用本报告的相关商业行为都将违反《中华人民共和国著作权法》和其他法律法规以及有关国际公约的规定。

## 免责条款

本报告中行业数据及相关市场预测主要为公司研究员采用桌面研究、行业访谈、市场调查及其他研究方法，并且结合艾瑞监测产品数据，通过艾瑞统计预测模型估算获得；企业数据主要为访谈获得，仅供参考。本报告中发布的调研数据采用样本调研方法，其数据结果受到样本的影响。由于调研方法及样本的限制，调查资料收集范围的限制，该数据仅代表调研时间和人群的基本状况，仅服务于当前的调研目的，为市场和客户提供基本参考。受研究方法和数据获取资源的限制，本报告只提供给用户作为市场参考资料，本公司对该报告的数据和观点不承担法律责任。



# 为商业决策赋能

EMPOWER BUSINESS DECISIONS



艾 瑞 咨 询