

Chapter 6 知识图谱
(研究进展、现状&趋势)

任务定义、目标和研究意义

知识图谱 (Knowledge Graph, KG) 旨在描述客观世界的概念、实体、事件及其之间的关系

- 概念是指人们在认识世界过程中形成对客观事物的概念化表示，如人、动物、组织机构等
- 实体是客观世界中的具体事物，如篮球运动员姚明、互联网公司腾讯等
- 事件是客观事件的活动，如地震、买卖行为等
- 关系描述概念、实体、事件之间客观存在的关联关系
- 以结构化的形式描述客观世界中概念、实体间的复杂关系，将互联网的信息表达成更接近人类认知世界的形式，提供了一种更好地组织、管理和理解互联网海量信息的能力。
- 给互联网语义搜索带来了活力，同时也在智能问答中显示出强大威力，已经成 为了互联网智能服务的基础设施
- 知识图谱与大数据和深度学习一起，已经成为推动人工智能发展的核心驱动力之一。

知识图谱技术

- 建立知识图谱中使用的技术，是融合认知计算、知识表示与推理、信息检索与抽取、自然语言处理与语义 Web、数据挖掘与机器学习等的交叉研究。
- 探索从互联网语言资源中获取知识的理论和方法
- 促进知识驱动的语言理解研究
- 研究从大数据中挖掘隐含的知识理论与方法，将大数据转化为知识，增强对互联网资源的内容理解
- 将促进当代信息处理技术从信息服务向知识服务转变

知识图谱应用

- 知识融合 — 当前互联网大数据具有分布异构的特点，通过知识图谱可以对这些信息资源 进行语义标注和链接，建立以知识为中心的资源语义集成服务
- 语义搜索 — 将用户搜索输入的关键词，映射为知识图谱中客观世界的概念和实体，搜索结果直接显示的满足用户需求的结构化信息内容，而不是互联网网页
- 问答系统 — 基于知识的问答系统将知识图谱看成一个大规模的知识库，通过理解将用户 的问题转化为对知识图谱的查询，直接得到用户关心问题的答案
- 大数据分析 & 决策 — 知识图谱通过语义链接可以帮助理解大数据，获得对大数据的洞察，提供决策支持。

研究内容、技术方法&研究现状

研究内容

- 知识表示
 - 研究客观世界的知识如何在计算机里表示和处理
 - 需要解决的问题
 - 建立什么样的知识表示形式能够准确地 映客观世界的知识
 - 建立什么样的知识表示可以具备语义表示能力
 - 知识表示如何支持高效知识推理和计算，从而使知识表示具有得到新知识的推理能力
 - 主要知识表示技术
 - 符号主义 — 人类认知和思维的基本单元是符号，而认知过程就是在符号表示上的运算
 - 联结主义 — 人的认知就是相互联系的具有一定活性值的神经单元所形成网络的整体活动，知识信息不存在于特定的地点，而是在神经网络的联结或者权重中
- 当前主要知识表示方法
 - 传统人工智能中基于符号逻辑的知识表示
 - 如：产生式系统、谓词逻辑、框架表示、语义网等
 - 逻辑表示法（最早使用，如 一阶逻辑、描述逻辑），产生式表示法和框架表示等
 - 缺点：生成规则的能力较弱，对数据质量要求较高，很少使用
 - 万维网内容的知识表示
 - 互联网资源的开放知识表示方法，如 XML（为内容置标）、RDF（通过三元组（主体，谓词，客体）描述）和 OWL（构建在 RDF 之上，是具有更强 力的语言）等
 - 在工业界得到大规模应用的基于三元组的知识图谱知识表示方法
 - 表示学习
 - 基于知识图谱的表示学习通过深度学习可以将知识表示成低维连续实值稠密的 向量空间，有助于实现高效的知识计算
 - 主要包括张量重构和势能函数的方法
 - 张量重构综合整个知识库的信息，但在大数据环境下张量维度很高，重构的计算量较大
 - 势能函数方法认为关系是头实体向尾实体的一种翻译操作
- 知识图谱构建
- 知识图谱应用
 - 研究如何利用知识图谱更好地解决实际问题
 - 终极目标：利用知识图谱，建立基于知识的系统并提供智能的知识服务
 - 基于知识的互联网资源的信息融合
 - 语义搜索
 - 利用具有良好语义定义的形式，以有向图的方式提供满足用户需求的结构化 语义内容。主要包括 RDF 和 OWL 的语义搜索引擎和基于链接数据的搜索等
 - 直接得到用户感兴趣的客观世界的实体和实体关系信息
 - 典型应用：谷歌的 Knowledge Graph 和国内的百度知心、搜狗的知立方等
 - 基于知识的问答系统
 - 通过对问句的语义分析，将非结构化问句解析成结构化的查询语句，在已有结构化的知识库上查询答案
 - 通过将用户的提问转换成对结构化知识图谱的查询可以 直接得到用户的答案，其中问题理解和基于推理的知识匹配是核心技术
 - 进展：近两年很多研究者开始研究基于深度学习的知识 库问答方法
 - 基于知识的大数据分析和挖掘 — 研究语义标注或者实体链接技术，实现不同资源类型、不同媒体类型的互联网资源的融合、管理与服务

总结&展望

知识图谱技术是知识表示和知识库在互联网环境下的大规模应用，显示出知识在智能系 统中重要性，是实现智能系统的基础知识资源

未来发展方向

- 融合符号逻辑和表示学习的知识表示
- 高精度大规模知识图谱构建
- 知识图谱平台技术
- 基于知识图谱的应用研究