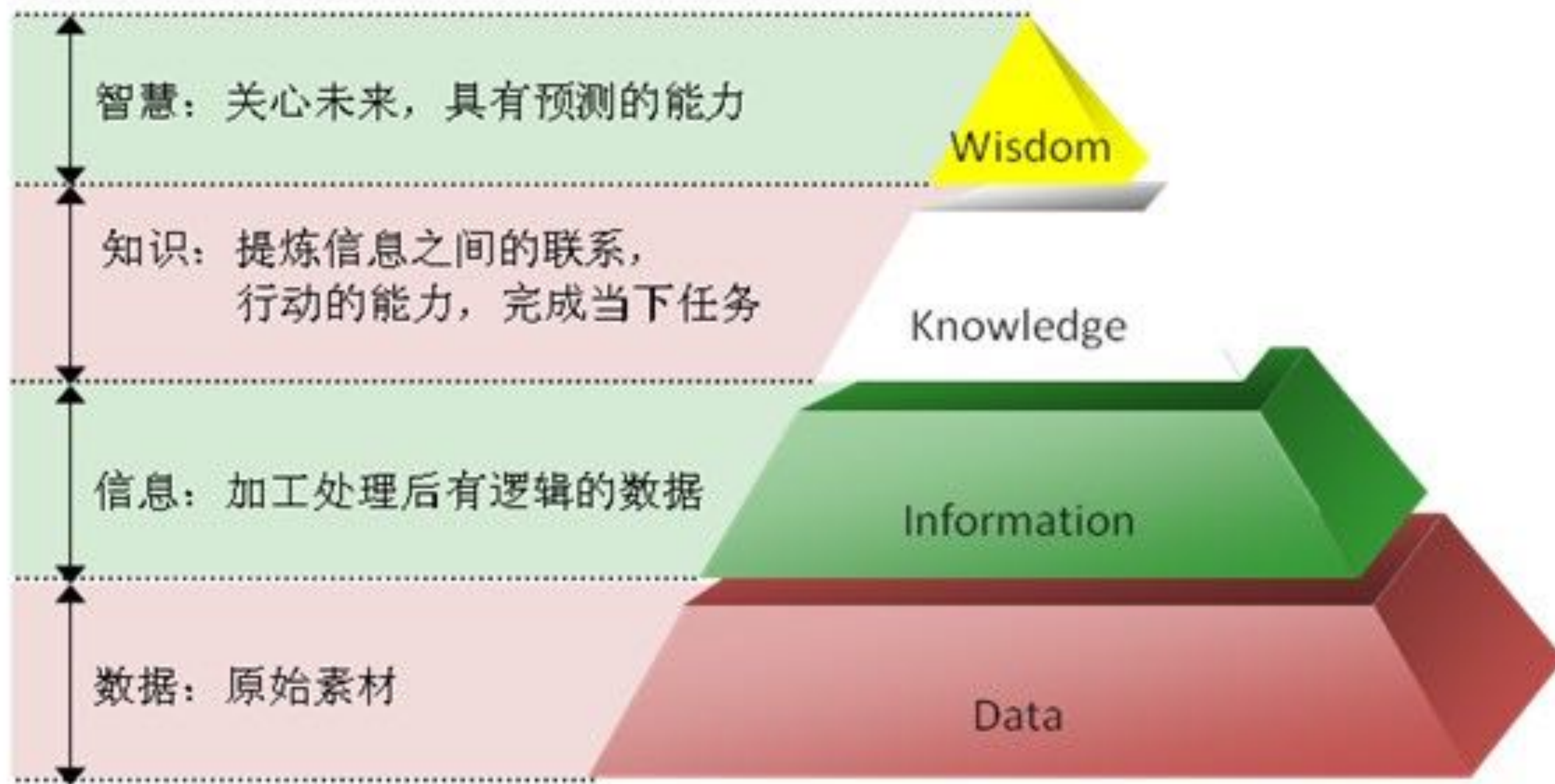


从知识到智慧： 知识图谱还要走多远？

Jane 上海成趣信息科技有限公司

- 知识图谱小动画
- 什么是知识图谱?
- 知识图谱能解决怎样的问题?
- 国内外通用知识图谱的构建方法是?
- 如何构建知识图谱?
- 知识图谱的开源框架有哪些?
- 一起用Cayley构建一个知识图谱
- 开放知识图谱

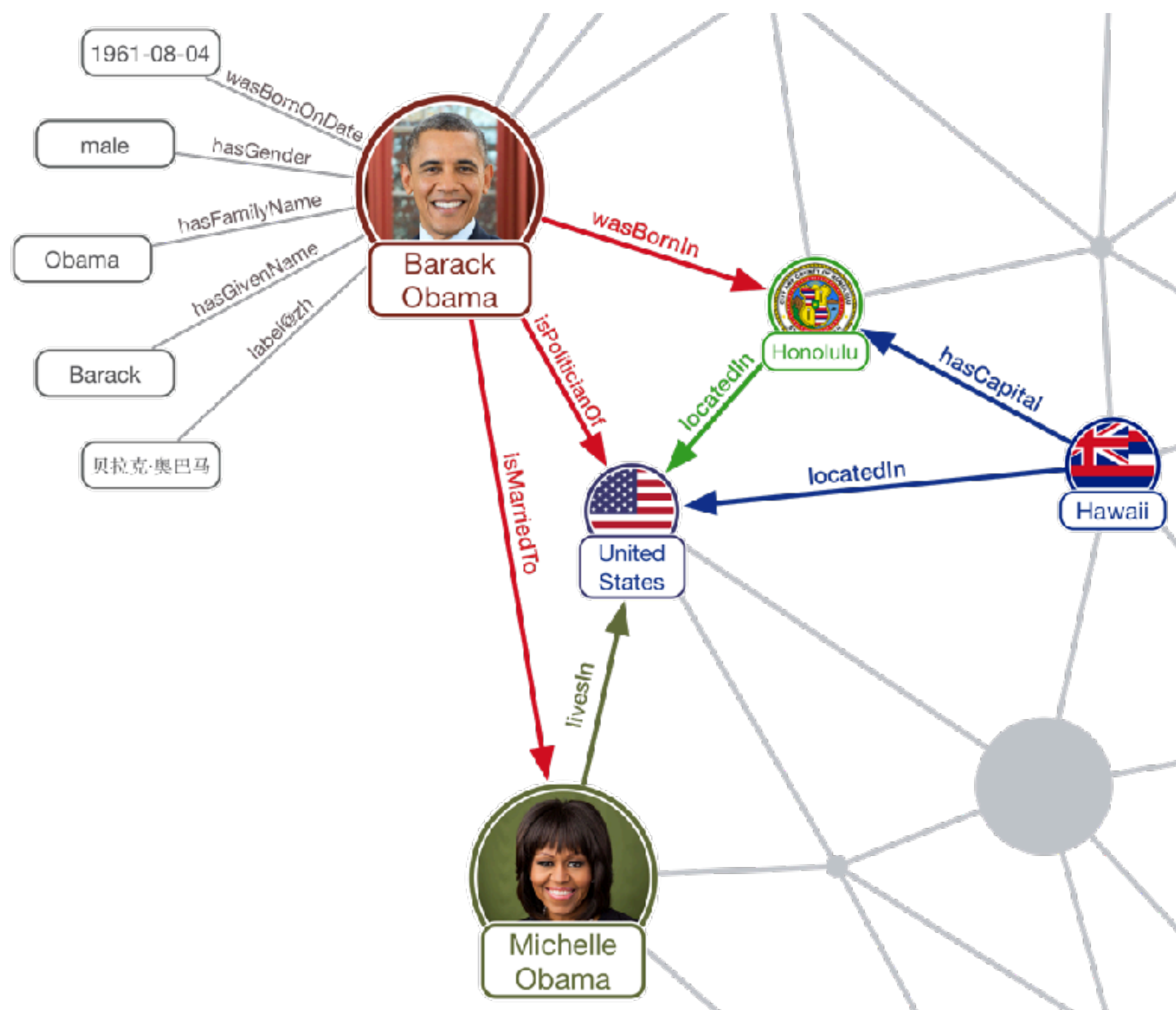


知识图谱小动画

- 知智一分钟 | 第九话：什么是知识图谱？
- 知识图谱通过建立数据之间的关联链接，将碎片化的数据有机的组织起来，让数据更加容易被人和机器理解 and 处理，并为搜索、挖掘、分析等提供便利，为人工智能的实现提供知识库基础。

什么是知识图谱?

- 语义网络
- 将现实世界映射到数据世界，由节点和边组成的语义网络
- 节点 (Point) : 物理世界中的实体/概念 (Entity/Concept)
- 边 (Edge) : 实体的属性或它们之间的关系 (Relationship)



- 概念
- 实体
- 关系

知识图谱的构成

- 数据层：知识的基础，存储在图数据库中，描述客观事实。
- 模式层：人类知识领域的概念层面，强调概念及概念关系的形式化表达。

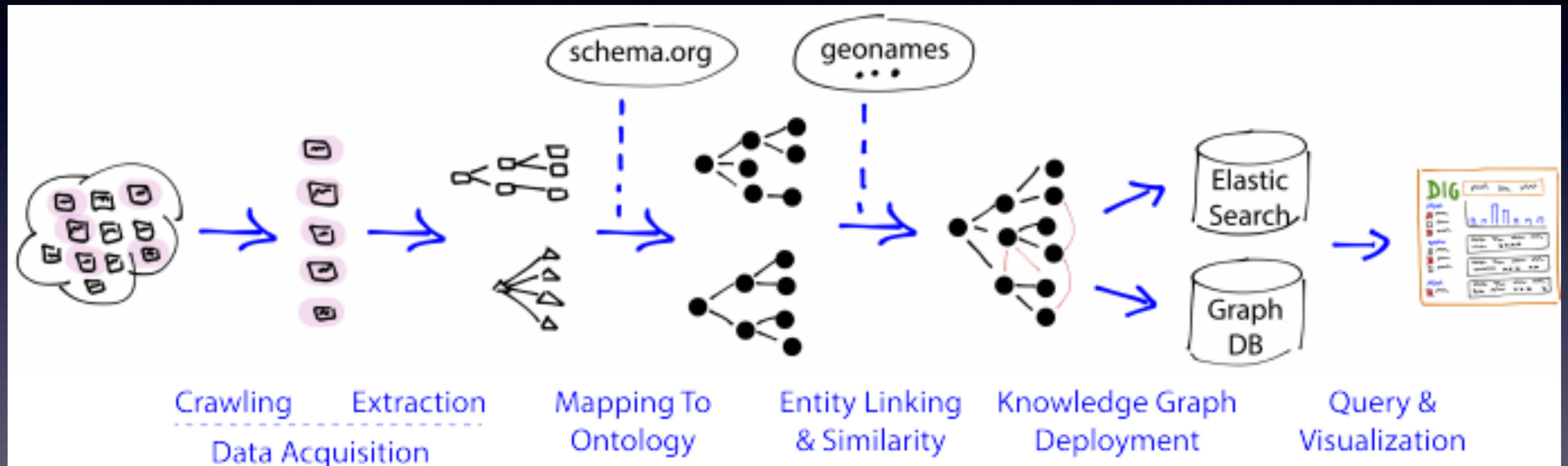
知识图谱能解决怎样的问题？

- 提高搜索质量
- 用户能够使用此功能提供的信息解决其查询的问题，而不必导航到其他网站并自己汇总信息
- 提供问答引擎,辅助深度理解人类的语言和支持推理

国内外构建通用知识图谱的方法

- 自顶向下：Freebase利用维基百科等数据源，提取本体和模式信息，加入到知识库中；
- 自底向上：自动抽取资源，从网页中识别出实体数据，再从实体数据归纳出数据模式，选择其中置信度较高的模式，经过人工审核，构建知识库；
- 所需技术：信息检索、自然语言处理、WWW、人工智能

如何构建知识图谱



- 本体：概念和关系的形式化描述，定义类和关系，以及类层次和关系层次
- 类似于数据库中的schema，OWL语言描述，包括类、属性、个体、数据类型(Datatype)等

本体

- Appendix B. OWL Examples in XML Syntax
- RDF 1.1 N-Quads N-Quads is a line-based, plain text format for encoding an RDF dataset.
- RDF-Resource D? Framework

知识图谱的框架

- 知识存储：
 - 基于RDF的图数据库： Apache Jena、Sesame、gStore、RDF-3X等
 - 原生图数据库： 包括Neo4j(商业的)、OrientDB、Titan和Cayley等
- 知识融合工具： Falcon-AO、Silk、PARIS、DEDUPE、LIMES和KnowledgeVault
- 语义搜索： ElasticSearch

Cayley

Cayley的目标是成为开发者在Linked Data 和图数据（语义网络、社交网络等）的工具之一。初步性能测试表明，在普通PC硬件和硬盘上，基于LevelDB存储的Cayley跑1.34亿三元组毫无压力，多跳交叉查询（比如有X和Y两位影星同时出演的电影）只需150ms。Cayley的主要特点如下：

- Go语言实现，可以作为Go的库
- 运行简单（三四条命令）
- RESTful API（或者REPL）
- 内置查询编辑器和可视化工具
- 支持多种查询语言：JavaScript、MQL
- 多种后端存储数据库：LevelDB（单机）、MongoDB或者内存
- 模块化设计，容易扩展
- 良好的测试覆盖
- 速度快

Cayley相关信息

- 官网: <https://cayley.io/>
- 项目地址: <https://github.com/cayleygraph/cayley>
- 快速安装: <https://github.com/cayleygraph/cayley/blob/master/docs/Quickstart-As-Application.md>
- 相关视频: Build a Small Knowledge Graph Part 2 of 3: Managing Graph Data With Cayley
- 在线平台: <https://cayley-graph.appspot.com/>
- 讨论区: <https://discourse.cayley.io/>

一起用Cayley构建知识图谱

- [Build a Small Knowledge Graph Part 2 of 3: Managing Graph Data With Cayley](#)

开放知识图谱

- <http://openkg.cn/> OpenKG.CN 中文开放知识图谱
- 旨在促进中文知识图谱数据的开放与互联，促进知识图谱和语义技术的普及和广泛应用。
- 涉及的技术领域包括：知识表示、自然语言理解、智能问答、知识抽取、链接数据、图数据库、图挖掘、常识推理等。
- 参考<http://blog.openkg.cn/%E9%99%88%E5%8D%8E%E9%92%A7-openkg%EF%BC%9A%E6%84%BF%E6%99%AF%E4%B8%8E%E6%8C%91%E6%88%98/>



杨晓春
上海成趣信息
科技有限公司
独立顾问

产品设计
技术开发、技术管理
人工智能、数据分析解决方案
物联网解决方案
医疗养老产品
DevHub开发者社区

