知识图谱可视化

知识图谱可视化



显然知识图谱可视化的两个主要步骤是——知识图谱的构建和可视化。

知识图谱 (Knowledge Graph) 并没有一种准确定义,一般我们认为它表示一组相互链接的实体描述——现实世界的对象,事件,情况或抽象概念。或者说,知识图谱由一些相互连接的实体和它们的属性构成,或者简单将知识图谱简单理解为多关系图(包含多种类型的节点和边的图)。

三元组是知识图谱的一种通用表示形式——

G = (E, R, S)

实体集合: E={e1,e2,...}

关系集合: R={r1,r2,...}

三元组集合: S⊆E×R×E

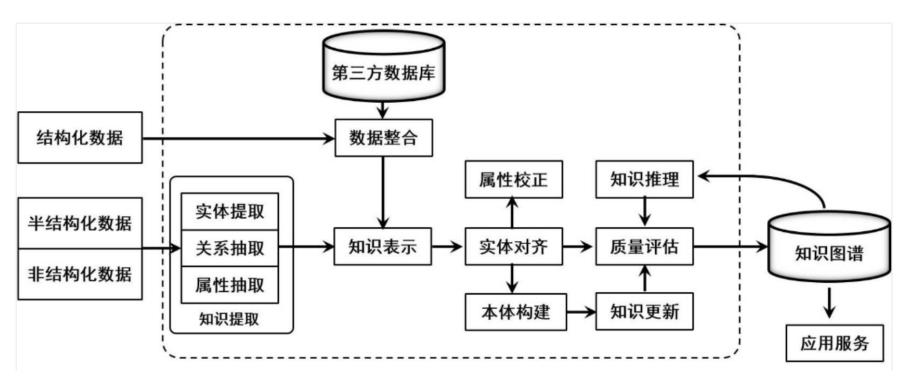
一个简单示例



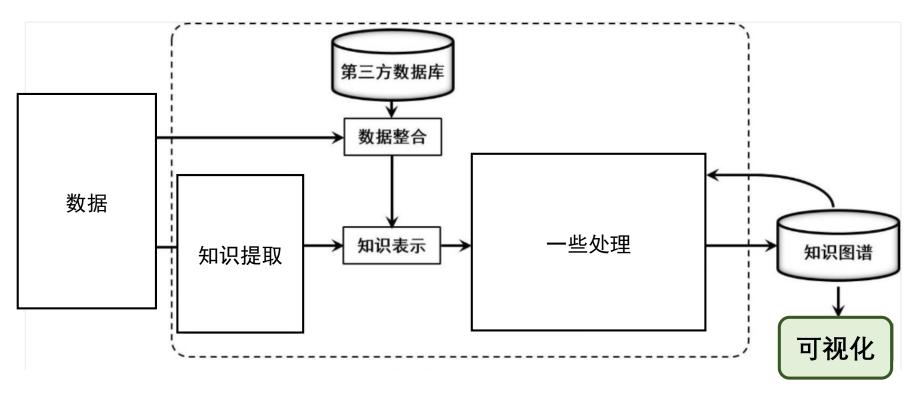
一个复杂示例



知识图谱的构建



知识图谱的构建



数据处理

结构化数据

结构化数据是指已经是表格等结构化形式存储的数据,通过映射将数据转化为所需形式即可。一般有Direct mapping和利用R2RML语言两种映射方法,后者更灵活也有现成工具可以进行使用。

数据处理

半结构/非结构化数据

此类数据需要进行知识提取(信息抽取)才能得到我们需要的数据格式。



数据处理

信息抽取

实体抽取

- 人工编写规则进行抽取
- 机器学习->抽取出与之具有相似上下文特征的实体再筛选
- 基于实体的语义特征->识别出命名实体,聚类算法对识别出的实体对象进行聚类

关系抽取

- 基于模板:人工构造语法和语义规则->模式匹配(基于触发词/字符串 | 基于依存句法)
- 机器学习: 少量人工标记数据(训练集)->实体关系分类模型->分类->抽取原型

*属性抽取

来自百科or文本,同关系

知识融合

用来消除歧义、冗余、错误

- 实体链接
 - 对比文本语言数据和知识库中的条目,将知识库中没有的知识补全到库中
 - 引用表构建: 引用表存储一个名字所有可能指向的实体
 - 例如: 名字-目标实体-次数
 - 实体知识构建
 - 根据实体知名度、实体上下文、实体语义关联度、文章主题等
 - 链接推理算法
- 知识合并
 - 合并外部数据库

知识存储

知识图谱有三种存储方式:关系型数据库、RDF数据库和图数据库,常用的是后两种方式。

关系型数据库

- 数据以数据表的形式存储,表的各个字段存储了不同的属性值
- 另有关系表映在存储实体之间的关系。
- 但是一旦关系越来越多,表的关联也会越来越多,关联条件也会错综复杂,因此只有在小数据量和实体的情况,可用关系型数据库进行存储
- 常用的关系型数据库有Mysql, DB2, Oracle等

RDF数据库

- RDF数据类型的体现是三元组, 三元组是由两个节点和一条边组成的。可以用于描述<实体, 属性, 属性实体>或<实体1, 关系, 实体2>这样的属性或关系
- 优点是简单直接,缺点则是整个知识图谱存在一张表中,单表规模特别大,查询成本和开销很大,实体无法存储较长的字符串, 因此无法存储较长的属性描述
- 常用的RDF数据库有jena等

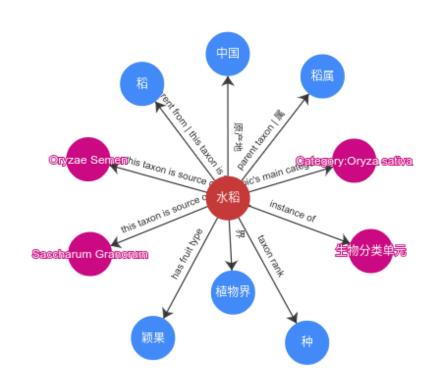
图数据库

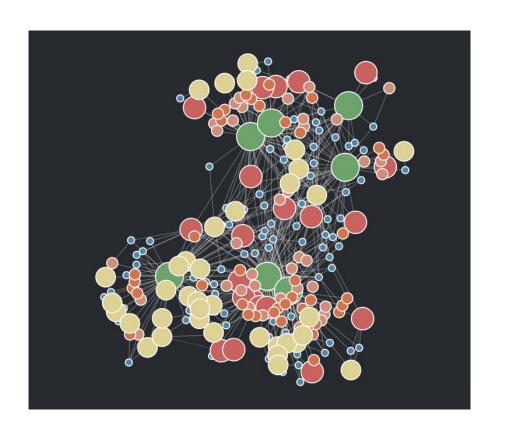
- 图数据库中,一个图数据模型包含两种数据类型,即节点和关系。
- 知识图谱以描述实体的属性与实体间关系为主。
- 图数据库的优势在于可以存储多种较长描述的属性值,不需要单独存储这种较长的属性值。
- 图数据库在工业界已经应用广泛,常用的图数据库有neo4j, OrientDB等。

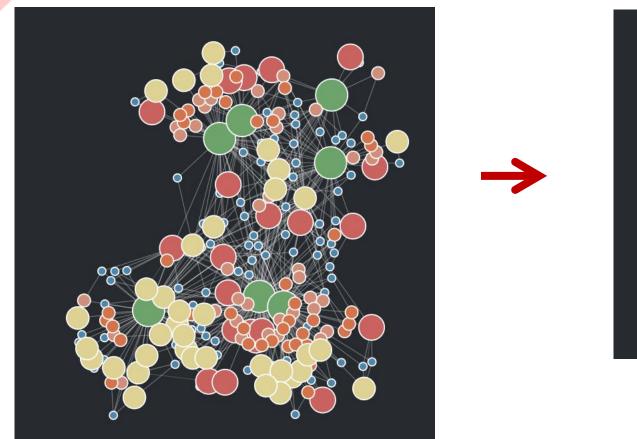
应用服务

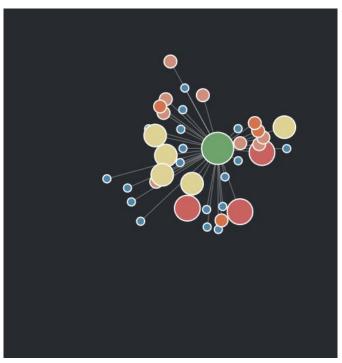
知识推理、知识检索、**可视化**......

知识图谱的可视化,最直观也最容易想到的就是以力导向图形式呈现。

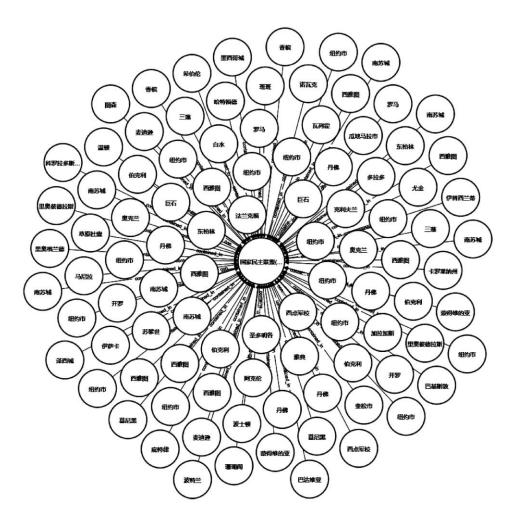


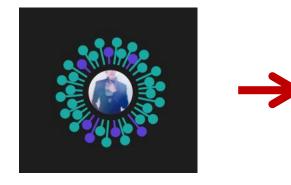


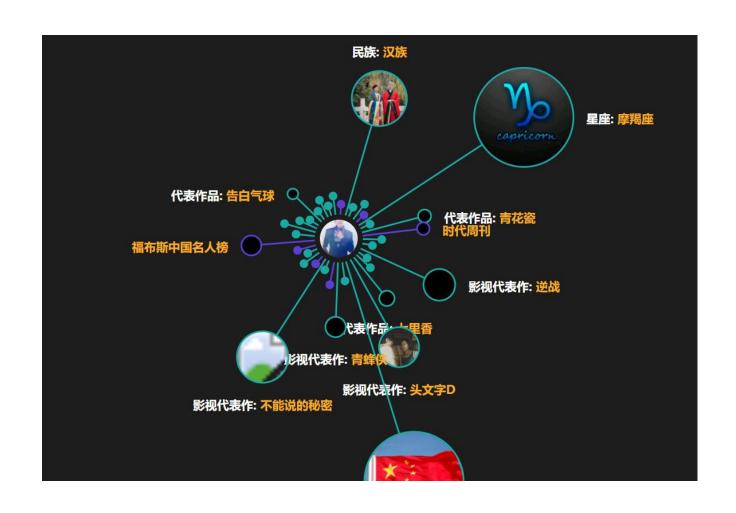




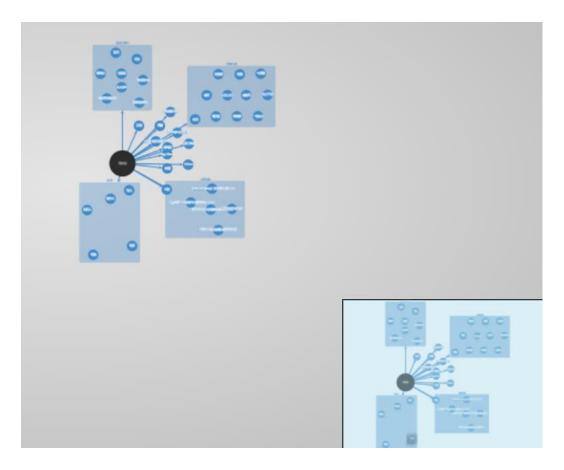
Visualization





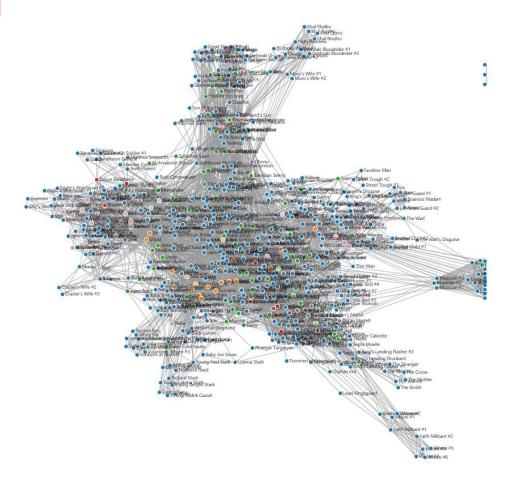


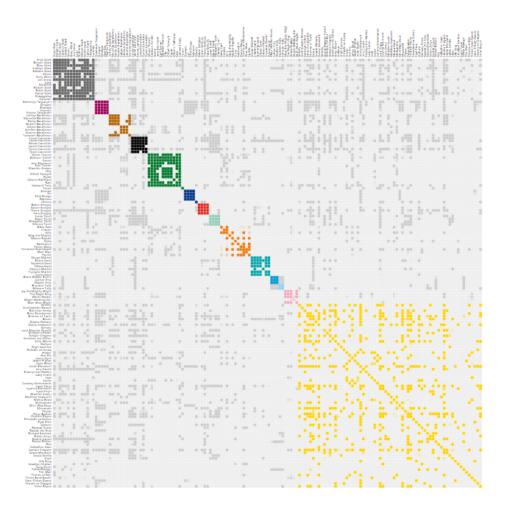
http://shuyantech.com/cndbpedia/kggraph?entity=%E5%91%A8 %E6%9D%B0%E4%BC%A6



http://kw.fudan.edu.cn/cndbpedia/search/?mention=%E5%91%A8 %E6%9D%B0%E4%BC%A6&entity=%E5%91%A8%E6%9D%B0%E 4%BC%A6

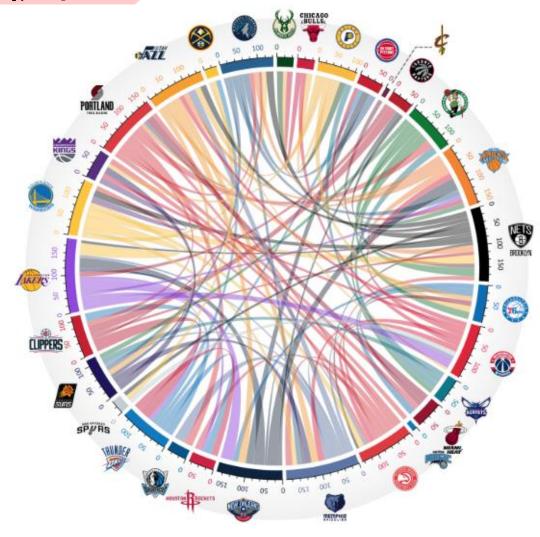






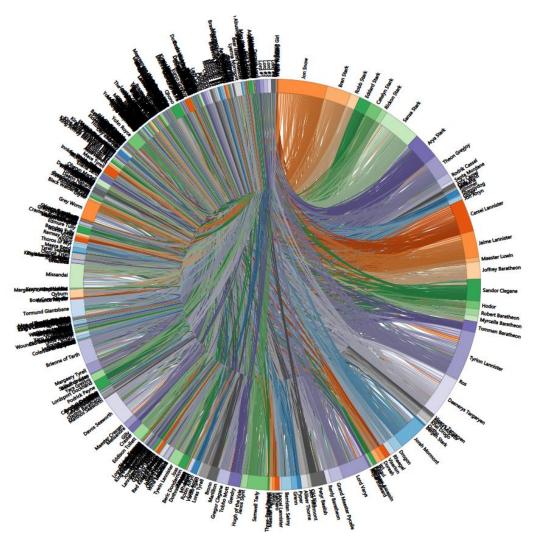
https://jeffreylancaster.github.io/game-of-thrones/force-directed/

https://jeffreylancaster.github.io/game-of-thrones/matrix/

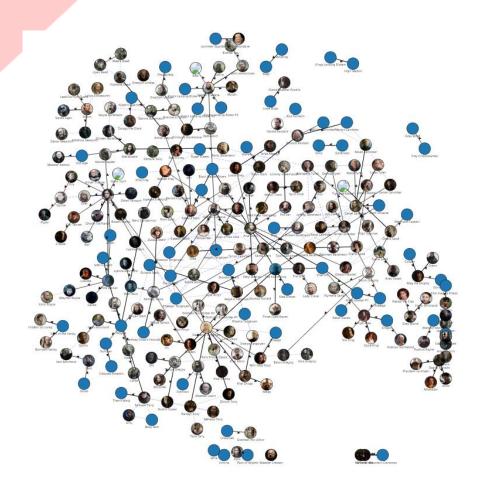


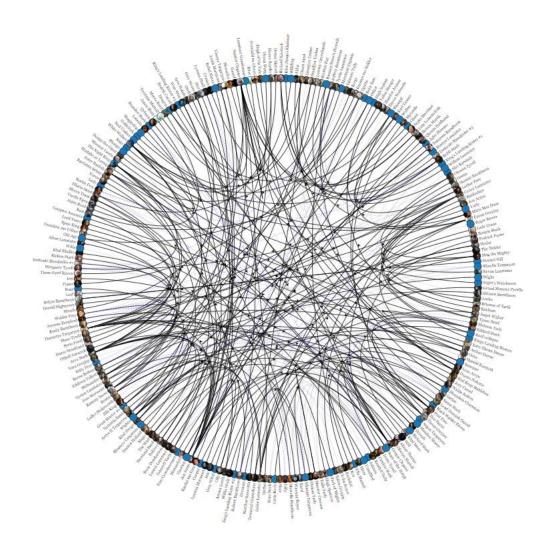
https://www.linkedin.com/pulse/3rd-spin-nba-trades-miguel-serrano/?trackingId=PFB0wUya2ysWPSTWPLNSMw%3D%3D





https://jeffreylancaster.github.io/game-of-thrones/matrix-chord/



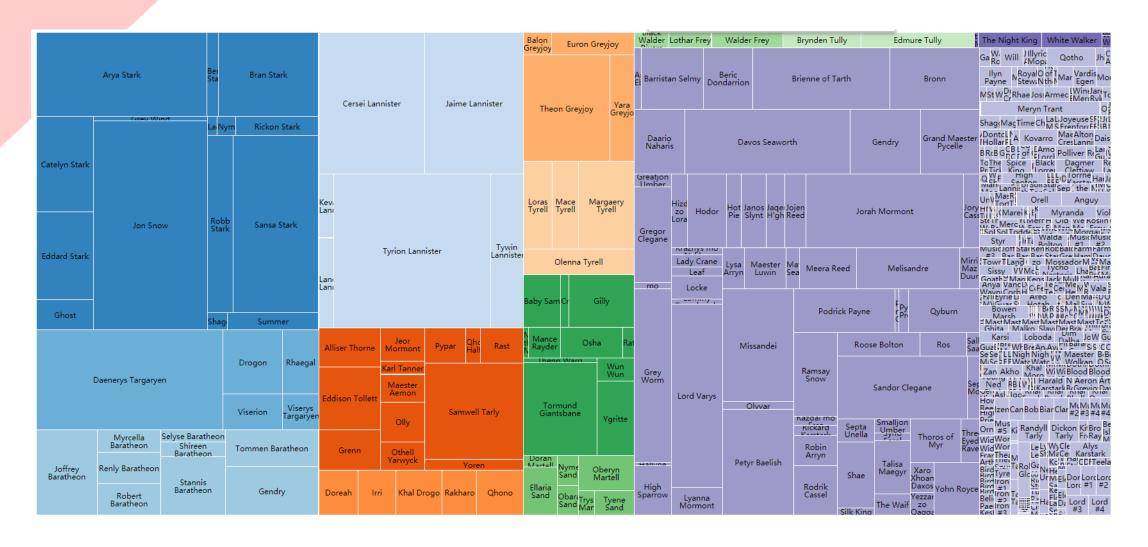


https://jeffreylancaster.github.io/game-of-thrones/relations-force/

https://jeffreylancaster.github.io/game-of-thrones/relations-circle/

Truck Drivetheysicians Bendsonal Filtranfitieds Miscellane dikip and Betrivate Defentiventer Trivethers Financial Extensional Filtranfitieds Servicians Sandrol Language Construction dustrial Candriers addentists Pest Controlla Mandale and English Construction of the Construc Janitors anthipping anthilitary, Barden Weite al Surveying anthipmant of Machineiry Agriculturalle and the Articultural Communication and the distribution of the Articultural Communication and the Articultural Communication an Computer Subjammenter and Subfamore Staladacers and evision Editi Ushers and ro Ticket Construction of the chanic, Indicated African and informatical and the change of the chanic of the change of the c Customer Lawyers and Midgengin Regustrial Stratingers | Brazingers Cottoner Sale Methods Stread Stromedical Bind ogical Nurse Anesthetilist Stel Phindresian's Bill and Adman Painters. Other Installationia Athinabative Landomotive Englineerand Shintor Intertallmens dintersency Management (Electron Management Miscellane Gusulfied) at the electron Service Farmers and distinctivities and accommunication in the communication in Pipelavers, Mechanical Emericanors, Machicale LotCuttine Worformsevors, Computer Official dividical deterioultural divertising Emericanishing affect Preparates see Therricanistical Michaeometer v Miscellanedteavy Vehicheckintlects Computer Colinishinks and Mandillines iffeieres Inspectionmenter authoid and Tolhishine Mackither Religional Desk Schricks and Mandillines tate Schricks and Mandillines Inspection Industrial Marking and Musicians afther Food Other Motor Weighickendiking centric ing Earliesonic Matternal Scilendering Marking managementing Marking School

Visualization



THANK YOU

