



知识图谱构建

报告人：王翔

CONTENT



论文改进思路



输入形式



模型架构



整体代码



分层 概念 规则
 属性 实体 关系

应用

知识搜索

知识展示

知识表示

图特征

知识平台

MySQL

Neo4j

JanuasGraph

Dgraph

ElasticSearch

MongoDB

NebulaGraph

HugeGraph

Jena

Faiss

图谱构建

Schema
管理工具

实体识别

关联挖掘

知识规整

人工
审核

事件抽取

上下位挖掘

实体对齐

关系抽取

同义词挖掘

知识质检

数据

结构化数据

半结构化数据

非结构化数据

公司结构化数据

百科

互联网文本

图片/PDF

开源知识库

专业网站

日志数据

非计算机载体文本

需求、数据和case分析

结构化数据

简单爬虫

人工打标

Schema初步方案

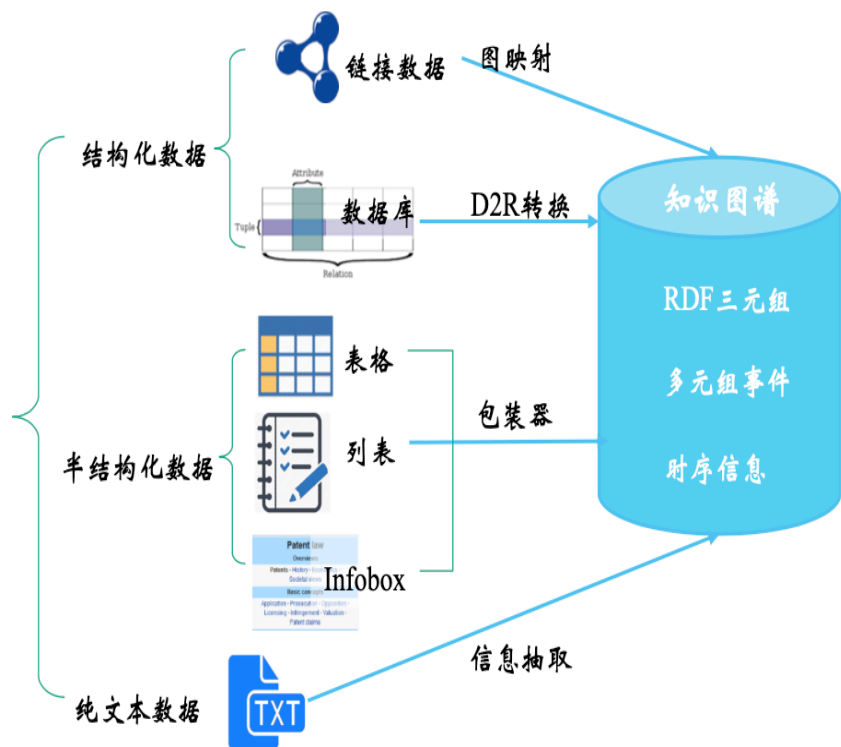
应用初步方案

测试用例/图谱构建指标

KG Demo

方案验证和Review





- 专业领域的新词发现

- 命名实体

- 关系抽取

- 事件抽取

- 上下位挖掘

- 同义词集挖掘

搜狗百科

姚明

中文名	姚明
出生日期	1980年9月12日
出生地	上海市徐汇区
场上位置	中锋
体重	140.6kg ^[1]
别名	明王、移动长城、小巨人、大姚
星座	处女座
血型	B型
主要荣誉	8次NBA全明星 (2003-2009+ 2011) E SPN全球最有潜力运动员奖、 展开
配偶	叶莉

NBA比赛中心 > NBA球员 > 姚明

姚明 (Yao Ming)	
	位置：中锋（11号） 身高：2.29米/7尺6 体重：141公斤/310磅 生日：1980-09-12 选秀：2002年第1轮第1顺位 国籍：中国
姚明生涯常规赛表现	姚明生涯季后赛表现

诊断原词	归一标准词
右肺结节住院	肺占位性病变
右肺结节转移可能大	肺占位性病变 + 肺继发恶性肿瘤 + 转移性肿瘤
前列腺癌骨髓侵犯	前列腺恶性肿瘤 + 骨髓继发恶性肿瘤 + 癌

词标准化/实体对齐



域分类



召回模型



排序模型

聚类模型

• 规则质检

finance:孙宏斌 rdf:type finance:公司

finance:孙宏斌 rdf:type finance:人

finance:人 owl:disjointWith finance:公司

因为用到了OWL词汇，一定要配置
OWL推理机！

打印不一致具体信息

存在不一致，如下：

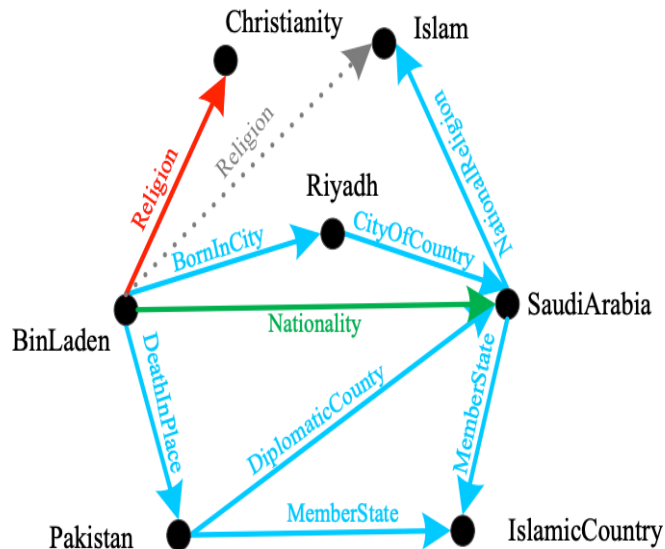
“finance:孙宏斌 rdf:type finance:公司.

finance:孙宏斌 rdf:type finance:人.”

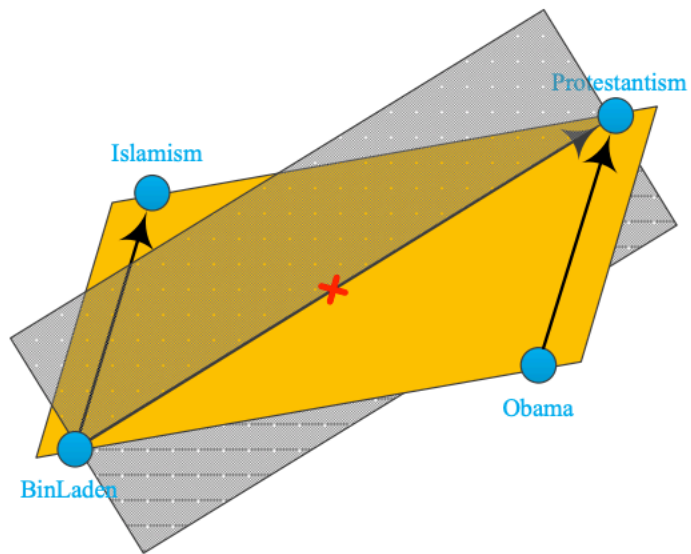
are not compatible with

“finance:人 owl:disjointWith finance:公司”

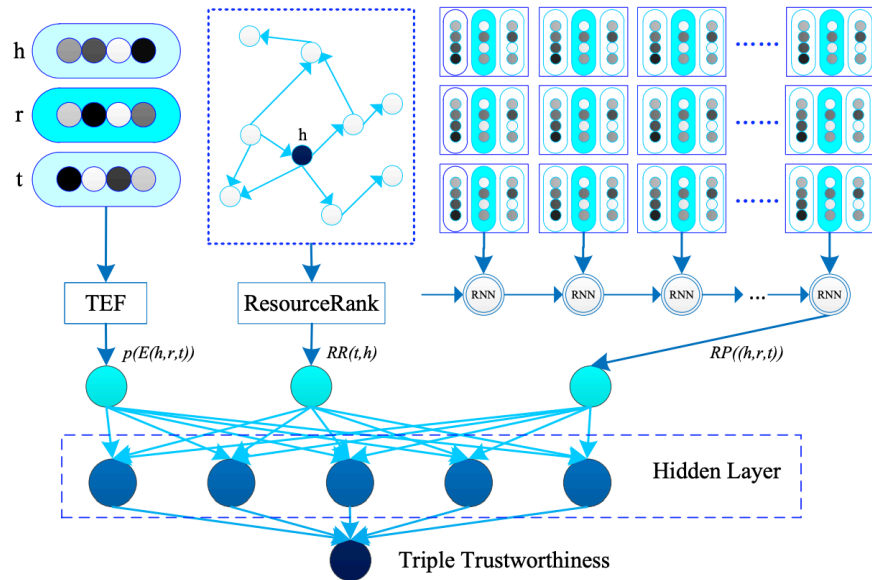
• 图质检



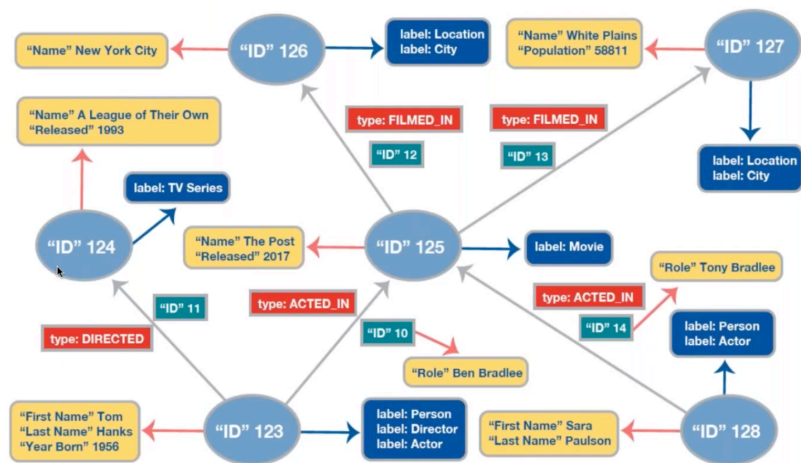
- 嵌入



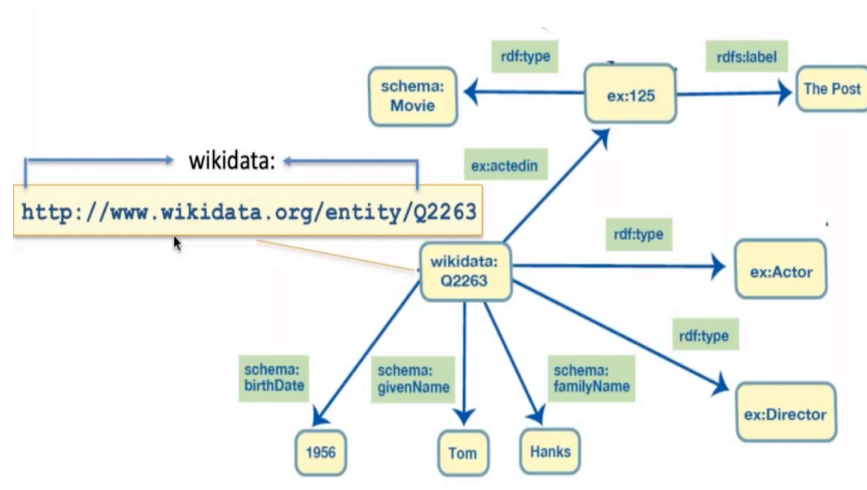
- 模型



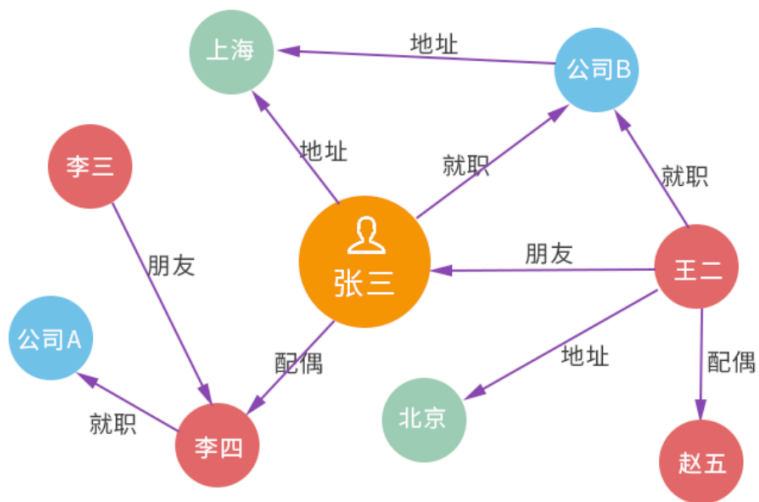
• 属性图



• RDF



张三



张三

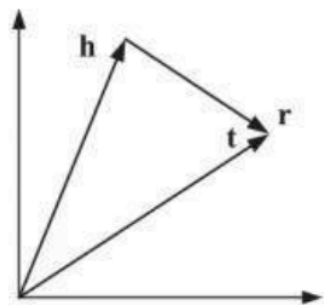
基本介绍

性别：男
民族：汉
地址：上海市卢湾区打浦路15号
邮箱：zhangsan@163.com
出生日期：1983年6月
学历/学位：在职研究生/法学博士
毕业院校：清华大学

设定分析步长

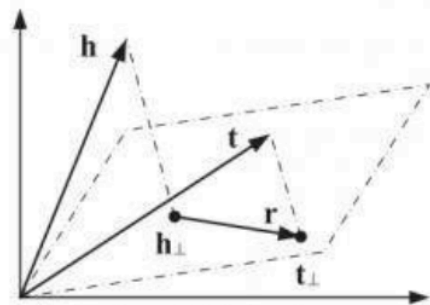
- 1步
- 2步
- 3步

- 地址
- 公司
- 人物



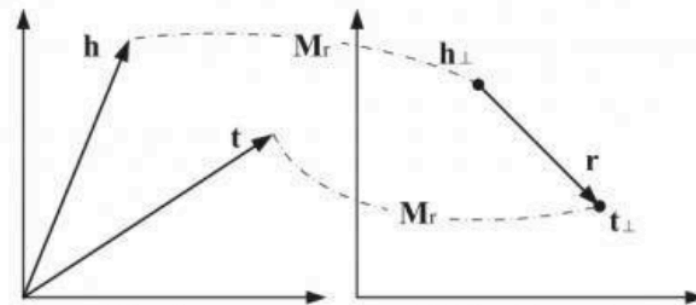
Entity and Relation Space

(a) TransE.



Entity and Relation Space

(b) TransH.



Entity Space

Relation Space of r

(c) TransR.

应用

知识搜索

知识展示

知识表示

图特征

知识平台

MySQL

Neo4j

JanuasGraph

Dgraph

ElasticSearch

MongoDB

NebulaGraph

HugeGraph

Jena

Faiss

图谱构建

Schema
管理工具

实体识别

关联挖掘

知识规整

人工
审核

事件抽取

上下位挖掘

实体对齐

关系抽取

同义词挖掘

知识质检

数据

结构化数据

半结构化数据

非结构化数据

公司结构化数据

百科

互联网文本

图片/PDF

开源知识库

专业网站

日志数据

非计算机载体文
本

- 概念或者实体的颗粒
- 层级构建
- 三元组质检



感谢您的聆听！

报告人：王翔