

# 行业知识图谱+ 知识建模与知识存储技术 及工具分享

链接数据 洞察价值  
Linked Data & Smart Data

---

# What is Happening

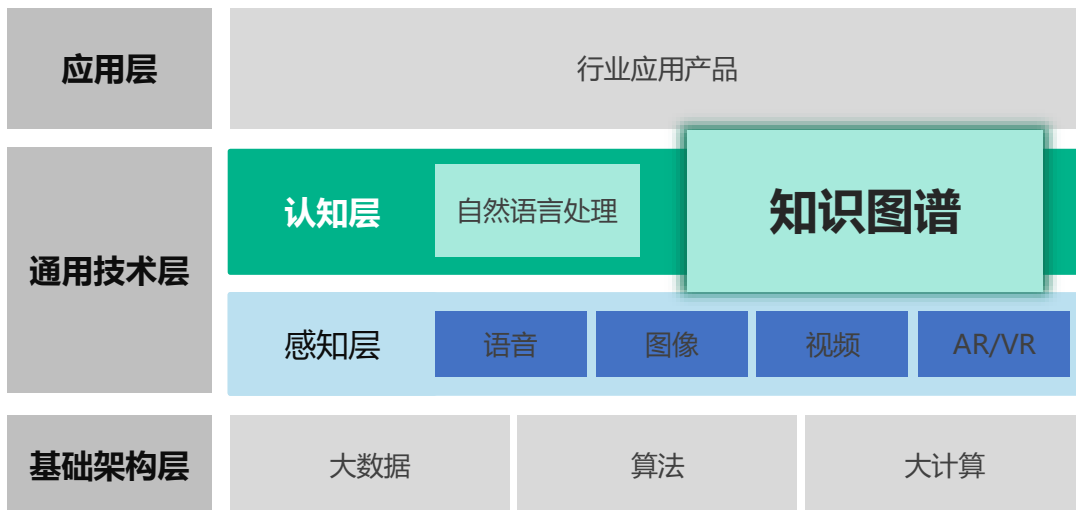
# AI浪潮下知识图谱的发展

## ● 起源



知识图谱(Knowledge Graph)于2012年5月首先由Google提出

## ● 定位



人工智能的架构

## ● 发展



Apple WATCH



微软小冰

WolframAlpha  
计算知识引擎



facebook  
Facebook兴趣图谱

Palantir

Baidu 百度  
百度知心



阿里巴巴  
商品知识图谱

搜狗搜索  
搜狗立知

Tencent 腾讯  
TopBase

# 政策助力

## 国务院《新一代人工智能发展规划》

2017年7月8日 国务院明确提出了

**建立新一代人工智能关键共性技术体系重点任务**

特别强调了

研究跨媒体统一表征、关联理解与知识挖掘、知识图谱构建与学习、知识演化与推理、智能描述与生成等技术，开发跨媒体分析推理引擎与验证系统的关键共性技术



共性关键技术中反复提到以知识为中心，从获取、表示、融合、计算到应用的全生命周期建设。**解决方案——知识图谱！**



# 通用知识图谱 VS 行业知识图谱

## 通用知识图谱



- 面向通用领域
- 以常识性知识为主
- “结构化的百科知识”
- 强调知识的广度
- 使用者是普通用户

## 行业知识图谱



- 面向某一特定领域
- 基于行业数据构建
- “基于语义技术的行业知识库”
- 强调知识的深度
- 潜在使用者是行业人员

# 行业态势

To C

行业应用

InnoTREE

ResearchSpace

DRUGBANK



量子魔镜

WATCH



出门问问



蚂蚁金服 新安娜

问财

产品

facebook

amazon



阿里巴巴  
商品知识图谱



IBM Watson

EVERSTRING

Baidu 百度

Palantir

STARDOG

解决方案

metaphacts



明略数据

明智系统

tamr

拓尔思 TRS

水晶球

To B

# 行业现状——缺少合适的知识建模产品

知识



schema



# 行业现状——缺少统一的实现各类信息抽取的现成工具

---

## CloseIE

- 面向特定领域抽取信息
- 预先定义好抽取的关系类型
- 基于领域专业知识抽取
- 规模小
- 精度比较高

## OpenIE

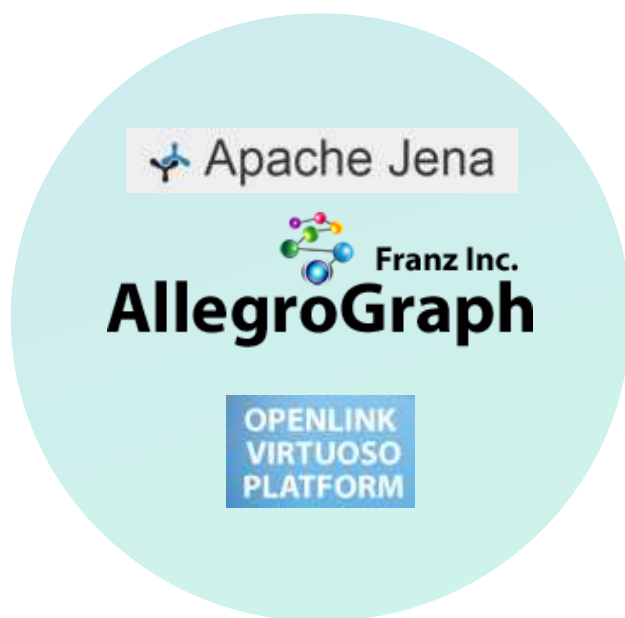
- 面向开放领域抽取信息
- 关系类型事先未知
- 基于语言学模式进行抽取
- 规模大
- 精度相对较低



# 行业现状——缺少合适的知识图谱专用存储产品

- 业界普遍采用 *RDF store* 及 *Graph Store*, 但各有问题

## RDF store



## Graph Store



No Size Fits All !

# 行业现状——缺少图谱统一消费的产品

---

## 知识计算

---



知识推理



图挖掘



动态事件演化

## 知识应用

---



语义搜索



智能问答



可视化辅助决策

# 行业知识图谱生命周期

---



——行业知识图谱构建与应用[CCKS 2017]

# 学界研究

---

a knowledge base management system (KBMS), which performs complex rule-based reasoning tasks over very large amounts of data and, in addition, provides methods and tools for data analytics and machine learning

KGMS, 即一个知识库管理系统 (KBMS), 它对大量数据执行复杂的基于规则的推理任务, 此外还提供数据分析和机器学习的方法和工具

# KG认知误区

---



limited to storing and analyzing social network data only

知识图谱仅限于存储和分析社交网络的数据



acquiring a graph database system and feeding it with data is sufficient to achieve a corporate knowledge graph

采用一个图数据库存储方式，然后向其提供数据就可以构建一个完整的企业知识图谱。



restricting a knowledge graph to RDF data only

知识图谱必然使用RDF三元组方式存储，而不是传统关系型数据。

**KGMS = KBMS + Big Data + Analytics**

# 我们的思考——KGMS整体架构描述



完整的KGMS应该包括：

KGStorage & Management +  KGRobot + KGBuilder + KGSensor + KGPro

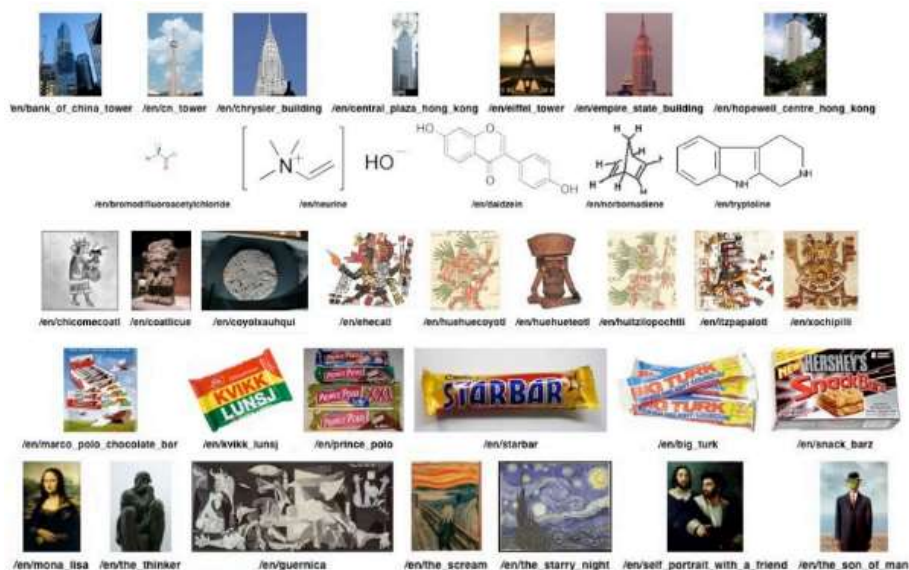
---

**Concept, entity, attribute..... then?**

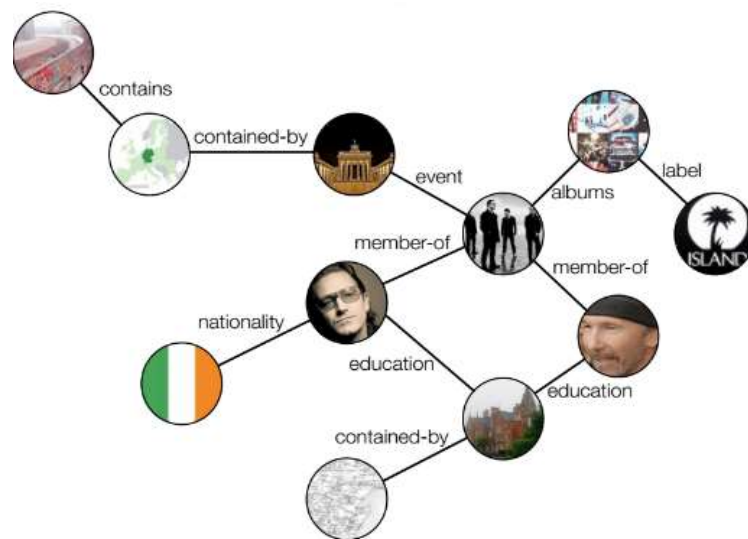
# 知识的组成

知识图谱旨在描述真实世界中存在的各种**实体或概念**，以及它们之间的**关联关系**。其中：

- 每个实体或概念用一个**全局唯一确定的ID**来标识，称为标识符；
- 每个**属性—值对**用来刻画实体的内在特性，而**关系**用来连接**两个实体**，刻画它们之间的关联。



现实世界中有多类型的事物



事物之间有多类型链接



# 知识建模的目标：实体-概念-属性-关系



雷军



人物



董事长



任职机构

# 知识建模的工作

---

## 知识建模需要完成以下几点工作：

- ① 以概念（实体类型）为主体目标，描述领域内的知识类型；
- ② 利用属性来表示不同数据源中针对实体的描述，形成对实体的全方位描述；
- ③ 利用关系来描述各类抽象成实体的数据之间的关联关系，从而支持关联分析；
- ④ 使用事件机制描述领域中动态发展，并利用时序描述事件的发展状况。

# 复杂知识建模 – 事件

---

- 事件：谓词性知识及其关系
- 现有的知识图谱普遍是以“概念及概念间的关系”为核心的，缺乏对“事件”知识的描述和应用。
- 事件间的演化关系多数是不确定的，而实体之间的关系基本是稳定的。
- 事件知识是表示-回答why/how类型问题的必要条件

# 复杂知识建模 – 原子事件

原子事件（也称为基本事件或简单事件）是一个可直接观察到的、最基本的不能再分解的事件。

——《概率论》

## 常用属性

发生  
时间

结束  
时间

关联  
实体

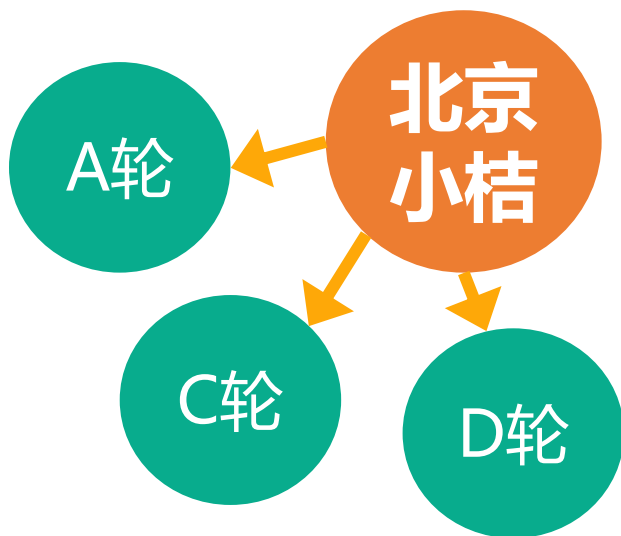
起因

结果

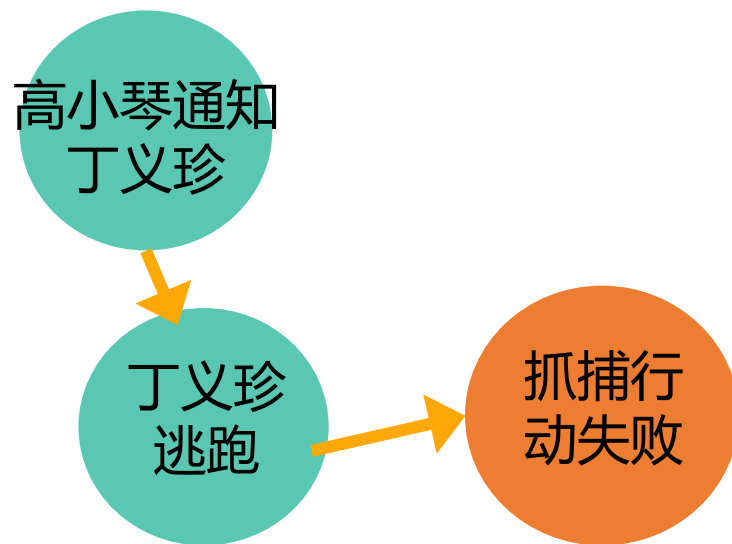
如：经典概率实验——抛硬币  
走路  
吃饭.....

# 复杂知识建模 – 动态事件

- 动态事件：若干个相关的原子事件的集合



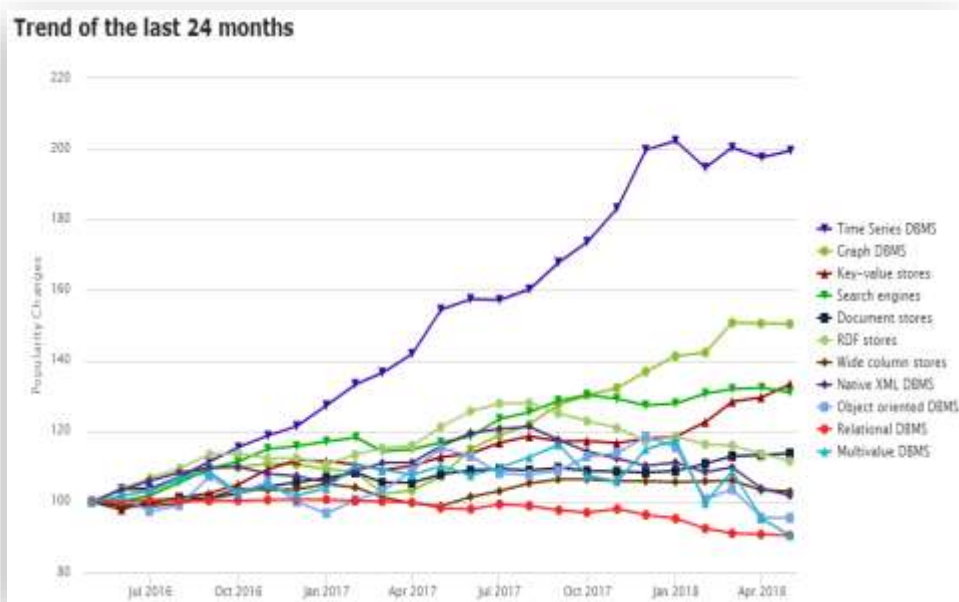
基于时序的集合



基于因果的集合

# 复杂知识建模 – 时序知识

- 时序特性是信息的客观存在。
- 知识和数据是不断更新的，这些变化的时间点同样代表了有意义的信息。
- 时态信息的需求与技术一直伴随着数据库技术的发展而产生和发展。



# 知识图谱中的时序信息

---

- **事件的生成时间**

如：PlantData完成天使轮融资的时间为2017年12月

- **某事件的有效时间段**

如：某人在某公司任一个职位时间段

- **某一对象在特定历史时刻的状态**

如：2018年5月PlantData融资状态

- **知识图谱在过去特定时间的版本**

如：工商知识图谱在去年12月的历史快照

# 事件基于时序演进的可视化

?

重新  
布局

返回  
上一片

天使投资人王刚

北京小桔科技有限公司

2-07-01 2012-07-01

^

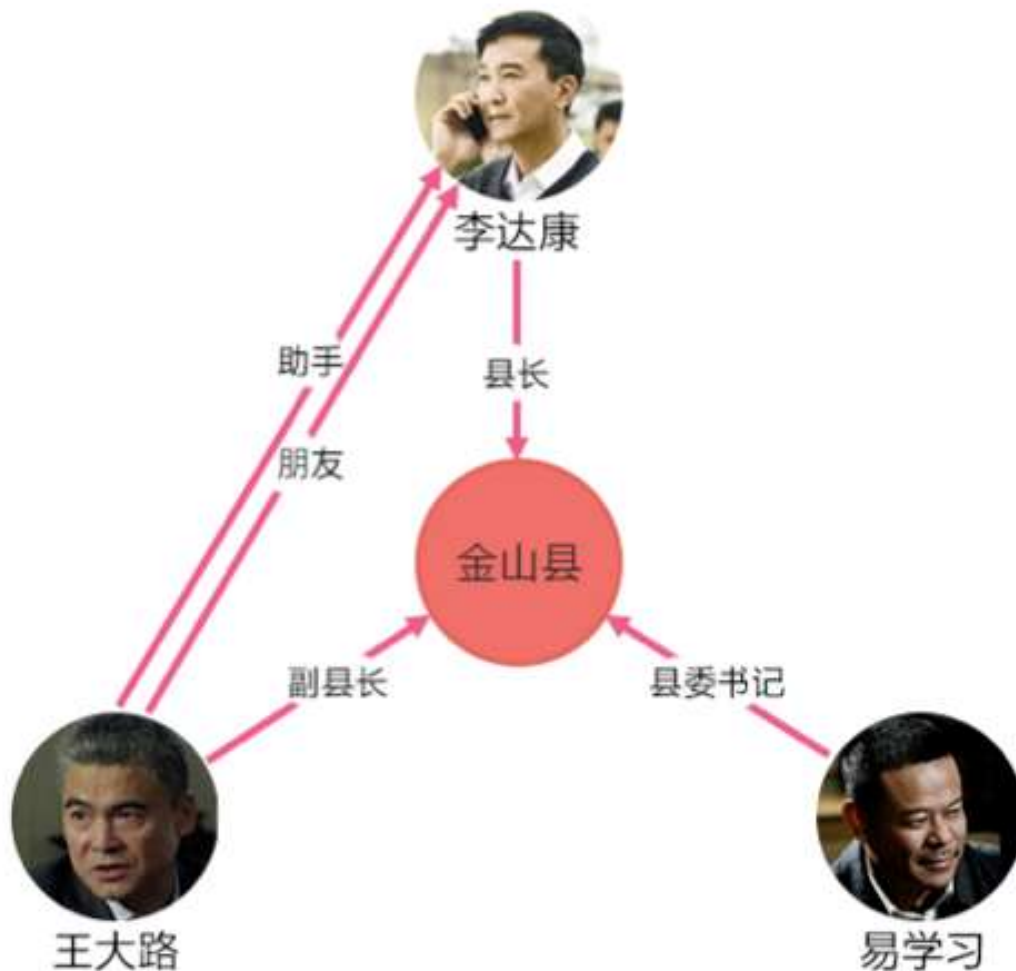


21年前

李达康为了修路，在全县搞强行摊派捐款

一位农妇被逼喝农药自杀

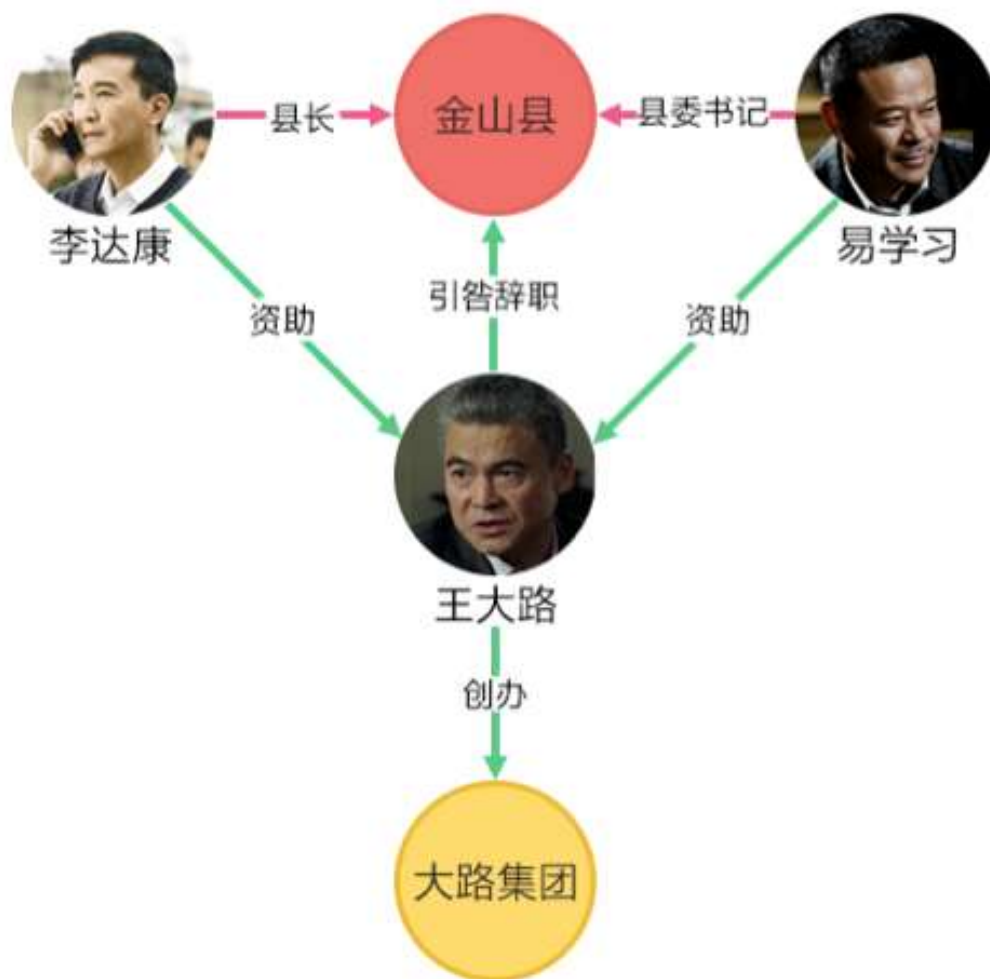
县政府被闹



王大路背锅，引咎辞职

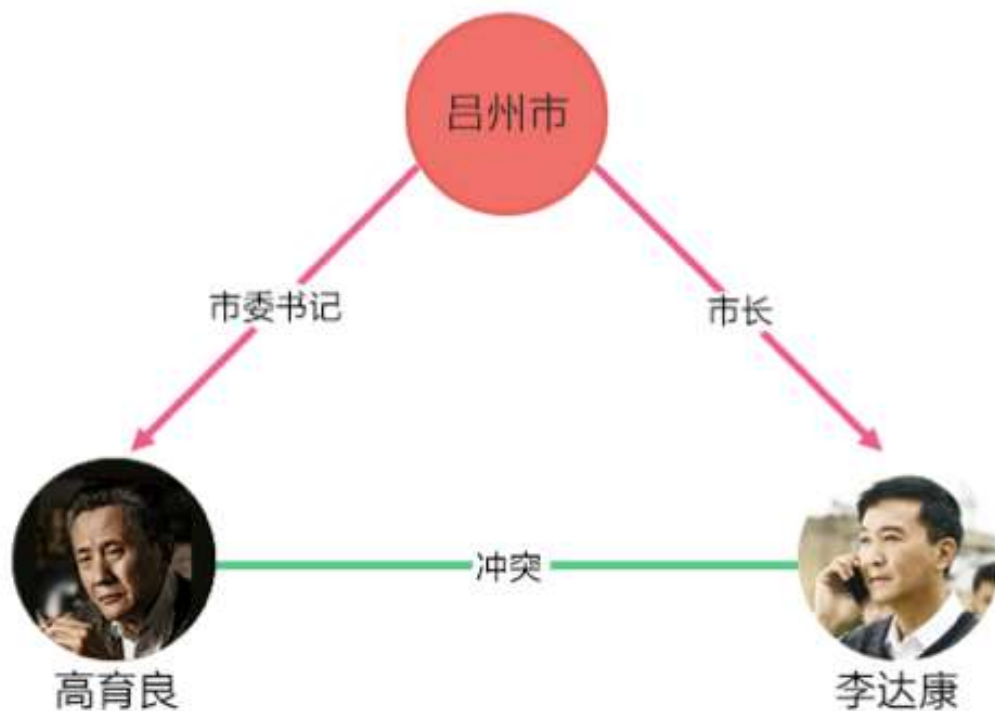
李达康与易学习  
资助王大路创业

王大路多年经营，  
拥有今天的大路集团



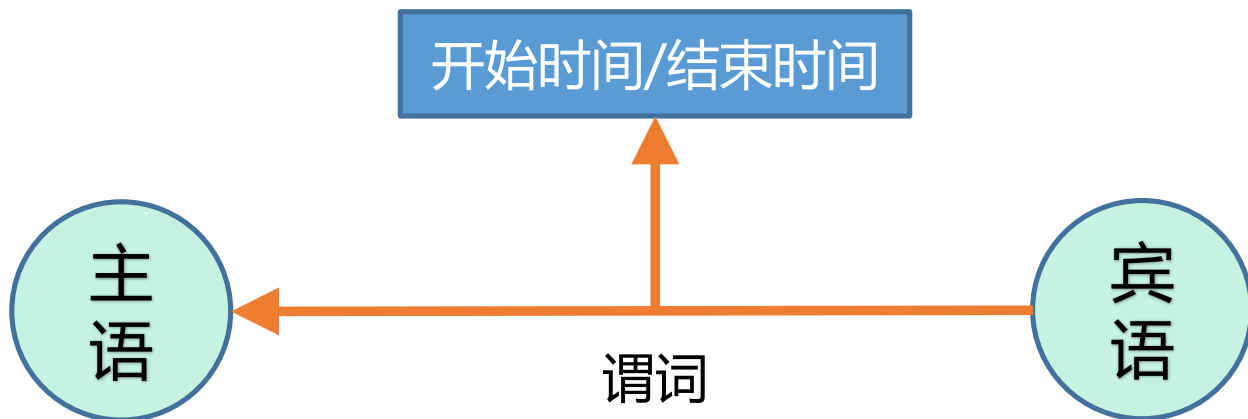
## 高、李冲突原因

- 赵立春要求吕州市长李达康批地
- 李达康拒接批地
- 赵立春调离李达康至林城
- 高育良批地



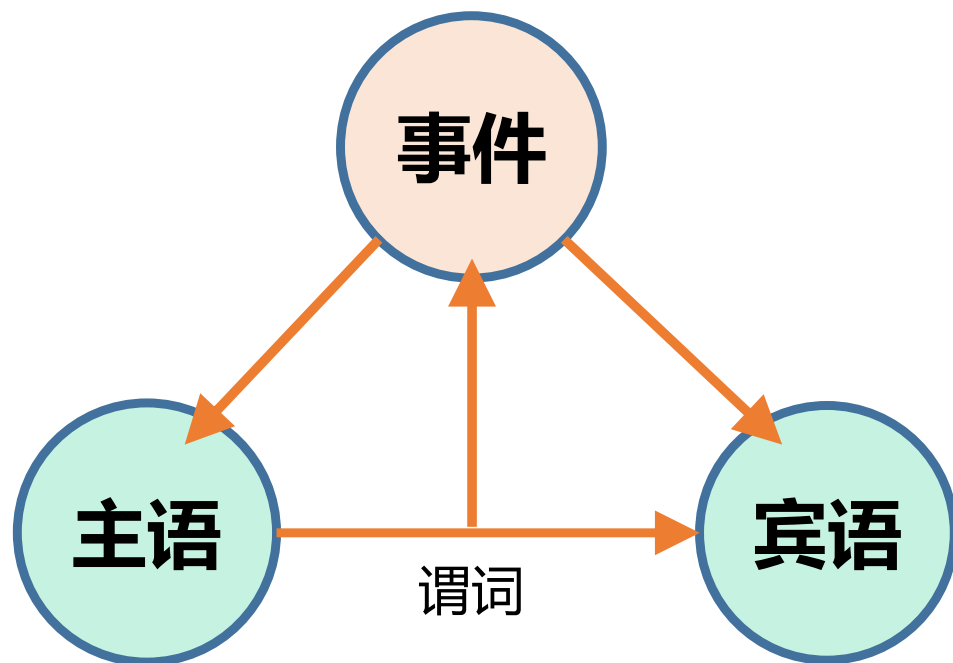
# 复杂知识建模 – 时序的表示方式

- 边属性



# 复杂知识建模 – 事件的表示方式

- 抽象类型实体



# 使用知识图谱对数据进行抽象建模

---

- 以**实体**为主体目标，实现对不同来源的数据进行映射与合并。（实体抽取与合并）
- 利用**属性**来表示不同数据源中针对实体的描述，形成对实体的全方位描述。（属性映射与归并）
- 利用**关系**来描述各类抽象建模成实体的数据之间的关联关系，从而支持关联分析。（关系抽取）
- 通过**实体链接**技术，实现围绕实体的多种类型数据的关联存储。（实体链接）
- 使用**事件**机制描述客观世界中动态发展，体现事件与实体间的关联；并利用**时序**描述事件的发展状况。（动态事件描述）

# 知识建模的方法

---

即建立知识图谱的数据模式。行业知识图谱的数据模式对整个知识图谱的结构进行定义，因此需要保证可靠性。

## 自顶向下的方法

专家手工编辑形成数据模式

## 自底向上的方法

基于行业现有的标准进行转换  
从现有的高质量行业数据源（如业务系统数据库表）中进行映射

# 知识建模关键技术与难点

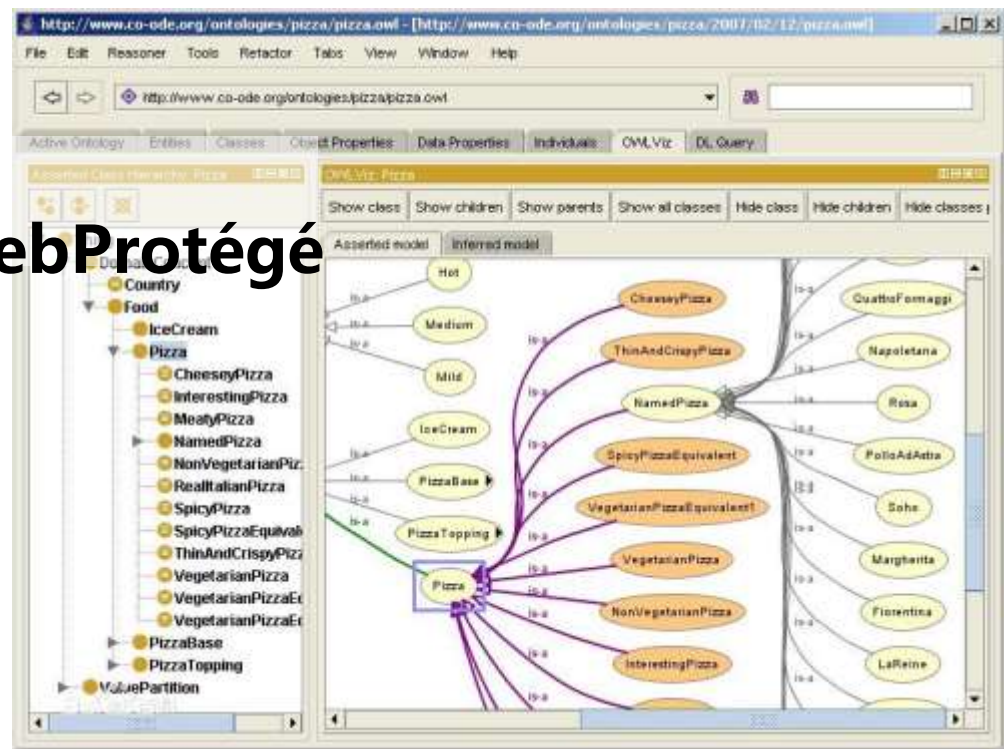
---

- 复杂知识建模
- 多人在线协同编辑，并且实时更新
- 能够导入集成使用现有的（结构化）知识
- 支持大数据量
- 能够支撑事件、时序等复杂知识表达
- 可以与自动算法进行结合，避免全人工操作



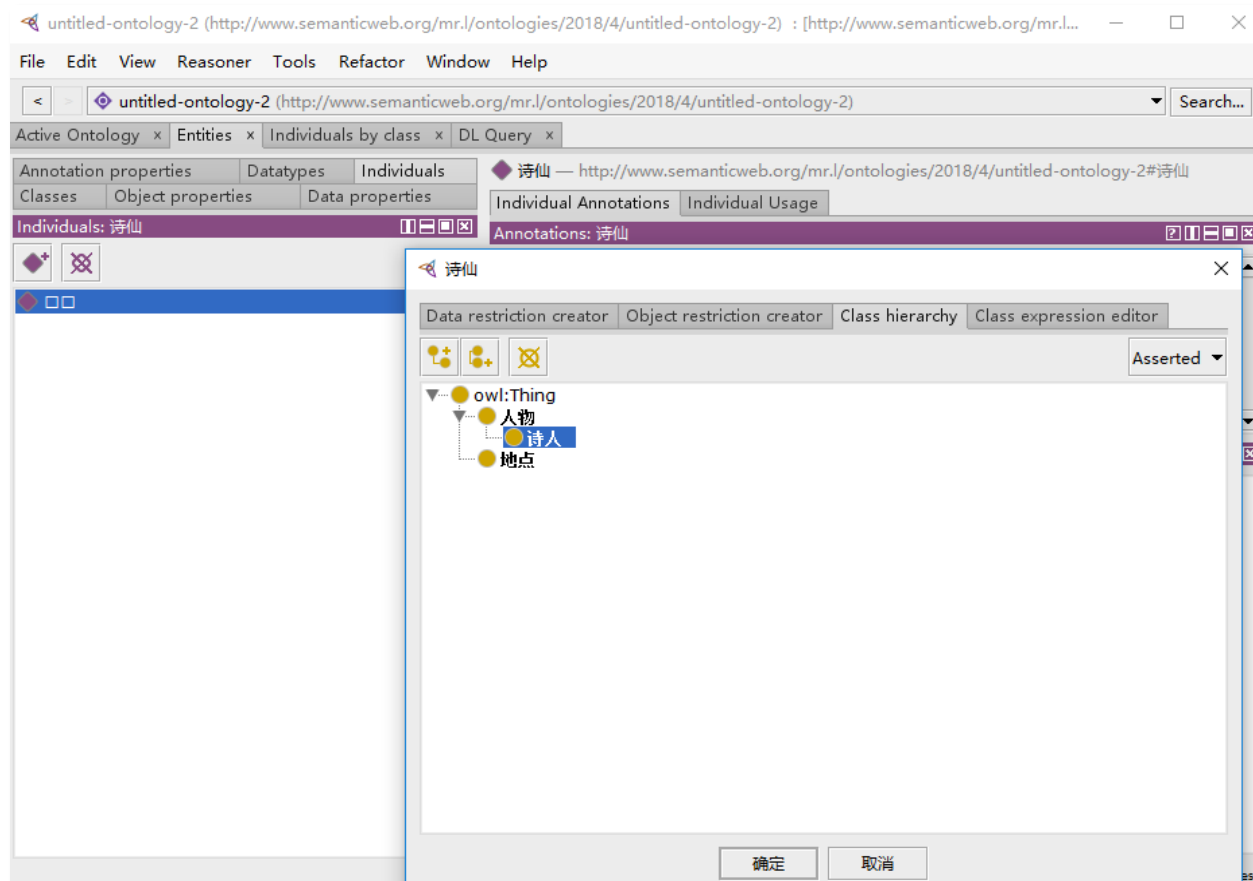
# 知识建模工具 Protégé

- 本体编辑器
- 基于RDF(S), OWL等语义网规范
- 图形化界面
- 提供了在线版本——WebProtégé
- 适用于原型构建场景



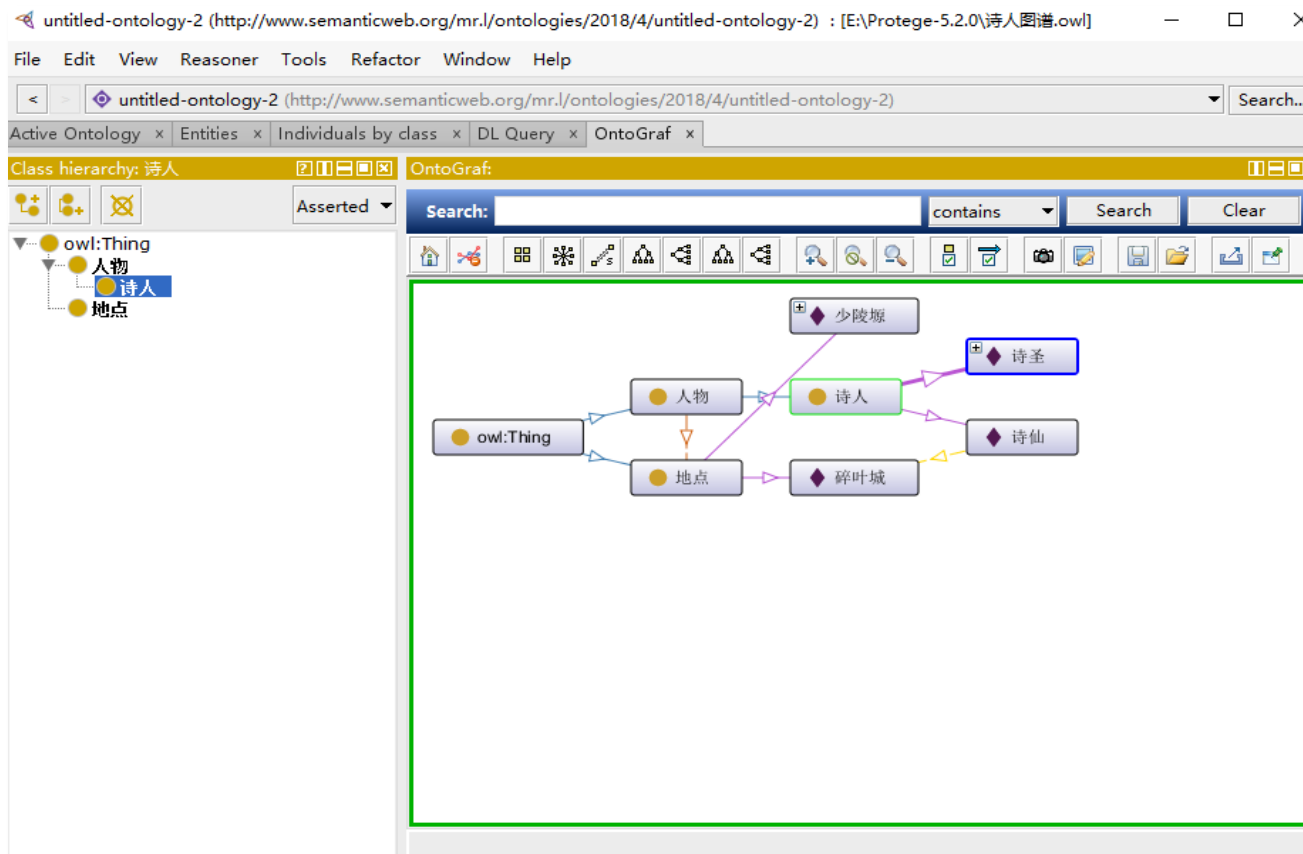
# 知识建模工具 Protégé

- 概念定义
- 属性定义
- 实例编辑



# 知识建模工具 Protégé

## ● 可视化



# 知识建模工具 WebProtégé

## ● 概念编辑

The screenshot displays the WebProtégé web application interface. At the top, there is a navigation bar with a logo, a 'Home' button, and a dropdown menu containing 'Project', 'Share', 'hx5266463', and 'Help'. Below this is a secondary navigation bar with tabs for 'Classes', 'Properties', 'Individuals', 'Comments', 'Changes by Entity', 'History', 'OWL Properties', and an 'Add tab' button. The main content area is divided into two panels. The left panel, titled 'Class Hierarchy', contains a tree view with a search bar and buttons for 'Create', 'Delete', 'Watch', and 'Search'. The tree shows a hierarchy starting with 'owl:Thing', which includes sub-classes like '产品', '人物', '企业' (highlighted in blue), '企业标签', '投资机构', and '融资事件'. The right panel, titled 'Class: 企业', is used for editing the selected class. It features an 'IRI' field with the value 'http://webprotege.stanford.edu/RXvduU5QCItIT6vrp1d9sj'. Below this are sections for 'Annotations' (showing 'rdfs:label' with a value of '企业'), 'Classes' (showing 'owl:Thing' as a superclass), and 'Relationships'. Each section has input fields for 'Enter property', 'Enter value', and 'lang'.

# 知识建模工具 WebProtégé

## ● 属性定义

The screenshot displays the WebProtégé web application interface. At the top, there is a navigation bar with a logo, a 'Home' button, and a menu containing 'Project', 'Share', 'hx5266463', and 'Help'. Below this is a secondary navigation bar with tabs for 'Classes', 'Properties' (which is active), 'Individuals', 'Comments', 'Changes by Entity', 'History', 'OWL Properties', and an 'Add tab' button. The main content area is divided into two panels. The left panel, titled 'Property Hierarchy', contains a tree view with a search bar and buttons for 'Create', 'Delete', 'Watch', and 'Search'. It shows a hierarchy starting with 'owl:topObjectProperty', followed by '开发商' (Developer), and then '所属地区' (Region). The right panel, titled 'Property: 开发商', shows the configuration for this property. It includes an 'IRI' field with the value 'http://webprotege.stanford.edu/R8kP0Nnc148afpziQ9jaBXE'. Below this are sections for 'Annotations', 'Domain', and 'Range'. The 'Annotations' section shows 'rdfs:label' with a value of '开发商'. The 'Domain' section shows '企业' (Company). The 'Range' section shows '产品' (Product).

创投 | Home Project Share hx5266463 Help

Classes Properties Individuals Comments Changes by Entity History OWL Properties Add tab

Property Hierarchy

Create Delete Watch Search

Object Properties Data Properties Annotation Properties

owl:topObjectProperty

开发商

所属地区

Property: 开发商

IRI

http://webprotege.stanford.edu/R8kP0Nnc148afpziQ9jaBXE

Annotations

rdfs:label 开发商 lang

Enter property Enter value lang

Domain

企业

Enter a class name

Range

产品

Enter a class name

# 知识建模工具 WebProtégé

## ● 实例编辑

The screenshot displays the WebProtégé web application interface. At the top, there is a navigation bar with a logo, a 'Home' button, and a 'Project' dropdown menu. Below this is a secondary navigation bar with tabs for 'Classes', 'Properties', 'Individuals' (which is currently selected), 'Comments', 'Changes by Entity', 'History', and 'OWL Properties'. An 'Add tab' button is located on the far right of this bar.

The main content area is divided into two primary sections. On the left, the 'Class Hierarchy' panel shows a tree structure starting with 'owl:Thing'. Under it, several classes are listed: '产品', '人物', '企业' (which is highlighted), '企业标签', '投资机构', and '融资事件'. Below the class hierarchy, there is a section titled 'Individuals by Class: 企业'. It includes a search bar and a list of individuals: '北京小桔科技有限公司', '北京拜克洛克科技有限公司' (highlighted), and '北京摩拜科技有限公司'.

On the right side of the interface, the 'Individual: 北京拜克洛克科技有限公司' panel is open. It contains several sections for editing the individual's properties:

- IRI**: A text field containing the URI 'http://webprotege.stanford.edu/R7gTDEePuY7eqSbybO8wqUh'.
- Types**: A section where the individual is assigned the type '企业'.
- Relationships**: A section showing a relationship with the property 'rdfs:label' and the value '北京拜克洛克科技有限公司'.
- Same As**: A section for defining the same-as relationship with another individual.

# 知识建模工具 Protégé 总结

---

## ● 建模特性层面

- 不支持**复杂事件及时序**的建模
- 不支持**边属性**定义
- 完全遵守RDF规范，所有元素都必须以**URI**的方式进行定义

## ● 使用层面

- 基本只提供单人编辑，在线版本的并发功能支持也不完善；并发编辑时需要通过文件共享来实现；
- 因为基于单机构建，因此对大数据量支持不够，会出现内存溢出；
- 完全依靠人工，难以实现与知识图谱构建（半）自动化过程的交互。

# 构建一个适用的建模工具（1）

- 在线并发编辑支持；
- 编辑的知识实时保存，当其它用户对当前用户正在编辑的内容有更新时，系统自动提示加载最新版本，因此能够有效地解决并发知识编辑冲突。



在线编辑



上下位关系定义



属性定义编辑



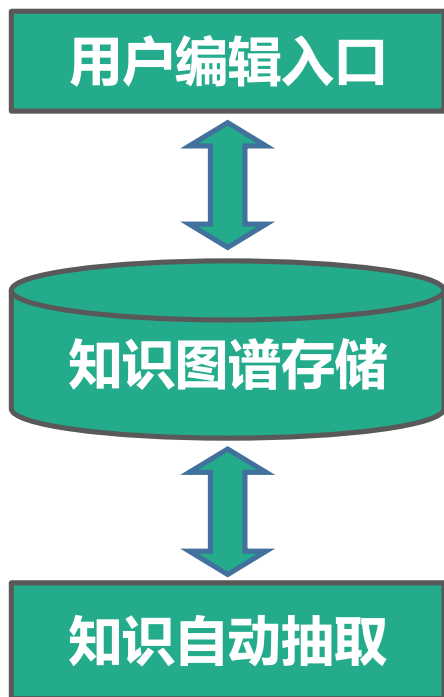
# 构建一个适用的建模工具（2）

- 提供导入、集成功能，能够把现有的知识通过导入功能进行集成；
- 以存储为桥梁，可以对自动算法的结果进行编辑。



The screenshot shows a dialog box titled '概念导入' (Concept Import) with three tabs: '概念导入' (selected), '实例导入' (Instance Import), and '实例关系导入' (Instance Relationship Import). The main content area contains the text: '下载模板，按模板标准构建多层次概念' (Download template, build multi-level concepts according to template standards) and a note: '(注意：编辑完成后请保存好表格，便于下一步将创建的概念树导入图谱)' (Note: After editing is complete, please save the table well for the next step of importing the created concept tree into the graph). There is a '下载模板' (Download Template) button. Below this is a text input field with the placeholder '待上传文件:import\_concept\_template.csv' and a '选择文件' (Select File) button. At the bottom right, there are two buttons: '开始导入' (Start Import) and '关闭' (Close).

知识导入



# 构建一个适用的建模工具（3）

- 支持对动态事件数据的建模，使用时态信息存储实现事件时间描述。

属性名	属性类型	定义域	操作选项
轮次	数值	融资事件	<a href="#">详情</a>
融资额	数值	融资事件	<a href="#">详情</a>
融资时间	数值	融资事件	<a href="#">详情</a>
投资方	对象	融资事件	<a href="#">详情</a> <a href="#">边属性编辑</a>
融资方	对象	融资事件	<a href="#">详情</a> <a href="#">边属性编辑</a>
融资事件标签	对象	融资事件	<a href="#">详情</a> <a href="#">边属性编辑</a>

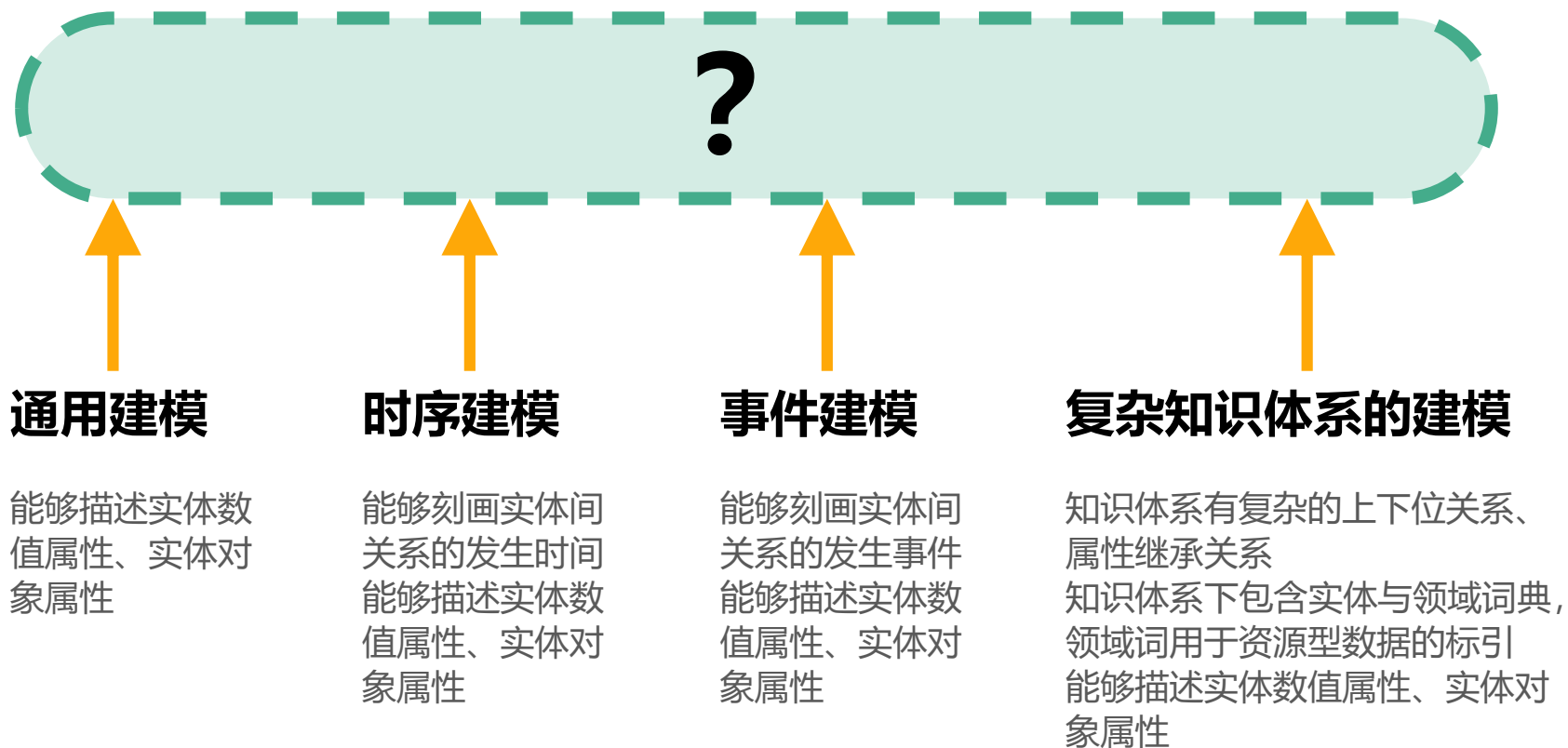
融资事件的建模

# 构建一个适用的建模工具（4）

---

- 支持大数据量的知识图谱编辑：编辑是基于底层的知识图谱存储的，每次编辑时加载到前端的仅为当前相关的数据，因此不会造成内存溢出等问题。

# 知识建模总结



需要统一的知识建模工具来解决所有建模需求！

---

# **RDF store or Graph store?**

---

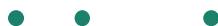
**知识图谱数据存储  $\neq$  使用图数据库**

——大规模知识图谱存储[2016]

# 知识图谱数据存储

---

**知识图谱最适合处理关联密集型的数据，因此首先需要需要存放的是节点和边的数据。**



# 知识图谱数据存储

---



在现实应用中，还有一些不适合放入知识图谱中的数据（日志、事件，需要频繁统计、计算）怎么办？

使用其它的存储和计算框架进行存储和处理，同时实现与知识图谱中数据的链接。



**No Size Fits All !**



# 知识图谱存储方案

---

**知识图谱是基于图的数据结构，其存储方式主要有两种方式**

**1. RDF存储**

**2. 图数据库(Graph Database)。**

# Native RDF 存储方案

---

六重索引：SPO、SOP、PSO、POS、OSP、OPS

## 优点：

- 三元组模式查询 (triple pattern) 的高效执行
- 任意两个三元组模式的高效归并连接 (merge-join)

问题：1、6倍空间开销，如果是 (S, P, O, C) 四元组呢？

2、更新维护代价大

# 基于Graph store的存储方案

---

- 图数据模型（属性图）
- 节点、边、节点属性、边属性
- **优点：**图查询语言、图挖掘算法
- **缺点：**
  - 存储缺点：分布式存储实现代价高，数据更新速度慢，大节点处理慢
  - 消费缺点：不支持推理、问答等语义类消费

# 现有产品：Virtuoso

---

OpenLink Virtuoso是一个混合模型数据库，实现了Web，文件和数据库服务器功能以及本机XML存储和通用数据访问中间件，提供单一服务器解决方案。

Virtuoso提供了扩展的对象关系模型，它提供了关系型数据访问，类型继承，运行时数据类型，迟绑定，基于id的访问等数据访问方式。

提供CSV、RDF文件形式的数据导入，支持从外部的数据库进行库、表、数据的关联。

缺点：

- 联合查询效率较低
- 图算法的支持不足

# 现有产品： neo4j

---

- Neo4j is a highly scalable native graph database that leverages data relationships as first-class entities, helping enterprises build intelligent applications to meet today' s evolving data challenges.

## ● 特点

- 节点存储 (node store)
- 关系存储 (relationship store)
- 属性存储 (property store)



## ● 不足

- 企业数据管理场景下不使用 Schema 会难以从整体把握数据
- 非企业版本受数据量、查询速度等方面的限制
- 不支持时态信息的存储
- 没有顶层Schema约束

# 现有产品：Titan

---

Titan是一个分布式的图数据库，支持横向扩展，可容纳数千亿个顶点和边。

Titan支持事务，并且可以支撑上千并发用户和 计算复杂图形遍历。

## Titan的特性

- 弹性与线性扩展、分布式架构、可容错
- 支持 ACID 和最终一致性.
- 支持多种存储后端，支持高可用和热备
- 支持位置、数字范围和全文检索
- 原生支持 TinkerPop 软件栈

- 停止维护2年后，于2015年2月被DataStax收购



# 时序知识存储

---

- **时间点**

基于点的时间元素 (Point - based) , 又称为时刻 (Time Points)

- **时间区间**

时间区间是指一段时间, 有固定的起止时间点

- **时间跨度**

指持续的一段时间, 表示时间的长度; 与时间区间不同的是, 时间跨度没有时间起点, 也没有时间终点。

- **时间元素Now**

Now是一个时间变元, 随着当前时间的变化而变化, 记录了随时间变化的信息, 它的有效值依赖于当前时间。

- **时间粒度**

指描述时间数据的最小时间单位, 是表示时间点之间离散化程度的因素; 时间粒度反映了时态信息系统中时间点描述的最小单位, 时间粒度越小, 离散的时间点越多, 描述的事件的变化信息越精细准确

。

# 四种时态数据库



## 快照数据库

以在特定的时刻的瞬间快照来建立模型，来考虑现实世界，现实世界是变化的，快照数据库只是反映了某一个瞬间的情况。



## 回滚数据库

支持事务时间，按事务时间进行编址，保存了每次事务提交，状态演变之前的状态。



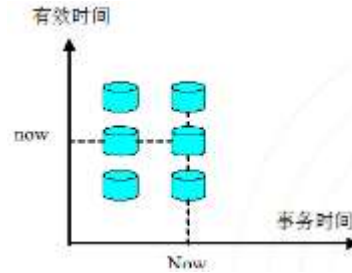
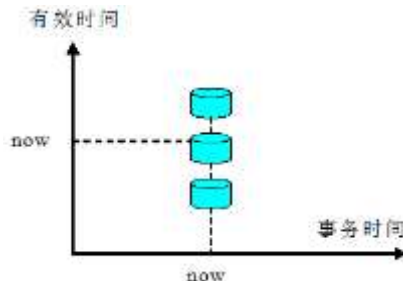
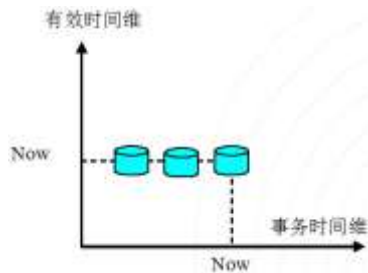
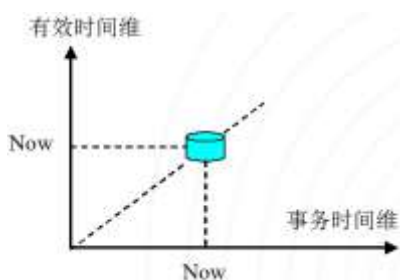
## 历史数据库

记录事实的有效时间的数据库称为历史数据库



## 双时态数据库

双时态数据库由时态关系组成，其时态关系是一个四维结构





# 知识图谱时态信息存储

---

- 在知识图谱存储中应用的为**历史数据库**，同时也借鉴**回滚数据库**的思想进行知识图谱的版本信息管理。
- **历史数据库**：记录事实的有效时间，用有限的数据冗余实现数据时态信息的应用。
- **回滚数据库**：在特定的时间点（可视为回滚点，但不用于真实的回滚操作）对知识图谱数据进行独立的数据备份，形成知识图谱的版本记录，实现对特定历史状态的查询，满足一些特定应用场景的使用需要。  
。

# 大规模知识图谱存储解决方案

## 指导思想

### 数据思维

依据数据特点进行数据存储结构  
选择与设计

### No Size Fits All

没有一种通用的存储方案能够解  
决所有问题



## 整体原则

基础存储支撑灵活  
基础存储可扩展、高可用  
按需要进行数据分割  
适时使用缓存和索引  
善于利用现有成熟存储  
保持图形部分数据的精简  
不在图中作统计分析计算  
在应用中进行扩充迭代

# 知识图谱存储方案设计思路

---

**关系型数据：**使用图形数据进行存储

**关联型数据：**作为记录型数据存入合适的存储中，通过实体链接与图谱中实体关联。

**属性型数据：**作为实体的数值属性存入知识图谱

**时态型数据：**使用基础存储上的时态处理中间件进行存储与查询

# 知识图谱数据存储方案设计思路

---



那应该如何选择RDF还是Graph呢

**视数据消费场景而定：**

对于不需要进行关系延伸计算的数据不放入图谱，这些数据可使用适应的存储并与知识图谱中实体作链接；  
对于结构固定、实体属性信息丰富的实体类，使用其它数据库存储更能体现优势；



**数据消费思维！**

# 现有知识图谱存储解决方案对比

- **RDF store**

无法满足遍历、路径等图计算消费场景

原因：存储结构决定了数据连接操作开销大

- **Graph store**

无法支持语义搜索、问答、推理等知识消费类场景

原因：没有顶层模式定义和约束，语义不明确

**Time  
series DB**



	知识消费	图计算消费	时序消费
RDF store	strong	weak	weak
Graph store	weak	strong	weak

**无法支撑知识图谱的统一消费！**

**知识消费**

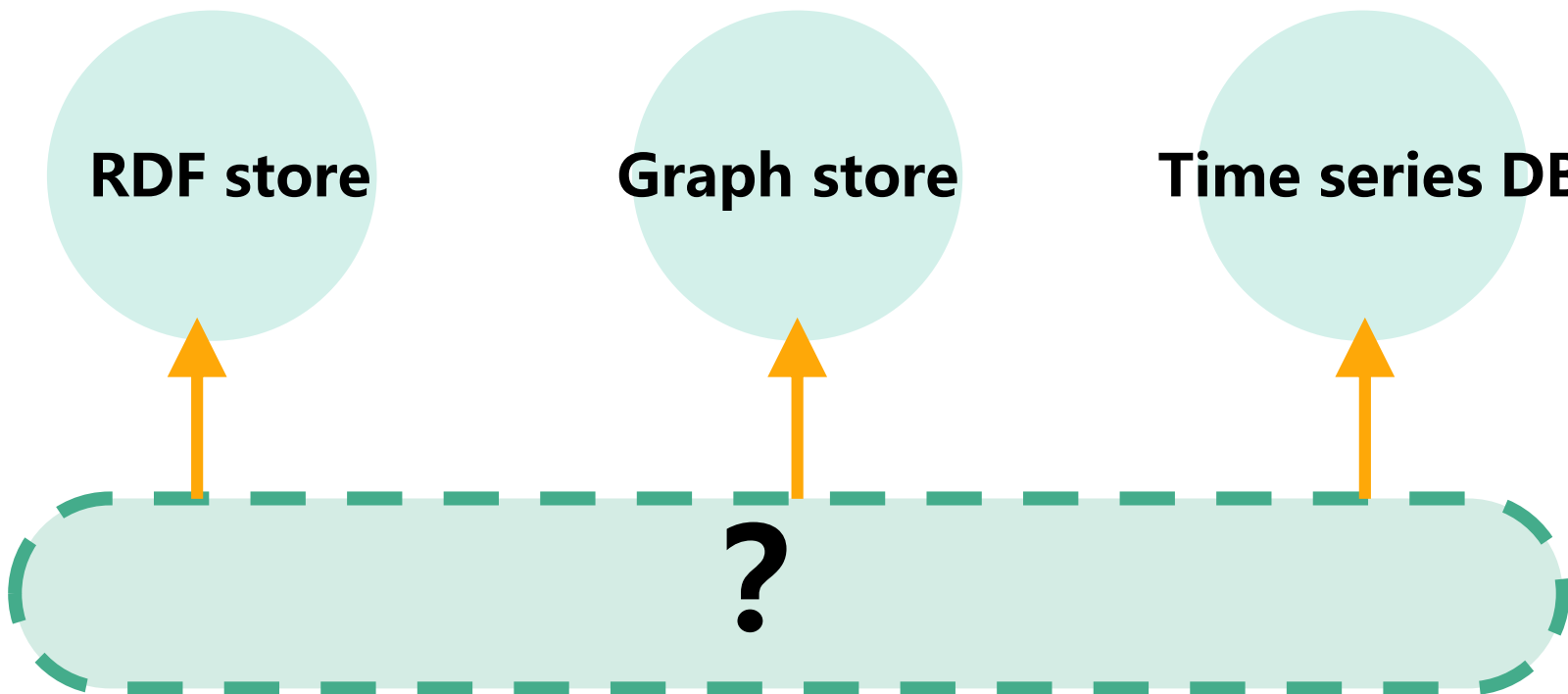
**图计算消费**

**时序消费**

**RDF store**

**Graph store**

**Time series DB**



**我们需要统一的知识图谱存储平台！**

---

# Best practice

# 现有案例数据场景总结

案例	数据场景	概念	实体	关系	属性	事件	时序
 舟山海洋数字图书馆 ZHOSHAN OCEAN DIGITAL LIBRARY	海洋知识体系、鱼类品种分类体系、论文、专利、书籍	√	√	√	√		
 NARI 国电南瑞 NARI-TECH	百科知识体系、论文、专利、标准、成果、实验室、图片、视频、人物、企业、科研机构、知识点	√	√	√	√		√
 同济大学图书馆 Tongji University Library	柴油发动机分类体系，产业链，专利、论文、研究机构、专家	√	√	√	√		√
 科创365	专利分类体系、专利、企业、人物、科研机构	√	√	√	√		√
 HUAWEI	知识点、专家、研究机构、企业、论文、专利	√	√	√	√		√
 农业知识图谱 Agriculture Knowledge Graph	水稻品种、产业链、病虫害、农作物、专利、论文、成果、政策法规、新闻	√	√	√	√		√



# 现有案例数据场景总结

案例	数据场景	概念	实体	关系	属性	事件	时序
	产业上下游体系、产品、企业、科研机构、人物	√	√	√	√		√
	商品分类体系、商品、品牌、店铺、分站、地域	√	√	√	√		
	人物、公司、行业、地域、股票	√	√	√	√	√	√
	产品、企业、投资机构、行业、融资事件、人物	√	√	√	√	√	√
	产业链、产品、公司、专利	√	√	√	√		
	开发商、人员、部门、客户、项目、合同、供应商、流程、地点、行业	√	√	√	√	√	

# 知识消费场景

实践案例	知识消费			图计算			时序消费
	知识问答	语义搜索	智能提示	图探索	路径	关联	时序图探索
 舟山海洋数字图书馆 ZHOSHAN OCEAN DIGITAL LIBRARY		√	√				
 NARI 国电南瑞 NARI-TECH		√	√	√	√	√	
 同济大学图书馆 Tongji University Library			√	√	√	√	√
 科创365		√	√	√	√	√	
 HUAWEI			√	√	√	√	√
 农业知识图谱 Agriculture Knowledge Graph		√	√	√	√	√	√

# 知识消费场景

实践案例	知识消费			图计算			时序消费
	知识问答	语义搜索	智能提示	图探索	路径	关联	时序图探索
 湖南工业大学 HUNAN UNIVERSITY OF TECHNOLOGY		√	√	√	√	√	√
 国网商城 www.esgcc.com.cn		√	√				
 量子数聚 QUANTUM DATA			√	√	√	√	√
 LengJing.io		√	√	√	√	√	
 科创云脑 inobrain.com		√	√	√			
Landray 蓝凌	√			√	√	√	√

# 知识建模案例-电商搜索

## 数据源



## 数据说明

商品明细 – 商品详情

商品属性 – 从商品1-3级分类共有的属性

商品分类 – 分1、2、3级分类

品牌 – 商品品牌、品牌简称、品牌英文、logo

店铺 – 店铺、店铺地域及店铺与商品的关系

订单 – 订单、订单与用户、订单与商品的关系

评论 – 评论、评论与商品、评论与用户的关系

用户 – 用户及其画像

专区 – 商品业务上的分区

分站 – 商品业务上的虚拟分站

- 商品体系复杂，多层次多类型，商品不同品类有公有属性、也有私有属性
- 商品属性需要做数据质量控制\需要通过图谱点概率影响搜索
- 搜索输入的联想及纠错

# 搜索意图识别

图谱具有搜索意图识别可优化精确搜索的结果

The diagram illustrates the process of search intent identification and optimization using a knowledge graph. It is divided into three main sections:

- Initial Search Results (Left):** A grid of 12 items. The top row contains four air conditioning units (AUX, Haier, Gree, GMCC). The bottom row contains four rice cookers (AUX, Midea, Panasonic, etc.). An orange arrow points from the text "我们要的!" (What we want!) to the top row of air conditioners.
- Optimization Process (Middle):** A large yellow arrow labeled "优化后" (After Optimization) points from the initial results to the refined results.
- Refined Search Results (Right):** A grid of 12 items, all of which are air conditioning units. An orange arrow points from the text "搜索结果精确!" (Search results are precise!) to this grid.
- Unwanted Results (Bottom Left):** A grey arrow points from the text "不准确的搜索结果" (Inaccurate search results) to the bottom row of rice cookers in the initial results section.

- 搜索词“奥克斯空调”在原系统返回结果中混合“空调、风扇、电饭煲”，结合电商知识图谱的系统，能正确返回所有“奥克斯的空调”

# 搜索结果优化

商品知识图谱拥有的概念属性支持更加丰富的搜索结果页筛选条件



笔记本电脑 商品筛选

相关分类：笔记本电脑配件 笔记本电脑电池 笔记本电脑

店铺：华硕电脑专卖店 京东商城 浙江物产电子商... 史泰博国际商城

品牌：戴尔/DELL 福正(FUZHENG) 惠普/HP 联想(Lenovo) 华硕/ASUS 海康威视/HIKVI... thinkpad 3M 双飞燕/A4 联想/LENOVO 范罗士 优越者/YOUYUE... TP-LINK

优化后

品类 笔记本电脑

品牌

店铺 惠普贸易(上海... 联想官方旗舰店 华硕电脑专卖店

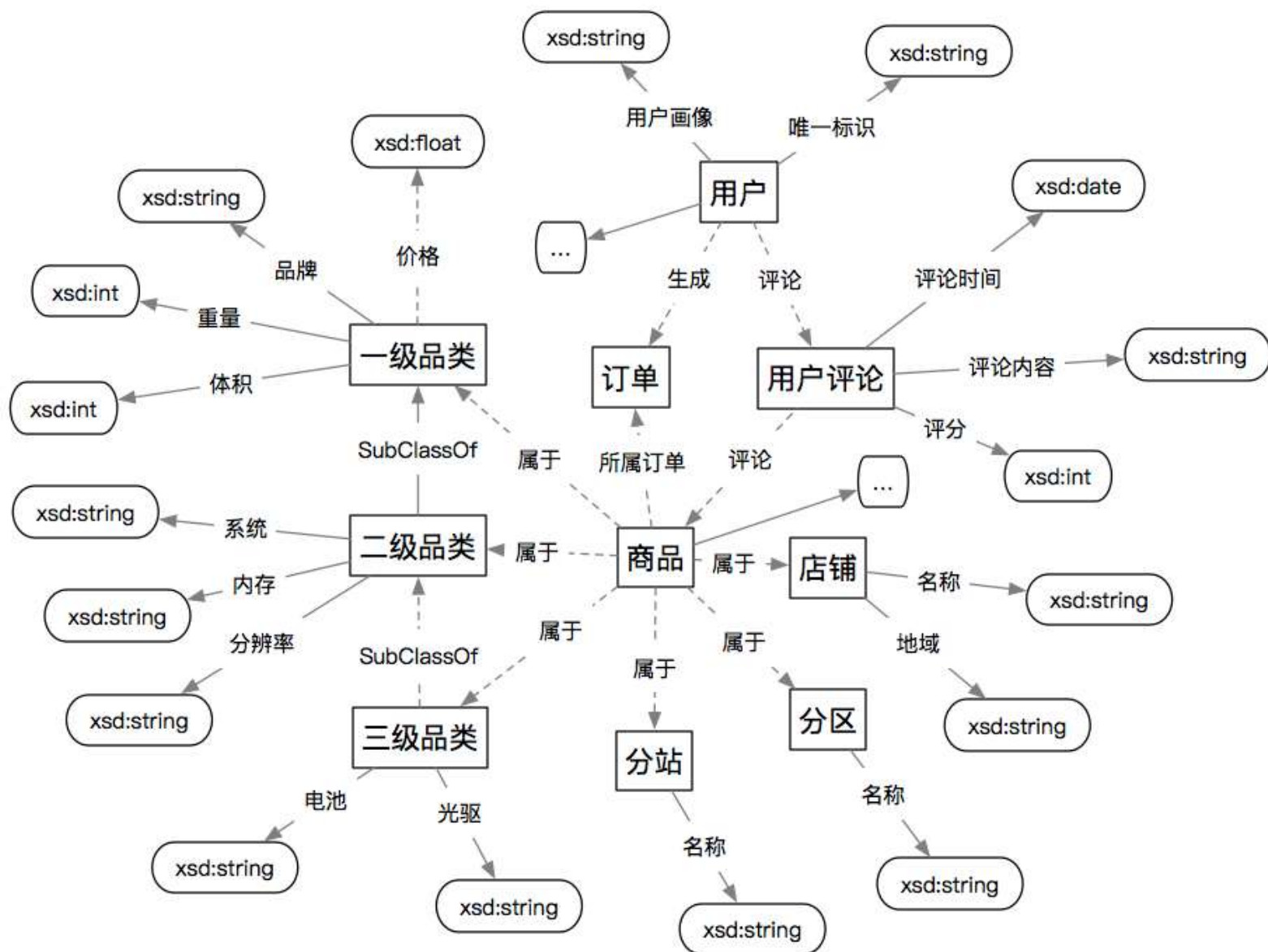
屏幕尺寸 13.3英寸 14英寸 16英寸

光驱 内置光驱 外置光驱 无

厚度 15.8mm 16.45mm 16.9mm 18.9mm 20.9mm

- 商品知识图谱支持商品品类层级及品类下：品牌、店铺、个性化属性的表达及质量控制。
- 搜索“笔记本电脑”，原系统只存在品类、店铺及品牌几个过滤筛选项，且品牌存在错误。在现系统中，支持品类、店铺、品牌（文字+logo）及属性（屏幕尺寸、光驱、厚度）的过滤筛选。

# 电商搜索-建模结果



# 知识建模案例-科技知识服务

## 数据源



## 数据说明

百科 - 从行业百科全书+百度、维基、互动百科构建知识体系

专利 - 专利数据, 分发明专利、实用新型、外观设计

论文 - 论文数据, 分期刊、学位等

成果 - 成果数据

标准 - 标准数据, 国标、行标、企标

图书 - 一般的书籍数据

图片视频 - 带文本描述的图谱、视频

专家学者 - 上述数据的作者 + 专家库

行业机构 - 上述数据的作者 + 行业机构库

企业 - 上述数据的作者 + 行业企业库

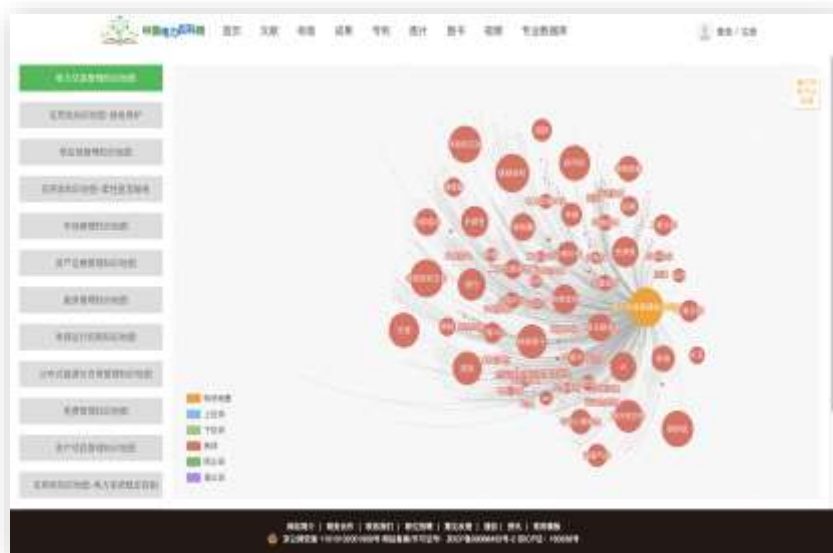
- 多源异构数据统一到一个知识体系上
- 不同类型搜索数据的个性化筛选项
- 决策支持-不同类型数据的统计,人物及资源合并



## 相关推荐

## 相关推荐

- 利用知识图谱的语义信息和大数据平台的计算能力，对行业知识进行分析和可视化展示
- 电力行业知识信息展示，包括知识点、行业热词、上下位推理、语义距离



# 搜索-一框式搜索



- 知识图谱能够辅助识别用户的查询意图，跨资源类型进行搜索
- 知识数据改善搜索体验，对搜索结果进行知识展示，包括图谱、电力行业知识体系、资源分类、时间统计等
- 用户能够根据标注结果、时间、资源类型进行筛选

# 决策支持

### 原电池

放电后不宜用充电方法使其再次获得放电能力,即反应是不可逆的化学电源。原电池是经常处于可工作状态,充分放电后只得丢弃的电池,又称为非贮备式电池、一次电池。

原电池的电极反应式可写成: 负极活性物质 | 电解质 | 正极活性物质 $\oplus$ , 式中“|”代表界面。当接通外电路时, 负极活性物质发生氧化反应, 释放出电子, 经外电路流至正极, 正极活性物质接受电子发生还原反应。这两种反应都发生在电

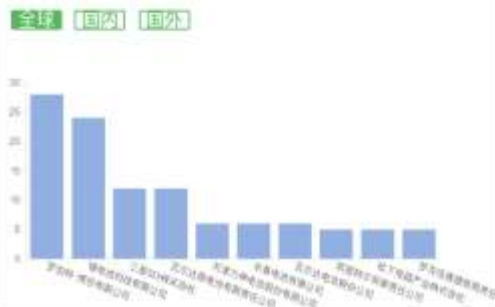
原电池 相关技术关键词词云



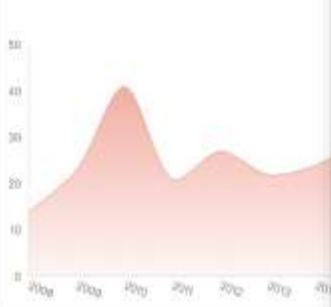
原电池 图谱



原电池 top10 专利数目机构

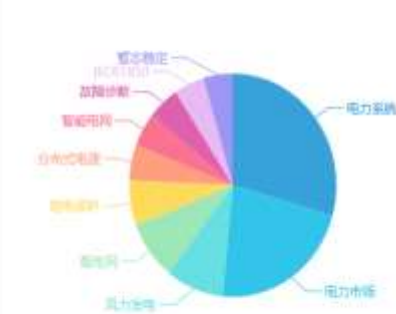


### 原电池 相关专利数量趋势



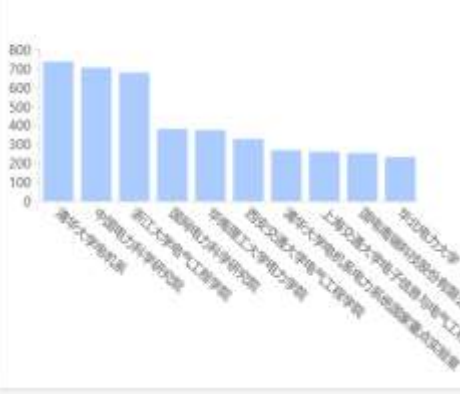
### 电力系统自动化 研究主题

发文量排行



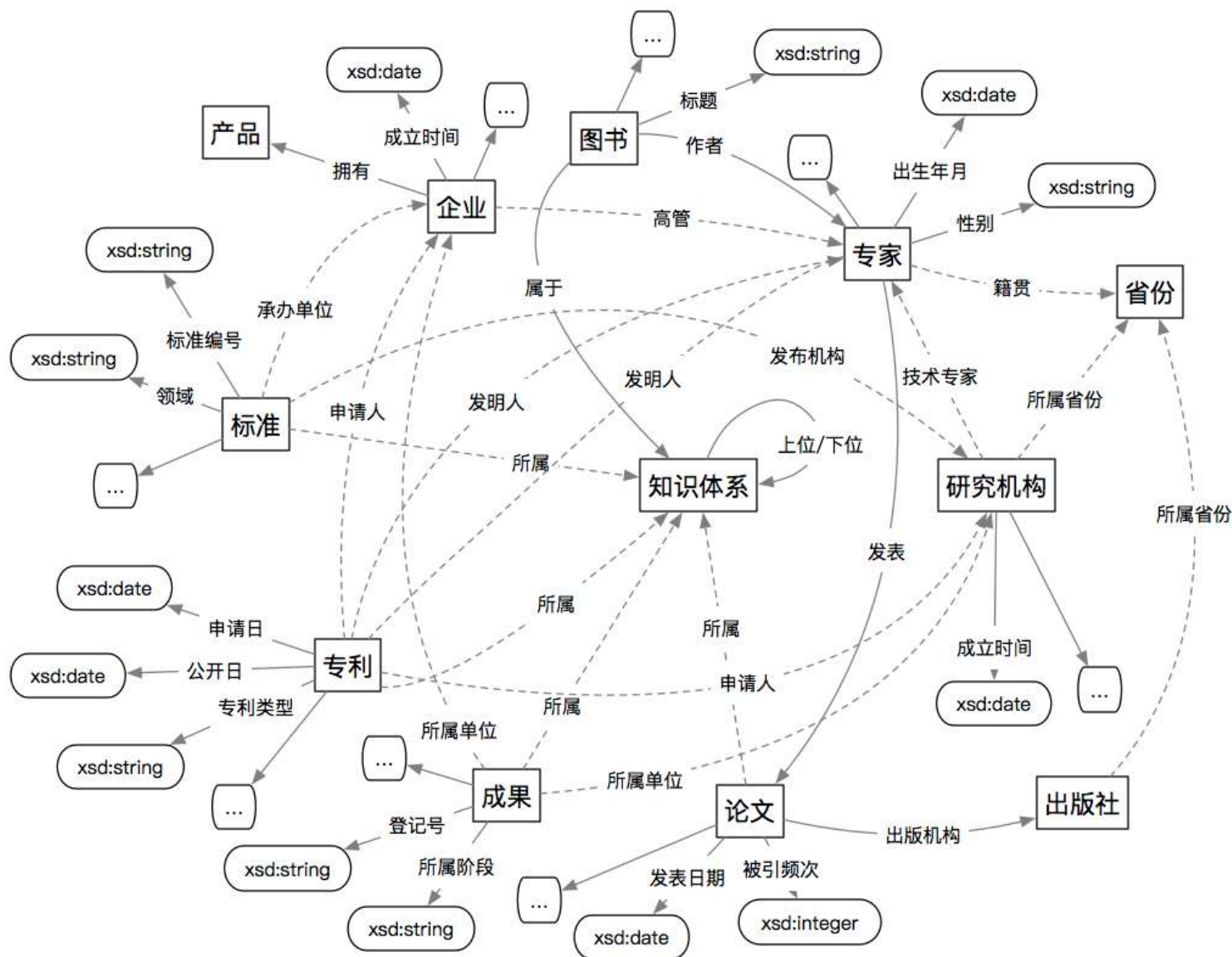
## 电力系统自动化 机构分析

黄文雷 徐行



- 通过对多源异构的知识数据进行融合，从而能够从多维度进行统一分析
- 从人物、机构、知识点、期刊等维度进行分析和可视化展示，辅助决策。

## 知识资源服务-建模结果



# 知识建模案例 创投分析

## 数据源



## 数据说明

产品 – 产品及其多种同义

投资机构 – 融资事件中的投资方

融资事件 – 企业的投融资事件

标签数据 – 刻画融资事件与人物的背景标签

企业 – 初创企业

行业 – 企业所在行业

人物 – 企业管理者

高校 – 高校，主要刻画人物教育背景

- 支持问答
- 支持会话式统计
- 能够对事件发生的时间进行表示
- 能够对事件发生的参与方进行表示

# 问答式创投关系发现

实体问答、知识问答、会话式统计分析



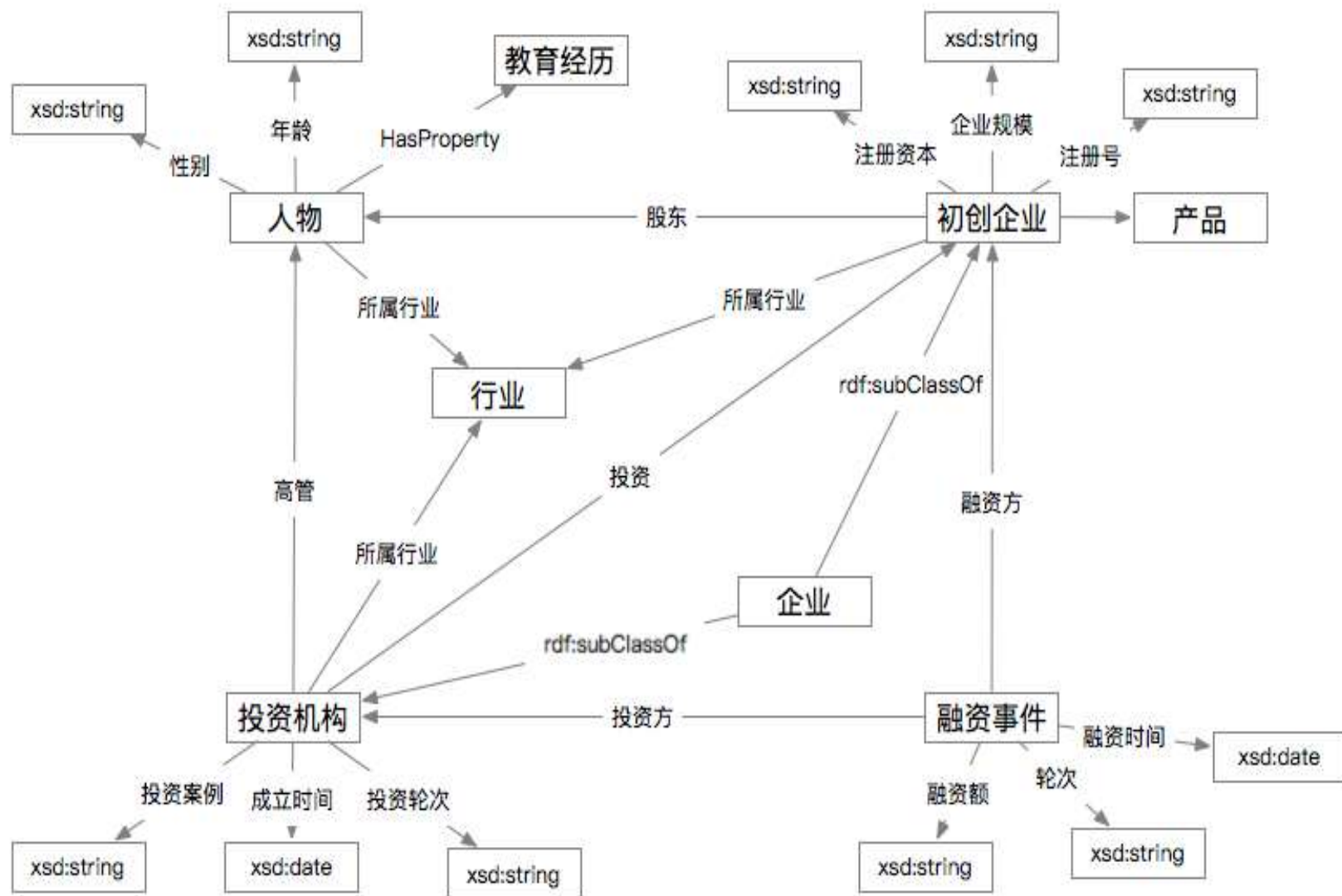


# 会话式创投关系发现

## 会话式路径发现、会话式关联分析



# 创投分析-建模结果





# 知识建模案例-金融

## 数据源



## 数据说明

专利数据 – 全国专利数据

公司 – 全国4000万家公司，包括上市公司

人物 – 企业法人、高管

失信 – 全国法院失信信息

诉讼 – 全国法院诉讼信息

招投标 – 上述公司招投标信息

新闻 – 上述公司新闻数据

招聘 – 上述公司招聘信息

### 建模难点：

- 行业分类体系定义
- 投资关系边属性定义
- 新数据源加入后，模式的动态扩充

# 风险评估

综合评估 ( 2015-05 )



综合信用等级

AAA



综合信用评价

低风险



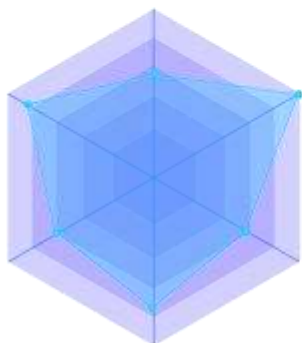
综合信用评分

757

身份特征

守法守规

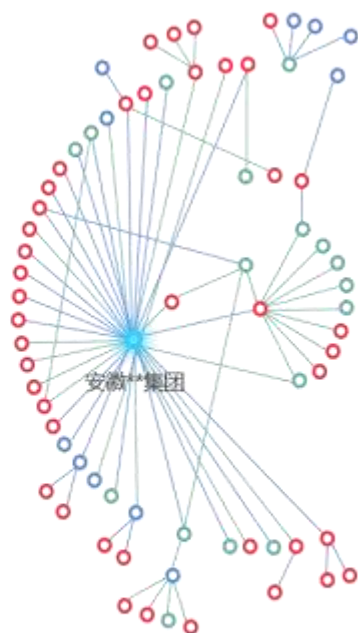
发展潜力



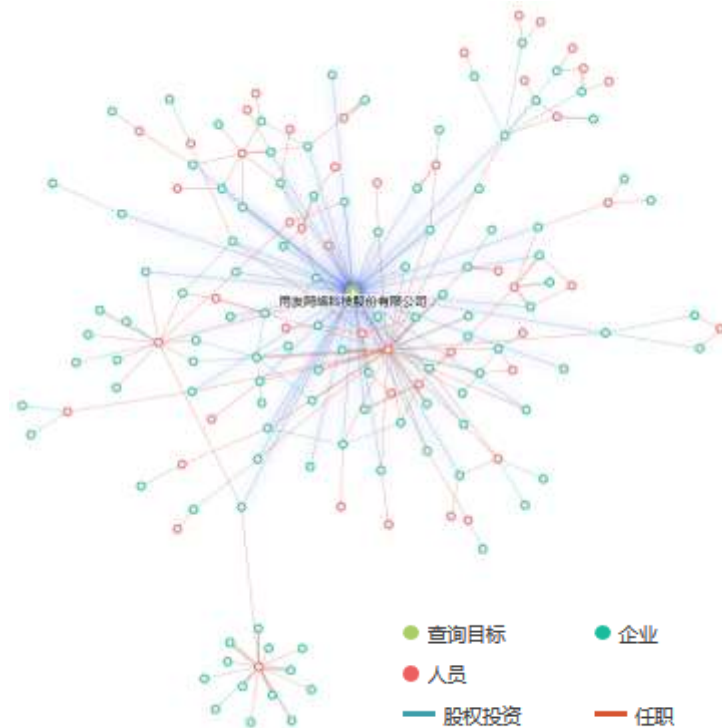
经营状态

创新资质

行为偏好



低风险 中等风险 高风险

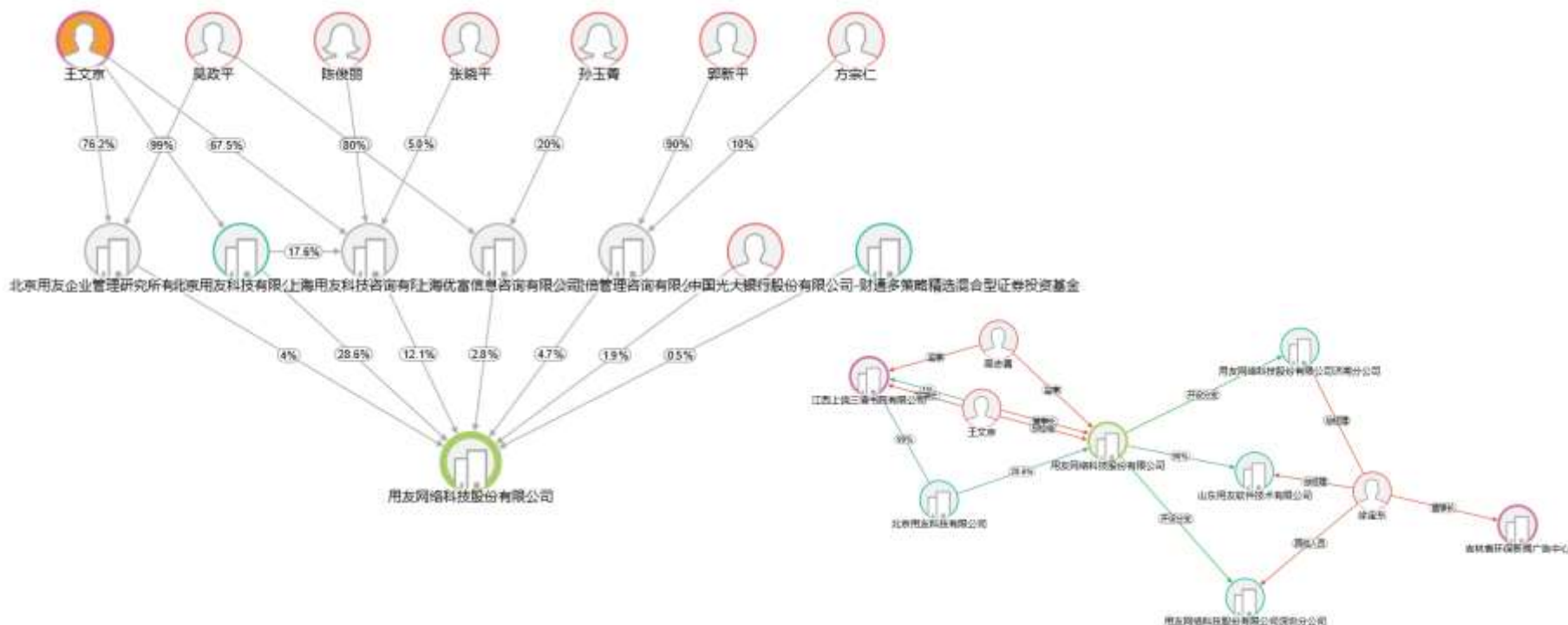


银行、投行、政府等机构可对客户的担保等行为，利用企业图谱进行**客户资源分类管理、信贷前期风险评估、采购企业风险审核、招投标企业资质评级**等。

基于投资、任职、专利、招投标、涉诉关系以目标企业为核心向外层层扩散形成一个网络关系图，直观立体展现企业关联。

## 企业最终控制人

## 利用图谱边属性记录的股权投资关系可追溯企业最终控制人

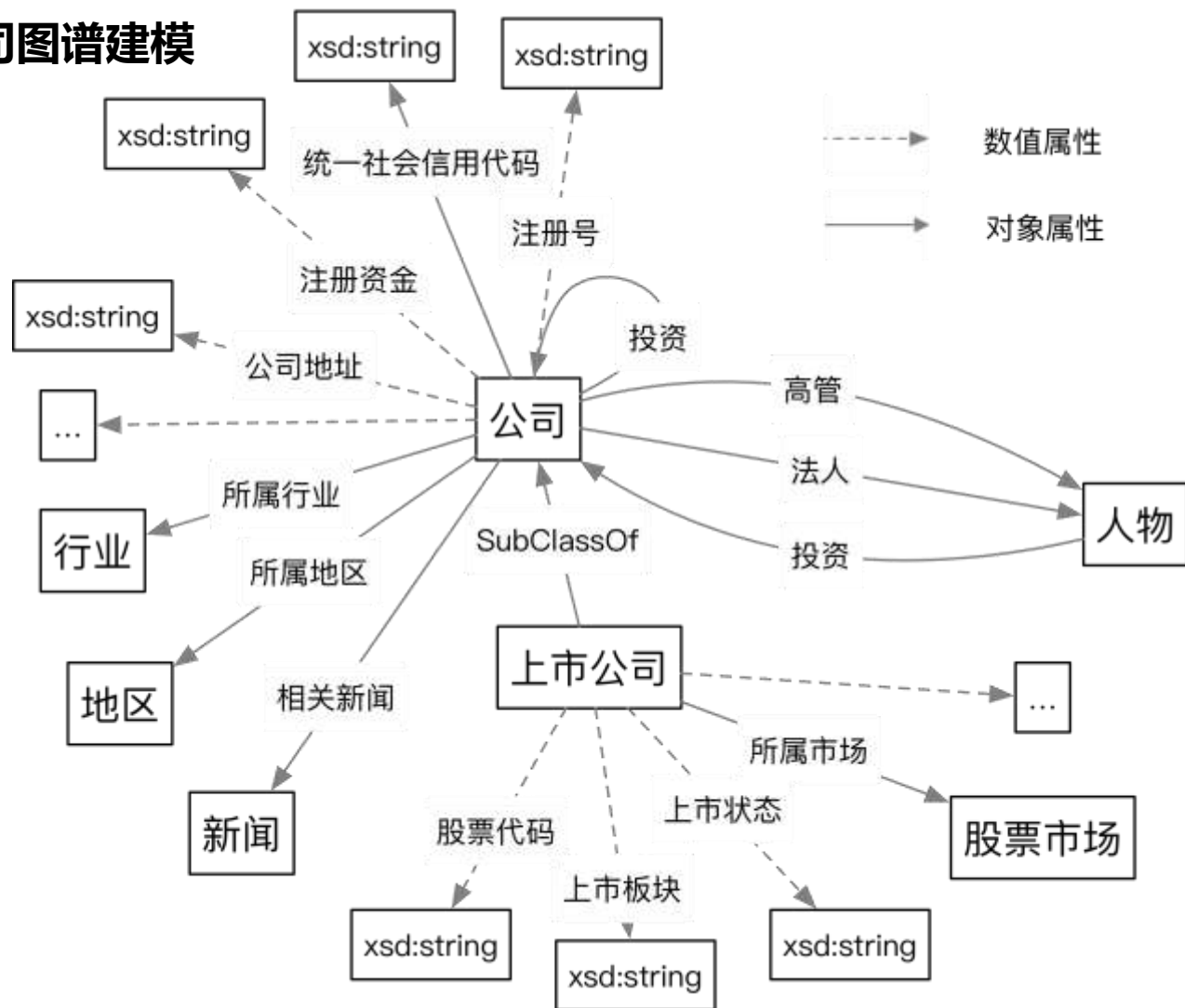


基于股权、任职、专利、招投标、涉诉关系形成的网络关系中，查询企业之间的最短关系路径，衡量企业之间的联系密切度。

基于股权投资关系寻找持股比例最大的股东，最终追溯至自然人或国有资产管理部門。

# 建模结果

## ● 上市公司图谱建模



# 总结

---

实体 关系  
概念 事件  
属性 时序

自主研发的知识图谱  
统一存储平台

统一知识建模

统一知识存储

```
graph TD; KGMS[KGMS] --> UKM((统一知识建模)); KGMS --> UKS((统一知识存储));
```

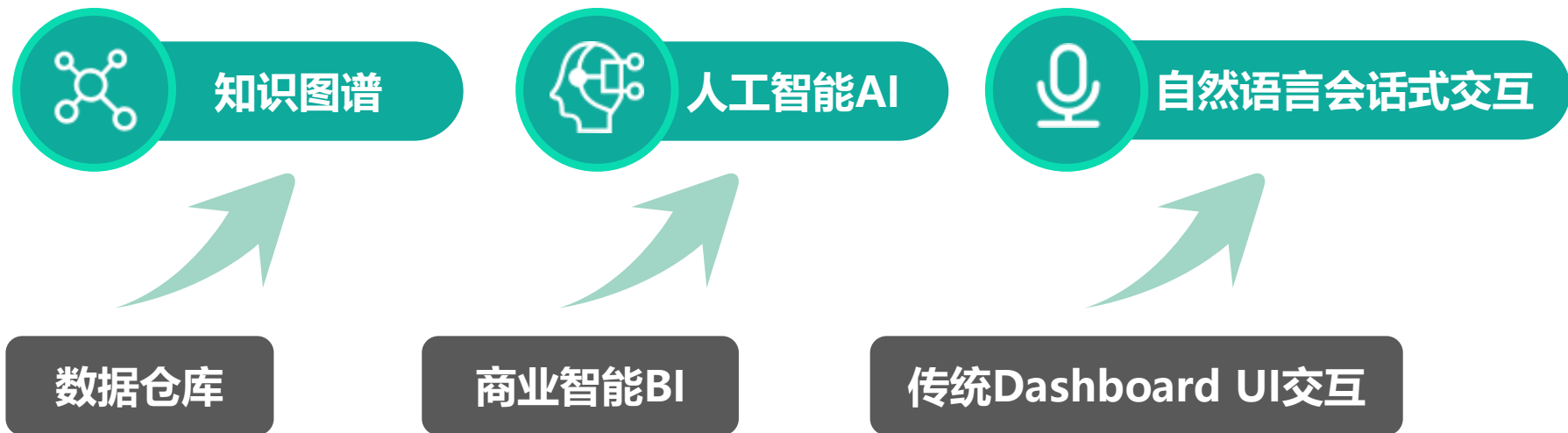
KGMS

---

# PlantData KGMS

# 知识图谱发展趋势

- 互联网的不断发展，链接了越来越多的事物，在此其中，不仅仅是数据或信息的爆炸，其本质是**知识的急剧增长**。PlantData KGaaS提供一套从知识构建到消费的解决方案，降低知识使用的门槛，发挥知识的最大价值。



# KGMS与PlantData KGaaS

## ● 产品体系及KGMS定位

KGMS 图谱管理工具集及图谱存储

KGBuilder 图谱自动化工具集

KGRobot 会话式图谱开放机器人

KGPro 图谱可视化深度分析工具

KGSenor 智能文档辅助编辑工具





# KGMS

---

- KGMS知识大脑的手工构建及运营管理问题
- 最简单直接的一站式**构建、管理、使用**知识图谱的产品，人人都能玩转知识图谱
- 使用场景



## 企业

快速构建针对业务场景的应用，做poc验证  
自动化流程后的数据运营



## 研究者

知识图谱构建、建模研究，可视化  
rdf导出，参与相关模型计算



## 开发者

二次开发，基于KGaaS的个性化应用

# KGMS的功能

---



## 知识建模工具

知识体系建模  
属性的定义



## 知识编辑工具

图谱构建  
数据修正



## 数据入图

以运营人员的视角  
完成外部数据到知  
识图谱的构建



## 数据标引

资源型数据的人工  
标引及修正

# KGMS的优势

---



—

**业界领先的知识图谱管理及存储工具**



—

**丰富的实战应用经验，知识图谱构建各环节针对性优化**



—

**JSSDK  
Restful SDK  
轻量级一站式图谱  
构建+应用**



—

**支持KGaaS  
支持私有化部署**

# KGMS 功能流程

- **KGMS提供一整套手工构建知识大脑的工具集**

知识建模、知识编辑、知识融合、数据接入、数据入图、数据标引及知识图谱的底层存储+sdk



# KGMS 我的图谱

- 图谱管理、图谱编辑、高级操作

功能	功能点
图谱管理	图谱的增删改
图谱编辑	概念编辑 属性定义 实例编辑 关系编辑 同义编辑 领域词典 实体融合 图谱可视化
高级操作	分析视图导出 实体导出、推荐问题导出

# KGMS 构建知识图谱的基本操作

---

- 1.概念编辑 – 定义知识体系
- 2.属性定义 – 定义数据的模式
- 3.实体编辑 – 添加实例数据、实例属性数据
- 4.关系编辑 – 添加实例关系
- 5.边关系编辑 – 添加实例边关系属性值

# KGMS 知识图谱编辑工具

---

- **概念编辑**

知识图谱的知识体系人工构建

- **场景**

多层级概念作为资源的统一分类

多层级概念属性继承

问答模式识别的需要

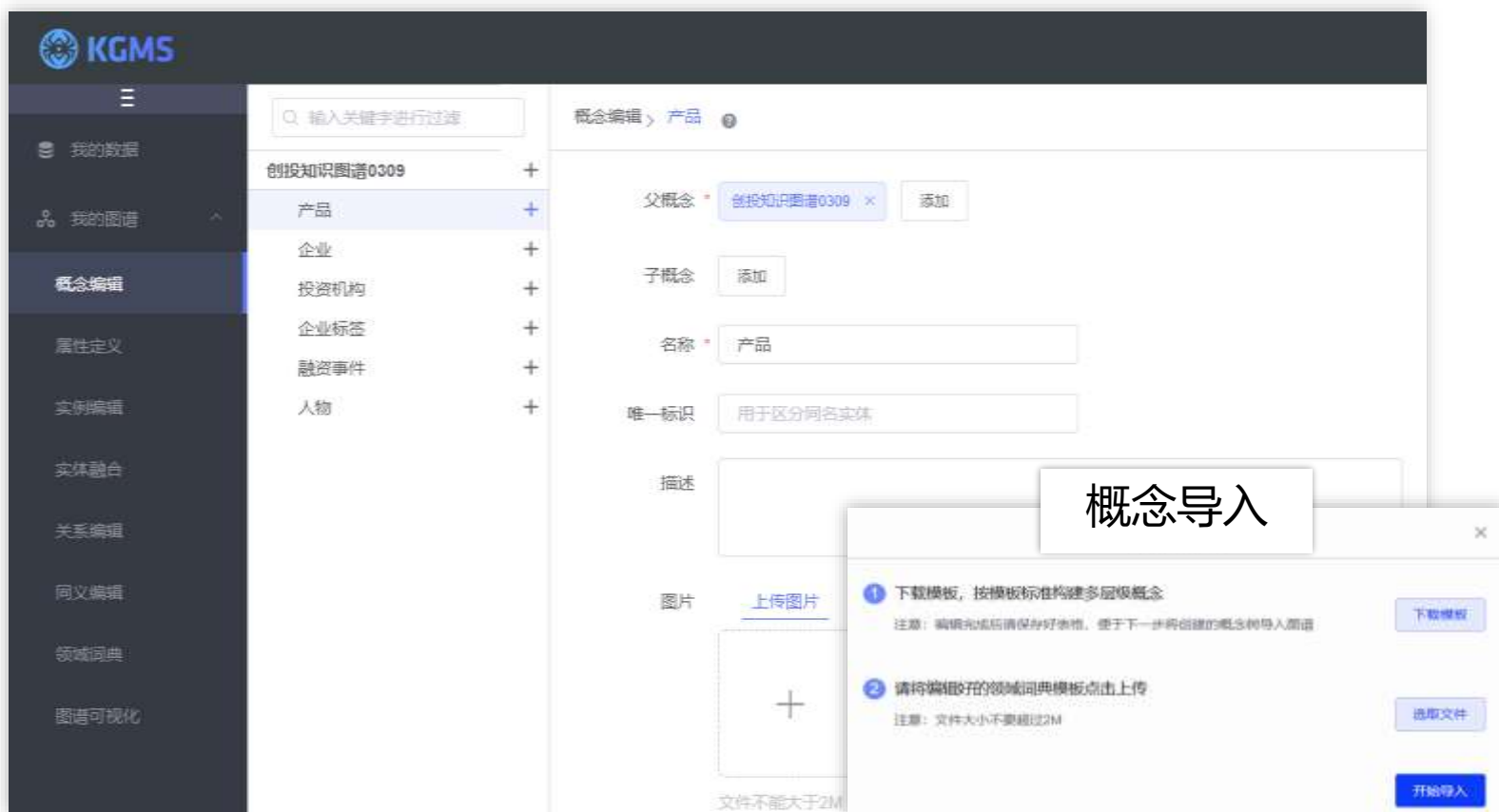
# KGMS 概念编辑

## ● 概念编辑

上下位关系定义

基本名称、唯一标示、描述、图片

文件导入





# KGMS 概念编辑

## ● 知识体系



# KGMS 概念编辑

## ● 资源搜索统一整合

标准行业知识体系分类是进行多资源数据整合的基础

知识体系

▼ 新能源发电(5390)

> 新能源发电概论(1933)

> 风力发电机组(3189)

▼ 输电与变电(52870)

> 交流变电站(51877)

▼ 电力综合知识(61384)

> 电力基本概念(20334)

> 电力标准化(7643)

> 电力可靠性管理(19363)

> 电力生产安全管理(24628)

> 电力系统(33236)

> 核能发电(1503)

⋮

按资源类型分类

• 成果(455)

• 专利(15737)

• 文献(45129)

• 标准(63)

检索历史

当前无相关数据

更多

相关度

### 1. 变电器检修的接地故障处理方法 期刊论文

全洪杰, 张景钰, 张政, 杨玉贵, 吕广辉, 薛乃良, 白云, 吕广臣

摘要: 变电器的接地故障检修工作由多个作业项目构成, 要求相关技术人员不仅要数量掌握多种维修技术, 同时需应对不同技术维修问题, 从而更为有效的组织变电器接地故障检修工作。为进一步提高变电器接地故障处理实际有效性, 本...

来源: 万方

< > 引用 批量引用 加入收藏 下载

### 2. 变电器检修的接地故障处理方法浅析 期刊论文

王少玄 《科技创新与应用》

摘要: 科学技术的日益进步, 对电力行业规模的不断扩大起到了积极的推动作用.为保证电气设备的良好运行, 各单位需要在综合评价的基础上, 制定系统的检修方案, 按期进行严格细致的检修.其中, 变电器检修虽然属于比较基础的工...

来源: 万方

< > 引用 批量引用 加入收藏 下载

### 3. 变电器检修的接地故障处理方法浅析 期刊论文

王少玄 《科技创新与应用》

摘要: 科学技术的日益进步, 对电力行业规模的不断扩大起到了积极的推动作用.为保证电气设备的良好运行, 各单位需要在综合评价的基础上, 制定系统的检修方案, 按期进行严格细致的检修.其中, 变电器检修虽然属于比较基础的工...

来源: 知网

< > 引用 批量引用 加入收藏 下载

### 4. 高低压变电器原理分析 期刊论文

郭德军

zdp

人物 朱德苹 安徽扬仕电子科技有限公司  
詹敦平 江苏飞格光电有限公司

机构 中电普瑞科技有限公司  
中电普瑞电力工程有限公司  
中电普华信息技术有限公司

# KGMS 概念编辑

## ● 概念体系是知识问答的基础

通过属性的搭配，能够回答如：

1. 2018年**融资事件**
2. 2018年大于1000万元的**融资事件**
3. 2018年新发布的**产品**
4. 2018年人工智能新发布的**产品**等



融资事件



产品

# KGMS 知识图谱编辑工具

---

- **属性定义**

- 数值属性

- 对象属性

- 边的数值属性

- 边的对象属性定义

- **场景**

- 数值属性，人的身高体重

- 对象属性，企业与投资机构的投资关系

- 边的数值，关系时间的刻画

- 边的对象属性，引起关系生的事件

# KGMS 数值属性定义

## ● 数值属性定义

属性名称

属性同义

数据类型

举例：

人的身高、体重

BMI

数据类型
整数值
浮点值
布尔值
日期
字符串
Map型
计算属性



# KGMS 数值属性定义

- 数值属性定义后，实例编辑自动生成表单

The screenshot displays the KGMS web application interface. On the left is a dark sidebar with a menu containing items like '我的数据', '我的图谱', '概念编辑', '属性定义', '实例编辑' (highlighted), '实体融合', '关系编辑', '同义编辑', '领域词典', and '图谱可视化'. The main content area has a top navigation bar with the KGMS logo and a search bar. Below the search bar is a breadcrumb trail: '返回 > 实例编辑 > 列表 > 北京拜克洛克科技有限公司'. A sub-menu on the left lists categories: '创建知识图谱0309', '产品', '企业', '投资机构', '企业标签', '融资事件', and '人物'. The '实例编辑' form for '北京拜克洛克科技有限公司' contains the following fields:

- 名称**: 北京拜克洛克科技有限公司
- 唯一标识**: 用于区分同名实体
- 简介**: 拜克洛克科技是一个无桩共享单车平台，用户通过微信服务号或移动APP输入车牌号即可获得密码解锁单车，随取随用，此外，还可以共享自己的单车给平台经营，从而获得所有ofo共享单车的免费使用权。
- 照片**: Includes a '上传图片' link and a placeholder box with a '+' sign. A note below states '文件不能大于2M'.
- 线上地址**: (Link)
- 父概念**: A dropdown menu showing '企业' with a close icon and a '添加' button.
- 所属地区**: 北京市海淀区北四环西路58号11层1102室
- 成立日期**: 2015-08-06
- 注册资本**: (Empty field)

At the bottom left of the main content area, there is a button labeled '实例导入 (小于2M)'.

# KGMS 对象属性定义

## ● 对象属性定义

属性名称

属性同义

数据类型

值域



The screenshot displays the KGMS (Knowledge Graph Management System) interface for defining object attributes. The interface is divided into three main sections: a left sidebar, a central list, and a right configuration panel.

- Left Sidebar:** Contains navigation links for '我的数据' (My Data), '我的图谱' (My Graphs), '概念编辑' (Concept Editing), '属性定义' (Attribute Definition - currently selected), '实例编辑' (Instance Editing), '实体融合' (Entity Fusion), '关系编辑' (Relationship Editing), '同义编辑' (Synonym Editing), '领域词典' (Domain Dictionary), and '图谱可视化' (Graph Visualization).
- Central List:** Shows a search bar with the text '输入关键字进行过滤' (Enter keywords to filter). Below it, a list of object types is displayed: '创投知识图谱0309', '产品' (highlighted), '企业', '投资机构', '企业标签', '融资事件', and '人物'.
- Right Configuration Panel:** Titled '对象属性编辑' (Object Attribute Editing) for the '产品' (Product) object. It includes a '返回' (Return) button and a breadcrumb '对象属性编辑 > 产品'. The '基本设置' (Basic Settings) section contains:
  - '名称' (Name): A text input field with the placeholder '请输入名称' (Please enter name).
  - '别称' (Alias): A button labeled '添加' (Add).
  - '值域' (Value Domain): A button labeled '添加' (Add).
  - '属性类型' (Attribute Type): Radio buttons for '非单值型' (Non-single-valued) (selected) and '单值型' (Single-valued).A note at the bottom states: '\*如果您需要更精确的校验设置请点击[高级设置](#)进行配置' (If you need more precise validation settings, please click [Advanced Settings](#) for configuration).

# KGMS 对象属性定义

- 对象属性定义后，关系编辑自动生成表单，并限定值域

The screenshot displays the KGMS web interface. On the left is a dark sidebar with navigation options: 我的数据, 我的图谱, 概念编辑, 属性定义, 实例编辑, 实体融合, 关系编辑 (highlighted), 同义编辑, 领域词典, and 图谱可视化. The main content area is titled '北京开壳洛克科技有限公司' and shows a list of relationships with their corresponding forms. The relationships and their values are:

- 法人**: 添加, 删除, 编辑
- 股东**: 添加, 删除, 编辑; 张巴丁, 删除, 编辑; 戴威, 删除, 编辑; 于信, 删除, 编辑; 王翔, 删除, 编辑; 刘德, 删除, 编辑
- 高管**: 添加, 删除, 编辑; 朱景士, 删除, 编辑; 杨品杰, 删除, 编辑; 肖敏, 删除, 编辑; 陈皓, 删除, 编辑; 戴威, 删除, 编辑; 朱景士, 删除, 编辑; 于信, 删除, 编辑; 张巴丁, 删除, 编辑; 薛勇, 删除, 编辑
- 标签**: 添加, 删除, 编辑; 共享出行, 删除, 编辑; 交通出行, 删除, 编辑; 共享经济, 删除, 编辑; 共享单车, 删除, 编辑; 出行工具, 删除, 编辑; 餐饮服务, 删除, 编辑; 出行, 删除, 编辑; 租车, 删除, 编辑; 汽车交通, 删除, 编辑; 骑行, 删除, 编辑; 移动app, 删除, 编辑

At the bottom left, there is a button labeled '实例关系导入 (小于2M)'.



# KGMS 属性定义



# KGMS 属性定义

- 电商属性的分类导航
- 知识图谱、搜索的知识卡片

The screenshot displays the KGMS (Knowledge Graph Management System) interface. On the left, a sidebar menu lists various attributes: 品类 (Category), 品牌 (Brand), 店铺 (Store), 屏幕尺寸 (Screen Size), 光源 (Light Source), and 厚度 (Thickness). A yellow thumbs-up icon highlights the 品牌 (Brand) attribute. The main content area shows a search results page for '北京拜克洛克科技有限公司' (Beijing Baidu Luoke Technology Co., Ltd.). A detailed information card for this company is displayed, including its introduction, address, establishment date, and key personnel. In the background, a knowledge graph is visible, showing relationships between entities like '共享单车' (Shared Bicycle), '共享出行' (Shared Mobility), '主刚' (Zhu Gang), '微信服务' (WeChat Service), '子信' (Zi Xin), and '移动' (Mobile).

**笔记本电脑属性定义表**

属性	值
品牌	hp, Lenovo, ASUS
店铺	惠普贸易 (上海...), 联想官方旗舰店, 华硕电脑专卖店
屏幕尺寸	13.3英寸, 14英寸, 16英寸
光源	内置光源, 外置光源, 无
厚度	15.8mm, 16.45mm, 16.9mm, 18.9mm, 20.9mm

**北京拜克洛克科技有限公司**

所属: 企业

**简介** 拜克洛克科技是一个无桩共享单车平台, 用户通过微信服务号或移动APP输入车牌号即可获得密码解锁单车, 随取随用, 此外, 还可以共享自己的单车给平台经营, 从而获得所有[查看全部>>](#)

**同义词** ofo共享单车, ofo, 小黄车

**所属地区** 北京市海淀区北四环西路58号11层1102室

**成立日期** 2015-08-06

**注册资本**

**法人** 戴威

**股东** 王刚 于信 戴威 张巳丁 刘坤 薛鑫

**高管** 于信 朱啸虎 戴威 陈鼎 陈靖 杨品杰 张巳丁 朱昱士 薛鑫 尚敬

# KGMS 属性定义

- 问答的高级功能支持

Q1: 2018年**融资额**大于1000万元的融资事件



Q2: Ofo的**高管**有哪些



Q3: idg融资事件统计按**轮次**



# KGMS 边属性定义

- 边属性定义的入口 – 属性定义列表
- 在已经定义的对象属性上，再定义边属性

The screenshot displays the KGMS 'Attribute Definition' (属性定义) interface for the 'Company' (企业) concept. The interface includes a sidebar with navigation options like '我的数据' (My Data), '我的图谱' (My Graph), and '属性定义' (Attribute Definition). The main area shows a table of attributes with columns for ID, Attribute Name, Concept, Type, Domain, and Actions. A red box highlights the 'Add Edge Attribute' (添加边属性) button for each attribute row.

ID	属性名	所属概念	属性类型	值域	操作
4	所属地区	企业	数值		编辑 添加边属性 删除
5	法人	企业	对象	人物、企业	编辑 添加边属性 删除
6	股东	企业	对象	人物、企业	编辑 添加边属性 删除
7	高管	企业	对象	人物	编辑 添加边属性 删除
8	成立日期	企业	数值		编辑 添加边属性 删除
9	注册资本	企业	数值		编辑 添加边属性 删除
10	标签	企业	对象	企业标签	编辑 添加边属性 删除

# KGMS 边属性定义

## ● 边数值属性定义

定义类型同数值属性，举例：投资关系的发生时间、发生金额

边数值属性定义

名称 \*

时间发生时间

数据类型

日期时间

取消

确定

## ● 边对象属性定义

定义类型同对象属性，举例：投资关系的发生事件对象

边对象属性定义

名称 \*

关联事件

值域 \*

添加

取消

确定

# KGMS 关系编辑

## ● 关系编辑的边属性编辑

根据边数值属性定义编辑

根据边对象属性定义编辑

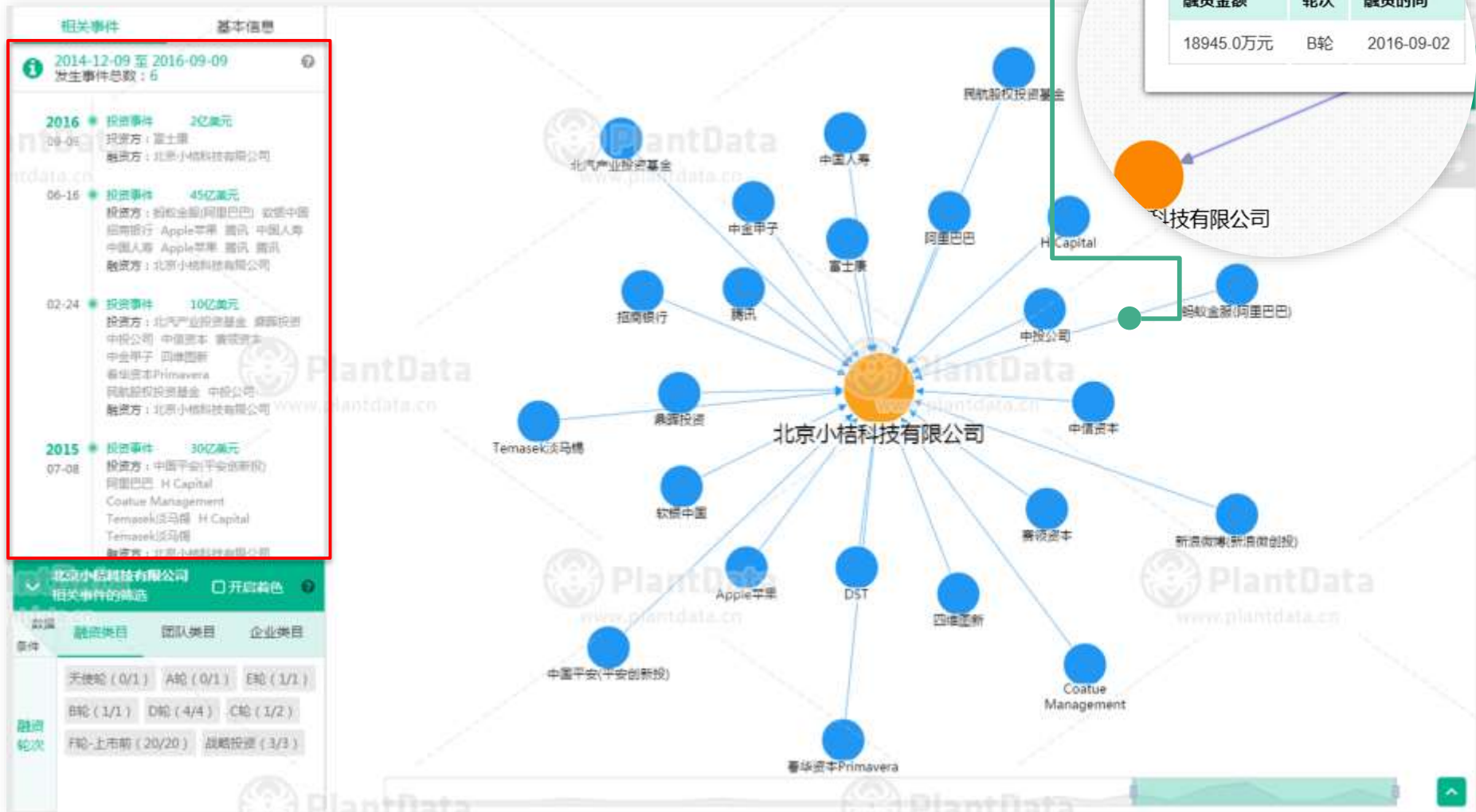
The screenshot displays the KGMS web application interface. On the left is a dark sidebar with a menu containing: 我的数据, 我的图谱, 概念编辑, 属性定义, 实例编辑, 实体融合, 关系编辑 (highlighted), 同义编辑, 领域词典, and 图谱可视化. The main content area has a top navigation bar with a search input and a breadcrumb path: 返回 > 关系编辑 > 投资机构 > 北京金沙江创业投资管理有限公司. Below the breadcrumb is a table of investment companies, with '青岛联创优内信息技术有限公司' selected. A modal dialog titled '北京拜克洛克科技有限公司边对象属性编辑' is open in the foreground. It contains four input fields: '起止时间 (时序)' with a range from '2016-09-02' to '2016-09-02', '轮次' set to 'B轮', '融资时间' set to '2016-09-02', and '融资金额' set to '18945.0万元'.

# KGMS 边属性定义

## ● 边属性定义的效果展示（定义边时间、定义边对象）

### 节点间边的事件关联描述

融资金额	轮次	融资时间
18945.0万元	B轮	2016-09-02



# KGMS 知识图谱编辑工具

---

- **实例编辑**

刻画现实社会中客观存在的唯一实体

- **关系编辑**

对象间的关系

- **场景**

企业、投资机构、产品、人

投资机构 – 投资 – 企业

企业 – 发布 – 产品

人 – 任职 – 企业



# KGMS 实例编辑

## ● 实例编辑

名称、唯一标识、简介、图片

从属概念

根据数值属性定义出的表单

支持文件导入

KGMS

18801610050

实例编辑

名称: 红杉资本

唯一标识: 用于区分同类实例

简介: 红杉资本于1972年在美国硅谷成立。在成立之后的30多年之中，红杉作为第一家机构投资者投资了如Apple, Google, Cisco, Oracle, Yahoo, LinkedIn等众多创新型的领导潮流的公司。在中国，红杉资本中国团队目前管理约20亿美元的母基金和近40亿人民币的国内基金。用于投资中国的成熟企业。红杉中国的合伙人及投资团队兼善国际经济宏观视野和本土创业企业经验，从2005年9月成立至今，在科技、消费服务、医疗、能源、金融和地产等领域投资了众多具有颠覆性创新的企业。红杉出

图片: 上传图片 线上地址

父概念: 投资机构

投资轮次:

投资类型: VC

总投资额度: 606

实例导入: (小于2M)

实例导入

1 选取您要导入实例的目标概念，然后下载模板

注意：此处仅支持实例及数值属性的导入。模板会根据不同的概念更换表头。

2 请将编辑好的实例模板点击上传

注意：文件大小不要超过2M

开始导入

# KGMS 关系编辑

## ● 关系编辑

对应对象属性下定义的关系添加

支持文件导入



☰

我的数据

我的图谱

概念编辑

属性定义

实例编辑

实体融合

关系编辑

同义编辑

领域词典

图谱可视化

Q 输入关键字进行过滤

创投知识图谱0309

产品

企业

投资机构

企业标签

融资事件

人物

返回 | 关系编辑 > 投资机构 > 北京金沙江创业投资管理有限公司 ?

保存

投资企业

青岛联创优内信息技术有限公司 编辑 删除

北京拜克浩克科技有限公司 编辑 删除

北京拜克浩克科技有限公司 编辑 删除

北京趣拿信息技术有限公司 编辑 删除

播思通讯技术（北京）有限公司 编辑 删除

北京一手科技有限公司 编辑 删除

北京伊电园网络科技有限公司 编辑 删除

上海易睦网络科技有限公司 编辑 删除

北京网游联动广告有限公司 编辑 删除

百姓网股份有限公司 编辑 删除

惠德时代能源科技（北京）有限公司 编辑 删除

恒原微电子（上海）有限公司 编辑 删除

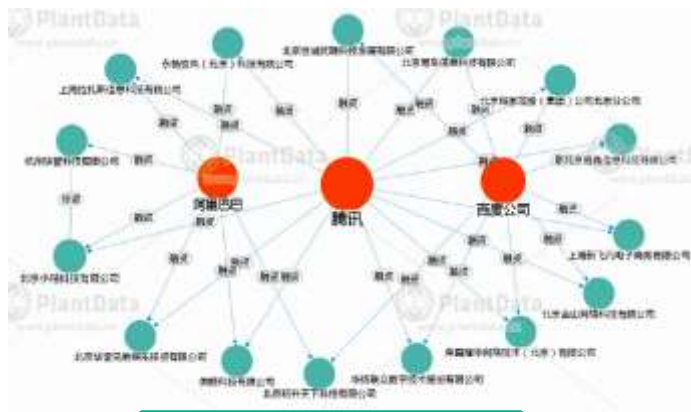
# KGMS 关系编辑



图探索



路径发现



关联分析

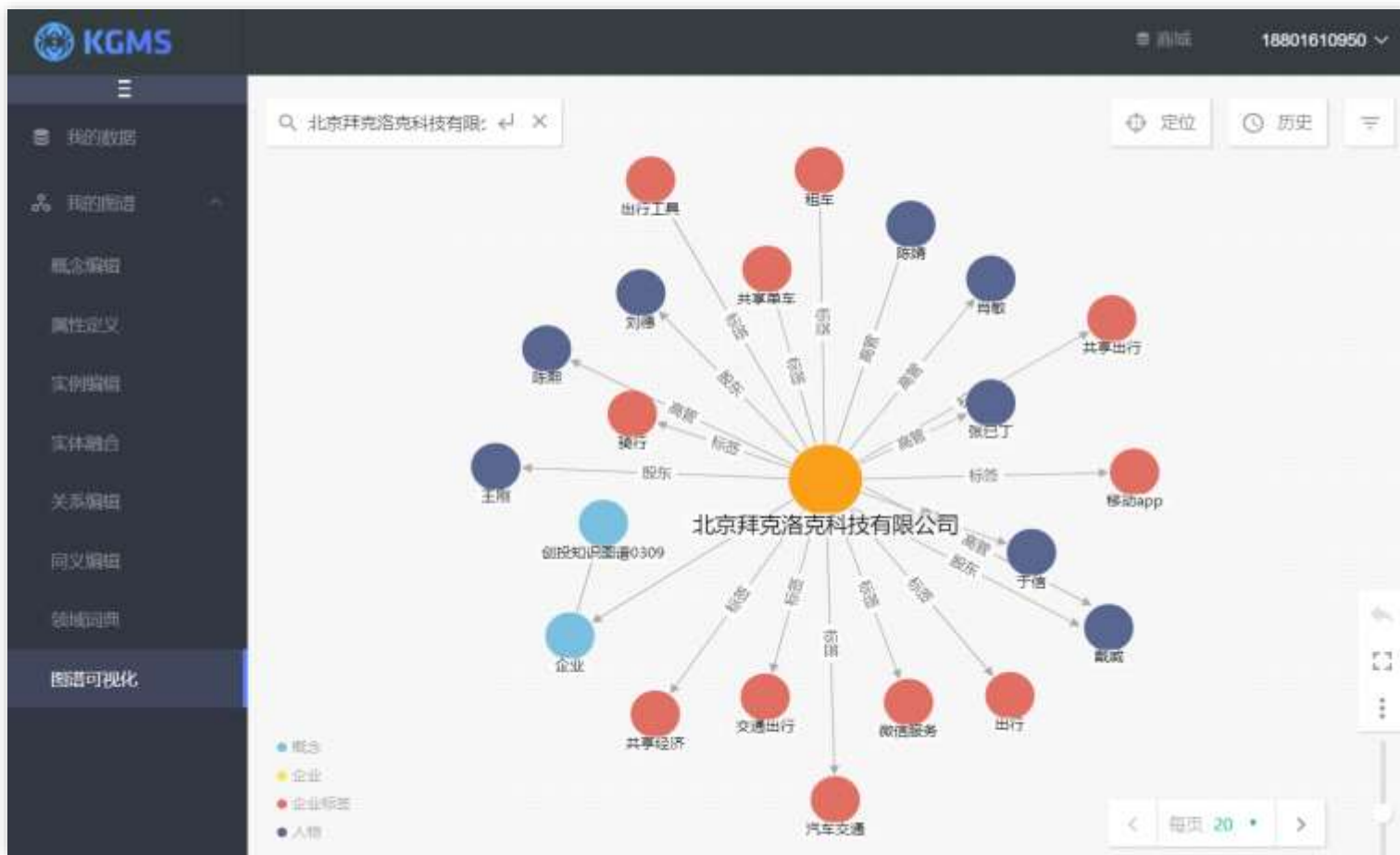


知识推荐

# KGMS 图谱可视化

## ● 图谱可视化

对于当前构建图谱的可视化展示



# KGMS 知识图谱编辑工具

---

- **同义编辑**

概念同义、属性同义、实例同义编辑

- **领域词典**

提示nlp能力的领域关键词、词性、词频

- **实体融合**

人工实体对齐、待合并实体候选集推荐、实体关联资源更新

- **场景**

搜索的同义扩展、语义识别

多资源的标引

同名不同机构作者合并 及 相关资源合并

# KGMS 同义编辑

## ● 同义编辑

概念同义编辑

实例同义编辑

属性同义编辑在属性定义的时候完成



☰

我的数据

我的图谱

概念编辑

属性定义

实例编辑

实体融合

Q 输入关键字进行过滤

创投知识图谱0309

产品

企业

投资机构

企业标签

融资事件

人物

同义编辑 > 企业 ?

概念同义

企业 创业公司 × 公司 × 添加

实例同义

北京拜克

北京拜克洛克... ofo共享单车 × ofo × 小黄车 × 添加

# KGMS 同义编辑

- 同义编辑效果



# KGMS 领域词典

## ● 领域词典

支持搜索、问答、标引

KGMS

冯莉

我的数据

我的图谱

概念编辑

属性定义

实例编辑

实体融合

关系编辑

同义编辑

领域词典

图谱可视化

输入关键字进行过滤

创投知识图谱0309

产品

企业

投资机构

企业标签

融资事件

人物

领域词典导入 (小于2M)

领域词典 > 列表

ID	领域词	所属概念	词性	频率	操作
143	股东	创投知识图谱0309	n	1	<a href="#">编辑</a> <a href="#">删除</a>
144	股东会	创投知识图谱0309	n	1	<a href="#">编辑</a> <a href="#">删除</a>
145	股东会议	创投知识图谱0309	n	1	<a href="#">编辑</a> <a href="#">删除</a>
146	股东大会	创投知识图谱0309	n	1	<a href="#">编辑</a> <a href="#">删除</a>
147	股东权益	创投知识图谱0309	n	1	<a href="#">编辑</a> <a href="#">删除</a>
148	股价	创投知识图谱0309	n	1	<a href="#">编辑</a> <a href="#">删除</a>
149	股价指数	创投知识图谱0309	n	1	<a href="#">编辑</a> <a href="#">删除</a>
150	股份	创投知识图谱0309	n	1	<a href="#">编辑</a> <a href="#">删除</a>
151	股份公司	创投知识图谱0309	n	1	<a href="#">编辑</a> <a href="#">删除</a>
152	股份制	创投知识图谱0309	n	1	<a href="#">编辑</a> <a href="#">删除</a>

<

>

前往 1 页



# KGMS 实体融合

● 实体融合

KGMS

☰

我的数据

我的图谱

概念编辑

属性定义

实例编辑

实体融合

关系编辑

同义编辑

领域词典

图谱可视化

17612950677

合并查询

合并推荐

执行合并

合并规则

合并属性值以主体为准

已选主体

自动推荐

武守远[中国电力科学研究院]

更换

待合并实体

添加

武守远[国网智能点网研究院]

移除

武守远[中电普瑞科技有限公司]

移除

武守远[电力部电力科学研究院]

移除

# KGMS 实体融合

## ● 实体融合

## 将同义实体合并

将人物合并，对标引关联的资源进行统一视角的计算



# KGMS 高级操作

## ● 图谱分析视图导出

用于统计、特征画像的导出，为后续计算服务。

## ● 生成数据导出

自动实体实体、联想提示、问题提示、输入纠错

KGMS



我的数据

我的图谱

高级操作

任务类型	开始时间	结束时间	任务状态
宽表导出	2018-05-08 15:04	2018-05-08 15:00	成功

# KGMS 我的数据

- 数据管理，对文档资源型数据进行管理

功能	功能点
数据集管理	数据集的增删改 录入数据的增删改 批量文件导入
数据入图	数据字段与图谱模式映射 数据内容校验 数据手工入图
数据标引	数据标引到概念 数据标引到实体 标引结果维护

# KGMS 数据导入

- 支持格式 xls、csv

将数据文件导入云端，为后续的处理做准备

1	id	detail_url	industry	persist_date	simple_name	full_name	money	round	e
2	1	http://w	医疗健康	2017/5/8	好孕帮	明祺思威	数千 万人民币	B轮	
3	2	http://w	旅游户外	2017/5/9	执惠旅游	北京执惠	2000 万人民币	A+轮	
4	3	http://w	硬件	#####	非常电	静信信息	数百 万人民币	种子轮	
5	4	http://w	硬件	#####	Hi电	上海重玩	1亿 人民币	A轮	
6	5	http://w	硬件	#####	小电	北京伊电	3.5亿 人民币	B轮	
7	6	http://w	房产	#####	安心公寓	上海悦讯	数亿 人民币	B+轮	
8	7	http://w	共享经济	#####	嗨了个球	嘉兴市单	1000 万人民币	Pre-A	



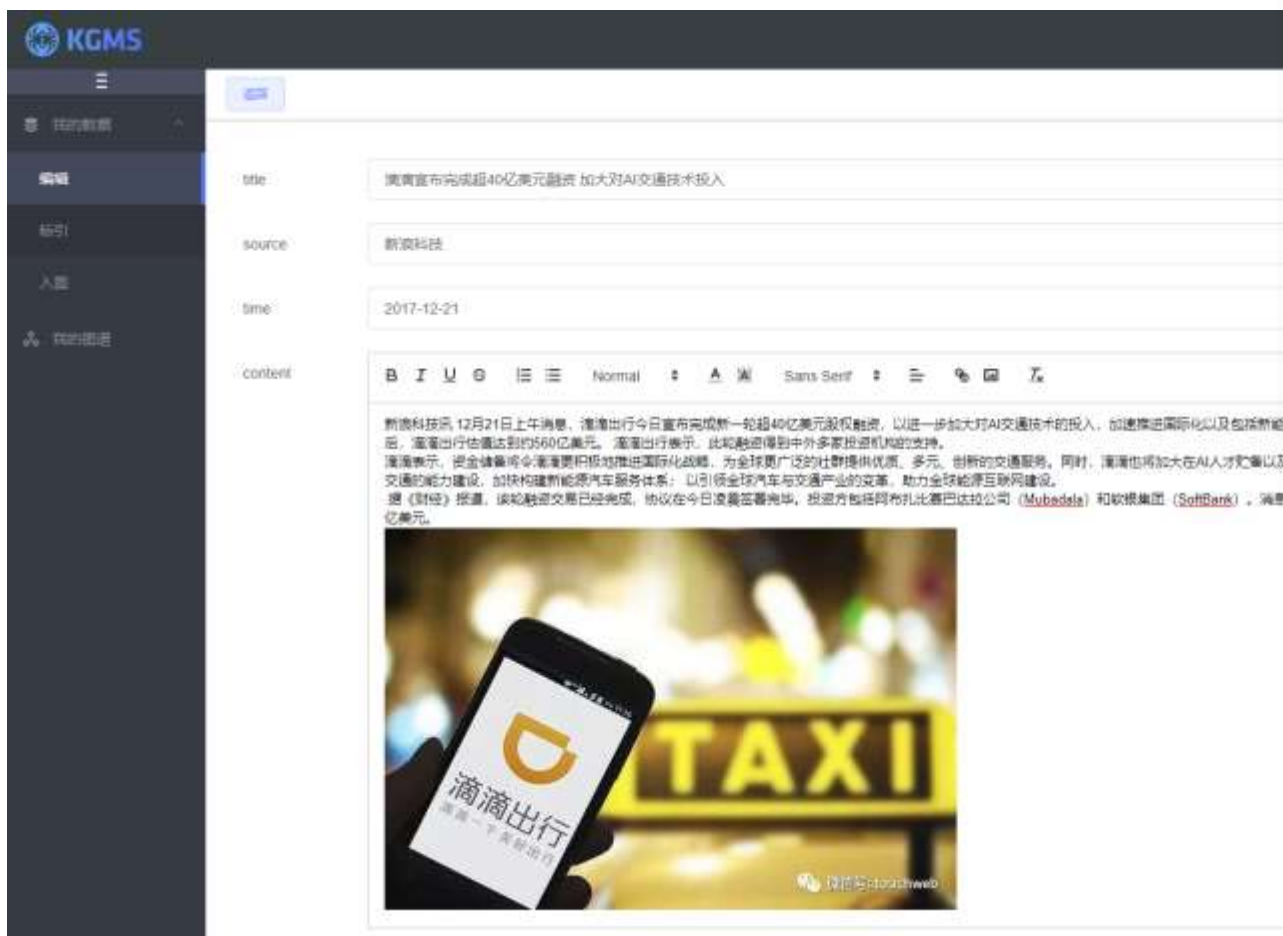
KGMS				
数据表: 批量导入测试				
全部 detail_url 输入关键词进行模糊查询				
	detail_url	persist_date	simple_name	money
1	https://www.ijuzi.com/company/1654...	2017/11/16	吾立方	数千万人民币
2	http://www.cyzone.cn/e/20171117/39...	2017/11/17	创想科技智能	100万人民币
3	http://www.cyzone.cn/e/20171117/39...	2017/11/17	泽生科技	5.07亿人民币
4	http://www.cyzone.cn/e/20171117/39...	2017/11/17	李锐和/青松	1.1亿人民币
5	https://www.ijuzi.com/company/1963...	2017/11/20	UNIBROWN由心咖啡	数百万人民币
6	https://www.ijuzi.com/company/50670	2017/11/20	火炭龟子	数千万人民币
7	https://www.ijuzi.com/company/68185	2017/11/20	大耳发/康威	224亿港元
8	http://www.cyzone.cn/e/20171120/39...	2017/11/20	聚数宏物资管	2.46亿人民币
9	http://www.cyzone.cn/e/20171120/39...	2017/11/20	爱乐奇-说宝量	3750万美元

# KGMS 数据编辑

## ● 编辑

对数据集进行编辑，支持多种形式的

如：短文本、长文本、富文本、时间、开关、可枚举值等



短文本

短文本

长文本

富文本

时间选择

日期选择

日期时间选择

开关

颜色选择

数字

下拉选择

标签输入

单选框

多选框

# KGMS 数据入图

## ● 入图

采集的数据需要通过  
人工校验审核  
导入知识图谱

## ● 入图步骤

1. 选择一个数据集
2. 选择一个知识图谱
3. 设定字段映射
4. 设定校验规则
5. 自动生成表单
6. 审核入图



KGMS 数据入图界面截图。左侧为深色侧边栏，包含“我的数据”、“编辑”、“标引”、“入图”（当前选中）和“我的图谱”等选项。右侧为表单区域，标题为“关联图谱: 创投知识图谱数据集”。表单包含以下字段：

- 名称: ofo融资事件
- 唯一标识: 用于区分同名实体
- 父概念: 融资事件 (可点击“添加”按钮)
- 融资额: 10000万
- 融资时间: 2017-01-01
- 投资方: (可点击“添加”按钮)
- 融资方: (可点击“添加”按钮)
- 标签: (可点击“添加”按钮)
- 地区: 北京
- 轮次: F轮

# KGMS 数据标引

## ● 标引

选择一个现有的知识图谱，手工选择概念、实例进行关联。

可支持脚本自动标引的结果修正。



我的数据

数据

**标引**

入库

我的图谱

滴滴宣布完成超40亿美元融资 加大对AI交通技术投入

返回

保存

新浪科技讯 12月21日上午消息，滴滴出行今日宣布完成新一轮超40亿美元股权融资，以进一步加大对AI交通技术的投入。加速推进国际化以及包括新能源汽车服务在内的创新业务。据彭博报道，融资完成后，滴滴出行估值达到约560亿美元。滴滴出行表示，此轮融资得到中外多家投资机构的支持。滴滴表示，资金储备将令滴滴更积极地推进国际化战略，为全球更广泛的社群提供优质、多元、创新的交通服务。同时，滴滴也将加大对AI人才储备以及技术上的投入规模。公司将持续提升智能驾驶与智慧交通的能力建设，加快构建新能源汽车服务体系；以引领全球汽车与交通产业的变革，助力全球能源互联网建设。据《财经》报道，该轮融资交易已经完成，协议在今日凌晨签署完毕。投资方包括阿布扎比穆巴达拉公司（Mubadala）和软银集团（SoftBank）。消息人士透露，本轮融资完成后，滴滴现金储备接近120亿美元。



图谱名称 创投知识图谱数据集

标引到概念 产品

产品 ×

标引到实例 选择实例

滴滴打车 × 北京小桔科技有限公司 ×



# KGMS 数据标引

## ● 标引

资源型数据不存在知识图谱里，通过标引，与知识图谱的概念及实例建立关联。

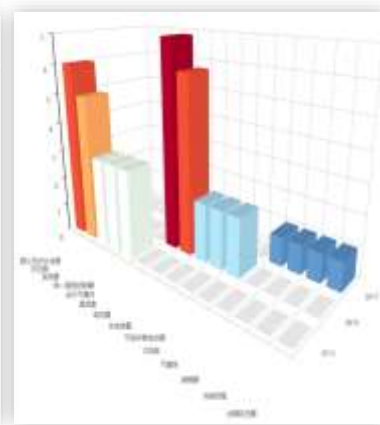
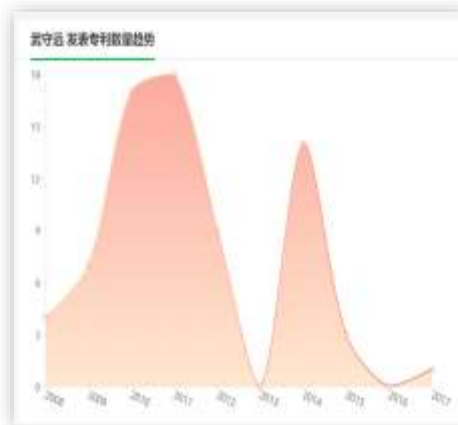
可用于支持语义搜索、各种图表统计。

论文、专利、文献、成果、标准

标引

知识  
图谱

统计



# KGMS 商店

- 商店，数据分享推广的平台



- 行业数据

知识图谱标准格式的行业数据。

- 行业知识图谱

知识图谱标准格式的行业数据。

- 行业知识模式

共享知识图谱的标准数据schema。

# KGMS JSSDK

Hiekn SDK v4.1.0

Docs

Gallery

vs code

Download

图谱探索

时序图谱

路径分析

关联分析

概念图谱

概念树

概念提示

实体提示

知识卡片

表格

资源

资源列表

资源统计

关联推荐

实体消歧

Schema

实体识别

语义标注

javascript

html

css

重置

运行

```
1 var settings = {
2   kgName: 'kg_ct',
3   selector: '#container',
4   ajaxSettings: {
5     baseUrl: 'https://kg.hiekn.com:8143/planetdata_console/sdk/',
6     dataFilter: function (data) {
7       console.log('common dataFilter');
8       return data;
9     }
10  },
11  schemaSettings: {
12    ajaxSettings: {
13      url: 'https://kg.hiekn.com:8143/planetdata_console/sdk/schema',
14      dataFilter: function (data) {
15        console.log('schema dataFilter');
16        return data;
17      }
18    }
19  },
20  nodeSettings: {
21    imagePrefix: 'https://opvdx0y.bkt.clouddn.com/',
22    images: buildImages([2, 7, 9, 10, 11], '../images/kg_ct'),
23    nodeColors: {
24      2: '#f45b5b',
25      7: '#3e3eb5',
26      9: '#8085e9',
27      10: '#1c7cd5',
28      11: '#1c7cd5',
29      12: '#06cae3',
30      13: '#2b908f',
31      16: '#ee891d',
32      17: '#ee891d',
33      18: '#ee891d',
34      1601: '#ee891d',
35      1602: '#ee891d'
36    },
37    // textColors: {
38    //   2: '#fff',
39    //   7: '#fff',
40    //   9: '#fff',
41    //   10: '#fff',
42    //   11: '#fff',
43    //   12: '#fff',
44    //   13: '#fff',
45    //   16: '#fff',
46    //   17: '#fff',
47    //   18: '#fff',
48    //   1601: '#fff',
49    //   1602: '#fff',
50    //   21: {
```

请输入关键词...

< 每页 20 >

历史

定位

投资机构

品牌

其他

人物

产品

一级行业

图例

初创企业

二级行业

# KGMS SDK

## ● Restful SDK

接口分类	具体接口	接口说明
通用接口	图谱schema接口	获取知识图谱底层schema定义
	搜索提示接口	对于用户输入，在知识图谱中进行实体或概念的提示
图数据及可视化	图数据探索接口	以一个中心实体为视角，以图谱可视化的方式查看与其存在关联关系的人物，机构，资源，技术点等
	时序图探索接口	以一个中心实体为视角，以图谱可视化的方式查看与其存在关联关系的人物，机构，资源，技术点等在每一个时间点的情况
	知识卡片接口	以infobox的形式展示一个实体或概念的相关信息
	语义关联接口	在知识图谱中获取与实体或概念相关联的实体或概念
网络分析	路径分析接口	在一个基于人物，机构，资源，技术点等实体形成的一个网络关系图中，查询任意两个主体的最短路径。
	关联分析接口	在一个于基于人物，机构，资源，技术点等实体形成的一个网络关系图中，查询任意多个主体的关联关系。

# KGMS SDK 图探索

## ● 图探索的配置项

以哪个图的数据作为计算依据

从中心节点出发步数的范围

从中心节点出发包含的关系类型

以哪个节点作为中心节点

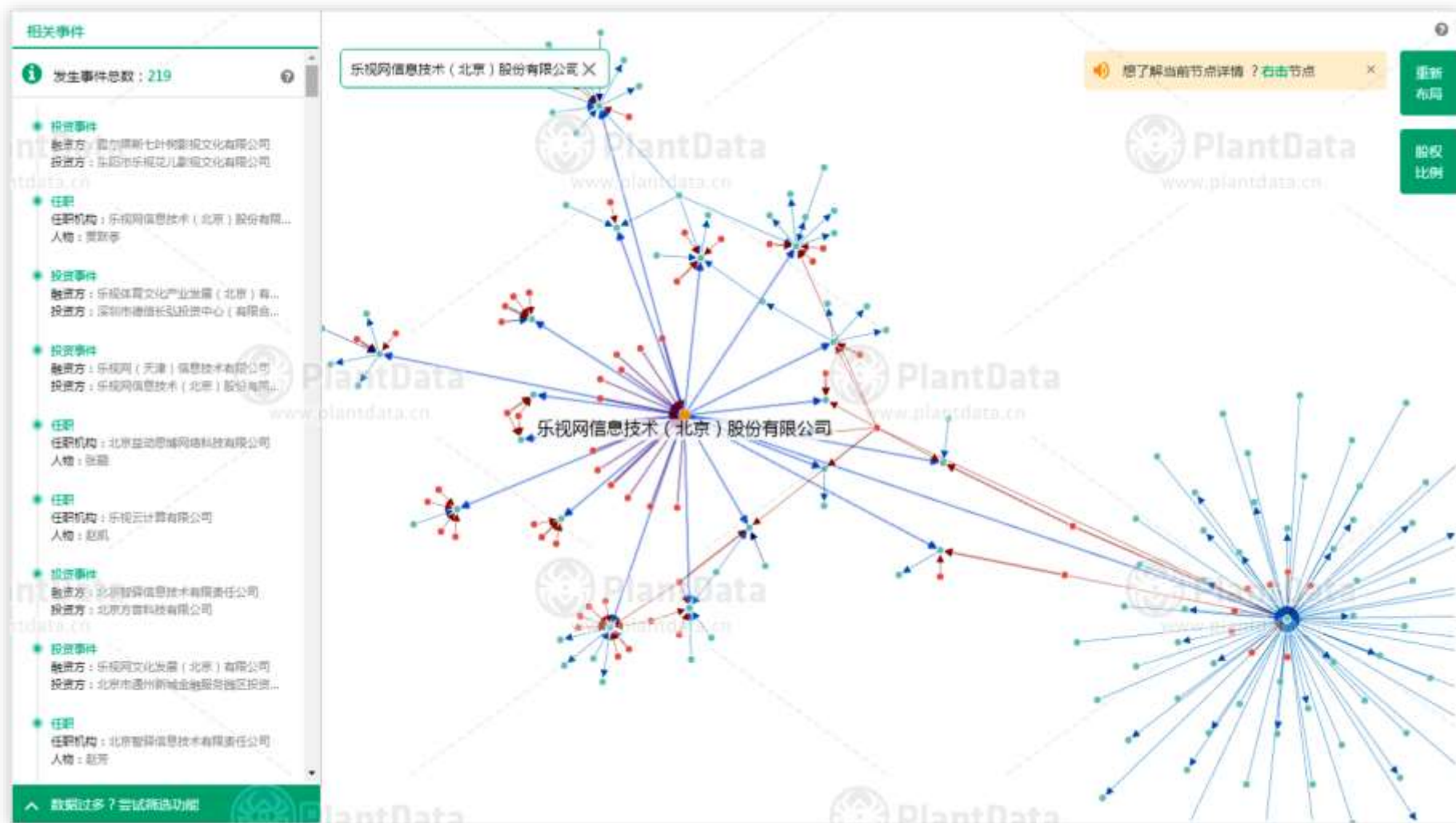
从中心节点出发包含的节点类型

是否合并两节点间的同类型关系





# KGMS SDK 图探索



# KGMS SDK 时序图探索

## ● 时序图探索的配置项

以哪个图的数据作为计算依据

从中心节点出发步数的范围

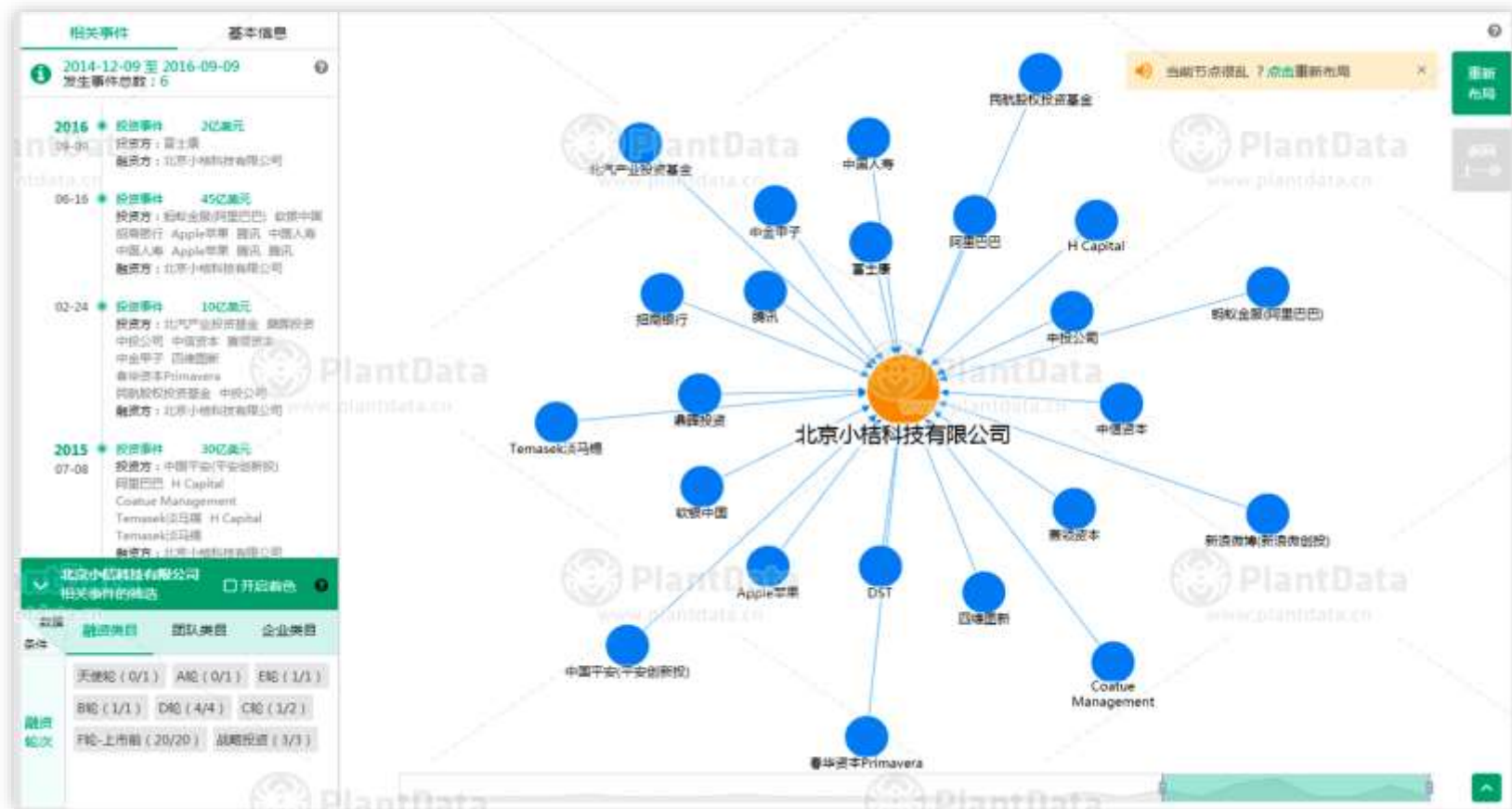
从中心节点出发包含的关系类型

该节点相关事件发生的时间区间

以哪个节点作为中心节点

从中心节点出发包含的节点类型

是否合并两节点间的同类型关系



# KGMS SDK 知识卡片

## ● 知识卡片的配置项

以哪个图的数据作为计算依据

指定卡片的核心节点A

是否读取节点A的数值属性

是否读取节点A的对象属性？读取哪些？

控制读取数据的量

PiantData 北京小桔科技有限公司	
所属： <a href="#">初创企业</a>	
简介	滴滴出行(原名滴滴打车)是一个手机智能叫车系统，帮助乘客随时随地、方便快捷的叫到出租车，帮助司机低成本更方便的接受预订。2015年与快的打车合并为滴滴快的集团。
同义词	-
企业规模	2000人以上
注册号	110108015068911
经营状况	开业
成立日期	2012-07-10
注册资本	1000 万人民币
经营范围	技术开发、技术咨询、技术服务、技术推广；基础软件服务；应用软件开发；设计、制作、代理、发布广告；软件开发；销售自 <a href="#">查看全部&gt;&gt;</a>
公司地址	北京市海淀区上地东路9号1幢5层北区2号
企业类型	有限责任公司(自然人投资或控股)
企业背景标签	<a href="#">阿里巴巴系</a>
企业行业标签	<a href="#">交通出行</a> <a href="#">汽车交通</a>
行业小类	<a href="#">交通出行</a>
行业大类	<a href="#">汽车交通</a>
企业其他标签	<a href="#">拼车</a> <a href="#">共享经济</a> <a href="#">打车</a> <a href="#">租车</a> <a href="#">代驾</a> <a href="#">O2O</a>



# KGMS SDK 资源统计

## ● 资源统计配置项

## 读取数据集的数据

## 对数据集的数据进行各维度统计

## 读取模型挖掘的计算结果

## 读取机器学习的计算结果



# KGMS SDK 知识推荐

## ● 知识推荐的配置项

读取图中节点A一步的任意类型关系

读取图中概念C下的实体

读取图中概念C下的子概念



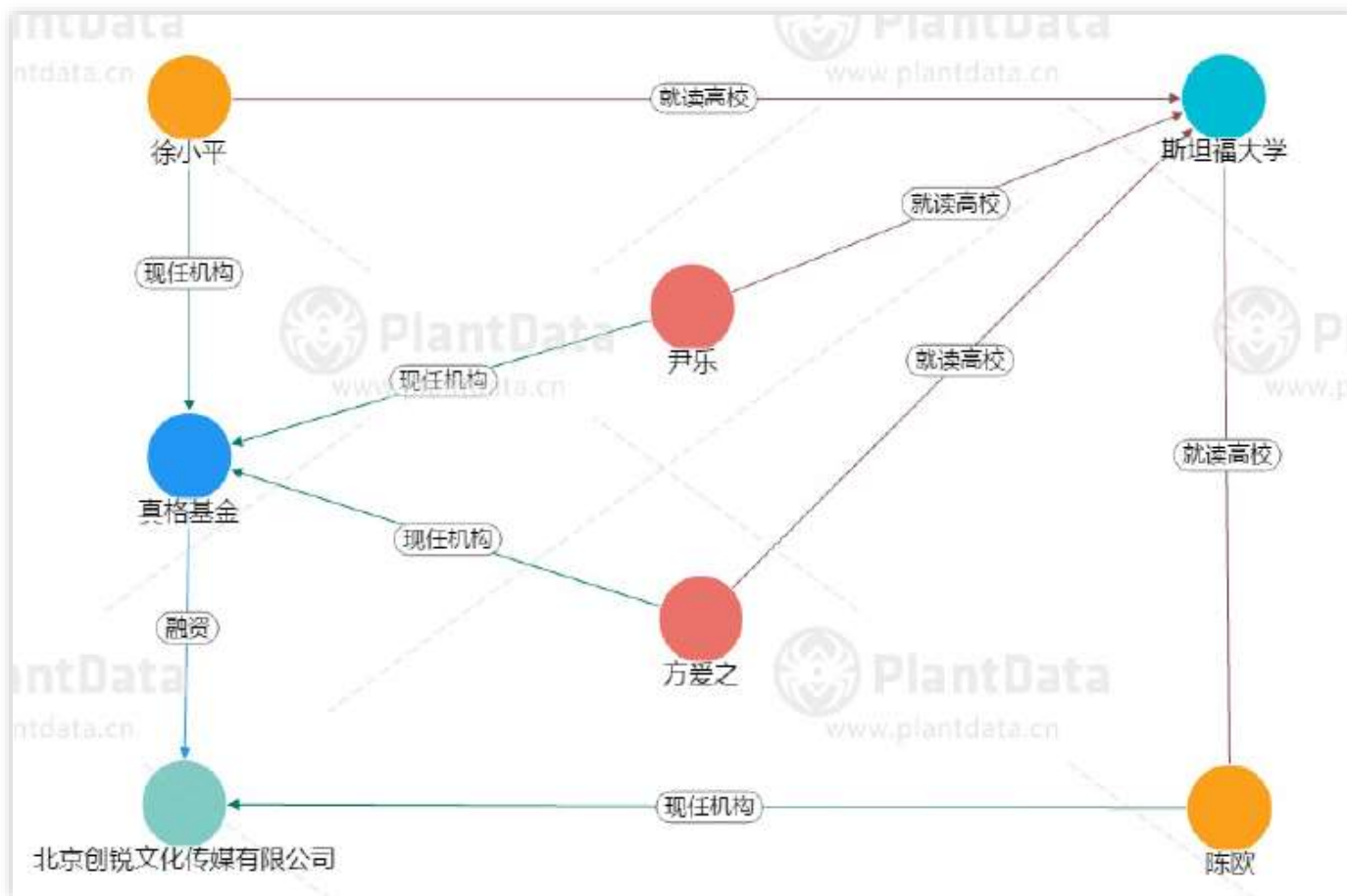
# KGMS SDK 路径分析

## ● 路径分析的配置项

- 以哪个图的数据作为计算依据
- 起始和终止节点是什么
- 路径步数的范围
- 路径中节点关系的方向
- 在计算路径时是否只计算最短路径
- 路径中包含的节点类型
- 路径中包含的边关系类型
- 对起始于终止节点进行关系解读
- 对指定的节点和边类型统计



# KGMS SDK 路径分析



# KGMS SDK 关系发现

## ● 关系发现配置项

以哪个图的数据作为计算依据

路径步数的范围

在计算路径时是否只计算最短路径

路径中包含的边关系类型

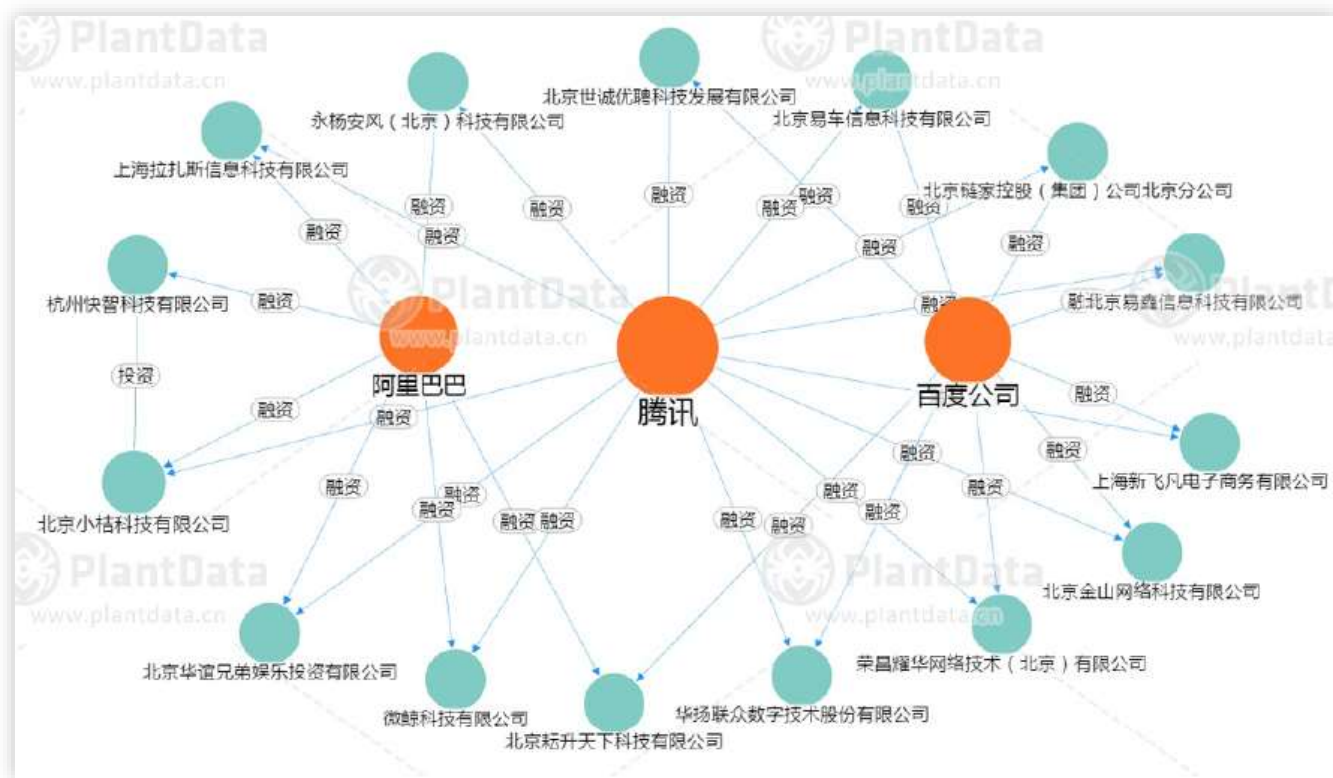
对指定的节点和边类型统计

起始和终止节点是什么

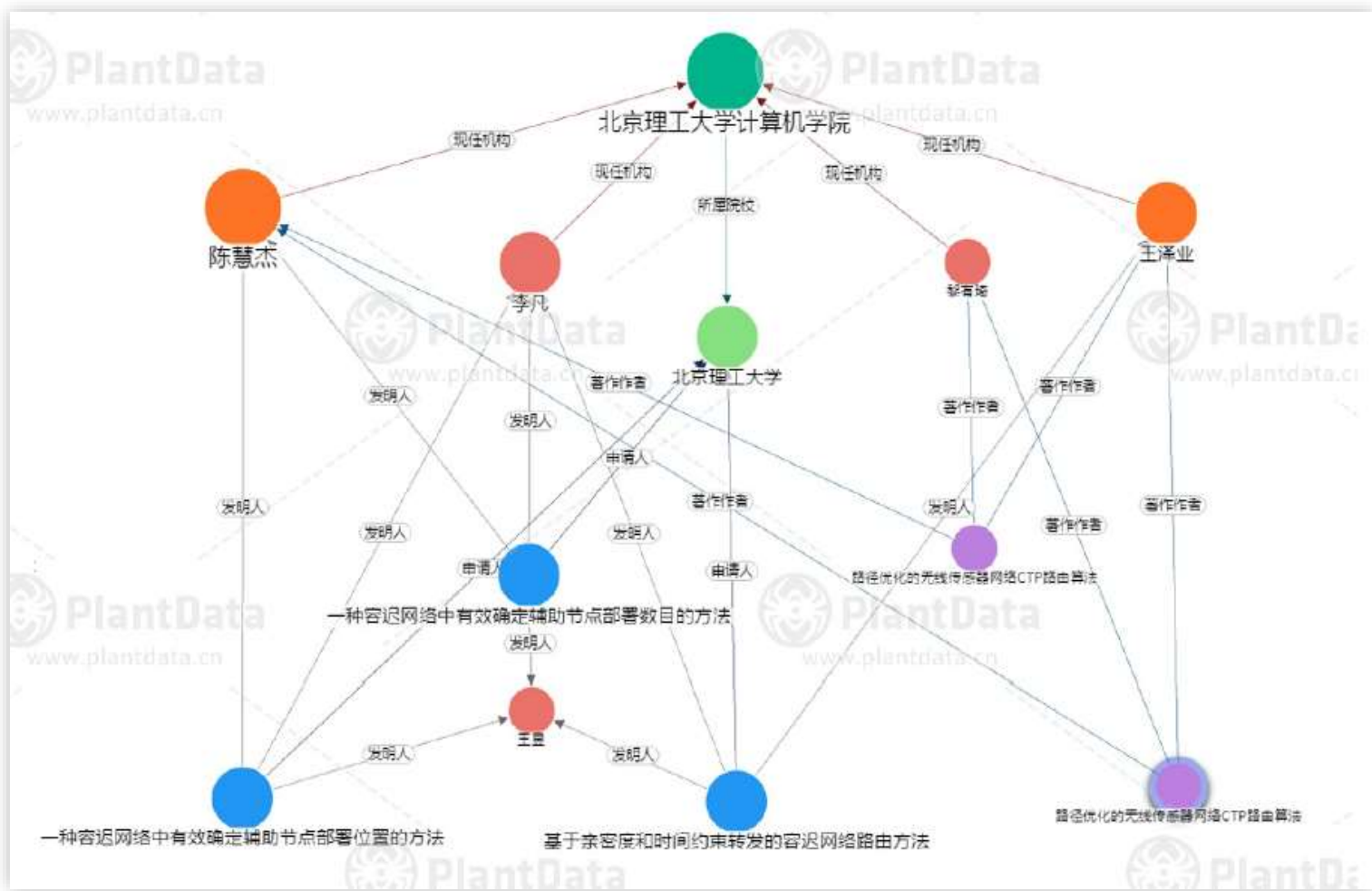
路径中节点关系的方向

路径中包含的节点类型

对起始于终止节点进行关系解读



# KGMS SDK 关系发现



# KGMS 服务方式

分类	功能	KGaaS	私有化基础版本	私有化高级版
资源控制相关	图谱个数	5	不限	不限
	数据集个数	5	不限	不限
	单图谱限制	1w	不限	不限
	单数据集限制	1w	不限	不限
数据管理	数据集管理	支持	支持	支持
	文件导入	支持	支持	支持
	数据修正	支持	支持	支持
	数据标引	支持	支持	支持
	数据入图	支持	支持	支持
知识图谱管理	知识图谱管理	支持	支持	支持
	文件导入导出	支持	支持	支持
	概念编辑	支持	支持	支持
	属性定义	支持	支持	支持
	实例编辑	支持	支持	支持
	关系编辑	支持	支持	支持
	同义编辑	支持	支持	支持
	领域词典	支持	支持	支持
	图谱可视化	支持	支持	支持
	实体融合	支持	支持	支持
	实体标引资源融合	不支持	支持	支持
	知识卡片	支持	支持	支持
	图谱探索	支持	支持	支持
SDK相关技能	时序探索	支持	支持	支持
	知识推荐	支持	支持	支持
	路径分析	支持	支持	支持
	关联分析	不支持	支持	支持
	搜索服务	不支持	支持	支持
	资源统计	不支持	支持	支持

# KGMS 演示

---

- 公有创投图谱数据查看
- 从零构建一个教学创投图谱
- 数据编辑、数据入图、数据标引的使用



# KGMS

---

- QA 公众号 活动推广