知识图谱的关键技术与应用

李征峻、孟静怡、王露、哈思娜、严文欣

目录

CONTENTS

知识图谱概述

PART.01

知识图谱概述

知识图谱的概念

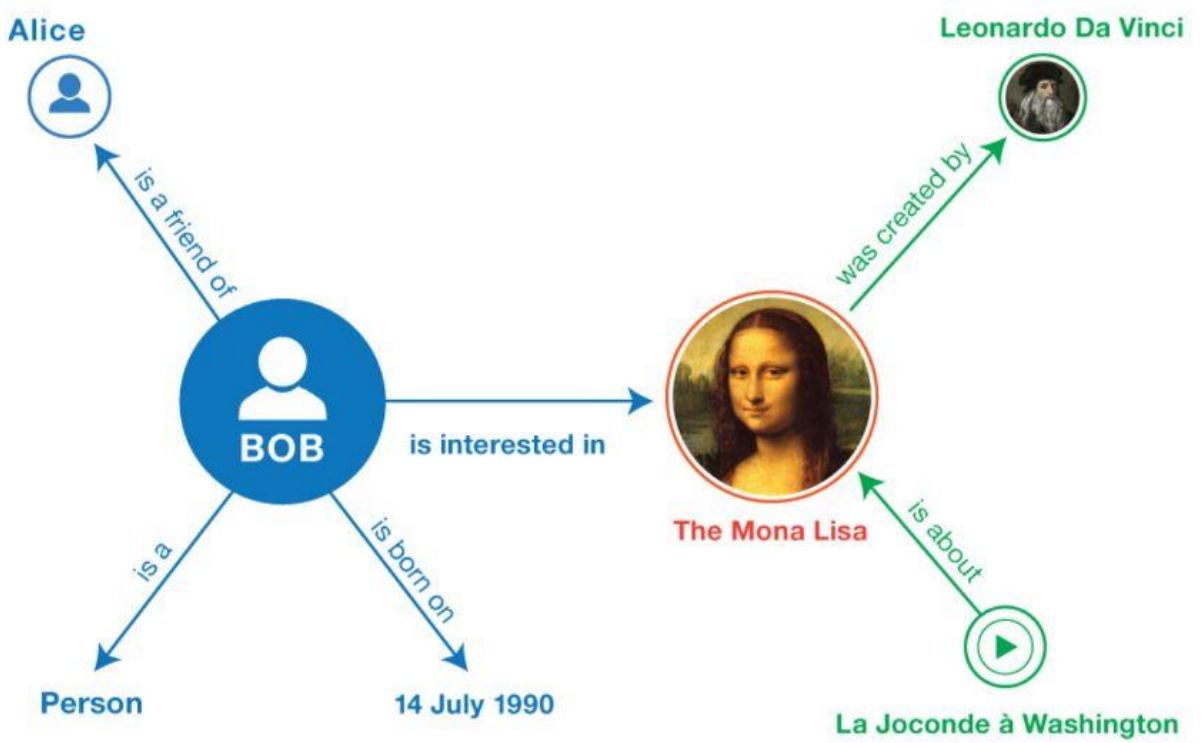
知识图谱的分类

知识图谱的历史

知识图谱的技术与应用

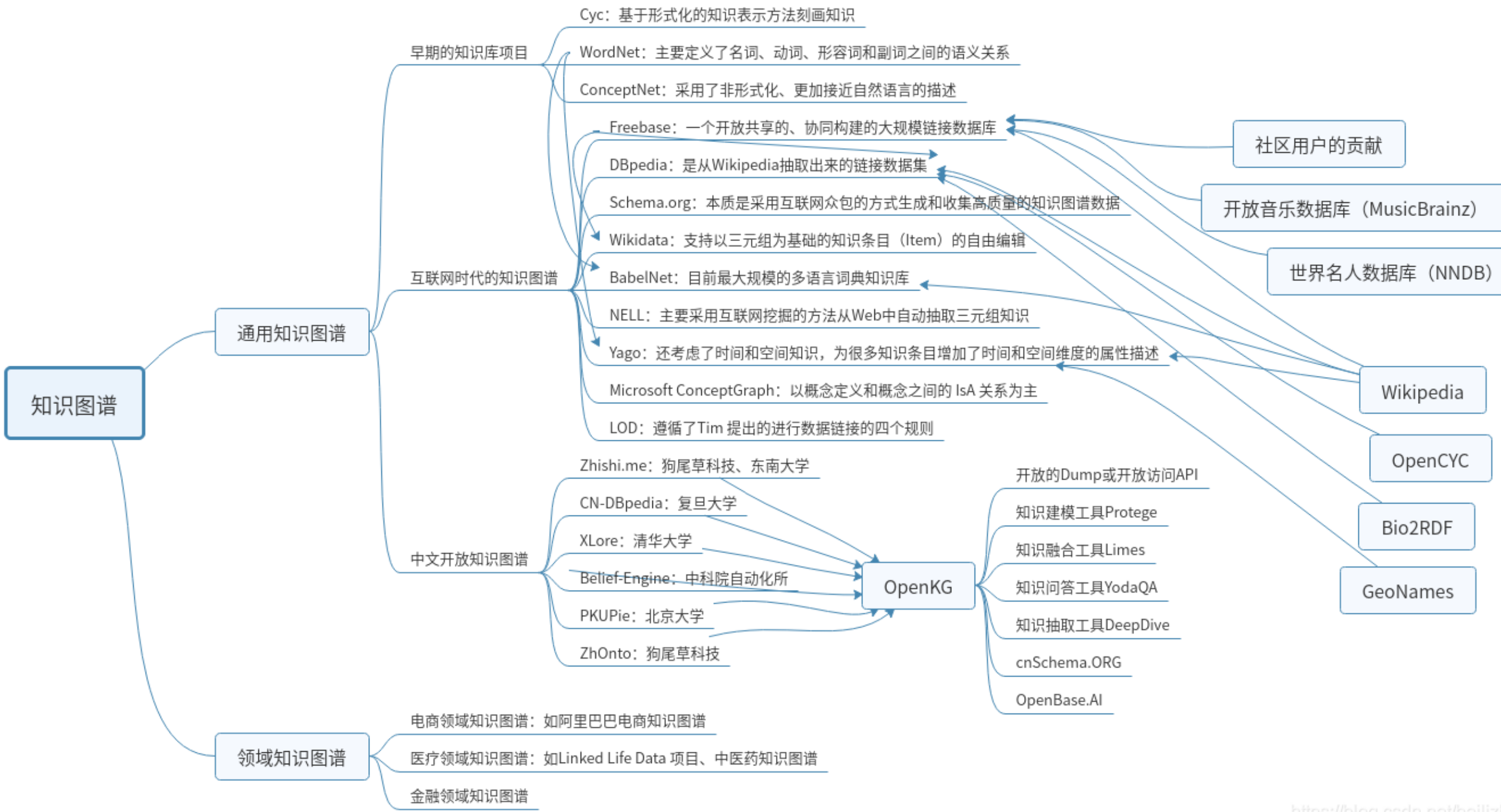
1. 知识图谱的概念

知识图谱旨在描述真实世界中存在的各种实体或概念。其中，每个实体或概念用一个全局唯一确定的 ID 来标识，称为它们的标识符。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。



1. 知识图谱的概念

- 在知识图谱里，我们通常用“实体（Entity）”来表达图里的节点、用“关系（Relation）”来表达图里的“边”。实体指的是现实世界中的事物比如人、地名、概念、药物、公司等，关系则用来表达不同实体之间的某种联系，比如人-“居住在”-北京、张三和李四是“朋友”、逻辑回归是深度学习的“先导知识”等等。



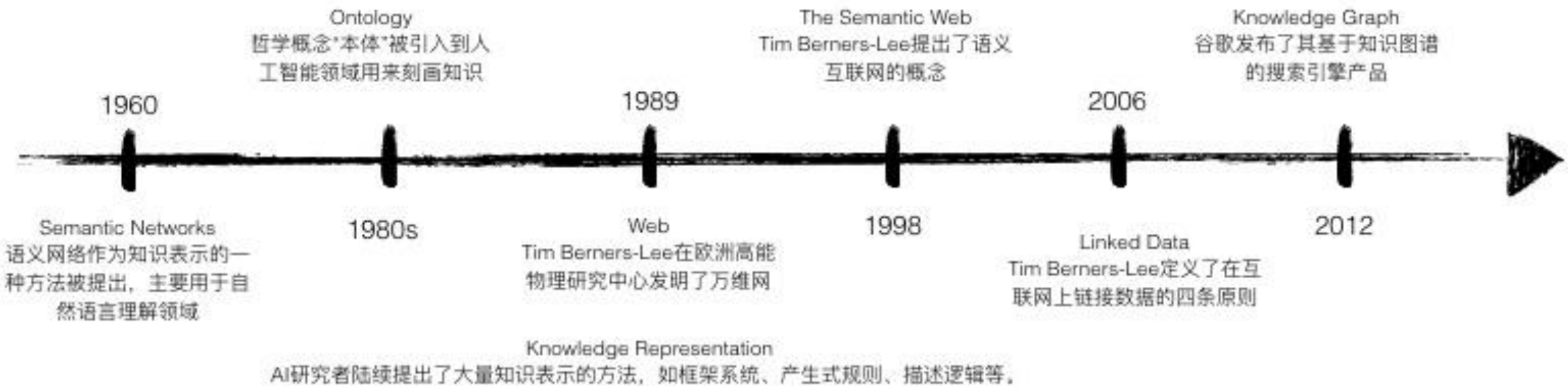
2. 知识图谱的分类：通用知识图谱与领域知识图谱的区别

分类 / 比较项目	通用知识图谱	领域知识图谱
知识来源及规模化	以互联网开放数据，如 Wikipedia 或社区众包为主要来源，逐步扩大规模	以领域或企业内部的数据为主要来源，通常要求快速扩大规模
对知识表示的要求	主要以三元组事实型知识为主	知识结构更加复杂，通常包含较为复杂的本体工程和规则型知识
对知识质量的要求	较多地采用面向开放域的 Web 抽取，对知识抽取质量有一定容忍度	知识抽取的质量要求更高，较多地依靠从企业内部的结构化、非结构化数据进行联合抽取，并依靠人工进行审核校验，保障质量
对知识融合的要求	融合主要起到提升质量的作用	融合多源的领域数据是扩大构建规模的有效手段
知识的应用形式	主要以搜索和问答为主要应用形式，对推理要求较低	应用形式更加全面，除搜索问答外，通常还包括决策分析、业务管理等，并对推理的要求更高，并有较强的可解释性要求
举例	DBpedia、Yago、百度、谷歌等	电商、医疗、金融、农业、安全等

3. 知识图谱的历史

- 知识图谱的概念是Google于2012年正式提出，但是知识图谱的发展却可以追溯到1960年的语义网络，中间经历了一系列的演变，才形成了今天的知识图谱.

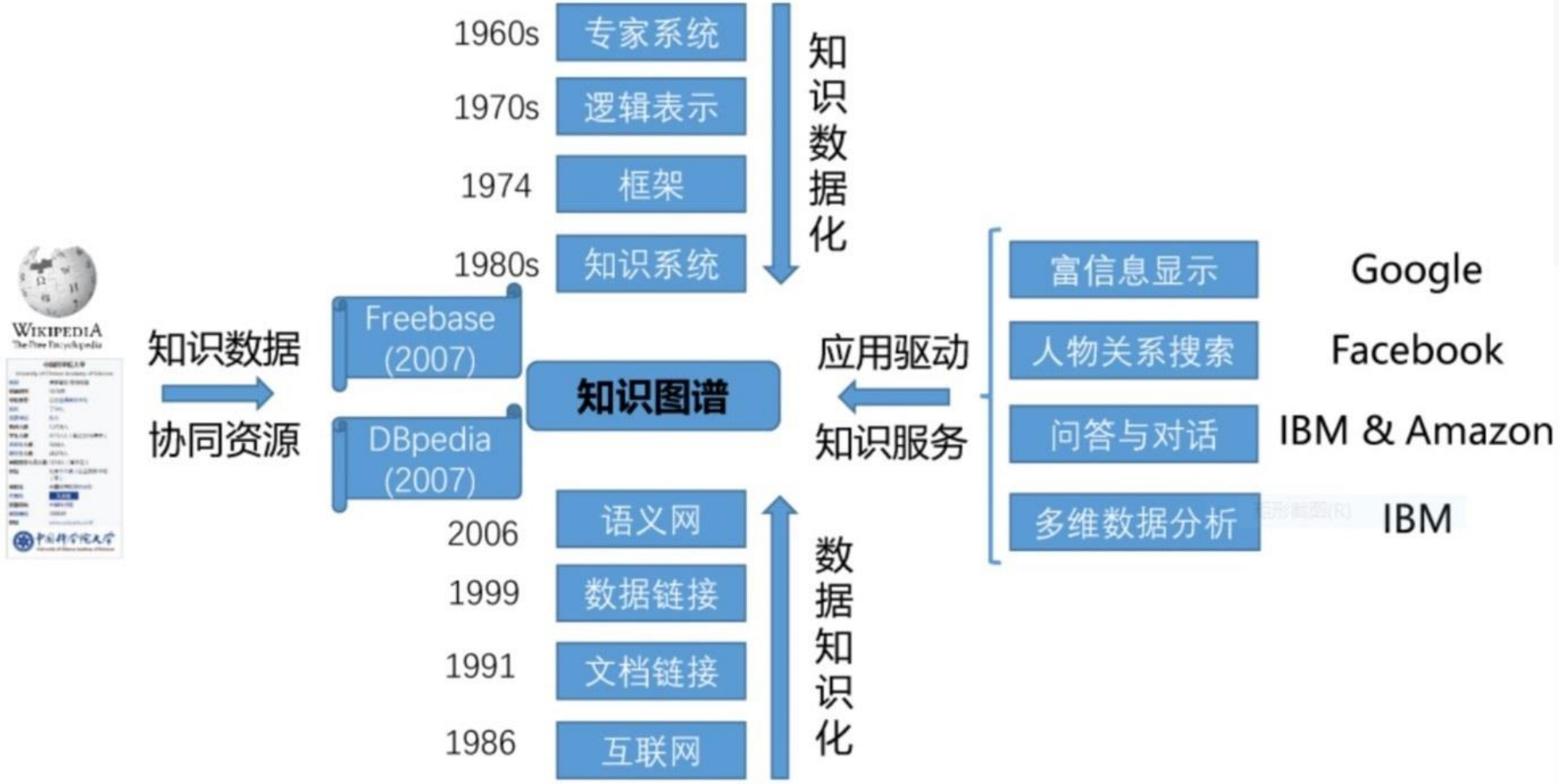
3. 知识图谱的历史


1960
Semantic Networks
语义网络作为知识表示的一种方法被提出，主要用于自然语言理解领域
Ontology
哲学概念"本体"被引入到人工智能领域用来刻画知识
1980s
1989
Web
Tim Berners-Lee在欧洲高能物理研究中心发明了万维网
The Semantic Web
Tim Berners-Lee提出了语义互联网的概念
1998
2006
Linked Data
Tim Berners-Lee定义了在互联网上链接数据的四条原则
Knowledge Graph
谷歌发布了其基于知识图谱的搜索引擎产品
2012
Knowledge Representation
AI研究者陆续提出了大量知识表示的方法，如框架系统、产生式规则、描述逻辑等。

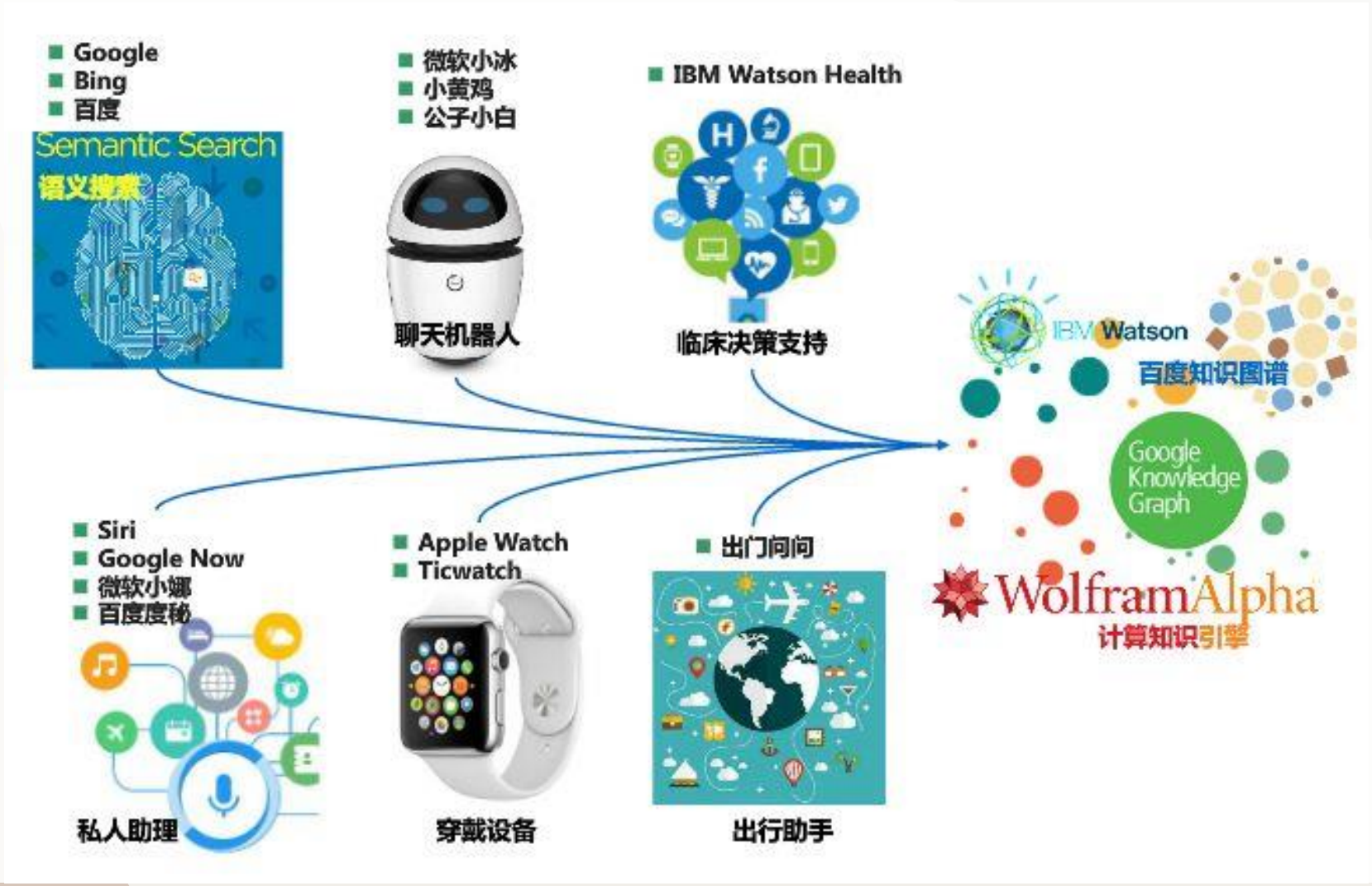
3. 知识图谱的历史

互联网公司知识图谱布局

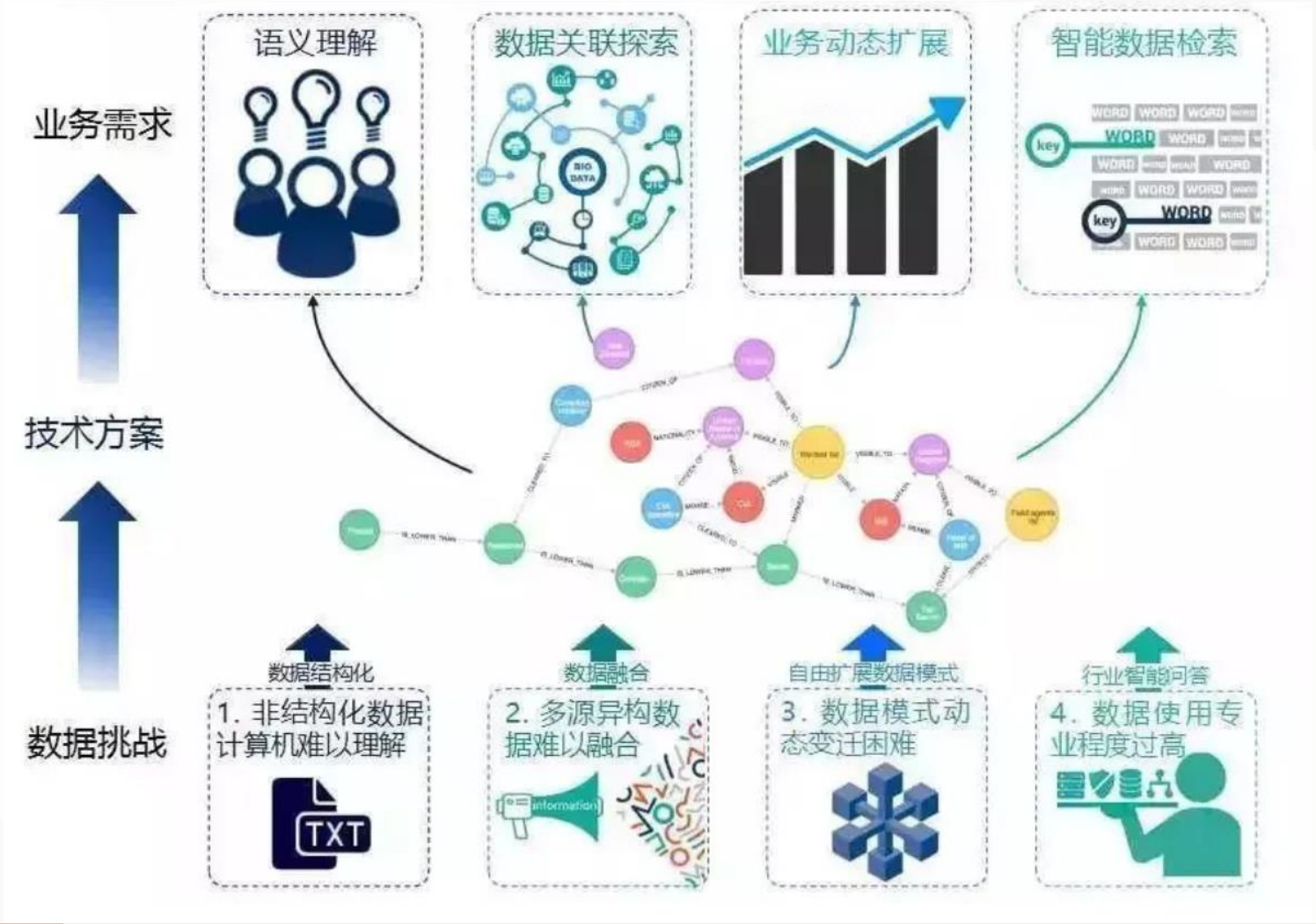
3. 知识图谱的历史


1960s
专家系统
1970s
逻辑表示
1974
框架
1980s
知识系统
知识数据化
WIKIPEDIA
知识数据
协同资源
Freebase (2007)
DBpedia (2007)
知识图谱
应用驱动
知识服务
富信息显示
人物关系搜索
问答与对话
多维数据分析
Google
Facebook
IBM & Amazon
IBM
2006
语义网
1999
数据链接
1991
文档链接
1986
互联网
数据知识化

4. 知识图谱的技术与应用



4. 知识图谱的技术与应用：知识图谱助力数据分析实现商业智能



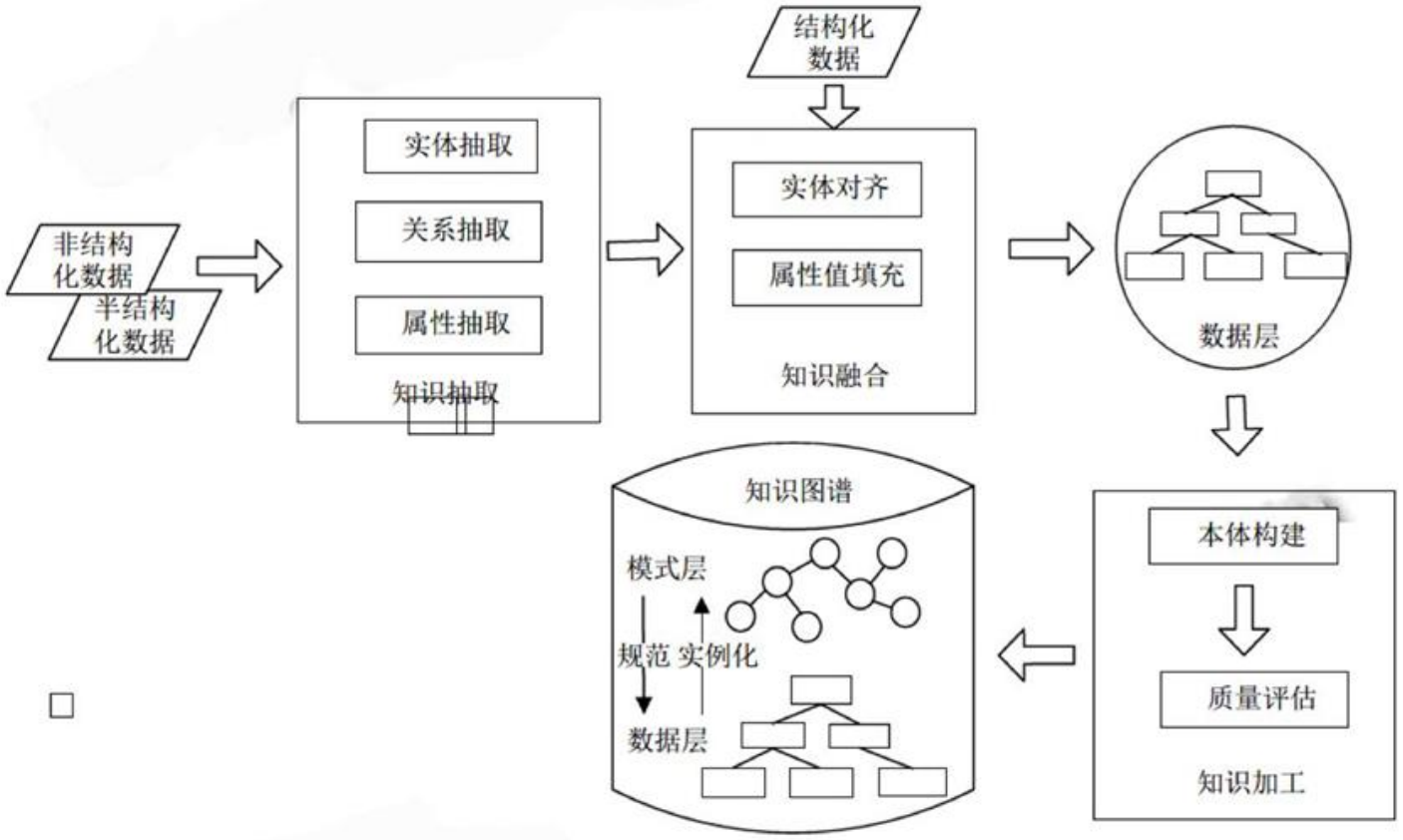
4. 知识图谱的技术与应用：开源知识图谱

	DBpedia	Freebase	OpenCyc	Wikidata	YAGO
Number of triples	411 885 960	3 124 791 156	2 412 520	748 530 833	1 001 461 792
Number of classes	736	53 092	116 822	302 280	569 751
Number of relations	2819	70 902	18 028	1874	106
No. of unique predicates	60 231	784 977	165	4839	88 736
Number of entities	4 298 433	49 947 799	41 029	18 697 897	5 130 031
Number of instances	20 764 283	115 880 761	242 383	142 213 806	12 291 250
Avg. number of entities per class	5840.3	940.8	0.35	61.9	9.0
No. of unique subjects	31 391 413	125 144 313	261 097	142 278 154	331 806 927
No. of unique non-literals in object position	83 284 634	189 466 866	423 432	101 745 685	17 438 196
No. of unique literals in object position	161 398 382	1 782 723 759	1 081 818	308 144 682	682 313 508

知识抽取与知识融合

PART.02

构建知识图谱——自底向上



自底向上的知识图谱构建流程

图片源于：百度百科

知识抽取的概念

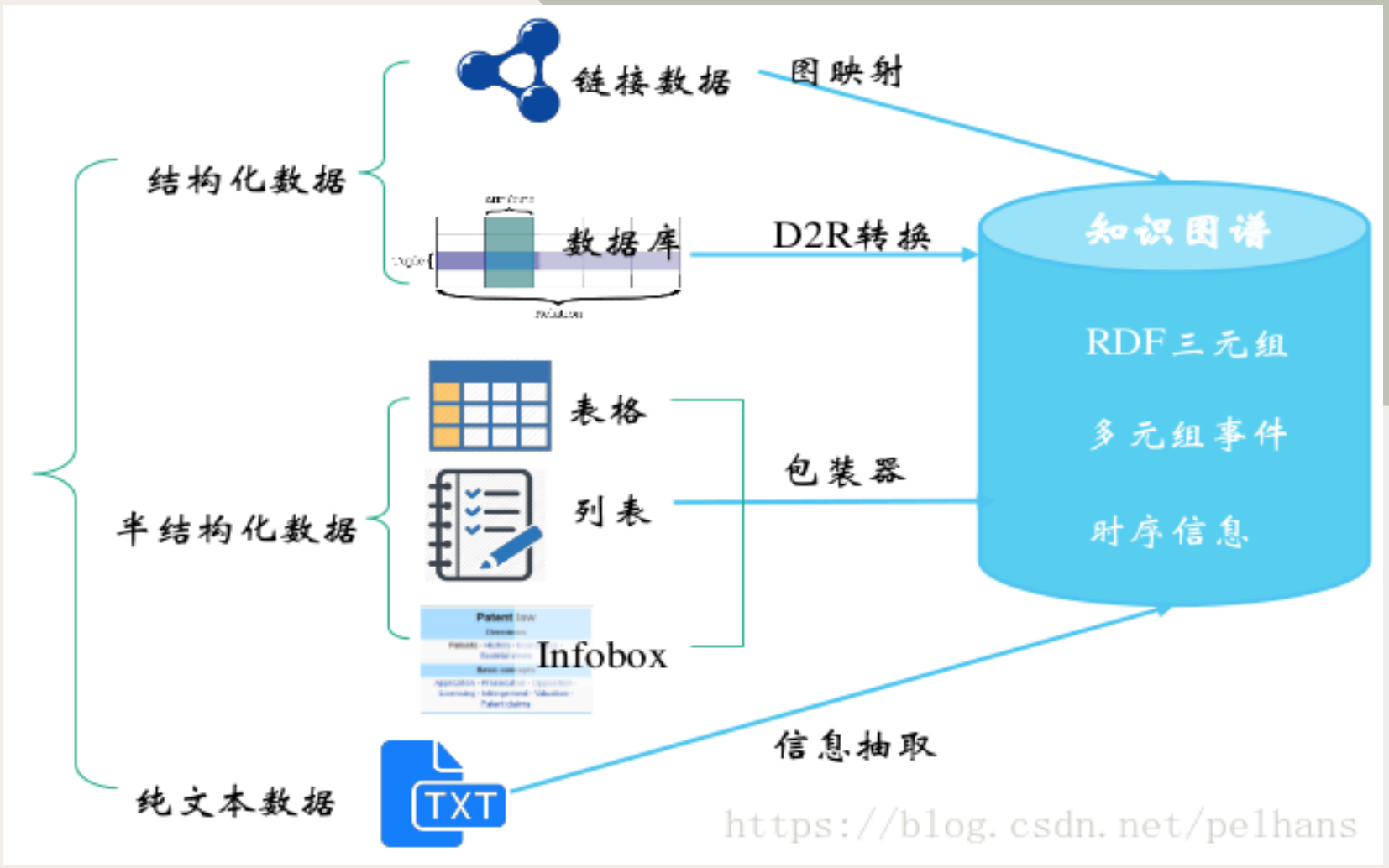
知识抽取是一种自动化地从半结构化和无结构数据中抽取实体、关系以及属性等结构化信息的技术。

在此基础上形成本体化的知识表达。

结构化数据

非结构化数据的不规则性和模糊行为
使得使用传统程序难以理解

半结构化数据
非结构化数据



知识抽取的种类

- 实体抽取
- 事件抽取
- 关系抽取
- 术语抽取
- 属性抽取
- 概念抽取

1. 实体抽取（命名体识别）

京东人工智能开放平台NeuHub

从文本数据集中自动识别出命名实体，实体抽取地质量（准确率和召回率）对后续的知识获取效率和质量的影响极大，因此是信息抽取中最为基础和关键的部分。

2. 关系抽取

京东人工智能开放平台NeuHub

文本语料经过实体抽取，得到的是一系列离散的命名实体，为了得到语义信息，还需要从相关预料中提取出实体之间的关联关系，通过关系将实体（概念）联系起来，才能够形成网状的知识结构。

3. 属性抽取

- 属性抽取的目标是从不同信息源中采集特定实体的属性信息。例如针对某个公众人物，可以从网络公开信息中得到其昵称、生日、国籍等信息。属性抽取技术能够从多种数据来源中汇集这些信息，实现对实体属性的完整勾画。
- 属性抽取较之关系抽取的难点在于，除了要识别实体的属性名还要识别实体的属性值，而属性值结构也是不确定的。因此大多研究都是基于规则进行抽取。
- 由于可以将实体的属性视为实体与属性值之间的一种名词性关系，因此可以将属性抽取问题视为关系抽取问题。

关于技术

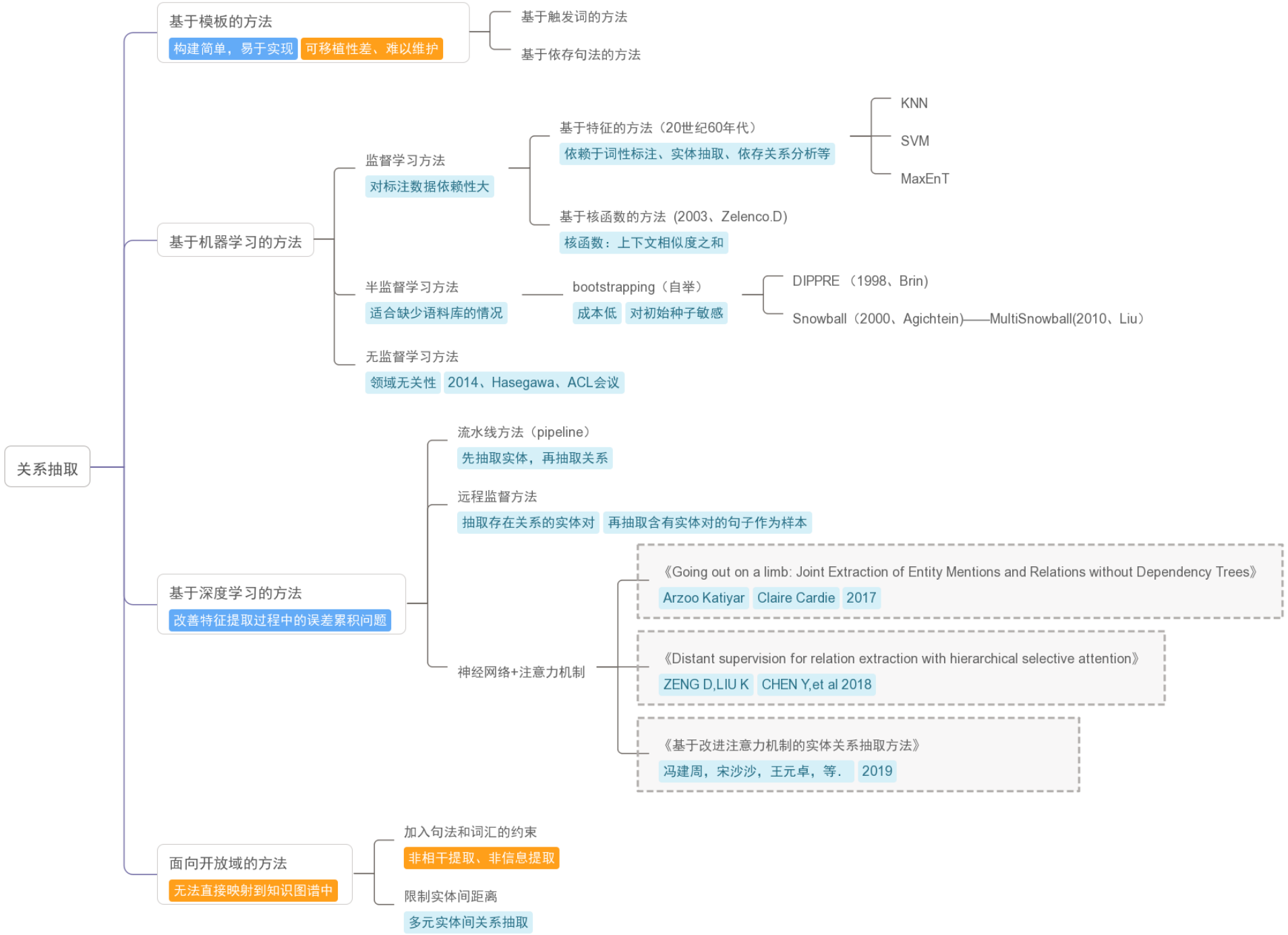
基于规则的方法

基于统计机器学习的方法

基于深度学习的方法

} 限定域  开放域

技术概览

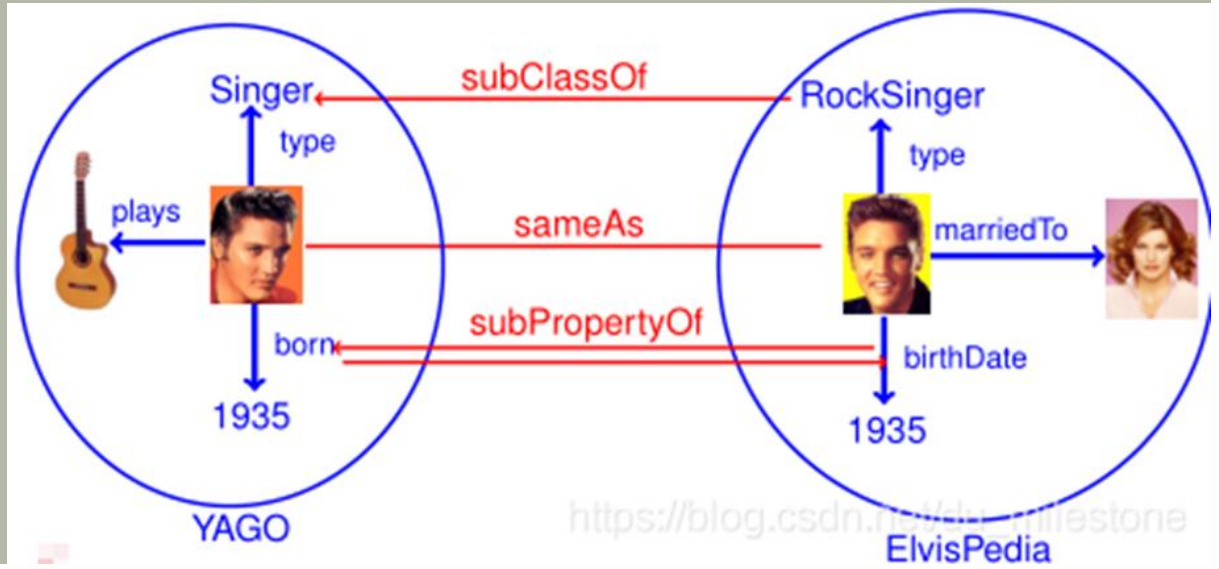

关系抽取
基于模板的方法
构建简单，易于实现
可移植性差、难以维护
基于触发词的方法
基于依存句法的方法
基于机器学习的方法
监督学习方法
对标注数据依赖性大
基于特征的方法（20世纪60年代）
依赖于词性标注、实体抽取、依存关系分析等
KNN
SVM
MaxEnT
基于核函数的方法（2003、Zelenco.D）
核函数：上下文相似度之和
半监督学习方法
适合缺少语料库的情况
bootstrapping（自举）
成本低
对初始种子敏感
DIPPRE （1998、Brin）
Snowball（2000、Agichtein)——MultiSnowball(2010、Liu）
无监督学习方法
领域无关性
2014、Hasegawa、ACL会议
基于深度学习的方法
改善特征提取过程中的误差累积问题
流水线方法（pipeline）
先抽取实体，再抽取关系
远程监督方法
抽取存在关系的实体对
再抽取含有实体对的句子作为样本
神经网络+注意力机制
《Going out on a limb: Joint Extraction of Entity Mentions and Relations without Dependency Trees》
Arzoo Katiyar
Claire Cardie
2017
《Distant supervision for relation extraction with hierarchical selective attention》
ZENG D,LIU K
CHEN Y,et al 2018
《基于改进注意力机制的实体关系抽取方法》
冯建周，宋沙沙，王元卓，等.
2019
面向开放域的方法
无法直接映射到知识图谱中
加入句法和词汇的约束
非相干提取、非信息提取
限制实体间距离
多元实体间关系抽取

知识融合

知识融合是高层次的知识组合。

可以消除概念的歧义，提出冗余和错误概念，

从而确保知识的质量。



将猫王从YAGO和ElvisPedia进行融合的例子

为什么要进行知识融合?

因为异构。

为什么会出现异构?

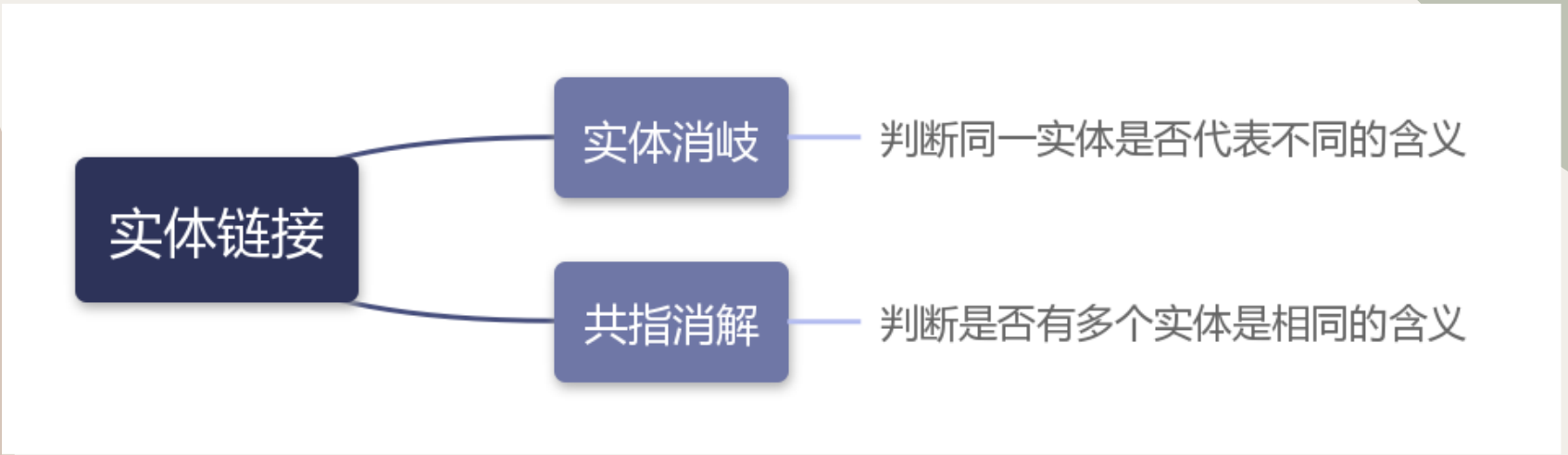
因为语言和概念层上的不匹配。

数据整合、消岐、加工、
推理验证、更新

数据、信息、方法、经
验以及人的思想的融合

要得到高质量的知识库!

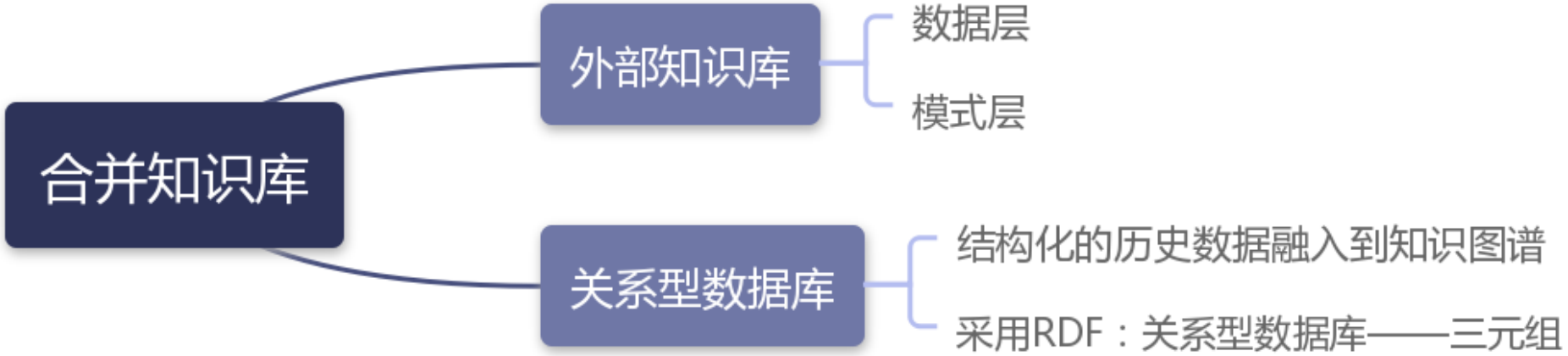
消岐



学术界对该问题有多种不同的表述
实体对齐，本体匹配.....
本质上一样

聚类、支持向量机、决策树.....
关键：相似性度量
用什么相似性函数?

合并


合并知识库
外部知识库
数据层
模式层
关系型数据库
结构化的历史数据融入到知识图谱
采用RDF：关系型数据库——三元组

知识表示、知识存储、知识推理

PART.03

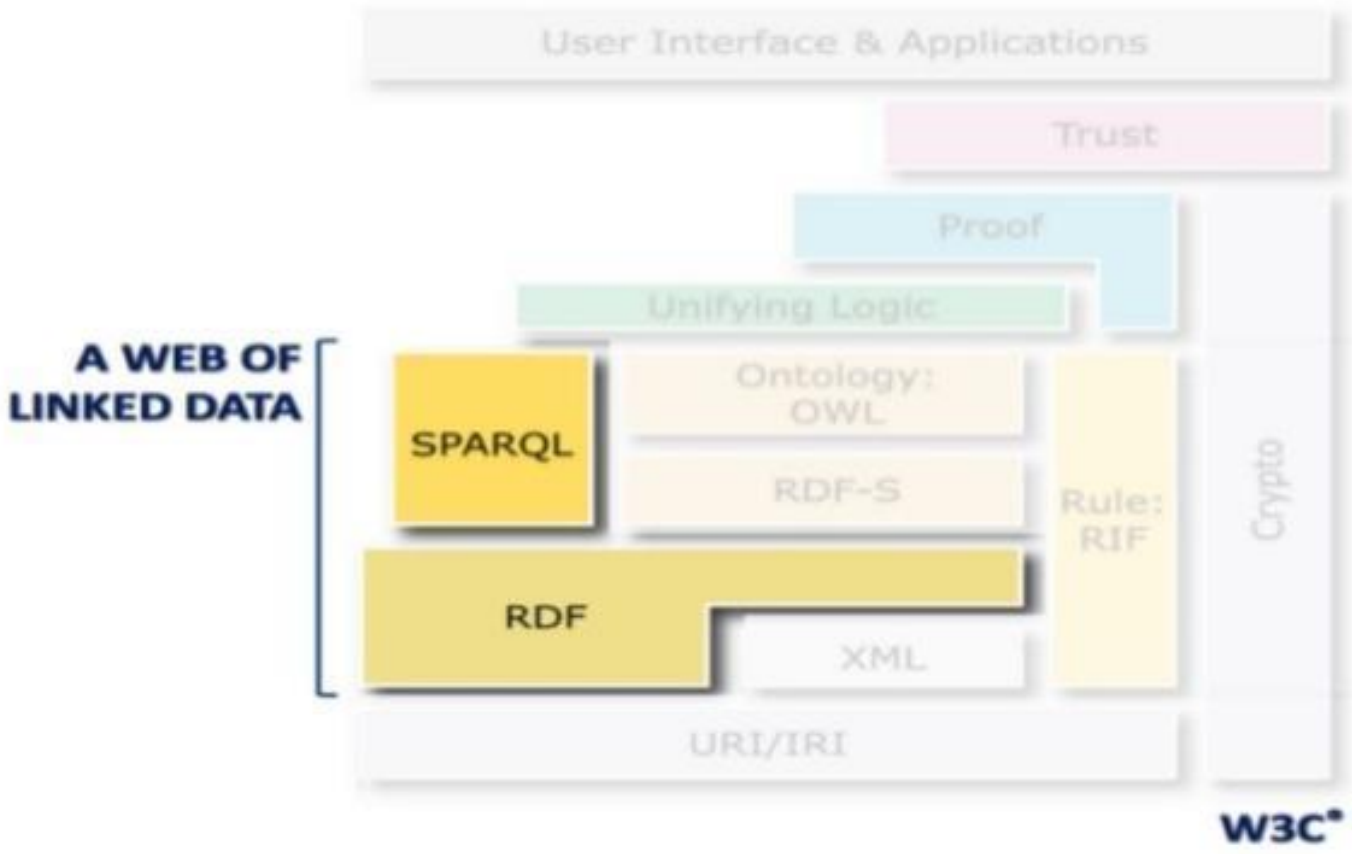
知识表示的概念

知识表示就是对知识的一种描述，或者说是对知识的一组约定，一种计算机可以接受的用于描述知识的数据结构。它是机器通往智能的基础，使得机器可以像人一样运用知识.

早期的知识表示方法：

- 一阶谓词逻辑
- 产生式系统
- 框架表示法
- 语义网络

基于语义网的知识表示框架

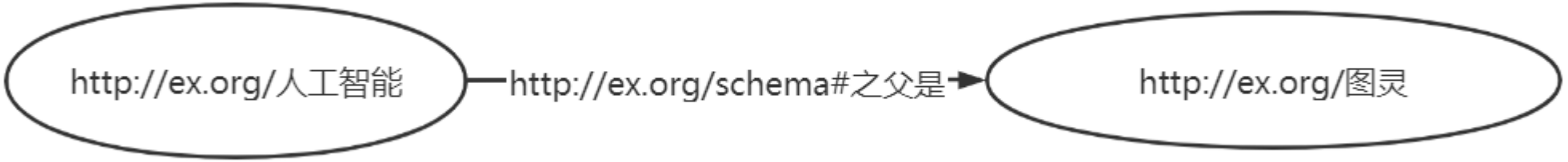


RDF（资源描述框架）

- R：页面，图片、视频等任何具有URI标识符的资源
- D：标识属性、特征和资源之间的关系
- f：标识模型、语言和这些描述的语法。

在RDF中，知识总是以三元组的形式出现，即每一份知识都可以被分解为：(subject, predicate, object)。

与此同时，RDF三元组可以看做是图模型的边和顶点（vertex, edge, vertex）。



RDF和RDFS

RDFS(RDF Schema)在RDF的基础上提供了一个术语、概念的定义方式，以及那些属性可以应用到哪些对象上。换句话说，RDFS为RDF模型提供了一个基本的类型系统。

RDFS支持推理功能

SPARQL

```
# prefix declarations
PREFIX foo: <http://example.com/resources/>
...
# dataset definition
FROM ...
# result
                    clause
SELECT ...
# query pattern
WHERE {
    ...
}
# query modifiers
ORDER BY ...
```

SPARQL是RDF的查询语言，它基于RDF数据模型，可以对不同的数据集撰写复杂的连接，同时还被所有主流的图数据库支持。

SPARQL与SQL的主要区别

变量,RDF中的资源,以“?”或者“$”指示

在WHERE子句中列示关联的三元组模板

SELECT子句中指示要查询的目标变量

```
SELECT ?student  ?email
WHERE {
    ?student exp:studies exp:CS328 .
OPTIONAL {
    ?student foaf:mbox ?email .
   }
}
```

知识存储的概念

知识存储，即获取到的三元组如何存储在计算机中。目前业内存储知识的方式有三种，第一种为三元组形式的RDF存储；第二种为传统关系型数据库存储；第三种为图数据库存储。

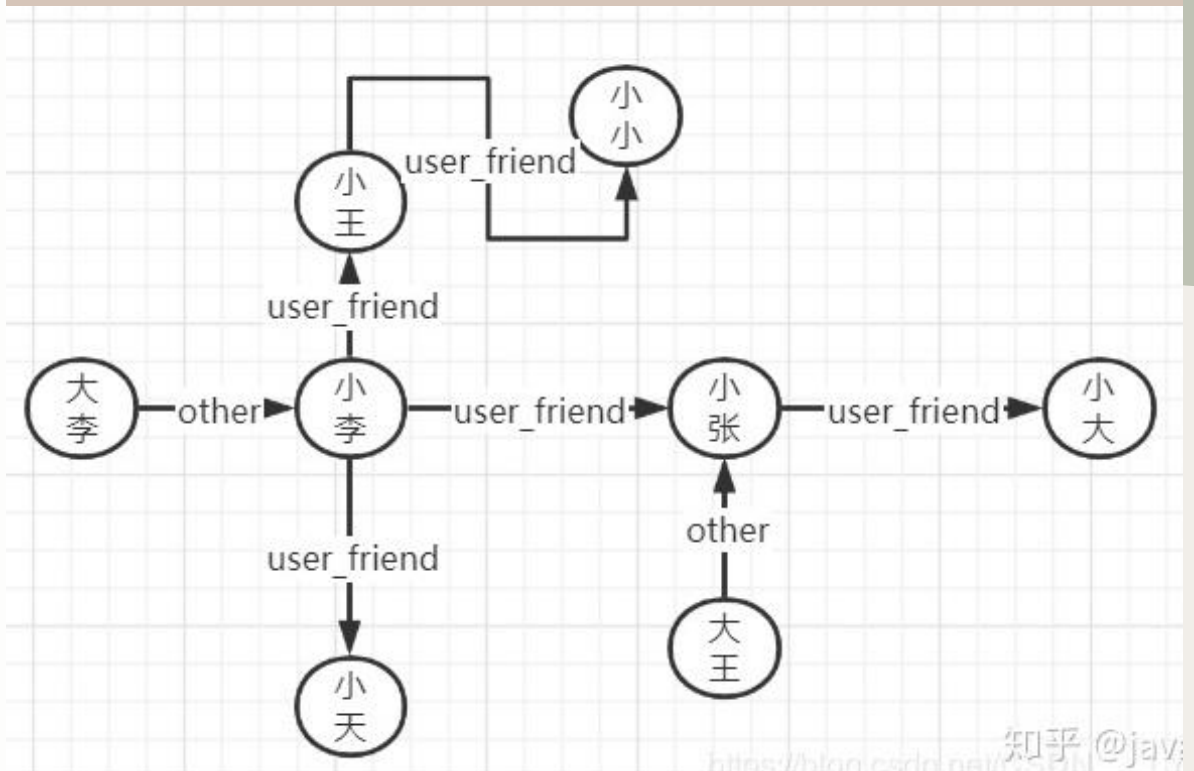
一个 RDF 数据集由一组相关的三元组的组成。由于这个三元组集合可以抽象为一张 graph，因此也称为 RDF graph。

图数据库

图形数据库是NoSQL数据库的一种类型，起源于欧拉理论和图理论，也可称为面向/基于图的数据库。

在图数据库中图将实体表现为节点，实体与其他实体连接的方式表现为联系（边）。

示例



图数据库

高性能

图模型固有的数据索引结构，使得它的数据查询与分析速度更快。

灵活

使用者可以根据业务变化随时调整数据模型。

敏捷

图数据库的图模型非常直观，支持测试驱动开发模式。

简便

很多图数据库提供了专业的分析算法、工具。

目前主流的图数据库有：Neo4j，Janusgraph，Dgraph，Giraph，TigerGraph等。

知识推理的概念

所谓的知识推理，就是在已有知识的基础之上，推断出未知的知识的过程。通过已经获取的知识，得到所蕴含的新的事实，或者从大量的已有知识中归纳，从个体知识推广到一般性的知识。

针对知识图谱特有的三元组存储形式,面向知识图谱的知识推理被定义为对三元组缺失部分的预测,更主要的是对实体和关系进行的预测,一般在<实体关系-实体>三元组中进行。

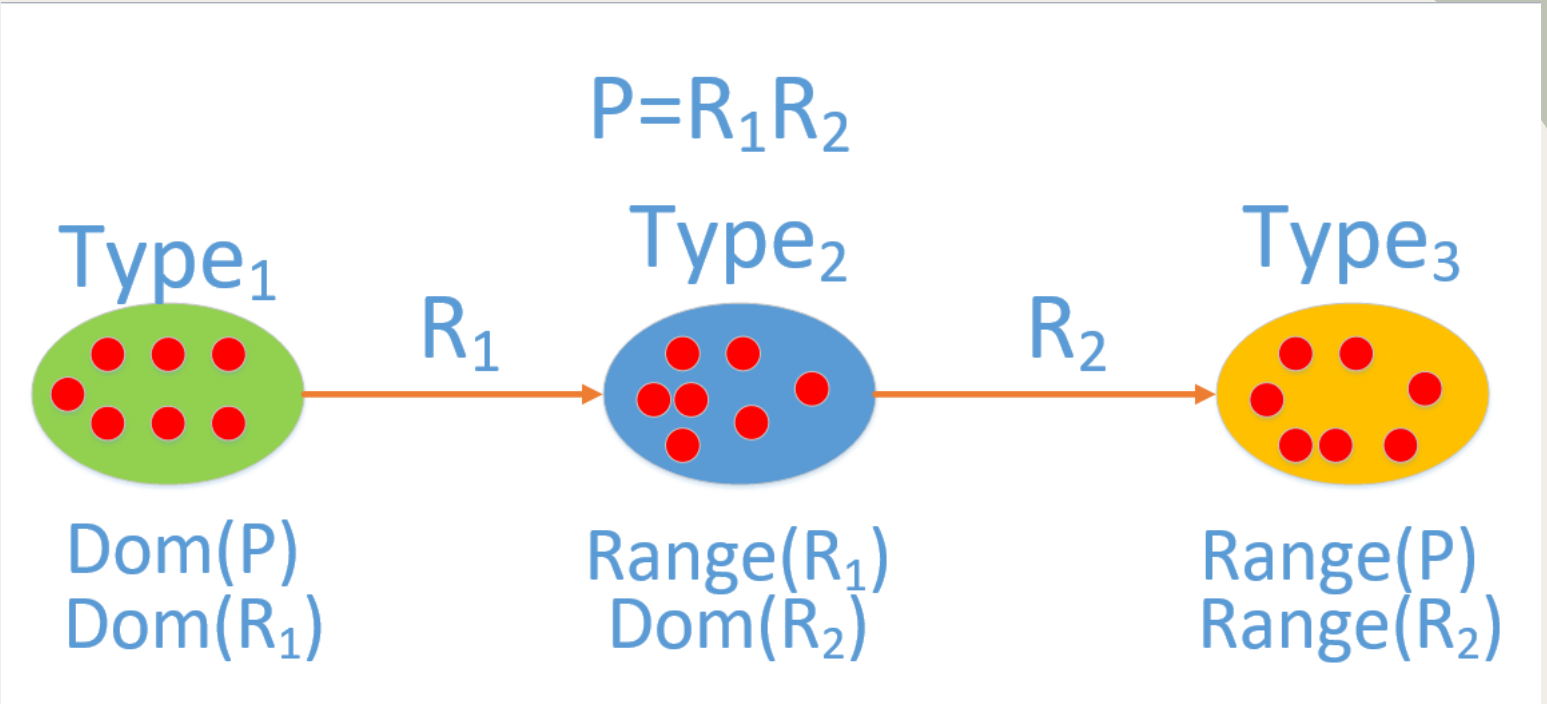
知识推理的分类

面向知识图谱的知识推理

- 基于图结构和统计规则挖掘的推理
- 基于知识图谱表示学习的推理
- 基于神经网络的推理
- 混合推理

基于图结构和统计规则挖掘的推理

PRA：路径排序算法

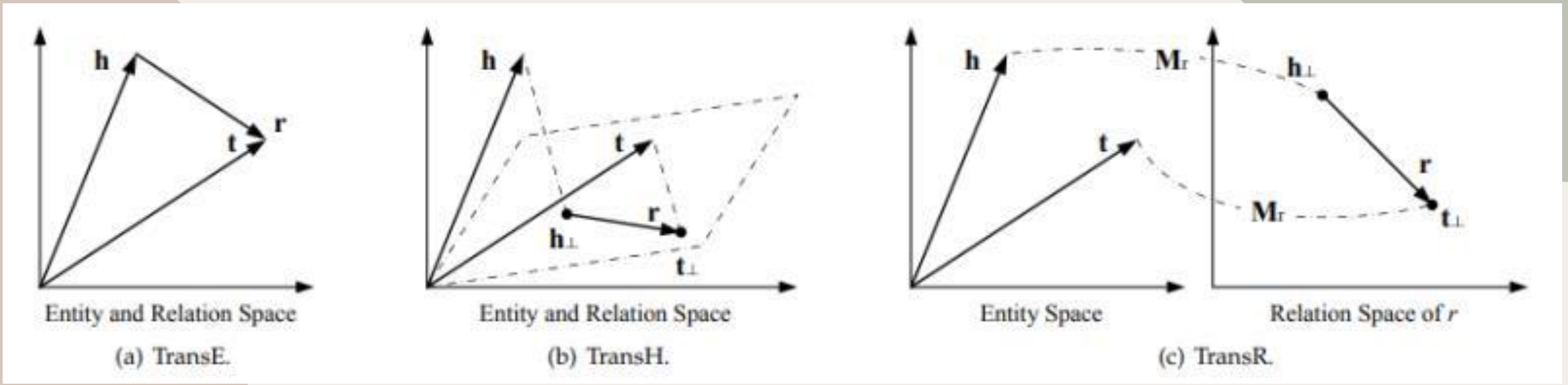


优点：可解释性强，能够自动发现推理规则

缺点：1. 处理低频关系效果差 2. 处理低连通图效果差3. 图较大时处理速度慢

基于知识图谱表示学习的推理

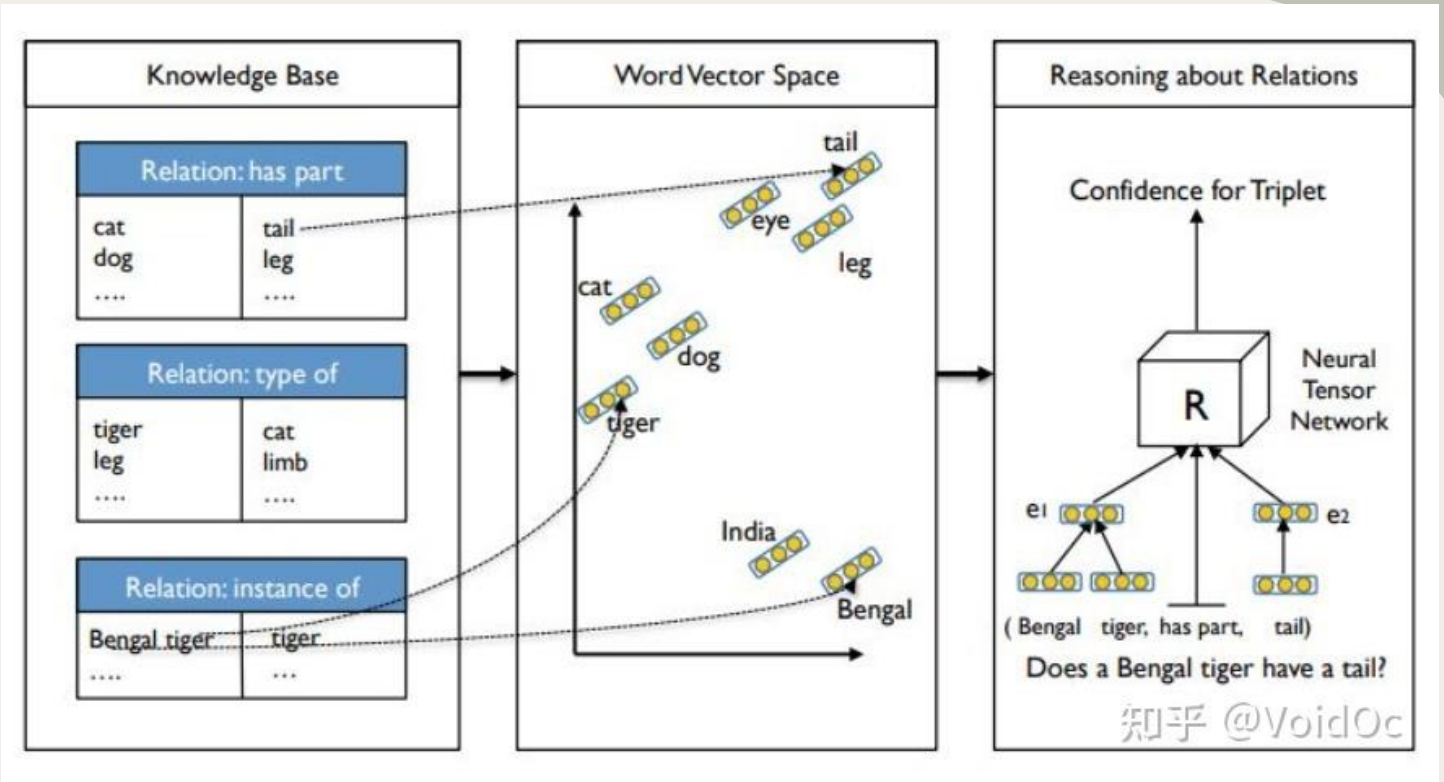
Trans 系列



Trans E：适合处理一对一的关系
Trans H：适合处理多对多的关系
Trans R：每个实体包含不同方面

基于神经网络的推理

NTN



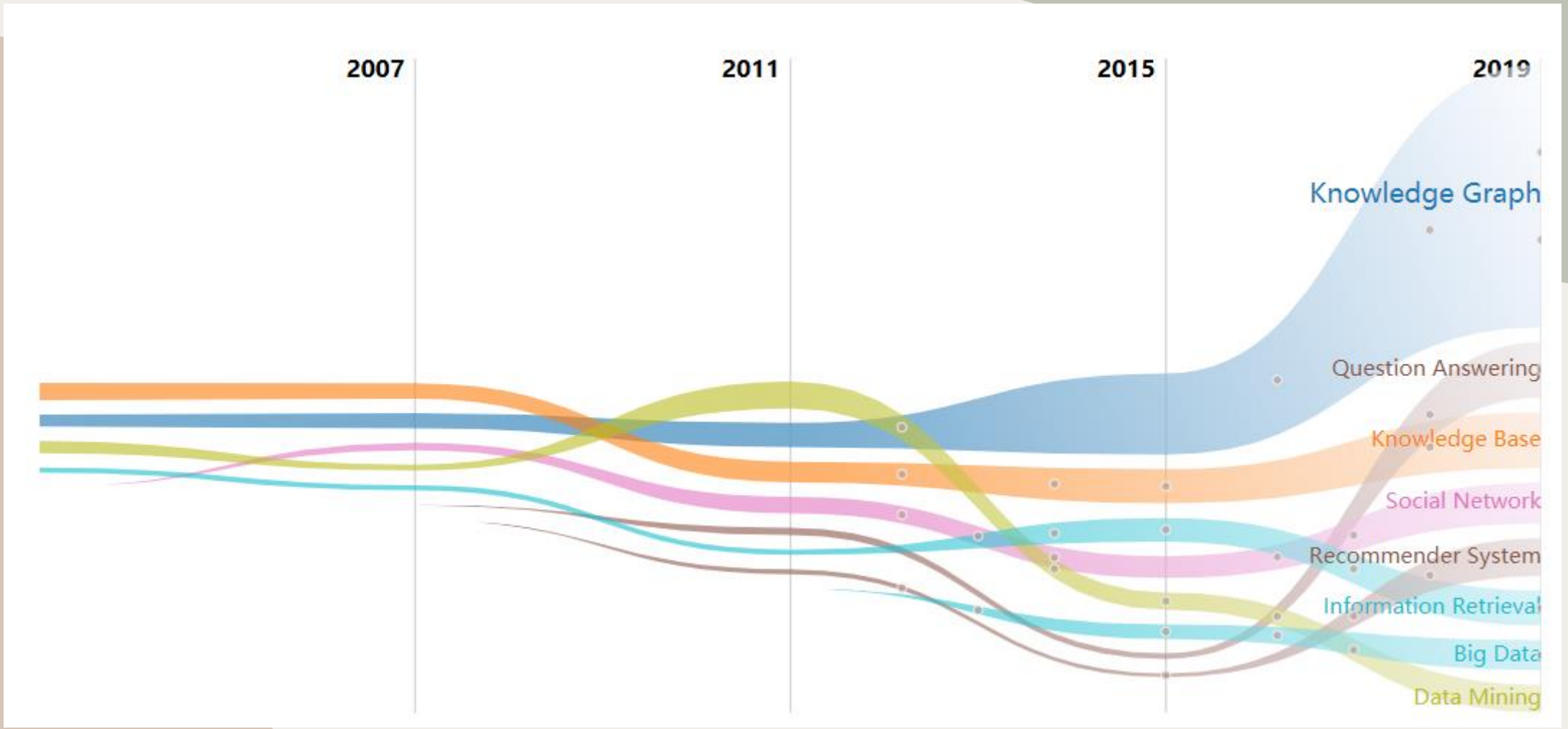
优点：实体中的单词数量远小于实体数量，可以重复利用单词向量构建实体表示

缺点：复杂度非常高在大规模稀疏知识图谱上的效果较差

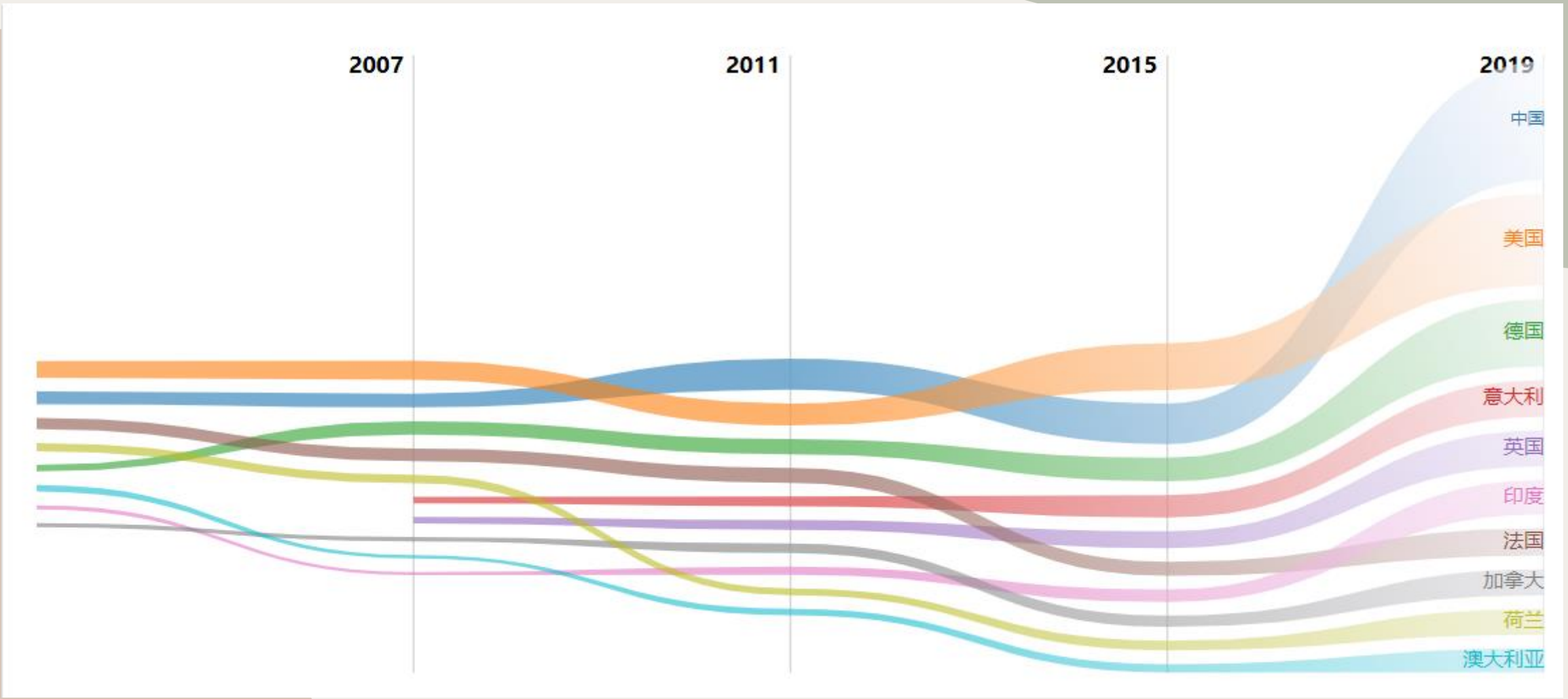
前沿进展

PART.04

知识图谱及其子领域研究趋势



各国知识图谱的研究情况



前沿工作：从顶会论文中发掘热点

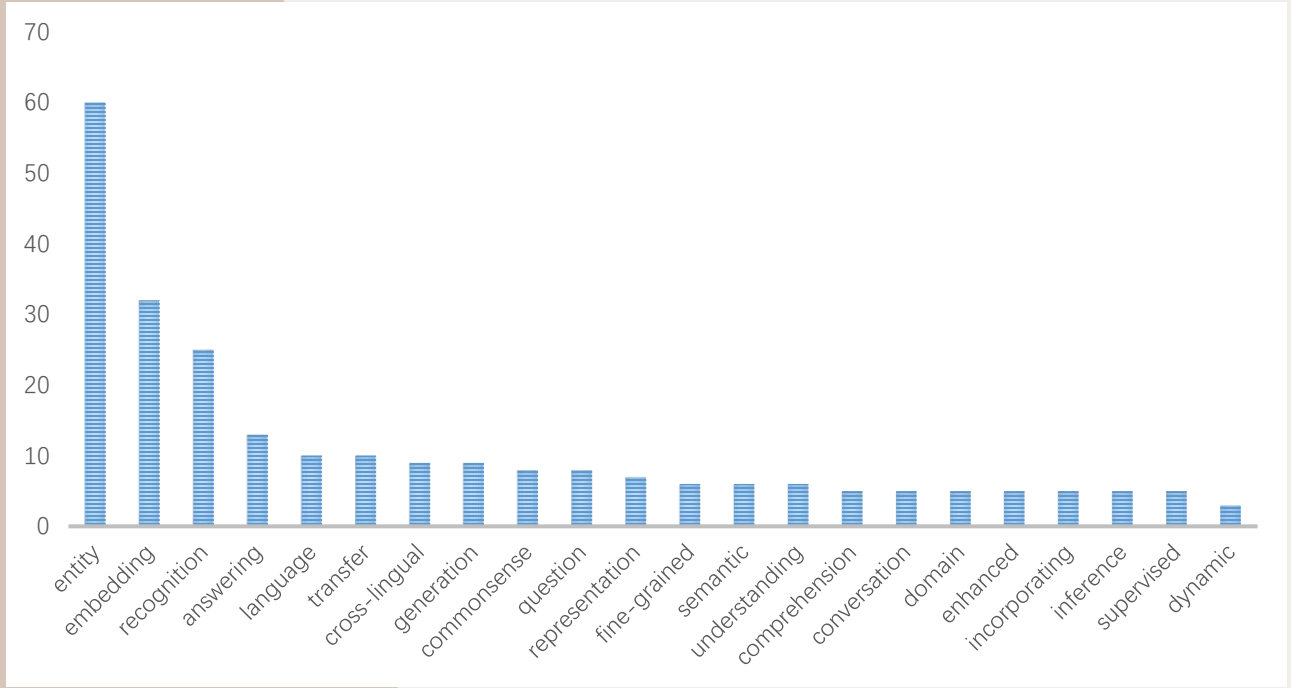
- 人工智能
- AAAI
- IJCAI
- 自然语言处理
- EMNLP
- ACL
- 学习表征
- ICLR
- 数据挖掘
- ICDM
- KDD

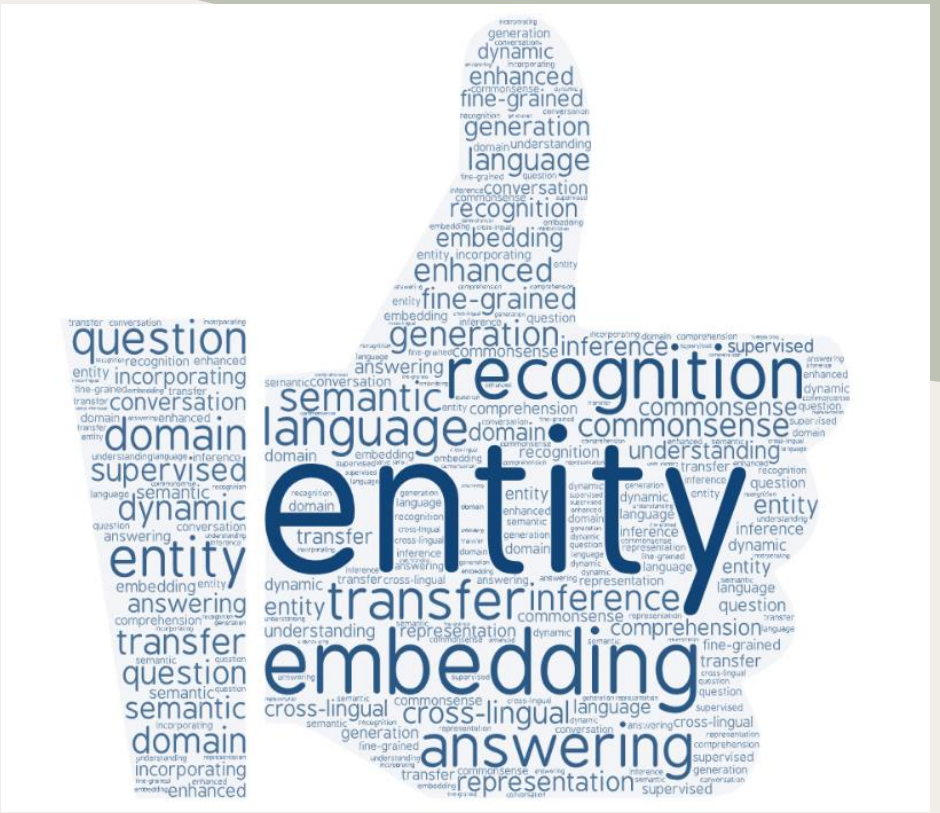
Publications in ACL2020 Groupd by Tracks

Name Entity Recognition

- Named Entity Recognition without Labelled Data: A Weak Supervision Approach
- Improving Multimodal Named Entity Recognition via Entity Span Detection with Unified Multimodal Transformer
- Code and Named Entity Recognition in StackOverflow
- A Unified MRC Framework for Named Entity Recognition
- An Effective Transition-based Model for Discontinuous NER
- Improving Low-Resource Named Entity Recognition using Joint Sentence and Token Labeling
- Multi-Cell Compositional LSTM for NER Domain Adaptation
- Pyramid: A Layered Model for Nested Named Entity Recognition
- Simplify the Usage of Lexicon in Chinese NER
- Bipartite Flat-Graph Network for Nested Named Entity Recognition
- Connecting Embeddings for Knowledge Graph Entity Typing
- Named Entity Recognition as Dependency Parsing
- Neighborhood Matching Network for Entity Alignment
- Single-/Multi-Source Cross-Lingual NER via Teacher-Student Learning on Unlabeled Data in Target Language
- FLAT: Chinese NER Using Flat-Lattice Transformer

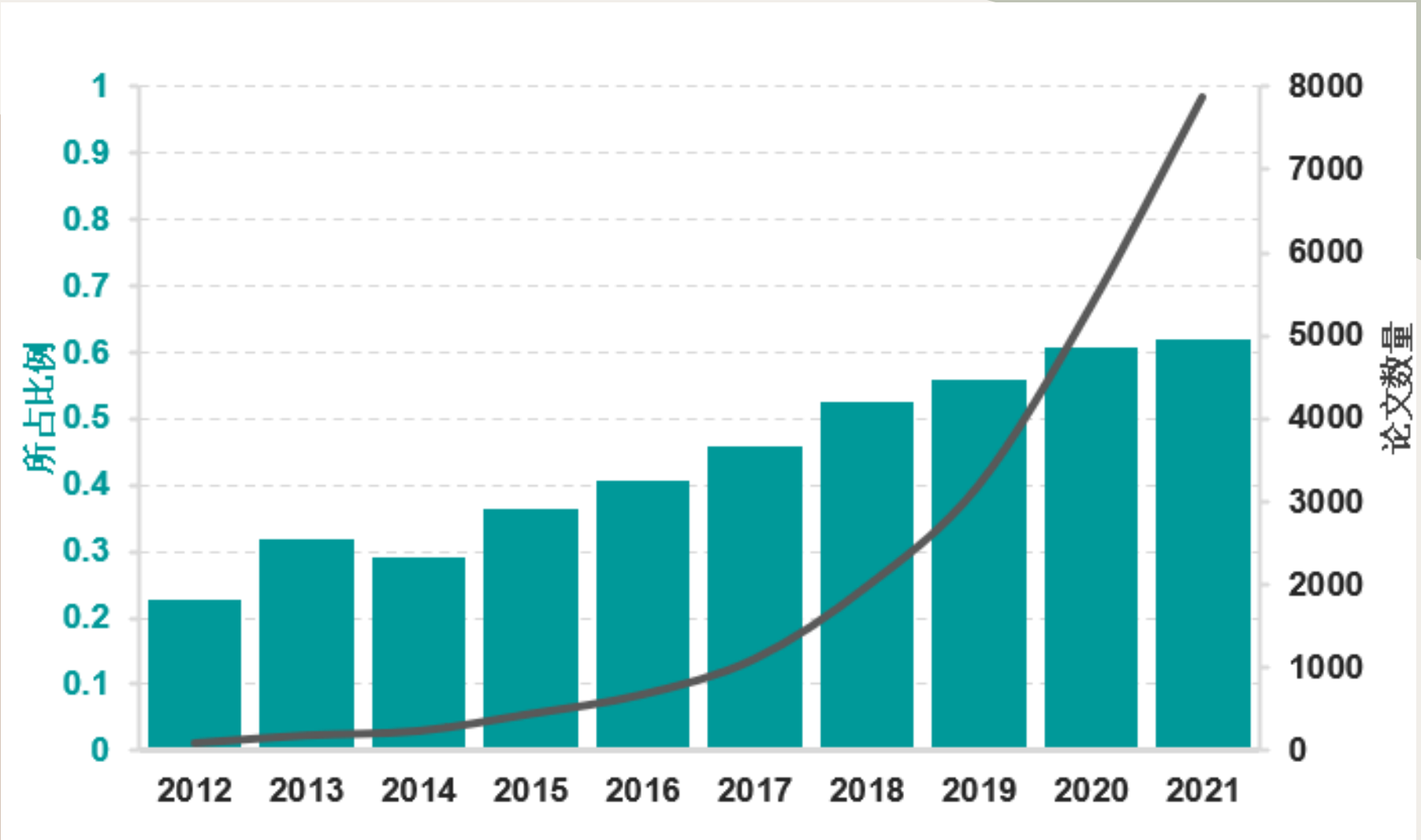
研究热点



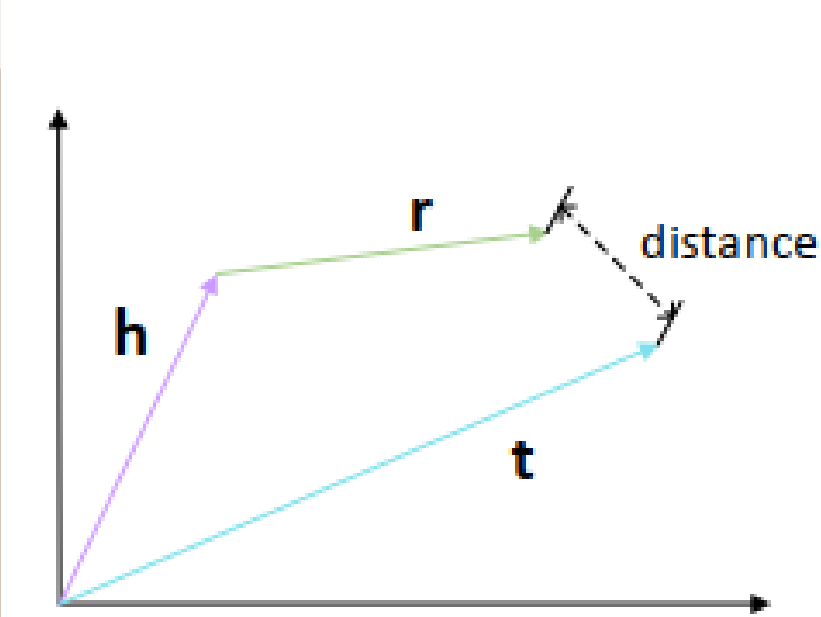


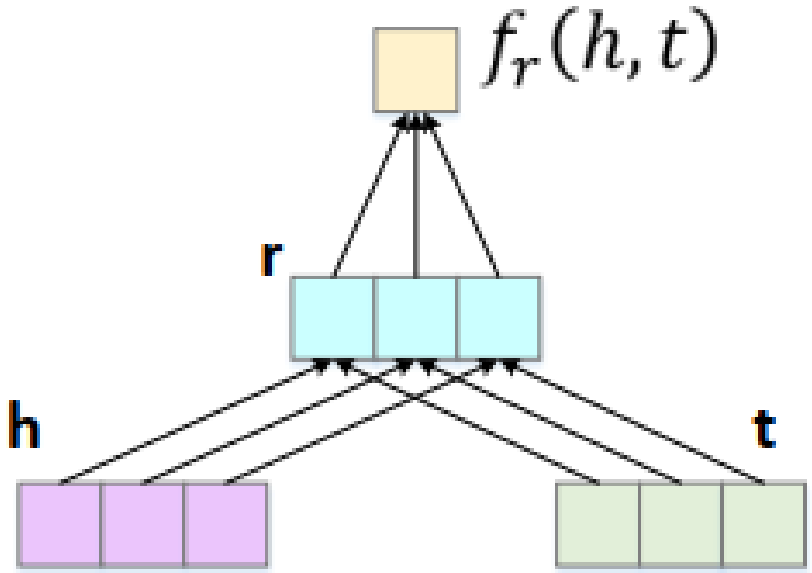
Embedding Explainable Transformer Dynamic

Embedding:从离散的符号化知识表示到连续的向量知识表示



Embedding:从离散的符号化知识表示到连续的向量知识表示





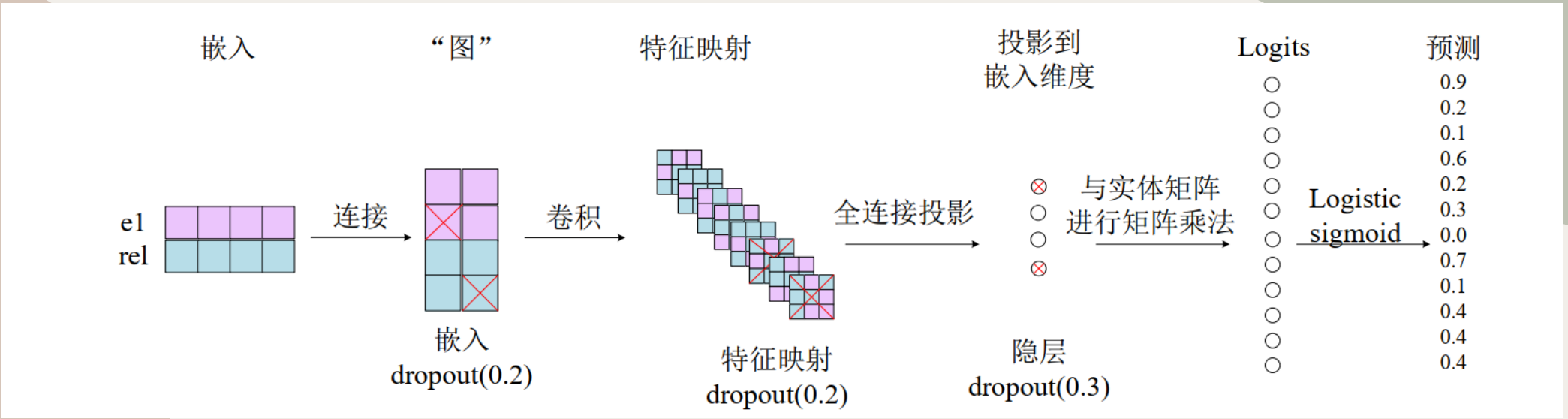
基于平移距离的TransE评分

基于语义相似度的DisMult评分

Antoine Bordes, Nicolas Usunier, Alberto García-Durán, Jason Weston, Oksana Yakhnenko. Translating Embeddings for ModelingMulti-relational Data. In: Proc. of the 27th Neural Information Processing Systems(NIPS). Lake Tahoe: Neural informationprocessing systems foundation, 2013. 2787-2795.

Bishan Yang, Wen-tau Yih, Xiaodong He, Jianfeng Gao, Li Deng. Embedding Entities and Relations for Learning and Inference inKnowledge Bases. In: Proc. of the 3th International Conference on Learning Representations(ICLR). San Diego: InternationalConference on Learning Representations, 2015.

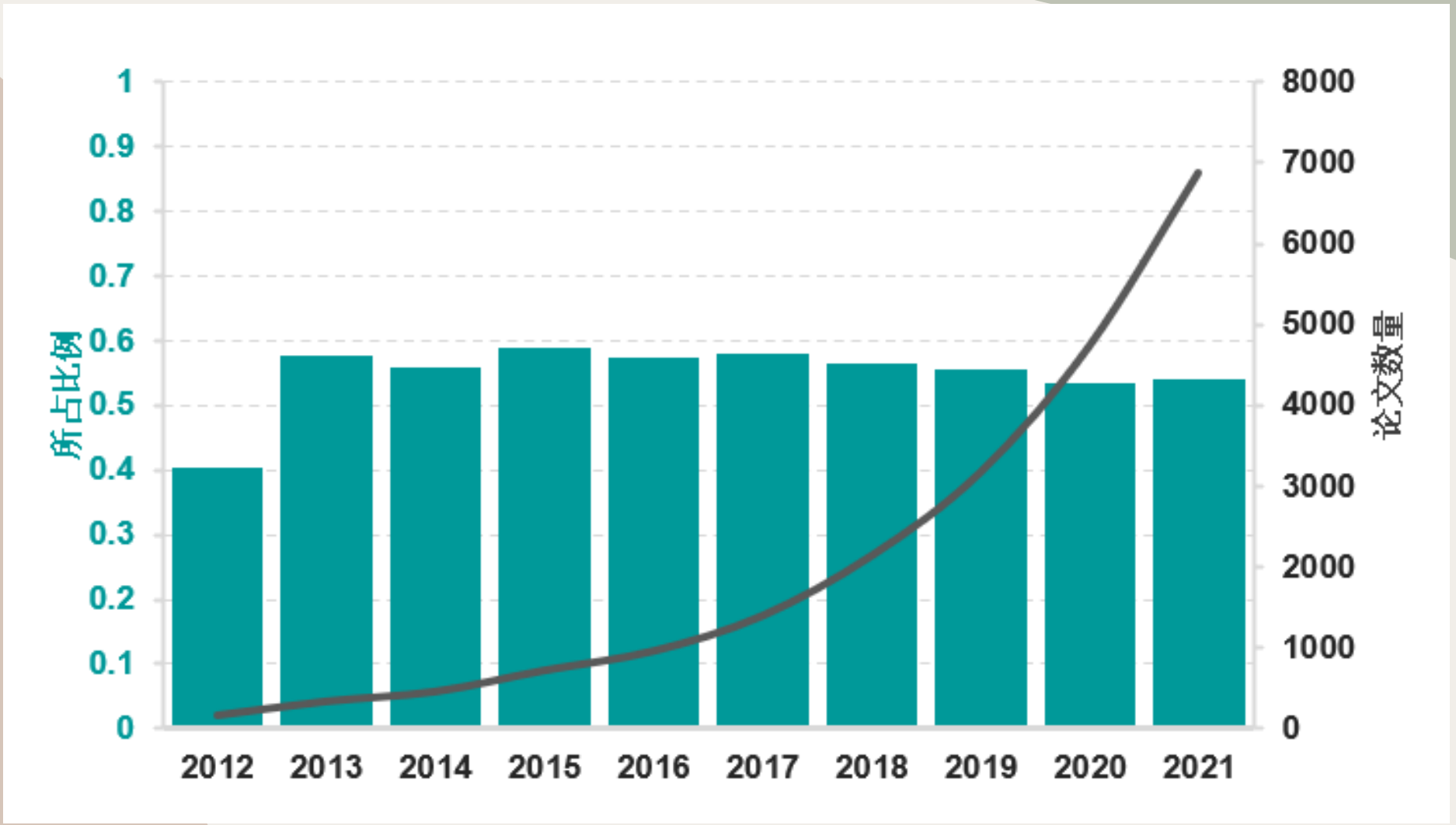
Embedding:从离散的符号化知识表示到连续的向量知识表示



convE：卷积的应用

Tim Dettmers, Pasquale Minervini, Pontus Stenetorp, Sebastian Riedel. Convolutional 2D Knowledge Graph Embeddings. In： Proc. of the 32th AAAI Conference on Artificial Intelligence(AAAI). New Orleans: AAAI press, 2018. 1811-1818.

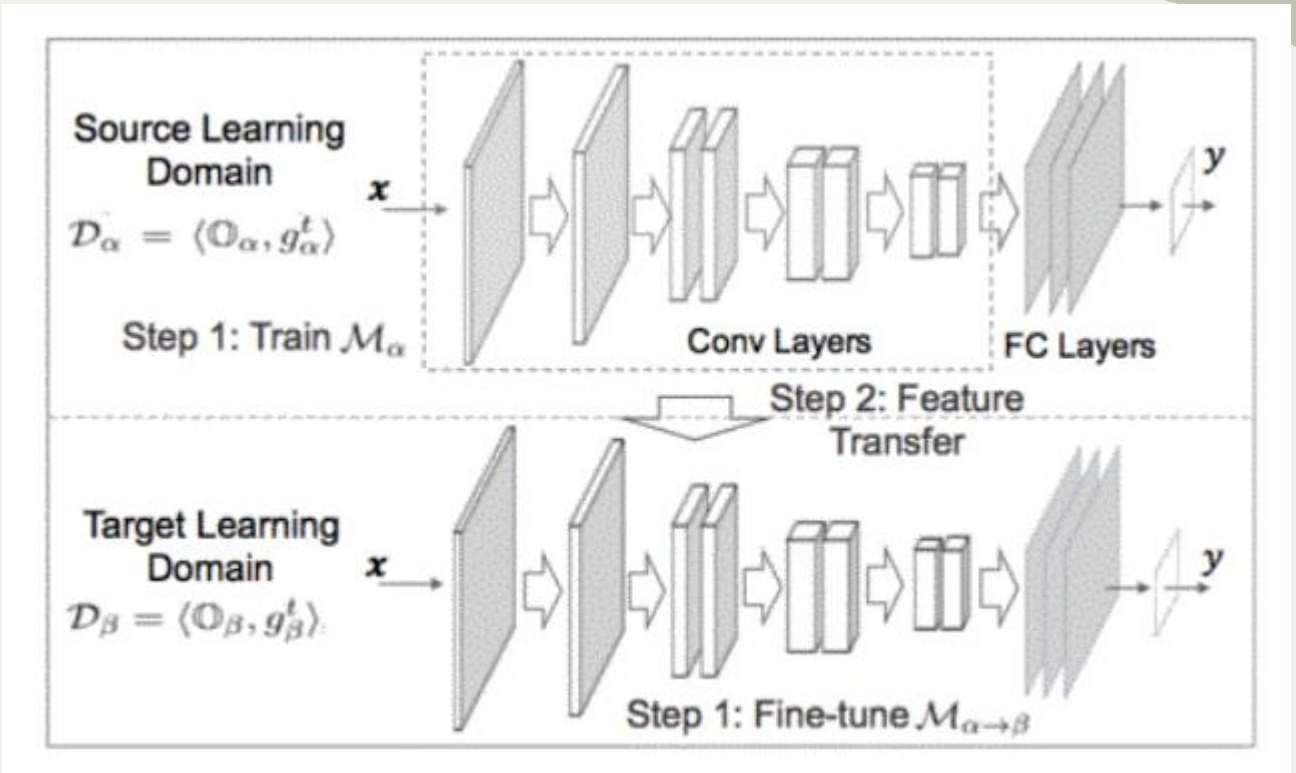
Explainable：黑盒的曙光



Explainable：黑盒的曙光

可解释性：将知识图谱作为先验知识融入深度学习模型，提升模型可解释性

ITransF



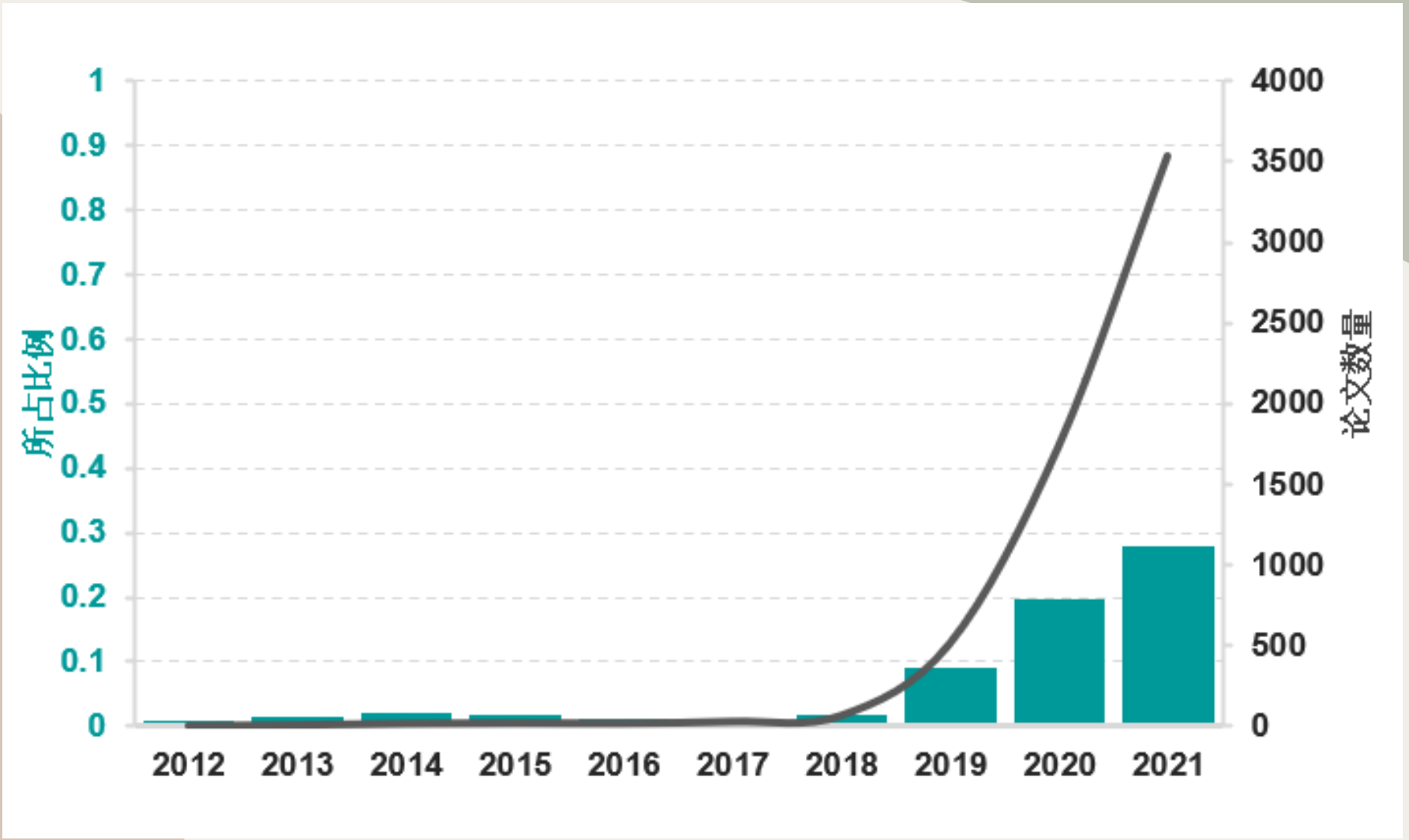
CrossE

可解释的AI：Explainable AI (XAI)

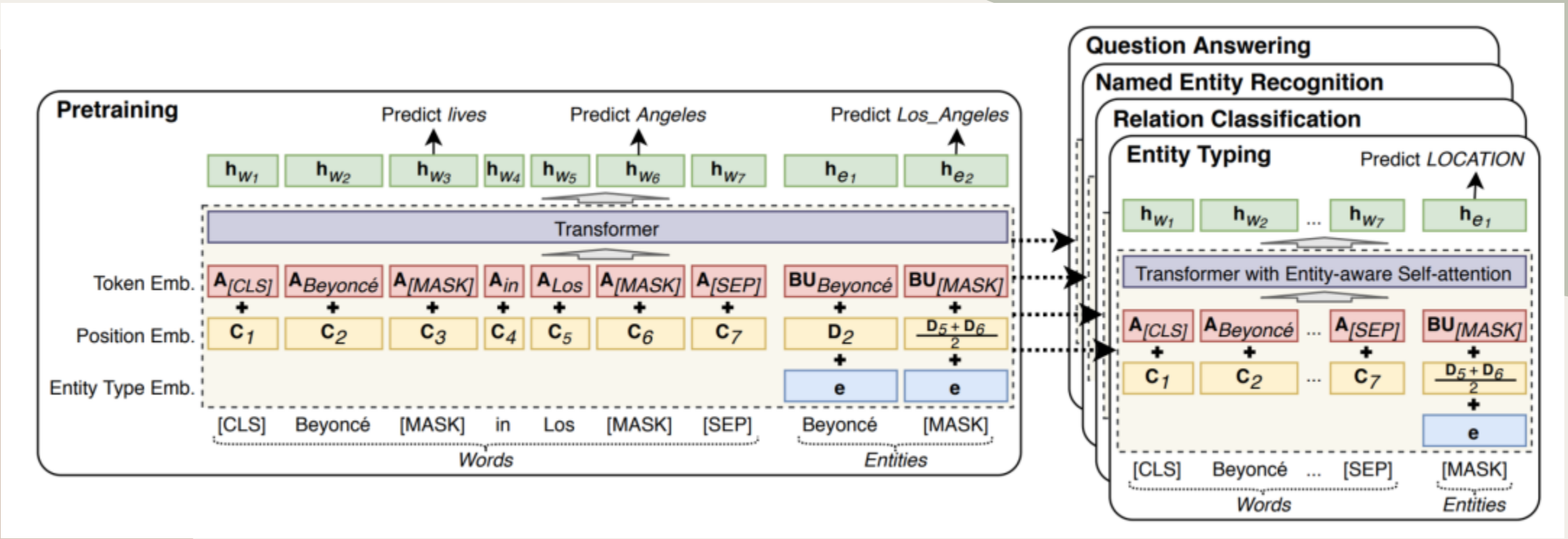
Qizhe Xie, Xuezhe Ma, Zihang Dai, Eduard H. Hovy. An Interpretable Knowledge Transfer Model for Knowledge Base Completion. In： Proc. of the 55th Conference of Association for Computational Linguistics(ACL). Vancouver: Association for Computational Linguistics (ACL), 2017. 950-962.

Wen Zhang, Bibek Paudel, Wei Zhang, Abraham Bernstein, Huajun Chen. Interaction Embeddings for Prediction and Explanation in Knowledge Graphs. In： Proc. of the 12th ACM International Conference on Web Search and Data Mining(WSDM). New York: Association for Computing Machinery, 2019. 96-104.

Transformer：知识图谱增强语言模型



Transformer：知识图谱增强语言模型

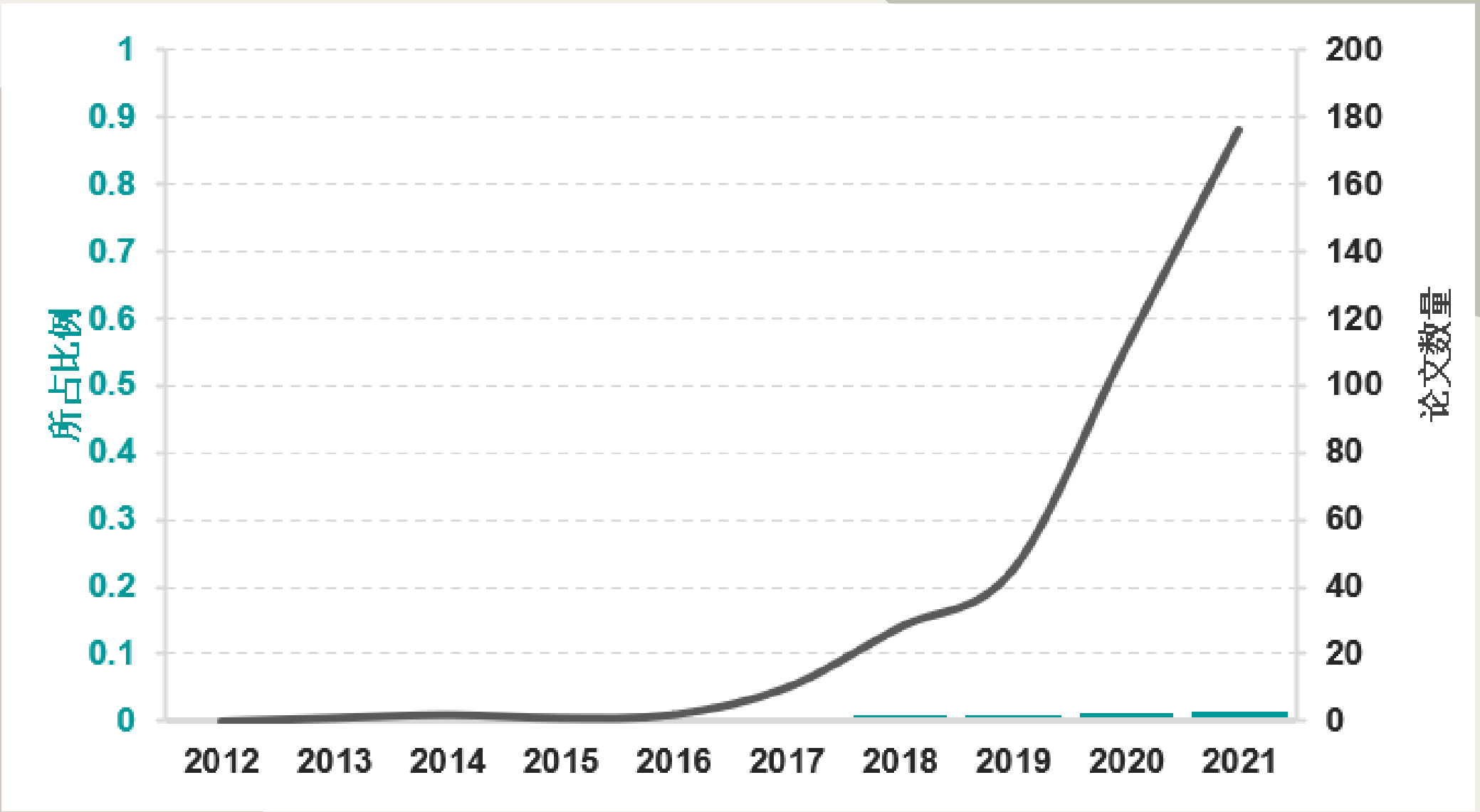


增强过后的Transformer

多语言嵌入仍是挑战…

LUKE: Deep Contextualized Entity Representations with Entity-aware Self-attention

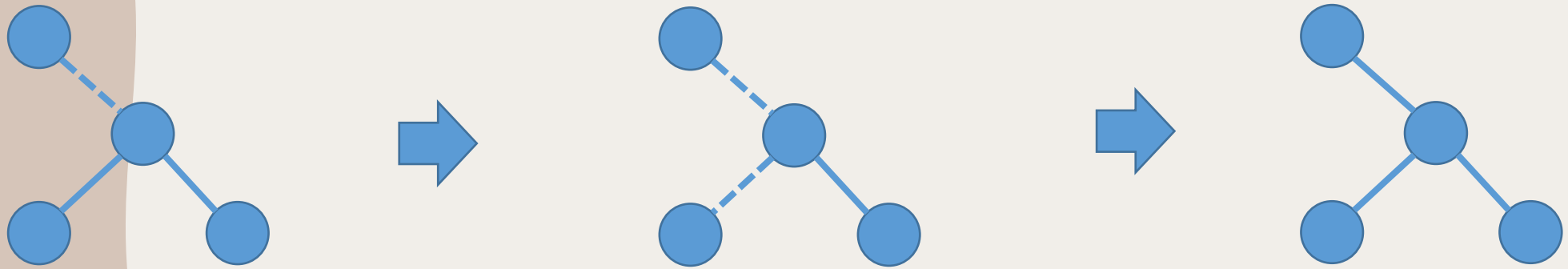
Dynamic：时序知识图谱



Dynamic：时序知识图谱

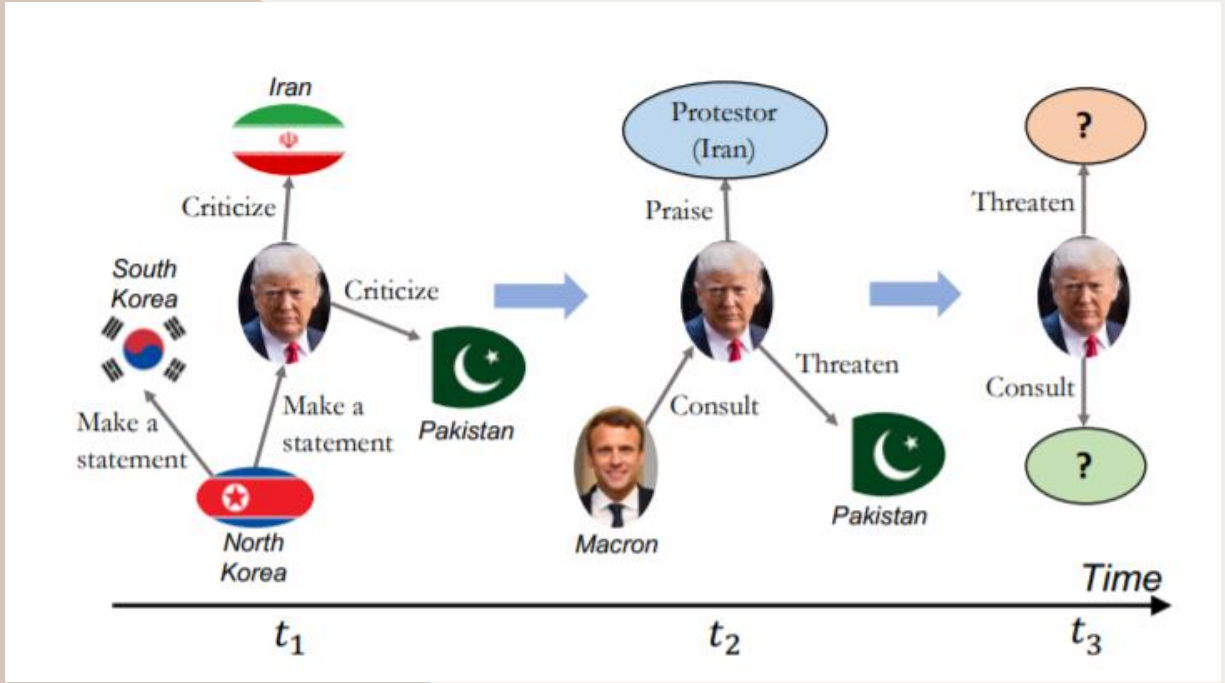
从（h，r，t）的三元组扩展为（h，r，t，τ）的时序四元组

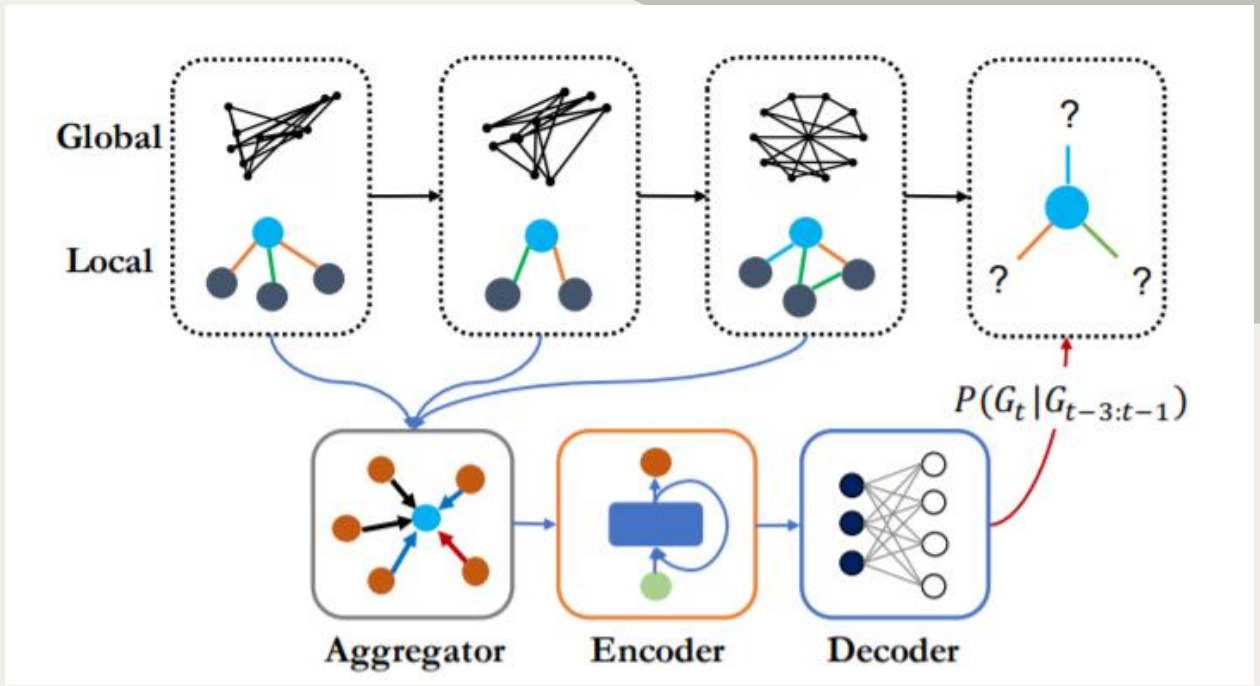
$$f_\tau(h,r,t) = -\|\mathbf{h} + \mathbf{r} + \tau - \mathbf{t}\|_{L_{1/2}}$$



J. Leblay and M. W. Chekol,“Deriving validity time in knowledge graph,” in WWW, 2018, pp. 1771–1776.

Dynamic：时序知识图谱





W. Jin, C. Zhang, P. Szekely,and X. Ren, “Recurrent event network for reasoning over temporal knowledge graphs,” in ICLR RLGM Workshop, 2019.

更多可能……

自动构建和动态变化：有监督学习

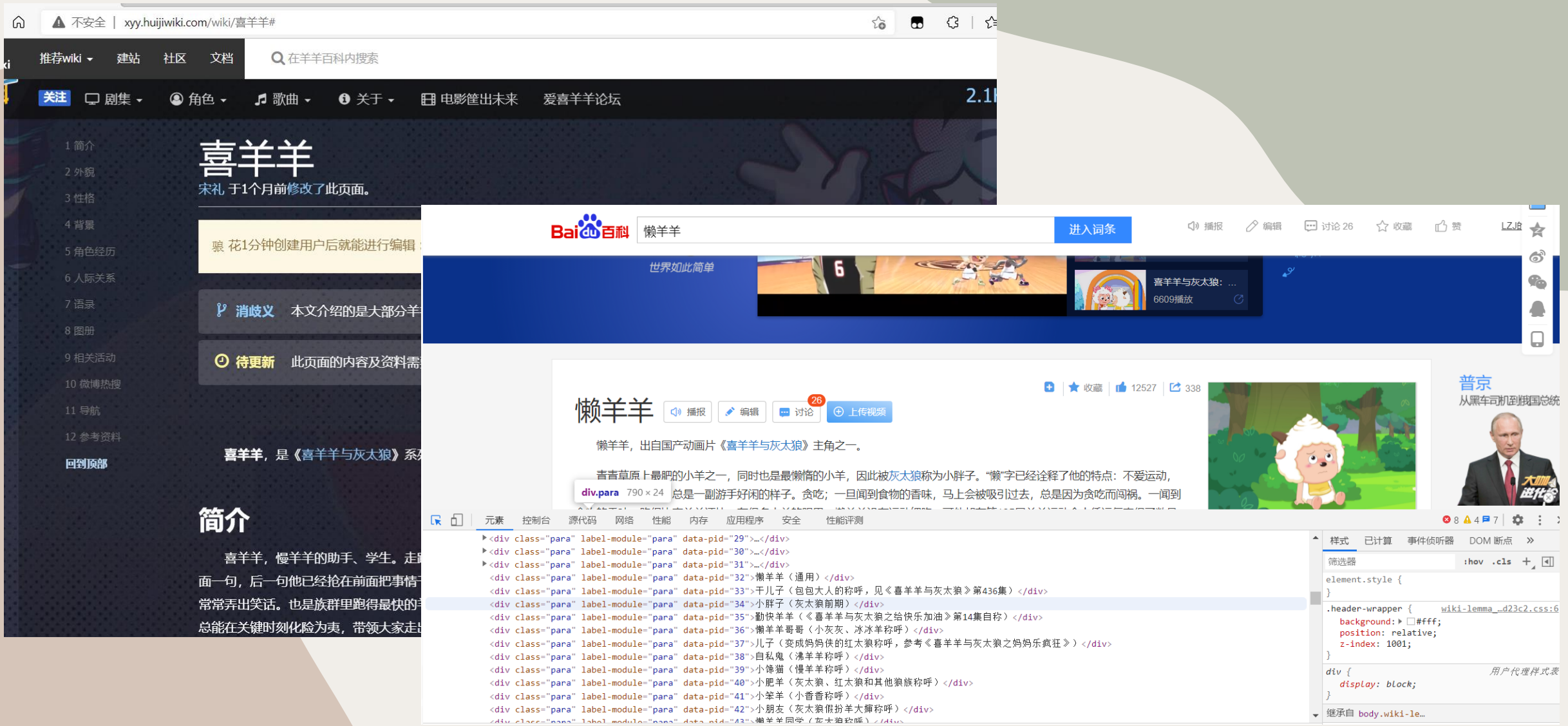
知识图谱X计算机视觉?
图片分类 鸢尾花分类
图片生成 多样性的问题

DEMO

PART.05

一、知识图谱的构建

知识抽取：结构化数据获取实体与三元组



知识抽取：结构化数据获取实体与三元组

喜羊羊.txt - 记事本

文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

```
中文名$$喜羊羊
外文名$$Weslie（主要）Pleasant GoatHappy（电影前2部）
别　名$$0803（威哥称）、喜太狼、喜羊羊哥哥（小灰灰、冰冰羊称）、儿子、喜先生（灰太狼称）、
配　音$$祖晴(普通话，主要)、杨阳洋(《羊年喜羊羊》幼年)、常蓉珊（《十万个冷笑话2》）、刘红韵、
性　别$$男
登场作品$$喜羊羊与灰太狼十万个冷笑话2
生　日$$羊历3506年5月25日
年　龄$$12岁（羊历3518年，《喜羊羊与灰太狼之奇幻天空岛》及以后的时间）
虚拟人物血型$$O型
象征颜色$$蓝色
出生地$$青青草原
院　校$$大肥羊学校
种　族$$羊族
父　亲$$智羊羊
母　亲$$丽羊羊
妹　妹$$冰冰羊 (无血缘关系)
表　哥$$瘦羊羊
```

第 23 行，第 1 列　100%　Windows (CRLF)　UTF-8

A186　喜羊羊

	A	B	C
236	喜羊羊	马莉	朋友
237	喜羊羊	黑大帅	朋友
238	喜羊羊	黑溜溜	朋友
239	喜羊羊	喜羊羊	本名
240	喜羊羊	丽羊羊	母亲
241	喜羊羊	青青草原	活动范围
242	喜羊羊	智羊羊	父亲
243	喜羊羊	羊	物种
244	喜羊羊	羊历3508	生日
245	喜羊羊	Goat（片	英文名
246	喜羊羊	O型	血型
247	喜羊羊	瘦羊羊	表哥
248	喜羊羊	祖晴、杨	配音演员
249	嘭恰恰	奶牛地带	出身地区
250	嘭恰恰	咚咚锵	弟弟

data

就绪　100%

知识抽取：非结构化数据获取实体与三元组

容易，终于又再进入羊村，可又被喜羊羊识破了。

煮煮。此时的众羊才知道灰太狼建造祭坛的真正目的。面对着将要被煮的危机，小羊这 次又会如何面对呢

西哥却被灰太狼给抓了。这可怎么办呢?

救。灰太狼为了抓到小羊们偷偷溜到了地面。小羊们怎么才能解决这些危机呢?

兵，而此时灰入狼又想出坏点子。眼看小羊就要胜利在望，灰人狼的坏上意会得逞吗?

羊非常生气，只好借助道具追赶，这回黑大帅和灰太狼能否逃脱呢? 喜羊羊他又能否成 功捉住他们呢?

发生什么事呢?

命名实体识别：BIOE标注方法

灰太狼用变大丸与变小丸，几经周折才成功穿越羊村闸门，逮住懒羊羊。聪明的懒羊 羊用激将法令得意洋洋的灰人狼自已
B-Person I-Person E-Person O O O O O O O O O O O O O O O O O O O O O O O O O B-Person I-Person E-Perso

灰太狼发现了一种昏睡果，便将其磨成汁，制出昏睡枪，灰太狼试验于红太狼，红太 狼睡着。灰太狼于是想到用昏睡果捉
B-Person I-Person E-Person O O O O O O O O O O O O O O O O O O O O O O B-Person I-Person E-Person O O

为纪念慢羊羊村长功积，众羊合力造了一个蜡像，前往蜡像途中经狼家附近。灰太狼 像变色龙一样，随周围的草、树而改
O O O B-Person I-Person E-Person O O O O O O O O O O O O O O O O O O O O O O O O O O O O B-Person I-P

羊村来了只专做坏事的“喜羊羊"。真的喜羊羊成了代罪羊，被逼离村，不慎落入狼窝。 喜羊羊以激将法，让灰太狼复制出
O O O O O O O O O O O B-Person I-Person E-Person O O O O B-Person I-Person E-Person O O O O O O O O O

沸羊羊钓鱼时钓上了潜于水底的灰入狼。因Z前多次谋吹警报哨，这次人家不信他， 等发现沸羊羊不见，他已快被下狼锅
B-Person I-Person E-Person O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O

隔岸观羊，令灰太狼夫妇眼馋馋的。终于机会来了，原来羊村要举行运动会。灰太狼 乔装成羊，骗过懒羊羊混进羊村，参
O O O O O O B-Person I-Person E-Person O O O O O O O O O O O O O O O O O O O O O O O O O B-Person I-P

灰太狼混进羊村播种迷糊草，并设计令懒羊羊吃下拌有迷糊草的沙律，岂料懒羊羊因 进食了过量的迷糊草而狂性大发，令
B-Person I-Person E-Person O O O O O O O O O O O O O O B-Person I-Person E-Person O O O O O O O O O O

灰太狼从广告中得知一种神奇的魔笛，听到笛声的动物都跟着走。但是买下了魔笛的 灰太狼并不知道笛子根本没有广告中
B-Person I-Person E-Person O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O B-I

	Precision	Recall	F1-Score
LSTM	0.542	0.5013	0.5209
BI-LSTM(I)	0.7447	0.6954	0.7192
BI-LSTM+CRF	0.7734	0.7634	0.7684

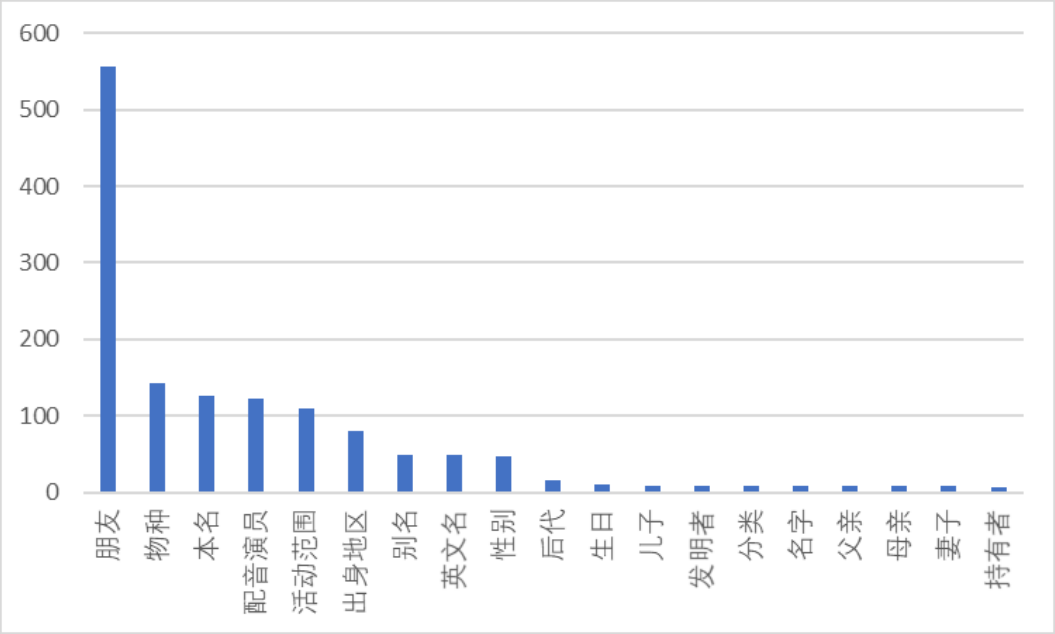
opennre

实体关系抽取

Model	ACC
wiki80_cnn_softmax	0.516
wiki80_bert_softmax	0.738
tacred_bert_softmax	0.619
tacred_bertentity_softmax	0.636
wiki80_bertentity_softmax	0.786
chinese-bert-wwm	0.919

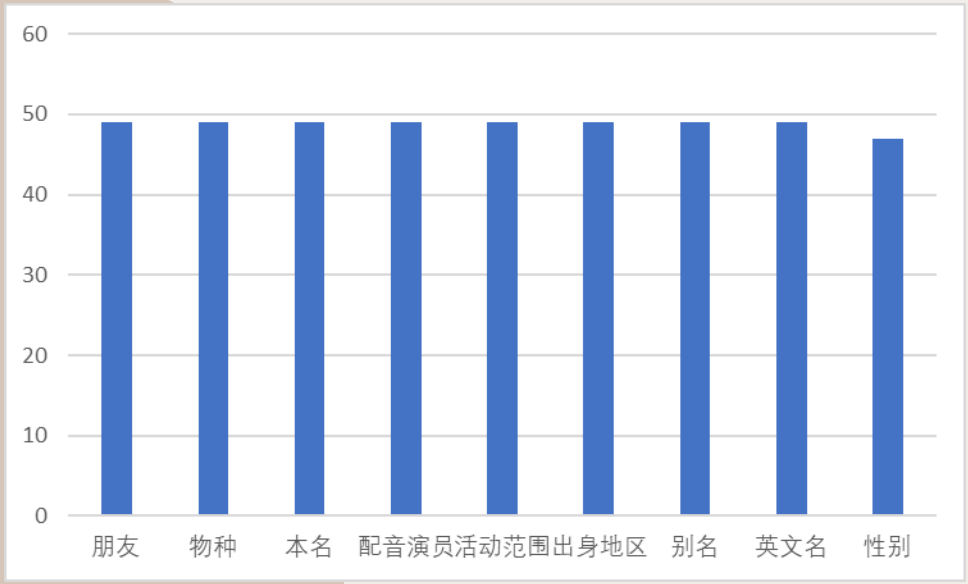
wiki80_cnn_softmax	沸羊羊	喜羊羊	said to be same
	美羊羊	懒羊羊	
	懒羊羊	喜羊羊	
tacred_bert_softmax	沸羊羊	喜羊羊	NA
	美羊羊	懒羊羊	
	懒羊羊	喜羊羊	
chinese-bert-wwm	灰太狼	懒羊羊	夫妻
	灰太狼	潇洒哥	
	灰太狼	喜羊羊	

DeepKE 实体关系抽取



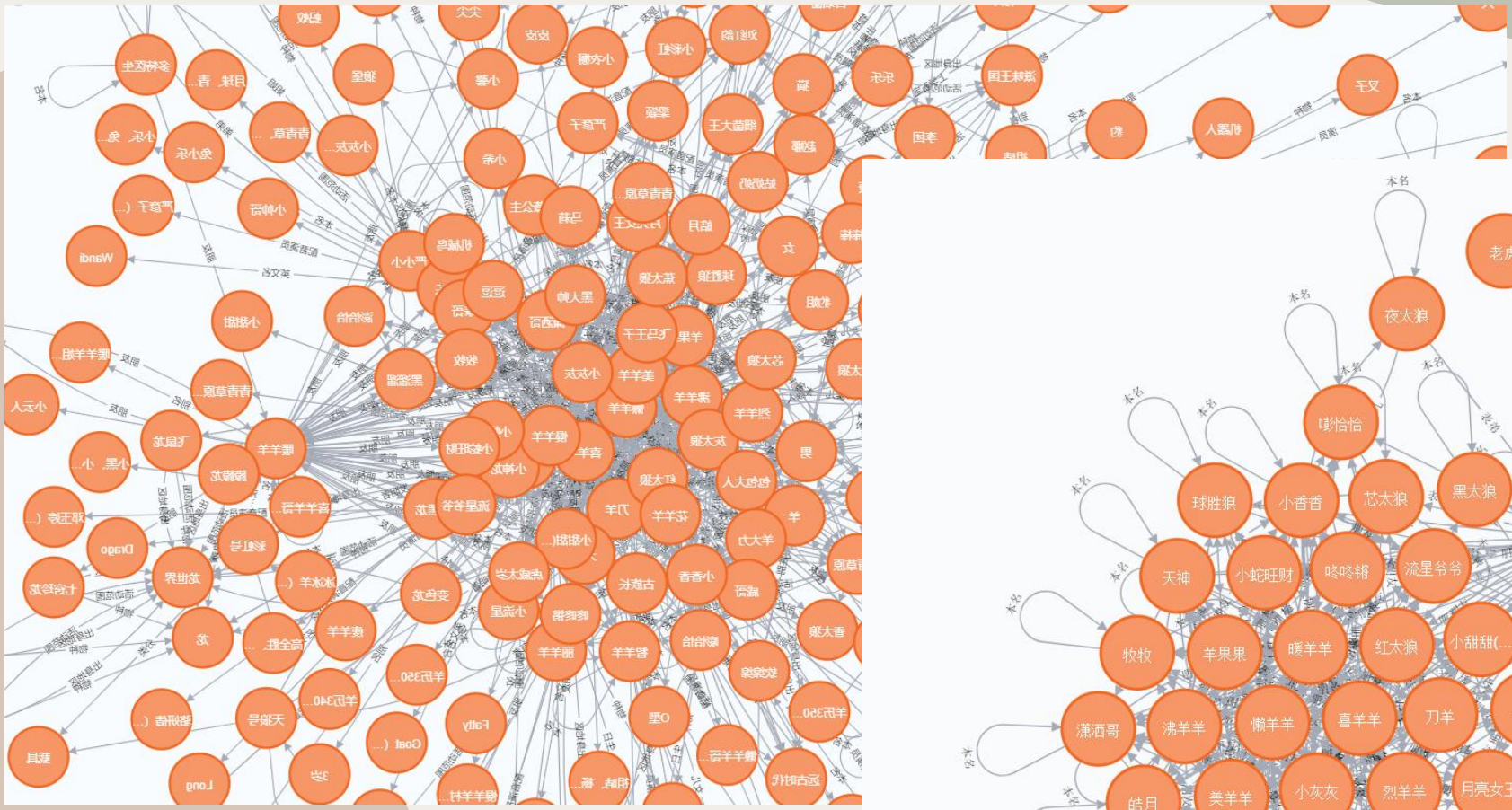
	PCNN	RNN	GCN	TRANSFORMER	BERT
验证集	80.11	83.87	55.91	82.26	**90.86**
训练集	**86.23**	85.64	63.15	**86.12**	**91.33**

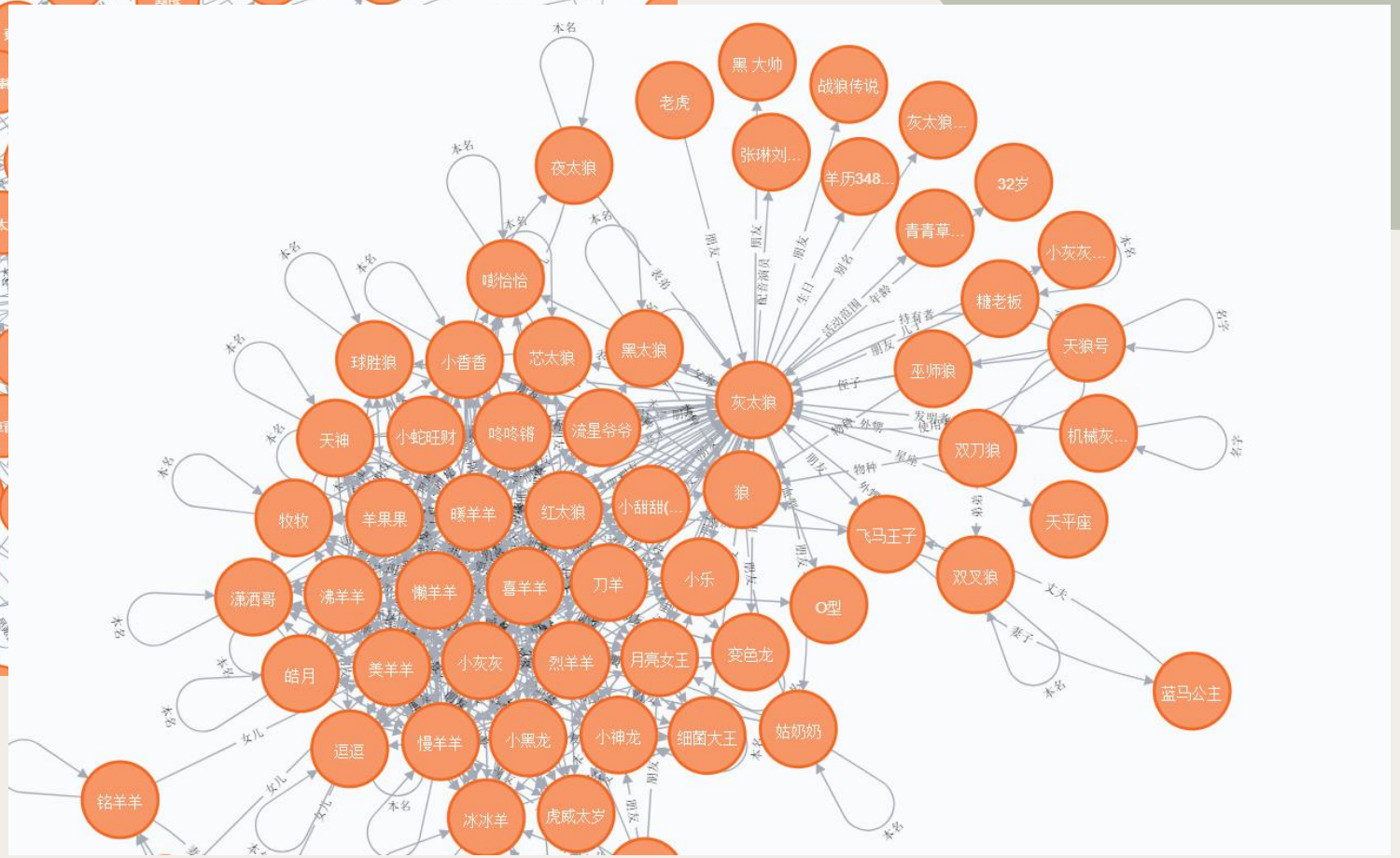
实体关系抽取



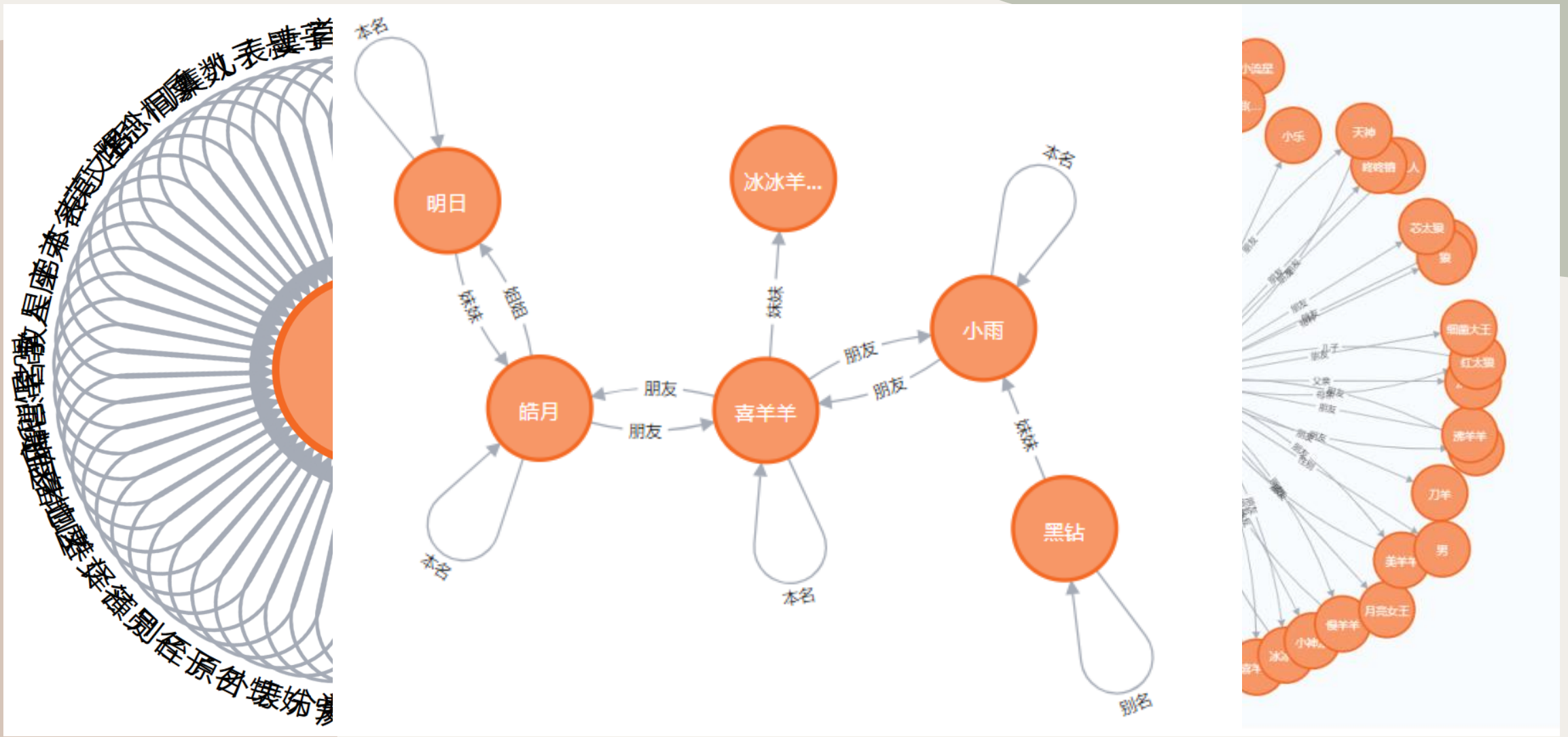
	PCNN	TRANSFORMER	PCNN	TRANSFORMER
验证集	80.11	82.26	77.1	78.56
测试集	86.18	86.18	80.61	82.39

知识存储:Neo4j



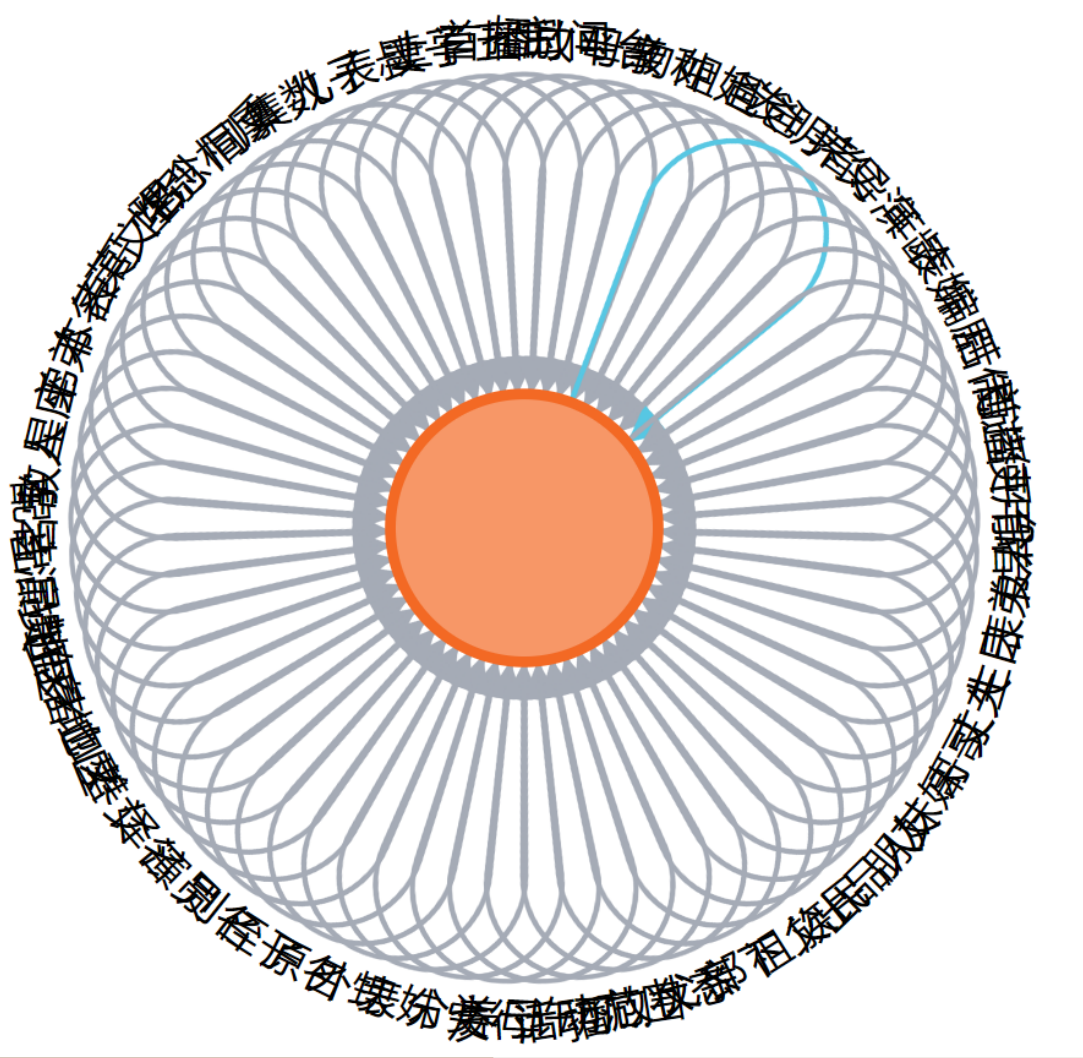

本名
黑 大帅
老虎
战狼传说
灰太狼...
张琳刘...
羊历348...
32岁
夜太狼
青青草...
小灰灰...
糖老板
哨怡怡
球胜狼
小香香
芯太狼
黑太狼
天狼号
巫师狼
灰太狼
天神
小蛇旺财
哈哈锵
流星爷爷
双刀狼
机械灰...
狼
牧牧
羊果果
暖羊羊
红太狼
小甜甜(...
飞马王子
天平座
双叉狼
潇洒哥
沸羊羊
懒羊羊
喜羊羊
刀羊
小乐
O型
皓月
美羊羊
小灰灰
烈羊羊
月亮女王
变色龙
蓝马公主
逗逗
慢羊羊
小黑龙
小神龙
细菌大王
姑奶奶
铭羊羊
冰冰羊
虎威太岁

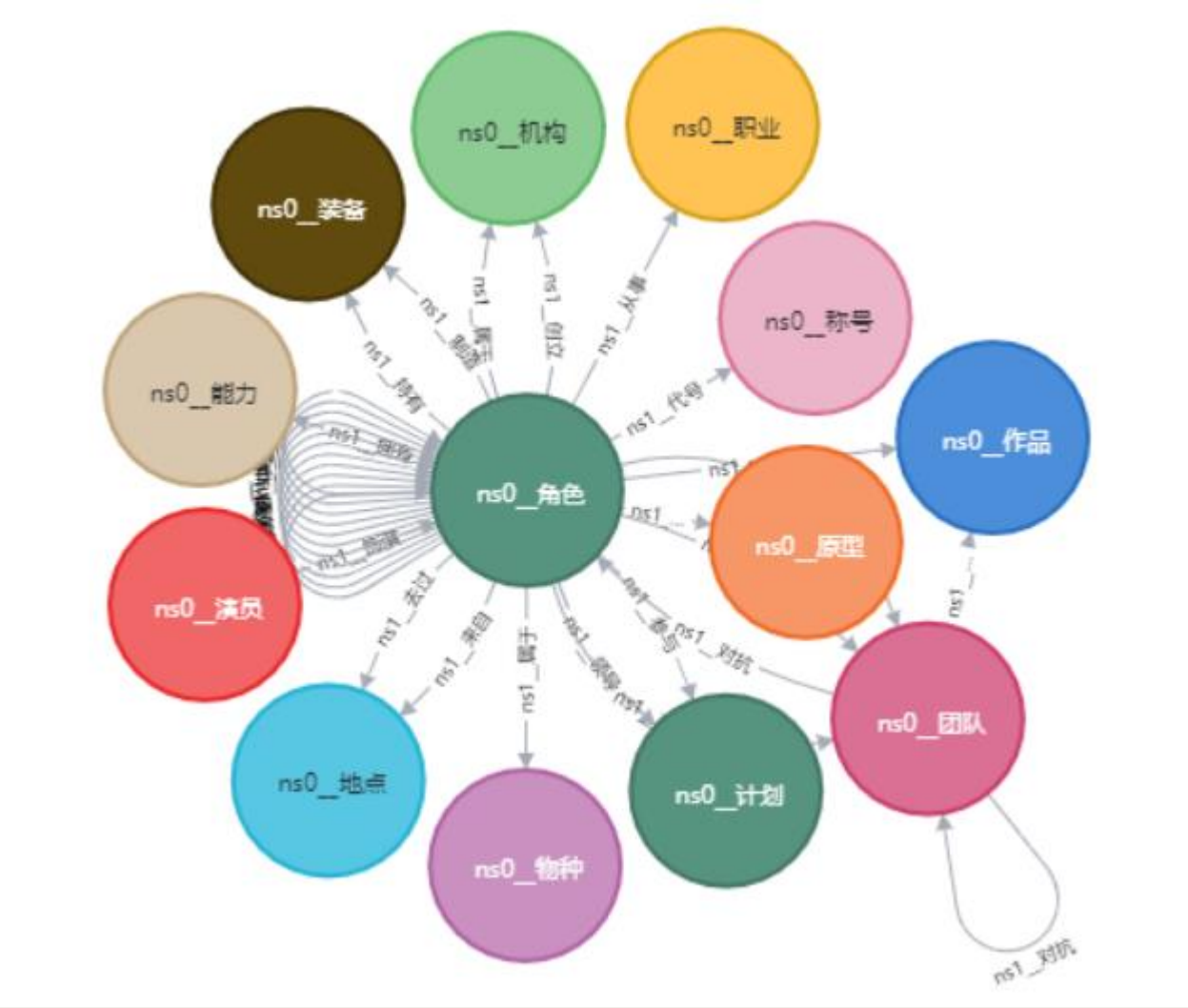
知识存储:Neo4j


本名
明日
妹妹
姐姐
冰冰羊...
妹妹
本名
小雨
朋友
朋友
皓月
朋友
喜羊羊
朋友
妹妹
黑钻
本名
本名
别名
小乐
天神
芝太狼
细菌大王
红太狼
沸羊羊
刀羊
美羊羊
男
月亮女王
慢羊羊

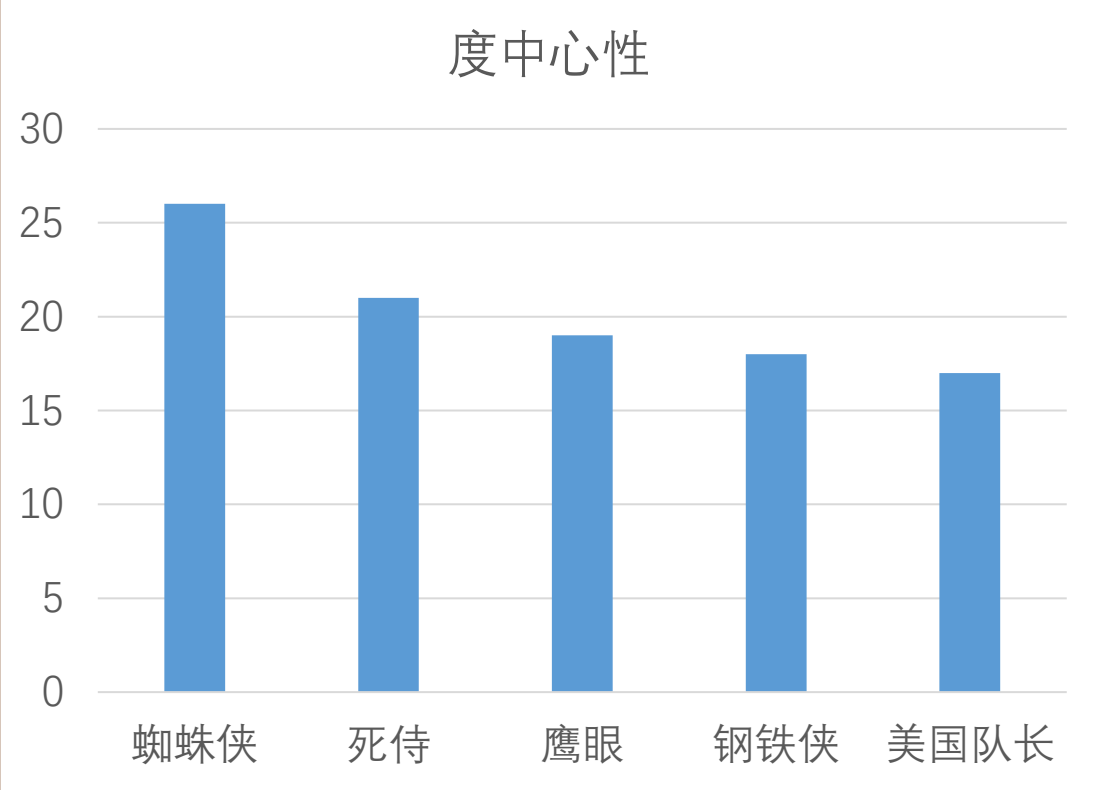
二、知识图谱的应用

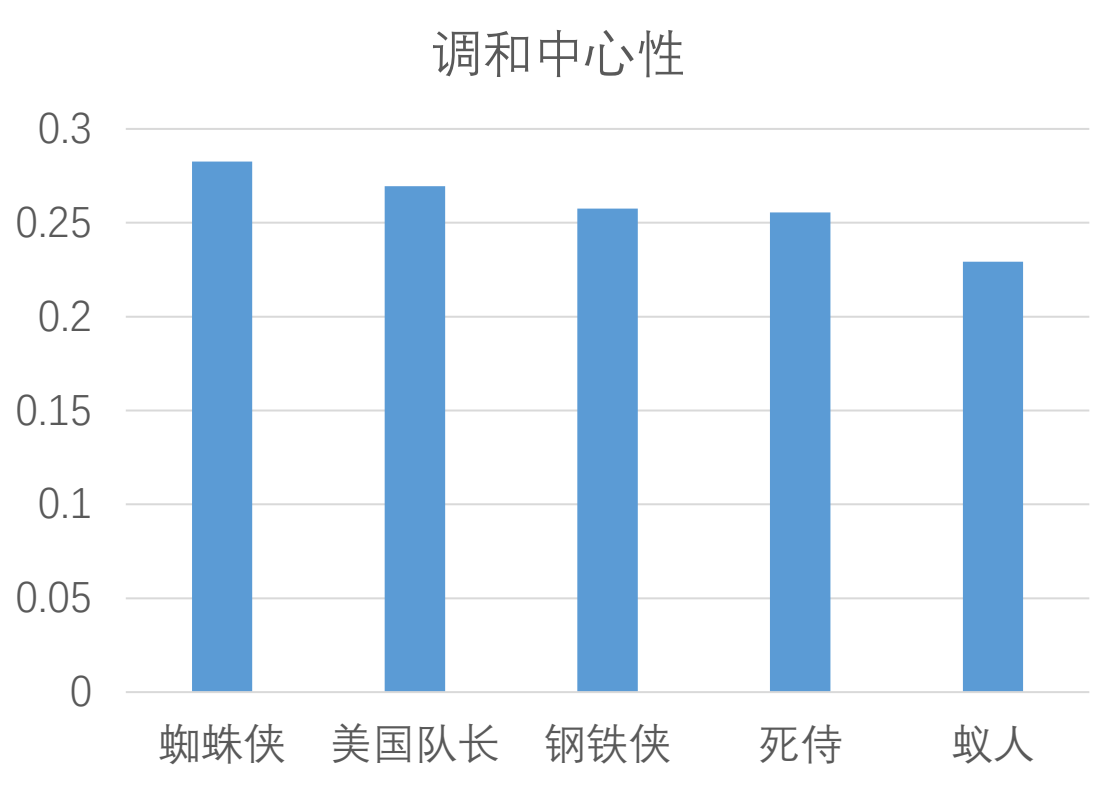
本体层建模




ns0__机构
ns0__职业
ns0__装备
ns0__称号
ns0__能力
ns0__作品
ns0__角色
ns0__原型
ns0__演员
ns0__团队
ns0__地点
ns0__计划
ns0__物种
ns1__对抗

角色网络分析：节点中心度分析





角色网络分析：社区检测

标签传播算法

知识问答

```
回答：  青青草原的小灰灰的爸爸是狼堡的灰太狼

用户：  小灰灰的妈妈是？
回答：  青青草原的小灰灰的爸爸是狼堡的红太狼

用户：  红太狼的妈妈是？
回答：  狼堡的的红太狼的妈妈是狼堡的粉红太狼

用户：  喜羊羊的爸爸是？
回答：  青青草原的的喜羊羊的爸爸是青青草原的智羊羊

用户：  喜羊羊的妈妈是？
回答：  青青草原的的喜羊羊的爸爸是青青草原的丽羊羊

用户：  智羊羊的妻子是？
回答：  青青草原的的智羊羊的妻子是青青草原的丽羊羊
```

```
========== RESTART: C:/Users/
灰太狼的爸爸是？
狼堡的灰太狼的爸爸是狼堡的黑太狼？
>>>
```

```
DeprecationWarning: invalid escape sequence \W
  sentence = re.sub('\W+', '', sentence).replace("_", '')
Paddle enabled successfully......
DEBUG:jieba._compat:Paddle enabled successfully......
match(p)-[r0:父亲]->(n:Person{Name:'灰太狼'}) return
p.Name,n.Name,p.cate,n.cate
 p.Name | n.Name | p.cate | n.cate
--------|--------|--------|--------
 小灰灰     | 灰太狼     | 青青草原    | 青青草原

青青草原的灰太狼的父亲是青青草原的小灰灰
```