

从人工智能到开放知识图谱

王昊奋

2018.7.6 南京



OpenKG.CN

开放的知识图谱社区

大纲

- 从人工智能看知识图谱
 - 知识图谱技术及应用概览
 - 开放知识图谱
-

大纲

- 从人工智能看知识图谱
 - 知识图谱技术及应用概览
 - 开放知识图谱
-

从AI顶级会议看KG的发展

- IJCAI2017 Best Paper
- 描述性数据分析: Declarative Data Analysis
- “It has recently been argued that data analysis should be declarative: users should describe what the desired output is, rather than how to compute it.”

Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17)

Foundations of Declarative Data Analysis Using Limit Datalog Programs

Mark Kaminski, Bernardo Cuenca Grau, Egor V. Kostylev, Boris Motik and Ian Horrocks

Department of Computer Science, University of Oxford, UK

{mark.kaminski, bernardo.cuenca.grau, egor.kostylev, boris.motik, ian.horrocks}@cs.ox.ac.uk

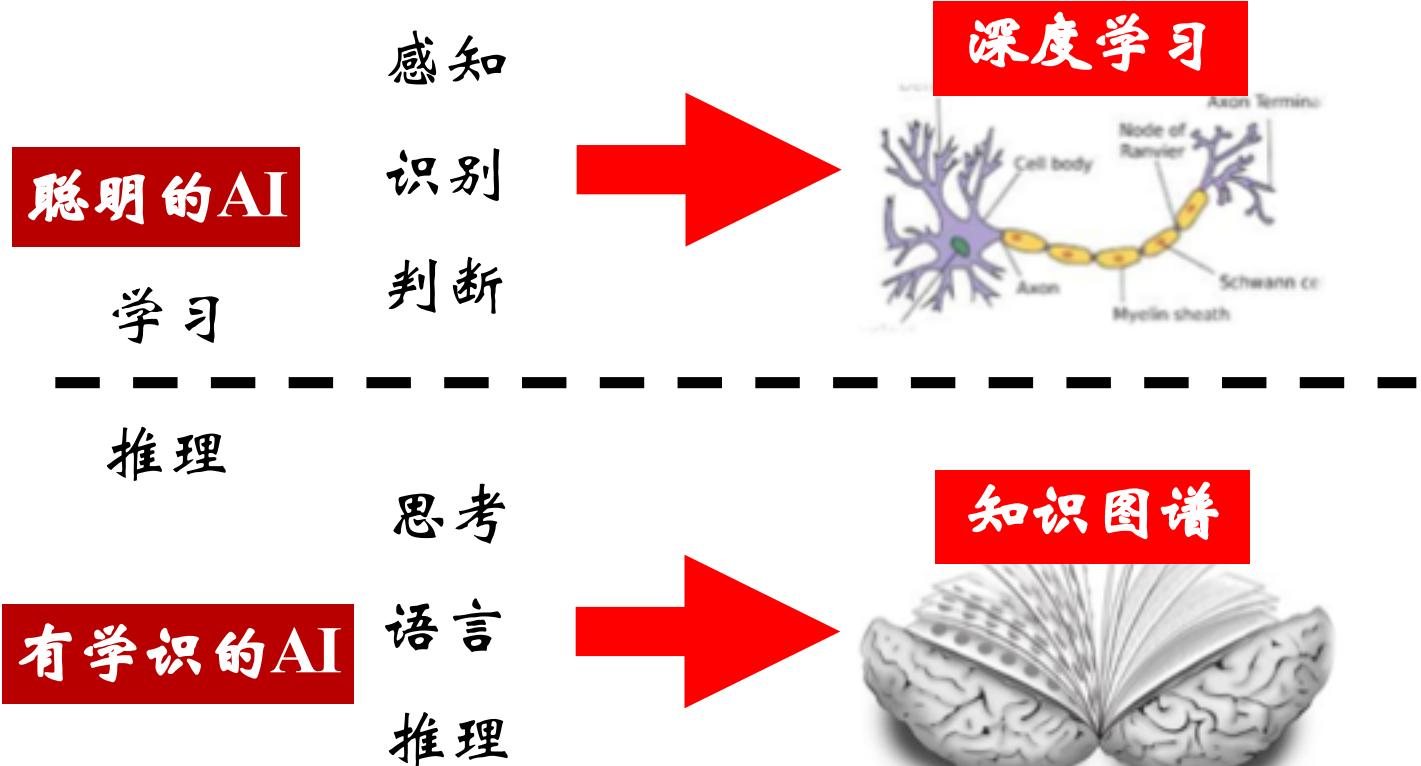
从AI顶级会议看KG的发展

- IJCAI2017 Prominent Papers
- BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network
- YAGO2: A spatially and temporally enhanced knowledge base from Wikipedia

Knowledge is important!

聪明的AI vs 有学识的AI

- 人的大脑依赖所学的知识进行思考、逻辑推理、理解语言



AI的未来

联结

+

符号

学习

+

推理

感知

+

认知

CoreAI : Learning + Reasoning

什么是知识表示

- 简单而言，知识表示（KR）就是用易于计算机处理的方式来描述人脑的知识。

 - KR不是数据格式、不等同于数据结构、也不是编程语言，对于人工智能而言，数据与知识的区别在于KR支持推理。
-

什么是知识表示（续）

- Entity ID, 概念模型, 支持推理, 易于计算, 人可理解

What is a Knowledge Representation?

Randall Davis
MIT AI Lab

Howard Shrobe
MIT AI Lab and Symbolics, Inc.

Peter Szabolts
MIT Lab for Computer Science

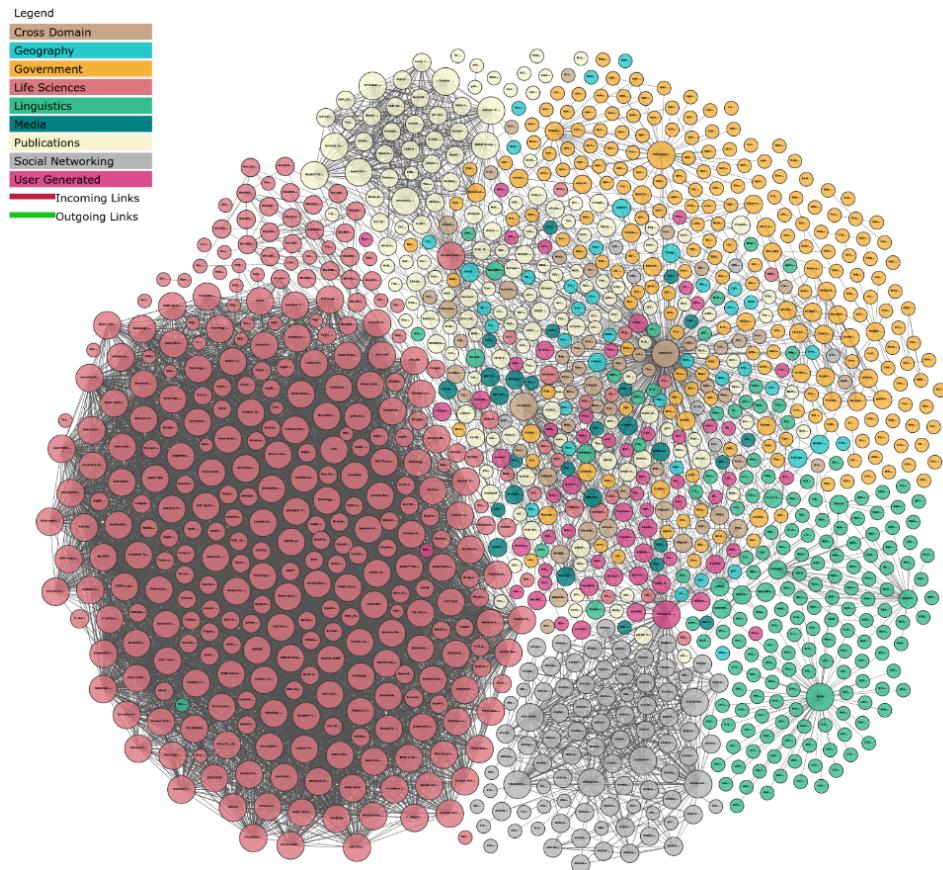
KR的五个用途



什么是知识表示（续）

□ 基于离散符号的知识表示

- RDF, OWL, 各种Rule Language等
- 显式知识、强逻辑约束、易于解释、推理不易扩展

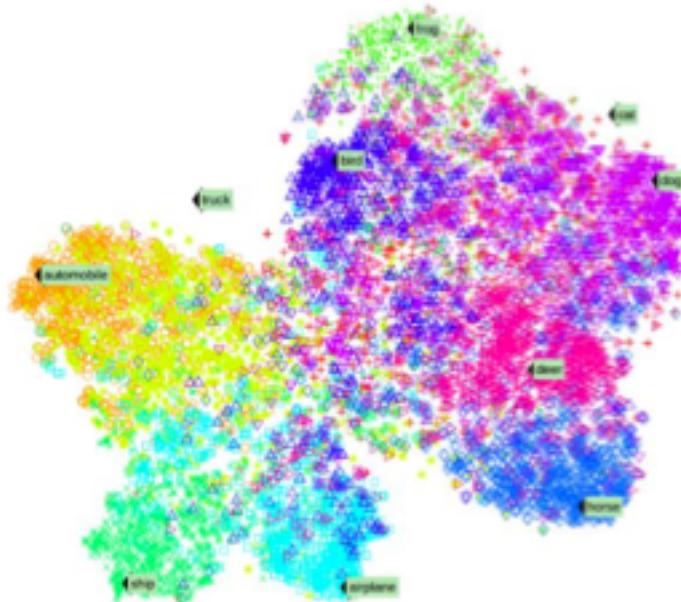


"Linking Open Data cloud diagram 2017, by Andrejs Abele, John P. McCrae, Paul Buitelaar, Anja Jentzsch and Richard Cyganiak. <http://lod-cloud.net/>"

什么是知识表示（续）

□ 基于连续向量的知识表示

- Tensor, 各种Embedding, 神经网络表示等
- 隐式知识、弱逻辑约束、不易解释、对接神经网络



谷歌知识图谱：Things not strings

despicable me 2

Web Images Maps Shopping News More Search tools

About 163,000,000 results (0.29 seconds)

Despicable Me 2 showtimes for San Francisco, CA

See showtimes for 3D

1 hr 38min - Rated PG - Animation
In summer 2013, get ready for more Minion madness in Despicable Me 2. Chris Meledandri and his acclaimed team ...
AMC Van Ness 14 - 1000 Van Ness Avenue, San Francisco, CA - Map
11:25am - 2:05 - 4:55 - 7:40 - 10:30pm
Century San Francisco Centre 9 and IMAX -
835 Market St., San Francisco, CA - Map
7:00 - 9:25pm
+ Show more theaters

Despicable Me 2
despicableme.com/

A short description of the movie, ratings, release date, directors, cast, etc.

★★★★ = Rating: 7.8/10 - 61,214 votes
Directed by Pierre Louis Patrie Coffin, Chris Renaud. With Steve Carell, Kristen Wiig, Benjamin Bratt, Miranda Cosgrove. Gru is recruited by the Anti-Villain ...
Release info: Full cast and crew · Videos · Version 3

Despicable Me 2 - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Despicable_Me_2

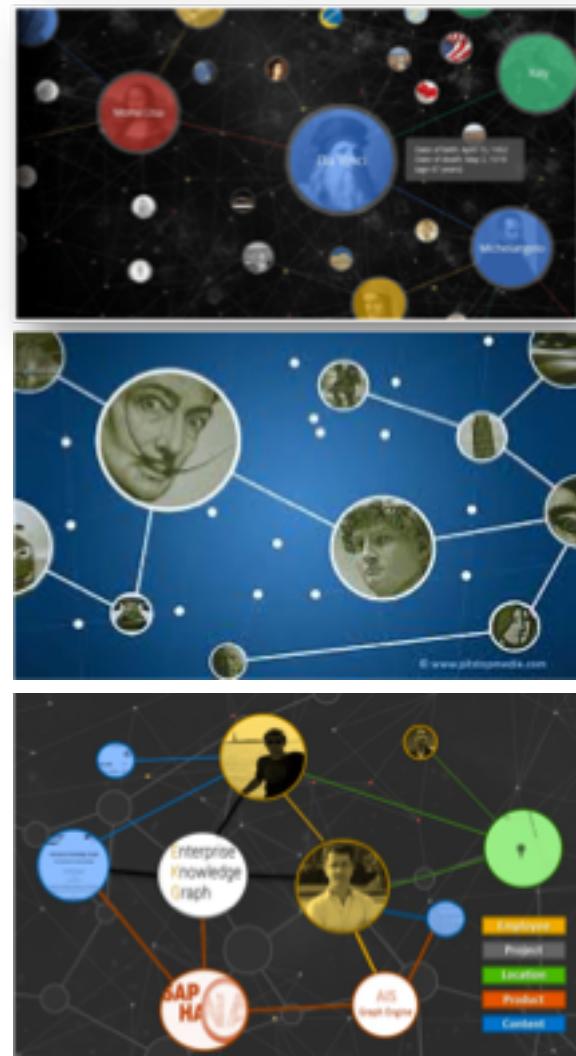
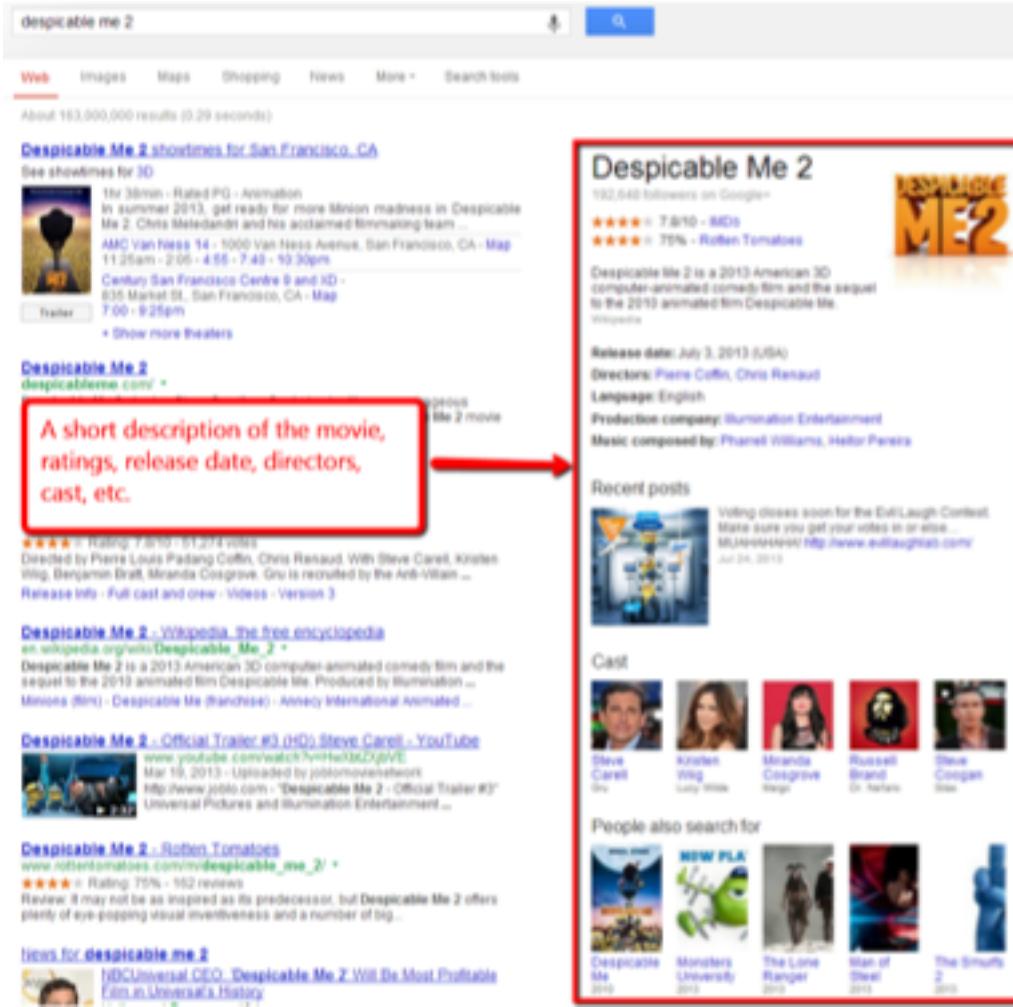
Despicable Me 2 is a 2013 American 3D computer-animated comedy film and the sequel to the 2010 animated film Despicable Me. Produced by Illumination ...
Minions (film) · Despicable Me (franchise) · Amcyc International Animation ...

Despicable Me 2 - Official Trailer #3 (HD) Steve Carell - YouTube
www.youtube.com/watch?v=vtHedzXQgVE
Mar 19, 2013 - Uploaded by johncarmack777
<http://www.joblo.com/~Despicable-Me-2-Official-Trailer-#3>
Universal Pictures and Illumination Entertainment ...

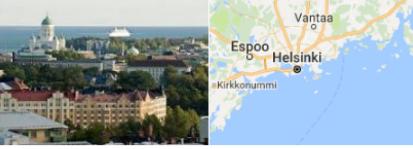
Despicable Me 2 - Rotten Tomatoes
www.rottentomatoes.com/m/despicable_me_2/

★★★★ = Rating: 75% - 162 reviews
Review: It may not be as inspired as its predecessor, but Despicable Me 2 offers plenty of eye-popping visual inventiveness and a number of big ...

News for despicable me 2
[NBCUniversal CEO: Despicable Me 2 Will Be Most Profitable Film in Universal's History](http://nbcuniversal.ceo/Despicable_Me_2_Will_Be_Most_Profitable_Film_in_Unclears_History)



搜索结果中体现关联



Helsinki
Capital of Finland

Helsinki, Finland's southern capital, sits on a peninsula in the Gulf of Finland. Its central avenue, Mannerheimintie, is flanked by institutions including the National Museum, tracing Finnish history from the Stone Age to the present. Also on Mannerheimintie are the imposing Parliament House and Kiasma, a contemporary art museum. Ornate red-brick Uspenski Cathedral overlooks a harbor.



Clash of Clans
Mobile game

4.6/5 Google Play 4.1/5 iTunes - Apple

Clash of Clans is a freemium mobile MMO strategy video game developed and published by Supercell. The game was released for iOS platforms on August 2, 2012, and on Google Play for Android on October 7, 2013. [Wikipedia](#)

Initial release date: August 2, 2012
Developer: Supercell
Publisher: Supercell
Platforms: Android, iOS



Supercell
Game developer

Supercell is a mobile game development company founded in June 2010 in Helsinki, Finland. The company's debut game was the browser game Gunshine.net, and after its release in 2011, Supercell started developing games for mobile devices.

Revenue: 2.11 billion EUR (2015)
CEO: Ilkka Paaninen (2010-)
Founded: June 2010, Helsinki, Finland
Owner: Tencent (84.3%)
Founders: Mikko Kodisjä, Lasse Leppinen, Niko Derome, Visa Forsten, Petri Styman, Ilkka Paaninen

Profiles
 Facebook  Twitter

Video games

 Clash of Clans 2012  Clash Royale 2016  Boom Beach 2013  Hay Day 2012



Ilkka Paaninen

Businessman

Born: 1978

People also search for: Mikko Kodisjä, Niko Derome, Lassi Leppinen, Visa Forsten, Petri Styman

Organizations founded: Supercell, Sumea

Tencent

Internet company



Tencent Holdings Limited is a Chinese investment holding company whose subsidiaries provide media, entertainment, internet and mobile phone value-added services and operate online advertising services in China. [Wikipedia](#)

Stock price: 0700 (HKG) HK\$202.60 -1.80 (-0.88%)
Nov 3, 2:30 PM GMT+8 - Disclaimer

Headquarters: Shenzhen, China

Revenue: 102.9 billion CNY (2015)

Founded: November 11, 1998

CFO: John Lo

Subsidiaries: Riot Games, WeBank, Epic Games, Tencent Pictures

Founders: Ma Huateng, Zhang Zhidong

Facebook兴趣图谱

The diagram illustrates the Facebook interest graph for Harvard University, showing the interconnected nature of user profiles based on shared interests.

Central Node: Harvard University (College & University)

Connections:

- From Harvard University:** A double-headed arrow connects to the "People who like Harvard University" page.
- From Harvard University:** A double-headed arrow connects to the "People who visited Harvard University" page.

People who like Harvard University:

- Paul McDonald:** Likes Harvard University, iRunFar.com and 182 others. Studied Computer Science at Harvard University '03. 3 mutual friends including Clodagh Chloe Takeuchi and Serkan ...
- Ekaterina Skorobogatova:** Likes Harvard University, Loves Company and 3,455 others. Studied Interactive Multimedia at New York University. 2 mutual friends: Amina Belghiti and Alexey Spiridonov
- Gary Johnson:** Likes Harvard University and 278 others. Studied at Wharton School, University of Pennsylvania '08. 5 mutual friends including Jen Holmstrom and Clodagh Chloe T...
- Greg Marra (马格雷):** Likes Harvard University, Emmy's Spaghetti Shack and 693 others. Studied Electrical and Computer Engineering at Franklin W. Olin ...

People who visited Harvard University:

- Florence Trouche:** Global Client Partner at Facebook. Visited Harvard University, Marché Poncelet and 317 other places. Studied at Rouen Business School '90. 35 mutual friends including Michelle Gilbert and Lisa Carucci
- Andrew Tulloch:** Machine Learning at Facebook. Visited Harvard University, City Beer Store and 90 other places. Studied Machine Learning at University of Cambridge. 21 mutual friends including Jason Weston and Nicolas Vasilache
- Joseph Barillari (joeb):** Software Engineer at Facebook. Visited Harvard University, Philz Coffee At Facebook and 990 others. Studied Computer Science at Harvard University '07. 10 mutual friends including Tudor Bosman and Jessica Traynor
- Sheryl Sandberg:** Chief Operating Officer at Facebook. Visited Harvard University and 423 other places. Studied at Harvard Business School. 14 mutual friends including Laurent Solly and Catalina Fries Sa...

Facebook兴趣图谱复杂示例

Structured search within the graph

The screenshot shows a Facebook search results page with a blue header bar containing the search query: "People who like Harvard University and Basketball and work at Facebook". Below the header, there are five search results, each showing a profile picture, name, title, and a brief bio. The profiles belong to Mike Vernal, Jared Morgenstern, Florin Ratu, Ning Zhang, and Zhongyuan Xu. Each profile has a "Add Friend" button and a "View on Facebook" link.

- Mike Vernal**
VP Engineering at Facebook
Likes Harvard University, Harvard Crimson and Facebook. Studied Computer Science at Harvard University '02. 10 mutual friends including Keith Adams and Philip Bohannon.
- Jared Morgenstern**
Product Manager / Ninja - Games ...
Likes Harvard University, Facebook Summer Basketball League. Studied Computer Science at Harvard University. 5 mutual friends including Céadoghan Takeuchi and Pierre ...
- Florin Ratu**
Software Engineer at Facebook
Likes Harvard School of Public Health, Stanford 8th Man and the ... Studied Management Science and Engineering at Stanford Univ ... 3 mutual friends including Alexey Spiridonov and Serkan Piantino
- Ning Zhang (张宁)**
Software Engineer at Facebook
Likes Harvard University, Basketball and 314 others ... Studied Computer Science at University of Waterloo '06. 6 mutual friends including Taylor Boeman and Yev Blayevoy
- Zhongyuan Xu (徐重远)**
Software Engineer at Facebook
Likes Harvard University, Basketball and 364 others ... Studied at Stony Brook University. 1 mutual friend: Ledell Wu.

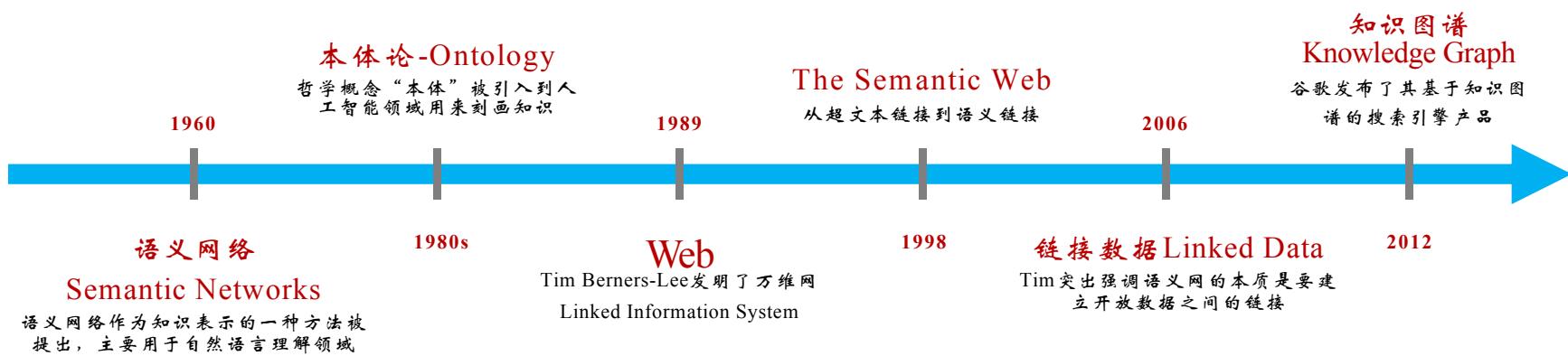
Introducing Graph Search

Q People who like Cycling and are from my hometown

The screenshot shows the official Facebook Graph Search landing page with a dark blue header and a large white title "Introducing Graph Search". Below the title is a search bar with the query "People who like Cycling and are from my hometown". The results are displayed in a grid of user profiles, each with a thumbnail, name, title, and a "View on Facebook" link. The profiles include Sharon Hwang, Morin Oluwole, Russ Maschmeyer, Peter Jordan, and Anish Bhasin.

| Name | Title | Location | Relationship |
|-----------------|--|--|------------------------------|
| Sharon Hwang | Product Designer at Facebook | All Lives in San Francisco, California | Relationship with Mike Matas |
| Morin Oluwole | Business Lead to VP, Global Marketing So... | All 13 mutual friends including Matt Brown | |
| Russ Maschmeyer | Interaction & User Experience Designer at... | | |
| Peter Jordan | Film Producer at Facebook | | |
| Anish Bhasin | Graphic Designer at... | | |

知识图谱的历史起源



知识表示与知识库- Knowledge Representation / Knowledge Base

人工智能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。

知识图谱得益于Web的发展（更多的是数据层面），有着来源于KR、NLP、Web、AI多个方面的基因。

KG的本质

Web视角

像建立文本之间的超链接一样，建立数据之间的语义链接，并支持语义搜索

NLP视角

怎样从文本中抽取语义和结构化数据

KR视角

怎样利用计算机符号来表示和处理知识

AI视角

怎样利用知识库来辅助理解人的语言

DB视角

用图的方式去存储知识

做好KG要兼容并蓄，综合利用好KR、NLP、
Web、ML、DB等多方面的办法和技术

大纲

- 从人工智能看知识图谱
 - 知识图谱技术及应用概览
 - 开放知识图谱
-

KG辅助搜索：KG4SEO

- Web的理想是万物的链接，搜索的理想是事物的搜索



 Freebase

社区协同构建



维基众包



网页嵌入语义数据

KG辅助问答: KG4BOTS

□ 机器人及IoT设备的智能化：给万物都挂接一个背景知识库

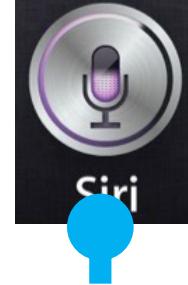
对话式的信息获取更加需要精准度和可靠度，知识图谱对于提升用户体验更加不可少。



DBpedia
Wordnet



True Knowledge/Evi



WolframAlpha
DBpedia



zhishi.me

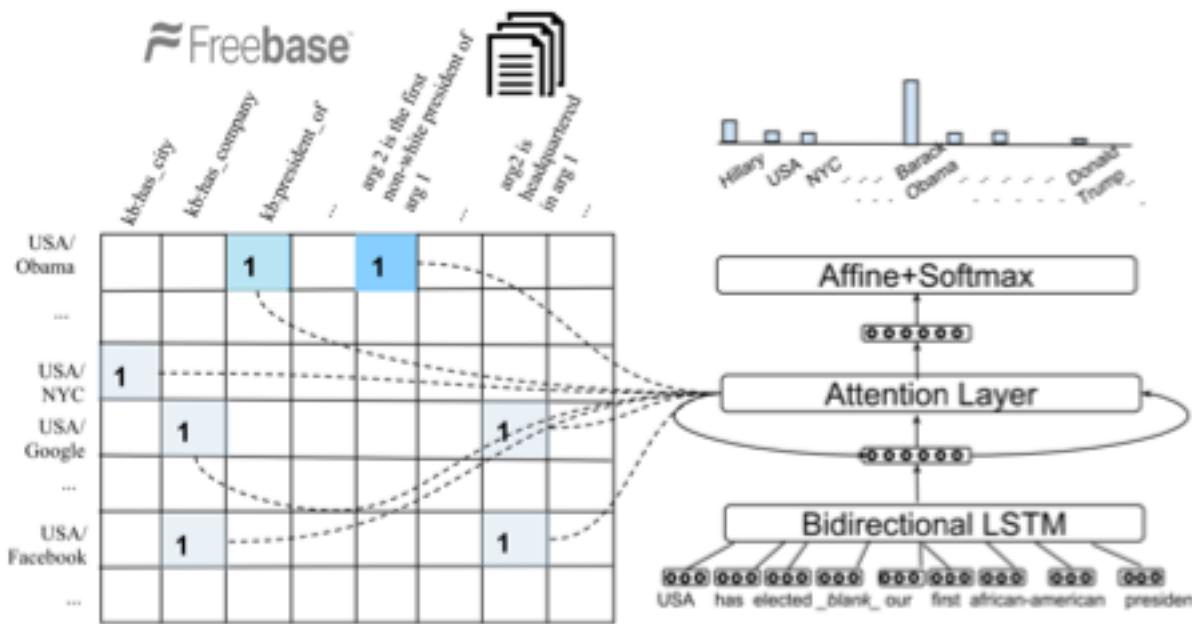


Conversational AI OS for everyone and everywhere



KG辅助问答: KBQA新的发展

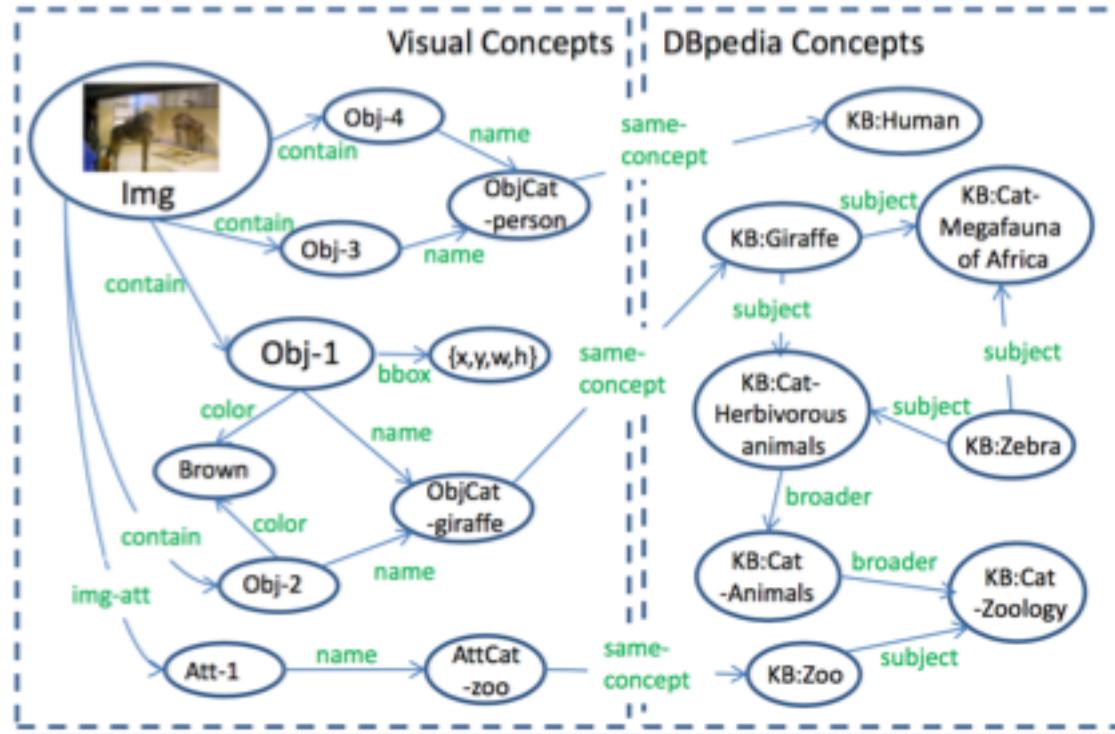
- 综合使用文本 + KB + 深度神经网络来实现问答



- Das, R., Zaheer, M., Reddy, S., & McCallum, A. (2017). Question Answering on Knowledge Bases and Text using Universal Schema and Memory Networks. ACL2017

KG辅助问答: Visual QA

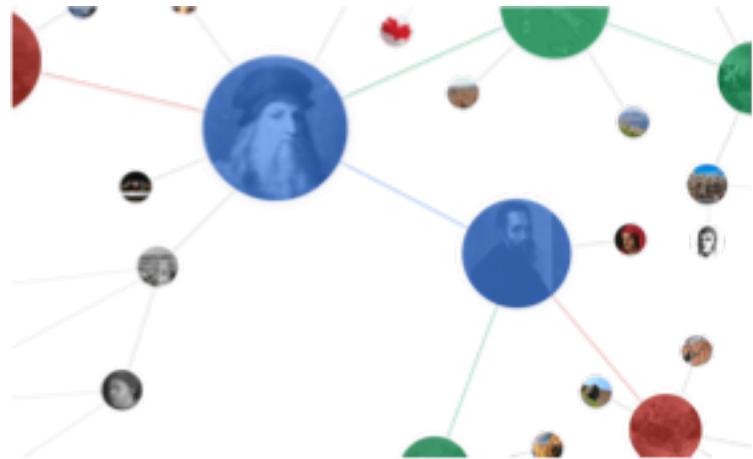
□ 利用KB来做Visual QA



- Wang, P., Wu, Q., Shen, C., Dick, A., & Hengel, A. Van Den. (2017). Explicit Knowledge-based Reasoning for Visual Question Answering, IJCAI2017

KG辅助决策

MORE MACHINE UNDERSTANDABLE



建立数据链接

预先抽取语义



文本



多媒体



传感器

大数据与可计算大数据

...我们现在触手可及的文本型健康医疗数据很多，能够获取的数据规模也都很大，但这是否就意味着这些数据可以供研究机构或商业机构做大数据分析？

.....要把自然语言的描述转换为**计算机可以计算的数据**，需要依赖自然语言处理中的很多基础技术，比如句子的**分词**，**实体识别**，**实体的归一化**和**链接等**

.....可以把**符号推理与统计学习结合起来**，让碎片化知识图谱的知识推理和深度学习的决策模型结合起来，我们称之为**Local Knowledge Powered Global Learning...**

KG辅助决策：预测性分析



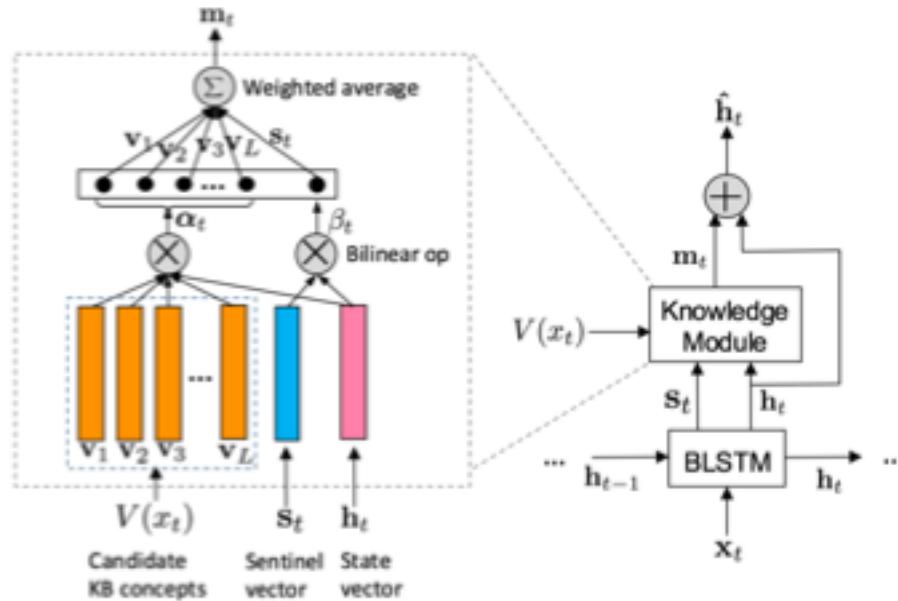
PAIANTIR

KENSHO



KG辅助语言理解：Machine Reading

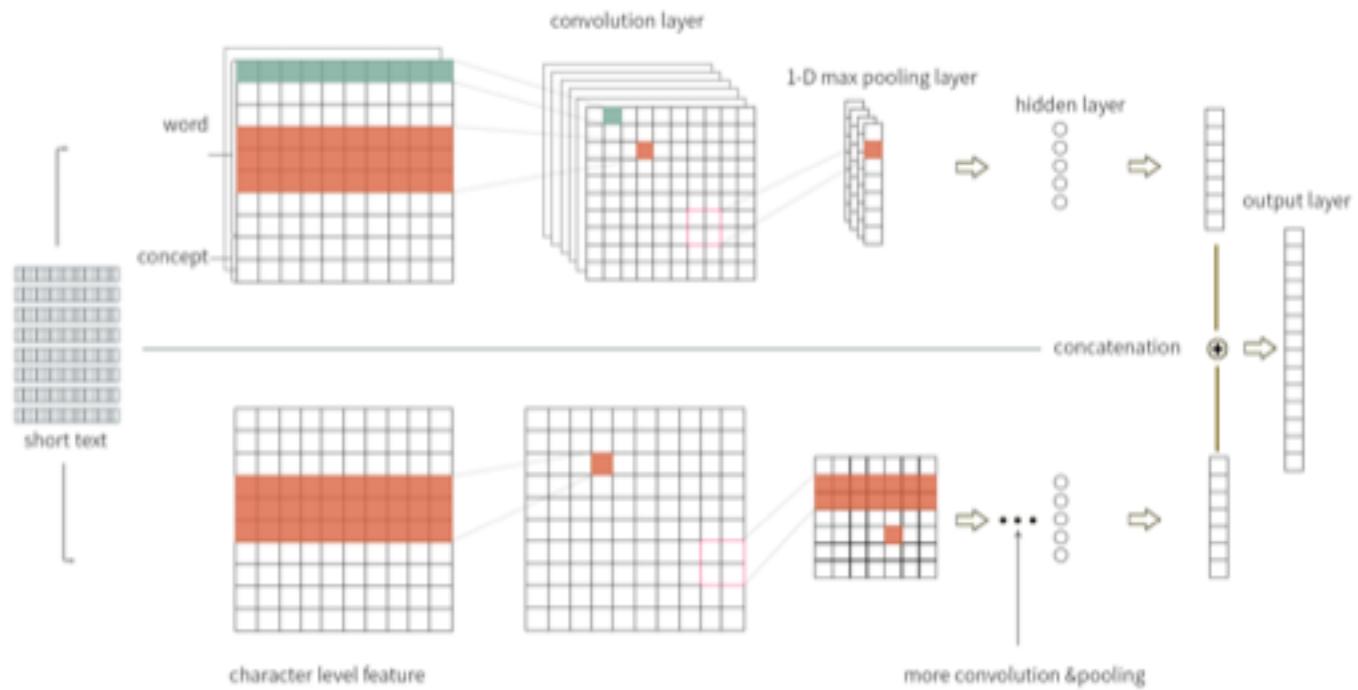
- 利用KB辅助深度神经网络提升实体抽取、事件抽取等机器阅读任务的效果。



- Yang, B., & Mitchell, T. M. (2017). Leveraging Knowledge Bases in LSTMs for Improving Machine Reading. *Association for Computational Linguistics*, 1436–1446.

KG辅助语言理解：Machine Reading

- 把概念性知识作为神经网络的输入，提升短文本分类的效果



- Wang, J., Wang, Z., Zhang, D., & Yan, J. (2017). Combining Knowledge with Deep Convolutional Neural Networks for Short Text Classification. IJCAI2017

KG辅助AI：常识推理

当一个人听到或看到一句话的时候，他使用自己所有的知识和智能去理解。这不仅包括语法，也包括他的词汇知识、上下文知识，更重要的，是对相关事务的理解。

—Terry Winograd (自然语言理解系统SHRDLU作者)

Winograd Schema Challenge

I. The trophy would not fit in the brown suitcase because **it** was too **big (small)**.

What was too **big (small)**?

Answer 0: the trophy Answer 1: the suitcase

II. The town councilors refused to give the demonstrators a permit because **they feared (advocated) violence**. Who **feared (advocated) violence**?

Answer 0: the town councilors Answer 1: the demonstrators

纯NLP: 50%



NLP+KB: 超过60%



及格线: 90%



KG辅助AI：未知的难题

- 人的符号记忆是连续的么?
 - ✓ Hinton：“...智能的精华是怎样实现推理，推理的关键是形式化逻辑...大脑中并没有符号和图像的存储，而是一些Big Activity Vectors之间相互发生作用。”
 - ✓ 知识的表示是否需要向量化和连续化？
 - 人的符号记忆是结构化的么?
 - ✓ 大脑分区域管理不同的文字语义，一个区域分管某个语义区域或一组相关的概念。
 - ✓ 大脑中知识组织的结构是文本化还是结构化？
 - 符号记忆的获取和推理过程又是什么?
 - ✓ 认知科学家道格拉斯·霍夫施塔特：“记忆，是高度重建的。在记忆中进行搜索，需要从数目庞大的事件中挑选出什么是重要的，什么是不重要的，强调重要的东西，忽略不重要的东西。这种选择过程实际上就是感知、认知的核心过程与感知的关系非常非常的密切。”
 - ✓ 符号推理的过程是符号搜索还是向量计算？
-

知识图谱工程



大纲

- 从人工智能看知识图谱
 - 知识图谱技术及应用概览
 - 开放知识图谱
-

各种知识图谱项目

Freebase

LINKINGOPENDATA
W3C SWEO Community Project

WordNet
A lexical database for English

ZHISHI^{me}

PKUBASE

NELL

schema.org
....the new SEO?

DBpedia

XLORE

WIKIDATA

WEB CHILD

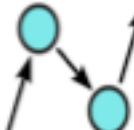
CN-DBpedia

ConceptNet

Herbnet

yAGO
select knowledge

LinkedGeoData.org

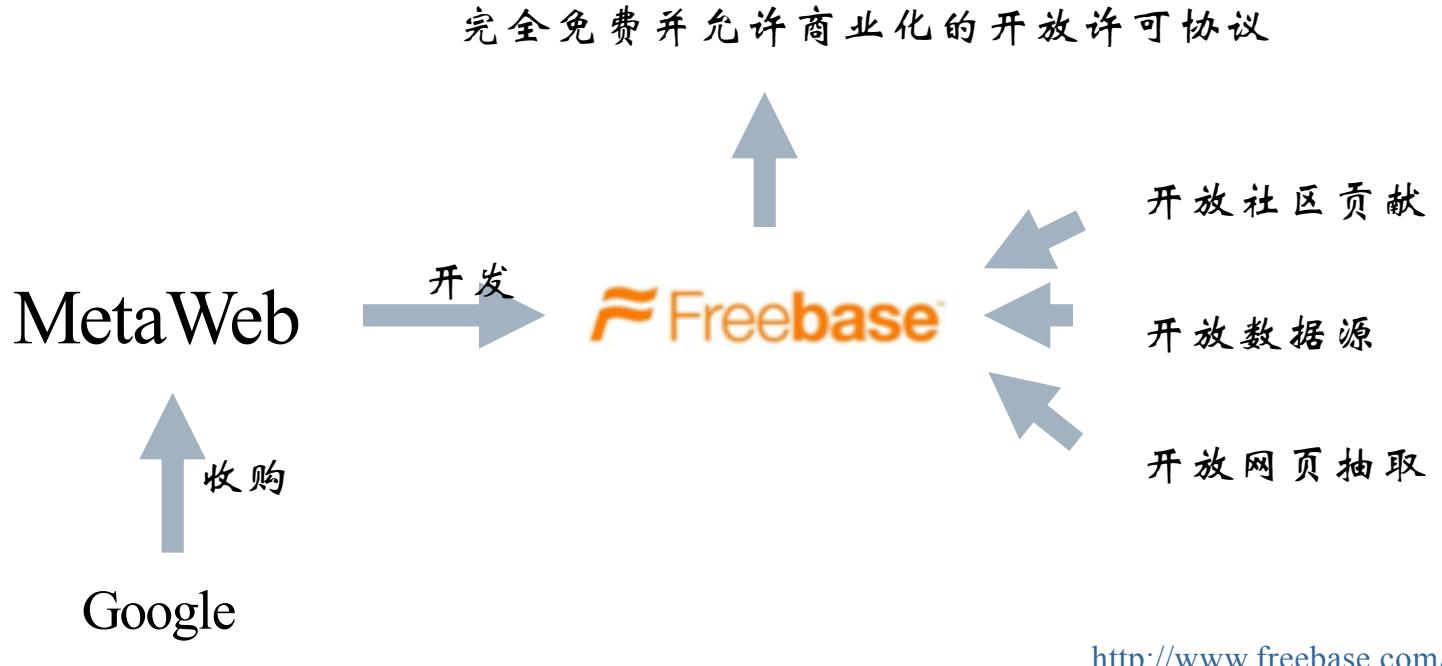


An open, multilingual knowledge graph

linked life data

WEBKB

KG 构建与获取: Freebase



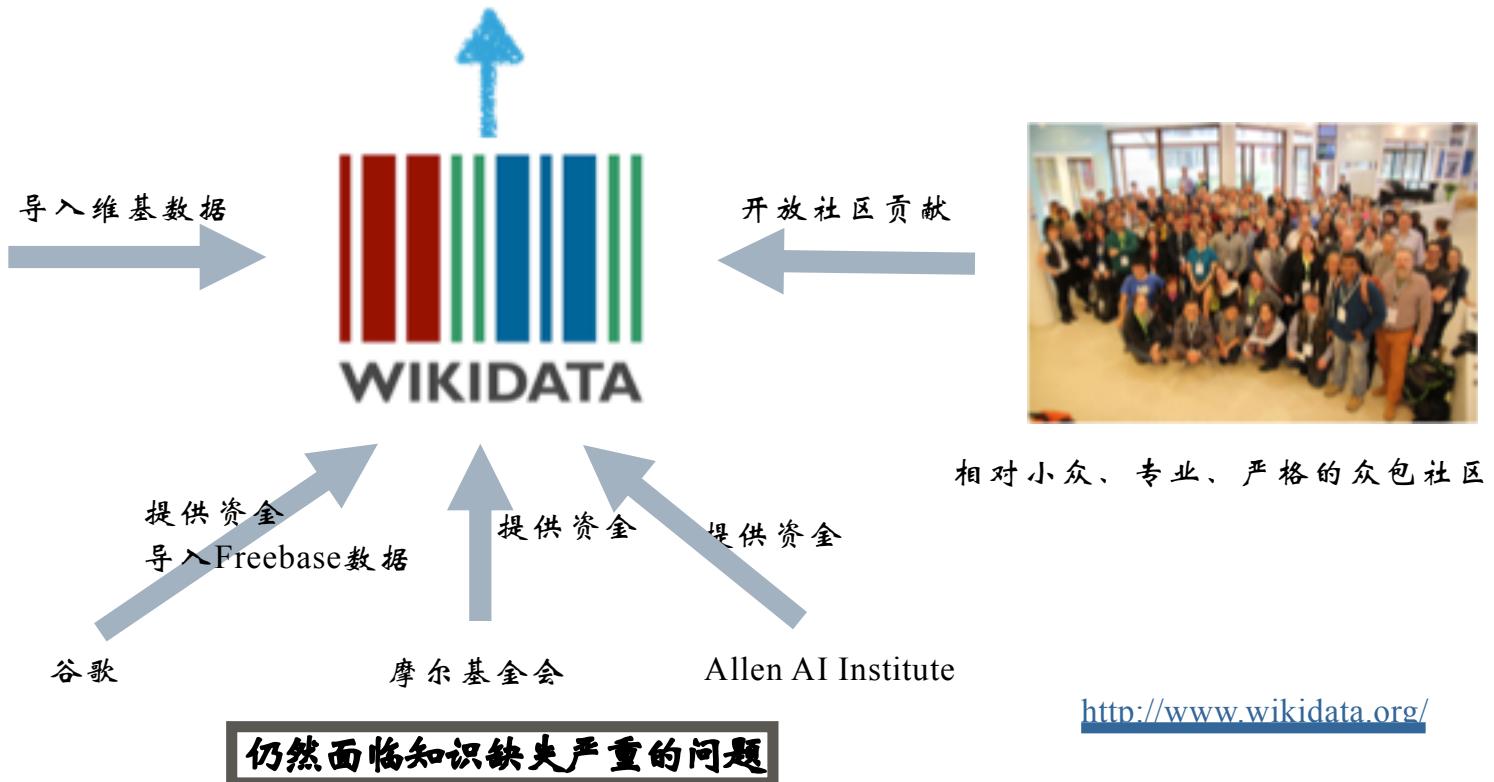
通过开源免费吸引用户贡献数据，增值的应用及技术服务收费

KG 构建与获取: WikiData

目标是构建全世界最大的免费知识库，采用CC0完全自由许可协议



维基基金会



KG 构建与获取: Schema.Org

Semantic Markup: 在网页、邮件、应用程序中嵌入语义数据

```
<p>Big (1988) is a fantasy comedy movie starting Tom Hanks and directed by Penny Marshall.</p>
```

RDFa, JSON-LD, HTML5 Microdata

```
<div itemscope itemtype="http://schema.org/Movie"><p>
  <span itemprop="name">Big </span> (1988) is a
  <span itemprop="genre">fantasy comedy</span> movie starting
  <span itemprop="actor" itemscope itemtype="http://schema.org/Person">
    <span itemprop="name">Tom Hanks</span></span> and directed by
    <span itemprop="director" itemscope itemtype="http://schema.org/person">
      <span itemprop="name">Penny Marshall</span></span>.</p></div>
```

搜索引擎优化

只针对开放域

已经包含642个类，992个关系



截止目前，全球约有1.2亿的网站，超过30%的网页已经嵌入有Schema.org的语义数据。谷歌2015年12月推出的Customizable Knowledge Graph，允许各网站基于Schema.org，以RDFa, JSON-LD, HTML Microdata等方式在网页、邮件等数据源中嵌入语义化数据，支持个人和企业定制自己的知识图谱信息。

KG 构建与获取: ConceptNet

起源于MIT Marvin Minsky 倡导的 Open Mind Common Sense 项目

采用CC BY SA 开放许可协议



LUMINOSO

早期版本主要来源于专家录入、众包和游戏。

新版本导入大量开放的结构化数据，如
DBpedia、Wikinary、Wordnet等

关于KG的几个事实

- 现有的大规模知识图谱都大量的依赖开放域的数据抽取和开放社区协作的方式构建

Freebase Yago DBpedia Wikidata ConceptNet ...

Zhishi.me CN-DBpedia X-Lore PKU-PIE



关于KG的几个事实

- 不能指望NLP和机器学习能自动的去构建你真正需要的知识图谱。但KG的确需要大量的使用NLP/ML的各种技术手段，来帮助冷启动、后期补全和质量提升。

NLP只是获取KG手段之一，文本也只是KG数据的一种来源

关于KG的几个事实

- 构建KG的成本非常高，单一机构无法构建覆盖全面的知识图谱，即便是金融这样的垂直领域也是如此。每个从事KG研发的机构都会也应该去寻求开放社区和外部数据的帮助。

所有KG的共同问题：知识缺失严重！

纯手工众包

内部结构化或文本数据

开放域的结构化/半结构化数据 开放域的文本抽取

关于Open的解读

OpenKG ≠ Open Data

开放域的数据

结构化并链接已经开放的数据



语义表达规范的数据更加容易被利用和重用

开源的数据
类似于开源的软件



开放是为了让更多人发现你，并给你带来收益

开放的数据



开放你所拥有的数据

开源的数据：可商业化的开放数据许可协议

□ 做KG生意绕不过开放的问题，并面临数据版权问题

■ Creative Common

□ CC0：完全自由，例如WikiData

□ CC-BY：署名许可，适应性最强的许可，鼓励商业化，如Yago

□ CC-BY-SA：署名且要求采用相同的许可条款，鼓励商业化，例如：ConceptNet

■ Open Knowledge Foundation

□ PDDL\ODC-By\ODbL

可能的KG商业模式（大佬企业除外）

KG开放或部分开放 + 增值技术服务/优先使用数据/高质量KG服务等

关于 OpenKG

- Open KG DATA：数据以百科知识为核心
 - Open Schema：以通用常识域为主
 - Openbase：人人都可编辑的中文开放知识图谱
 - Open Bots：主要技术为QA Testbeds
 - Open Models：主要技术为Analytics Testbeds
-

OpenKG工作选择性介绍

- OpenKG.CN：开放图谱 + 开放工具 + 开放文献
 - OpenKG.LINK：链接的中文百科知识图谱
 - cnSchema.Org：开放的知识图谱Schema
-

关于OpenKG

□ 开放的资源库



关于OpenKG

□ 知识建模工具：Protege

/ 机构 / OpenKG / 斯坦福本体构建工具Protege

斯坦福本体构建工具
Protege

追从者
0

● 追随

机构

OpenKG.cn
中文开放知识图谱联盟

OpenKG

没有关于此机构的描述

社会

Google+

Traffic

工具 分类 活动流 管理

斯坦福本体构建工具Protege

Protégé软件是斯坦福大学医学院生物信息研究中心基于Java语言开发的本体编辑和知识获取软件，或者说这是本体开发工具，也是基于知识的编辑器，属于开放源代码软件。这个软件主要用于语义网中本体的构建，是语义网中本体构建的核心开发工具，现在的最新版本为5.0.0版本。Protégé提供了本体概念类，关系，属性和实例的构建，并且屏蔽了具体的本体描述语言，用户只需在概念层次上进行领域本体模型的构建。

数据与资源

Protege网址
A free, open-source ontology editor and framework for building intelligent...
浏览

其他信息

| 域 | 价值 |
|------|--------------------------------|
| 状态 | active |
| 最近更新 | 2017年9月9日, 下午4点14分 (UTC+08:00) |
| 创建的 | 2017年9月9日, 下午4点14分 (UTC+08:00) |

关于 OpenKG

□ 知识抽取工具：DeepDive

The screenshot shows a web interface for a knowledge graph platform. At the top, there is a navigation bar with links for '首页' (Home), '机构' (Institution), '浙江大学' (Zhejiang University), and '支持中文的deepdive：斯坦福大学的开源知识抽取工具（三元组抽取）' (Supports Chinese deepdive: Stanford University's open-source knowledge extraction tool (Triplet Extraction)). Below the navigation, there are tabs for '工具' (Tools), '分类' (Categories), and '活动流' (Activity Stream). A '管理' (Management) button is also present.

The main content area features a large title: '支持中文的deepdive：斯坦福大学的开源知识抽取工具（三元组抽取）'. To the left, there is a sidebar with the following information:

- 支持中文的deepdive：
斯坦福大学的开源知识
抽取工具（三元组抽
取）
- 追从者
1
- 追随
- 机构

Below the title, there is a detailed description: 'deepdive是由斯坦福大学InfoLab实验室开发的一个开源知识抽取系统。它通过弱监督学习，从非结构化的文本中抽取结构化的关系数据。本项目修改了自然语言处理的model包，使它支持中文，并提供中文tutorial。后续将持续更新一些针对中文的优化。'

Under the description, there are two sections: '数据与资源' and '其他信息'. Under '数据与资源', there are two items:

- 支持中文处理的DeepDive
- 中文tutorial 中文版用例说明

Under the '其他信息' section, there is a '知识抽取' button.

关于OpenKG

□ 知识存储工具：北京大学gStore

/ 机构 / 北京大学 / 北京大学图数据库-gStore

北京大学图数据库-gStore

追从者 0

● 追随

机构



北京大学

北京大学图数据库-gStore

分类 活动流 管理

北京大学图数据库-gStore

一种基于图的RDF存储和SPARQL查询系统(triplestore)；支持标准的RDF数据n3文件导入；标准SPARQL 1.1查询语句（含Union, Optional, Filter等）；单机支持5亿规模的三元组数据；在5个Benchmark数据集（3-5亿三元组）上的测试，平均性能优于Virtuoso 7.2。

数据与资源

| | |
|----------------------------|----|
| gStore-github 主页 | 浏览 |
| gStore项目主页 | 浏览 |
| gStore-v0.4.0-source codes | 浏览 |
| gStore-v0.4.0-test report | 浏览 |
| gStore-v0.4.0-用户手册-英文 | 浏览 |

关于 OpenKG

□ 知识问答工具

The screenshot shows a web interface for managing knowledge graphs. At the top, there's a navigation bar with links for '机构' (Institution), 'OpenKG', and 'YodaQA-开源QA工具'. Below the navigation, there's a sidebar with sections for '追随者' (Followers) showing '0' and a green '追溯' (Trace) button; '机构' (Institution) with a logo for 'OpenKG.cn - 中文开放知识图谱联盟' and a link to 'OpenKG'; '社会' (Social) with links for 'Google+' and 'Twitter'; and '其他信息' (Other Information). The main content area has tabs for '工具' (Tools), '分类' (Categories), and '活动流' (Activity Stream). The '工具' tab is selected, displaying the title 'YodaQA-开源QA工具'. The page describes YodaQA as an open-source Factoid Question Answering system that can produce answers from databases and text corpora using on-the-fly information extraction. It mentions that by default, open-domain question answering is performed on top of the Freebase and DBpedia knowledge bases as well as the texts of enwiki articles. It also notes the practicality and extensibility of the design, current alpha status regarding accuracy and speed, future deductive capabilities, and personal assistant-style conversation capabilities. The YodaQA source code is linked to under '数据与资源' (Data & Resources).

关于OpenKG

□ 知识融合工具

机构 / OpenKG / Limes:实体链接发现框架

Limes:实体链接发现框架

追从者
0

● 追随

机构

OpenKG.cn 中文开放知识图谱联盟

OpenKG
没有关于此机构的描述

社会

Google+ Twitter

工具 分类 活动流 管理

Limes:实体链接发现框架

OpenKG搜集和整理知识图谱相关的技术工具，并将组织开展技术评测。Limes是一个基于度量空间的实体链接发现框架，结合了数学统计，前缀后缀，位置过滤来计算相似率。这个相似率用来过滤掉不符合匹配条件的实体对，以此减少比较时间复杂度，提高效率。

数据与资源

limes工具 源代码 浏览

知识链接

其他信息

| 域 | 价值 |
|------|---|
| 源 | https://github.com/AKSW/LIMES-dev |
| 状态 | active |
| 最近更新 | 2017年3月17日, 上午9点46分 (UTC+08:00) |

关于OpenKG

□ 链接的中文百科知识图谱



关于OpenKG

□ 链接的中文百科知识图谱



关于OpenKG

- 清华大学开放百科KG



XLORE

xlore

知识图谱以结构化的形式描述客观世界中概念、实体及其之间的关系，将互联网信息表达成更认知世界的形式，以更好地组织、管理和理解互联网海量信息。

关于OpenKG

□ 东南大学、狗尾草科技Zhishi.Me百科KG

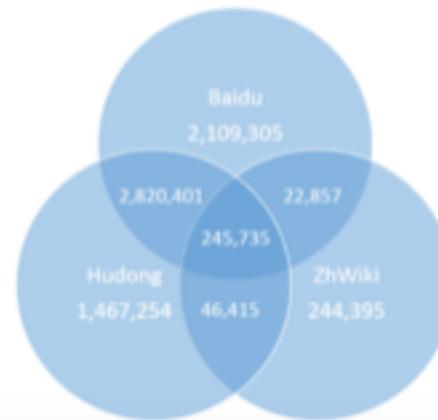


[About](#)
[Lookup](#)
[SPARQL](#)
[Publications](#)
[Credits](#)
[API](#)

Zhishi.me is an effort to build Chinese Linking Open Data. Currently, it covers three largest Chinese encyclopedias: Baidu Baike, Hudong Baike and Chinese Wikipedia.

Dataset overview:

- 5,198,298 entities from Baidu Baike
- 4,579,805 entities from Hudong Baike
- 559,402 entities from Chinese Wikipedia (zhwiki)



关于OpenKG

□ 复旦大学开放百科KG：CN-DBpedia

• 介绍

- CN-DBpedia是由复旦大学知识工场实验室（kw.fudan.edu.cn）研发并维护的大规模中文通用百科知识图谱，CN-Probase大规模中文IsA知识图谱
- 主要针对单数据源中文百科类网站（如百度百科、互动百科、中文维基百科等）进行深入挖掘，经知识抽取、知识清洗、知识填充以及知识更新等操作后，最终形成一个质量高、知识多、更新快的中文通用百科知识图谱，供机器和人访问

• 特点

- 结构化知识
- 高质量知识
- 知识丰富
- 实时更新

• CN-DBpedia规模（2017.08）

- 百科实体数 1600万+
- 百科关系数 2.1亿+

• DUMP规模（2015.07）

- 百科实体数 900万+
- 百科关系数 6700万+
 - mention2entity: 110万+
 - triples: 6600万+
 - 摘要: 400万+
 - 标签: 1980万+
 - infobox: 4100万+

CNProbase 规模

- 4600万+的IsA关系
- 其中实体数量2200万+
- 概念数量217万+

<http://openkg.cn/dataset/cndbpedia>

<http://openkg.cn/dataset/cnprobbase>

关于 OpenKG

□ 北京大学百科KG



PKUBase

PIE

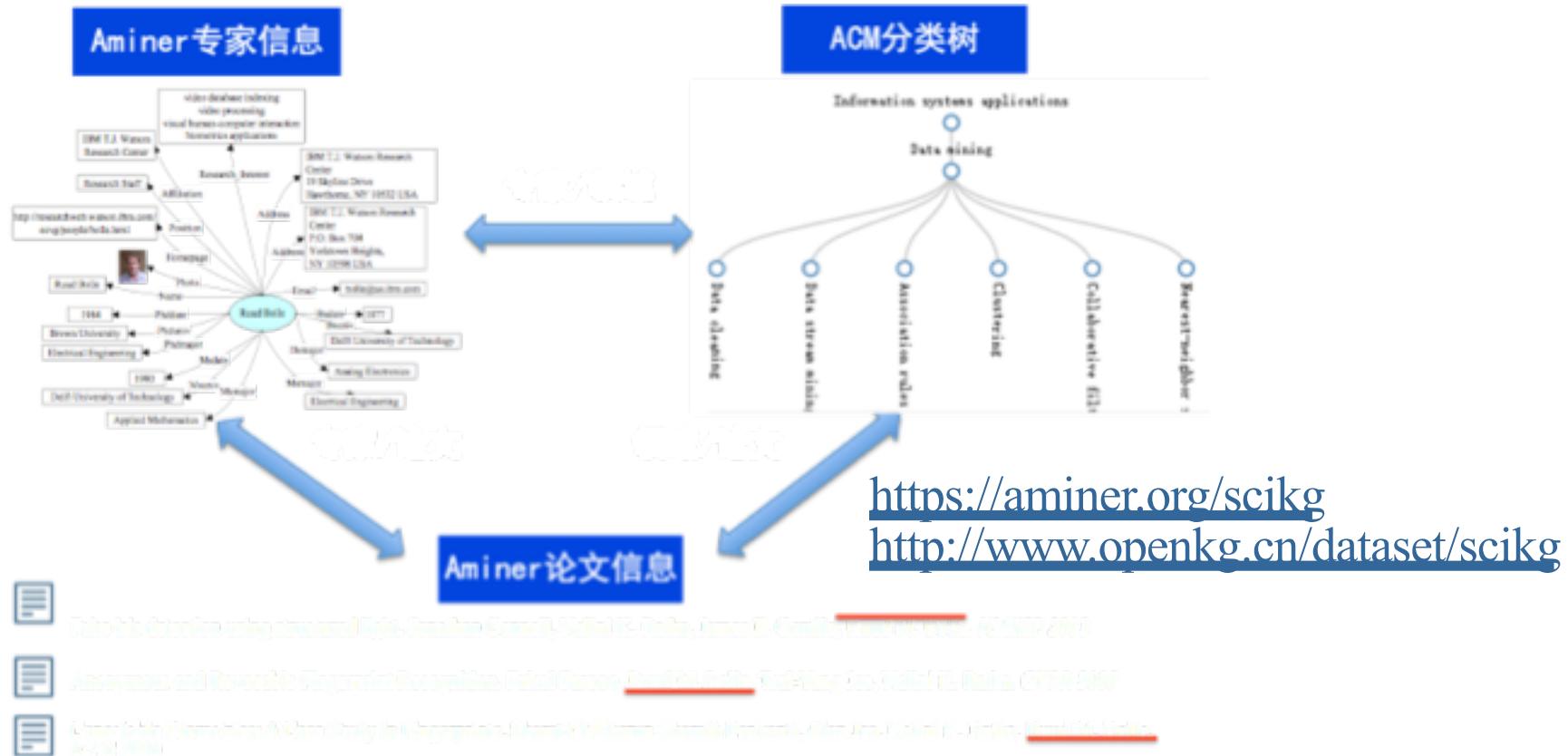
- 以多种网络百科为核心
 - 融合百度百科、维基百科、[DBpedia](#)、Freebase 等网络百科资源
- 包含 900 多万中文实体，4000 多万知识条目
 - RDF 三元组形式；支持标准结构化查询引擎
 - 动态更新
- 相对完整的中文类别体系框架
 - 中文特色；与[DBpedia](#)、维基等类别体系有交叠
 - 手工校正
- 丰富的谓词体系
 - 与类别体系紧密结合 (Domain/Range)



北京大学

关于OpenKG

□ 其它开放图谱：清华大学科技知识图谱



关于OpenKG

□ 其它开放图谱：微软亚洲研究院ConceptGraph

骨骼 / 机构 / 微软亚洲研究院 / Microsoft Concept Graph

Microsoft Concept Graph

数据集 分类 活动流 管理

追从者 2

追随

机构

Microsoft

微软亚洲研究院
没有关于此机构的描述

社会

Google+

Microsoft Concept Graph

微软推出Concept Graph以更好理解自然语言。背景知识是人类与机器在语言理解上的主要差别之一。Probbase是一个知识数据库，微软在该产品上开发了相当一段时间。该数据库是开放工具Microsoft Concept Graph（微软概念图）的基础。Probbase包含了540万个概念，超过了如提供12万个概念的Cyc等其他知识数据库。

数据与资源

Microsoft Concept Graph 浏览

ConceptGraph是以概念层次体系为中心的知识图谱。与Freebase等知识图谱不同，ConceptGraph中的实体之间的IsA关系为主。...

其他信息

| 域 | 价值 |
|------|---|
| 源 | https://concept.research.microsoft.com/ |
| 状态 | active |
| 最近更新 | 2017年2月16日,下午4点12分 (UTC+08:00) |
| 创建的 | 2017年1月17日,晚上9点49分 (UTC+08:00) |

关于OpenKG

□ 其它开放图谱：文因互联-公众公司人物图谱

/ 机构 / 北京文因互联科技有限公司 / 文因互联-公众公司人物图谱

文因互联-公众公司人物图谱

追从者
0

● 追随

机构

MEMECT
文因互联
—记录世界的一刻—

北京文因互联科技有限公司
没有关于此机构的描述

数据集 分类 活动流 管理

文因互联-公众公司人物图谱

本数据包含主板、三板公司31w高管人物简历信息，以API方式提供访问。

数据与资源

公众公司列表
HTTP GET 参数: 无 URI例子: http://openapi.fingo.memect.cn/resume_company 浏览

公司高管列表
HTTP GET 参数: company_code 公司代码... 浏览

高管简历信息
HTTP GET 参数: company_code 公司代码 leader_name 高管姓名... 浏览

金融

其他信息

| 域 | 价值 |
|---|----|
|---|----|

关于OpenKG

□ 其它开放图谱：上海图书馆名人手稿档案关联开放数据

备 / 机构 / 上海图书馆 / 上海图书馆名人手稿档案关联开放数据集

上海图书馆名人手稿档案关联开放数据集

追从者 0

追随

机构


上海图书馆
上海图书馆名人手稿档案关联开放数据集

上海图书馆

上海图书馆，简称“上图”，上海市最大的图书馆，是中国大陆规模仅次于中国国家图书馆（位于北京市）的第二大图书馆。国家一级图书馆，古籍特

数据集 分类 活动流 管理

上海图书馆名人手稿档案关联开放数据集

本数据集包括上海图书馆馆藏的24万余种手稿及档案的元数据。其数量丰富、类型多样：包括私人信函电报12万余通，珍贵的创作手稿1万7千余种，谕旨奏折公牍条约1万4千余种，合同章程4千余种，照片及音像资料1万8千余种，等等。涉及名人近5万、地点2千余个，时间跨度为19世纪初至20世纪末近200年。数据经过语义化的清洗、加工、转换进行知识组织之后，不再是简单的关于文献的描述性元数据，而存在着丰富的人与人、人与文献、文献与文献间的关联关系，是研究近现代历史、人文、经济、社会等问题的宝贵资料。

数据与资源

| | | |
|---|--|---|
|  | 网站首页 |  |
|  | API接口说明文档 |  |
|  | 上图名人手稿本体 |  |
|  | 实例API接口 http://data.library.sh.cn/sg/resource/instance/ |  |

关于OpenKG

□ 其它开放图谱：荷兰阿姆斯特丹自有大学医学知识图谱

/ 机构 / Vrije University of Amsterdam / 乳腺癌知识图谱

乳腺癌知识图谱

1 追从者

● 追随

机构

VU VRIJE UNIVERSITEIT AMSTERDAM

Vrije University of Amsterdam

没有关于此机构的描述

社会

Google+ Twitter

数据集 分类 活动流 管理

乳腺癌知识图谱

乳腺癌知识图谱集成了乳腺癌相关的知识/数据资源，包括乳腺癌临床试验数据，乳腺癌医学指南，乳腺癌电子病历，乳腺癌临床试验数据语义标注，乳腺癌医学文献等，数据规模超过两千两百万三元组。

数据与资源

乳腺癌知识图谱 浏览

临床试验 乳腺癌 医学指南 医学文献 电子病历 知识图谱 语义标注

其他信息

| 域 | 价值 |
|------|---|
| 源 | http://wasp.cs.vu.nl/BreastCancerKG/ |
| 作者 | 黄智生 |
| 维护者 | 黄智生 |
| 状态 | active |
| 最后更新 | 2016年12月20日 18时19分35秒 UTC+08:00 |

关于OpenKG

□ 其它开放图谱：华东理工大学的中文症状库

/ 机构 / 华东理工大学 / 中文症状库

中文症状库

数据集 分类 活动流 管理

追从者 0

追随

机构 华东理工大学

没有关于此机构的描述

社会 Google+

中文症状库

这是一个包含症状实体和症状相关三元组的数据集。中文症状库的数据来自8个主流的健康咨询网站、3个中文百科网站和电子病历。它还包含了中文症状与UMLS中概念的链接结果。我们为该数据集提供了关键词查询和SPARQL查询功能 (<http://dsc.nlp-bigdatalab.org:8081/search.html>)。

数据与资源

| | |
|--------------|----|
| 中文症状知识库 | 浏览 |
| 数据集在DCAT中的描述 | 浏览 |
| 数据示例 | 浏览 |
| 数据集质量评估结果 | 浏览 |
| 症状知识库的模式图 | 浏览 |

关于OpenKG

□ 其它开放图谱：广州索答菜谱知识图谱

索答 / 机构 / 广州索答信息科技有限公司 / 索答菜谱本体信息

索答菜谱本体信息

追从者 1

◆ 追随

机构



SUMMBA
Summary-Based Answers

广州索答信息科技有限公司

没有关于此机构的描述

数据集 分类 活动流 管理

索答菜谱本体信息

索答50w菜谱本体信息，每个菜谱包含菜名，食材，味道，烹饪时间等属性。

数据与资源

索答菜谱本体信息

索答菜谱本体信息，包含50万个实体，每个菜谱包含菜名，食材，味道，烹饪时间等属性。

浏览

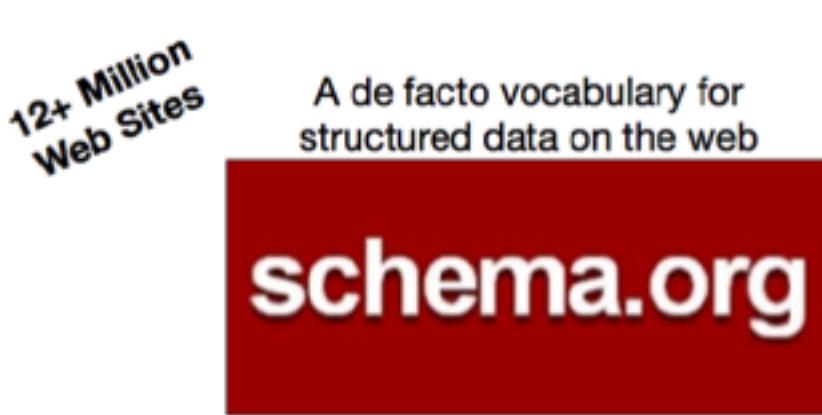
味道 烹饪时间 菜名 菜谱 食材

其他信息

| | |
|------|---------------------------------|
| 域 | 价值 |
| 作者 | yunbiao |
| 维护者 | yunbiao |
| 状态 | active |
| 最近更新 | 2017年1月17日,下午2点31分 (UTC+08:00) |
| 创建的 | 2017年1月17日,下午12点51分 (UTC+08:00) |

回顾 Web Schema 的进化

- 标准优先 (源于1996) ? MCF, RDF, OWL, FOAF, RSS, ...
- 数据优先 (源于2004) ? Swoogle, DBpedia, Linked Data, Open Government Data
- 消费者优先 (源于2008) ! searchMonkey, Facebook Open Graph, Schema.org (2011)



schema.org 简介

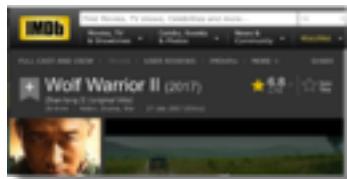
FIGURE 4A: MAJOR SITES THAT HAVE PUBLISHED SCHEMA.ORG

| CATEGORY | SITES |
|----------------|---|
| News | nytimes.com, guardian.com, bbc.co.uk |
| Movies | imdb.com, rottentomatoes.com, movies.com |
| Jobs / Careers | careerjet.com, monster.com, indeed.com |
| People | linkedin.com, pinterest.com, familysearch.org, arch |
| Products | ebay.com, alibaba.com, sears.com, cafepress.com, |
| Video | youtube.com, dailymotion.com, frequency.com, vin |
| Medical | cvs.com, drugs.com |
| Local | yelp.com, allmenus.com, urbanspoon. |
| Events | whereevent.com, meetup.com, zillow.c |
| Music | last.fm, myspace.com, soundcloud.co |



Schema.org运行机理

1. 发布网页



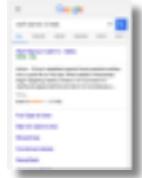
2. 检查Movie schema



3. 扩充页面语义标注



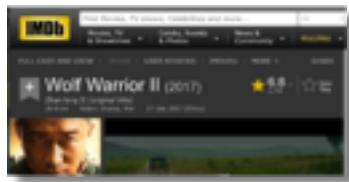
4. 消费rich snippet



A screenshot of the IMDB movie page for "Wolf Warrior II". The page features a large image of the movie poster on the left and a video player showing a scene from the trailer on the right. Below the poster, there is a summary of the movie's plot. At the bottom, there is a call-to-action button for getting showtimes and tickets.

Schema.org运行机理

1. 发布网页



The screenshot shows the schema.org homepage with a red header containing the site's name. Below the header, there is a search bar and navigation links for 'Home', 'Schemas', and 'Documentation'.

2. 检查Movie schema



Movie
Canonical URL: <http://schema.org/Movie>

Thing > CreativeWork > Movie

A movie.

3. 扩充页面语义标注



Usage: Between 10,000 and 50,000 domains

[more...]

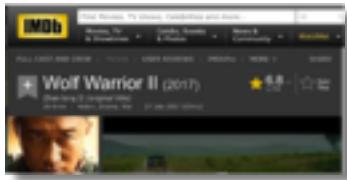
| Property | Expected Type | Description |
|---------------------------------|-------------------------|---|
| Properties from Movie | | |
| actor | Person | An actor, e.g. in tv, radio, movie, video games etc., or in an event. Actors can be associated with individual items or with a series, episode, clip. Supersedes actors . |
| countryOfOrigin | Country | The country of the principal offices of the production company or individual responsible for the movie or program. |
| director | Person | A director of e.g. tv, radio, movie, video gaming etc. content, or of an event. Directors can be associated with individual items or with a series, |

4. 消费rich snippet



Schema.org运行机理

1. 发布网页



2. 检查Movie schema



3. 扩充页面语义标注



4. 消费rich snippet



Google Structured Data Testing Tool

http://www.imdb.com/title/tt7131870/

NEW TEST

Movie

PREVIEW 4 ERRORS 5 WARNINGS

ID: http://www.imdb.com/title/tt7131870/pagecontent

| @type | Movie |
|---------------|---|
| @id | http://www.imdb.com/title/tt7131870/pagecontent |
| name | Wolf Warrior II (2017) |
| duration | PT2H6M |
| genre | Action |
| genre | Drama |
| genre | War |
| datePublished | 2017-07-27 |
| image | https://images-eu.ssl-images-amazon.com/images/M/Mv5BMTY0NjU4NzE0NzE5MjQ4LnBuZw.. |

1000 tcoosat="tt7131870" data-caller-name="title">
1001 <div class="wlb_dropdown_box" data-tcoosat="tt7131870" data-size="large" data-caller-name="title" data-type="primary"></div>
1002 <div class="wlb_dropdown_list" style="display:none"></div>
1003 <div class="title_wrapper">
1004 <h1 itemprop="name" class="">Wolf Warrior
1005 II2017</h1>
1006 <div class="originalTitle">Zhan
1007 lang II (original
1008 title)</div>
1009 <div class="subtext">
1010 <time itemprop="duration" datetime="PT2H6M">
1011 2h 6min
1012 </time>
1013 |
1014 Action,
1015 Drama..

Schema.org运行机理

1. 发布网页



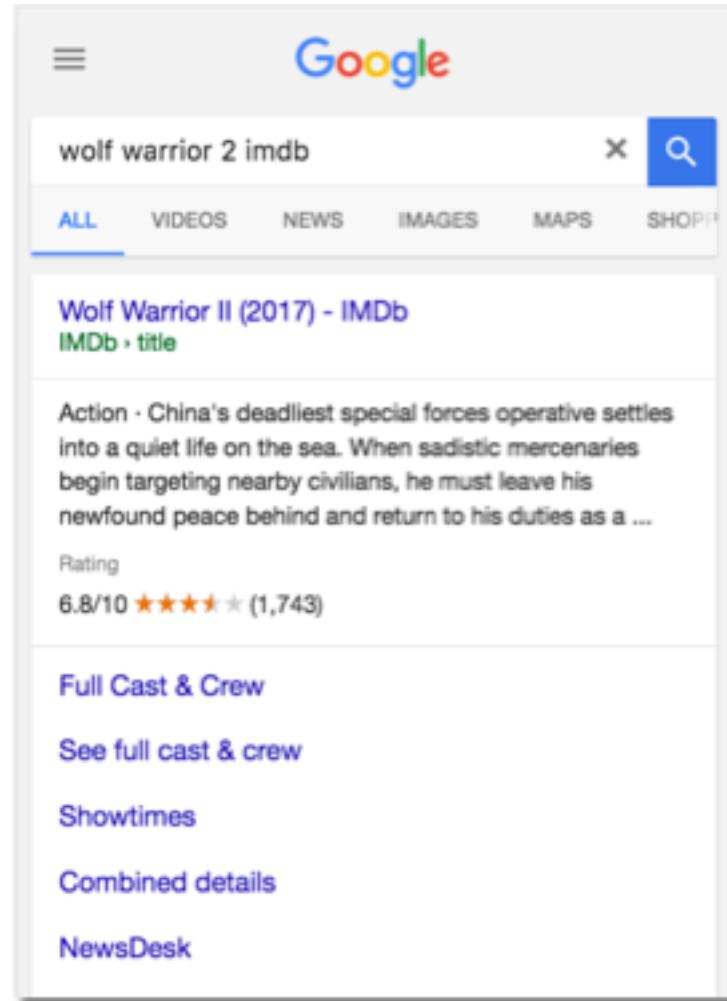
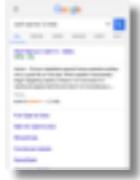
2. 检查Movie schema



3. 扩充页面语义标注



4. 消费rich snippet



The screenshot shows a Google search result for the query "wolf warrior 2 imdb". The top result is a link to the IMDb page for "Wolf Warrior II (2017) - IMDb". Below the link, there is a rich snippet summary. The summary includes the movie's title, a short plot description ("Action · China's deadliest special forces operative settles into a quiet life on the sea. When sadistic mercenaries begin targeting nearby civilians, he must leave his newfound peace behind and return to his duties as a ..."), its rating ("Rating 6.8/10 ★★★★☆ (1,743)"), and several interactive options: "Full Cast & Crew", "See full cast & crew", "Showtimes", "Combined details", and "NewsDesk".

有趣的Schema定义

what is schema



{ Money.

Money is useless unless you spent it

and

Schema is useless unless use it with data

cnSchema

- cnSchema是中文知识图谱的“普通话”“标准
 - cnSchema 面向中文信息处理，利用先进的知识图谱、自然语言处理和机器学习技术，融合结构化与文本数据，支持快速领域知识建模，支持跨数据源、跨领域、跨语言的开放数据自动化处理，为智能机器人、语义搜索、智能计算等新兴应用市场提供schema层面的支持与服务。
-

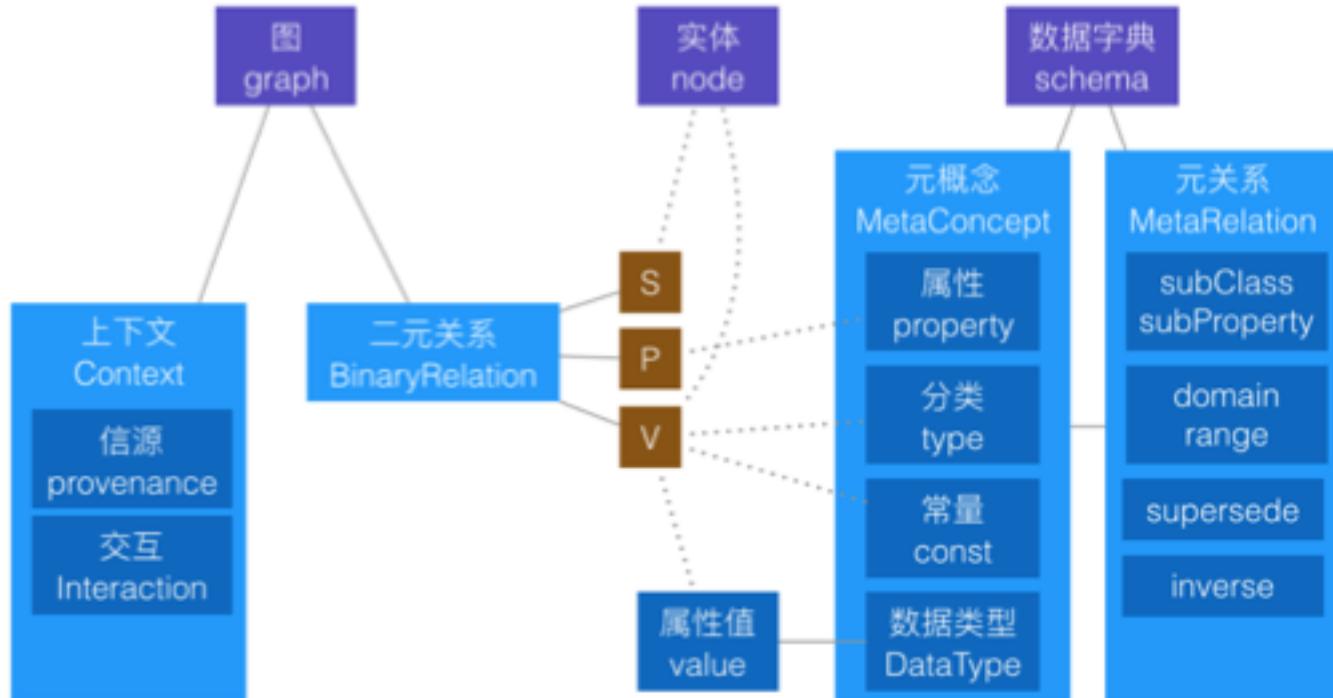
cnSchema.Org: KR for KG

通用

cnSchema.org是一个基于社区维护的开放的知识图谱Schema标准。cnSchema的词汇集包括了上千种概念分类(classes)、数据类型(data types)、属性(properties)和关系(relations)等常用概念定义，以支持知识图谱数据的通用性、复用性和流动性。结合中文的特点，我们复用、连接并扩展了Schema.org, Wikidata, Wikipedia等已有的知识图谱Schema标准，为中文领域的开放知识图谱、聊天机器人、搜索引擎优化等提供可供参考和扩展的数据描述和接口定义标准。

复用

流动



cnSchema主要贡献者

cnSchema.org是由来自清华大学、浙江大学、北京大学、复旦大学、东南大学、南京大学、英国阿伯丁大学等十多所国内外高校的计算机科学专家，以及微软亚洲研究院、海知智能、狗尾草科技、文因互联等企业所共同发起、建立并维护。cnSchema也得到Schema.org的负责人R.V. Guha和Dan Brickley的指导和支持。cnSchema是完全开放的，有关它的讨论也是完全公开的。

cnSchema贡献者

邹磊(北京大学) 于彤(中国中医科学院信息研究所) 张鹏(清华大学) 徐波(复旦大学)

侯磊(清华大学) 胡伟(南京大学) 吴天星(东南大学) 徐常亮(新华智云)

张大卫(微软亚洲研究院) 徐艺(微软亚洲研究院) 仲亮靓(东南大学) 王梁(浙江大学)

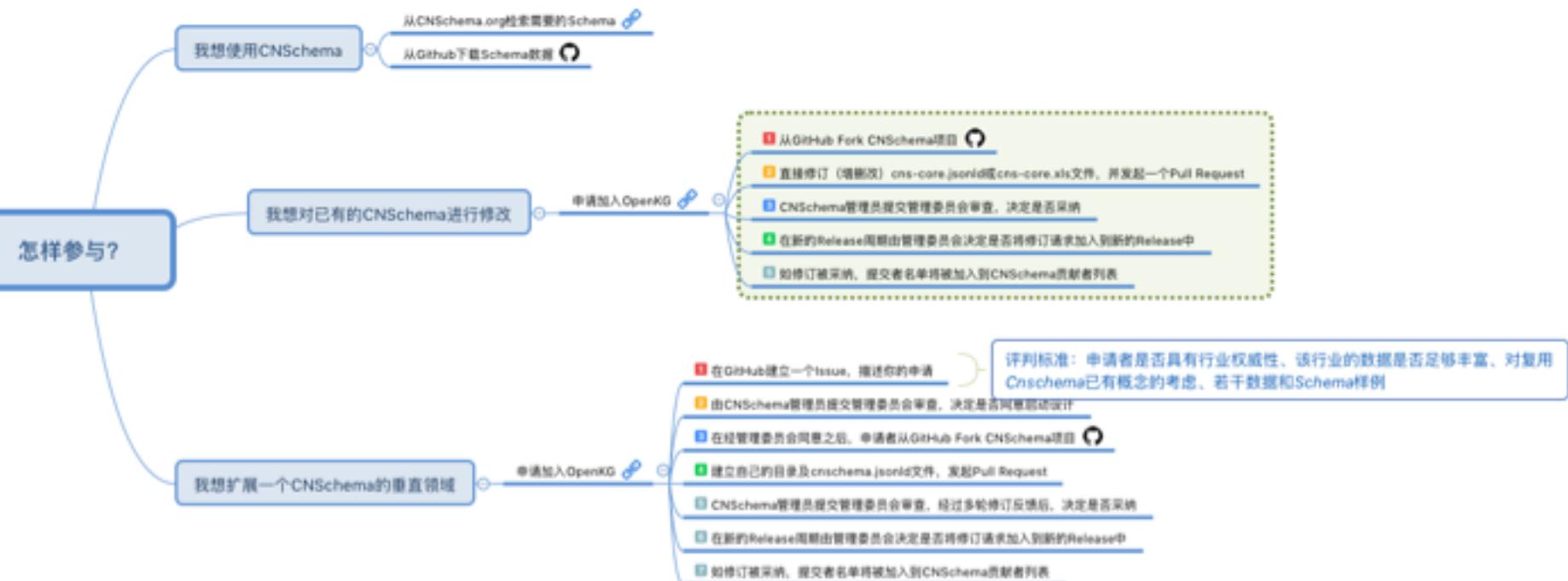
张宇轩(浙江大学) 邓淑敏(浙江大学) 王宇(上海海知智能) 杨平京(上海海知智能)

孙娜 董孙竑路 陈旭 郭唯 Kevin Xin

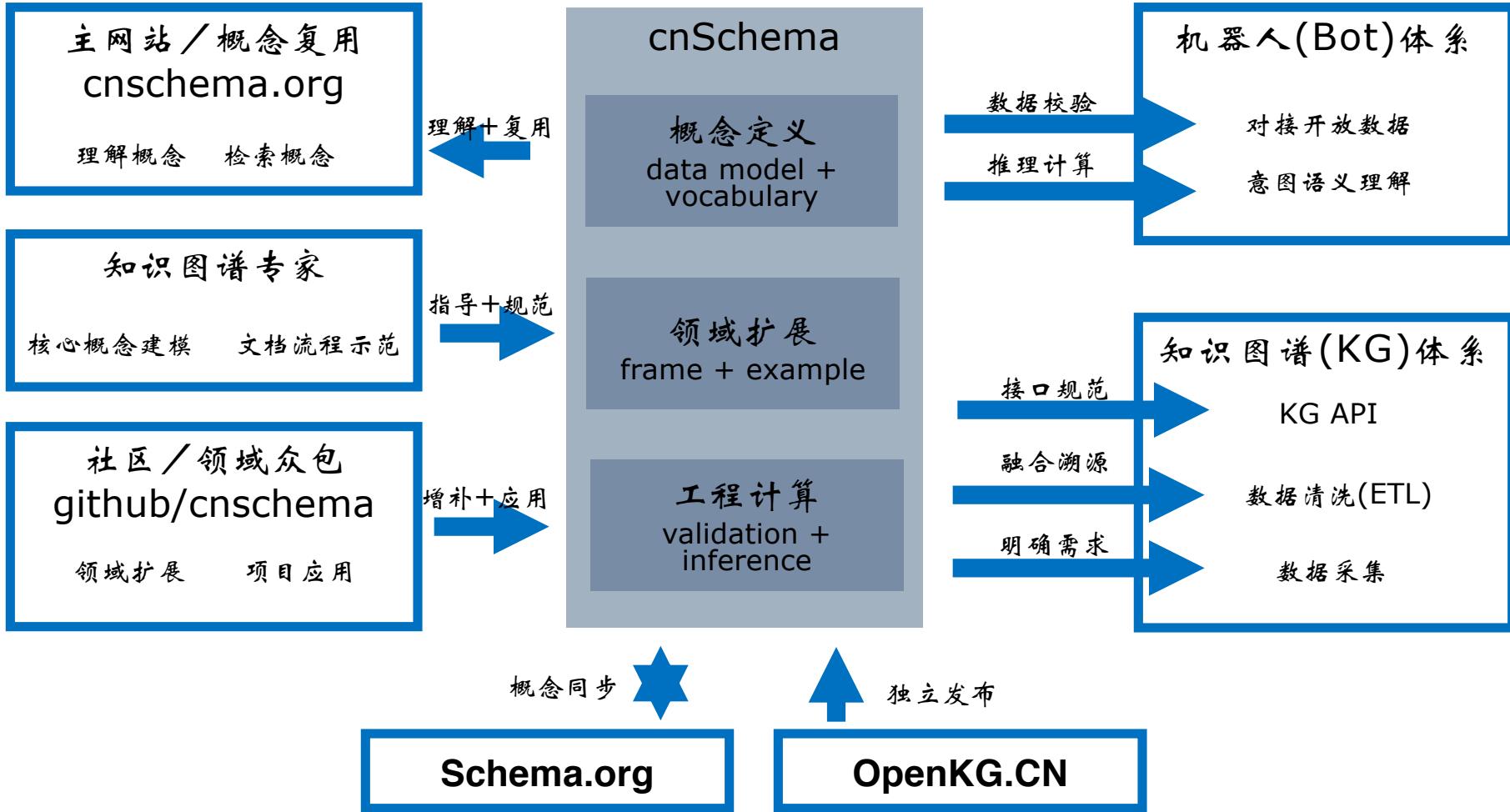
cnSchema发起人

| | |
|----------|---------------|
| 丁力 | 上海海知智能 CTO/博士 |
| 陈华钧 | 浙江大学 教授 |
| 漆桂林 | 东南大学 教授 |
| 王昊奋 | 深圳狗尾草 CTO/博士 |
| 谢殿侠 | 上海海知智能 CEO |
| 李涓子 | 清华大学 教授 |
| 闫峻 | 微软亚洲研究院 研究员 |
| 肖仰华 | 复旦大学 副教授 |
| 鲍捷 | 文因互联 CEO/博士 |
| 曾毅 | 中科院自动化所 研究员 |
| Jeff Pan | 英国阿伯丁大学 教授 |

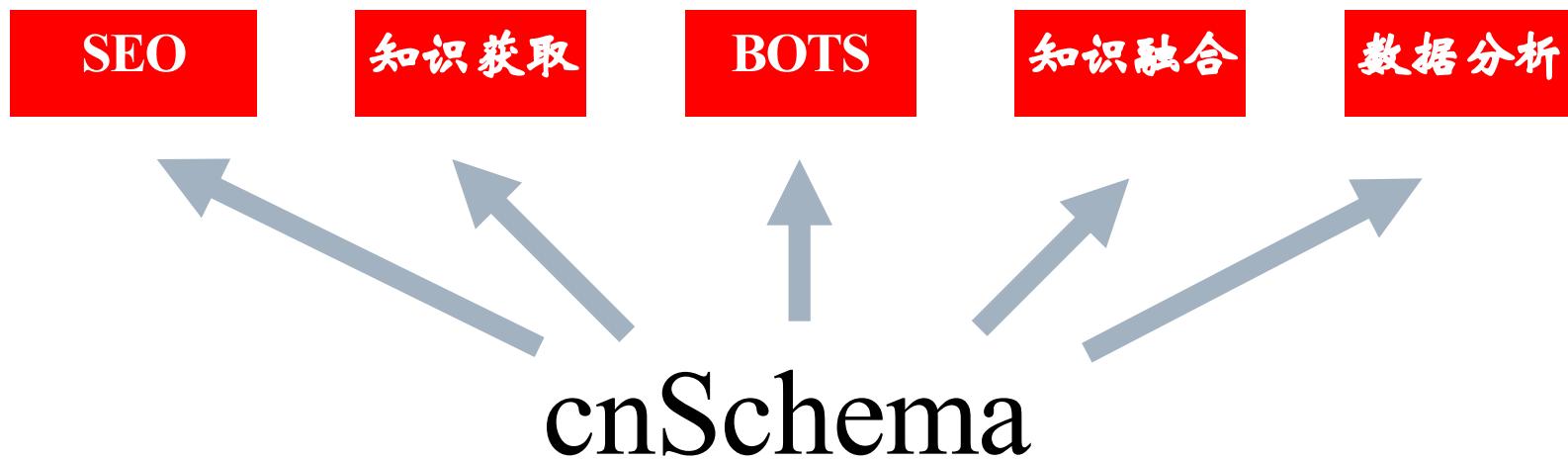
怎么参与



cnSchema 生态体系



cnSchema 的价值



用于机器人和KG

产品经理

“用户关心什么问题？”

“我的产品需要哪些数据？”

开发者

“接口字段怎么理解？”

“KG存储结构和查询怎么写？”

”

数据发布者／采集者

“我的数据哪里需要改进？”

“让人找到并用到我的数据？”

学习问答意图模版

苏东坡的父亲是谁？

@kg.person 的 @cns.father 是谁？

cnSchema 项目 文档 词汇表 关于 OpenKG.CN

成员

membershipNumber(会员编号)
A unique identifier for the membership.(或其身份的唯一标识)
<http://cn-schema.org/membershipNumber>

musicGroupMember(乐团成员)
A member of a music group—for example, John, Paul, George, or Ringo.(一个音乐团体的成员。例如John, Paul, George, Ringo)
<http://cn-schema.org/musicGroupMember>

查属性名定义

描述数据源结构



五月天 Mayday

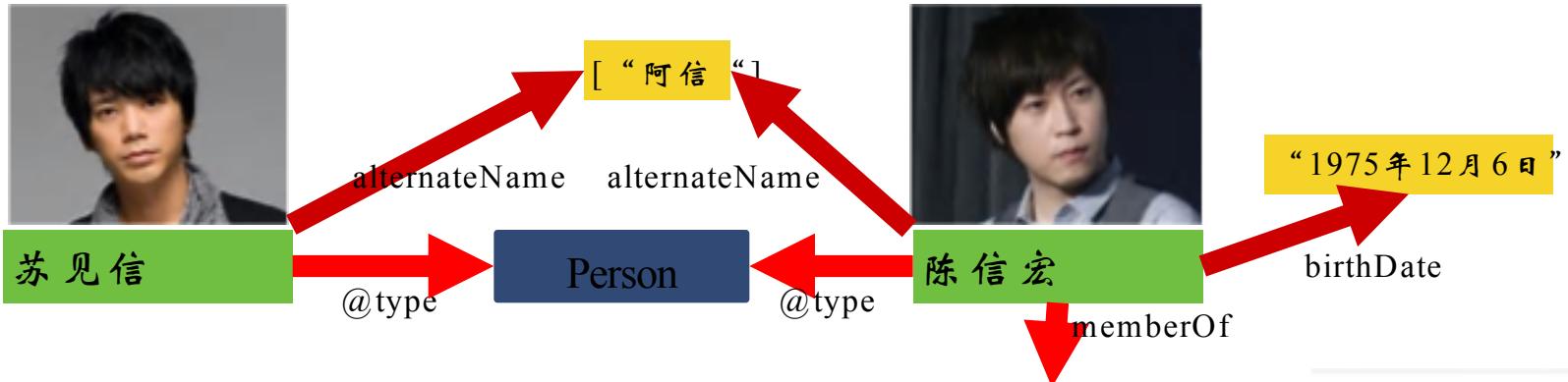
成立时间: 1997-01-01

国家/地区: 台湾

厂牌: 滚石唱片(Rock)/相信音乐

乐队成员: 阿信/怪兽/玛莎/石头/冠佑

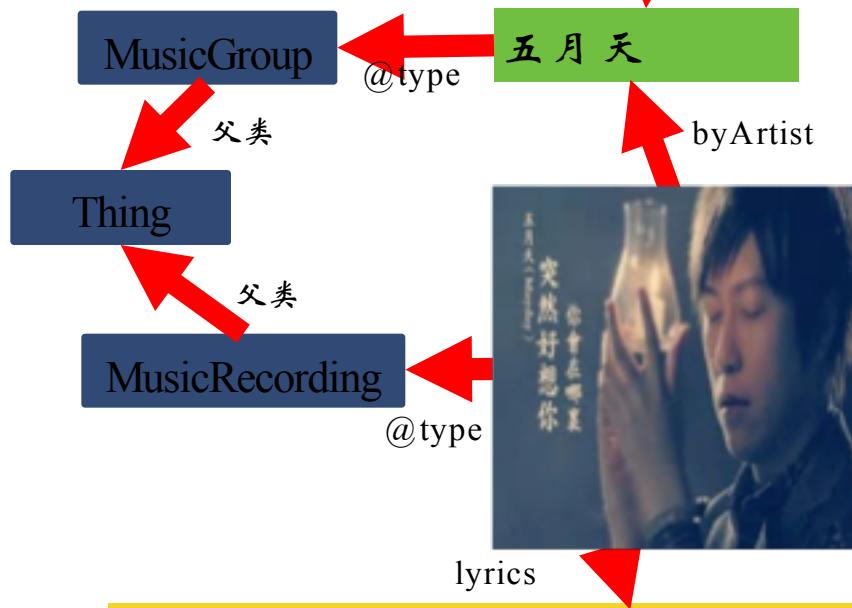
用于实体链接、查询和任务补全



Entity Linking using web statistics

P(苏见信) 92,300 RESULTS
P(陈信宏) 53,600 RESULTS
P(陈信宏, 阿信) 25,800 RESULTS
P(苏见信, 阿信) 16,200 RESULTS

- 我要听“突然好想你”
- 我要听“阿信”的歌
- 放一个“突然好想你”你会在哪里过得快乐或委屈
- “阿信”是“五月天”里最老的吗？



| 五月天的热门歌曲 TOP100 | | |
|-----------------|-----------------------------------|----------|
| * 全部播放 | | |
| 01 | 突然好想你 | 19621194 |
| 02 | What If... | 15903899 |
| 03 | 后来的我们 | 13603860 |
| 04 | Haze, After Us | 13433860 |
| 05 | 做不到让你一个人 | 87433860 |
| 06 | 好好 (把你爱说一首歌) | 86201080 |
| 07 | 筑梦指南《你的名字。》中文主题曲 / Song About You | 86201080 |

“...突然好想你 你会在哪里，过得快乐或委屈...”

用于Web数据抽取与融合



采集+清洗



```
{ "statedIn": "虾米",  
  "": 16948661,  
  ...}
```

使用统计

- * 信息来源、用户播放数等数据不能融合
- * 融合结果溯源，支持数据可信度分析

融合



采集+清洗



```
{ "statedIn": "网易",  
  "playCount": 2731000,  
  ...}
```

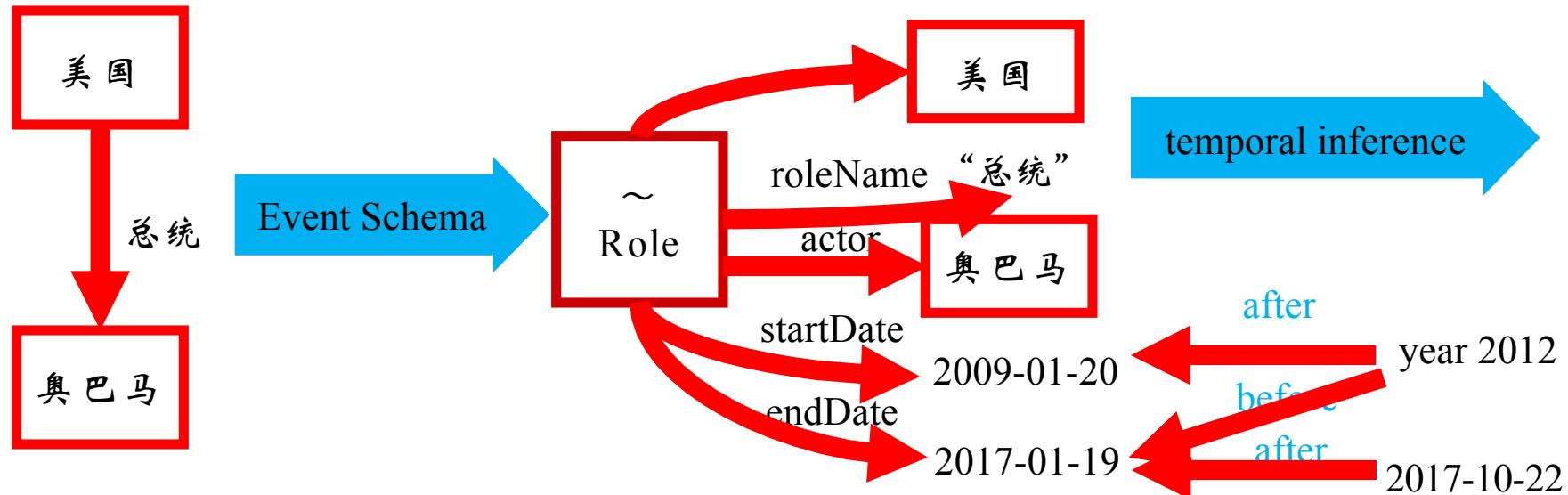
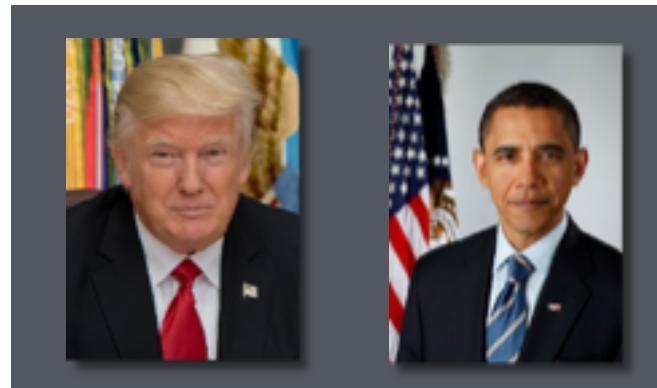
```
{ "": [  
    { "statedIn": "虾米",  
      "playCount": 16948661,  
      ...},  
    { "statedIn": "网易",  
      "playCount": 2731000,  
      ...}],  
  ...  
}
```

来源

• “十年在虾米的评价如何？”

用于KG数据建模与推理

- 现在的美国总统是谁？
- 2012年美国总统是谁？



扩展中文概念

- 拿“籍贯”这个属性来说，只有中国人有，所以schema.org没有收录。wikidata有收录，<https://www.wikidata.org/wiki/Property:P66>，因此使用ancestralHome
- 有些中文属性更难翻译到贴切的英文概念，只能采用拼音。

| C | D | E | F | G | H |
|------------------------|------------|---------------|---|------------------------|-------------|
| 规范属性名 | 中文属性名 | schema.org属性 | wikidata属性 | w | cnschema属性名 |
| name | 姓名 | name | | | |
| alternateName | 别名 | alternateName | | | |
| description | 简介 | description | | | |
| image | 图片 | image | | | |
| keywords | 标签 | keywords | | | |
| birthPlace | 出生地 | birthPlace | P19 | | |
| birthDate | 出生日期 | birthDate | P569 | | |
| deathDate | 死亡日期 | deathDate | P570 | | |
| deathPlace | 死亡地 | deathPlace | P20 | | |
| placeOfBurial | 墓地、安葬地 | | P119 | placeOfBurial | |
| homeLocation | 家庭地址 | homeLocation | | | |
| alumniOf | 毕业院校 | alumniOf | P69 | | |
| ancestralHome | 籍贯 | | P66 | ancestralHome | |
| occupation | 职业、身份 | | P106 | occupation | |
| cnProfessionalTitle | 职称 | | | cnProfessionalTitle | |
| fieldOfWork | 领域、专业 | | P101 | fieldOfWork | |
| academicMajor | 高校专业 | | P812 | academicMajor | |
| ethnicGroup | 民族、民族族群 | | P172 | ethnicGroup | |
| nobleFamily | 家族 | | https://www.wikidata.org/wiki/Property:P119 | nobleFamily | |
| religion | 宗教信仰、信仰 | | P140 | religion | |
| memberOfPoliticalParty | 政党、党派 | | P102 | memberOfPoliticalParty | |
| courtesyName | 字、表字 | | P1782 | courtesyName | |
| artName | 号、自号、别号、又号 | | P1787 | artName | |
| templeName | 庙号 | | P1785 | templeName | |
| posthumousName | 谥号、私谥 | | P1786 | posthumousName | |
| pseudonym | 化名 | | P742 | pseudonym | |
| birthName | 原名 | | P1477 | birthName | |
| familyName | 姓氏 | familyName | P734 | familyName | |

谢谢