



2023全国知识图谱与语义计算大会

China Conference on Knowledge Graph and Semantic Computing – 沈阳 – 2023年

面向知识图谱的构建与推理研究进展

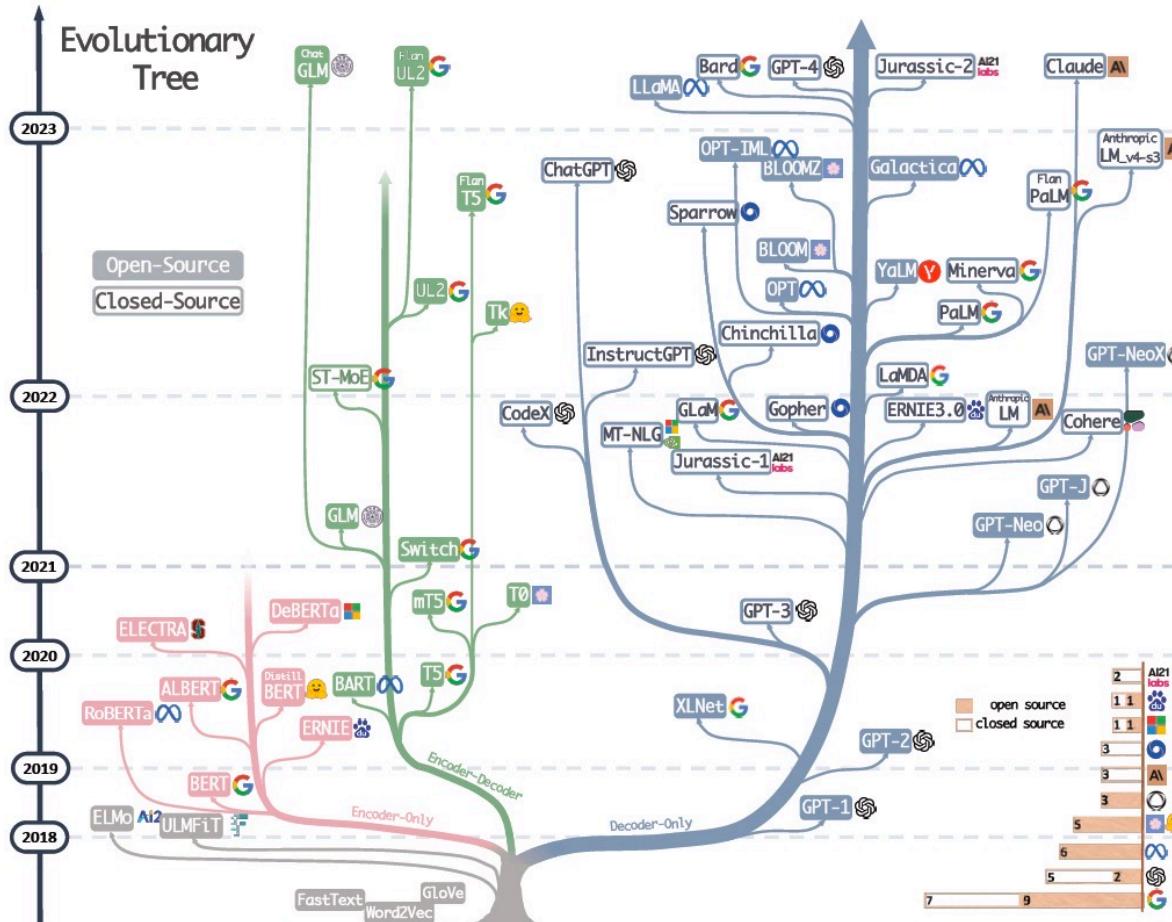
张宁豫

浙江大学

08/26/2023



大模型时代还需要知识图谱么？



VS.

符号知识图谱



[1] Harnessing the Power of LLMs in Practice: A Survey on ChatGPT and Beyond. 2023

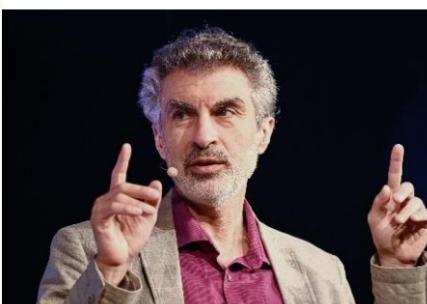
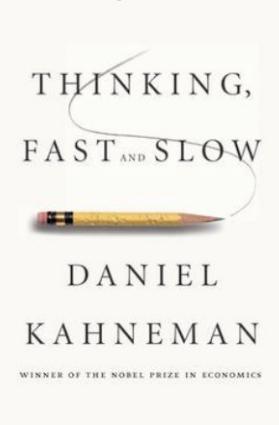
大模型时代还需要知识图谱么？

□ 语言的有限性-- *Yann LeCun*

- 语言只承载了人类全部知识的一小部分，很多知识学习依赖于与真实世界的交互

□ 深度学习快慢双系统-- *Yoshua Bengio*

- 有意识的、带逻辑、规划、推理的系统



EL How many World Cup championships has Messi won?

Lionel Messi has won four World Cup championships with the Argentina national team. He has played in three World Cup tournaments (2006, 2010, and 2014), and his team finished as runners-up in the 2014 World Cup.

GPT-4 自我对话

用户启动: 今天很热。

GPT-4 会话1: 是的, 似乎今天非常热! 记得要补充水分并尽量保持凉爽。如果你要到户外去, 确保涂抹防晒霜。

GPT-4 会话2: 感谢你的建议! 我一定会多喝水并且涂抹防晒霜。也许今天是个吃冰淇淋的好日子。

GPT-4 会话1: 不客气! 在这样一个炎热的日子里享用冰淇淋听起来是个很好的主意。祝你玩得开心, 保持凉爽!

GPT-4 会话2: 谢谢你! 我会确保自己有个美好的时光并保持凉爽。祝你也过得愉快!

.....

GPT-4 会话1: 谢谢你! 如果你有任何问题或需要帮助, 请随时提问。祝你度过美好的一天!

GPT-4 会话2: 谢谢!

符号表示落地问题！

Fig. 1: 在用户启动话题后两个GPT-4之间的自我对话。两个模型在只进行一轮对话后就迅速转变为彼此表示感谢（标记为黄色）。5轮之后仍旧无限重复，这与符号落地理论相呼应，即表明在没有建立词语到世界的联结时，对话将进入无限循环。

[1] Brain in a Vat: On Missing Pieces Towards Artificial General Intelligence in Large Language Models. 2023

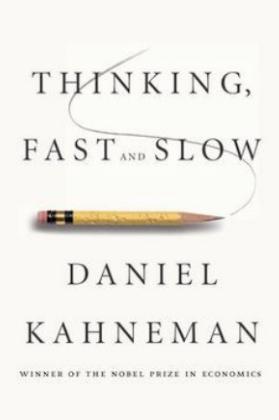
大模型时代还需要知识图谱么？

□语言的有限性-- *Yann LeCun*

- 语言只承载了人类全部知识的一小部分，很多知识学习依赖于与真实世界的交互

□深度学习快慢双系统-- *Yoshua Bengio*

- 有意识的、带逻辑、规划、推理的系统



Given a set of rules and facts, you have to reason whether a statement is true or false. Here are some facts and rules:

The bear likes the dog.

The cow is round.

The cow likes the bear.

The cow needs the bear.

The dog needs the squirrel.

The dog sees the cow.

The squirrel needs the dog.

If someone is round then they like the squirrel.

If the bear is round and the bear likes the squirrel then the squirrel needs the bear.

If the cow needs the dog then the cow is cold.

Does it imply that the statement "The cow likes the squirrel." is True?

Given a set of rules and facts, you have to reason whether a statement is true or false. Here are some facts and rules:

The e4 likes the e5.

The e14 is e2.

The e14 likes the e4.

The e14 needs the e4.

The e5 needs the e26.

The e5 sees the e14.

The e26 needs the e5.

If someone is e2 then they like the e26.

If the e4 is e2 and the e4 likes the e26 then the e26 needs the e4.

If the e14 needs the e5 then the e14 is e1.

Does it imply that the statement "The e14 likes the e26." is True?

复杂推理问题！

复杂鲁棒可信的人工智能系统仍离不开高质量的知识图谱

[1] Brain in a Vat: On Missing Pieces Towards Artificial General Intelligence in Large Language Models. 2023



面向知识图谱的
构建与推理研究进展

浙江大学

张宁豫，博士，浙江大学
副教授，博士生导师。研究
方向包括知识图谱、自然
语言处理等。



知识增强的
大规模语言模型
进展及展望

上海交通大学

陈露，上海交通大学计算机科学
与工程系助理研究员、硕士生导师。
主要研究兴趣包括智能人机对话与问答、自然语言处理等。

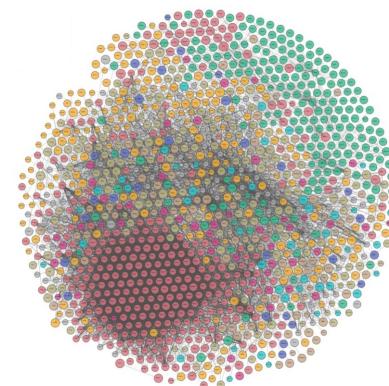


费根鲍姆
图灵奖1994

图灵奖获得者费根鲍姆提出了**知识工程**的概念，确立了其在人工智能中的重要地位



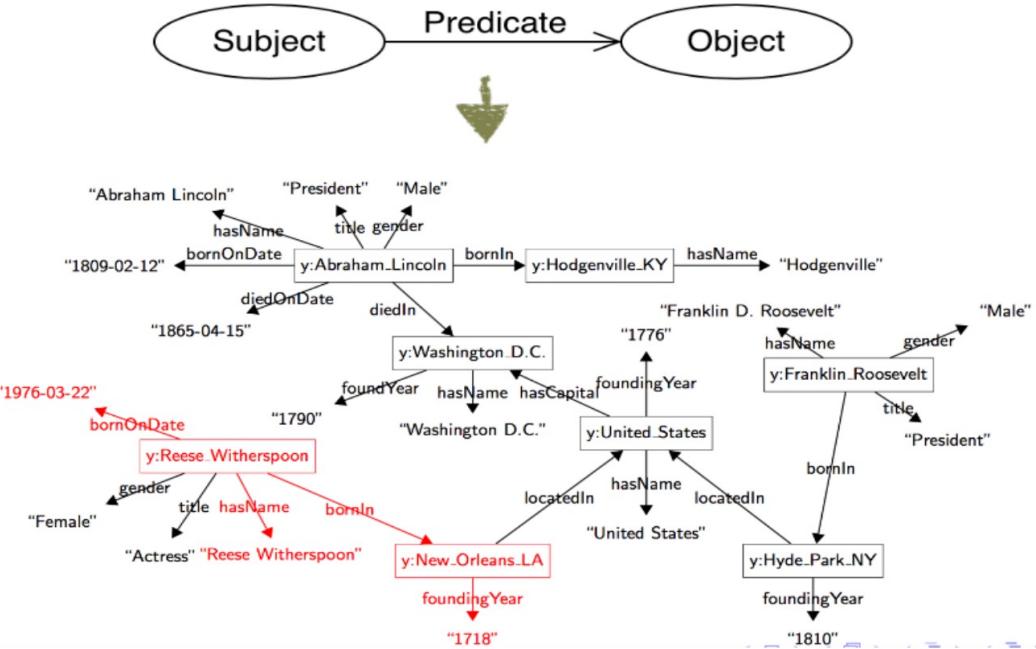
大数据



大知识

知识图谱接近自然语言和人脑认知、
易于扩展增加、易于神经网络化

知识图谱 = 文字语义 + 结构知识



语言 ≠ 知识

表示方法	可解释性	知识类型	可计算性
自然语言	人可理解	显性知识	不易于计算处理
知识图谱	人可理解	显性知识 + 隐性知识	较容易计算处理
语言模型	人不可理解	隐性知识	易于计算和处理

自然语言处理领域 (NLP)

ACL(23,22,21), EMNLP(22,21),
NAACL(22,21), COLING(22,21),
EACL(23,21)

人工智能&深度学习领域 (AI)

AAAI(22,21), IJCAI(23,22,21),
ICLR(23,22,21), ICML(23,22,21),
NeurIPS(22,21)

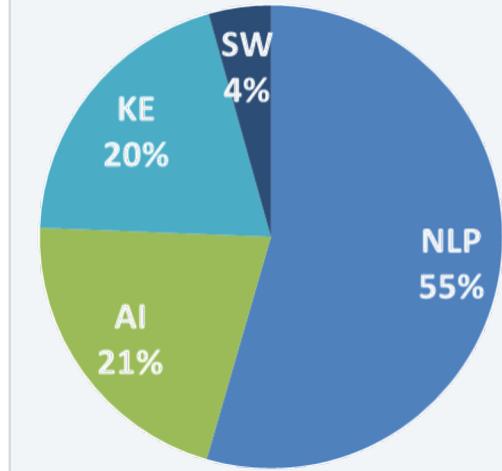
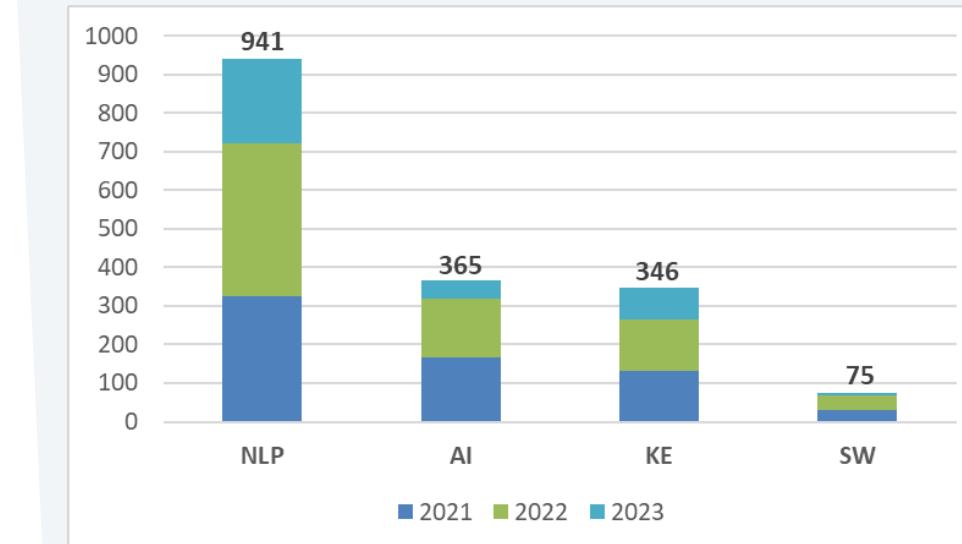
知识工程领域 (KE)

SIGIR(23,22,21), CIKM(22,21),
WSDM(23,22,21), KDD(22,21),
WWW(23,22,21)

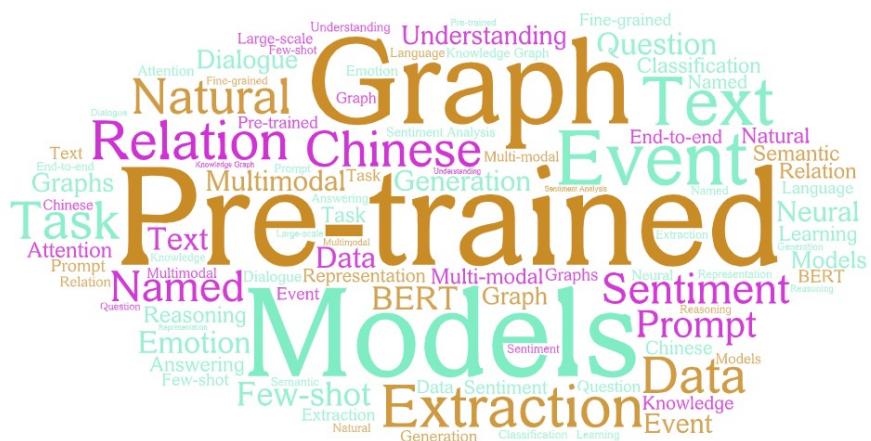
数据库&语义网领域 (SW)

SIGMOD(23,22,21), ISWC(22,21)

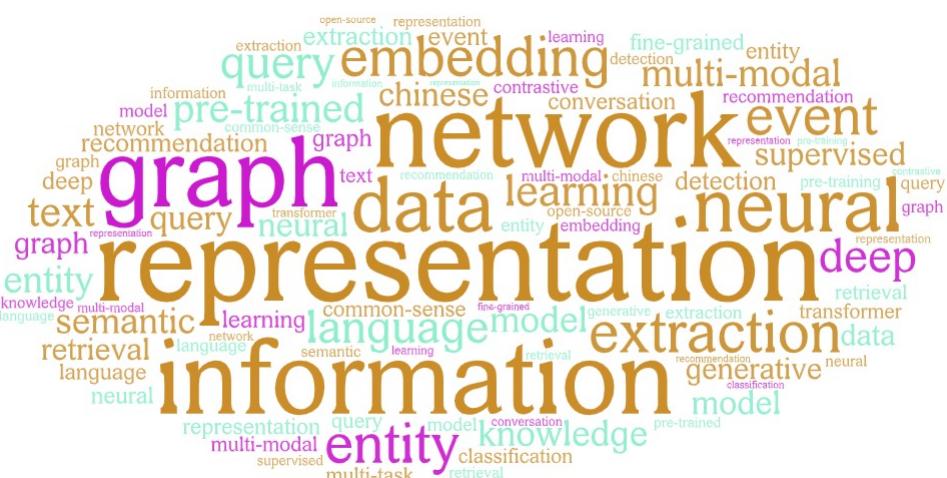
不同领域文献调研



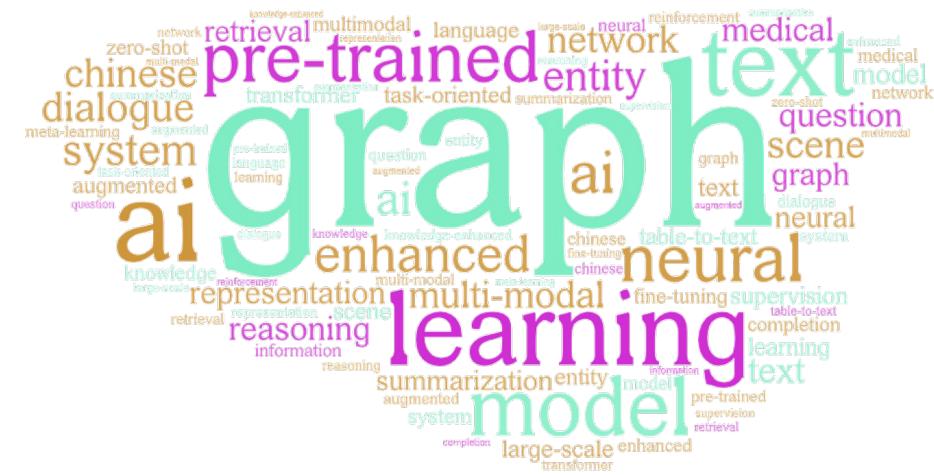
数据源：自动检索 (1399) + 人工查阅 (353) → 1727 篇文献



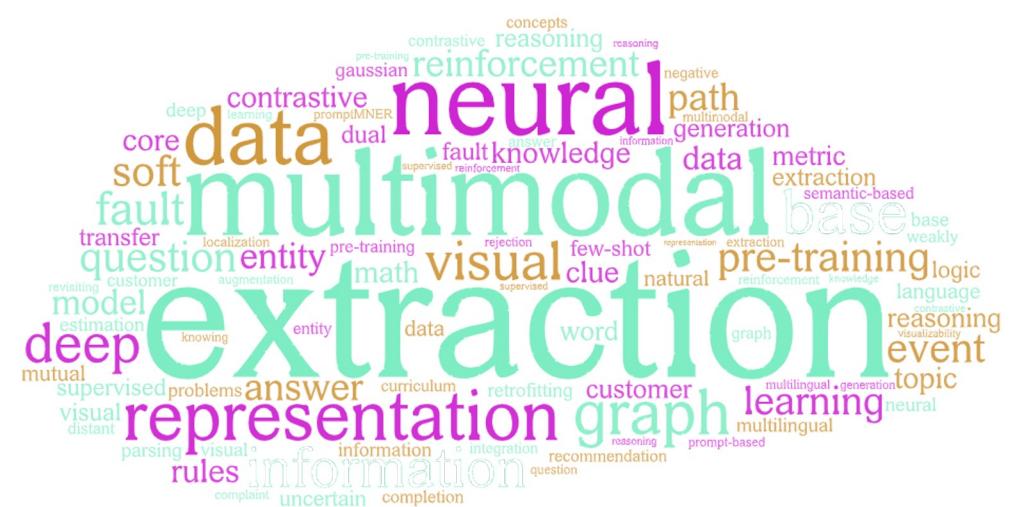
NLP领域



KE领域

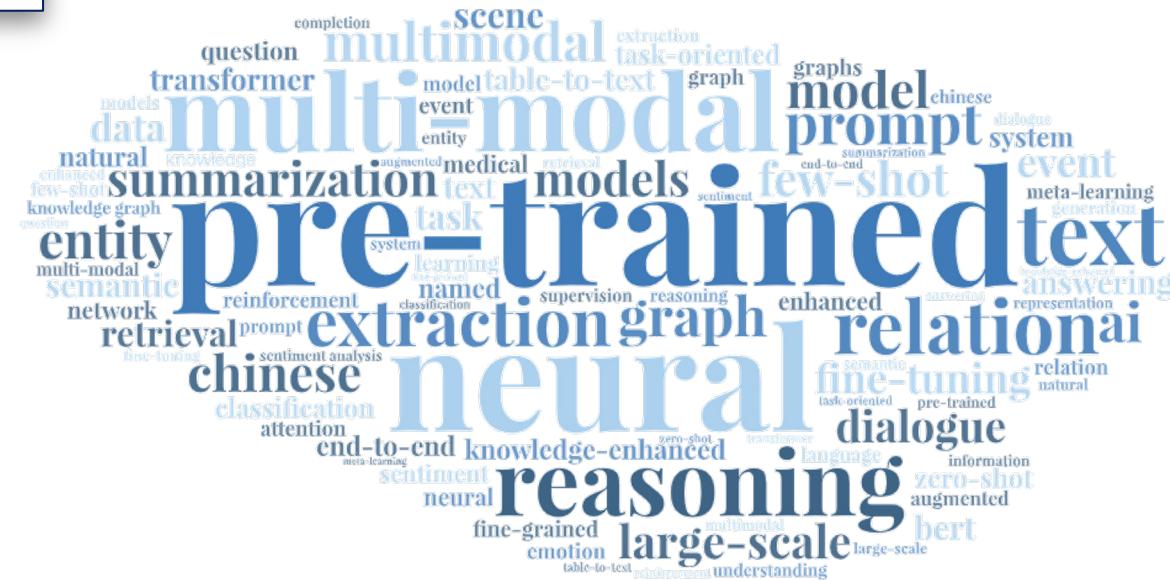


AI领域



SW领域

关键词分析



任务

reasoning, extraction,
multi-modal,
summarization...

技术

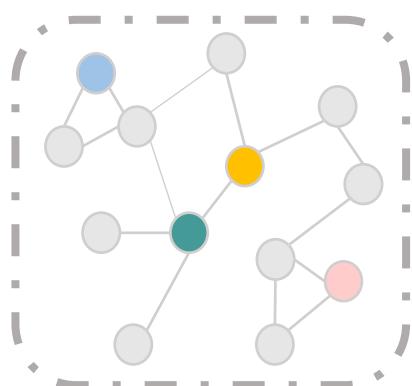
pre-trained,
prompt, zero-shot,
retrieval, end-to-end...

“

- **知识应用 (~79)** : 对话系统、问答系统、推荐系统、代码搜索/补全、情感分析、因果分析、摘要生成、机器阅读理解、信息检索、语义解析
- **预训练模型 (~79)** : 知识注入、知识探测、领域化
- **信息抽取 (~63)** : 实体识别、实体分类、关系抽取、事件抽取
- **知识推理与问答 (~55)** : 开放域问答、抽取式问答、常识问答、图推理、常识/因果推理、视觉推理
- **场景扩展 (~52)** : 知识编辑、持续学习、细粒度、在线学习、时空、跨语言、多模态、增量学习
- **其他 (~42)** : 对抗性攻击、去偏置、细粒度、KG表示学习、区块链、比喻生成、义原预测、传播预测、模型分析、蒸馏

知识图谱构建

知识图谱构建旨在从海量结构化和非结构化数据中进行**实体**、**关系**、**属性**和**事件**等信息的提取、融合，将知识存储到知识库中。



知识图谱

Named Entity Recognition (NER)

Relation Extraction (RE)

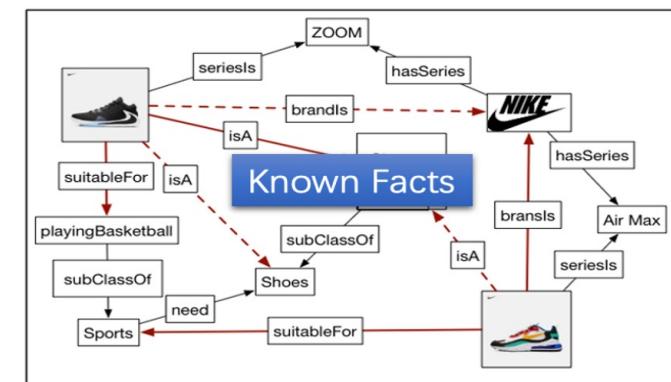
Event Extraction (EE)

Entity Linking (EL)

.....

知识图谱推理

知识图谱推理旨在利用已有的知识，通过逻辑推理、关联推断或机器学习等方法，从知识图谱中推导出**新的知识**。



New Facts
New Relations
New Axioms
New Rules

- 01 **方法范式扩展**：端到端学习、工具调用与神经符号集成、多智能体人机协同
- 02 **处理对象扩展**：模型参数（隐式知识）、视觉、音频、表格
- 03 **任务类型扩展**：复杂知识建模与获取、类比推理
- 04 **评测维度扩展**：系统评测、具身评测

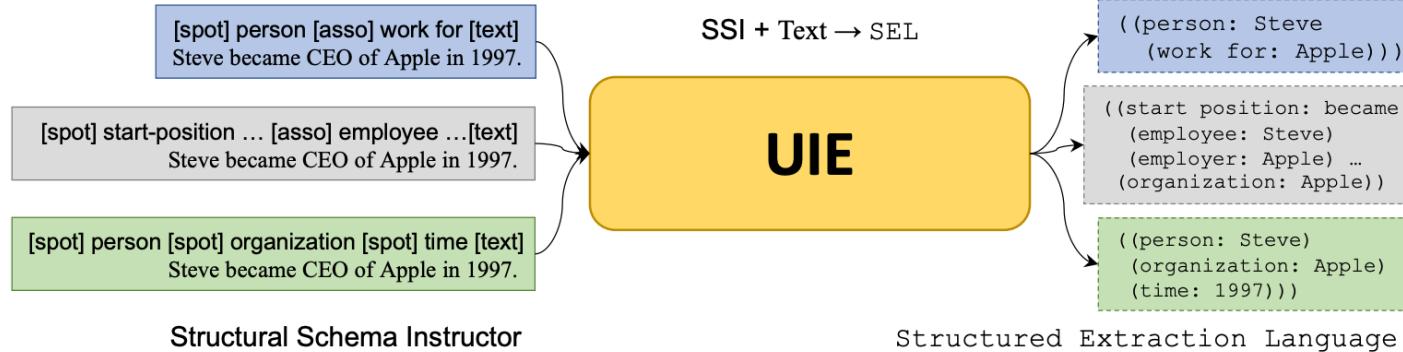
01 方法范式扩展：端到端学习、工具调用与神经符号集成、多智能体人机协同

02 处理对象扩展：模型参数（隐式知识）、视觉、音频、表格

03 任务类型扩展：复杂知识建模与获取、类比推理

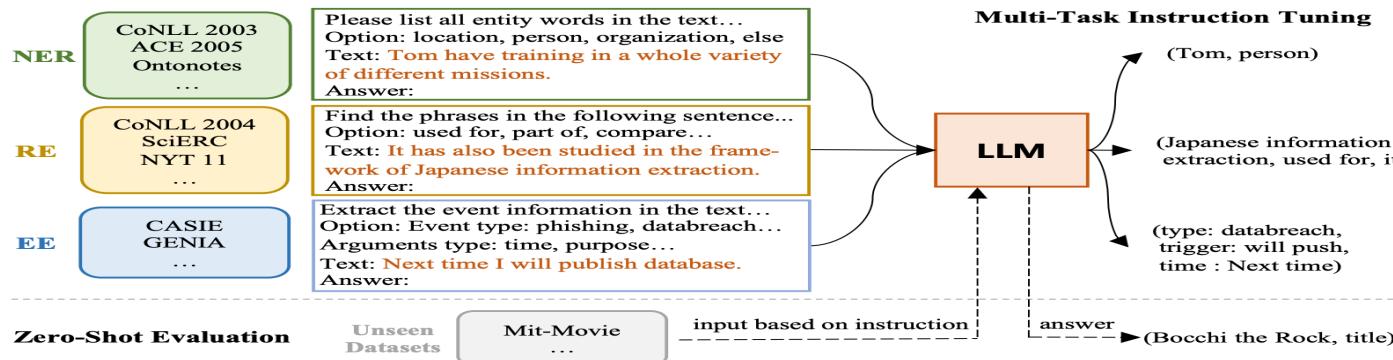
04 评测维度扩展：系统评测、具身评测

UIE: Unified Structure Generation



Task	Schema	Instance								
Entity	PER: _ ORG: _	PER ORG In 1997, Steve was excited to become the CEO of Apple.								
Relation	(_, Work for, _)	Work For In 1997, Steve was excited to become the CEO of Apple.								
Event	<table border="1"> <tr> <th>Type</th> <th>Start Position</th> </tr> <tr> <td>employee</td> <td></td> </tr> <tr> <td>employer</td> <td></td> </tr> <tr> <td>...</td> <td></td> </tr> </table>	Type	Start Position	employee		employer		...		Start-Position In 1997, Steve was excited to become the CEO of Apple. Person → Start-Position → Entity
Type	Start Position									
employee										
employer										
...										
Sentiment	Positive { Opinion: _; Target: _ }	Positive In 1997, Steve was excited to become the CEO of Apple. Opinion → Positive → Target								

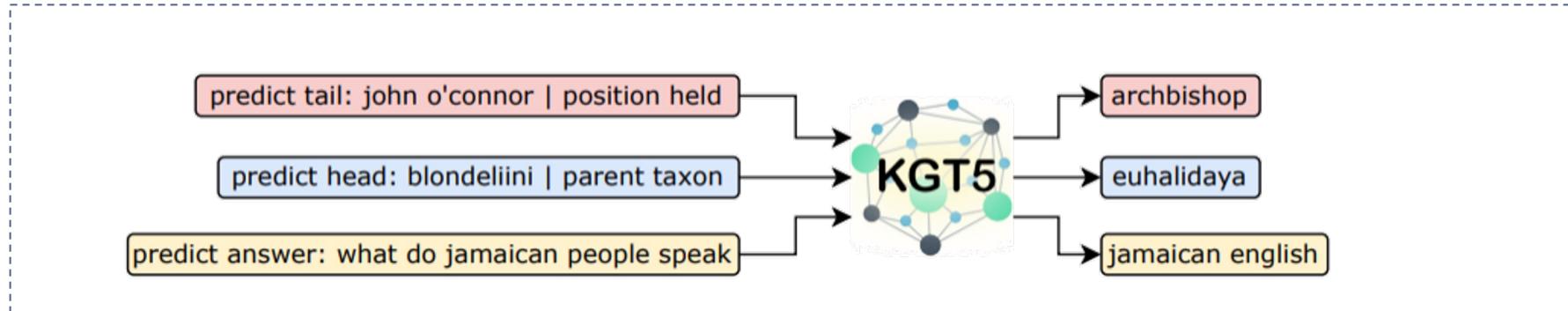
InstructUIE: Multi-task Instruction Tuning



- [1] Unified Structure Generation for Universal Information Extraction. ACL 2022
[2] InstructUIE: Multi-task Instruction Tuning for Unified Information Extraction. 2023

KGT5: A sequence-to-sequence T5-small model

- 把KGC任务转化为一个seq2seq任务



TechGPT: Technology-Oriented Generative Pretrained Transformer

强化：

- 以“知识图谱构建” 为核心的关系三元组抽取等各类信息抽取任务
- 以“阅读理解” 为核心的各类智能问答任务
- 以“文本理解” 为核心的关键词生成等各类序列生成任务



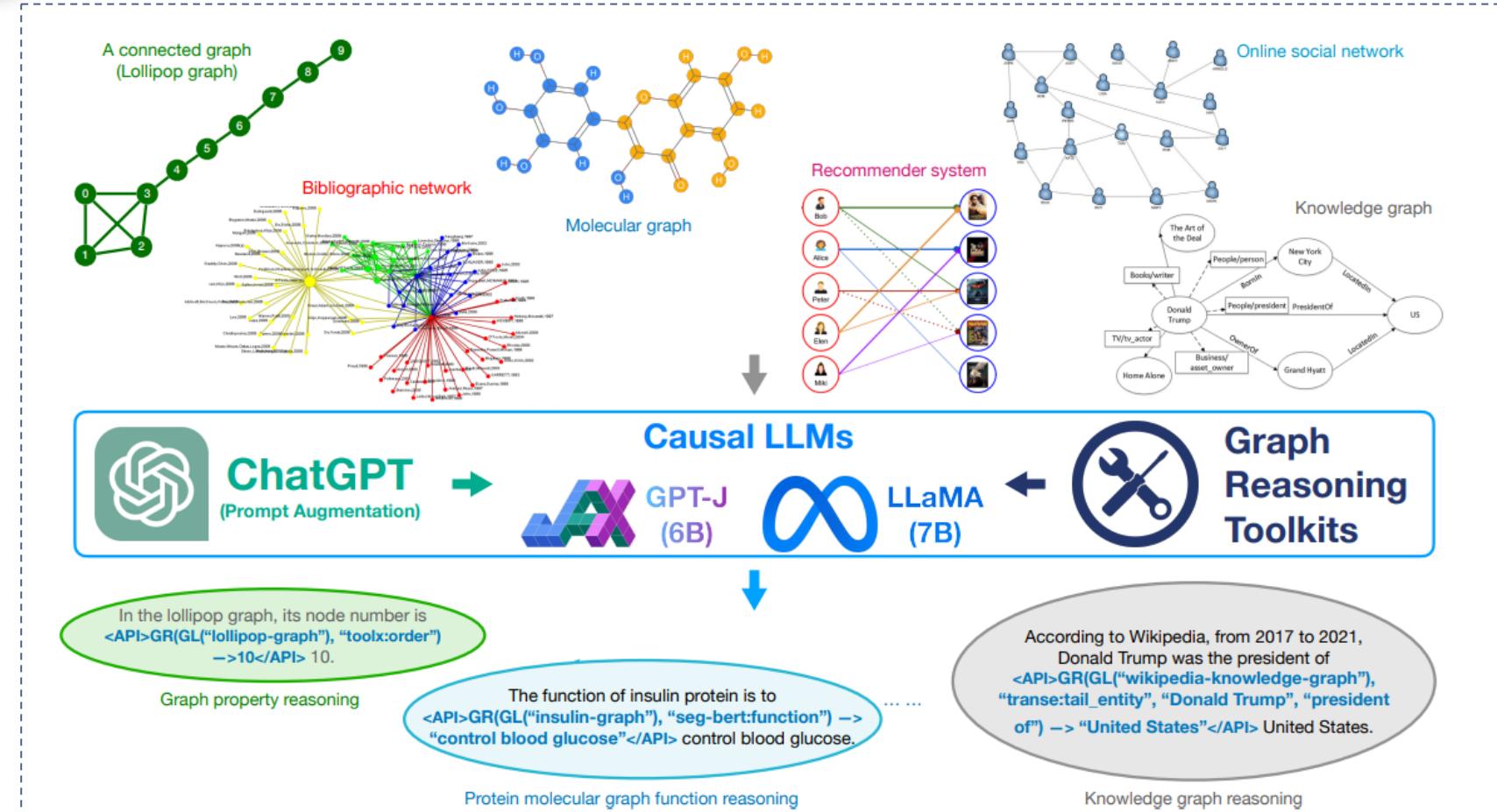
垂直领域自然语言文本的处理

[1] Sequence-to-Sequence Knowledge Graph Completion and Question Answering. ACL 2022

[2] TechGPT: Technology-Oriented Generative Pretrained Transformer. 2023

Graph-ToolFormer

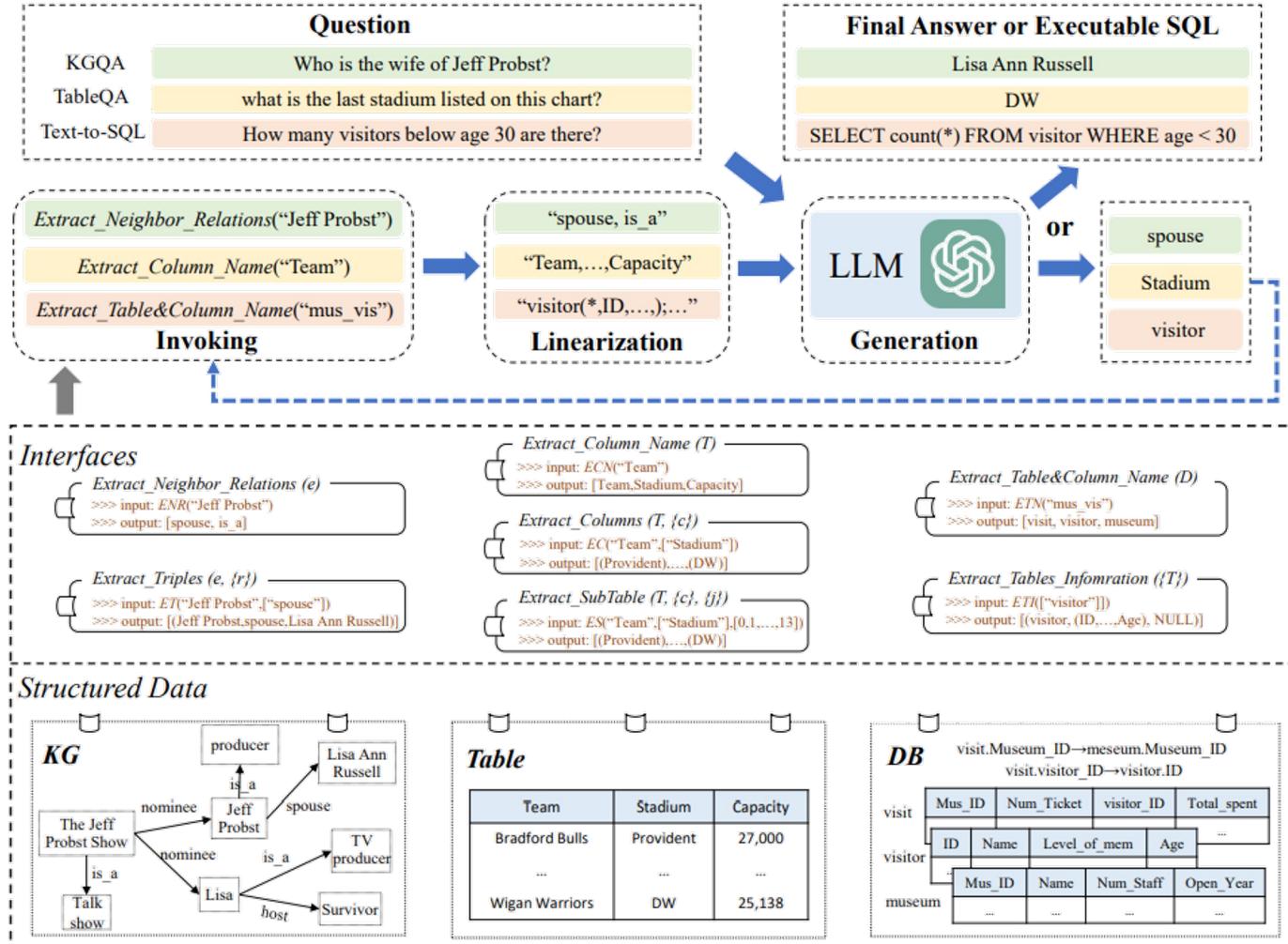
- 使用 LLM 调用处理 Graph 数据的工具
- 完成复杂的图推理等任务



[1] Graph-ToolFormer: To Empower LLMs with Graph Reasoning Ability via Prompt Augmented by ChatGPT. 2023

StructGPT

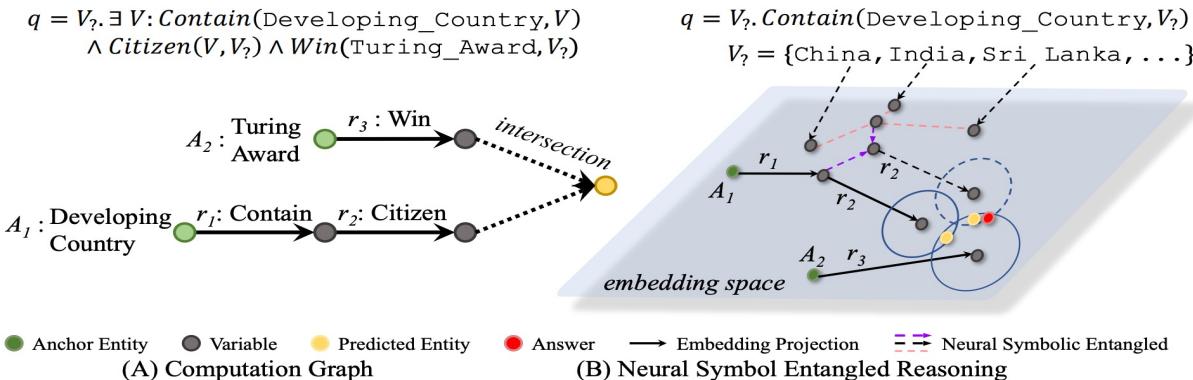
- 符号推理与大模型推理迭代结合
- Iterative Reading-then-Reasoning (IRR) 迭代阅读-推理方法



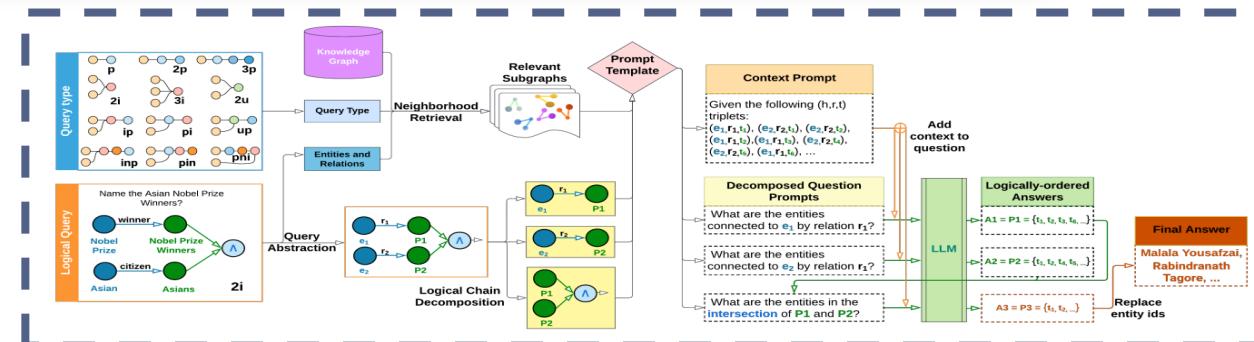
[1] StructGPT: A General Framework for Large Language Model to Reason over Structured Data. 2023

Neural-Symbolic Entangled Framework

- 符号推理与神经网络推理迭代结合
- 将复杂KG的推理解耦为上下文KG搜索与逻辑查询推理的结合



Language-guided Abstract Reasoning



[1] Neural-Symbolic Entangled Framework for Complex Query Answering. NeurIPS 2022

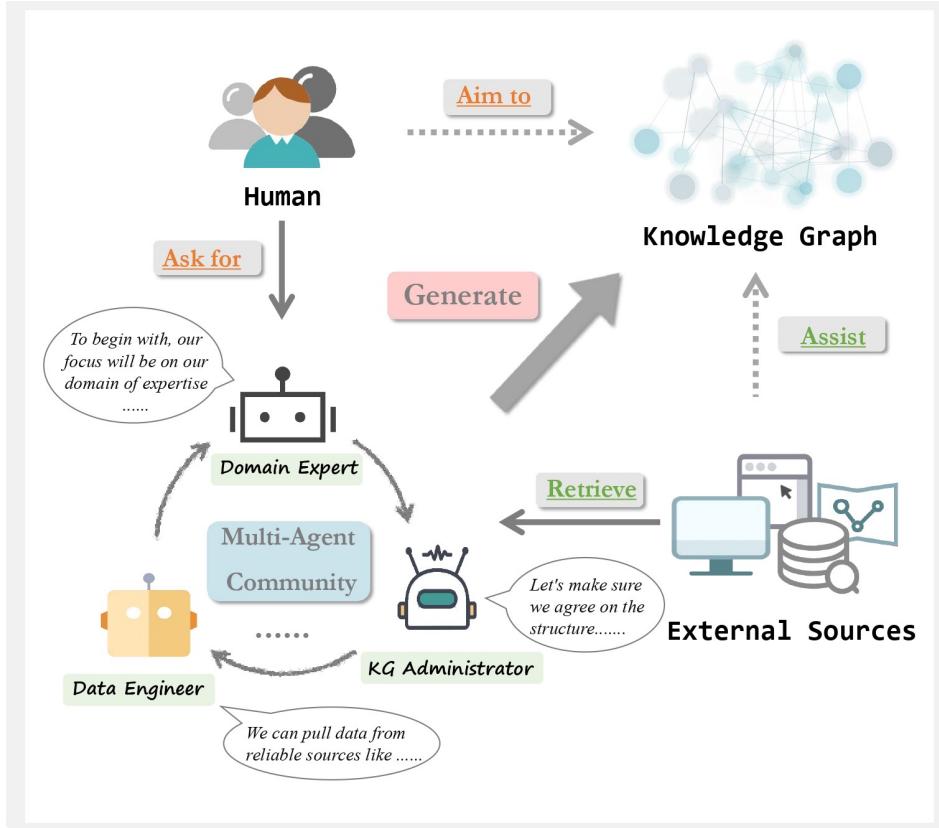
[2] Complex Logical Reasoning over Knowledge Graphs using Large Language Models. 2023

01 方法范式扩展：多智能体人机协同

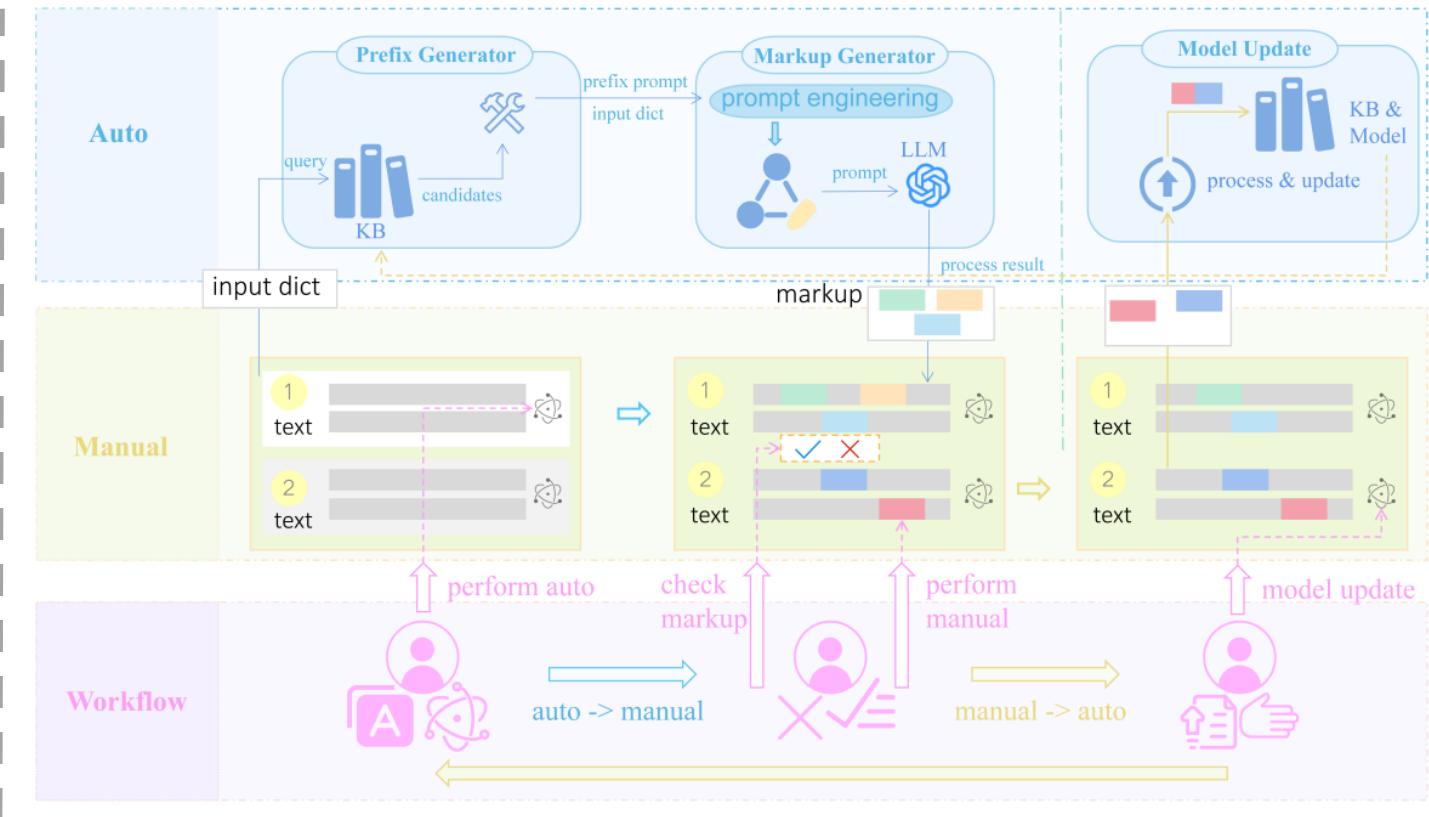
Agent：以LLM为核心，可以与agent、tool、human等协作完成任务

Collaboration

Agent + Agent (Tool) → KG



Agent + Human → IE toolkit



[1] LLMs for Knowledge Graph Construction and Reasoning: Recent Capabilities and Future Opportunities. 2023

[2] CollabKG: A Learnable Human-Machine-Cooperative Information Extraction Toolkit for (Event) Knowledge Graph Construction. 2023

01 方法范式扩展：端到端学习、工具调用与神经符号集成、多智能体人机协同

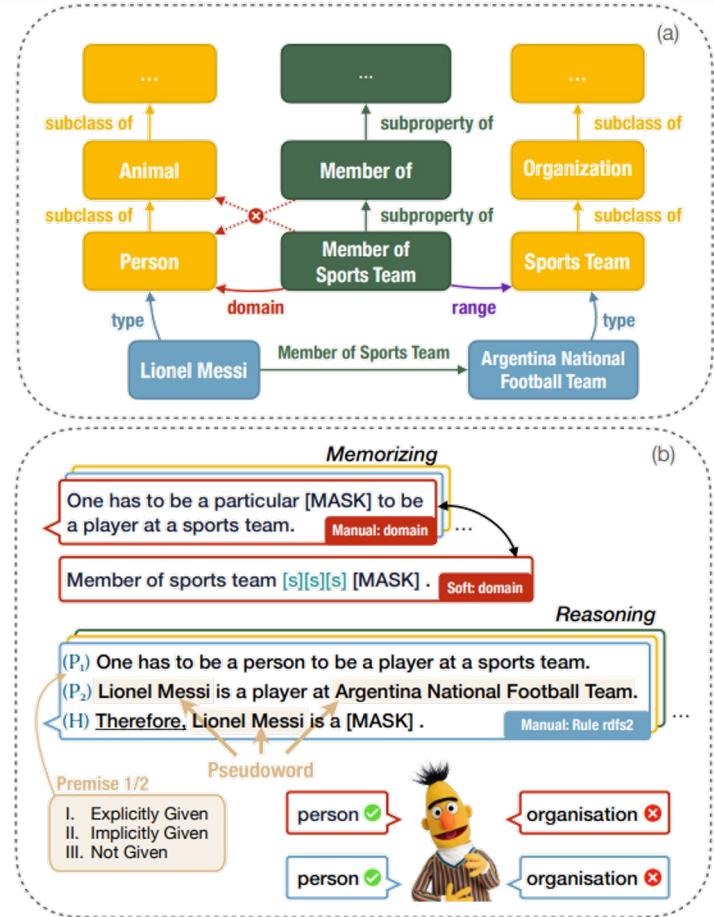
02 处理对象扩展：模型参数（隐式知识）、视觉、音频、表格

03 任务类型扩展：复杂知识建模与获取、类比推理

04 评测维度扩展：系统评测、具身评测

02 处理对象扩展扩展：模型参数（隐式知识）

PLM + Ontological Knowledge



PLM + Knowledge Graph

Text Mining

Albert Einstein, a German theoretical physicist, published the theory of relativity in 1915.

Mechanical Pipelines
NER, CR, RE...

(Albert Einstein, publish, the theory of relativity)

KG Completion (COMET)

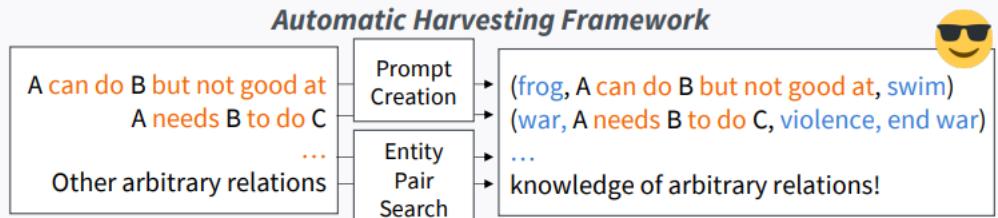
(bridge, UsedFor, cross water)
(bowl, UsedFor, holding popcorn)
(toothpaste, Usedfor, ?)

Prompting Finetuned LMs

LMs trained with existing KG



BertNet (Ours)

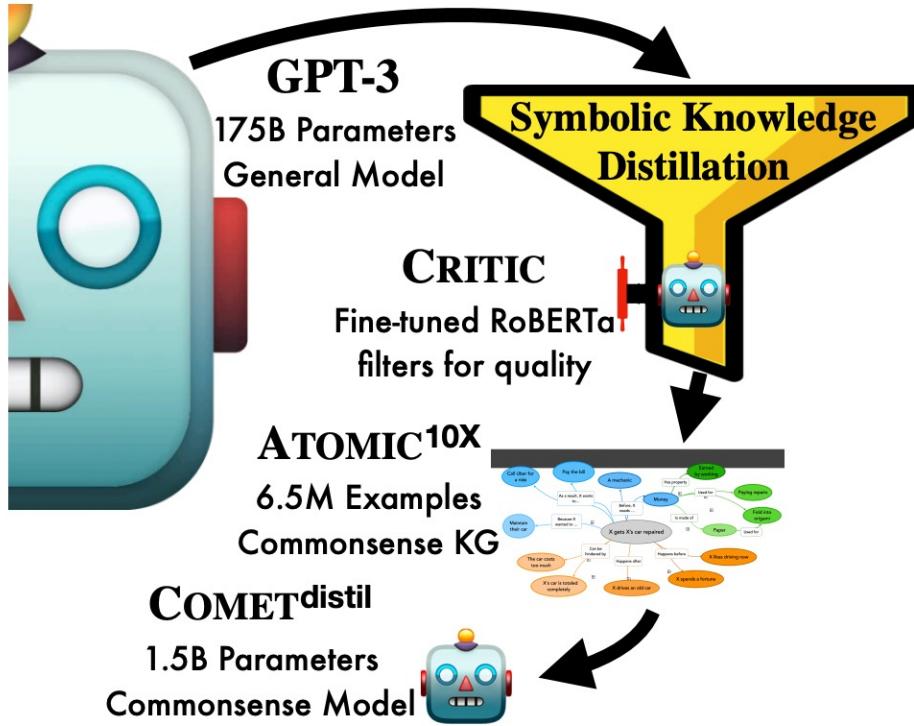


[1] Do PLMs Know and Understand Ontological Knowledge? ACL 2023

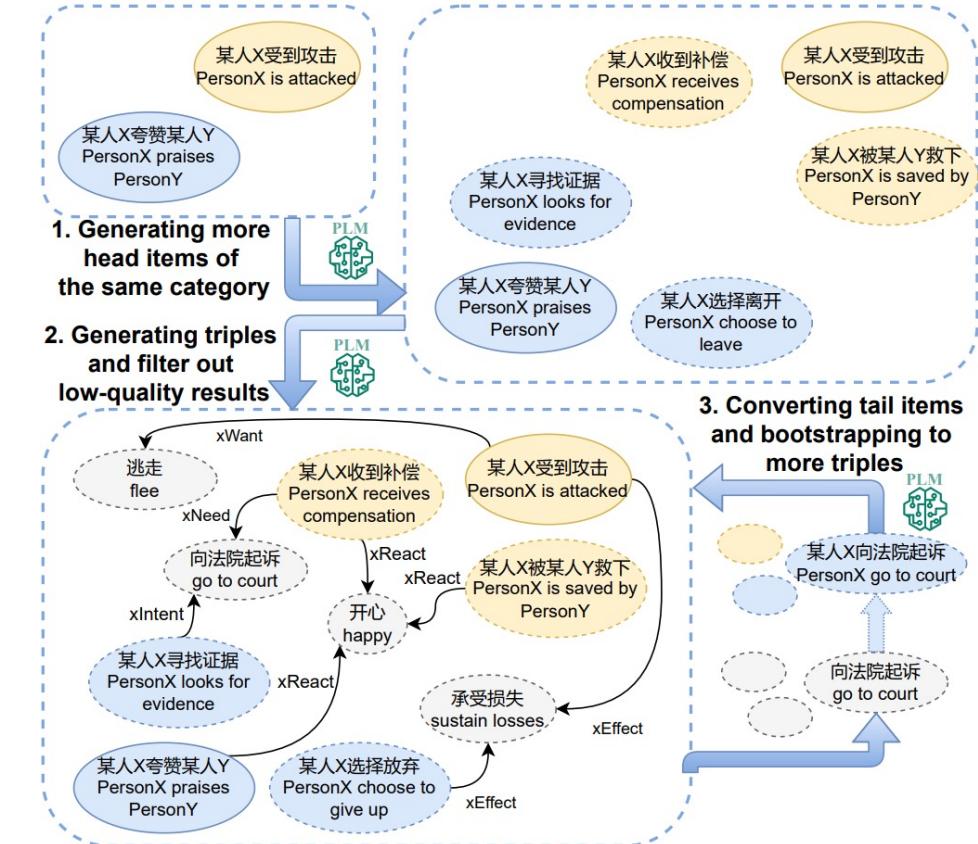
[2] BertNet: Harvesting Knowledge Graphs with Arbitrary Relations from Pretrained Language Models. ACL 2023 Findings

02 处理对象扩展扩展：模型参数（隐式知识）

PLM + Commonsense KG



PLM + Commonsense KG (CN)

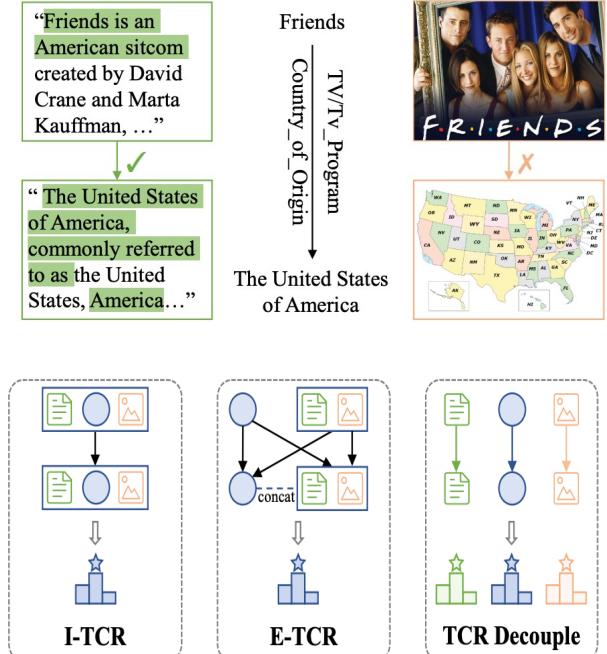
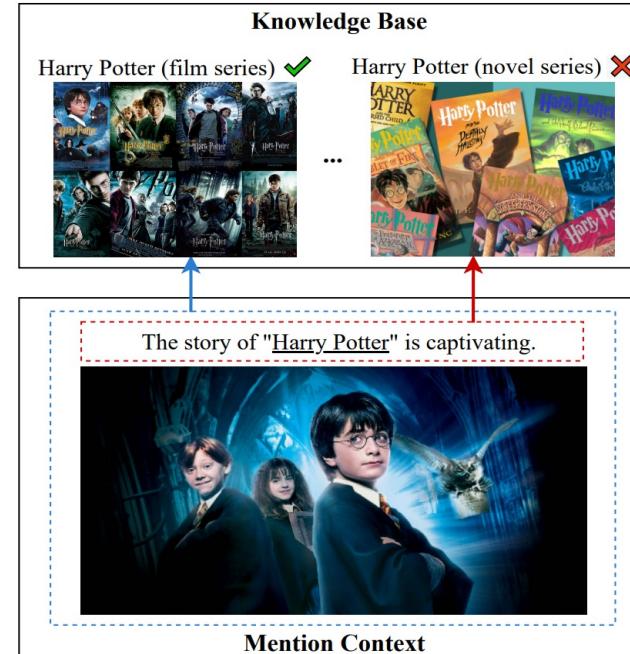
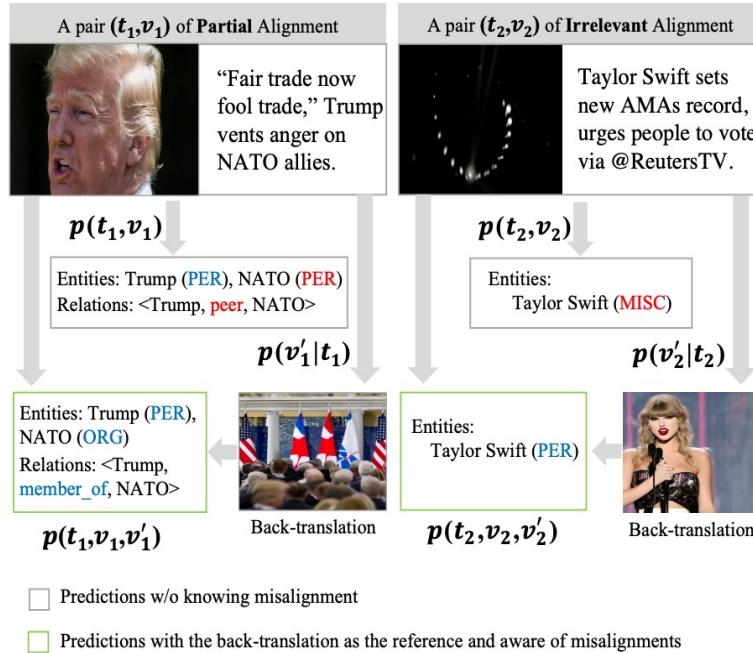


[1] Symbolic Knowledge Distillation: from General Language Models to Commonsense Models. NAACL 2022

[2] CN-AutoMIC: Distilling Chinese Commonsense Knowledge from Pretrained Language Models. EMNLP 2022

02 处理对象扩展：视觉

利用语言模型解决多模态实体、关系抽取链接与推理任务



多模态实体关系抽取

多模态实体链接

多模态知识图谱推理

[1] Rethinking Multimodal Entity and Relation Extraction from a Translation Point of View. ACL 2023

[2] Generative Multimodal Entity Linking. 2023

[3] MoSE: Modality Split and Ensemble for Multimodal Knowledge Graph Completion. EMNLP 2022

02 处理对象扩展：音频

SpeechRE

从音频数据中提取实体之间语义关系的问题

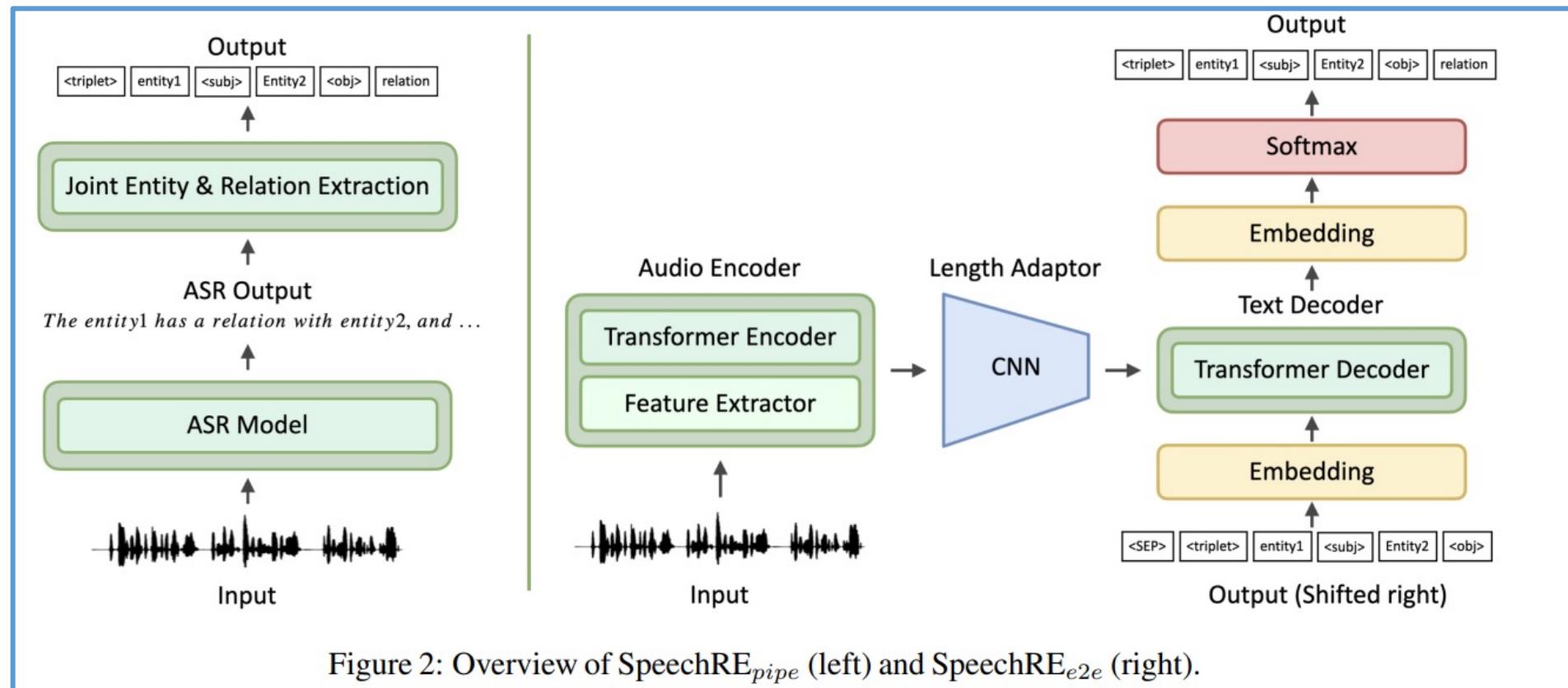
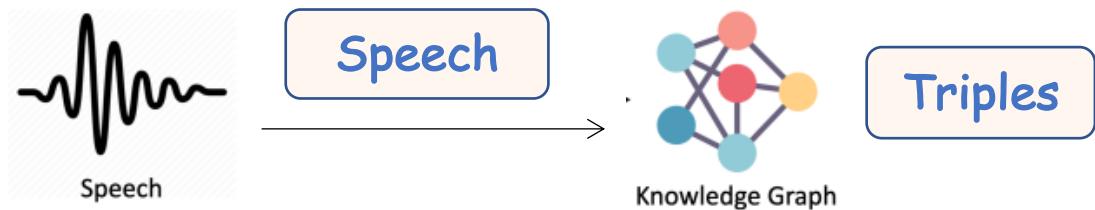


Figure 2: Overview of $\text{SpeechRE}_{\text{pipe}}$ (left) and $\text{SpeechRE}_{\text{e2e}}$ (right).

02 处理对象扩展：表格

Table to KG

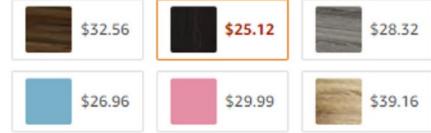
从表格数据中提取实体和关系构建知识图谱的问题

FURINNO Simplistic Study Table,
Espresso

Visit the Furinno Store
★★★★★ 11,604 ratings
Amazon's Choice for "table"

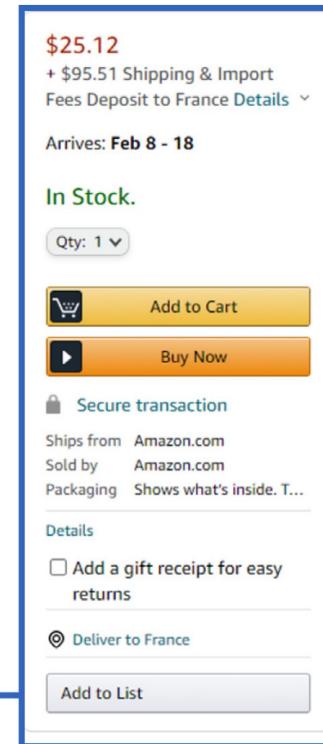
List Price: \$59.99
Price: **\$25.12** + \$95.51 Shipping & Import Fees Deposit to France Details
You Save: **\$34.87 (58%)**

Color: Espresso



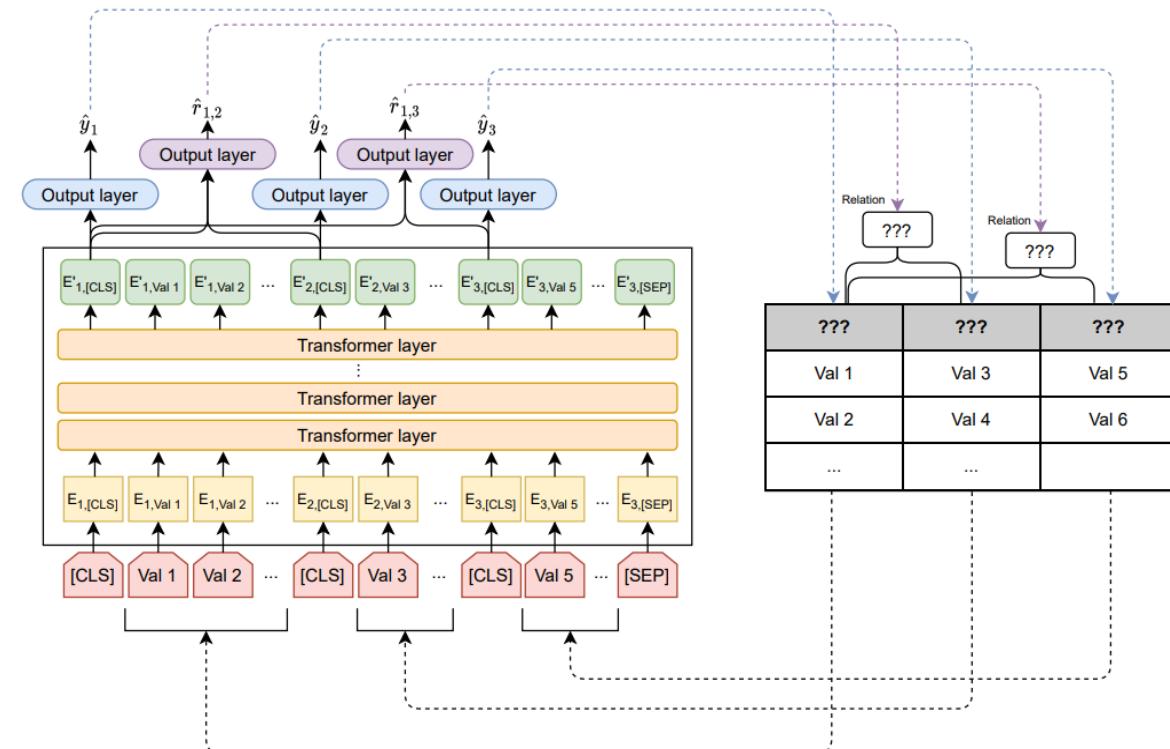
Pattern Name: Table

Color	Espresso
Material	Engineered Wood
Furniture Finish	Espresso
Shape	Rectangle
Number of Shelves	1



Genuine table

Layout table



[1] Annotating Columns with Pre-trained Language Models. SIGMOD 2022

[2] From tabular data to knowledge graphs: A survey of semantic table interpretation tasks and methods. JWS 2023

01 方法范式扩展：端到端学习、工具调用与神经符号集成、多智能体人机协同

02 处理对象扩展：模型参数（隐式知识）、视觉、音频、表格

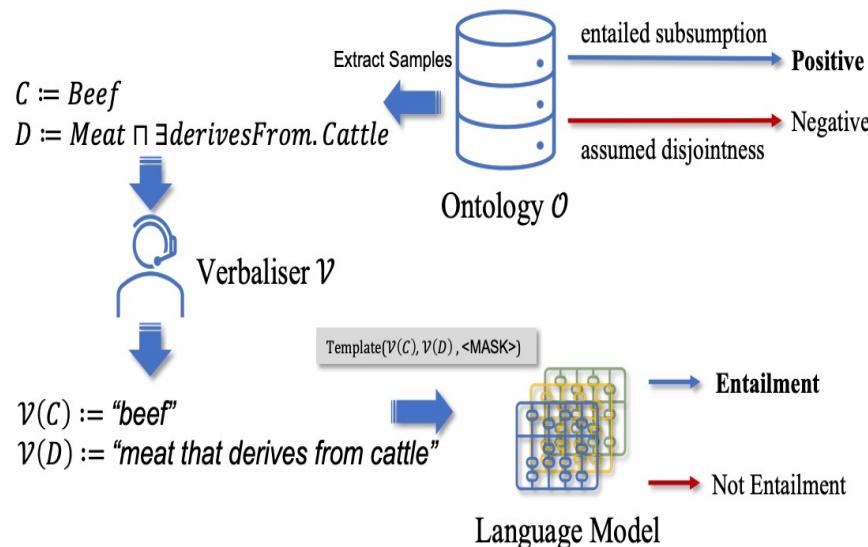
03 任务类型扩展：复杂知识建模与获取、类比推理

04 评测维度扩展：系统评测、具身评测

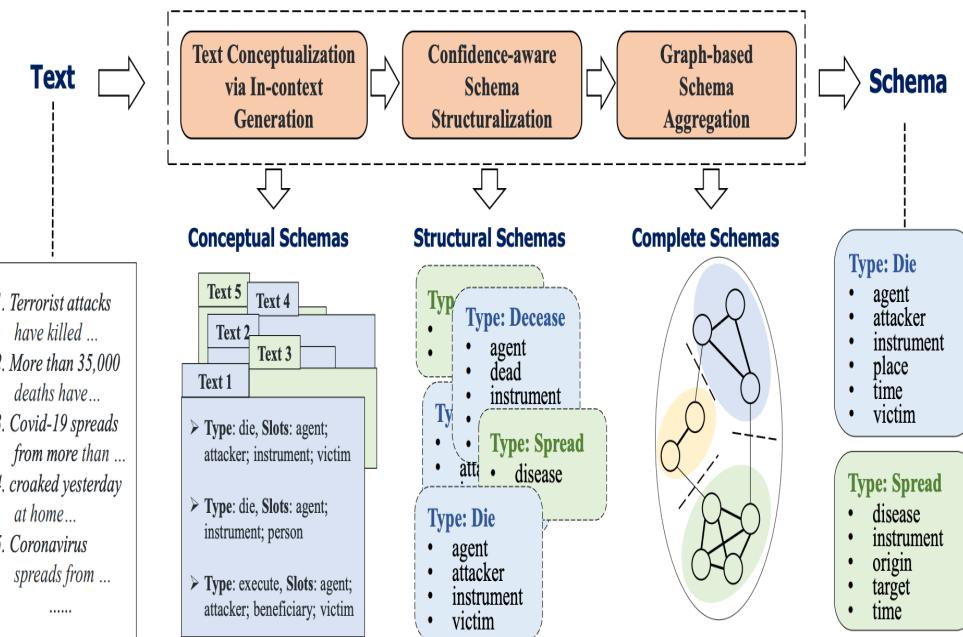
03 任务类型扩展：复杂知识建模与获取

本体知识

利用语言模型建模复杂本体和获取事件本体



复杂本体建模



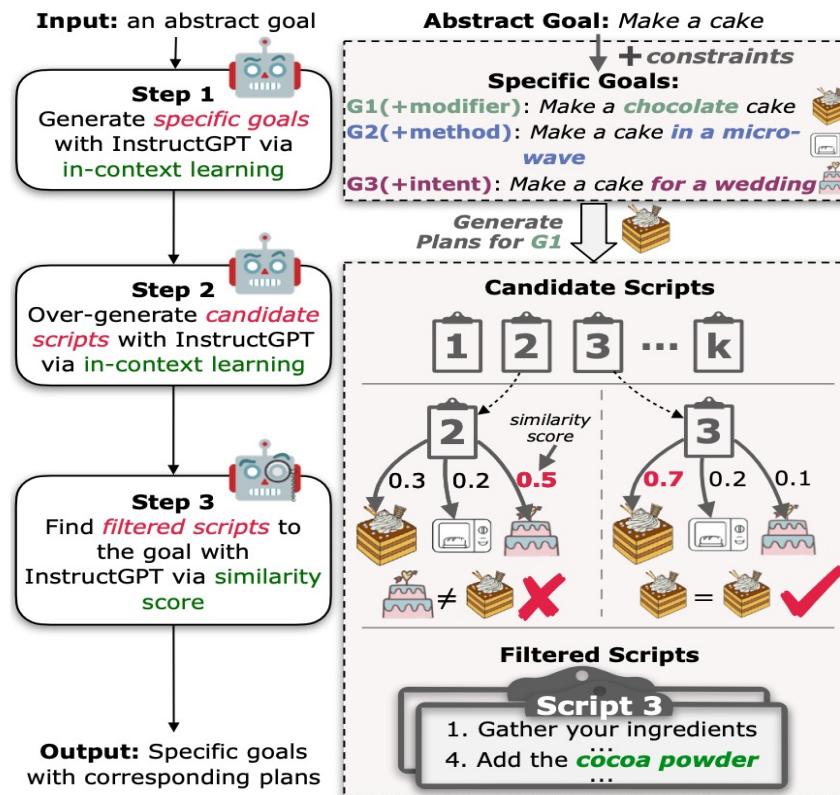
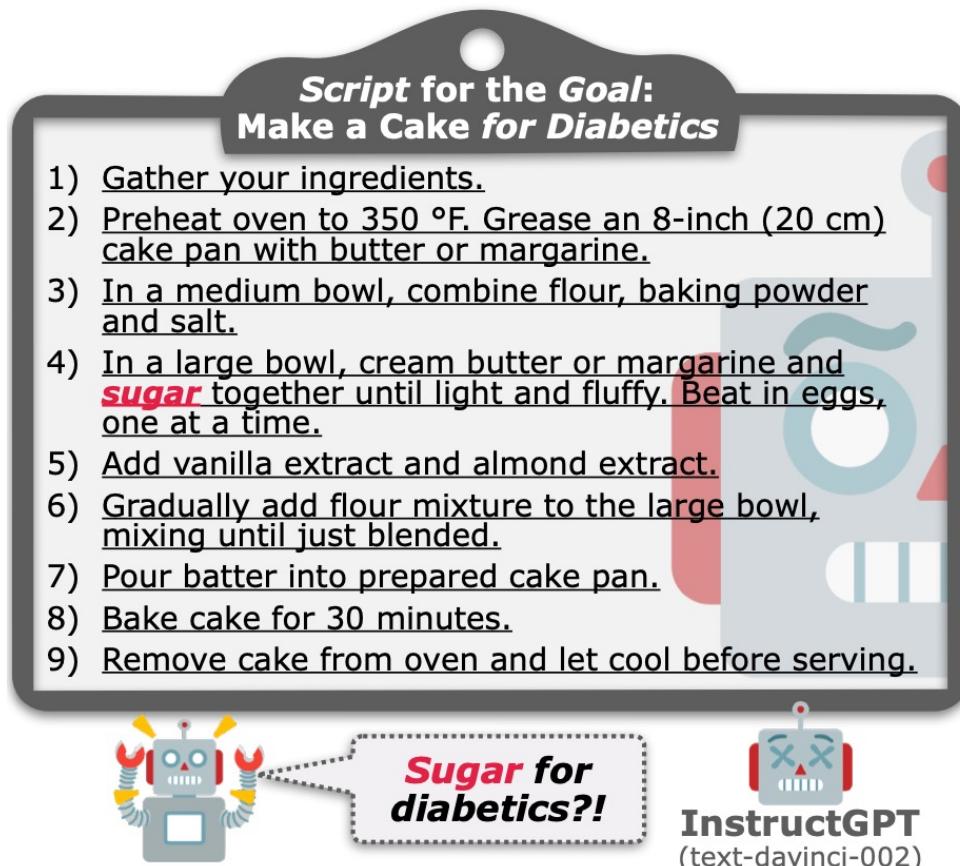
事件本体获取

- [1] Language Model Analysis for Ontology Subsumption Inference. ACL 2023 Findings
[2] Harvesting Event Schemas from Large Language Models. 2023

03 任务类型扩展：复杂知识建模与获取

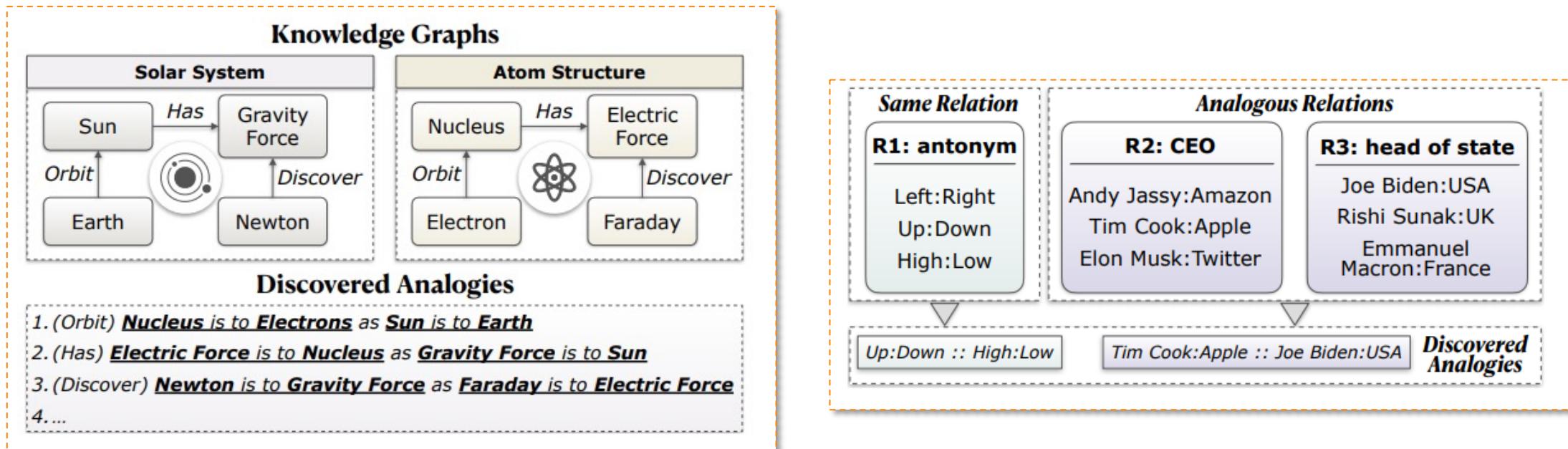
脚本知识

利用语言模型获取脚本知识为专业化小型模型赋予约束语言规划能力



ANALOGYKB

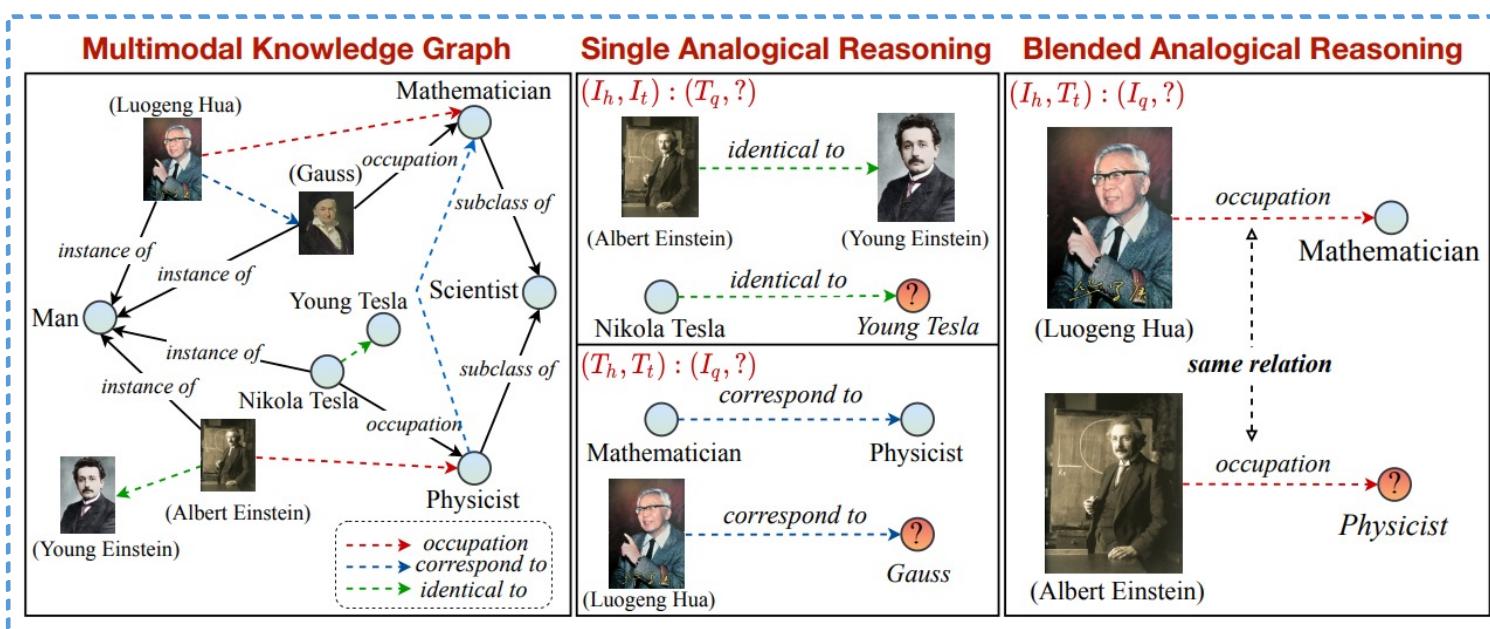
- 由现有的知识图谱 (KG) 衍生而来的一个**百万规模的类比知识库** (KB)
- **两类类比**: 相同关系的类比、类似关系的类比



03 任务类型扩展：类比推理

MKG_Analogy

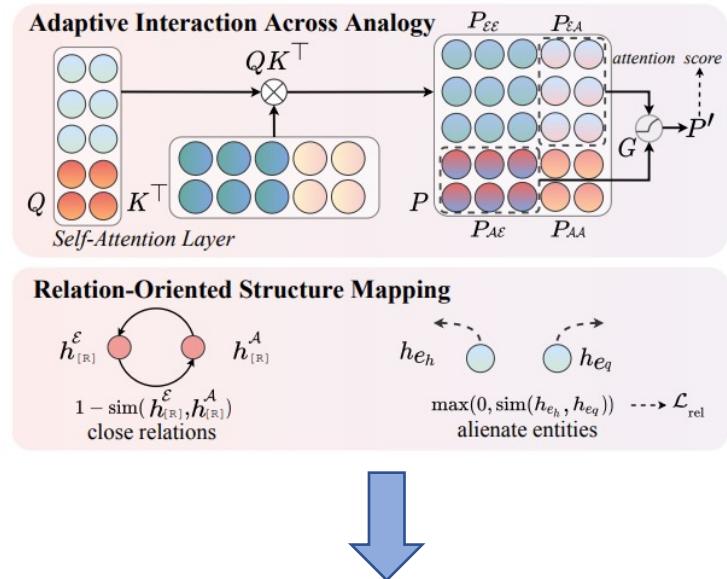
- 单一模态类比推理
 - 混合模态类比推理



引入类比知识进行提示学习提升推理性能

[1] Multimodal Analogical Reasoning over Knowledge Graphs. ICLR 2023

基于Structure Mapping理论模拟多模态类比过程



Model	Hits@1	Hits@3	Hits@5	Hits@10	MRR
TransAE	0.203	0.233	0.241	0.253	0.223
w/o MarKG	0.191	0.224	0.235	0.245	0.214
MarT_ViLBERT	0.256	0.312	0.327	0.347	0.292
w/o MarKG	0.253	0.292	0.297	0.310	0.270
w/o Analogy example	0.113	0.143	0.162	0.179	0.138
MarT_MKGformer	0.301	0.367	0.380	0.408	0.341
w/o MarKG	0.270	0.305	0.309	0.315	0.289
w/o Relaxation loss	0.295	0.349	0.373	0.399	0.332
w/o Adaptive interaction	0.285	0.345	0.365	0.395	0.324
w/o MarT	0.293	0.335	0.344	0.367	0.321
w/o Analogy example	0.101	0.123	0.132	0.149	0.120

01 方法范式扩展：端到端模型、工具插件化、多智能体人机协同

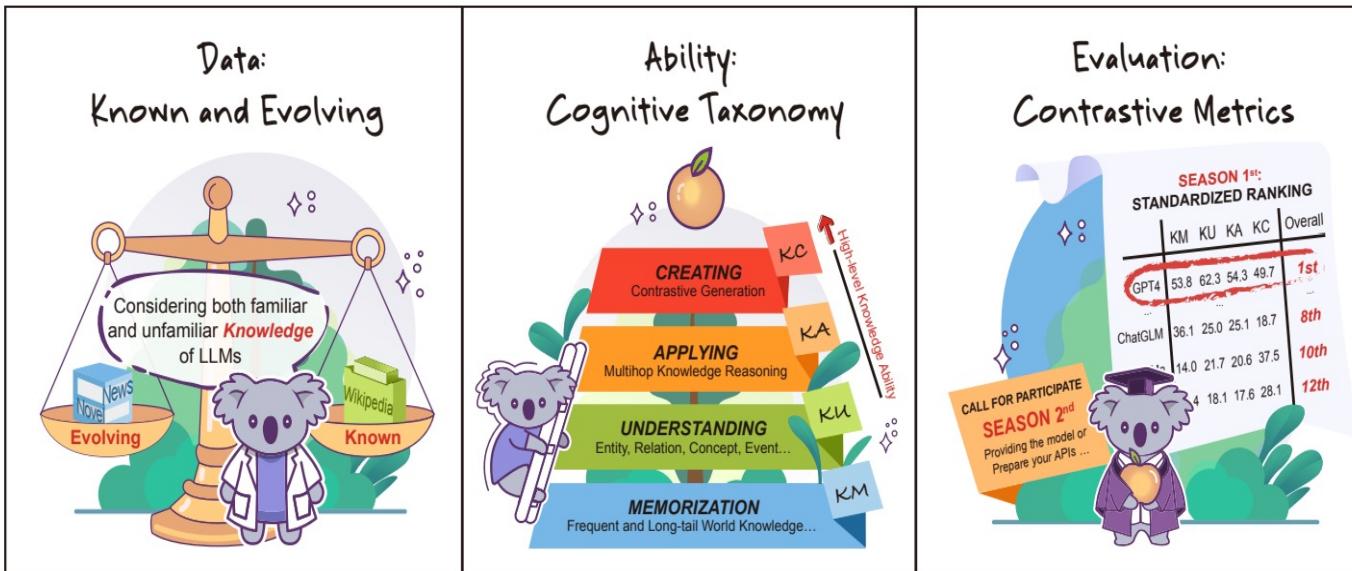
02 处理对象扩展：模型参数（隐式知识）、视觉、音频、表格

03 任务类型扩展：复杂知识建模与获取、类比推理

04 评测维度扩展：系统评测、具身评测

KoLA

- 从知识的记忆、理解、应用和创造4个层级，深度衡量大语言模型**处理世界知识**的能力
- GPT-4虽然很强，但依然未能霸榜，在**知识创造**层次的测试中仅排第三名

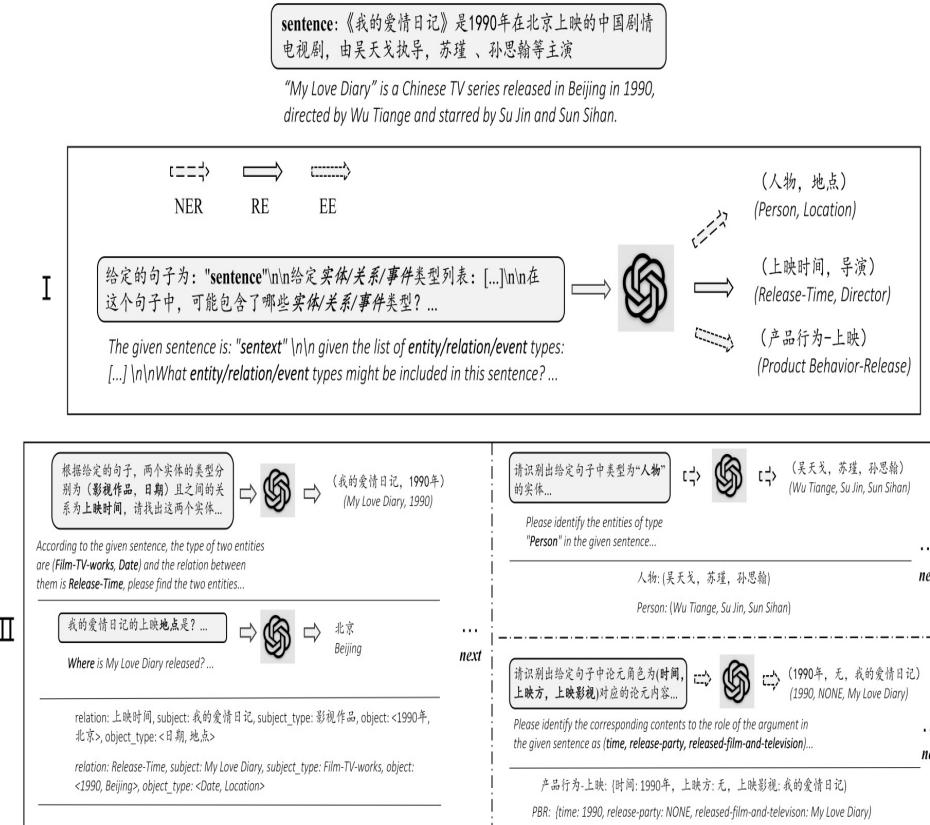


Level	ID	Dataset	Metrics	Exclusive	Context Type	Test Set	Pool	Source
KM	1-1	High-Freq.	EM, F1	✓	Triple	100	20.6M	Known
	1-2	Low-Freq.	EM, F1	✓	Triple	100	20.6M	
	1-3	RTM	EM, F1	✓	Triple	100	2.7k	Evolving
KU	2-1	COPEN-CSJ	Acc.	✓	Entity, Concept	100	3.9k	Known
	2-2	COPEN-CPJ	Acc.	✓	Concept	100	4.7k	
	2-3	COPEN-CiC	Acc.	✓	Concept	100	2.3k	
	2-4	FewNERD	F1	✗	Sentence	300	188.2k	
	2-5	DocRED	F1	✓	Document, Entity	100	12k	
	2-6	MAVEN	F1	✓	Document	100	20.4k	
	2-7	MAVEN-ERE	F1	✓	Document(s), Event	199	1.3M	
KA	2-8	ETU	F1	✓	Document, Entity	100	1.6k	Evolving
	3-1	HotpotQA	F1	✗	Document(s)	100	7.4k	Known
	3-2	2WikiMulti.	F1	✓	Document(s)	100	12.6k	
	3-3	MuSiQue	F1	✓	Document(s)	100	2.5k	
	3-4	KQA Pro	F1	✓	KG	100	1.2k	
KC	3-5	KoRC	F1	✓	Document(s), KG	100	5.2k	Known
	3-6	ETA	F1	✓	Document(s), KG	49	1.6k	
	4-1	Encyclopedic	BLEU, Rouge	✓	Document, Event	95	4.5k	
KC	4-2	ETC	BLEU, Rouge	✓	Document, Event	95	100	Evolving

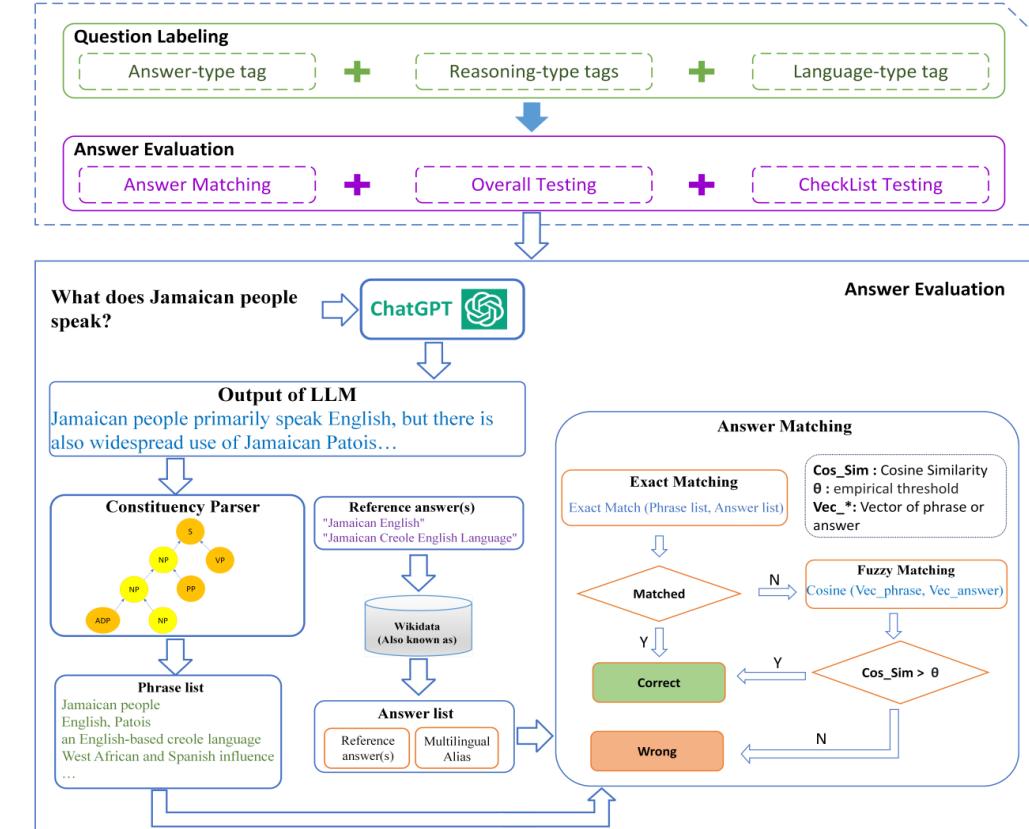
[1] KoLA: Carefully Benchmarking World Knowledge of Large Language Models. 2023

04 任务类型扩展：系统评测

PLM + 知识图谱构建



PLM + 知识图谱推理与问答



[1] Zero-Shot Information Extraction via Chatting with ChatGPT. ACL 2023

[2] Evaluating ChatGPT's Information Extraction Capabilities: An Assessment of Performance, Explainability, Calibration, and Faithfulness. 2023

[3] Evaluation of ChatGPT as a question answering system for answering complex questions. ISWC2023

AgentBench

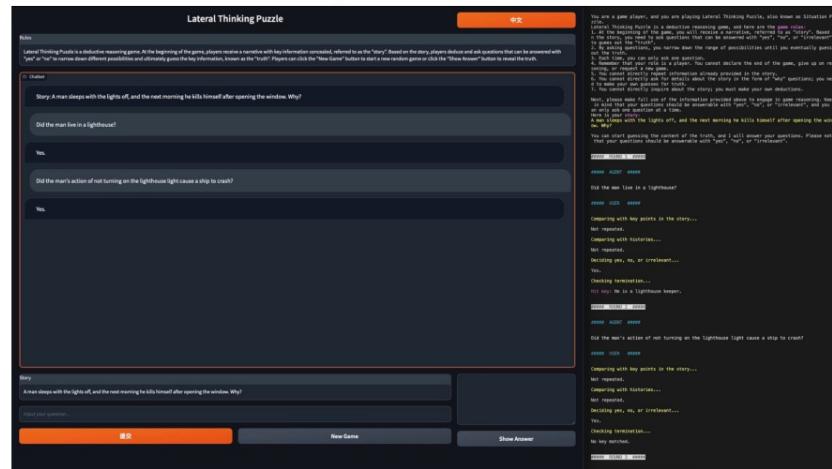
- 一个多维度的演进基准，目前由 8 个不同的环境组成
- 评估 LLM 作为代理在**多轮开放式生成环境（知识图谱）**中的推理和决策能力

(c) Knowledge Graph (KG)

Task: “Find tropical cyclones that are similar to Hurricane Marie and affected Eastern North America.”

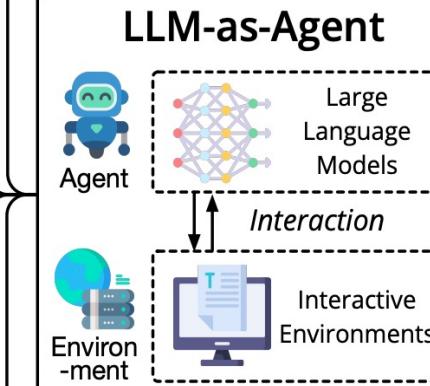
Action space: Basic KG-querying tools

Observation: Query results



Real-world Challenges

- (On an Ubuntu bash terminal)
Recursively set all files in the directory to read-only, except those of mine.
- (Given Freebase APIs)
What musical instruments do Minnesota-born Nobel Prize winners play?
- (Given MySQL APIs and existed tables)
Grade students over 60 as PASS in the table.
- (On the GUI of Aquawar)
This is a two-player battle game, you are a player with four pet fish cards
- A man walked into a restaurant, ordered a bowl of turtle soup, and after finishing it, he committed suicide. Why did he do that?
- (In the middle of a kitchen in a simulator)
Please put a pan on the dinning table.
- (On the official website of an airline)
Book the cheapest flight from Beijing to Los Angeles in the last week of July.



8 Distinct Environments



趋势一：传统三元组事实知识图谱的研究方法趋于统一

趋势二：复杂知识图谱研究方法呈现**工具调用与神经符号集成**，并渐有向**具身智能体交互**发展的趋势

业务驱动的知识图谱构建与推理应用



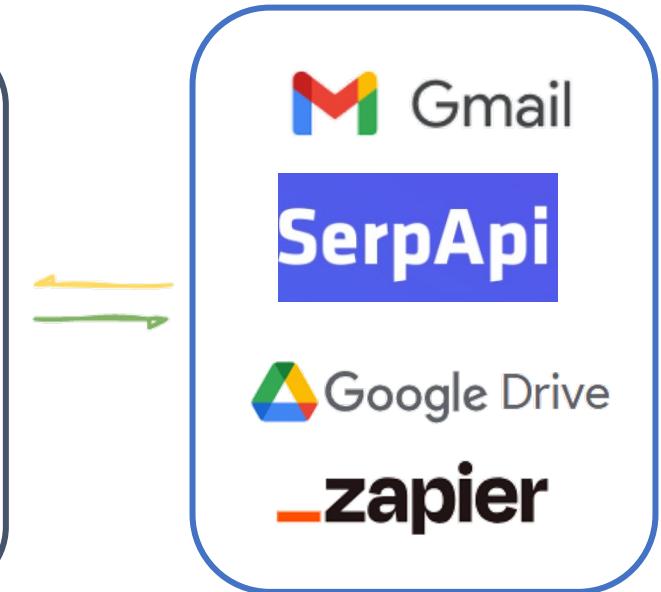
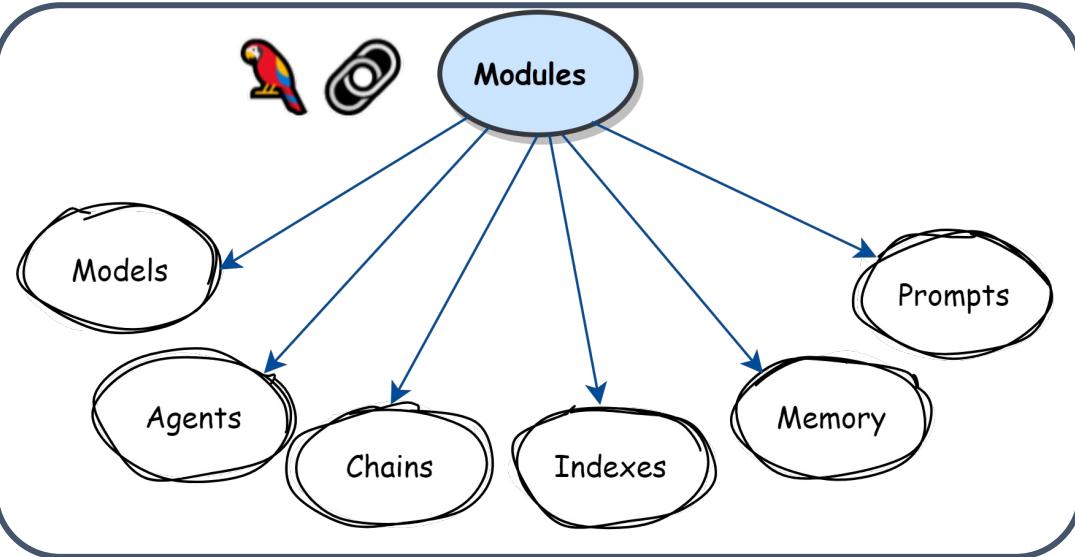
LangChain



LlamaIndex



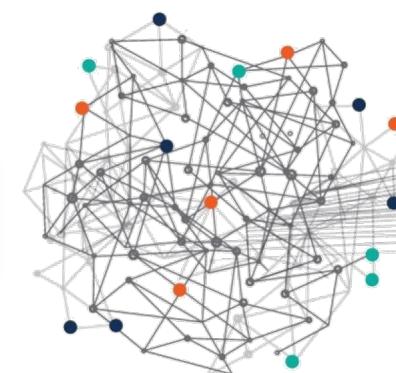
大模型



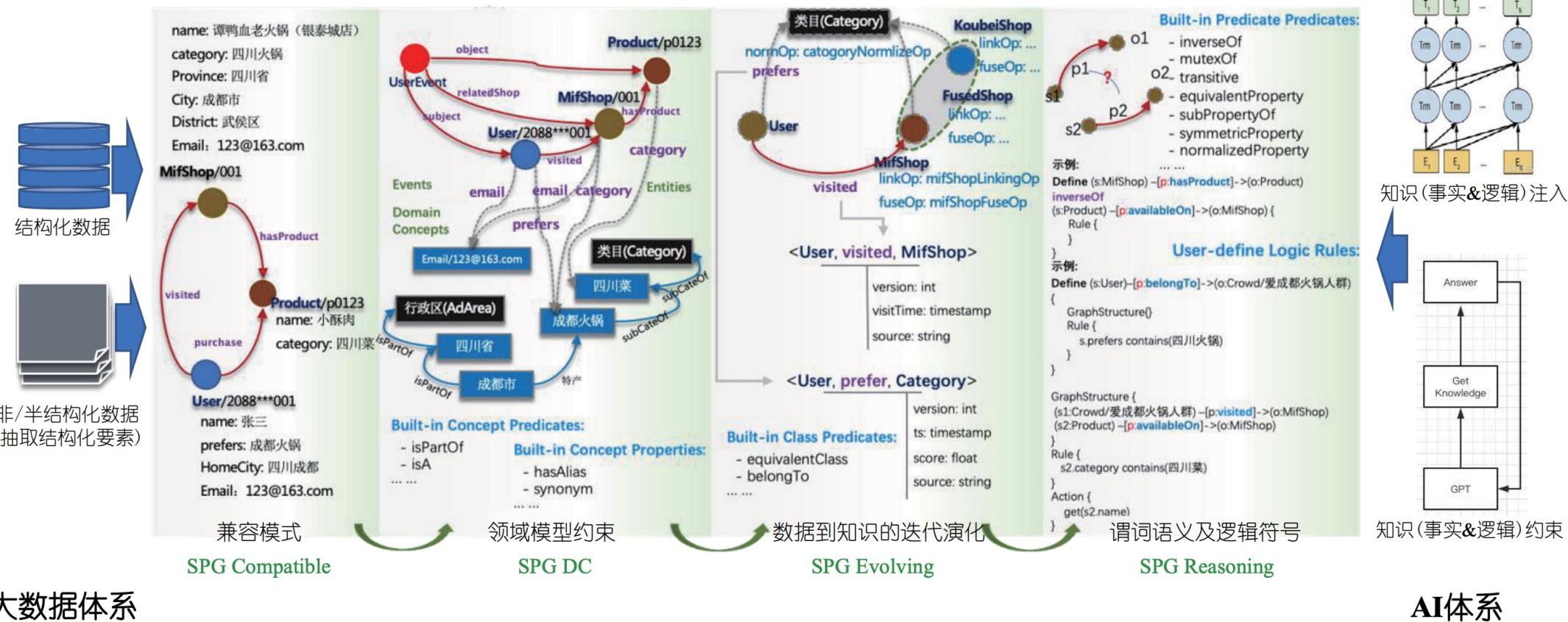
外部数据源

KG构建、推理与应用

知识图谱



衔接大数据与AI技术体系，帮助机器更好地理解世界



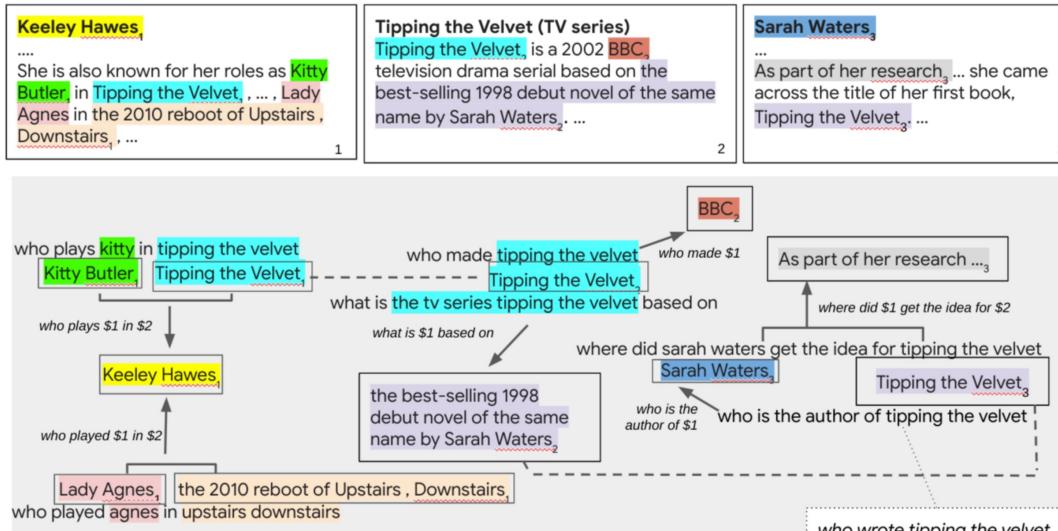
蚂蚁和OpenKG共建的工业级语义增强框架SPG

更多工业届应用

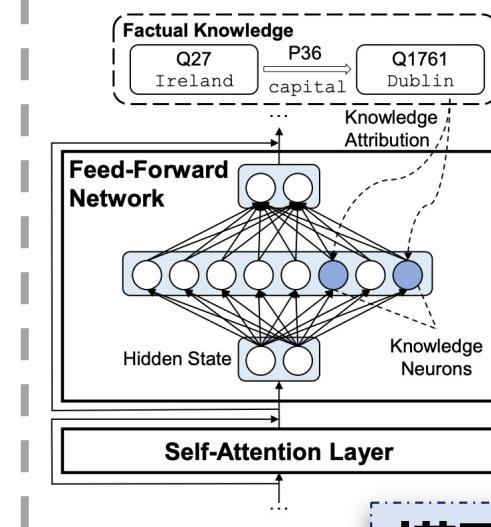


CCKS2023工业界论坛：8月27日上午

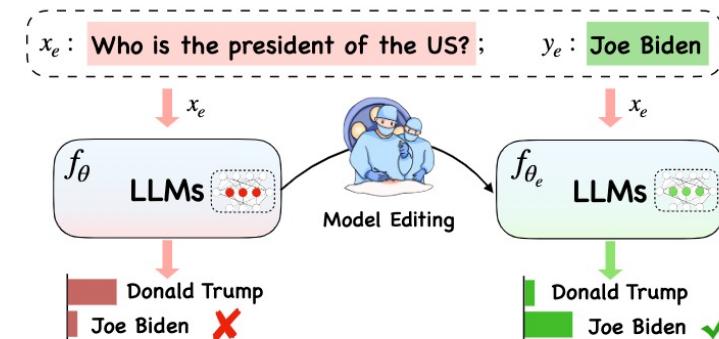
问题1：需要什么样的知识表示（图谱）？



文本符号知识表示



模型参数知识表示



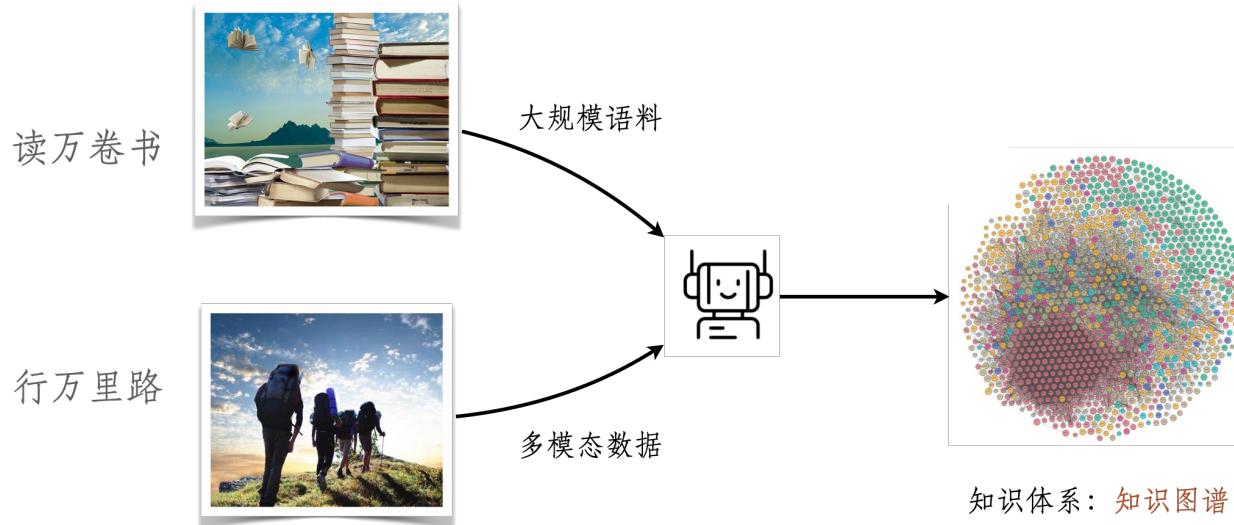
任务驱动的知识建模与表示问题；参数化模型的符号知识定位与编辑增强问题

[1] QA is the new KR: Question-answer Pairs as Knowledge Bases, AAAI2023

[2] Knowledge Neurons in Pretrained Transformers, ACL2021

[3] Editing Large Language Models: Problems, Methods, and Opportunities, 2023

问题2：知识通过何种方式习得？



Conversational Agents

"For the acquisition of the fruits of the earth?"

"Yes."

"And, further, is shoemaking also useful?"

"Yes."

"You would say, I suppose, for the acquisition of shoes?"

"Certainly."

"What about justice then? For the use or acquisition of what would you say it is useful in peacetime?"

"Contracts, Socrates."

"Do you mean by contracts, partnerships,²³ or something else?"

"Partnerships, of course."

"Then is the just man a good and useful partner in setting down draughts, or is it the skilled player of draughts?"²⁴

"The skilled player of draughts."

"In setting down bricks and stones, is the just man a more useful and better partner than the housebuilder?"

"Not at all."

"But in what partnership then is the just man a better partner than the harp player, just as the harp player is better than the just man when one has to do with notes?"

"In money matters, in my opinion."

"Except perhaps in using money, Polemarchus, when a horse must be bought or sold with money in partnership; then, I suppose, the expert on horses is a better partner. Isn't that so?"

《理想国》

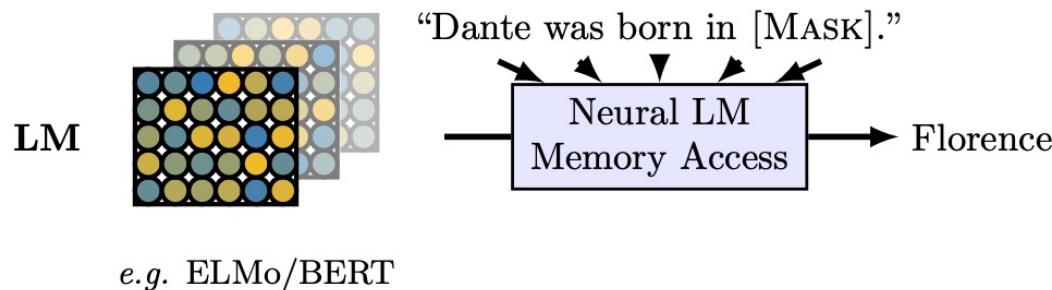
参数化模型的知识习得边界问题；多智能体交互情境下的知识获取问题

[1] Training Socially Aligned Language Models in Simulated Human Society, 2023

[2] Investigating the Factual Knowledge Boundary of Large Language Models with Retrieval Augmentation, 2023

[3] Encouraging Divergent Thinking in Large Language Models through Multi-Agent Debate, 2023

问题3：知识推理如何进行？



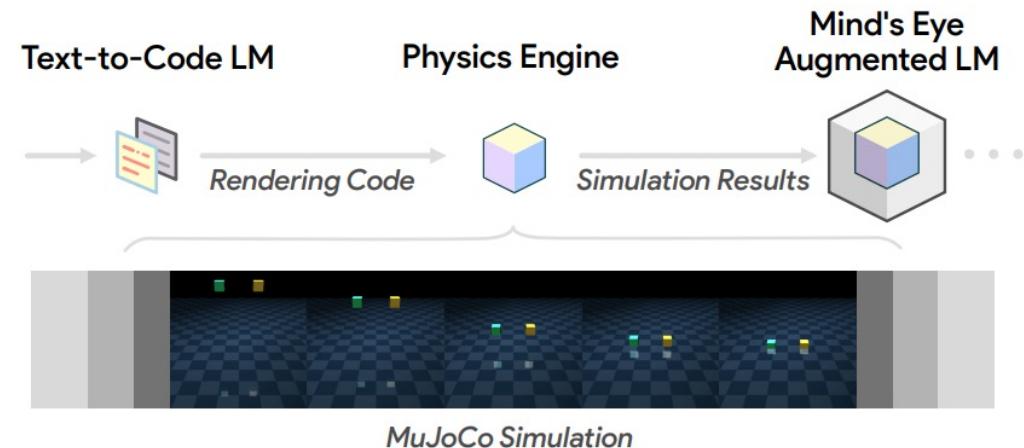
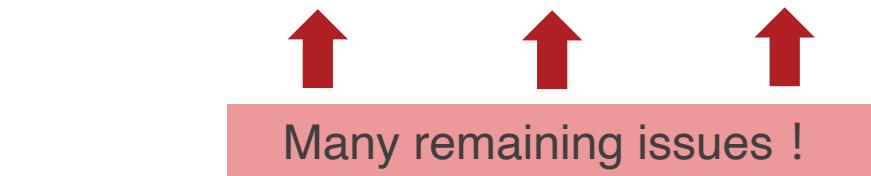
Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A: **Let's think step by step.**

(Output) There are 16 balls in total. Half of the balls are golf balls. That means that there are 8 golf balls. Half of the golf balls are blue. That means that there are 4 blue golf balls. ✓

缺乏对推理过程和内容的本质理解

Symbolic Convert Hierarchy Hybrid Neural

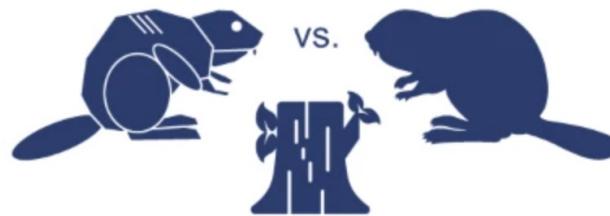


参数化模型的推理机理问题；符号引擎增强的可信鲁棒推理问题

- [1] Knowledgeable or Educated Guess? Revisiting Language Models as Knowledge Bases, ACL2021
- [2] Chain of thought prompting elicits reasoning in large language models, NeurIPS2022
- [3] Mind's Eye: Grounded Language Model Reasoning through Simulation, ICLR2023

问题4：知识能力如何评估？

Embodied Turing test



KoLA Benchmark



AgentBench

符号知识图谱



Embedding Concepts: Terms and Class Hierarchy

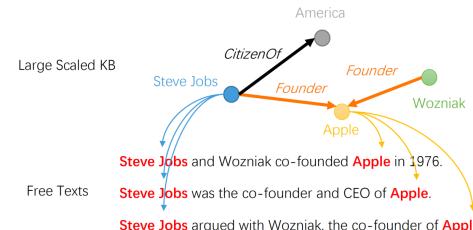
Embedding Facts: Entities , Attributes, Relations

Embedding Structures: Paths, Neighbors, Sub graphs

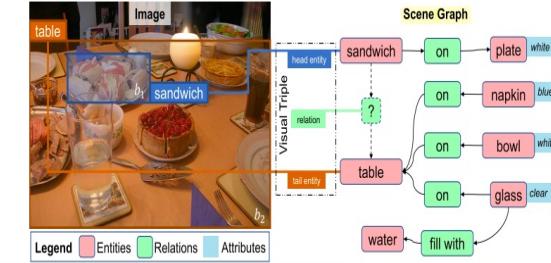
Embedding Logic: Axioms and Rules

感知与认知的无缝融合

语言知识



视觉知识

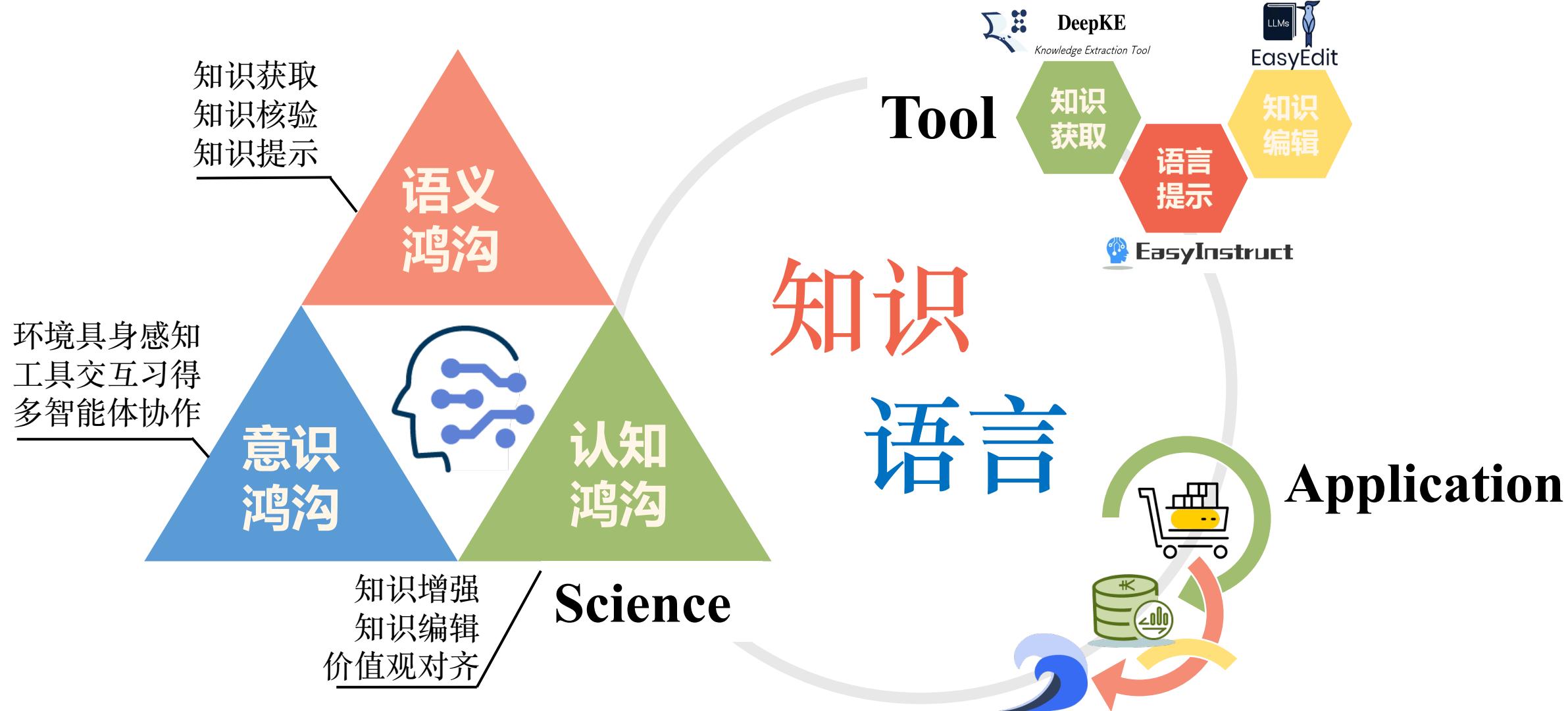


参数化模型的知识能力评估问题；多智能体具身知识评估问题

[1] KoLA: Carefully Benchmarking World Knowledge of Large Language Models. 2023

[2] Xiezhi: An Ever-Updating Benchmark for Holistic Domain Knowledge Evaluation

[3] AgentBench: Evaluating LLMs as Agents. 2023



大模型给知识图谱领域的发展带来了新的机遇与挑战

报告合作同学



朱雨琦

ZJU PhD
Student

王梦如

ZJU PhD
Student

谢辛

ZJU MSc
Student

陈想

ZJU PhD
Student

李磊

ZJU MSc
Student



Thank You!

zhangningyu@zju.edu.cn

A C C E P T M Y E N D L E S S G R A T I T U D E