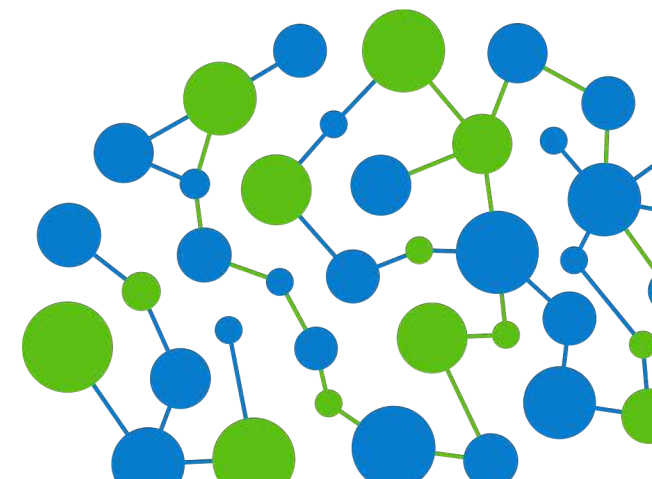


Towards Intelligent Systems Driven by Knowledge Graph and Large Language Model

Haofen Wang
CNLP 2023



- **The Rise of LLM**
- The Development of KG
- The Emerging Paradigm
- Outlook

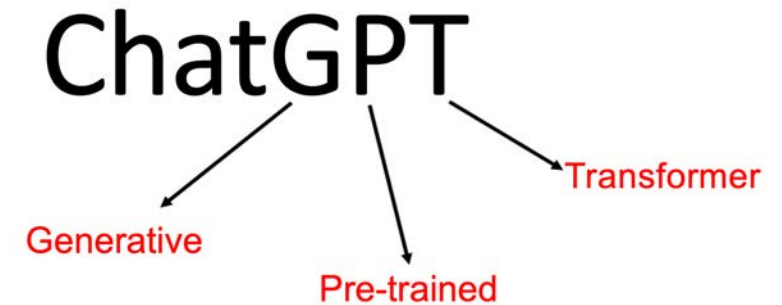


The Wave towards AGI by LLM

ChatGPT		
Examples	Capabilities	Limitations
"Explain quantum computing in simple terms" →	Remembers what user said earlier in the conversation	May occasionally generate incorrect information
"Got any creative ideas for a 10 year old's birthday?" →	Allows user to provide follow-up corrections	May occasionally produce harmful instructions or biased content
"How do I make an HTTP request in Javascript?" →	Trained to decline inappropriate requests	Limited knowledge of world and events after 2021

ChatGPT is a **generative conversational pre-trained large language model** launched by OpenAI in November 2022.

It interacts through conversations, which enables it to **answer follow-up questions, acknowledge mistakes, challenge incorrect assumptions, and refuse inappropriate requests.**



The Wave towards AGI by LLM

ChatGPT

Input

Andrew is free from 11 am to 3 pm, Joanne is free from noon to 2 pm and then 3:30 pm to 5 pm. Hannah is available at noon for half an hour, and then 4 pm to 6 pm. What are some options for start times for a 30 minute meeting for Andrew, Hannah, and Joanne?

Output

Here is a list of their availability:

- Andrew: 11 am to 3 pm
- Joanne: noon to 2 pm, and 3:30 pm to 5 pm
- Hannah: noon to 12:30 pm, and 4 pm to 6 pm

Based on their availability, there is a 30-minute window where all three of them are available, which is from 4 pm to 4:30 pm. So, the meeting can be scheduled at 4 pm.

GPT-4

Input

Andrew is free from 11 am to 3 pm, Joanne is free from noon to 2 pm and then 3:30 pm to 5 pm. Hannah is available at noon for half an hour, and then 4 pm to 6 pm. What are some options for start times for a 30 minute meeting for Andrew, Hannah, and Joanne?

Output

Andrew: 11 am - 3 pm

Joanne: 12 pm - 2 pm, 3:30 pm - 5 pm

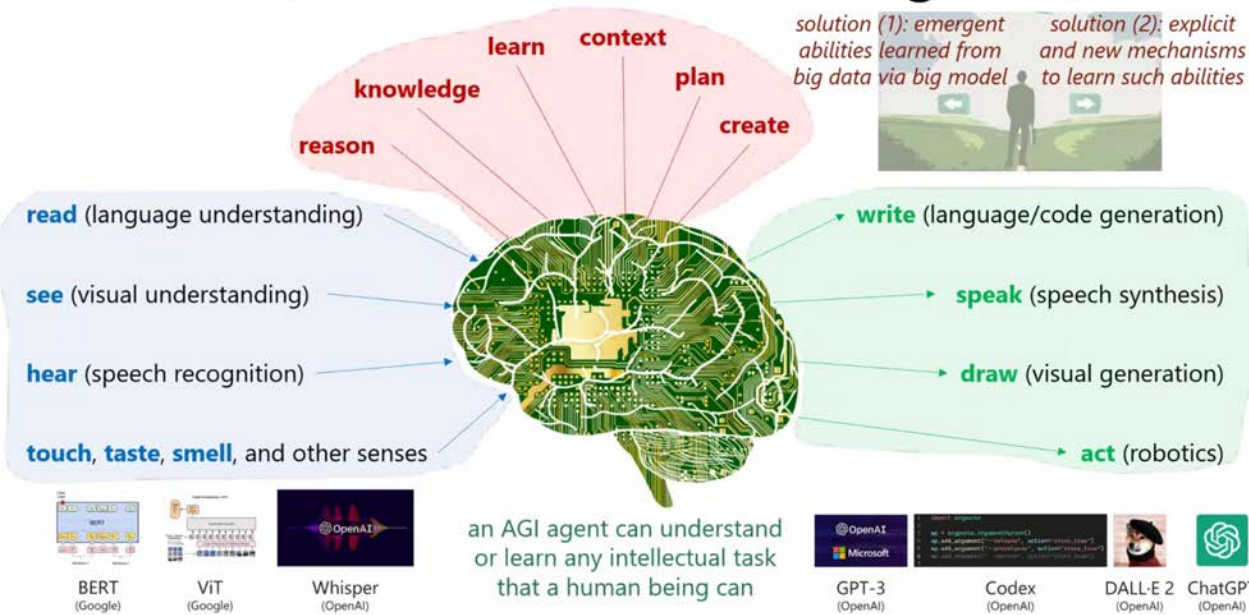
Hannah: 12 pm - 12:30 pm, 4 pm - 6 pm

Common availability for a 30-minute meeting: 12 pm - 12:30 pm

- **ChatGPT** has gained attention for providing detailed and clear answers in many areas, but it has some drawbacks such as **inaccurate facts and poor timeliness**.
- On March 15th, 2023, **GPT-4** was launched, which is **more fluent, more accurate, and supports image understanding**.

The Wave towards AGI by LLM

AGI (Artificial General Intelligence)



Credit by Nan Duan, MSRA

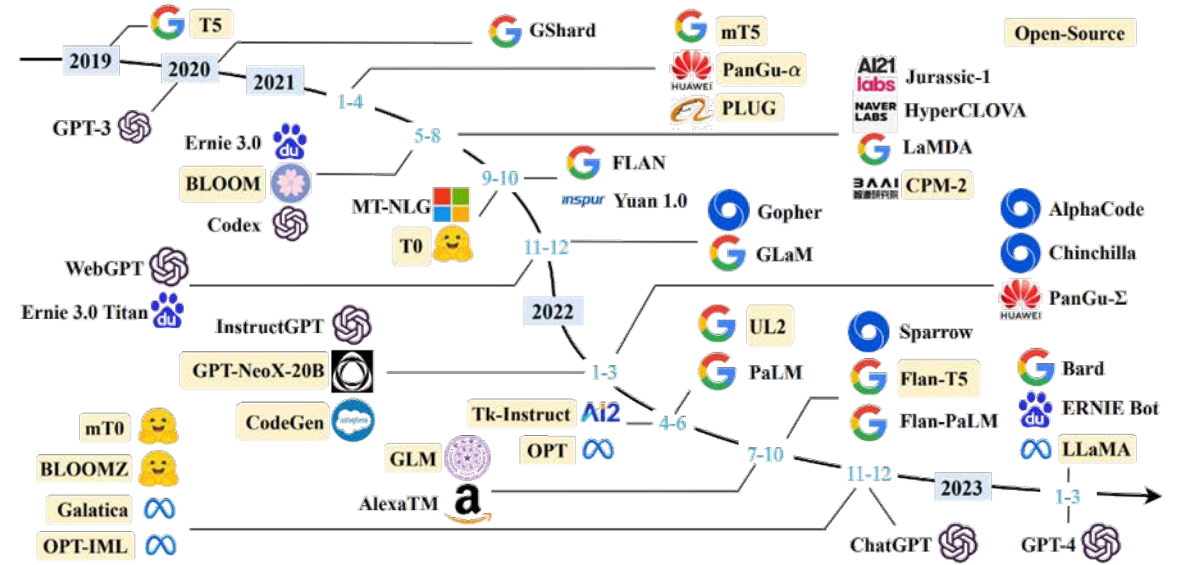
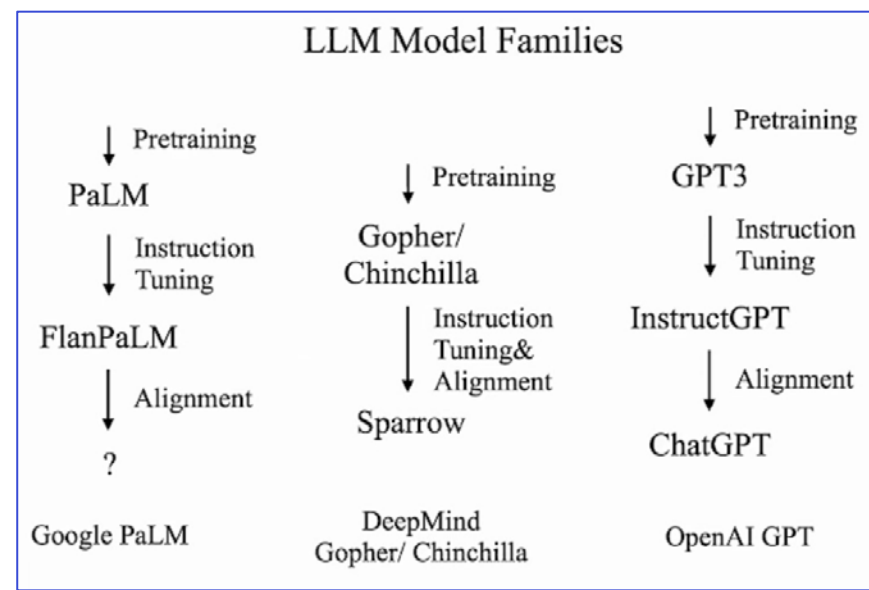
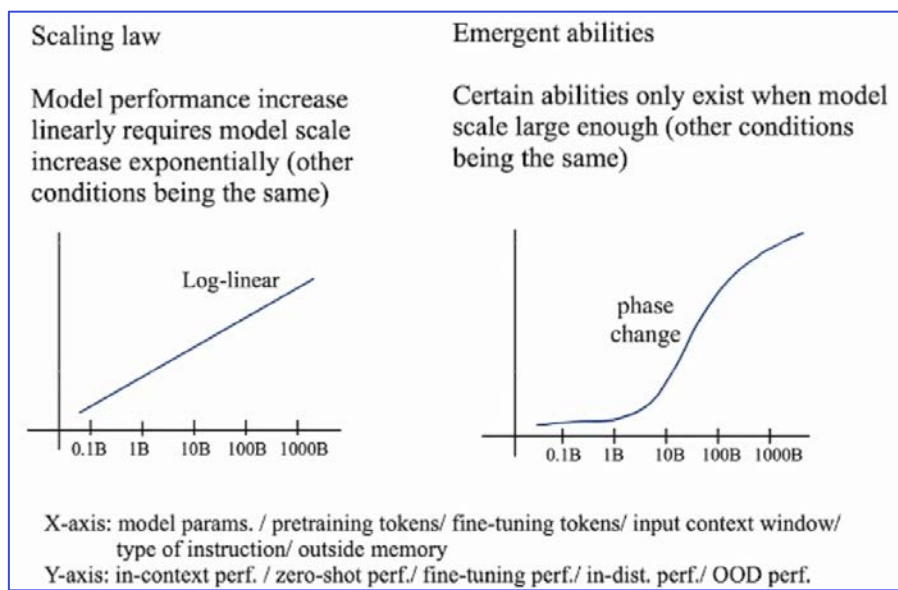


Fig. 1. A timeline of existing large language models (having a size larger than 10B) in recent years. We mark the open-source LLMs in yellow color.

A Survey of Large Language Models, 2023

The Emergent Abilities of LLM

- It is only when the size of a model reaches a certain threshold that its capabilities such as **cross-domain transfer** and **complex reasoning** emerges.
- The evolution of large language models at Google, DeepMind, and OpenAI generally involves three stages: **pre-training**, **instruction tuning**, and **alignment**.



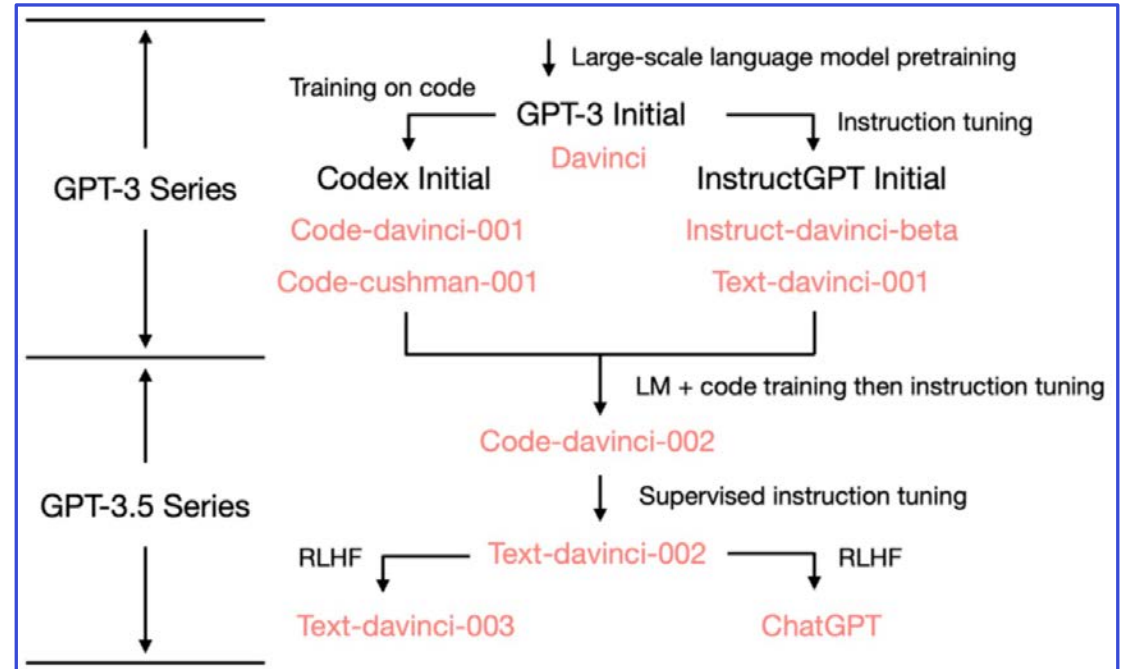
Pre-training, Instruction Tuning, Alignment, Specialization: On the Source of Large Language Model Abilities

Take a Closer Look at GPT Series

The Timeline of GPT

Time	Corp	Paper/Model	Milestones
2017.6	Google	Attention is all you need	Transformer
2018.6	OpenAI	Improving Language Understanding by Generative Pre-Training	GPT-1
2019.2	OpenAI	Language Models are Unsupervised Multitask Learners	GPT-2
2020.5	OpenAI	Language Models are Few-Shot Learners	GPT-3
2022.2	OpenAI	Training language models to follow instructions with human	InstructGPT
2022.11.30	OpenAI	ChatGPT	/
2023.3.15	OpenAI	GPT-4	/

The Roadmap of GPT Tech Tree



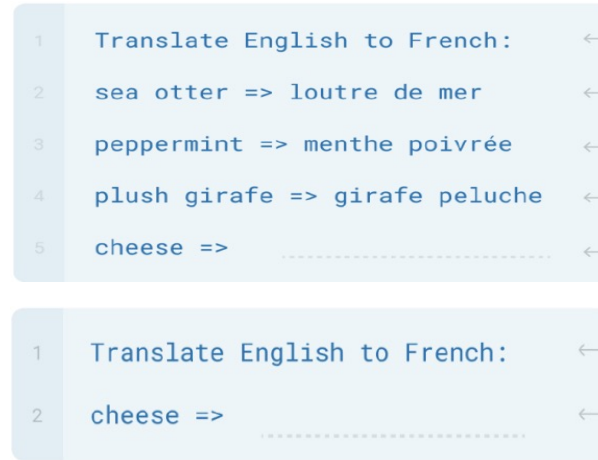
Key Tech Development



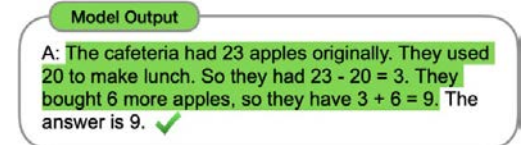
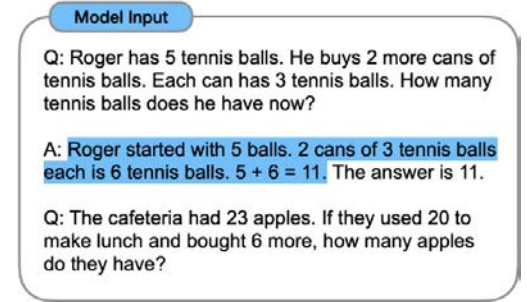
Take a Closer Look at Different Key Technologies



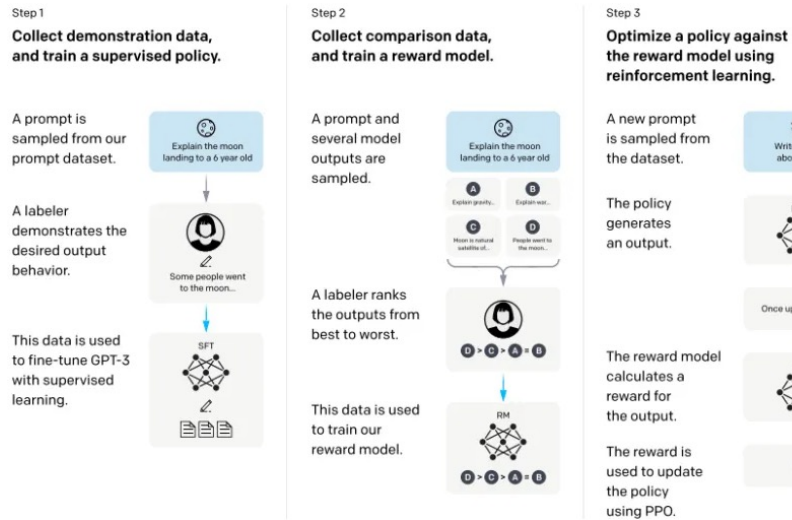
Fine-tuning



Few-shot prompting

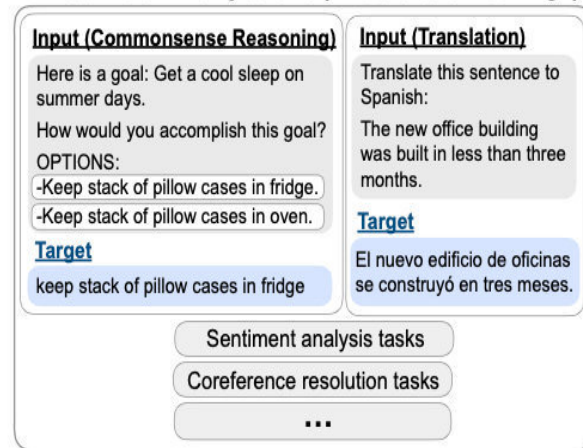


CoT prompting



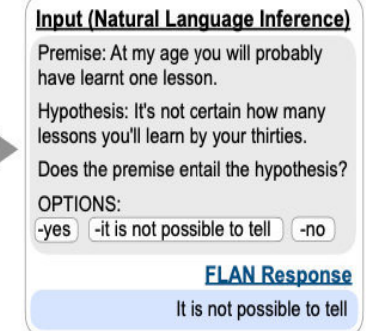
RLHF

Finetune on many tasks ("instruction-tuning")



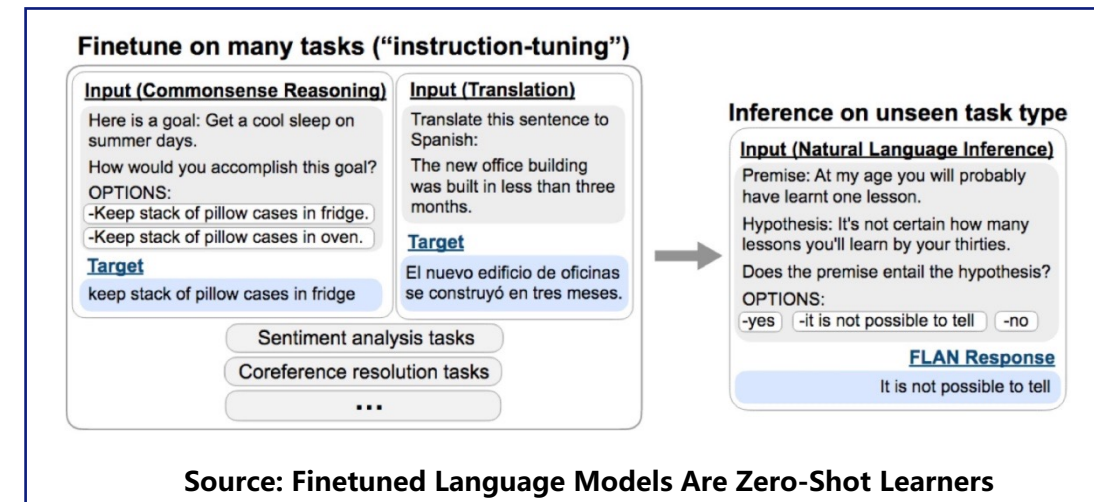
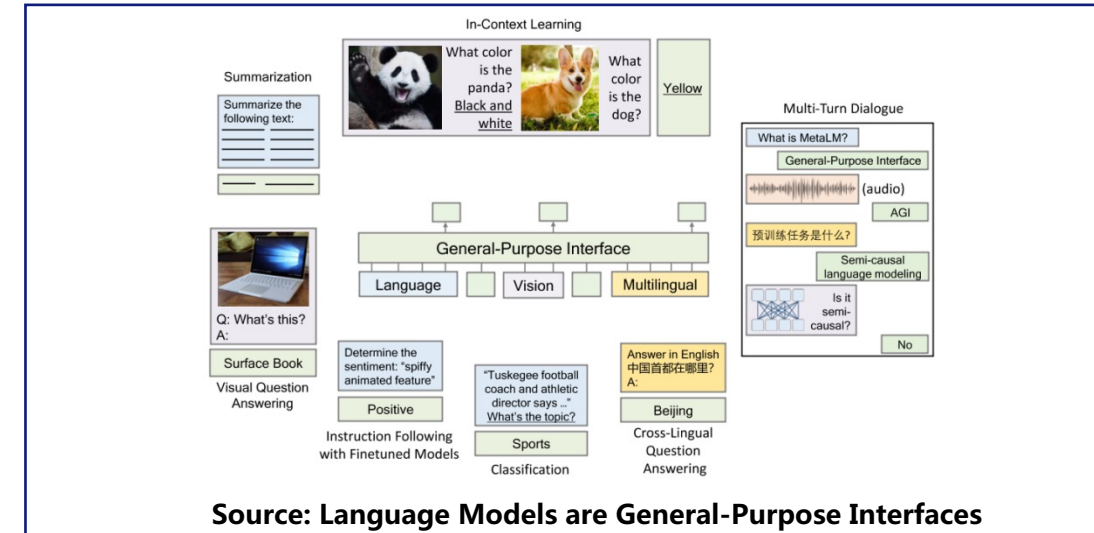
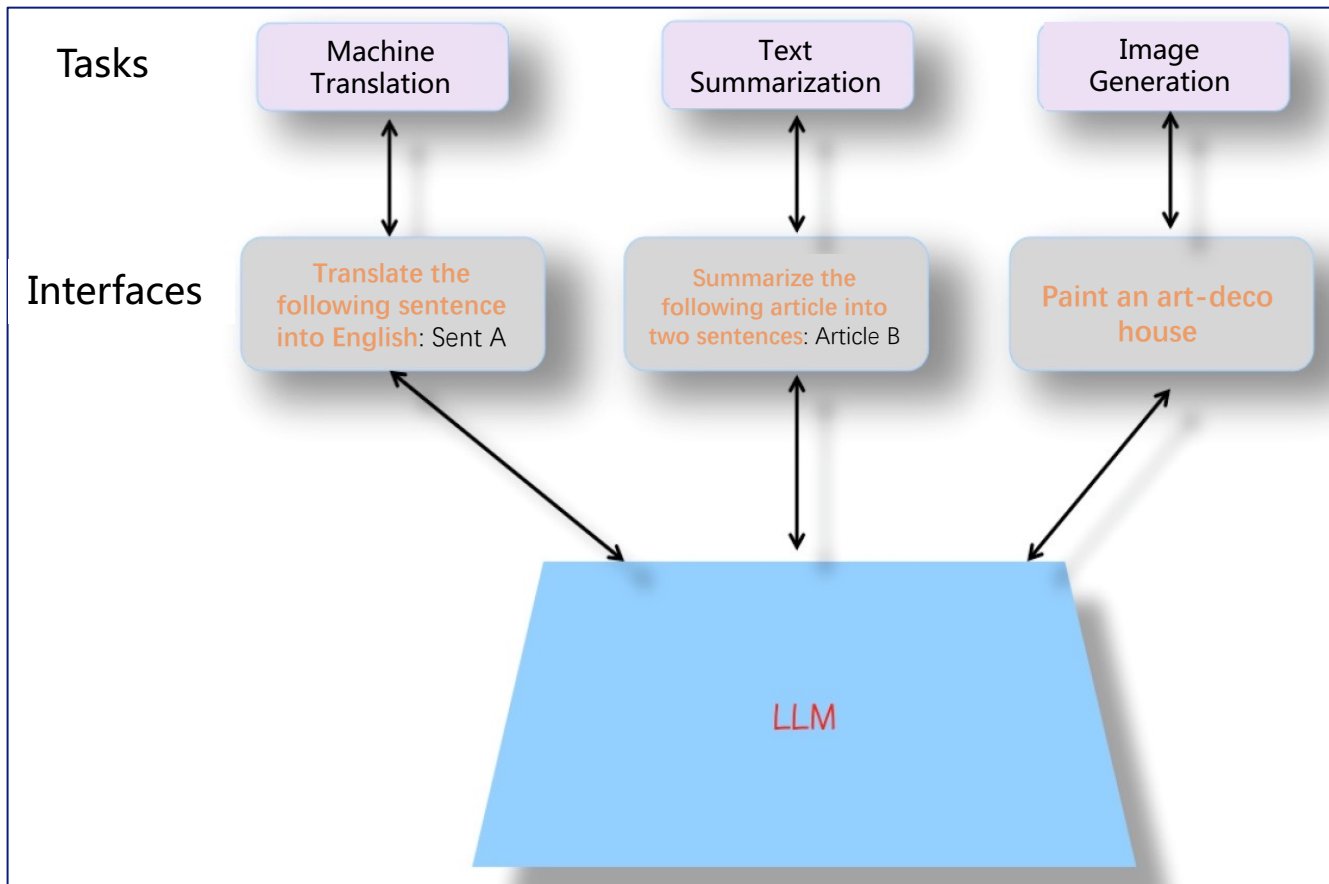
Instruction-tuning

Inference on unseen task type



LLM as a Basis for Providing Interfaces

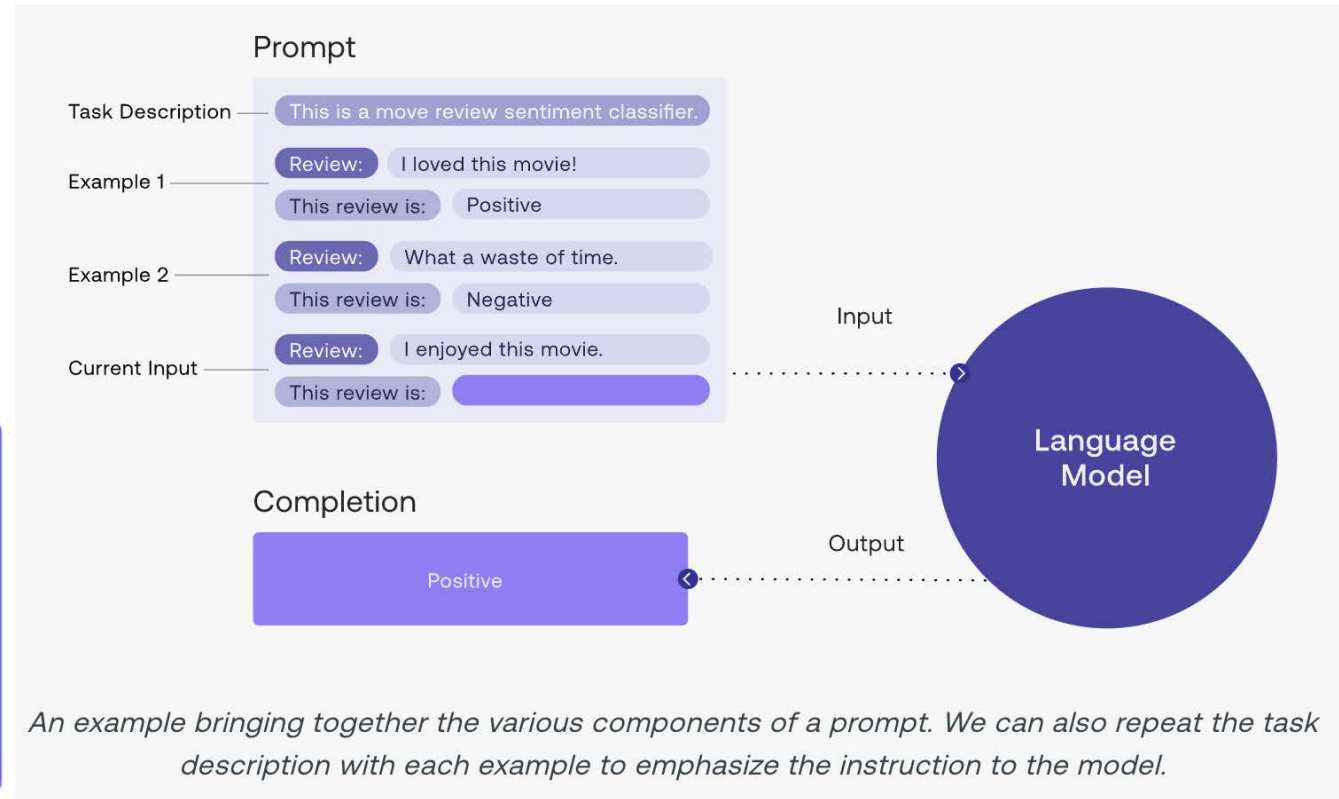
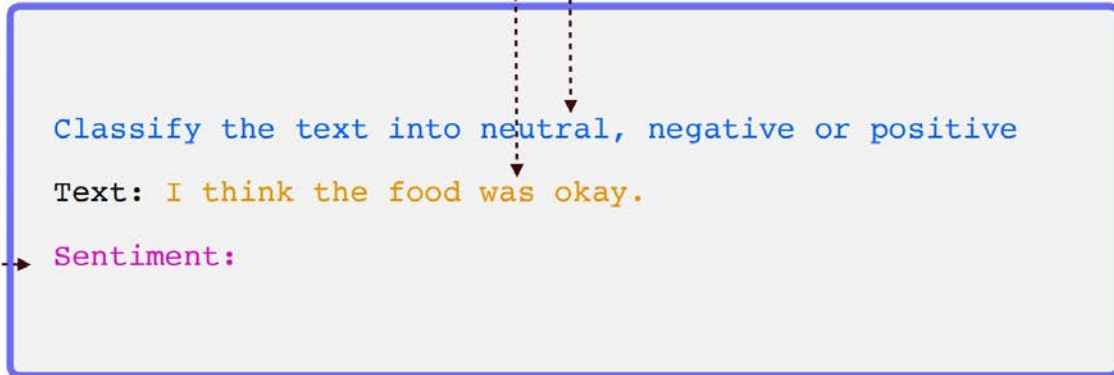
- Zero/few shot prompting
- In Context Learning
- Instruction



The Programming of LLM: Prompt Engineering

- A prompt is composed with the following components:

- **Instructions**
- Context
- **Input data**
- **Output indicator**



The Programming of LLM: Prompt Engineering

Standard Prompting

Q: Natalia sold clips to 48 of her friends in April, and then she sold half as many clips in May. How many clips did Natalia sell altogether in April and May?

T: The answer is :

LM

72

Few-shot Prompting

Q: There are 3 cars in the parking lot and 2 more cars arrive. How many cars are in the parking lot?
A: The answer is 5.

.....

Q: Natalia sold clips to 48 of her friends in April, and then she sold half as many clips in May. How many clips did Natalia sell altogether in April and May?

A: The answer is :

LM

72

Chain-of-Thought Prompting

Q: There are 3 cars in the parking lot and 2 more cars arrive. How many cars are in the parking lot?

C: There are 3 cars in the parking lot already. 2 more arrive. Now there are $3 + 2 = 5$ cars.

A: The answer is 5.

.....

Q: Natalia sold clips to 48 of her friends in April, and then she sold half as many clips in May. How many clips did Natalia sell altogether in April and May?

C: Natalia sold $48 / 2 = 24$ clips in May. Altogether, Natalia sold $48 + 24 = 72$ clips in April and May.

A: The answer is 72.

The Programming of LLM: Prompt Engineering

Zero-Shot

Chain-of-Thought

- Involves adding "Let's think step by step" to the original prompt

I went to the market and bought 10 apples. I gave 2 apples to the neighbor and 2 to the repairman. I then went and bought 5 more apples and ate 1. How many apples did I remain with?

11 apples

I went to the market and bought 10 apples. I gave 2 apples to the neighbor and 2 to the repairman. I then went and bought 5 more apples and ate 1. How many apples did I remain with?

Let's think step by step.

First, you started with 10 apples. You gave away 2 apples to the neighbor and 2 to the repairman, so you had 6 apples left. Then you bought 5 more apples, so now you had 11 apples. Finally, you ate 1 apple, so you would remain with 10 apples.

Self-Consistency

Q: There are 15 trees in the grove. Grove workers will plant trees in the grove today. After they are done, there will be 21 trees. How many trees did the grove workers plant today?
A: We start with 15 trees. Later we have 21 trees. The difference must be the number of trees they planted. So, they must have planted $21 - 15 = 6$ trees. The answer is 6.

Q: If there are 3 cars in the parking lot and 2 more cars arrive, how many cars are in the parking lot?
A: There are 3 cars in the parking lot already. 2 more arrive. Now there are $3 + 2 = 5$ cars. The answer is 5.

...

Q: Olivia has \$23. She bought five bagels for \$3 each. How much money does she have left?

A: She bought 5 bagels for \$3 each. This means she spent 5

Q: When I was 6 my sister was half my age. Now I'm 70 how old is my sister?
A:

When I was 6 my sister was half my age, so she was 3. Now I am 70, so she is $70 - 3 = 67$. The answer is 67.

When the narrator was 6, his sister was half his age, which is 3. Now that the narrator is 70, his sister would be $70 - 3 = 67$ years old. The answer is 67.

When the narrator was 6, his sister was half his age, which is 3. Now that the narrator is 70, his sister would be $70 - 3 = 67$ years old. The answer is 67.

Output 1
Output 2
Output 3

The Programming of LLM: Prompt Engineering

Chain-of-Thought (Wei et al., 2022)

Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: Roger started with 5 tennis balls. 2 cans of 3 tennis balls each is 6 tennis balls. $5 + 6 = 11$. The answer is 11.

Q: The bakers at the Beverly Hills Bakery baked 200 loaves of bread on Monday morning. They sold 93 loaves in the morning and 39 loaves in the afternoon. A grocery store returned 6 unsold loaves. How many loaves of bread did they have left?

Model Output

A: The bakers started with 200 loaves. They sold 93 in the morning and 39 in the afternoon. So they sold $93 + 39 = 132$ loaves. The grocery store returned 6 loaves. So they had $200 - 132 - 6 = 62$ loaves left.
The answer is 62.



Program-aided Language models (this work)

Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: Roger started with 5 tennis balls.

```
tennis_balls = 5
```

2 cans of 3 tennis balls each is

```
bought_balls = 2 * 3
```

tennis balls. The answer is

```
answer = tennis_balls + bought_balls
```

Q: The bakers at the Beverly Hills Bakery baked 200 loaves of bread on Monday morning. They sold 93 loaves in the morning and 39 loaves in the afternoon. A grocery store returned 6 unsold loaves. How many loaves of bread did they have left?

Model Output

A: The bakers started with 200 loaves

```
loaves_baked = 200
```

They sold 93 in the morning and 39 in the afternoon

```
loaves_sold_morning = 93
```

```
loaves_sold_afternoon = 39
```

The grocery store returned 6 loaves.

```
loaves_returned = 6
```

The answer is

```
answer = loaves_baked - loaves_sold_morning  
- loaves_sold_afternoon + loaves_returned
```

```
>>> print(answer)
```

```
74
```



The Programming of LLM: Prompt Engineering

(1) Hotspot QA

Question: Aside from the Apple Remote, what other device can control the program Apple Remote was originally designed to interact with?

(1a) Standard

Answer: iPod



(1b) CoT (Reason Only)

Thought: Let's think step by step. Apple Remote was originally designed to interact with Apple TV. Apple TV can be controlled by iPhone, iPad, and iPod Touch. So the answer is iPhone, iPad, and iPod Touch.

Answer: iPhone, iPad, iPod Touch



(1c) Act-Only

Act 1: Search[Apple Remote]

Obs 1: The Apple Remote is a remote control ...

Act 2: Search[Front Row]

Obs 2: Could not find [Front Row]. Similar: ...

Act 3: Search[Front Row (software)]

Obs 3: Front Row is a discontinued media center software ...

Act 4: Finish[yes]



(1d) ReAct (Reason + Act)

Thought 1: I need to search Apple Remote and find the program it was originally designed to interact with.

Act 1: Search[Apple Remote]

Obs 1: The Apple Remote is a remote control introduced in October 2005 by Apple ... originally designed to control the Front Row media center program ...

Thought 2: Apple Remote was originally designed to control the Front Row media center program. I need to search Front Row next and find what other device can control it.

Act 2: Search[Front Row]

Obs 2: Could not find [Front Row]. Similar: ['Front Row Seat to Earth', 'Front Row Motorsports', 'Front Row (software)', ...]

Thought 3: Front Row is not found. I need to search Front Row (software) .

Act 3: Search[Front Row (software)]

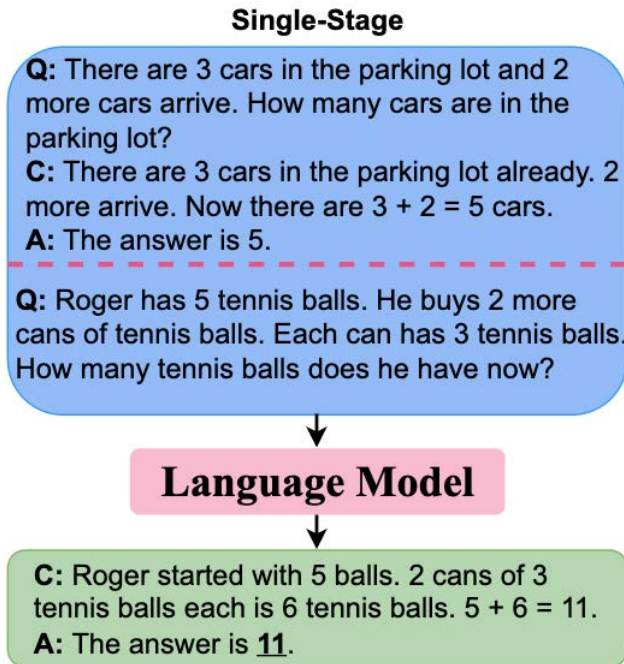
Obs 3: Front Row is a discontinued media center software ...

Thought 4: Front Row (software) is controlled by an Apple Remote or the keyboard function keys. So the answer is keyboard function keys.

Act 4: Finish[keyboard function keys]



The Programming of LLM: Prompt Reasoning



- Sensitivity of in-context learning
- complexity, diversity, explicit

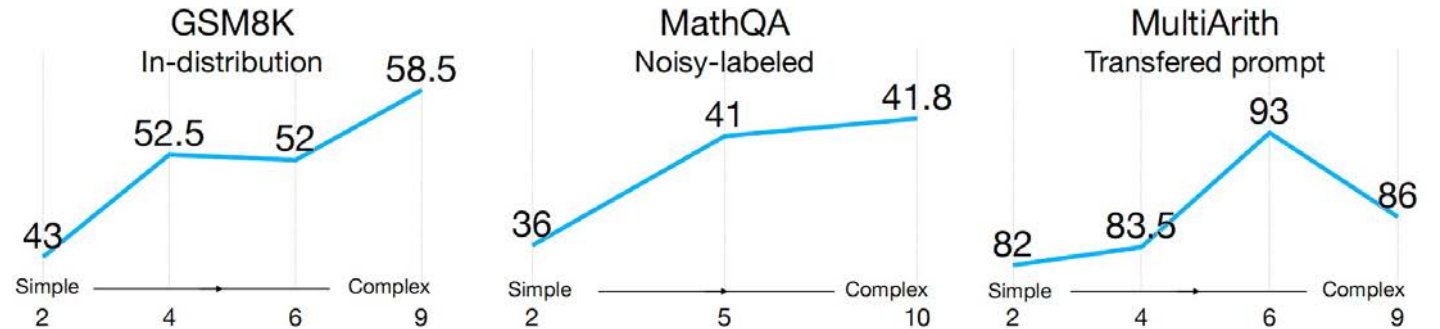
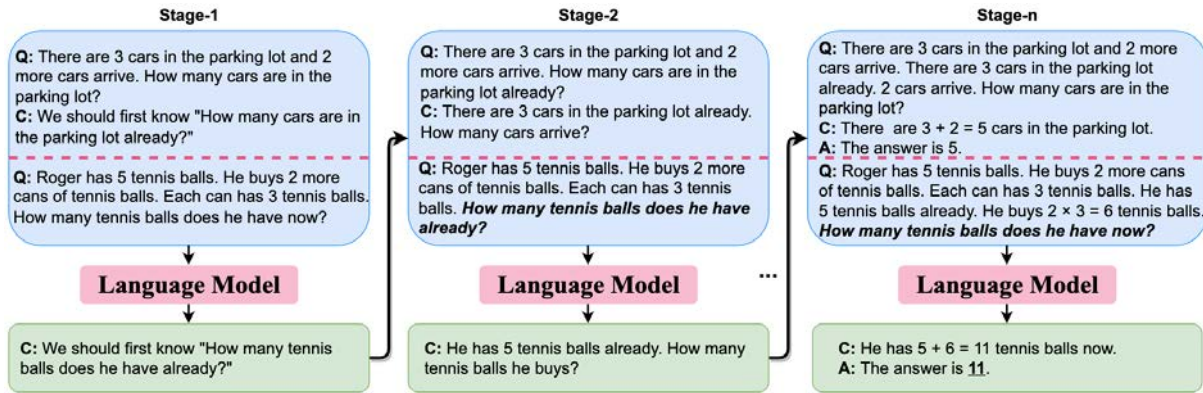


Figure 2: Validation set performance. X-axis means reasoning steps and y-axis means accuracy. More reasoning steps in prompts overall achieve higher accuracy when prompts are in-distribution (left), noisily labeled (middle), and out of distribution (right).

A simple and direct improvement on CoT methods. Existing work has shown that in-context learning is very sensitive to the order and quality of the examples contained in it. A very small change may cause a large decline in the effect of the model. To further optimize CoT, an intuitive method is to make its examples more complex, diverse or discriminative.

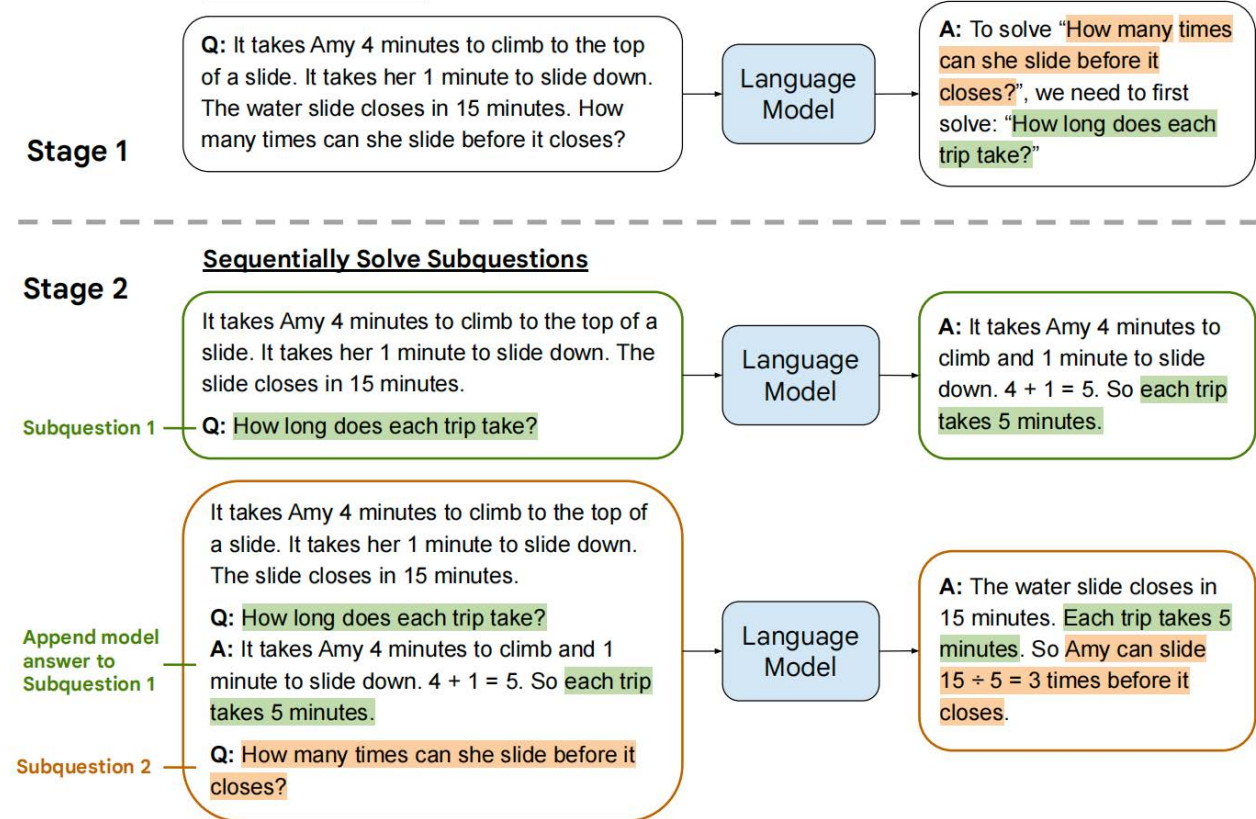
The Programming of LLM: Prompt Reasoning

Multi-stage



Just like human-like reasoning, we always need to iterative reason given a complex problem. Inspired by it, a multi-stage method reduces a problem into sub-questions and sequentially solve them.

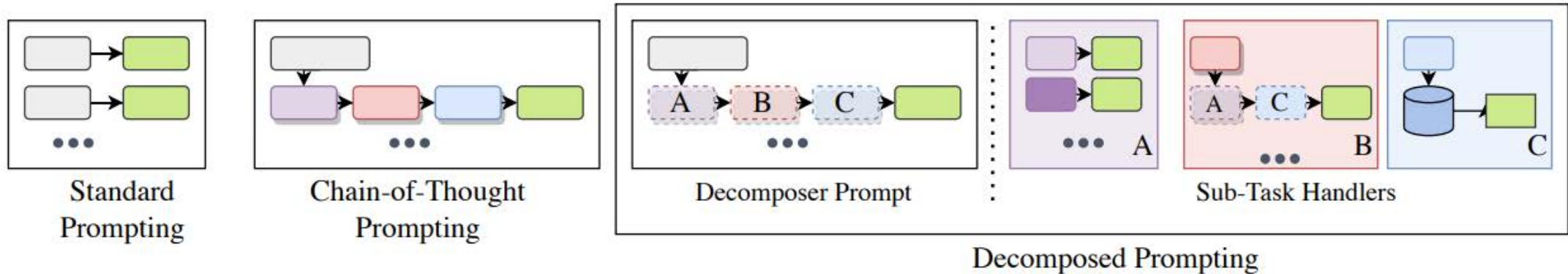
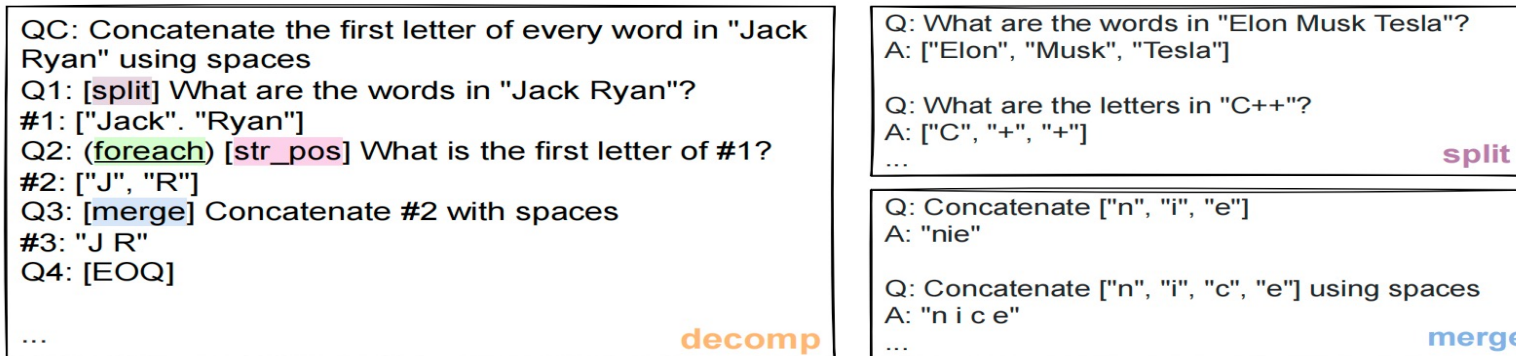
Problem Reduction



Least-to-Most Prompting Enables Complex Reasoning in Large Language Models 2022

The Programming of LLM: Prompt Reasoning

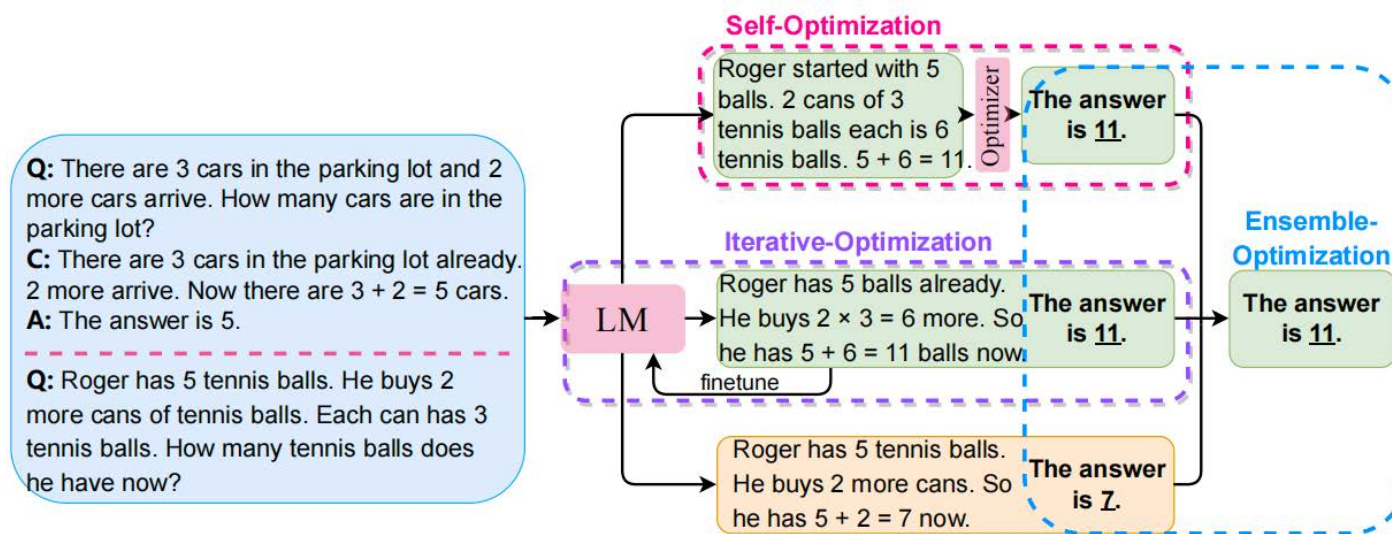
Decomposed Prompting



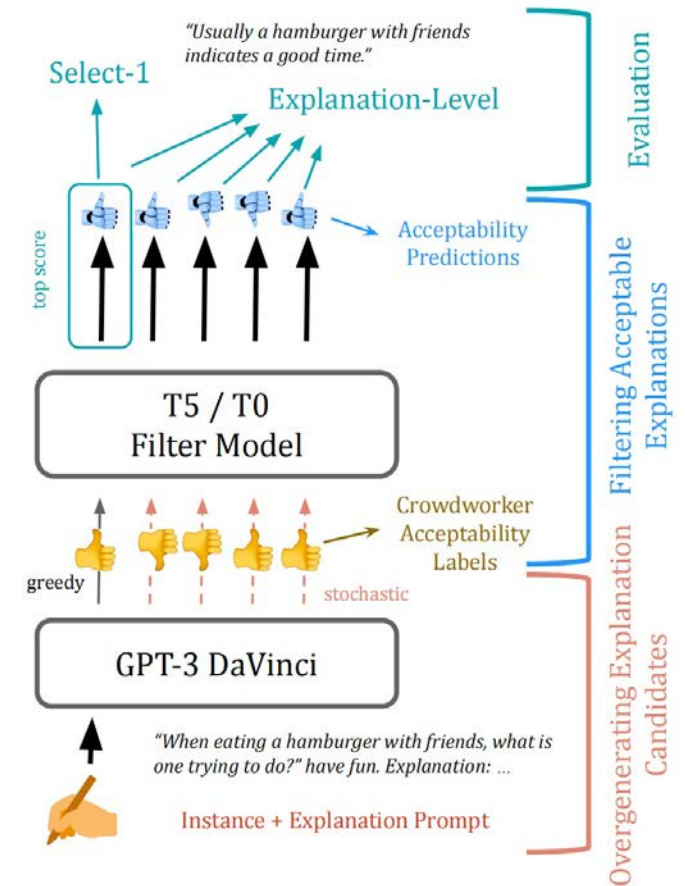
For different types of sub-questions, design specific prompts so that each type can be solved more precisely

The Programming of LLM: Prompt Reasoning

Self-optimization

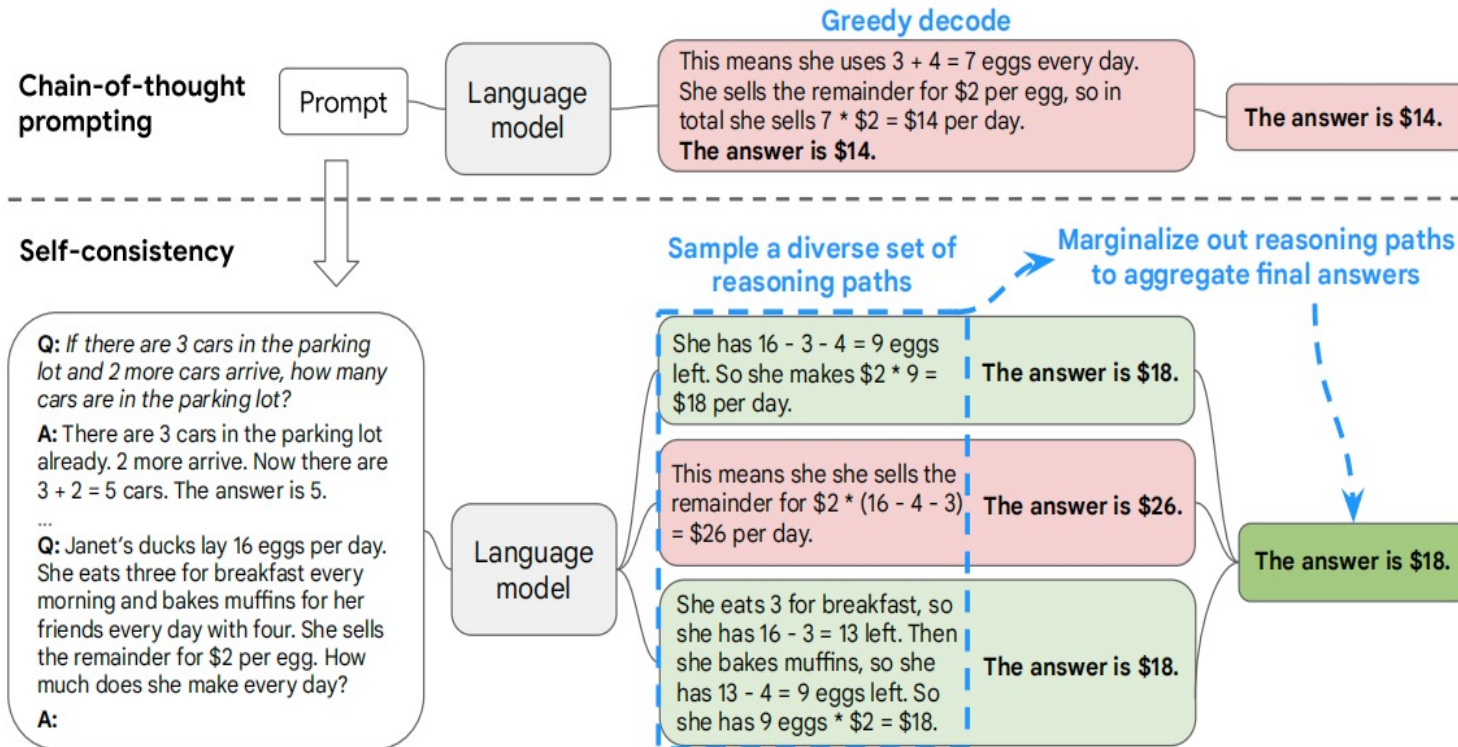


Both single-stage and multi-stage methods focus on the input optimization during prompting, we can also optimize the output of LLM. Since reasoning is the key of CoT prompting, self-optimization involves refinement or filter modules to optimize the reasoning paths during post-processing.



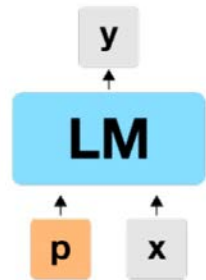
The Programming of LLM: Prompt Reasoning

Ensemble-Optimization

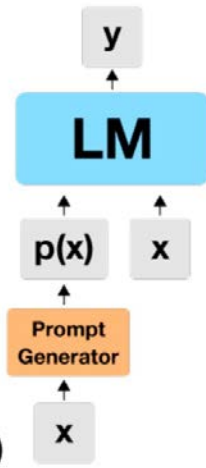


All roads lead to Rome, reasoning problems often have more than one path to the correct answer, and there are often multiple solutions to a problem, thus resulting in an ensemble optimization method, represented by Google's self-consistency. It uses sampling methods commonly used in the generative decoding process, such as temperature, top-k, etc., to allow the language model to generate multiple paths, and summarize the answers of all paths to generate the final answer

The Programming of LLM: Design Patterns of Prompts

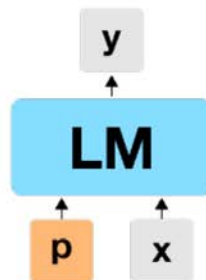


(a)



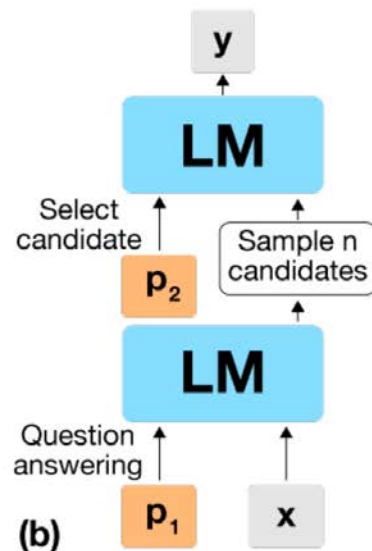
(b)

input-dependent prompt tuning

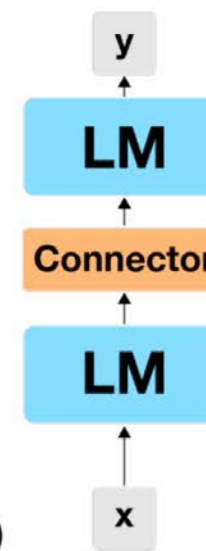


(a)

textual recursive LM

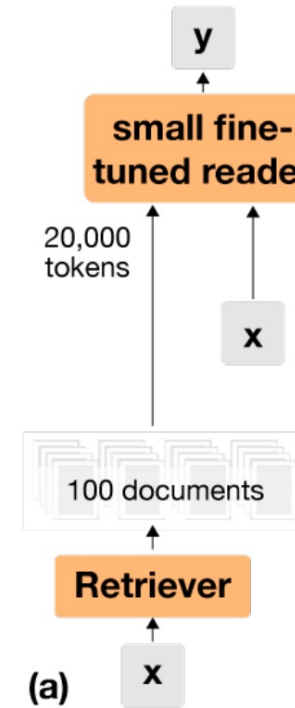


(b)



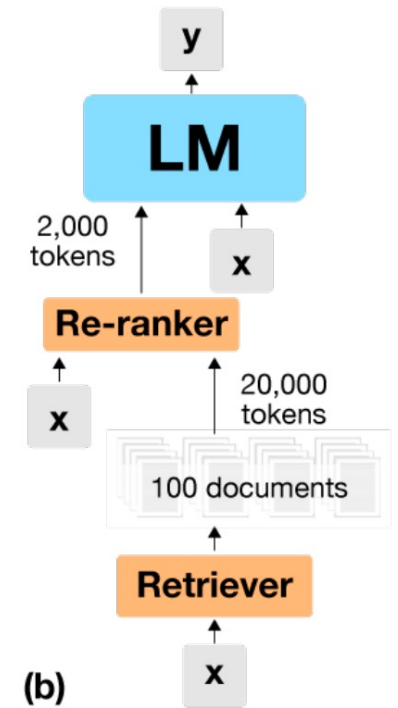
(c)

neural recursive LM



(a)

retrieve-read



(b)

re-rank

The Ecosystem of LLM

ChatGPT Plugins



Expedia

Bring your trip plans to life—get there, stay there, find things to see and do.



FiscalNote

Provides and enables access to select market-leading, real-time data sets for legal, political, and regulatory data and information.



Instacart

Order from your favorite local grocery stores.



KAYAK

Search for flights, stays and rental cars. Get recommendations for all the places you can go within your budget.



Klarna Shopping

Search and compare prices from thousands of online shops.



Milo Family AI

Giving parents superpowers to turn the manic to magic, 20 minutes each day. Ask: Hey Milo, what's magic today?



OpenTable

Provides restaurant recommendations, with a direct link to book.



Shop

Search for millions of products from the world's greatest brands.



Speak

Learn how to say anything in another language with Speak, your AI-powered language tutor.



Wolfram

Access computation, math, curated knowledge & real-time data through Wolfram|Alpha and Wolfram Language.

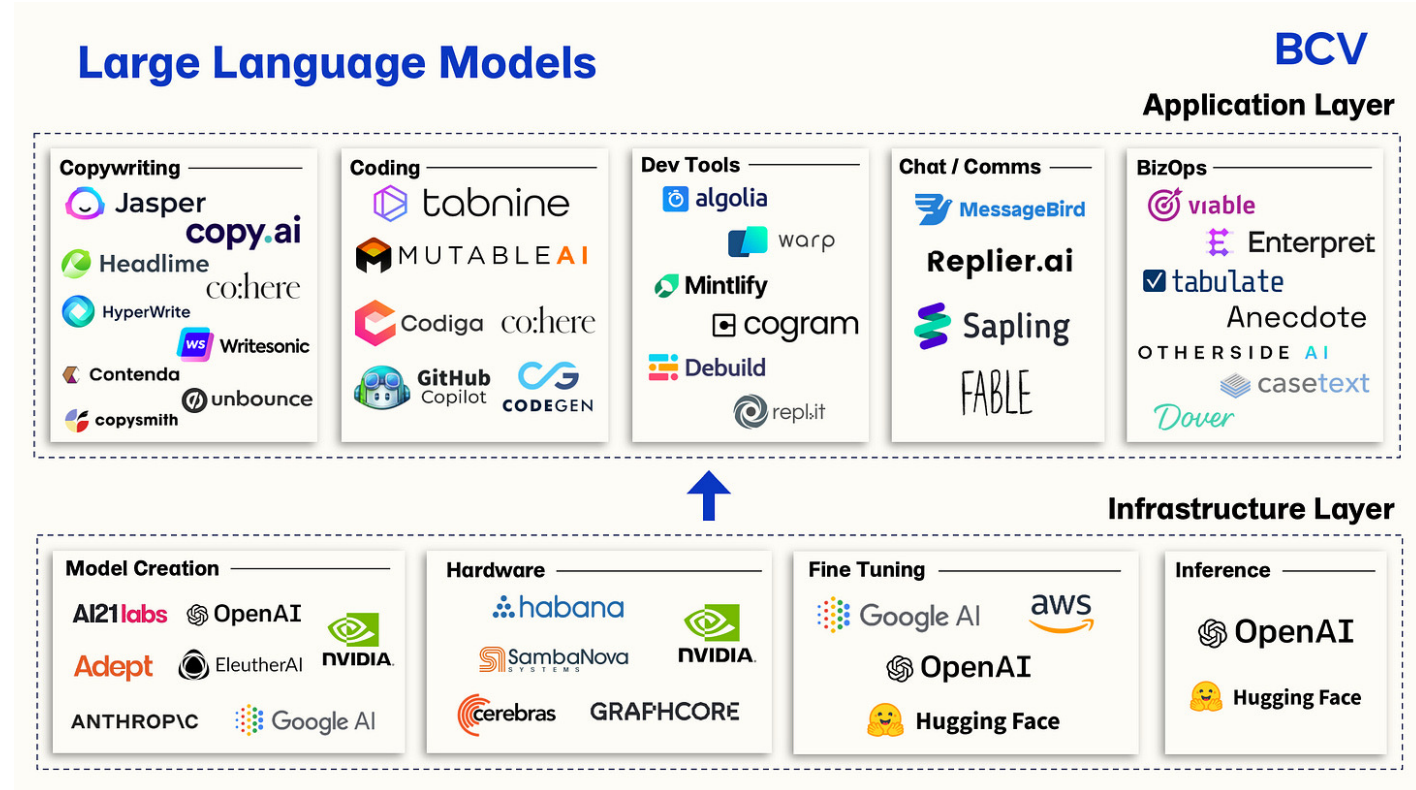
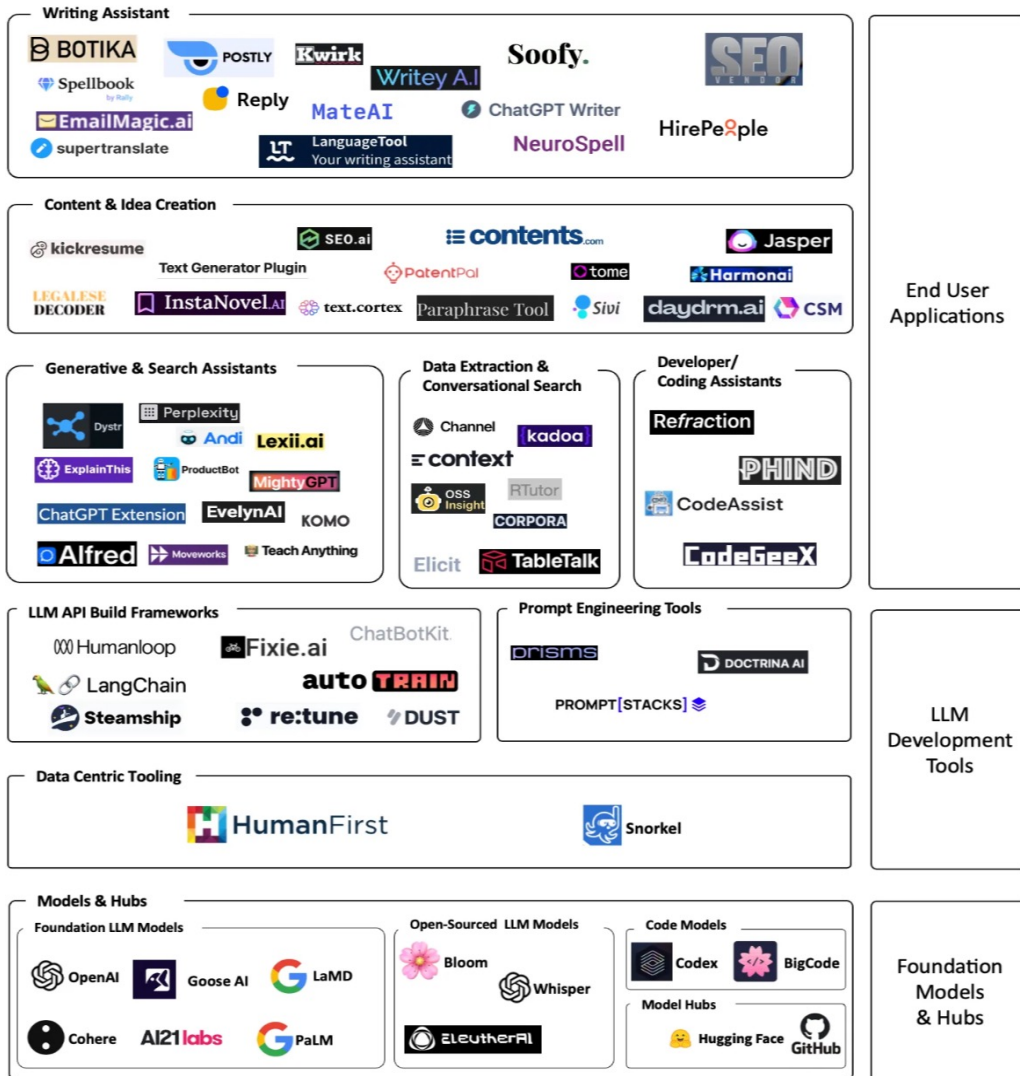


Zapier

Interact with over 5,000+ apps like Google Sheets, Trello, Gmail, HubSpot, Salesforce, and more.

The Ecosystem of LLM

Foundation Large Language Model Stack



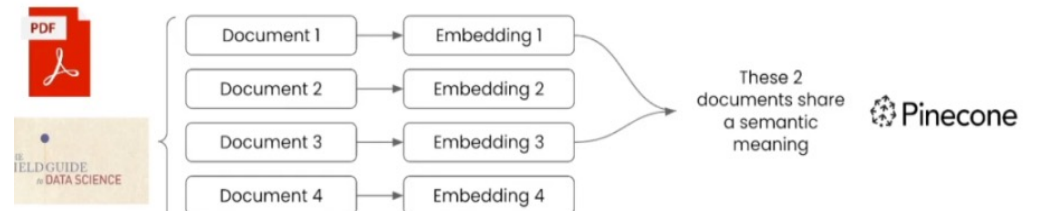
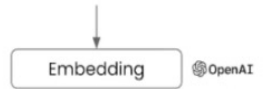
The Ecosystem of LLM: Open-source Players

Chat language models tracker

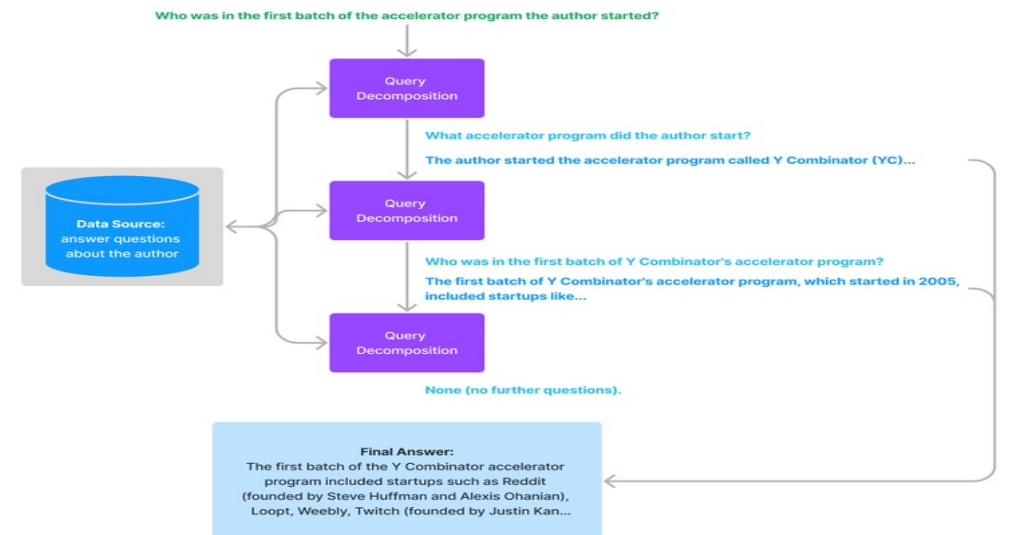
Model	HuggingFace	License	Training Data
Foundation Models (Recent)			
Cerebras-GPT	13B	Apache 2.0	
LLaMa	7-65B	Restricted	CCNet, C4, GitHub, Wikipedia, Books, ArXiv
Chat Models			
OpenChatKit	NeoX-20B	Apache 2.0	LAION OIG (Apache 2.0)
ChatRWKV	RWKV-4-Raven-v7-Eng	Apache 2.0	Alpaca, CodeAlpaca, Guanaco, GPT4All (OIG, GP3 samples and more)
Alpaca (LLaMa)	Weights no released	Restricted	OpenAI GPT3 davinci-003 (52K samples)
Alpaca 13B	13B (tweet)	Restricted	OpenAI GPT4 (>20k samples)
Alpaca w GPT4 data	13B	Restricted	Alpaca, GPT4All, CodeAlpaca, OIG
Alpaca w GPT4All and more	7B	Restricted	Alpaca (GPT3), ChatGPT, Quora, StackOverFlow and MedQuAD
Baize (LLaMa)	7-30B	Restricted	
Vicuna (LLaMa)	13B (LLaMa delta), 13B model	Restricted	ShareGPT/ChatGPT (70K samples)
Koala (LLaMa)	Weights no released	Restricted	ChatGPT, LAION OIG , Alpaca (GPT3) and more
BLOOM-Alpaca	7B (tweet)	RAIL	Alpaca (GPT3)
Flan-Alpaca	220M-11B	Apache 2.0	Flan + Alpaca/GPT4All/ShareGPT/Vicuna
Dolly (GPT-J-Alpaca)	6B	Apache 2.0	Alpaca (GPT3)
GPT4All (LLaMa)	7B	Restricted	GPT4All (OIG, GP3 samples and more)
GPT4All (GPT-J)	Coming soon	Open	GPT4All (OIG, GP3 samples and more)
ColossalChat (LLaMa)		Restricted	InstructionWild



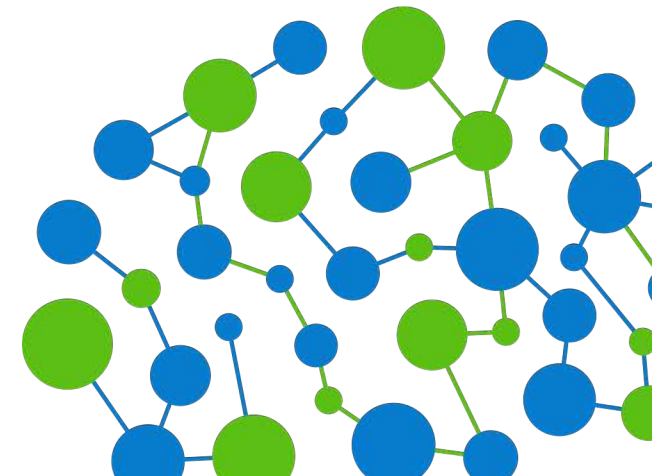
"What is a data scientist?"



Welcome to LlamaIndex (GPT Index)!

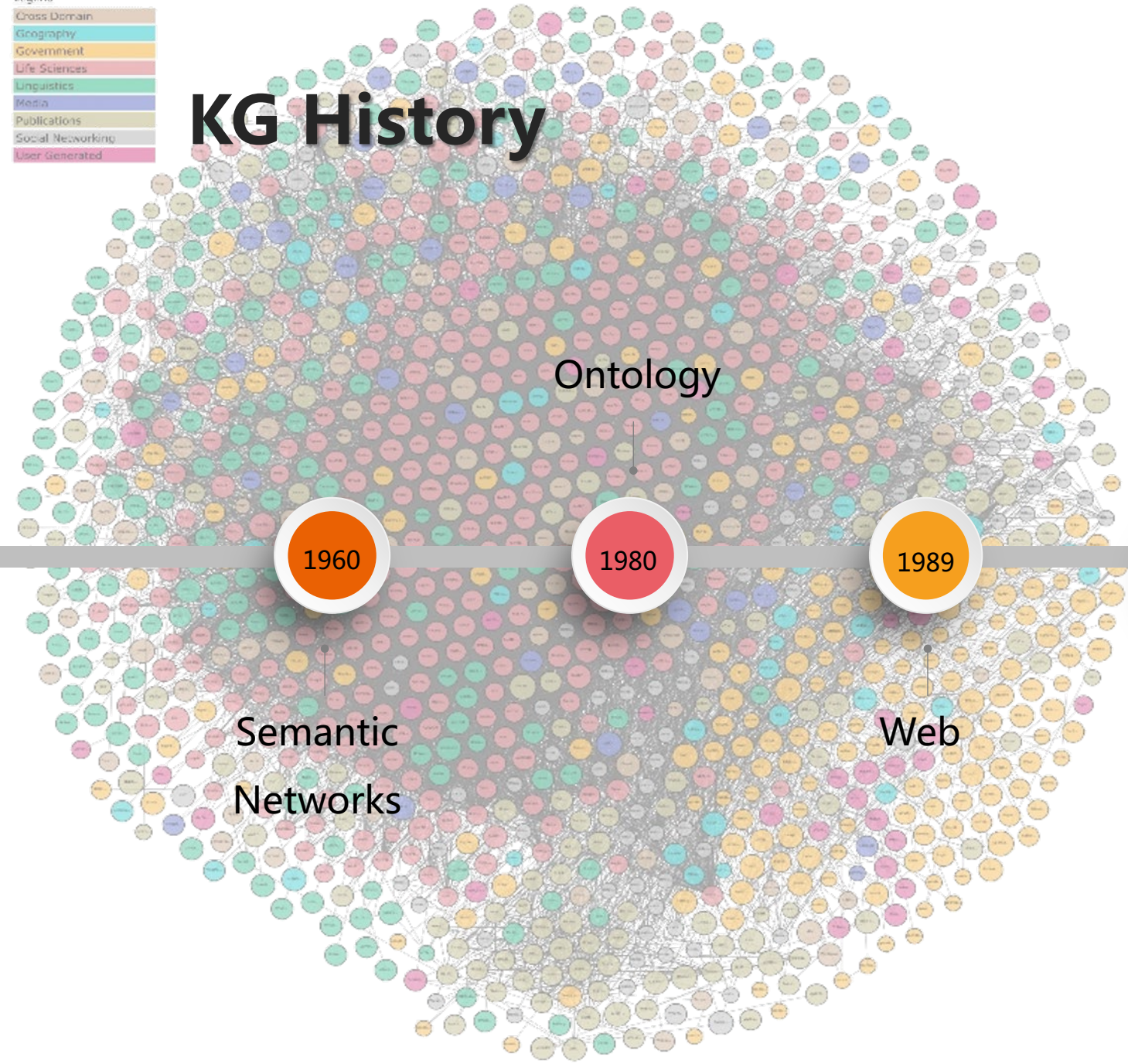


- The Rise of LLM
- **The Development of KG**
- The Emerging Paradigm
- Outlook



- Legend
- Cross Domain
- Geography
- Government
- Life Sciences
- Linguistics
- Media
- Publications
- Social Networking
- User Generated

KG History



Semantic
Web

Knowledge
Graph

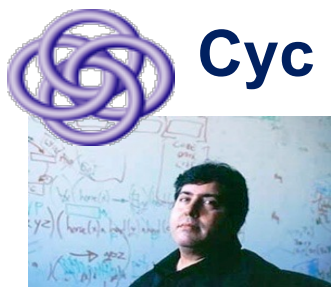


Linked data

Linking Open Data cloud diagram 2022-11-03
 Andrejs Abele, John P. McCrae, Paul Buitelaar,
 Anja Jentzsch and Richard Cyganiak.
<http://lod-cloud.net/>



What is Knowledge Graph (KG) – Popular KBs and Characteristics



WordNet



By Human
For Human

guitarist \subset {player, musician}
 \subset artist
 algebraist
 \subset mathematician
 \subset scientist

Wikipedia



4.5 Mio. English articles
 20 Mio. contributors

$\forall x$: human(x) \Rightarrow
 $(\exists y$: mother(x,y) \wedge
 $\exists z$: father(x,z))
 $\forall x,u,w$: (mother(x,u) \wedge
 mother(x,w)
 $\Rightarrow u=w$)

1985

1990

2000

2005

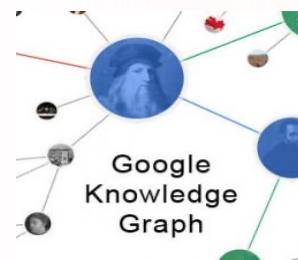
2010

2015

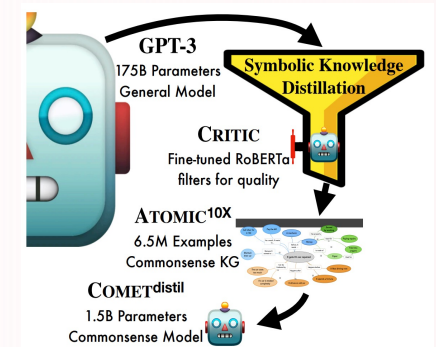
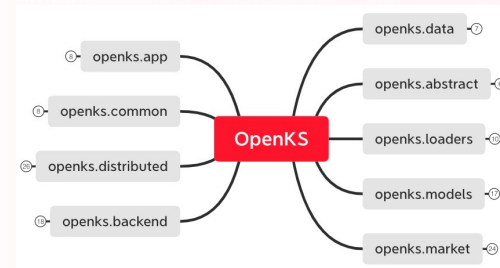
2023



By Algorithm
For Machine



Human Machine
Collaboration



Key Features

In the early stage, KG is **High-quality, manually-built, and for human consumption**; in the middle age, KG is **constructed by algorithms** and used **to enhance the understanding capability of machines**; nowadays KG is evolving towards **multi-modality** and **subsymbolic** representations

Knowledge Graph

Knowledge Graph (KG) is an explicit representation of human knowledge, which is stored in the form of graph and used for reasoning and computing.

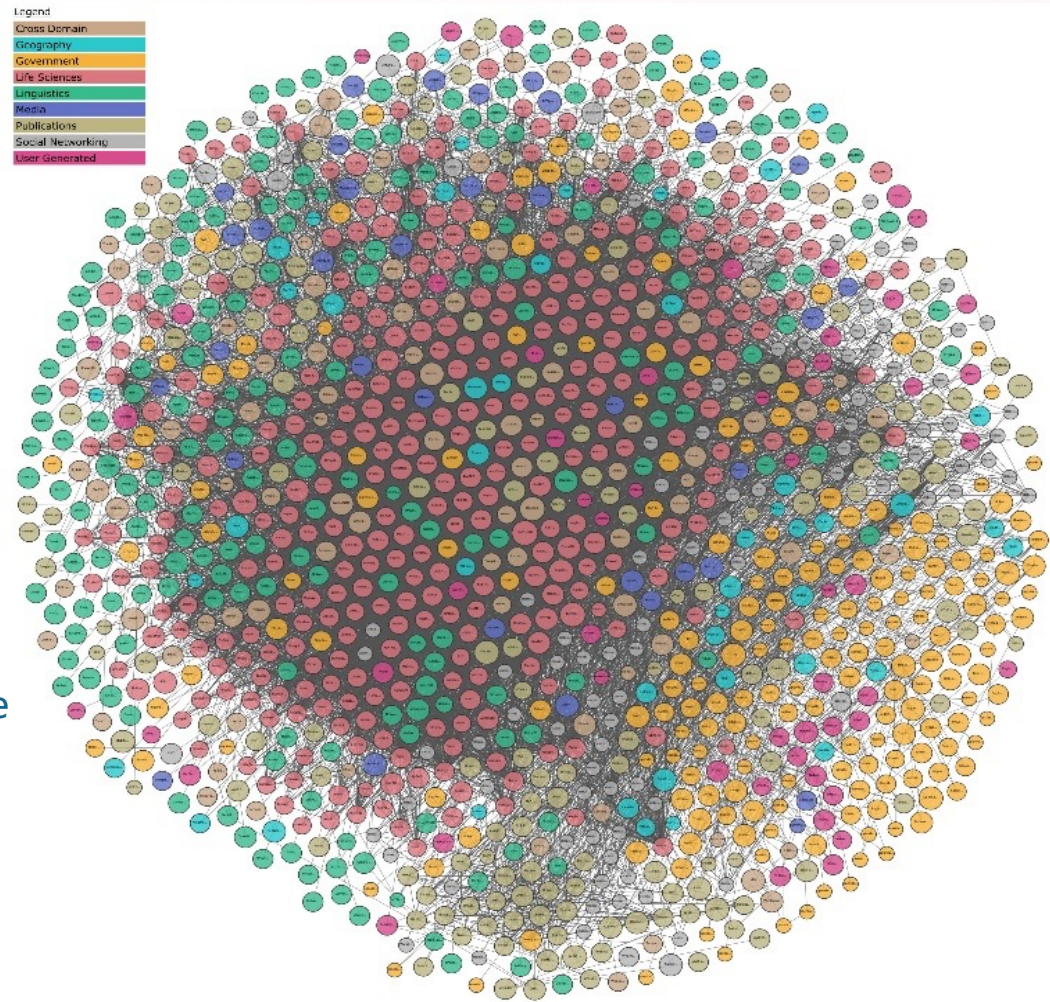
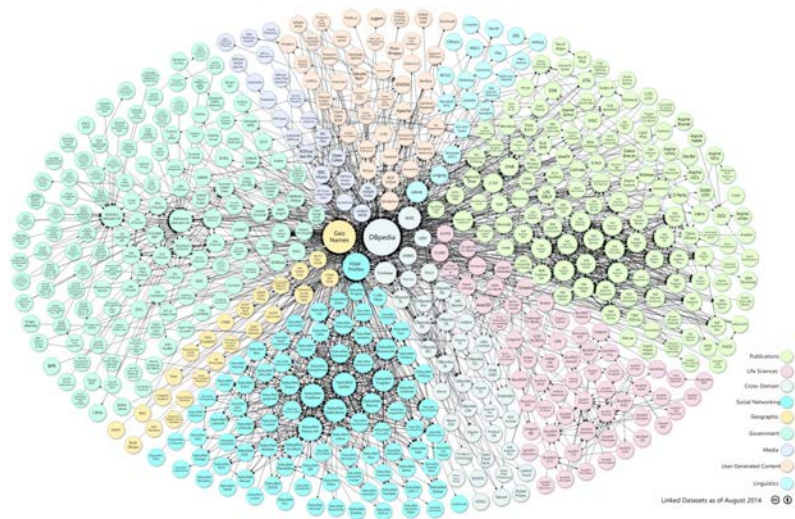


- General domain oriented
- Commonsense knowledge
- Structured encyclopedia knowledge
- Emphasize the breadth of knowledge
- For general users



- Industrial domain oriented
- Industrial data
- Semantic industrial knowledge base
- Emphasize the depth of knowledge
- For industry users

What is Knowledge Graph (KG) – Rapid Growth and Lower Cost



Manual

High Order Logic

2014

Red nodes represent knowledge graph datasets in life science



Other pharmacological data resources in an integrated, interoperable infrastructure

Explore.

Researchers can use Open PHACTS to access vast amounts of pharmacological data, all from a single, simple interface

Build.

Developers get free access to the Open PHACTS API, to query the pharmacological data resources in our integrated triple store

Join.

Members of the Open PHACTS Foundation get prioritised access to data, support and updates, as well as training opportunities

CYC - \$5.71 per statement
 Freebase - \$2.25 per statement
 NELL - 14.25 ¢ per statement
 DBPedia - 1.85 ¢ per statement
 Yago - 0.83 ¢ per statement

How much is a Triple?

ISWC2018

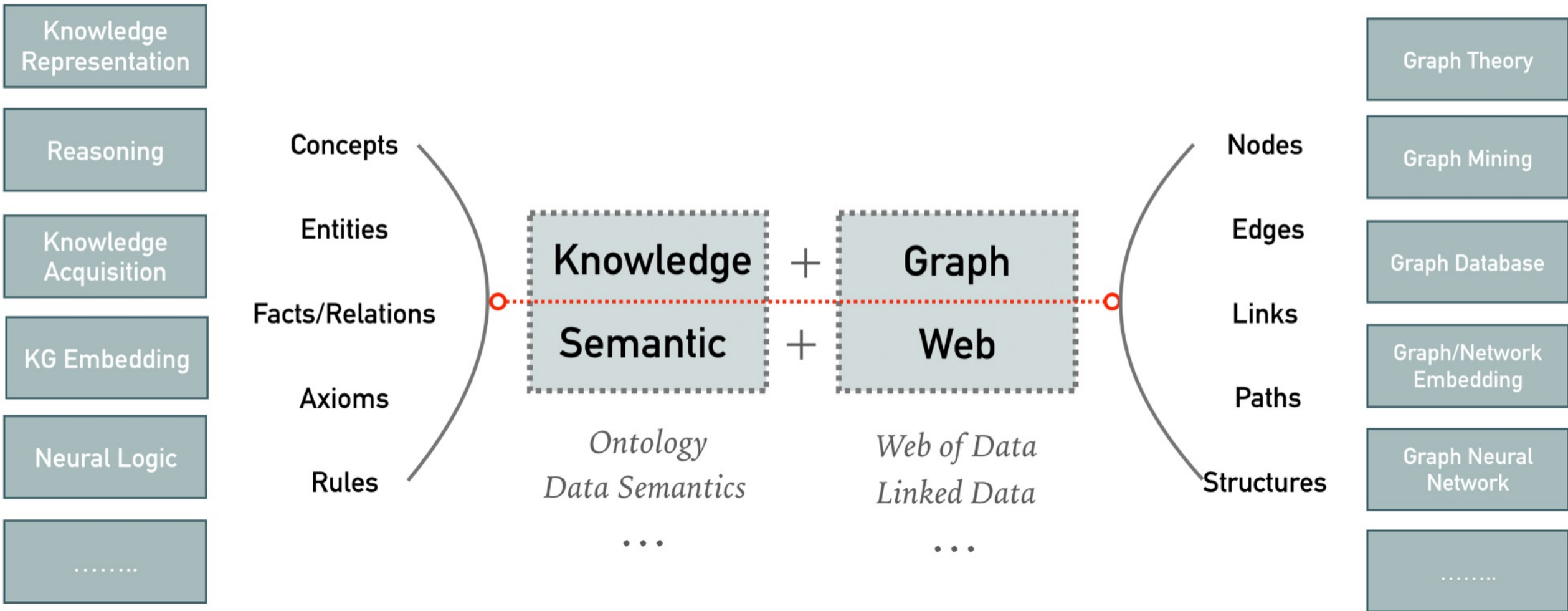
Automatic

Triples

Implication

From open to vertical domains, the **scale** of interlinked KGs has been grown **hundreds to thousands of times** in the past 15 years, the **cost** of extracting knowledge **is gradually decreasing**, improving **the quality** of extracted knowledge while **continuously** increasing the scale of knowledge is the main trend in the future

Knowledge Graph is more expressive than pure Graph but less complex than formal logic.



What is Knowledge Graph (KG) – Perspective and Implication

KG as a World Model

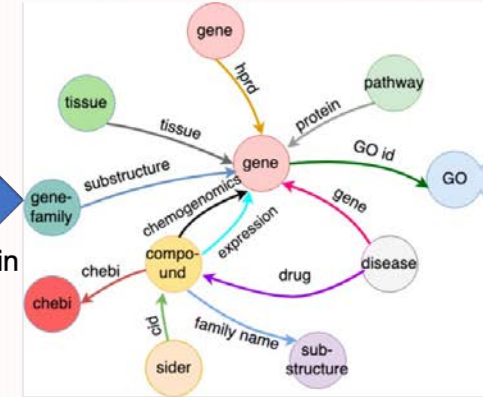


Text as Knowledge Base



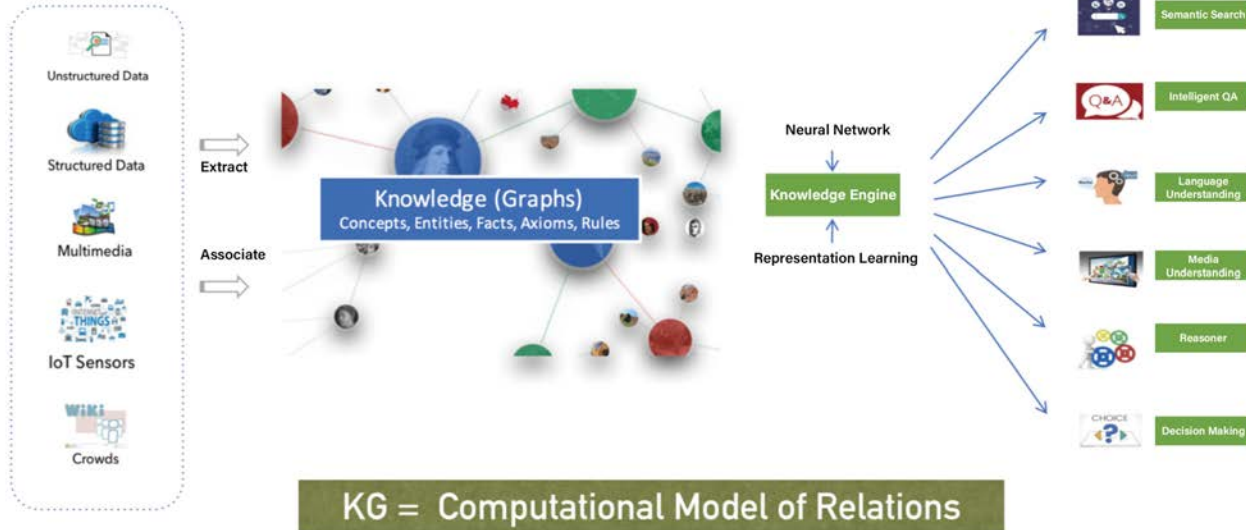
Text2Knowledge

Language Pretrain
BERT, GPT.....



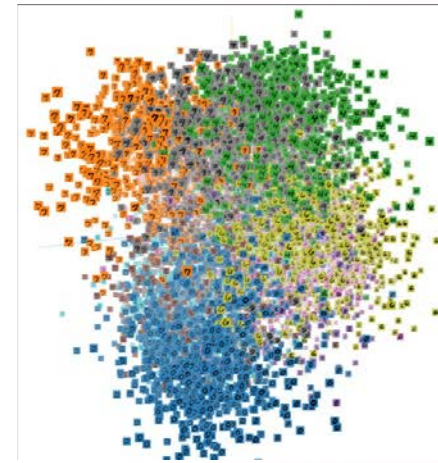
The Good Old Fashioned AI The Semantic Web & Linked Knowledge The Knowledge Graph

Graph Structure as Knowledge Base

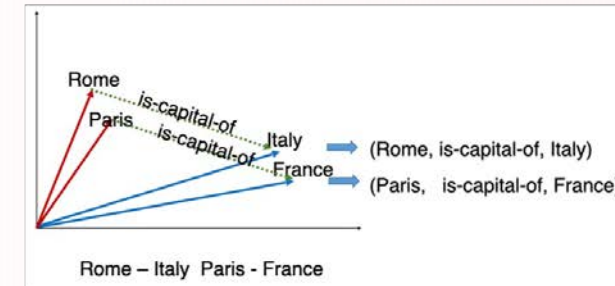


Embeddings as Knowledge Base

Embeddings : Distributed Vector Representation



- Text : Learn a vector of each word in a sentence
- KG: Learn a vector for each entity or property
- Image/Video : Learn a vector for each visual object



Implication

Knowledge Graph originates from how machines represent knowledge, use the graph structure to describe the relationship between things, developed in the rise of Web technologies, and landed in application fields such as search engine, intelligent QA, and recommender systems.

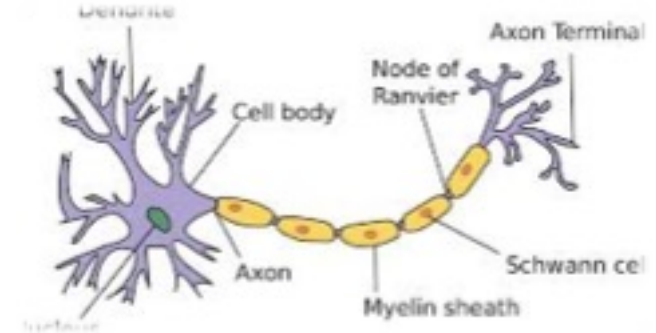
Smart AI vs. Knowledgeable AI

Smart
AI

perception
recognition
judgment



Deep Learning



Human brain can conduct reasoning and understanding based on acquired knowledge

Knowledgeable
AI

thinking
language
reasoning

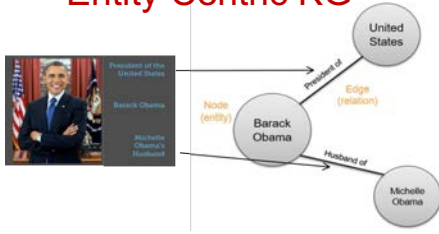


Knowledge Graph



SOTA and Trend of KG – Knowledge Representation and Reasoning

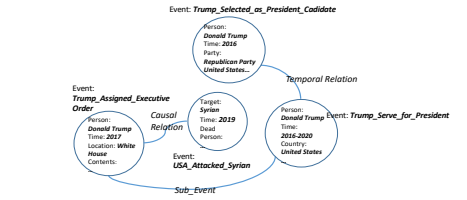
Entity Centric KG



(Barack Obama, Spouse, Michelle Obama)

Triples in form of (s, p, o)

Event Centric KG



Event Frames
 Birth: birth date, birth place, name
 Marriage: date, location, male, female
 EndPosition: time, company, position

Event Relations
 Causal relation, Temporal relation, Co-reference relation, Sub-class relation...

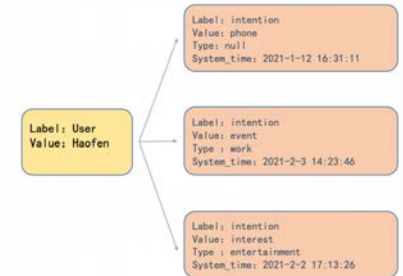
Prompt KG

```

1 Translate English to French:
2 sea otter => loutre de mer
3 peppermint => menthe poivrée
4 plush girafe => girafe peluche
5 cheese =>
    
```

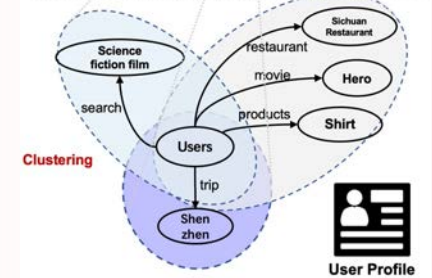
LLM as knowledge base and use prompts to elicit the corresponding facts

Personal KG



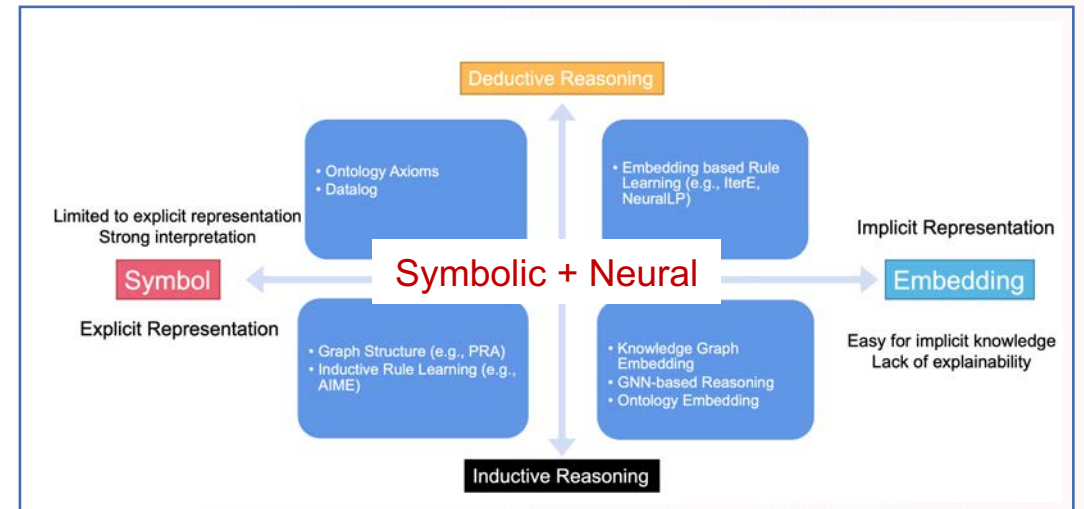
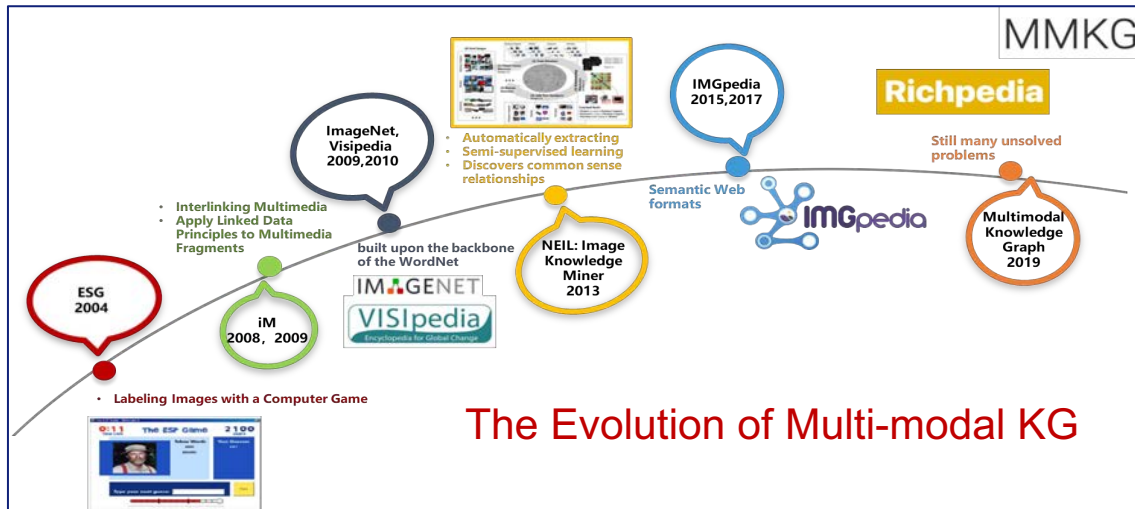
Graph composing events, interests, behaviors of an individual under the protection of privacy and security

Spatiotemporal KG



Expand knowledge representation level in time series and space dimensions

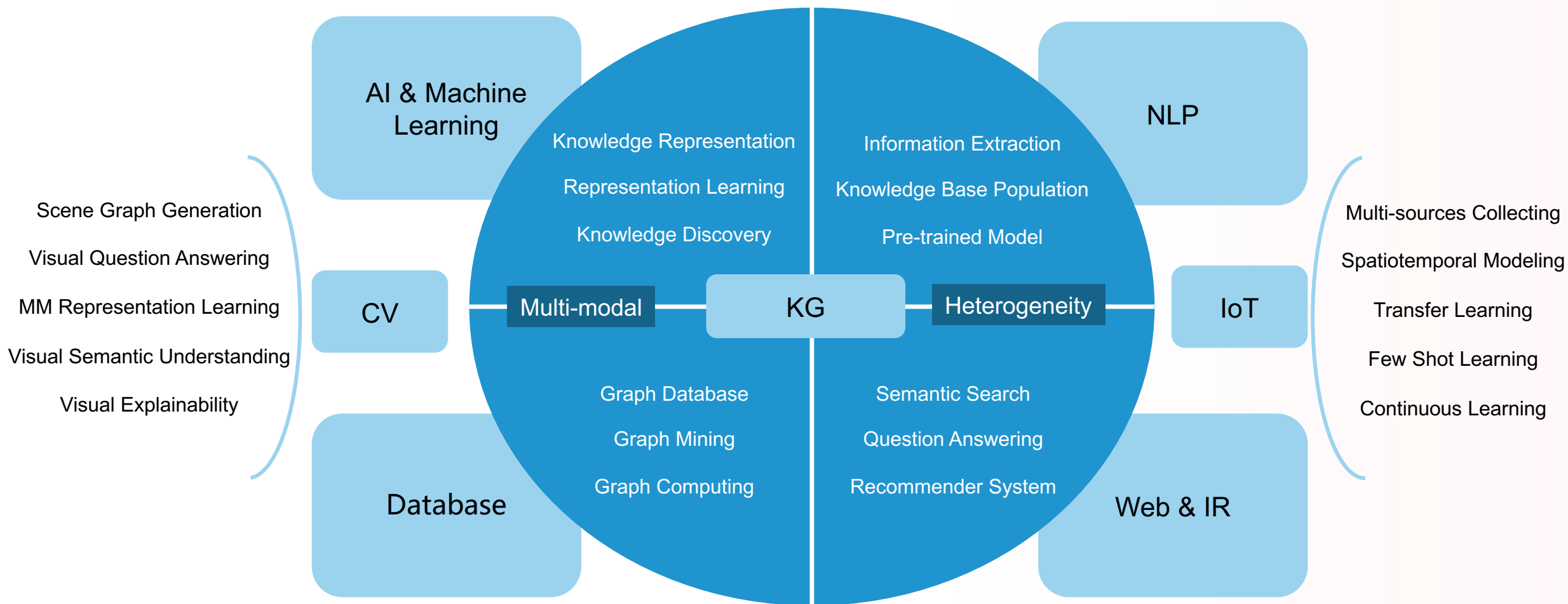
Knowledge types: simple -> complex, static -> dynamic, community -> personal, plain -> spatiotemporal



Challenges

Traditional symbolic knowledge representation methods are difficult to accurately represent complex knowledge such as **dynamics**, **processes**, and **cross-modalities**. At the same time, how to **combine symbolic reasoning** methods based on knowledge graphs and **neural reasoning** methods is extremely challenging.

SOTA and Trend of KG – Interdisciplinary



The life cycle of KG construction: more types/sources, advanced techs, rapid updates, and widely used applications

Challenges

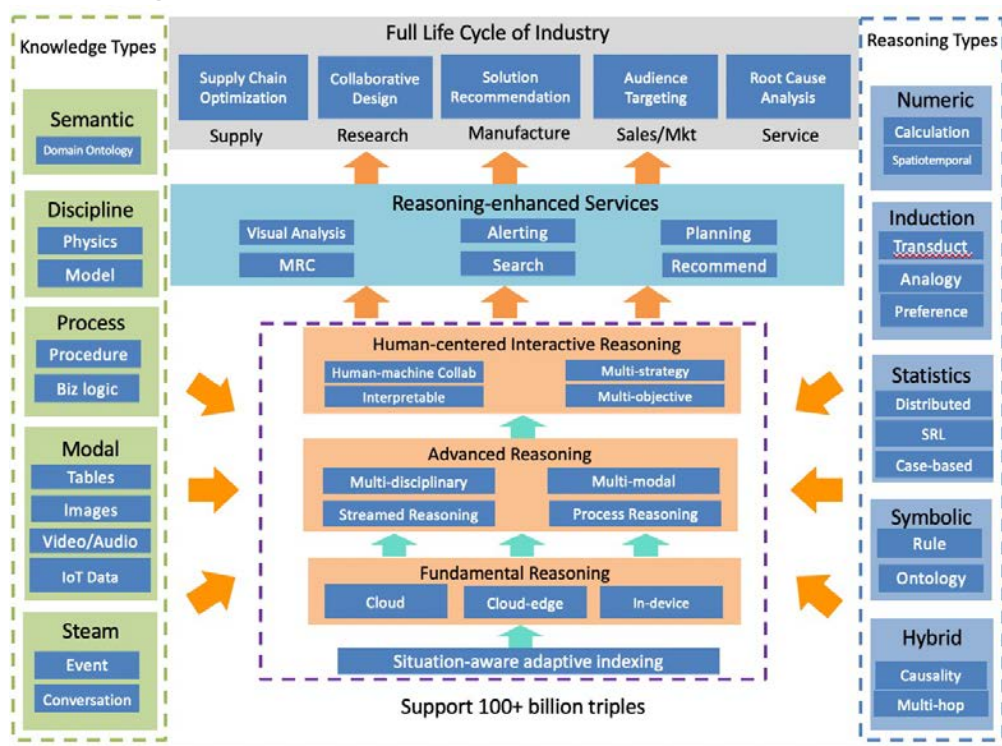
The **multi-scale**, **multi-modal**, and **multi-disciplinary** characteristics of data have put forward new requirements for knowledge representation, collection, extraction, storage, computing, and application. Among them, it is necessary to overcome **few shots**, **explainability**, and **domain adaptation** issues. At the same time, how to realize **knowledge update at a low cost** is also extremely

SOTA and Trend of KG – System Engineering View

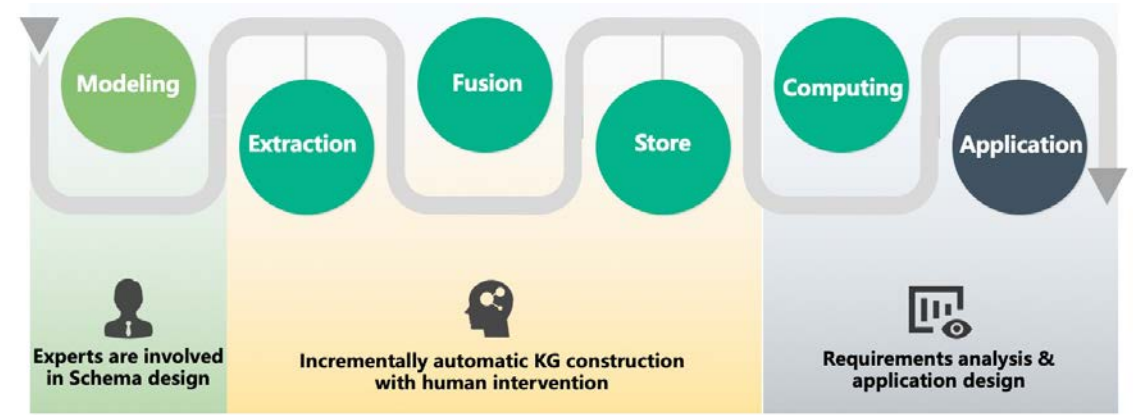


AI should focus on small data and **data centric AI**. Especially in the manufacturing industry, we must rely on **domain knowledge**

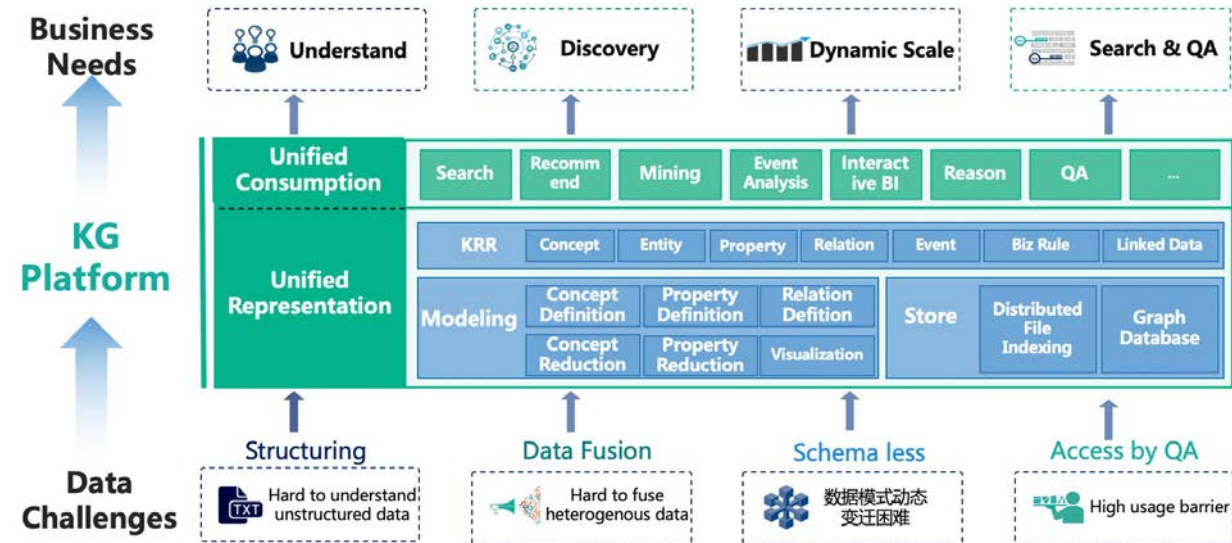
Andrew Ng



Knowledge Graph **System Architecture** in Industry



Different computing manners, "offline - near real time - real-time", depend on the type of knowledge



New Paradigm of Technology with Knowledge as the Core

Challenges

Data characteristics and knowledge differences in different fields lead to **low knowledge coverage**, **intensive labor input**, **shallow usage** in applications, **poor computing efficiency**, **difficult & weak sustainable operation** and **long time cost**

Trends of the Interdisciplinary Development of KG – Applications

Search

The image shows two Google search results. The first search is for 'tim berners lee', showing a knowledge card with a photo and a search results snippet. The second search is for 'how old is yao ming's wife', showing a result for 'Ye Li / Age' with the answer '38 years' and a photo of her.

Search Results for 'tim berners lee':

- Knowledge card: **Tim Berners-Lee**, Engineer. Born: June 8, 1955 (age 65 years), London, United Kingdom. Education: The Queen's College (1973-1976), MORE. Awards: Turing Award, Millennium Technology Prize, MORE. Parents: Conway Berners-Lee, Mary Lee Woods. Children: Ben Berners-Lee, Alice Berners-Lee.
- Search results snippet: **Tim Berners-Lee - World Wide Web Consortium**. Biography. Sir **Tim Berners-Lee** invented the World Wide Web in 1989. He is the co-founder and CTO of Inrupt.com, a tech start-up which uses, promotes and helps develop the open source Solid platform. Solid aims to give people control and agency over their data, questioning many assumptions about how the web has to work.
- Recommendation: **Famous computer scientists**. Alan Turing, Larry Page, Vint Cerf, Ada Lovelace, Robert E. Kahn, Grace Hopper, Steve Wozniak.

Search Results for 'how old is yao ming's wife':

- Result: **Ye Li / Age**. **38 years**. November 20, 1981. Photo of Ye Li.
- People also search for: Yao Ming (39 years), Yao Qinlei (10 years), Fang Fengdi (72 years).

Question Answering

Machine Reading Comprehension

Mary journeyed to the den.
Mary went back to the kitchen.
John journeyed to the bedroom.
Mary discarded the milk.

Q: Where was the milk before the den?
A: Hallway

Brian is a lion.
Julius is a lion.
Julius is white.
Bernhard is green.

Q: What color is Brian?
A: White

Sam walks into the kitchen.
Sam picks up an apple.
Sam walks into the bedroom.
Sam drops the apple.

Q: Where is the apple?
A: Bedroom

Multi-modal QA

At the base of a muddy ditch is the first primrose of my spring - glowing in the grey, a little spot of hope, brave, beautiful and perfect.



Hi Chris, wow well spotted with the beautiful flower, I love walking alongside the river where there is a bluebell way



Love every photo. Especially the weeping willow.

Thanks. It's nice to enjoy the wildlife nature and walk all the way to Winchester's great scenery.



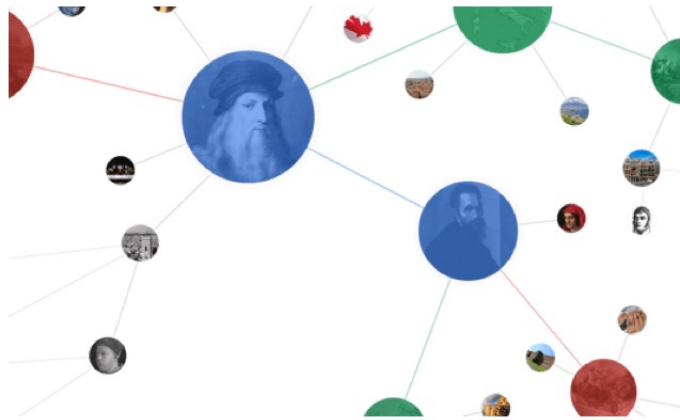
I live in Scotland. We have woods opposite with bluebells but not as thick as yours, but have a river with kingfisher, Heron and dipper.

Challenges

To build a multi-source and multi-modal knowledge graph, not only quality but also coverage should be considered. In the process of model training, **the alignment of heterogeneous and multimodal knowledge** is the difficulty of knowledge fusion and learning

Trends of the Interdisciplinary Development of KG – Applications

MORE MACHINE UNDERSTANDABLE



Data link

Knowledge representation

Computable data

Semantic extraction

Coarse data



text

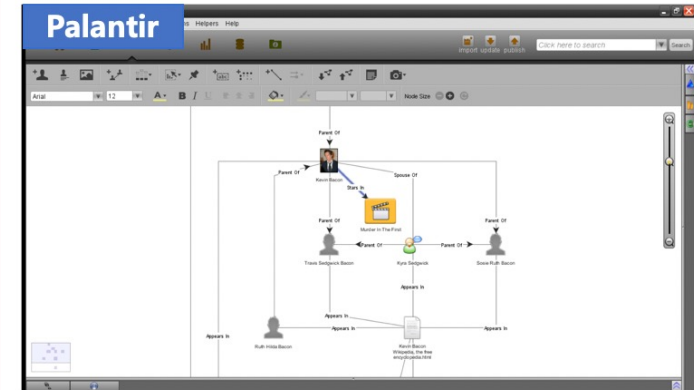
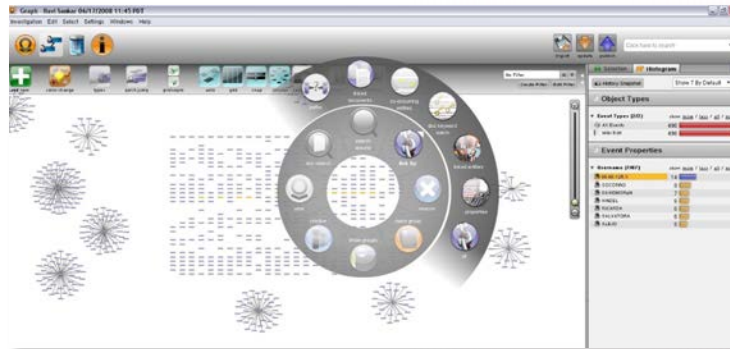
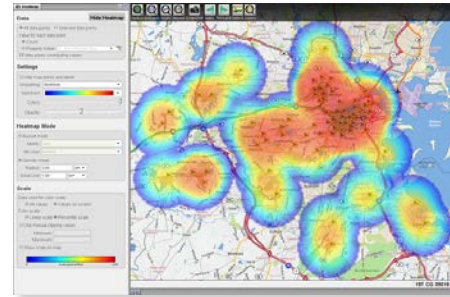
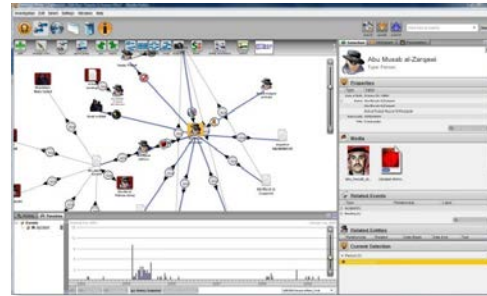


multimedia



sensor

Knowledge Graphs for Decision Making



<https://www.palantir.com/>



<https://www.kensho.com/>

Challenges

In each specific field, the explainability of the model and predictions are the most important to realize the application value. How to balance the advantages and disadvantages of **symbolic models** and **neural networks**, and learn from each other is a hot topic in academia and industry.

Life cycle of Knowledge Graph

- Top-down method
- Bottom-up method

- Schema graph fusion
- Data graph fusion

- Graph computing
- Ontology reasoning
- Rule-based reasoning

Reasoning is important !

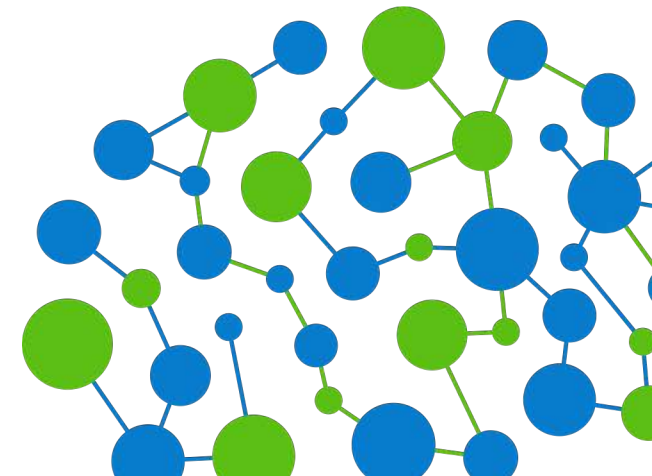


- Linked data: graph mapping
- Structured data: D2R
- Semi-structured data: wrapper
- Text: information extraction

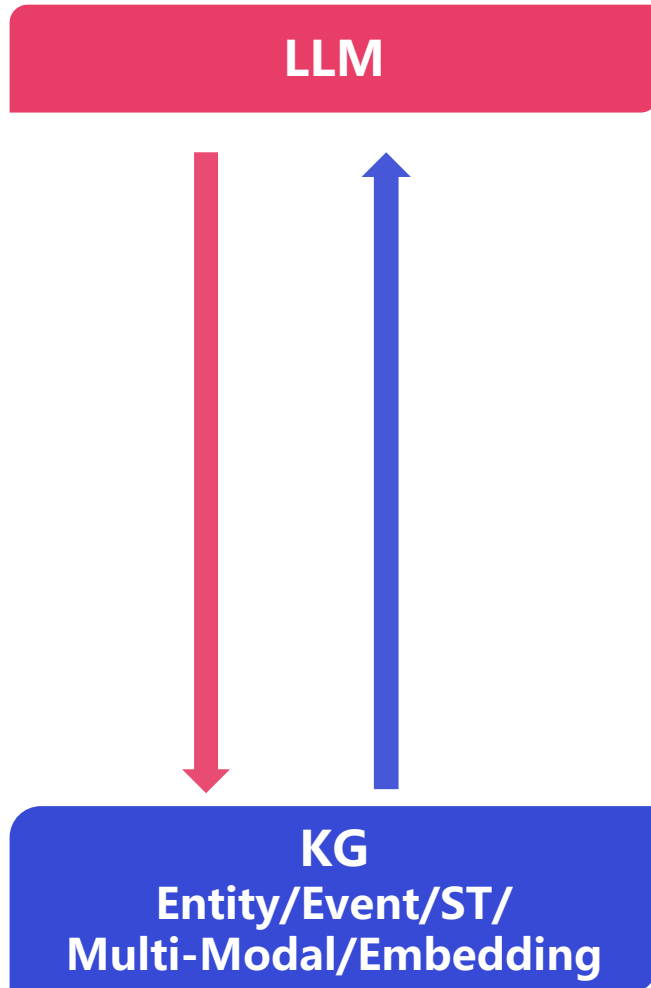
- Triples
- Event information
- Temporal information
- Multi-modal

- Semantic search
- Question answering
- Recommendation
- Assistant decision

- The Rise of LLM
- The Development of KG
- **The Emerging Paradigm**
- Outlook



When KG meets LLM: Two Sides of a Coin



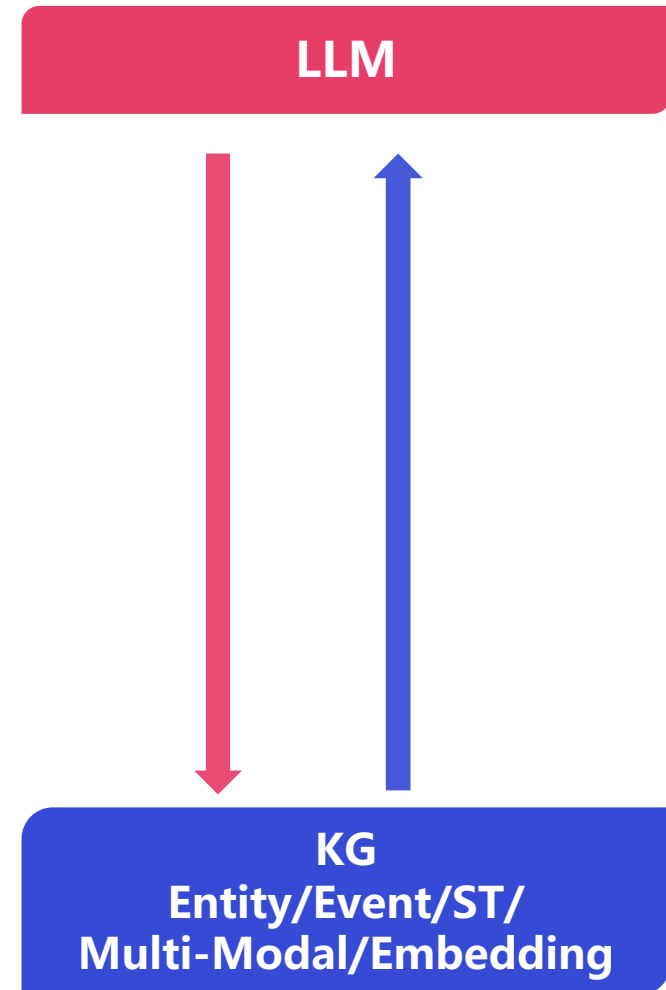
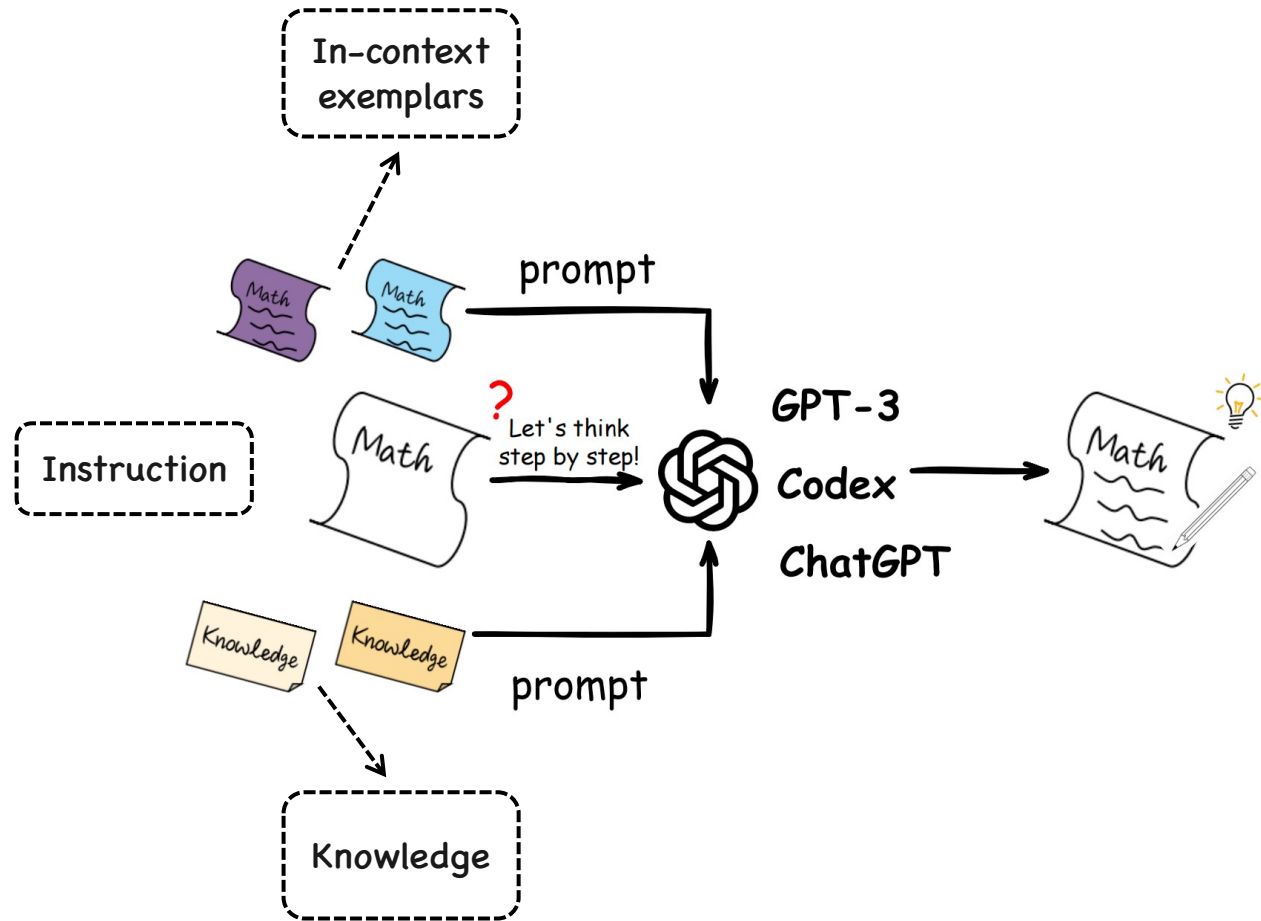
Boost Rapid KBP: Knowledge Extraction/Fusion

- Leverage the **understanding** and **generation** capabilities of LLM for automatic **extraction** and **fusion** of knowledge

Knowledge enhanced PTM/Prompt Engineering/Complex Knowledge Reasoning/Provenance/Integrating Dynamic Knowledge

- **Knowledge guided** automatic prompt engineering
- The emergent ability of LLM especially CoT , combined with **KG reasoning** to solve complex tasks
- Knowledge in form of triples, instructions, rules or codes can be injected into LM for data augmentation and bring more **trustworthy** and **interpretability**
- Linking generated responses of LLM to KG for **references, provenance or fact checking**

When KG meets LLM: Knowledge-guided Prompting



When KG meets LLM: Knowledge-guided Prompting

Implicit Knowledge

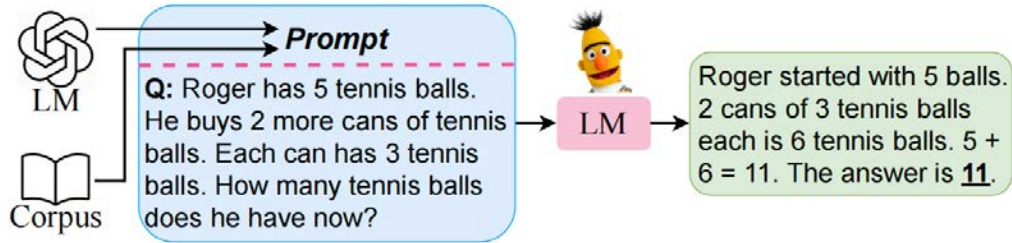
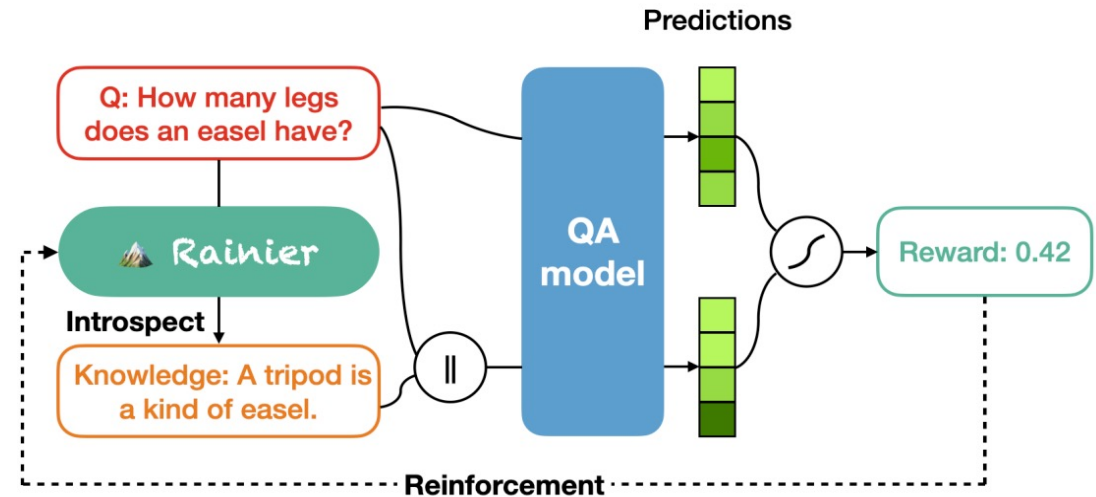


Figure 6: Knowledge Enhanced Reasoning (§3.2). Prompts are generated by LMs (**Implicit Knowledge**) or retrieved from external resources (**Explicit Knowledge**).

- LLM contains massive implicit knowledge
- Knowledge distillation



Few-shot prompting over GPT for downstream tasks like QA, and introspect knowledge via reinforcement learning

When KG meets LLM: Knowledge-guided Prompting

Explicit Knowledge

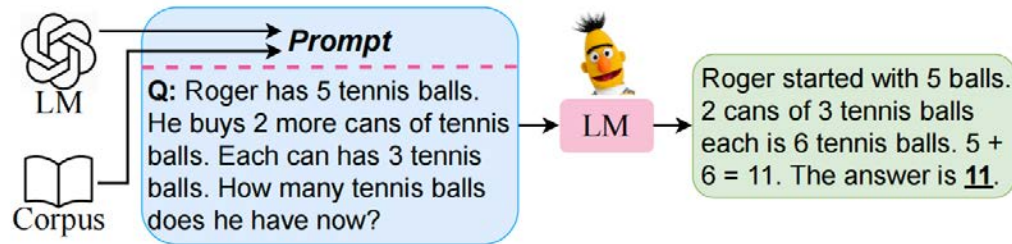
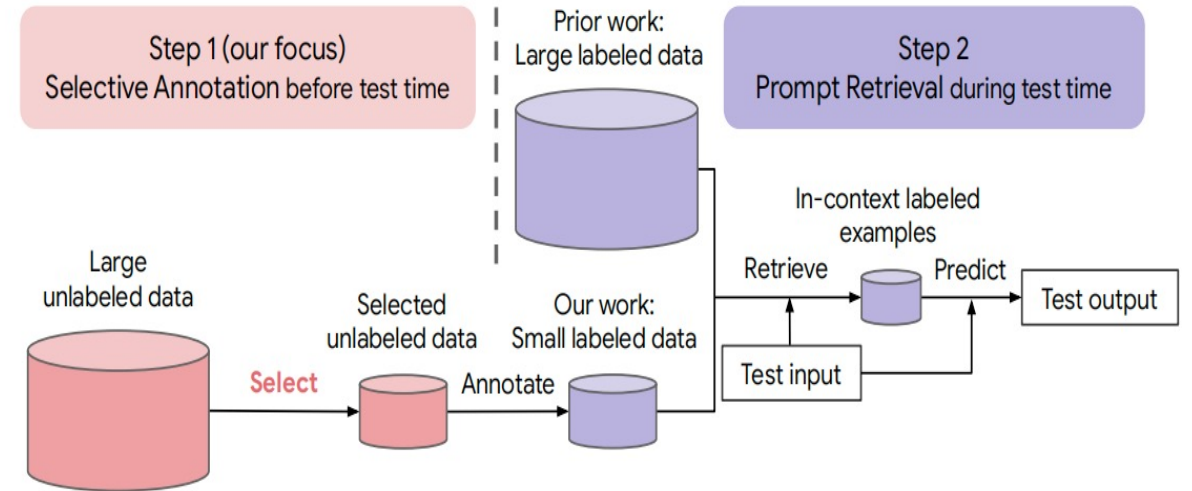


Figure 6: Knowledge Enhanced Reasoning (§3.2). Prompts are generated by LMs (**Implicit Knowledge**) or retrieved from external resources (**Explicit Knowledge**).

- Despite exhibiting strong generative abilities, LLM tends to hallucinate and generate inconsistent knowledge
- Retrieved from external explicit knowledge



Develop a selective annotation framework to avoid reliance on retrieval of large-scale annotated corpora

When KG meets LLM: Reasoning Capability Comparison

LLM Reasoning

- **Code pre-training:** enhance LLM reasoning during training
- **Prompt Engineering:** eliciting LLM reasoning during inference

KG Reasoning

- Graph computing
- Rule-based reasoning
- Ontology reasoning
- Spatial-temporal reasoning
- KG embedding/GNN

LLM Reasoning

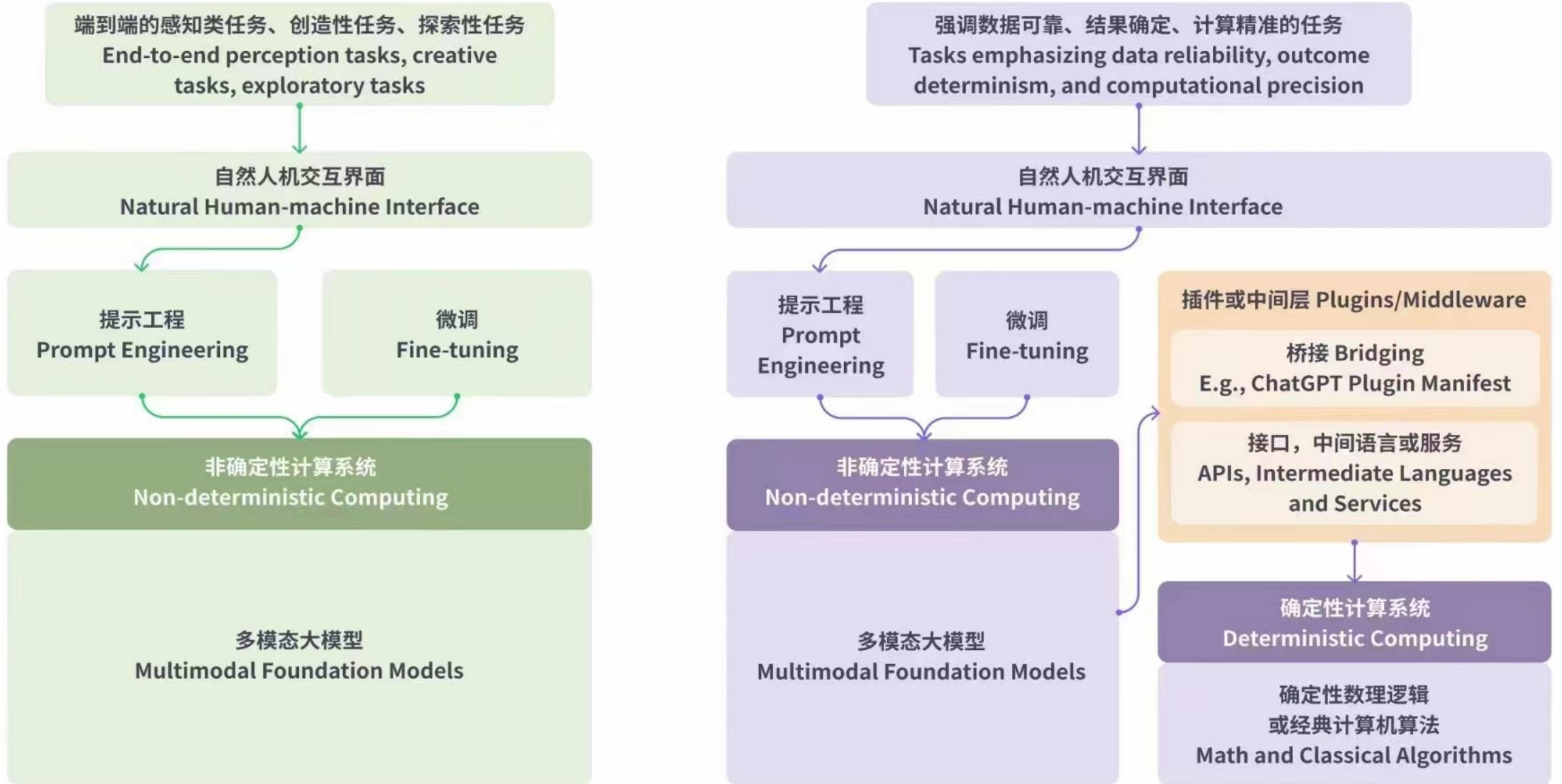
- zero-shot prompting
- Few-shot prompting
- CoT prompting
- Instruction



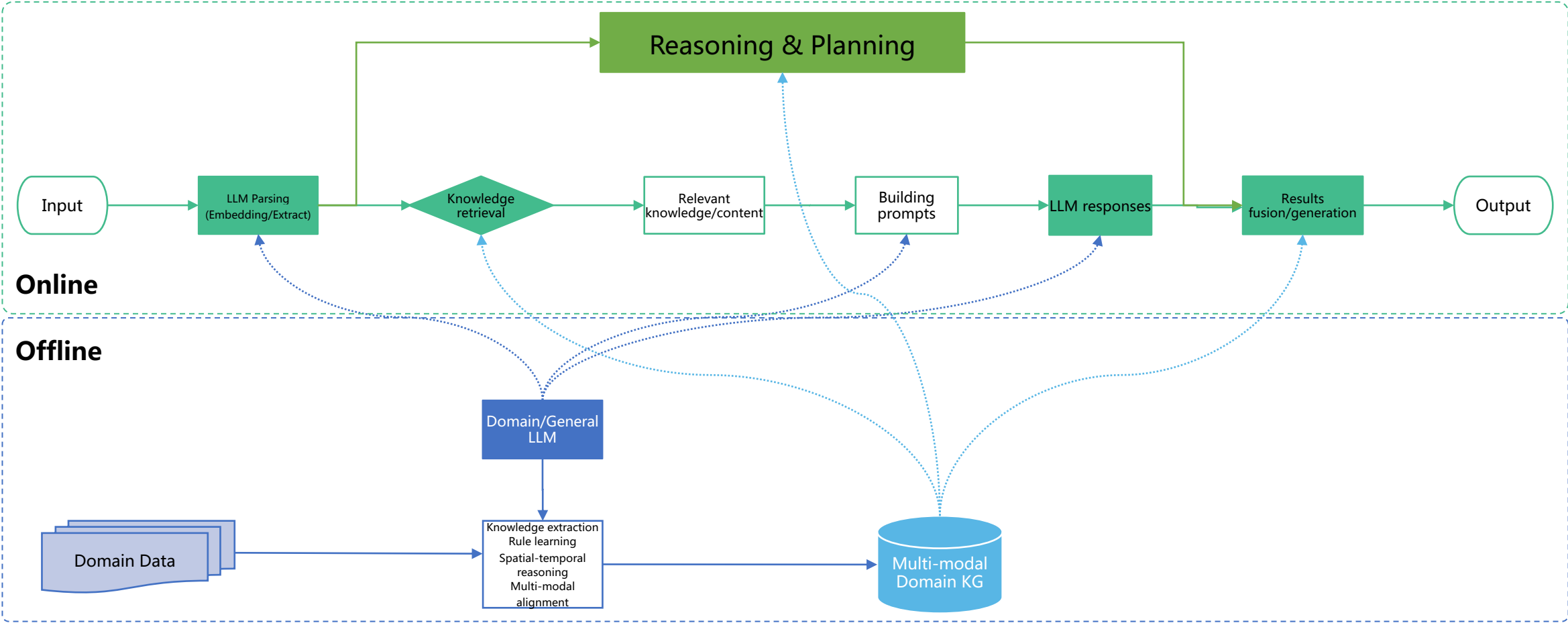
KG Reasoning

- Graph computing
- Rule-based reasoning
- Ontology reasoning
- Spatial-temporal reasoning
- KG embedding/GNN

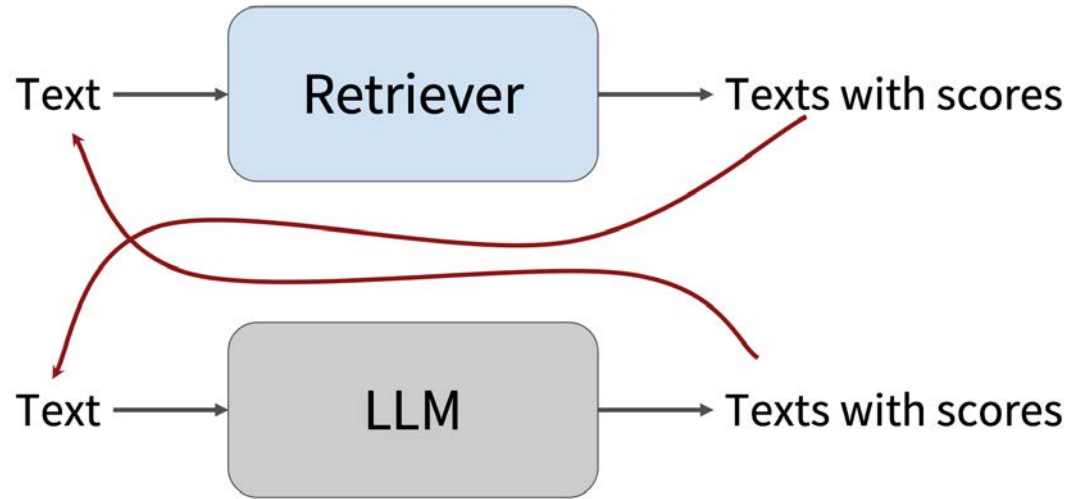
When KG meets LLM: Two Kinds of Tasks



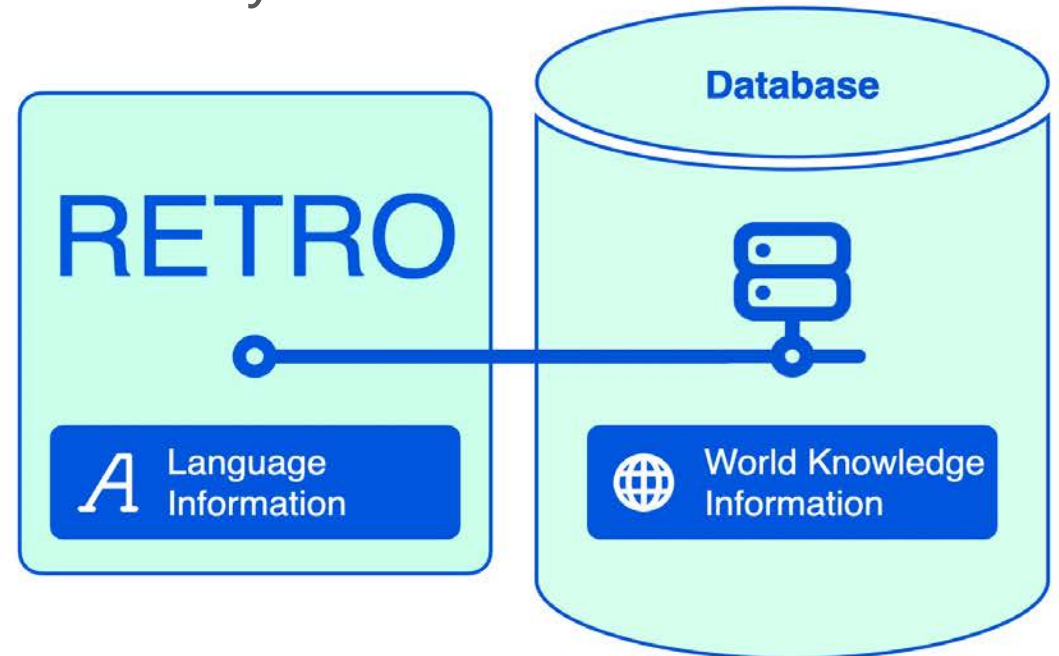
When KG meets LLM: Knowledge Augmentation



When KG meets LLM: Retrieval & Knowledge Augmentation



- Efficiency
- Updateability
- Provenance
- Effectiveness
- Synthesis



DeepMind's RETRO (**R**etrieval-**E**nhanced **T**Ransf**O**rmer)

When KG meets LLM: Knowledge Augmentation



Which 2013 Los Angeles Galaxy player transferred in from the team with 12 international titles ?

Consolidate evidence from external knowledge

WIKIPEDIA Transfers

No.	Pos.	Player	Transferred from	Fee/No
1	GK	Carlo Cudicini	Tottenham Hotspur	Free
	MF	Juninho	São Paulo	Undisck

2013 Los Angeles Galaxy season

Juninho (footballer, born January 1989)

Juninho grew up in the city of São Paulo and played for the São Paulo youth team, winning the U-17 Paulista Championship side in 200 made one appearance for the São Paulo side in 2007. He was sent out on loan to Major League Soccer team Los Angeles Galaxy in 2008 along with fellow Brazilians from the club Al Cazumba and Leonardo.^[1] He made his deb for the team on 27 March 2010, in Galaxy's opening game of the 2010 MLS season agai New England Revolution,^[2] and scored his f goal for the Galaxy in a 2-0 win over AC St. Louis

São Paulo FC is one of the most successful clubs in Brazil, having won a total of 30 domestic honours, in addition to their 12 international successes. It is the Brazilian club with the most international titles.

Worldwide		
Intercontinental Cup	2	1992, 1993
FIFA Club World Cup	1	2005

Continental		
Competitions	Titles	Seasons
Copa Libertadores	3	1992, 1993, 2005

Revise response via automatic feedback

Candidate response:
Jaime Penedo is transferred in from C.S.D. Municipal, a team with 12 international titles.

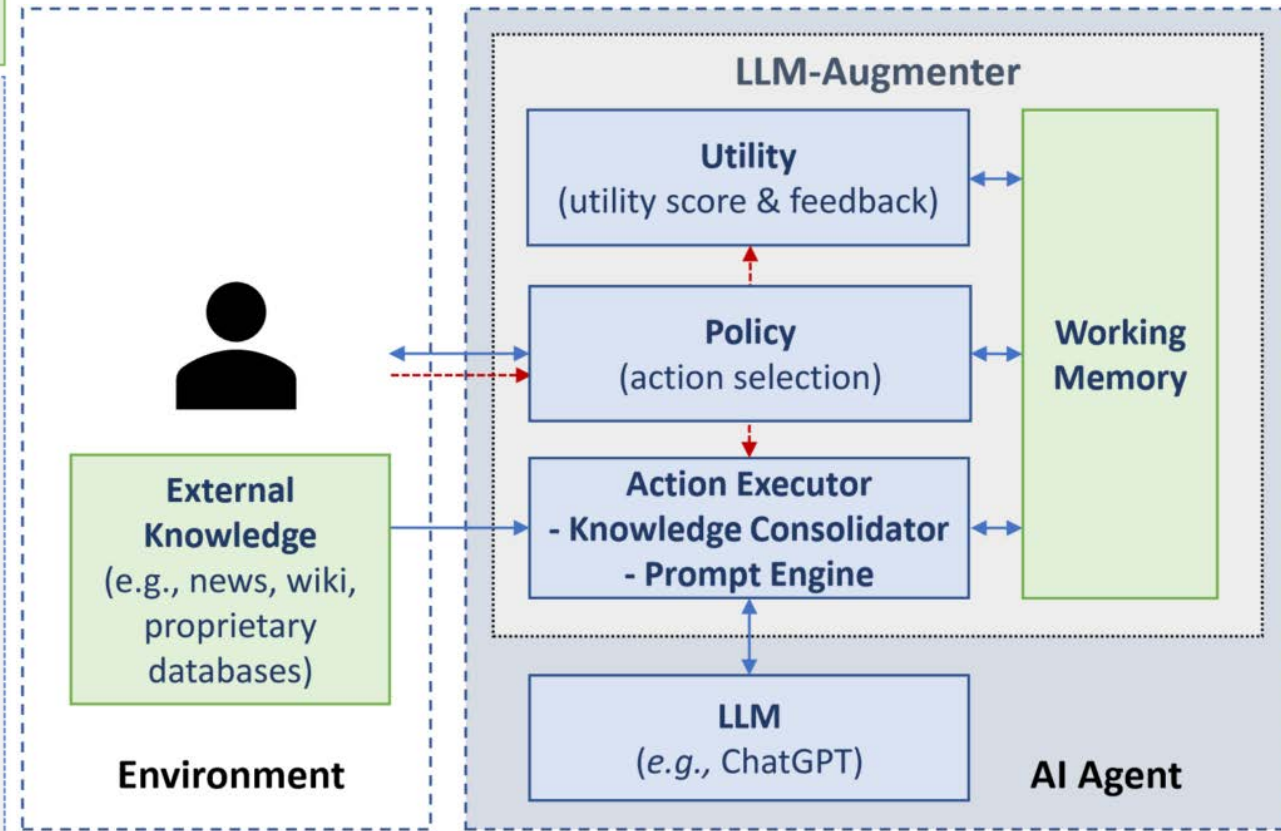
Feedback:
The player Jaime Penedo is transferred in from C.S.D. Municipal, but there is no information about the number of international titles of this team.

Revised candidate response:
Juninho is transferred in from São Paulo, a team with 12 international titles.

AI Agent (LLM-Augmenter + LLM)



Juninho is transferred in from São Paulo, a team with 12 international titles.

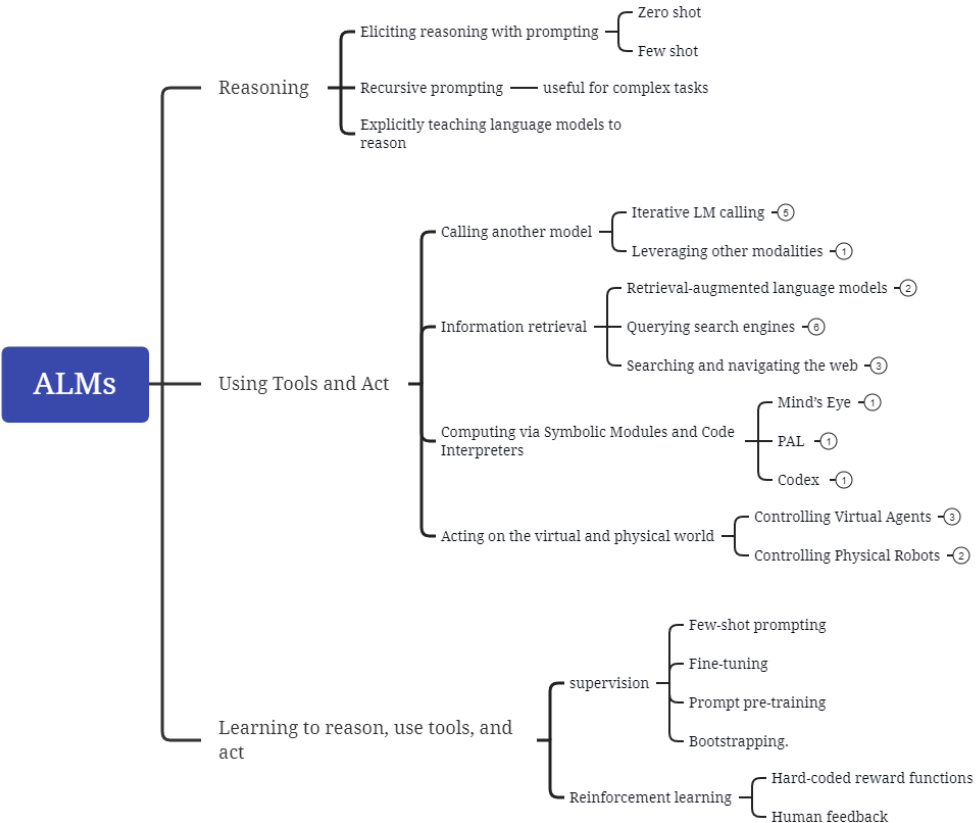


Check Your Facts and Try Again: Improving Large Language Models with External Knowledge and Automated Feedback 2023

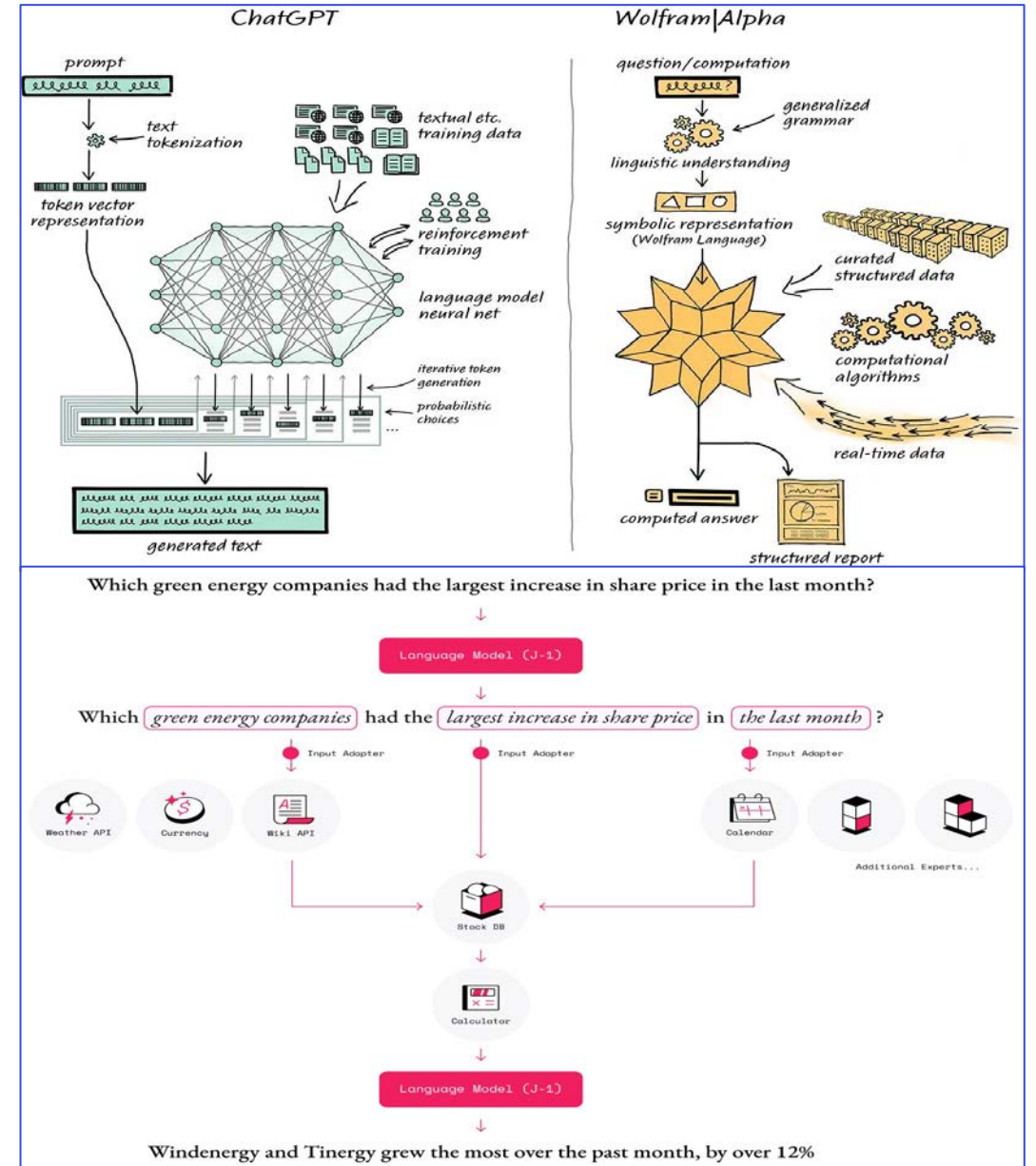
Augmented Language Models

Important Note: This week, the WL will be op

ToolFormer



Source: Augmented Language Models: a Survey (Yann Lecun et al.)

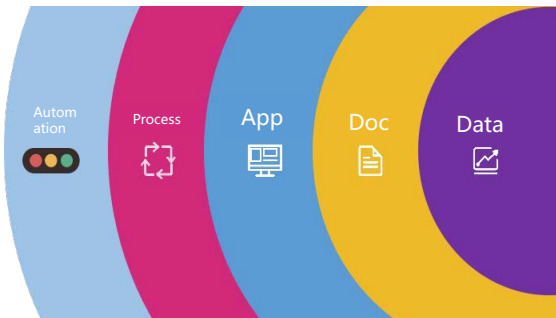
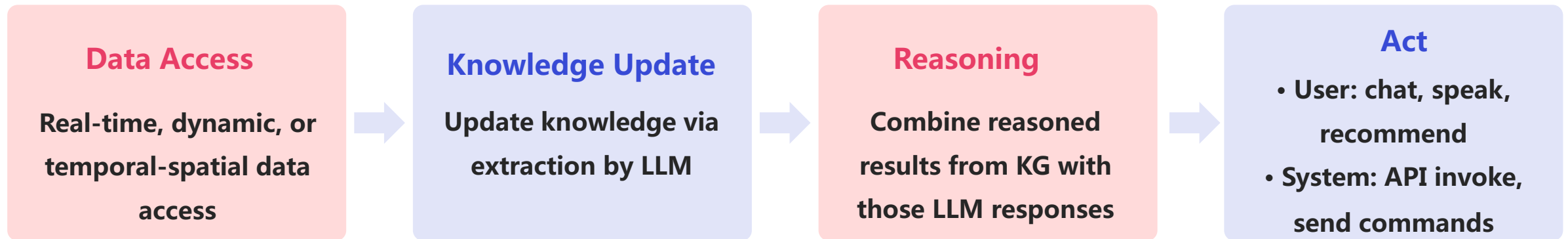


When KG meets LLM: Interoperability & Orchestration

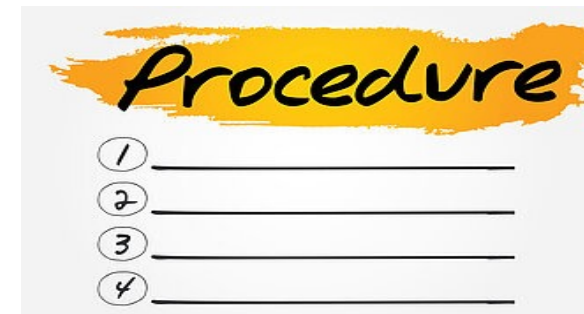
KG uses ontology to realize the representation of domain data, knowledge and interaction, and completes the automation of the whole process from real-time data access, knowledge update to user interaction.

Interoperate & Orchestrate

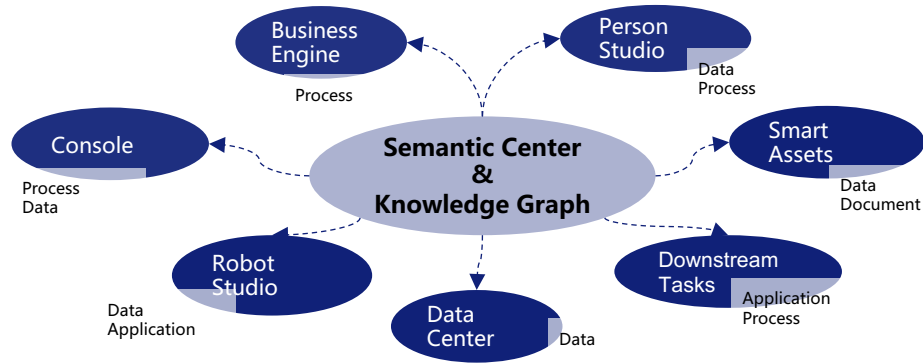
Integrate the following steps into a whole end-to-end process for hyper-automation



- Task mining
- Consistency checking
- Process discovery
- Modeling via simulation
- Case prediction
- Suggestion via history data

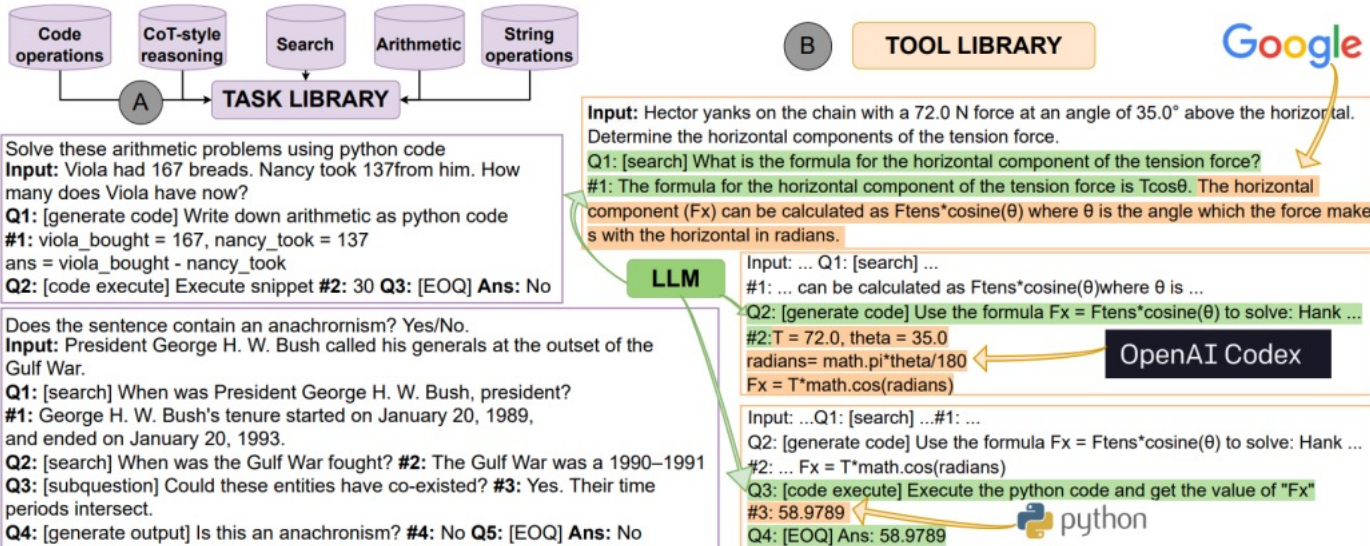


When KG meets LLM: Interoperability & Orchestration

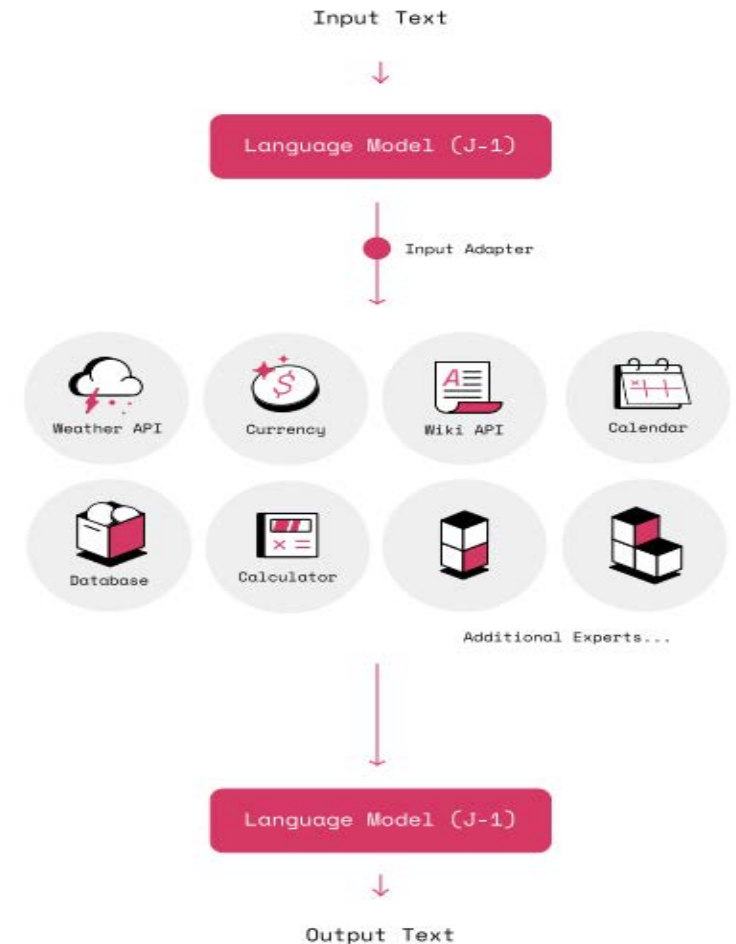


RPA + Tool use

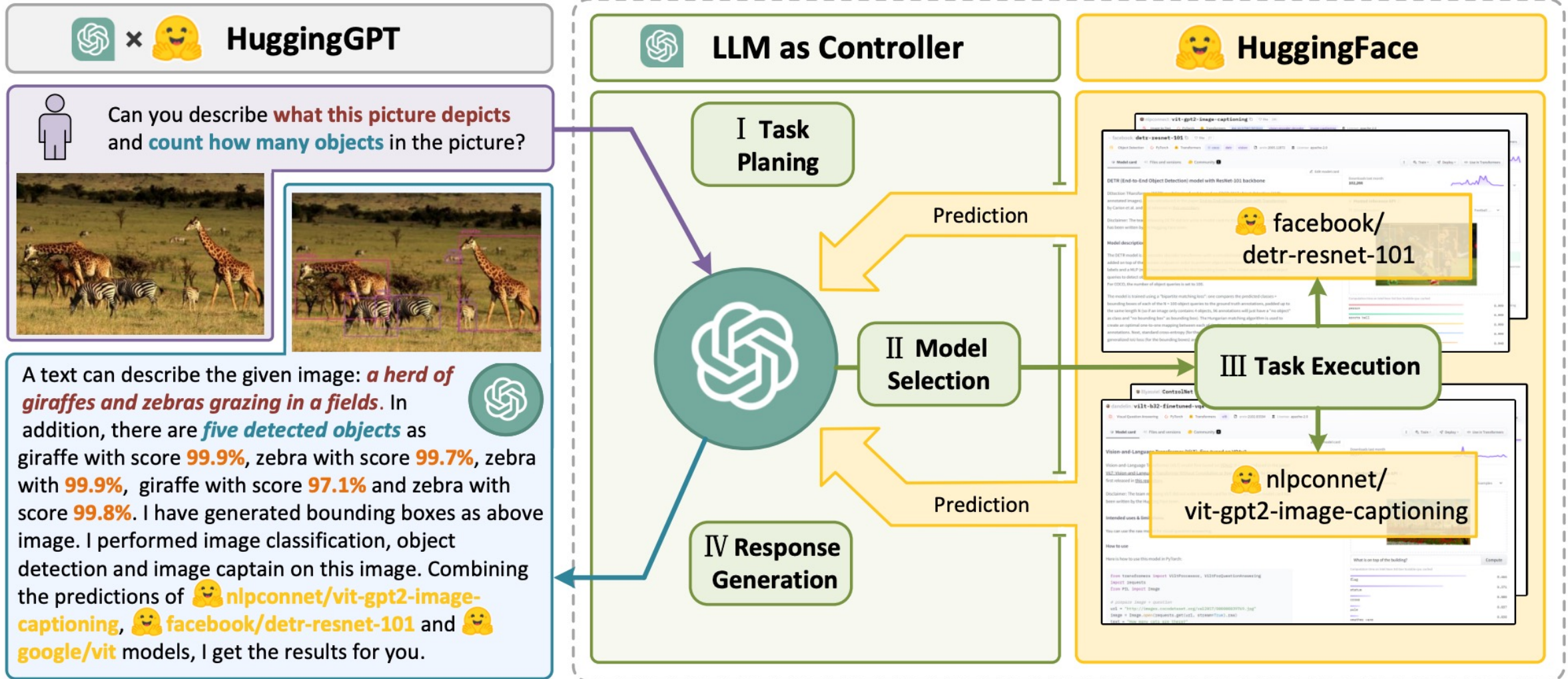
New Task (Physics QA) Answer this high-school physics question
Input: Hector yanks on the chain with a 72.0 N force at an angle of 35.0° above the horizontal. Determine the horizontal components of the tension force.



Adaptive Routing

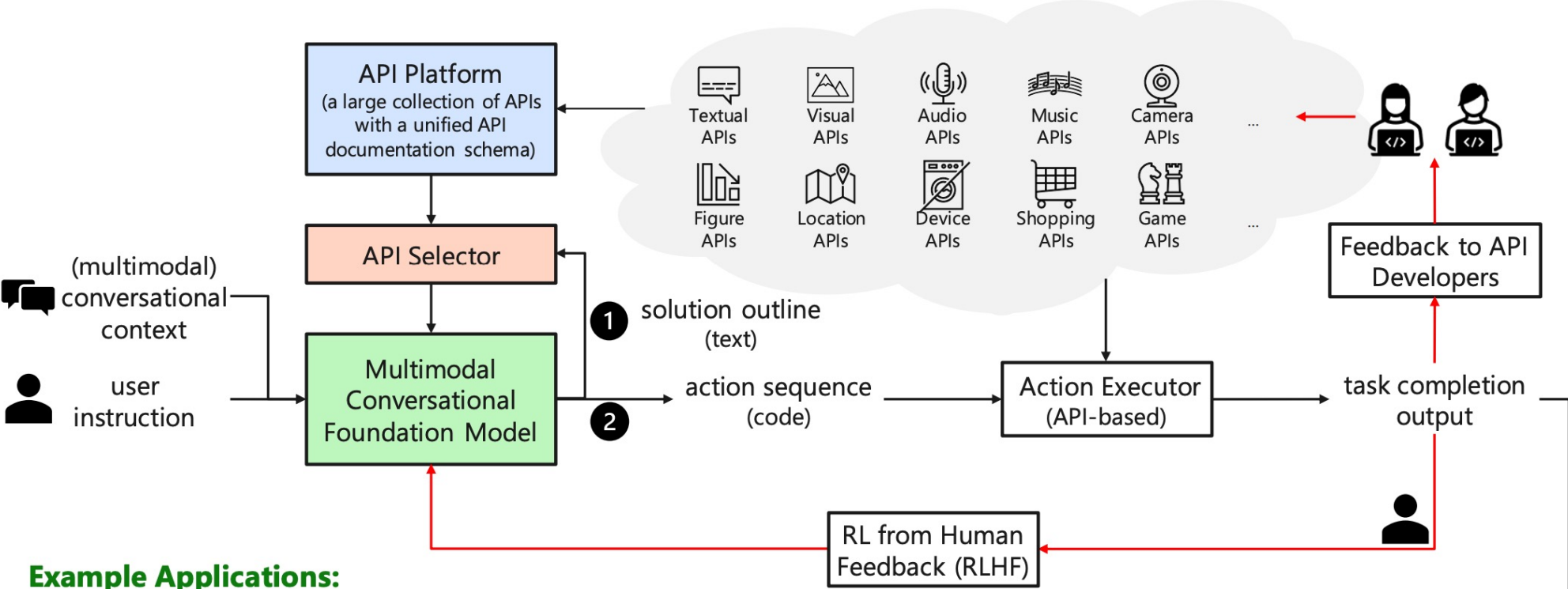


When KG meets LLM: Interoperability & Orchestration



HuggingGPT: Solving AI Tasks with ChatGPT and its Friends in HuggingFace 2023

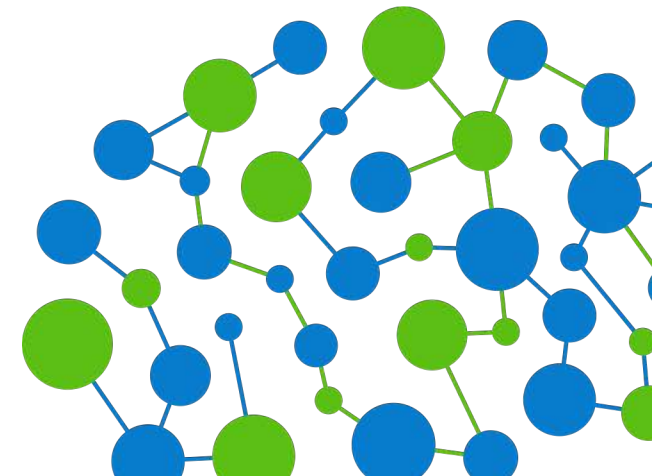
When KG meets LLM: Interoperability & Orchestration



Example Applications:

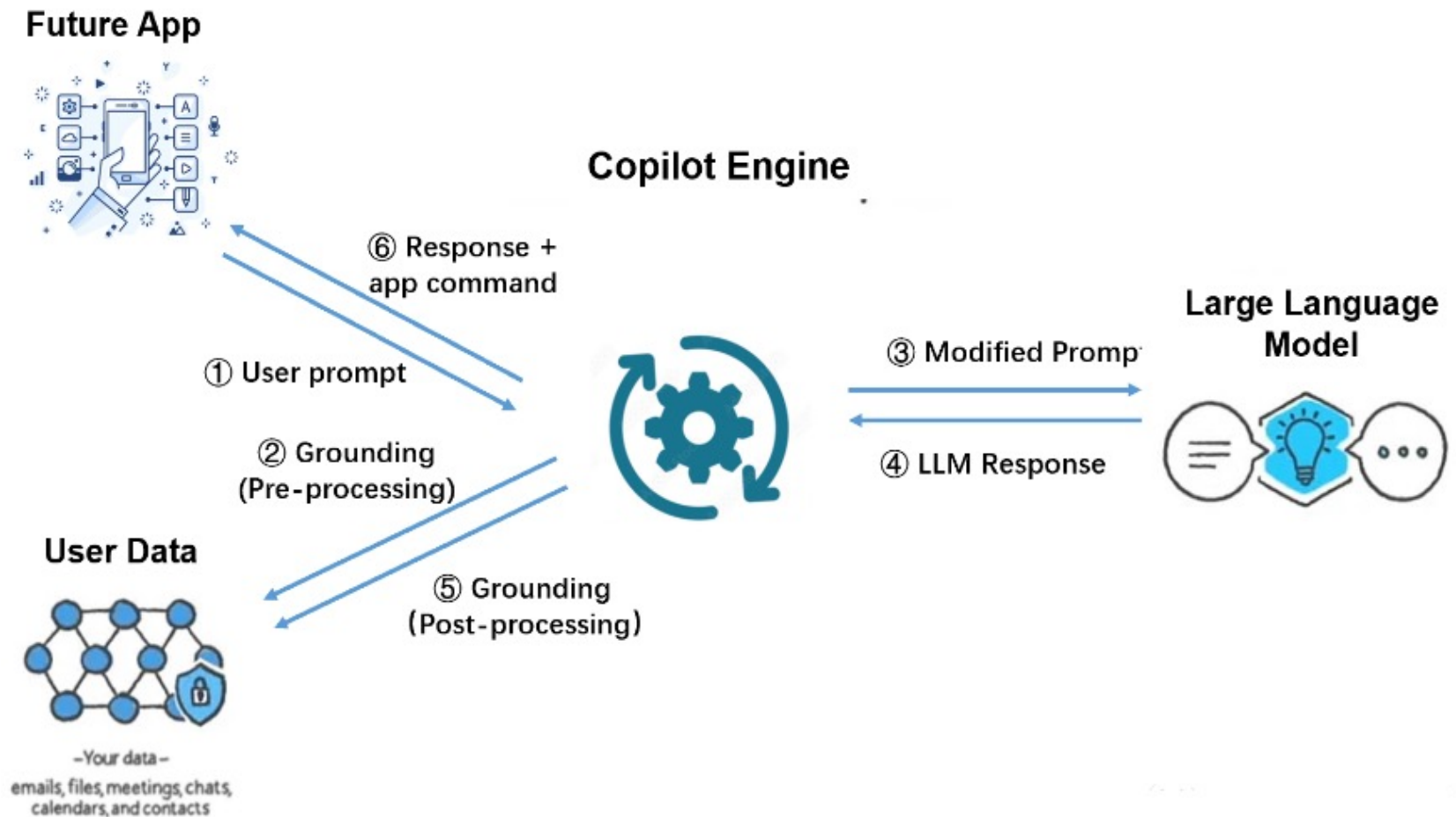
AI for Robotics AI for Office Automation AI of Things (AIoT) AI Copilots

- The Rise of LLM
- The Development of KG
- The Emerging Paradigm
- **Outlook**



Augmented Language Models: Copilot Engine

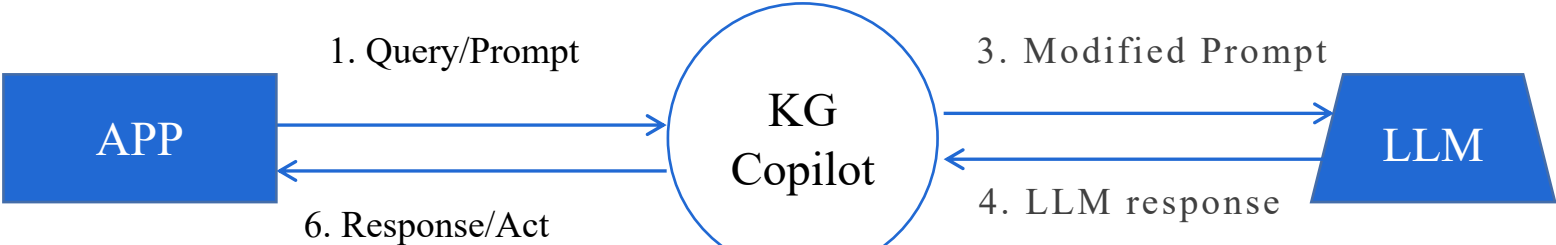
- ❖ Grounding (pre-processing): use Microsoft Graph to extract relevant user data, and treat them as a part of prompts
- ❖ Modified prompt: combine user prompt with data, and optimize the prompt for richer contexts to be fed into LLM
- ❖ Grounding (post-processing): perform post-processing for results from LLM with further validation and referencing
- ❖ Response + App command: convert results into commands for applications like Word, Excel, or PPT. For example, Office supports VBA and the output of a command can be any code.



Augmented Language Models: KG Copilot Engine

❖ **KG for unified knowledge modeling**

- ❖ Multi-modal data
- ❖ Multi-source data



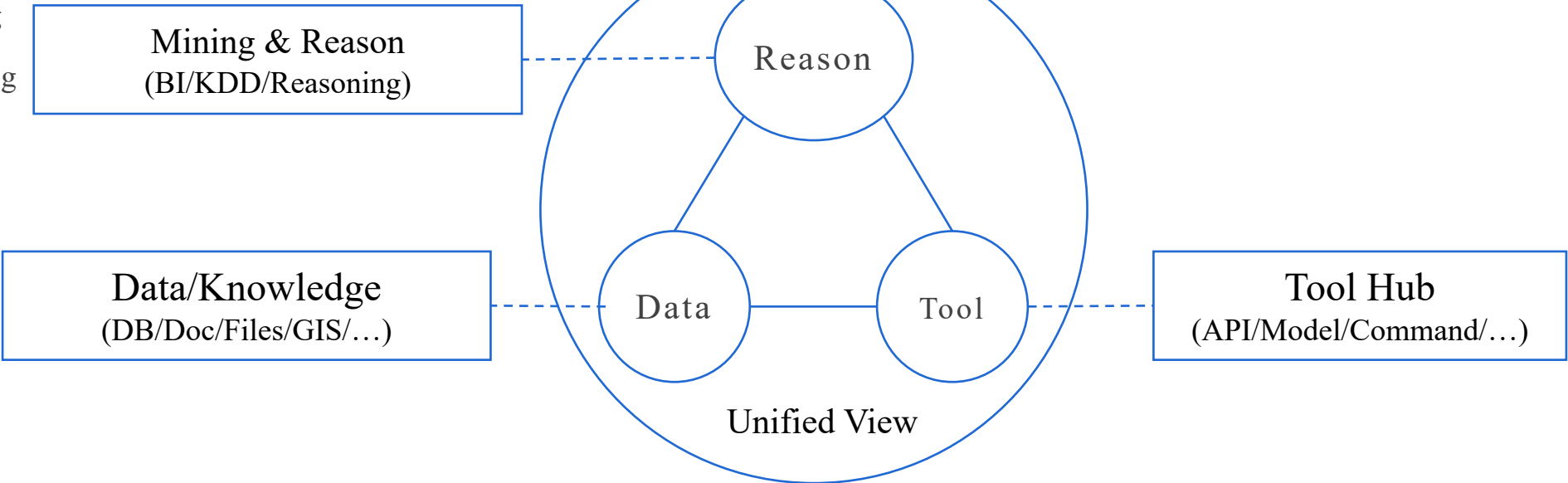
❖ **KG for unified programming as a semantic gateway**

- ❖ LLM + KG + Tool

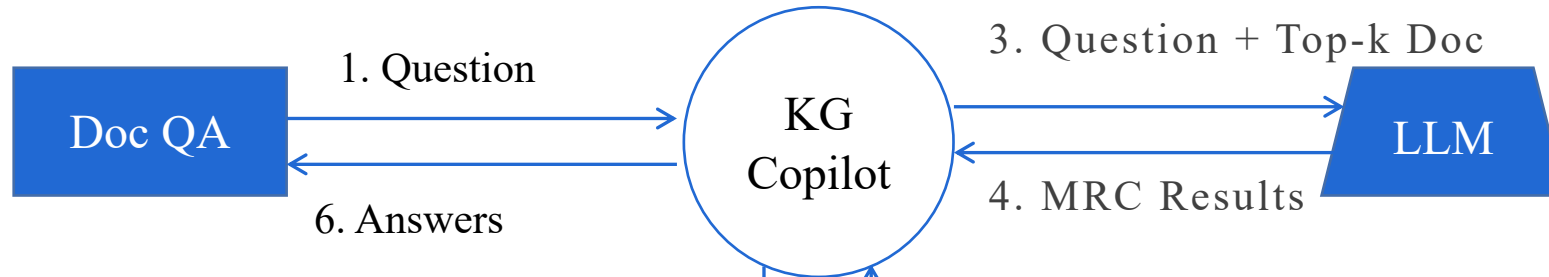
- a. Prompt build
- b. refinement
- a. Fact checking
- b. Data query
- c. Tool invoke
- d. Reasoning
- e. Result fusion

❖ **KG for automatic prompt building and result checking**

- ❖ Pre-processing grounding
- ❖ Post-processing grounding



KG Copilot Engine Example: ChatDocument



- a. Question Embedding
- b. Embedding Search
- c. + Similar Fragment

2. Pre-processing

5. Post-processing

- a. Referencing : Paragraph & Section Information
- b. Result Merging

GU

招标管理岗位的定位是什么? 请从以下文本中回答:

第二章 岗位定位

第四条 岗位定位

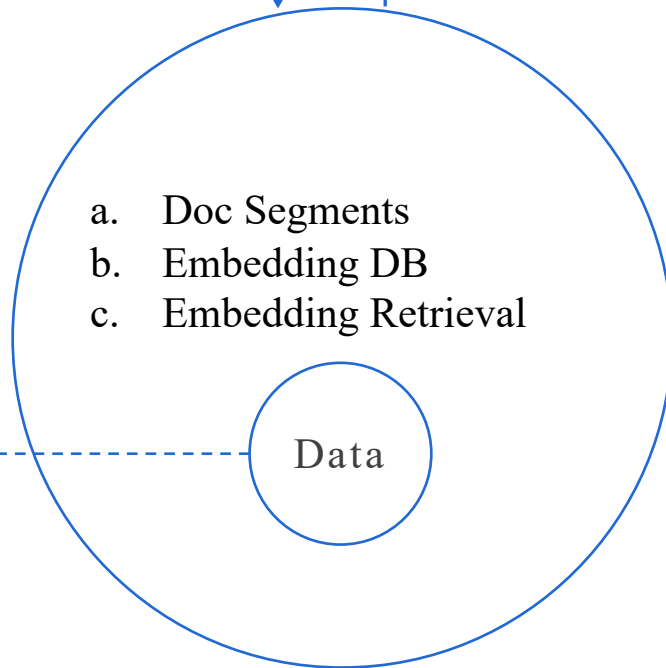
招标管理工作属于合约职能线的重要基础工作, 确保全景计划实现的关键业务。招标管理岗位是招标管理工作的具体执行人, 需要协调项目合约经理、项目工程经理级设计、报建等部门共同推进招标分判工作。

第五条 业务协同

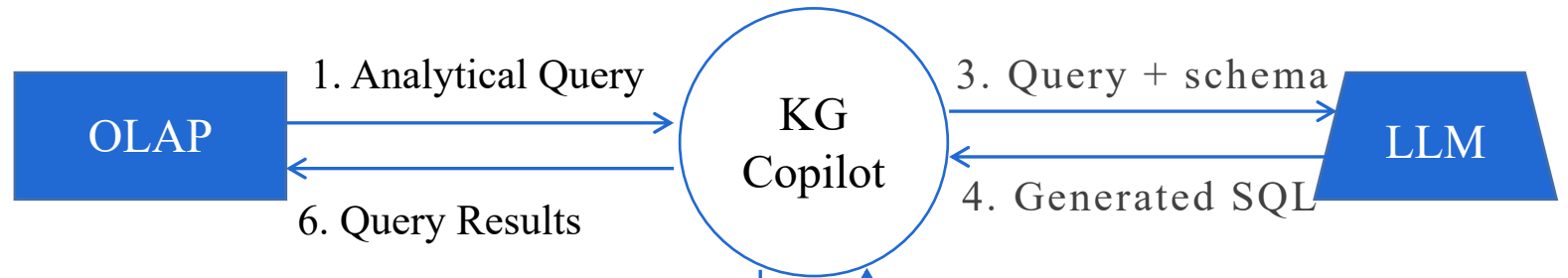
全景计划及成本合约专项计划是招标管理工作开展的计划执行依据。对于计划节点中, 受前置任务影响较大、跨专业沟通困难、易耽误项目发展的关键环节, 应及时反馈至合约管理部门经理及项目合约经理。与报建相关的招标工作应积极协调相关部门。



招标管理岗位的定位是负责招标管理工作的具体执行人, 属于合约职能线的重要基础工作。他们需要协调项目合约经理、项目工程经理级设计、报建等部门共同推进招标分判工作, 以确保全景计划实现的关键业务。同时, 招标管理岗位还需要与其他部门积极协同, 如及时反馈关键环节信息至合约管理部门经理及项目合约经理, 以及协调与报建相关的招标工作。



KG Copilot Engine Example: ChatTable




- a. Lookup Tables
- b. Schema2Prompt

2. Pre-processing

5. Post-processing

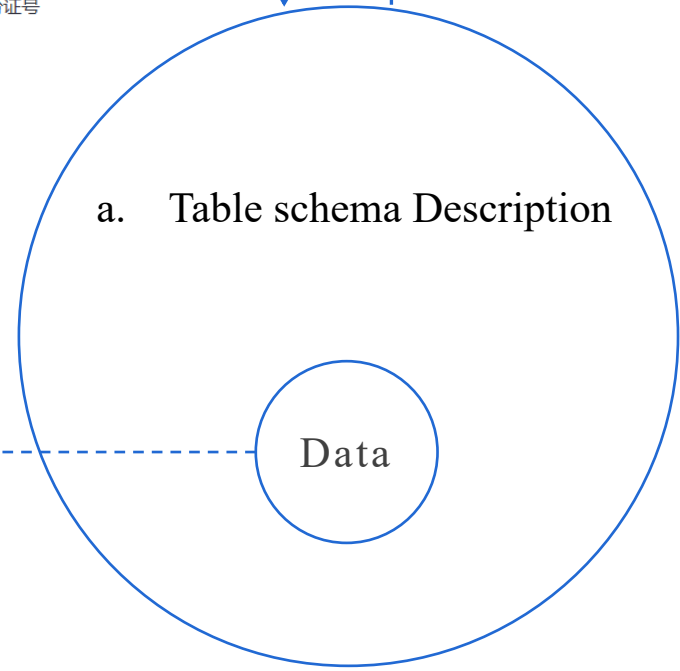
a. SQL Querying

GU 数据库表公民信息表 citizen, 表格的列为“唯一标识id”, “姓名name”和“身份证号ID_number”
写一个SQL查询: 查找出所有“身份证号相同但姓名不同”的记录

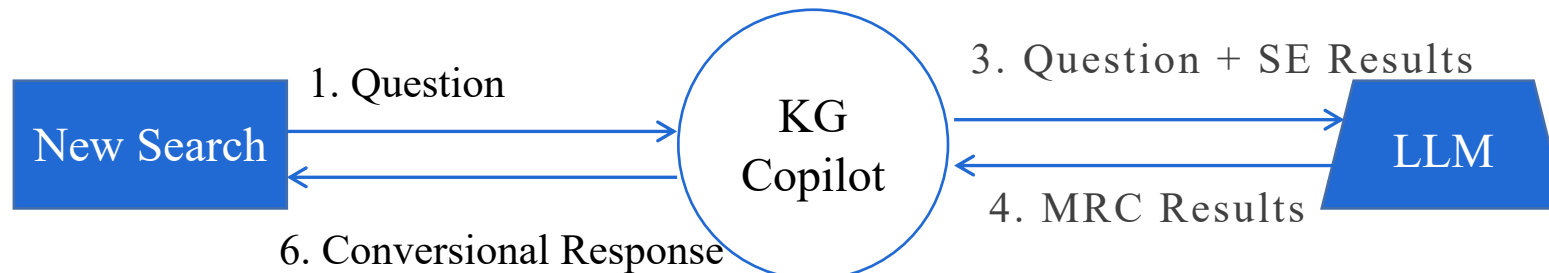
 要查询所有身份证号相同但姓名不同的公民信息记录, 可以使用子查询来选择具有多个不同名称的唯一身份证号码, 然后在唯一身份证号码上将子查询与原始表进行联接:

```
sql Copy code  
  
SELECT c.id, c.name, c.ID_number  
FROM citizen c  
INNER JOIN (  
    SELECT ID_number  
    FROM citizen  
    GROUP BY ID_number  
    HAVING COUNT(DISTINCT name) > 1  
) c2 ON c.ID_number = c2.ID_number  
ORDER BY c.ID_number, c.name
```

Tables in DB/DW/BD



KG Copilot Engine Example: ChatWeb

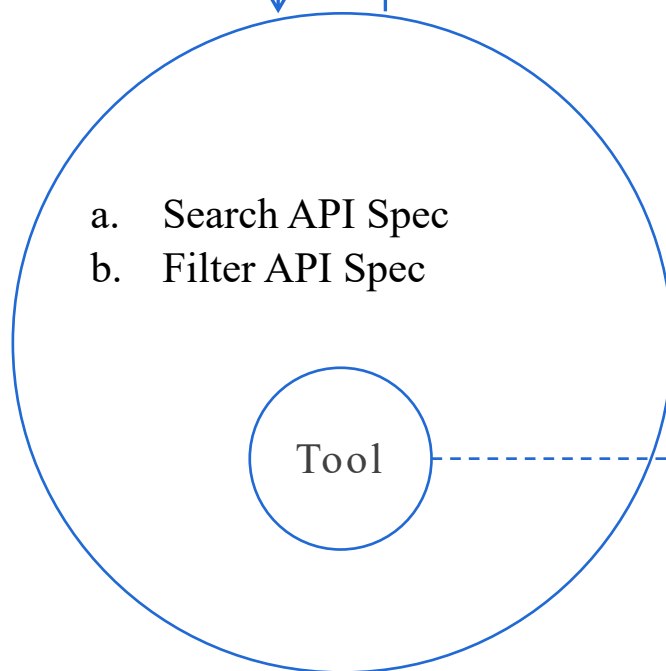


- a. Search & Filter
- b. SearchResult2Prompt

2. Pre-processing

5. Post-processing

- a. Referencing : Web Page Provenance
- b. Result Merging



事件1: 对台军售案

- 时间: 2021年8月4日
- 卖方: 美国政府
- 买方: 台湾
- 金额: 7.5亿美元
- 售卖内容: 40套M109A6自行榴弹炮系统和相关设备

事件2: 对台军售案

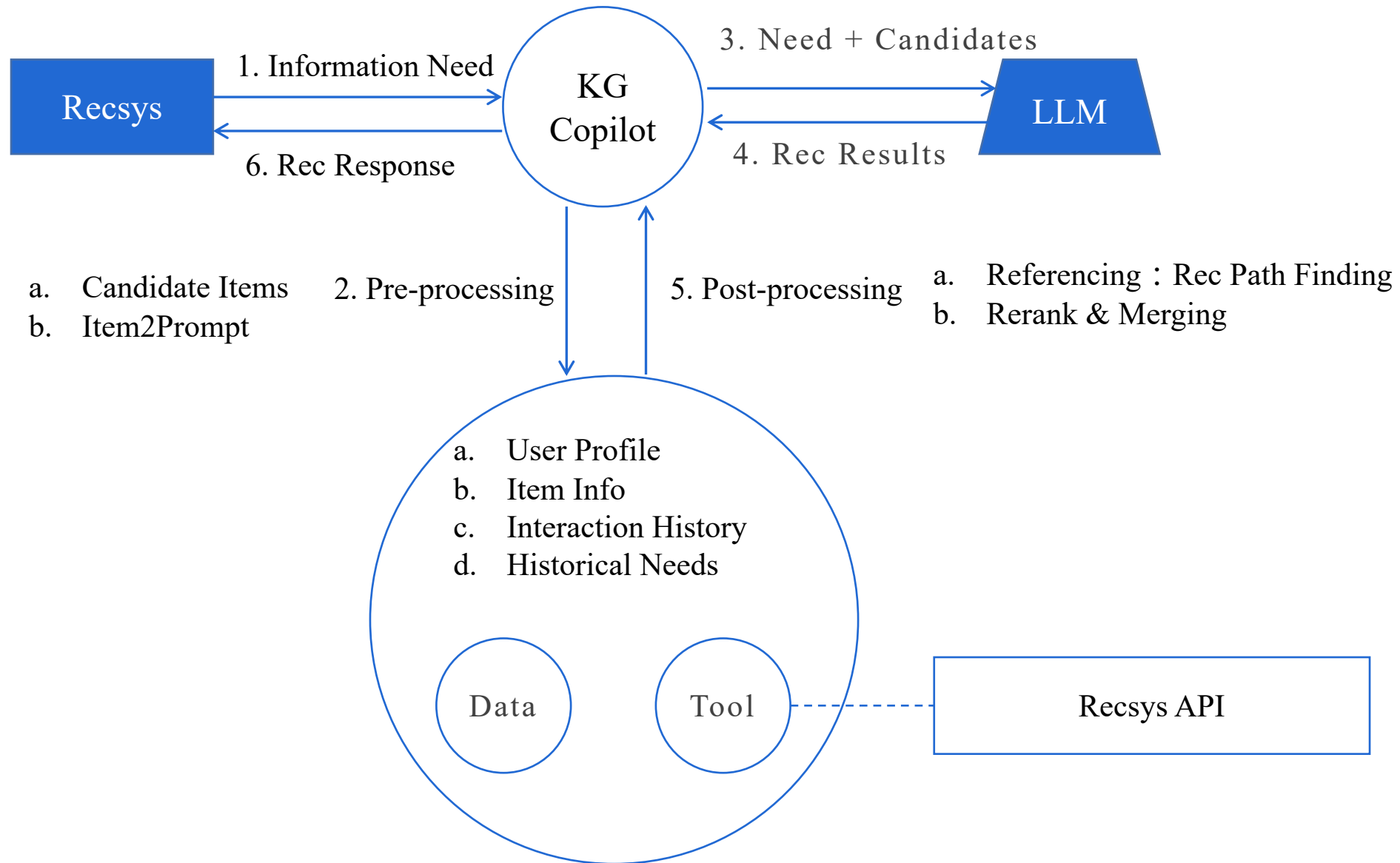
- 时间: 2022年前7个月
- 卖方: 美国政府
- 买方: 台湾
- 金额: 1.95亿美元
- 售卖内容: 对台湾购买的“爱国者”导弹进行评估保养, 并派遣人员提供直接技术支持



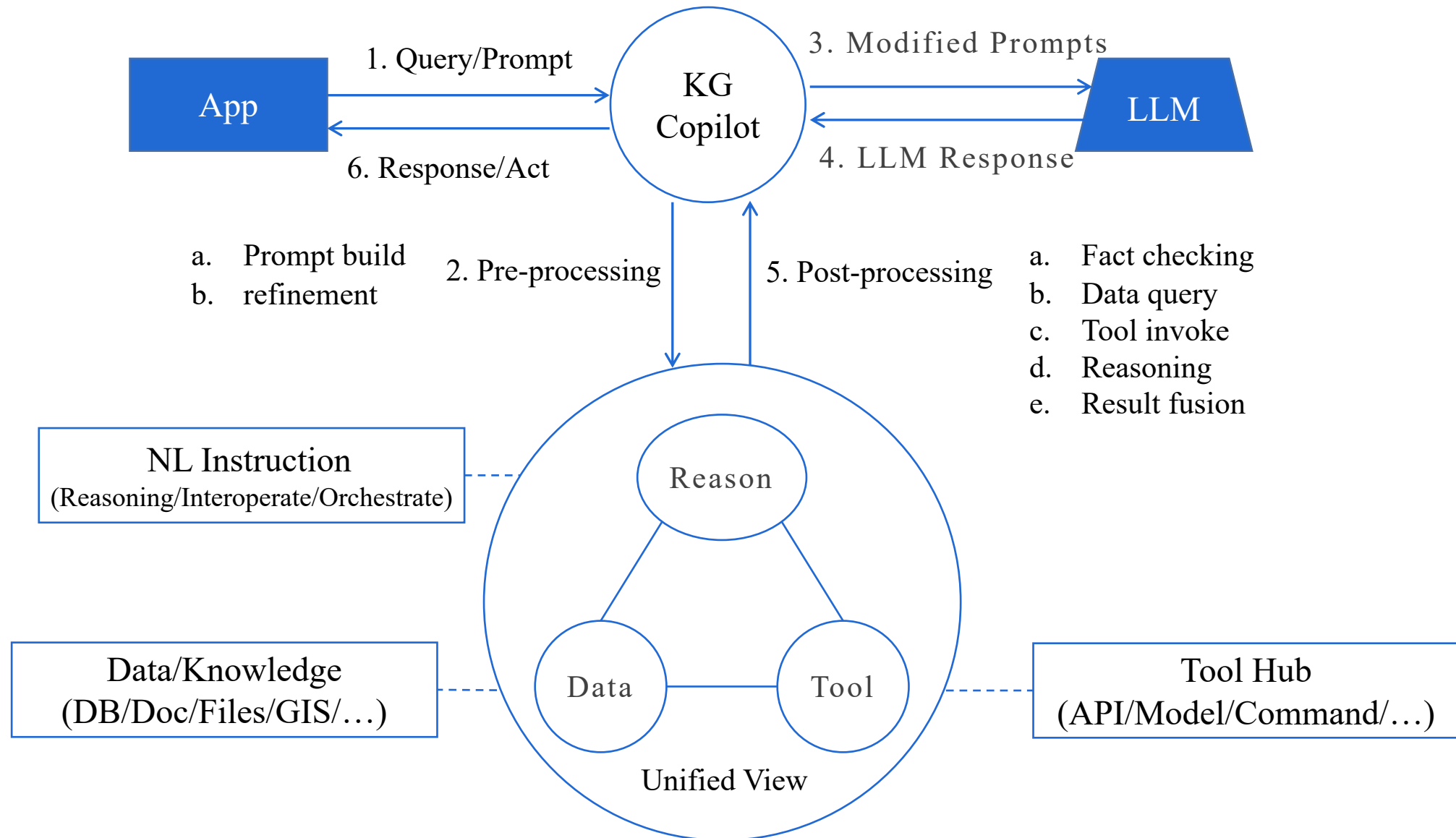
任务类型: 事件抽取
 事件类型: 军售
 事件论元角色: [时间, 卖方, 买方, 金额, 售卖内容]
 请从如下文本中进行事件抽取, 多个事件依次列出来, 没有时间的去除:
 中国社会科学院西亚非洲研究所研究员贺文萍指出, 美国媒体披露的这项新的对台军售案, 价值总额、规模均远超以往, 且多以美军现役攻击性武器为主, 具有较强的实用性。
 尽管3名美方知情人士都称军售方案仍处于策划阶段, 没有透露拜登计划何时向国会提出这项军售案, 但英国路透社8月25日报道称, 一名美国官员透露, 拜登最快可能在今年9月宣布启动此项新军售项目。按照美国相关法律, 拜登政府正式向国会发出对台军售通知后, 参议院外交关系委员会与众议院外交事务委员会的两党议员, 将讨论并决定是否批准此项军售案。据美国政治新闻网Politico的报道, 拜登极力推动此案, 国会方面也没有特别的阻力, 因而, 美国国会秋季复会后, 两党议员大概率会批准这项军售案。
 如果此项军售案获国会议员签署批准, 将是拜登2021年1月上台以来的第六次对台军售。2021年8月4日, 拜登政府就职半年多后宣布首次对台军售案, 对台出售总价7.5亿美元的40套M109A6自行榴弹炮系统和相关设备。2022年前7个月, 拜登政府已4次批准对台军售: 2月7日和4月5日, 美国国防安全合作局先后批准对台湾购买的“爱国者”导弹进行评估保养, 并派遣人员提供直接技术支持, 总预算为1.95亿美元; 6月9日, 美方再次批准对台出售价值1.2亿美元的军用船舶、舰载系统维修零部件和后勤技术协助; 7月15日, 美国国务院国防安全合作局 (DSCA) 批准售台“零附件采购与技术协助案”, 金额约为1.08亿美元。

Search Engine API

KG Copilot Engine Example: ChatRec



KG Copilot Engine Example: Chat & Analyze All Data





TONGJI UNIVERSITY
COLLEGE OF DESIGN AND
INNOVATION

同济大学设计创意学院

Thank you



 OpenKG.CN
开放的中文知识图谱

<http://www.openkg.cn/>