

그래프 기반 지식 강화형 인공지능 (Part 2)

w/ Graph & Language Intelligence

Byungkook Oh (오병국)

Department of Computer Science and Engineering

bkoh509@gmail.com / bkoh509@gmail.com

<https://bkoh509.github.io>

<https://gli.konkuk.ac.kr>

CONTENTS

1. Today
2. Background
3. Large Language Models
4. Knowledge Graphs
5. Graph-enhanced Training
6. Graph-enhanced Reasoning
7. LLM-based Recommendation

What's Next?

❑ Knowledge Graphs? or Large Language Models?

Structured Knowledge Models (Knowledge Graphs)

- ✓ Explicit knowledge
- ✓ Latest/Expert knowledge
- ✓ Domain-specific knowledge
- ✓ Accuracy
- ✓ Determinateness
- ✓ Interpretability

- ⊖ Incompleteness
- ⊖ Lacking language
- ⊖ Understanding
- ⊖ Unseen facts

↔
complementary
relationship^[10]

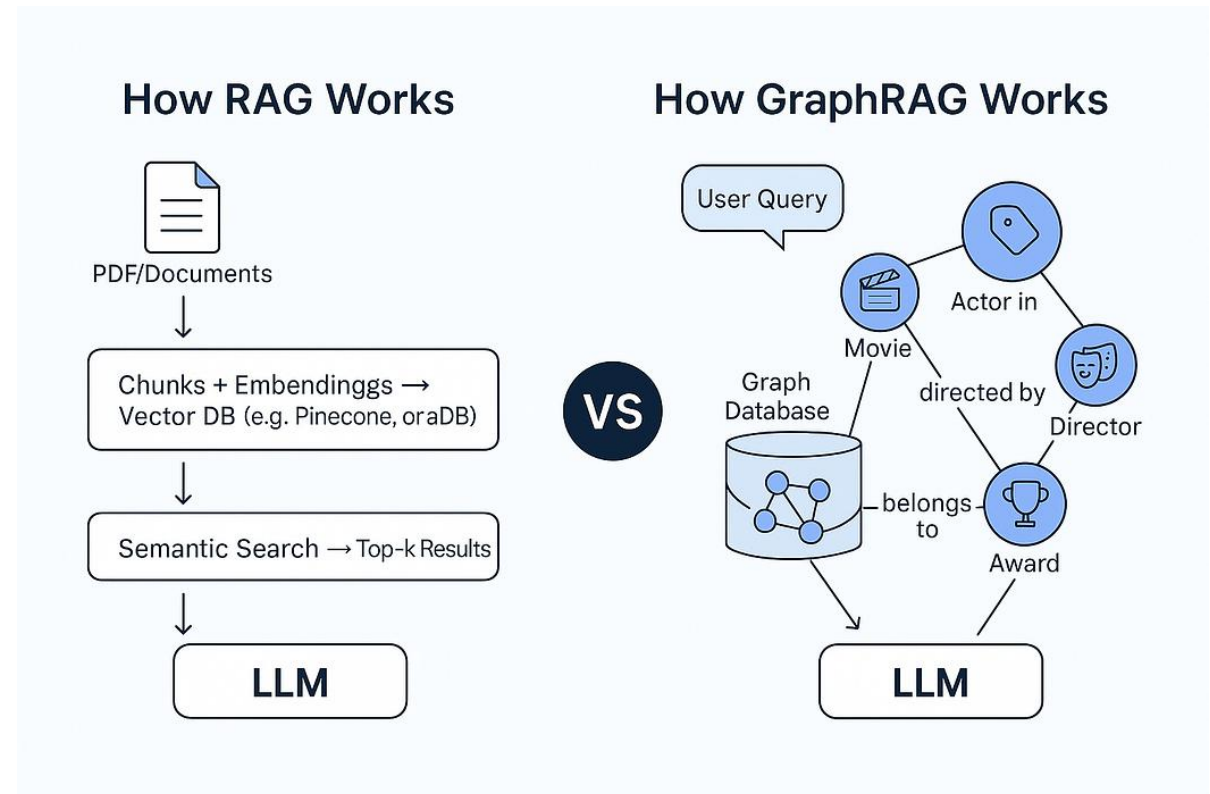
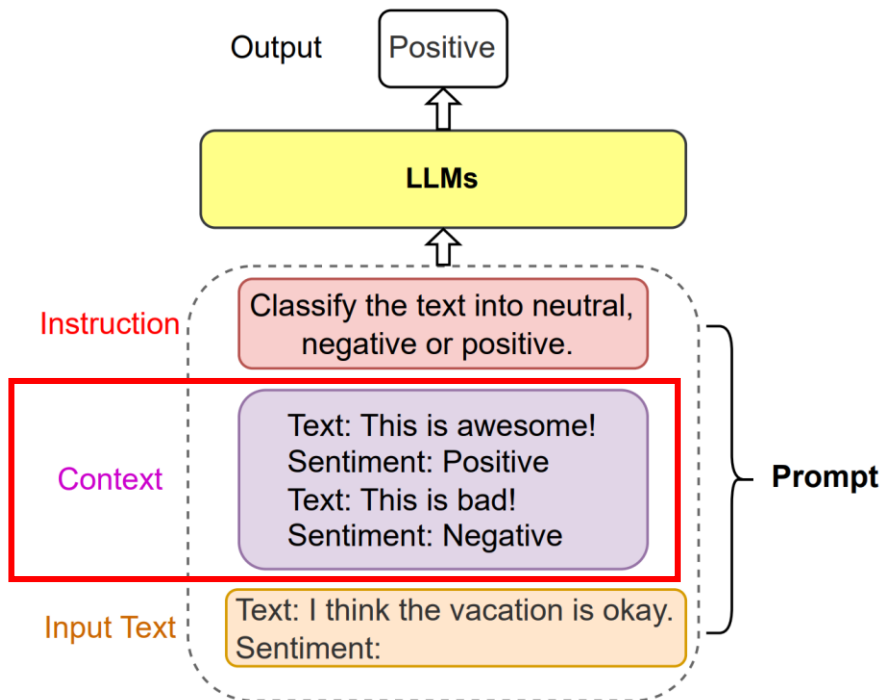
Large Language Models (Text)

- ✓ Generalizability
- ✓ General Knowledge
- ✓ Language Processing

- ⊖ Implicit knowledge
- ⊖ Hallucination
- ⊖ Indeterminateness
- ⊖ Black-box
- ⊖ Lacking domain-specific Knowledge
- ⊖ Lacking new knowledge

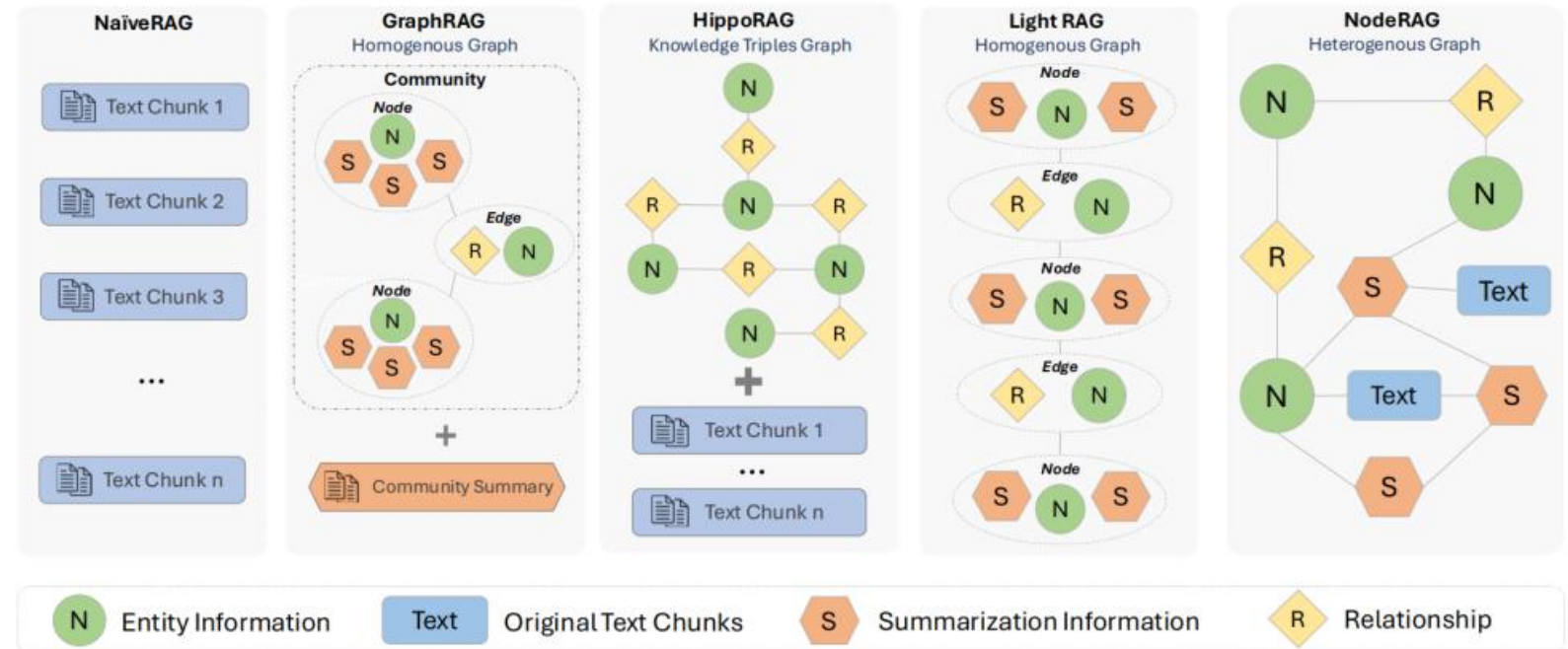
What's Next? – Augmentation

- ❑ 다양한 소스의 최신 지식을 실시간으로 검색·활용하여 생성모델의 환각을 줄이고 사실 검증 가능성을 높이는 것을 목표
 - ✓ 기존 방식: Query와 가장 유사한 chunk만을 개별적으로 검색하기 때문에 여러 문서에 흩어진 관련 정보들 사이의 연결성을 포착하지 못함



What's Next? – Augmentation

- ❑ 다양한 소스의 최신 지식을 실시간으로 검색·활용하여 생성모델의 환각을 줄이고 사실 검증 가능성을 높이는 것을 목표
 - ✓ 전역적 이해 (Global Understanding)
 - ✓ 다중 홉(Multi-hop) 추론
 - ✓ 맥락 연결
 - ✓ 중복 제거 및 일관성



What's Next? – Integration

□ 다양한 유형의 지식을 **멀티모달 생성모델**(또는 **추천/이상탐지 모델**)이 보다 명확하게 이해할 수 있도록 하는 것을 목표

✓ 기존 방식: 복잡한 문장 의도를 정확하게 해석하기 어렵게 함

사용자: 안녕! 내가 키우는 강아지 3마리에 대해 설명해 줄게. 내 질문에 답해줘.

시스템: 물론이지! 강아지들에 대한 설명을 해주면 그에 맞추어 답변해 줄게.

사용자: **첫번째**와 **세번째**는  에서 각각 탁자 위와 의자 위에
앉아있는 흰색 강아지야. 둘 다 곱슬거리는 흰색의 털을 가지고 있어.
세 마리 각각 , , 그리고  소리를 내.
두번째는  로, **세번째**와 똑같이 주둥이가 길고 다리가 짧아.
특히, **첫번째**와 **세번째** 동물은 비슷하게 생겼어.

그림 1. 기존의 선형적 멀티모달 상호작용

사용자: 안녕! 내가 3마리의 동물을 설명해 줄게. 그다음에 내 질문에 답해줘.

시스템: 물론이지! 동물에 대한 설명을 해주면 그에 맞추어 답변해 줄게.

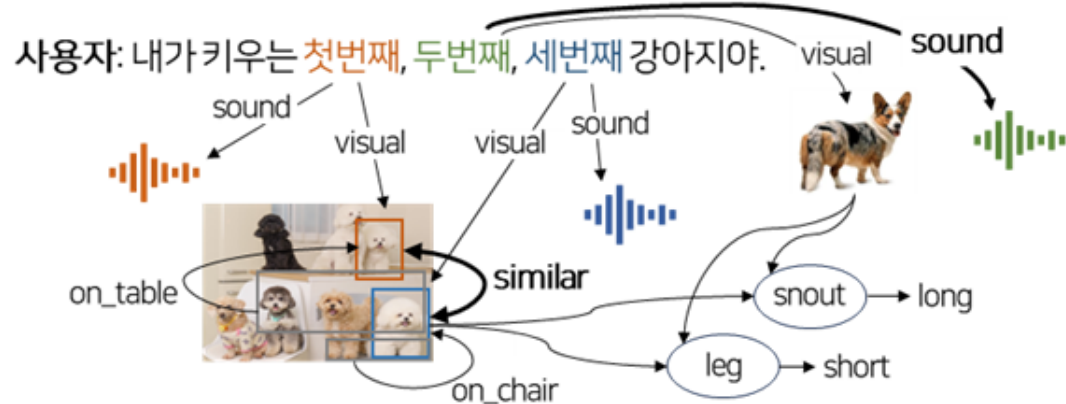


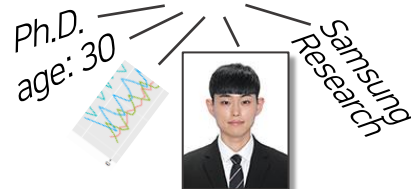




그림 2. 제안하는 그래프 구조 기반의 멀티모달 상호작용

What's Next? – Integration

❑ KG-LLM Synergized Model (GenAI)

ex) KG-augmented Generation
ex) KG Construction
ex) KG Question Answering

	Standard LLM Generation/QA	KG-LLM Synergized Model	
		KG-Augmented Generation	KG Question Answering
Input	Text, Image <i>ex) "Which fruit does Byungkook like,  or  ?"</i>	Text, Multi-Modal KG <i>ex) "Which fruit does <u>Byungkook</u> like, <u>apples</u> or <u>bananas</u>?"</i> <div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;">  </div> <div style="text-align: center;">  </div> <div style="text-align: center;">  </div> </div>	
		✓ Refer to <u>what each entity in text really wants</u> from KG (data + schema)	
Output	Text (probabilistic)	Text (probabilistic)	KG Entity (probabilistic + deterministic)
		✓ Lead to <u>more explicit and domain-specific answer</u>	

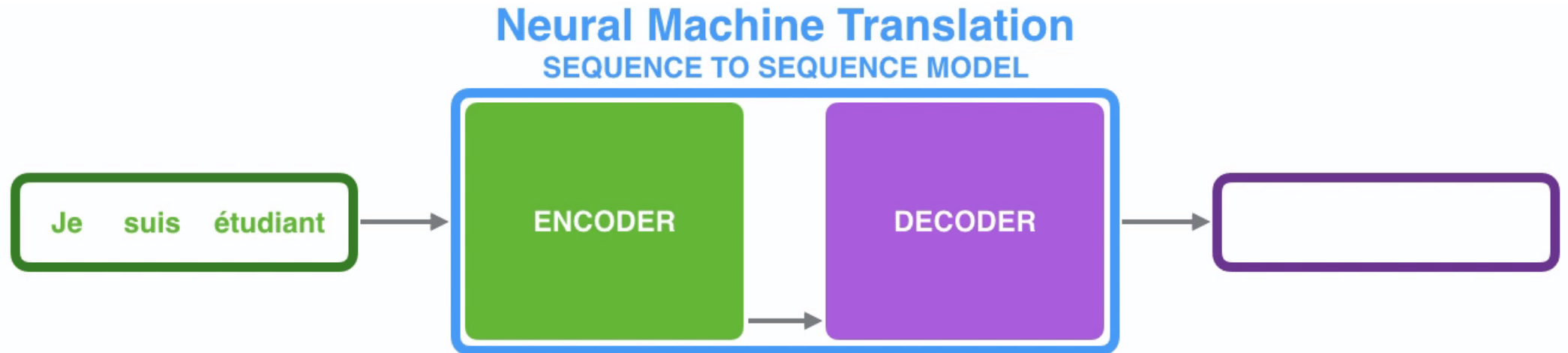
CONTENTS

1. Today
2. Background
3. Large Language Models
4. Knowledge Graphs
5. Graph-enhanced Training
6. Graph-enhanced Reasoning
7. LLM-based Recommendation

Sequence-to-Sequence

❑ Neural Machine Translation

- ✓ Single end-to-end neural network -> subcomponents 별로 따로 최적화할 필요가 없음
- ✓ 2개의 RNN으로 구성된 Sequence-to-Sequence model을 활용함
 - 2014년 구글에 의해 제안된 encoder-decoder 구조의 모델
 - 가변 길이의 문장 출력을 가능하게 한 모델

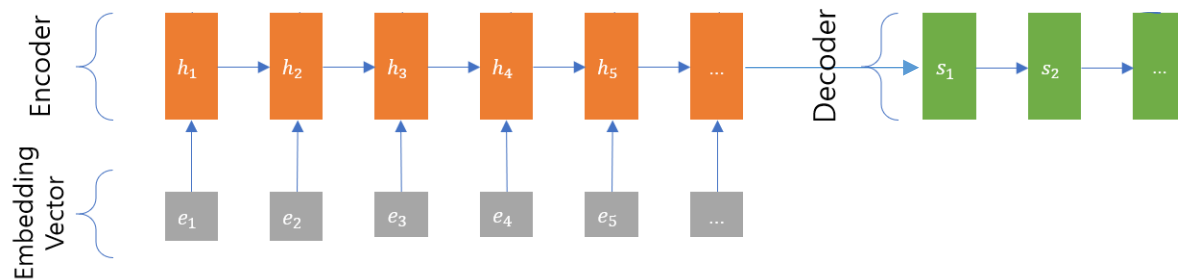


**BUT,,, 입력 문장을 단일 벡터로 표현
(bottleneck problem)**

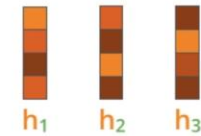
Sequence-to-Sequence with Attention

□ Attention

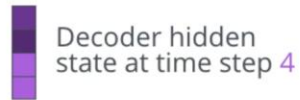
- ✓ 기존 sequence-to-sequence 모델들의 Bottleneck 문제를 해결하기 위해 제안됨
- ✓ decoder의 각 step에서 **입력 문장의 특정 부분에 '집중'**할 수 있도록 connection을 추가함
 - 연관성이 높은 토큰에 가중치를 높여 학습



1. Prepare inputs



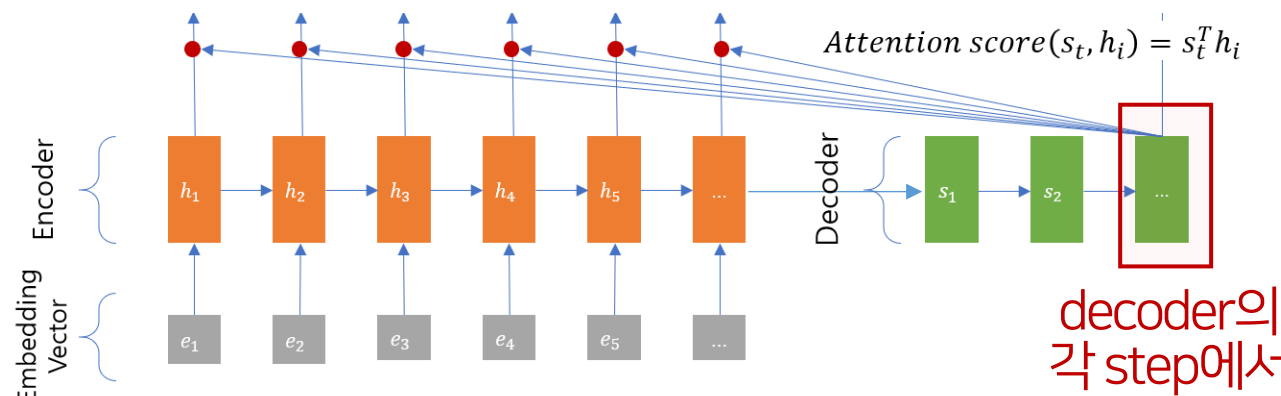
Encoder hidden states



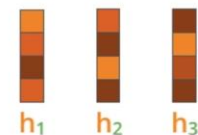
Sequence-to-Sequence with Attention

□ Attention

- ✓ 기존 sequence-to-sequence 모델들의 Bottleneck 문제를 해결하기 위해 제안됨
- ✓ decoder의 각 step에서 **입력 문장의 특정 부분에 '집중'**할 수 있도록 connection을 추가함
 - 연관성이 높은 토큰에 가중치를 높여 학습



1. Prepare inputs



Encoder hidden states



Decoder hidden state at time step 4

2. Score each hidden state

13	9	9
----	---	---

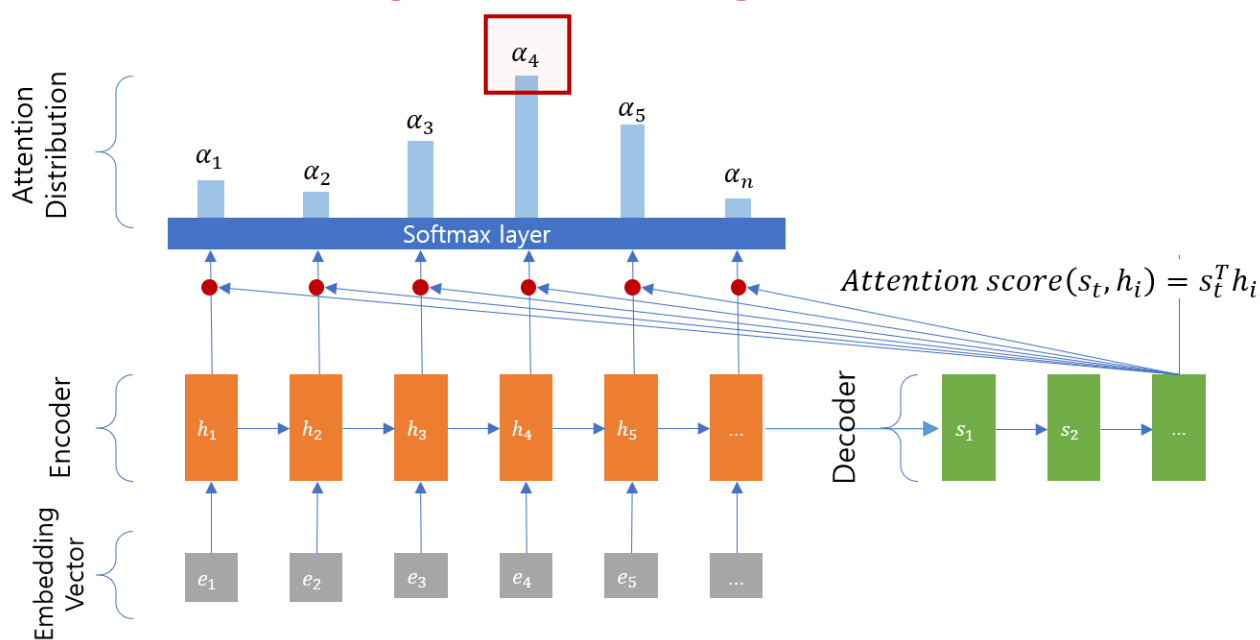
scores
Attention weights for decoder time step #4

Sequence-to-Sequence with Attention

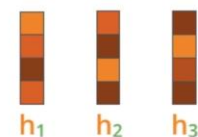
□ Attention

- ✓ 기존 sequence-to-sequence 모델들의 Bottleneck 문제를 해결하기 위해 제안됨
- ✓ decoder의 각 step에서 **입력 문장의 특정 부분에 '집중'**할 수 있도록 connection을 추가함
 - 연관성이 높은 토큰에 가중치를 높여 학습

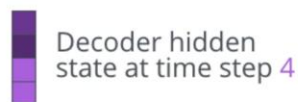
입력 문장의
특정 부분에 대한 집중



1. Prepare inputs



Encoder hidden states



2. Score each hidden state

13	9	9
----	---	---

scores
Attention weights for decoder time step #4

3. Softmax the scores

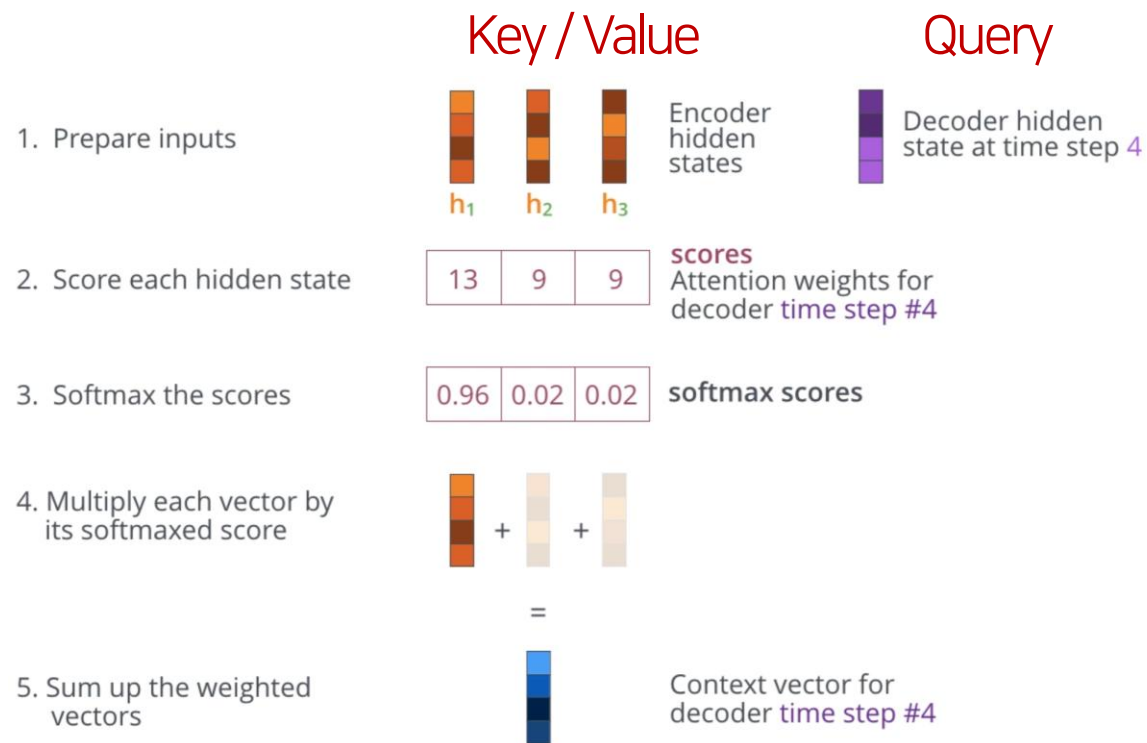
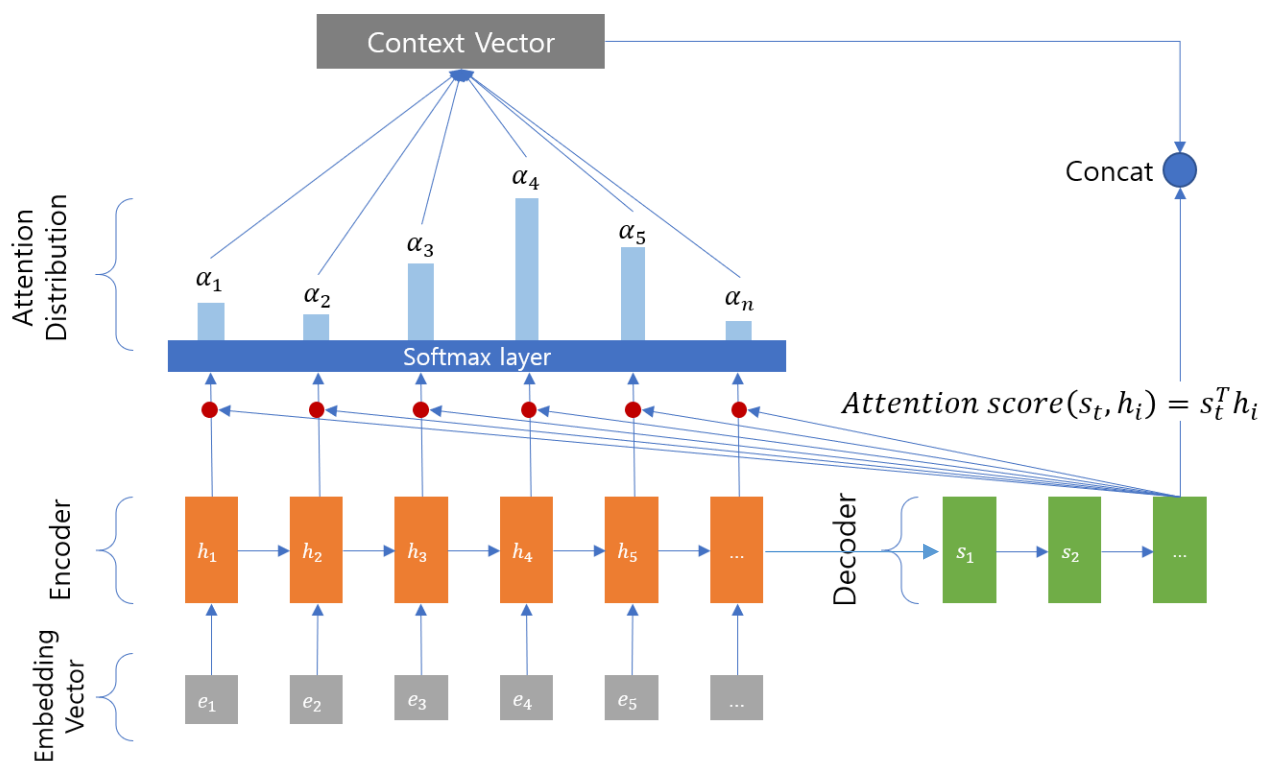
0.96	0.02	0.02
------	------	------

softmax scores

Sequence-to-Sequence with Attention

□ Attention

- ✓ 기존 sequence-to-sequence 모델들의 Bottleneck 문제를 해결하기 위해 제안됨
- ✓ decoder의 각 step에서 **입력 문장의 특정 부분에 '집중'**할 수 있도록 connection을 추가함
 - 연관성이 높은 토큰에 가중치를 높여 학습



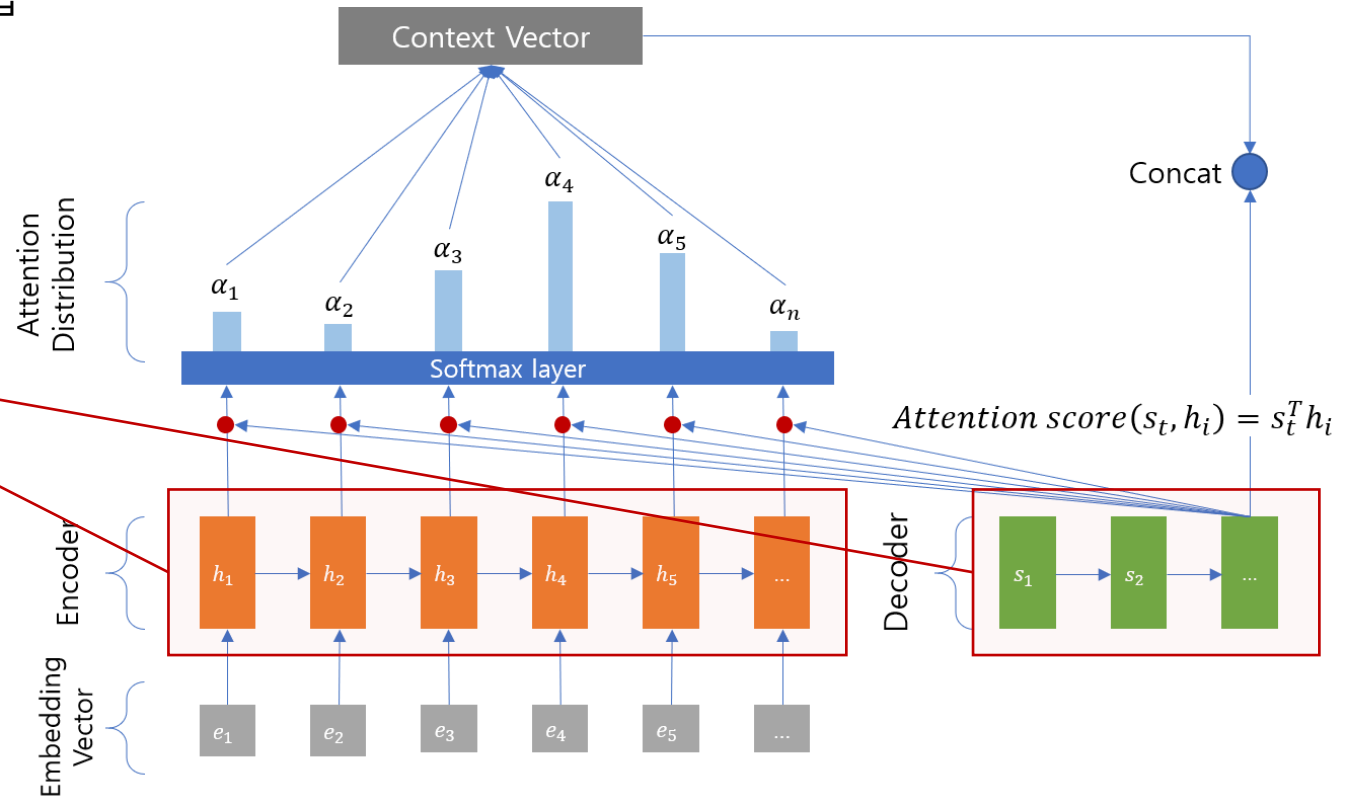
Sequence-to-Sequence with Attention

□ Attention

- ✓ 기존 sequence-to-sequence 모델들의 Bottleneck 문제를 해결하기 위해 제안됨
- ✓ decoder의 각 step에서 **입력 문장의 특정 부분에 '집중'**할 수 있도록 connection을 추가함
 - 연관성이 높은 토큰에 가중치를 높여 학습

BUT,,,
여전히 순차적으로 입력/출력 문장 처리

- gradient 불안정
- 이전 문장 정보들이 압축
- 병렬화 불가
- ...

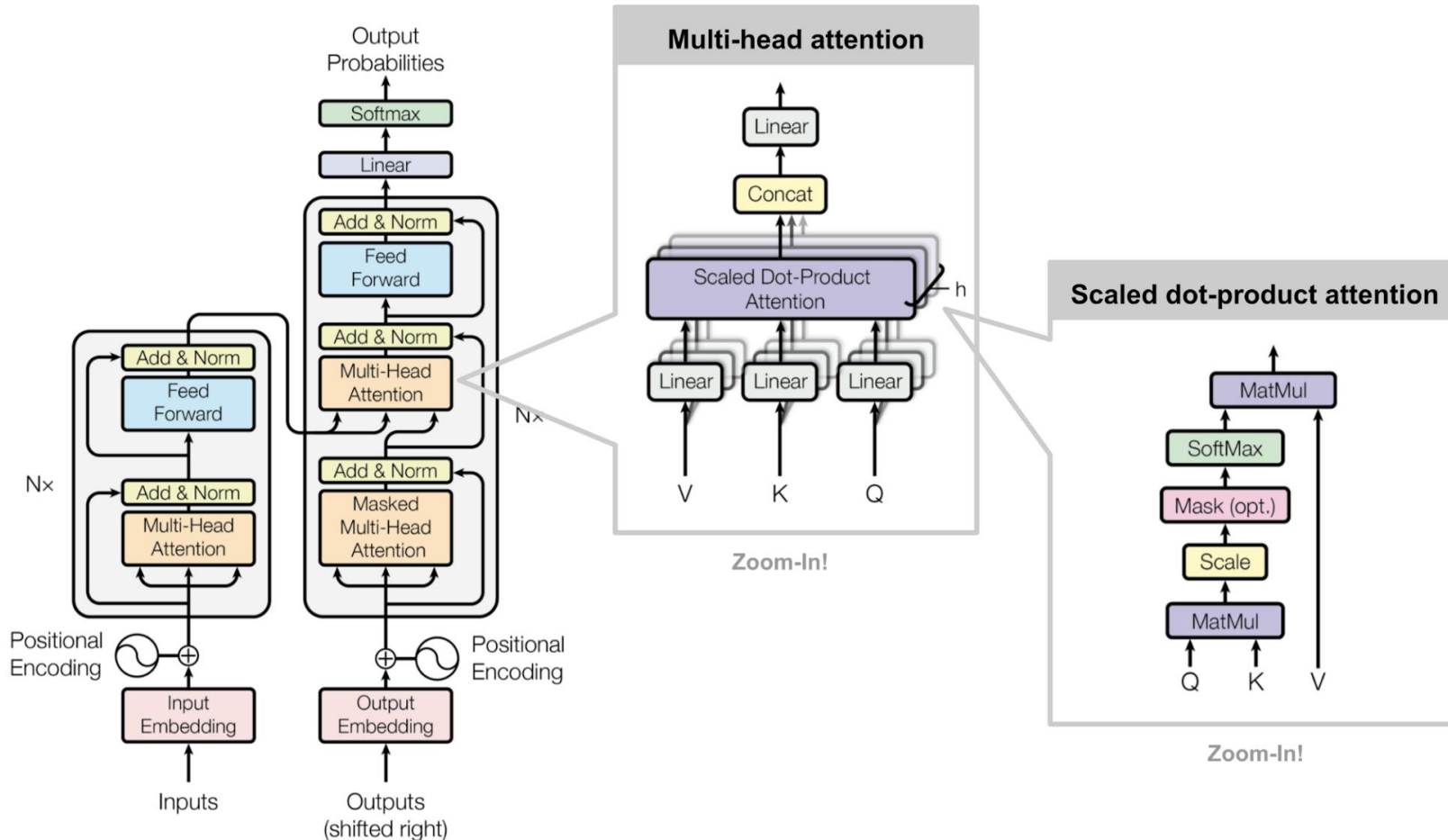


Sequence-to-Sequence with Attention

Transformer

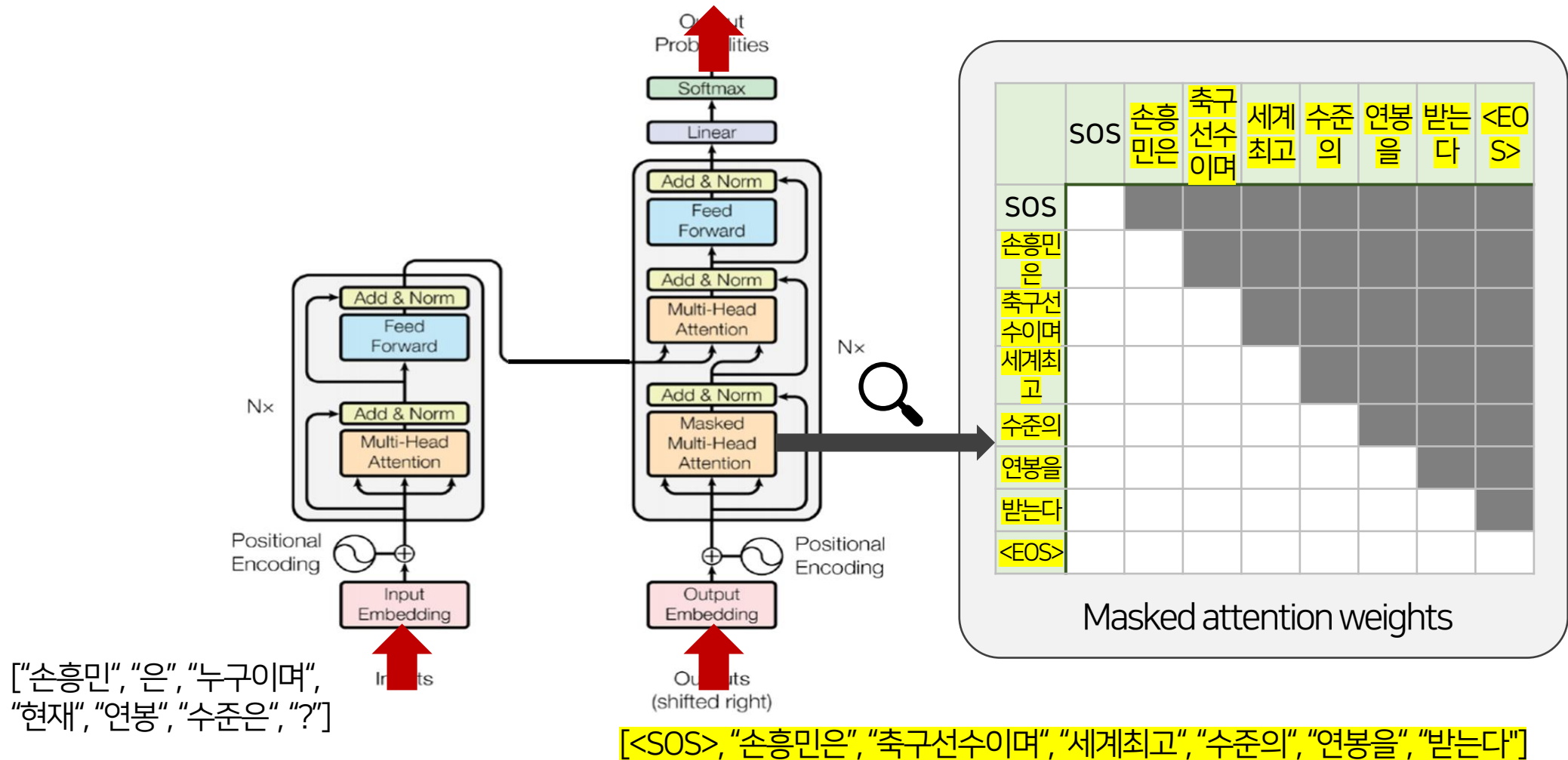
- ✓ Attention의 병렬적 사용을 통해 효율적인 학습이 가능한 구조의 언어 모델

입력 문장을 순차적으로 집어넣어서
 h_1, h_2, h_3 와 context vector를 생성하지 않고
입력 문장을 통째로 처리



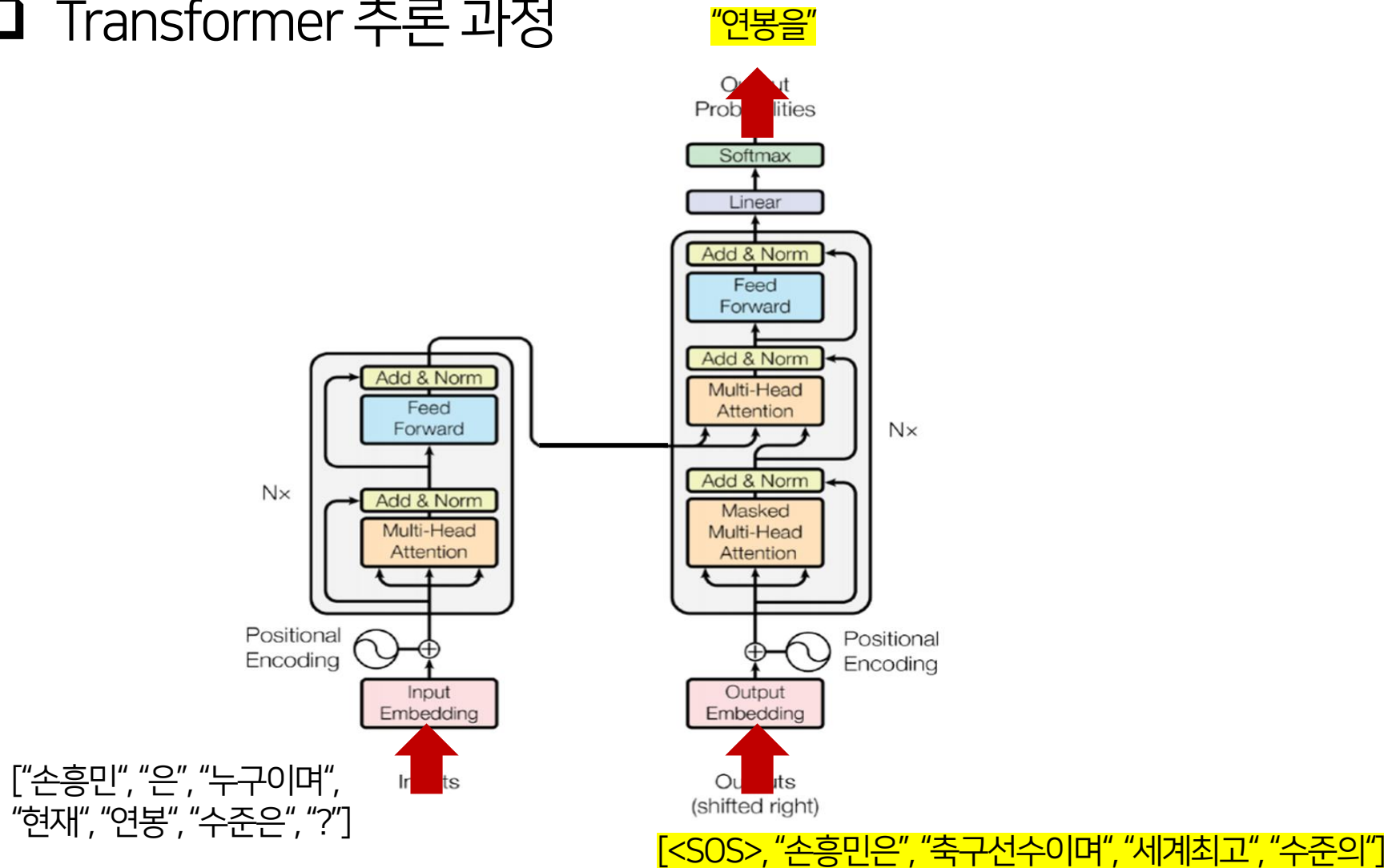
Sequence-to-Sequence with Attention

Transformer 학습 과정



Sequence-to-Sequence with Attention

□ Transformer 추론 과정

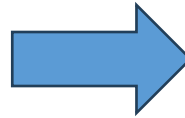
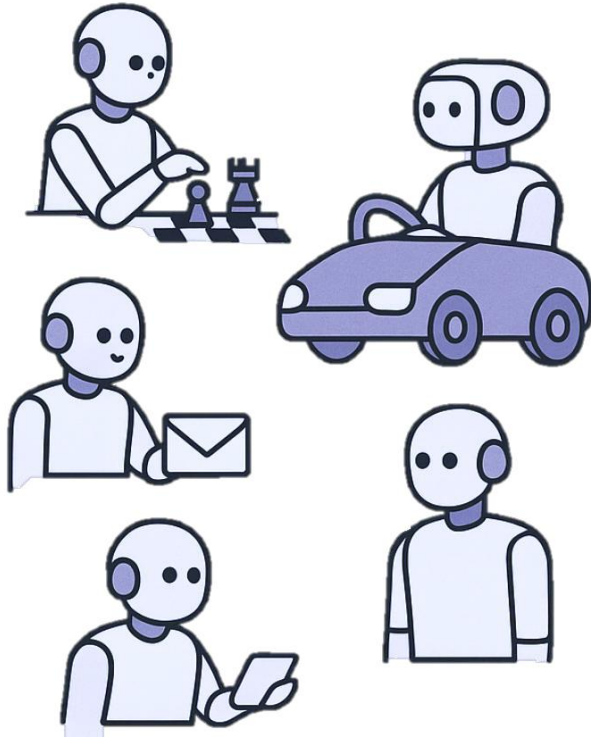


CONTENTS

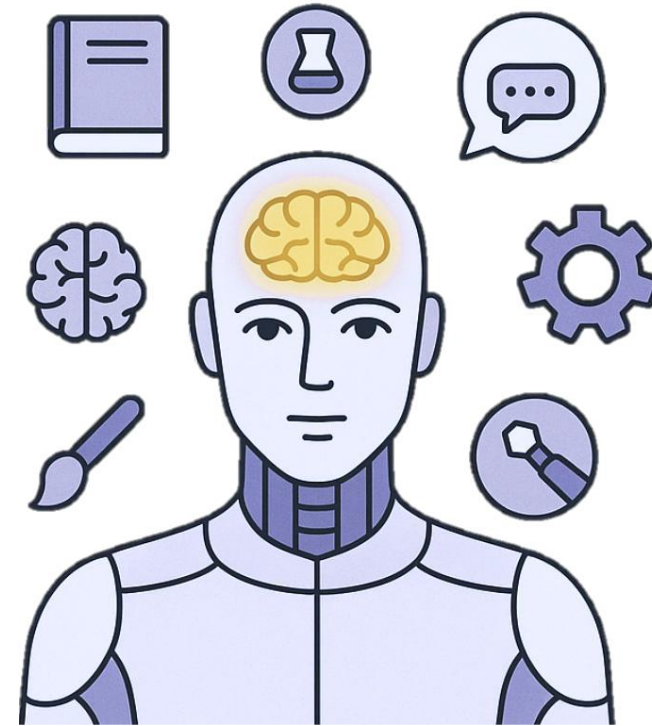
1. Today
2. Background
- 3. Large Language Models**
4. Knowledge Graphs
5. Graph-based Augmentation
6. Graph-LLM Integration
7. LLM-based Recommendation

What is AGI?

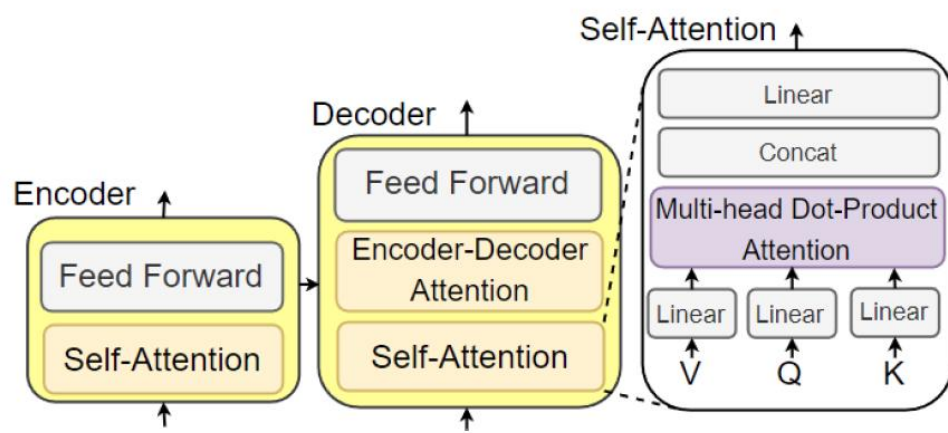
Artificial Intelligence (AI)
Specialized intelligence



Artificial General Intelligence (AGI)
Human-level general intelligence

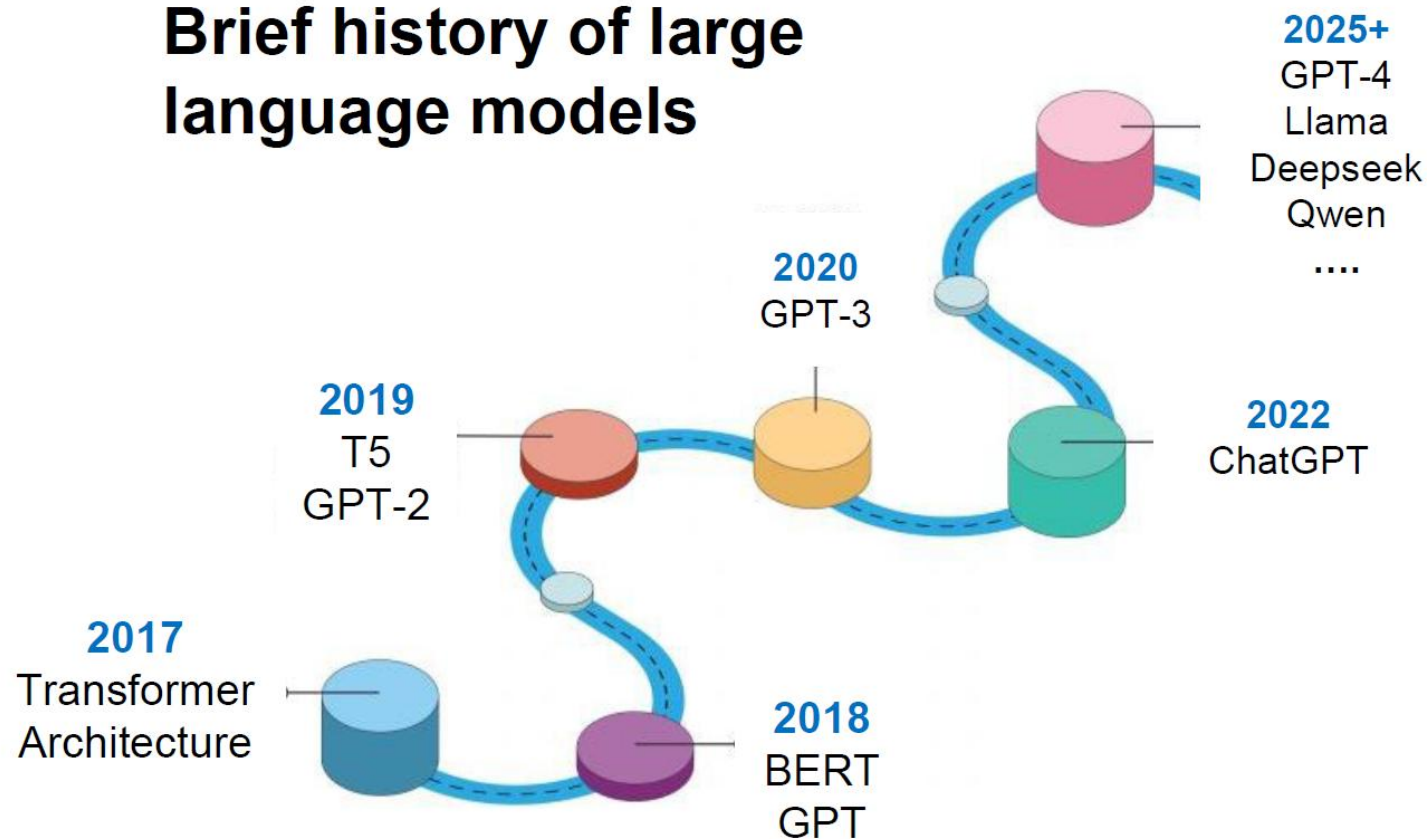


LLMs as AGI



Transformer architecture

Brief history of large language models



LLMs as AGI

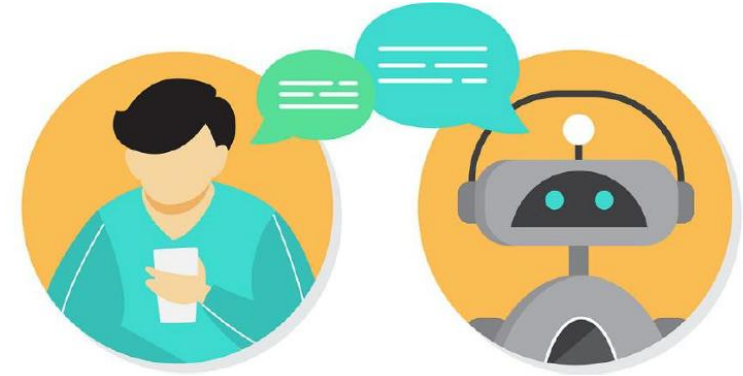
- ❑ LLMs achieve surprising performance across many tasks.



LLMs



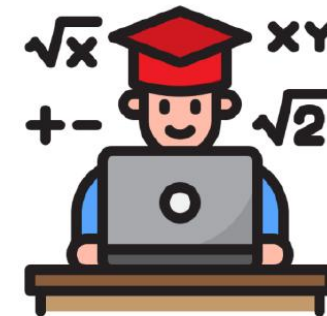
Translation



Conversation

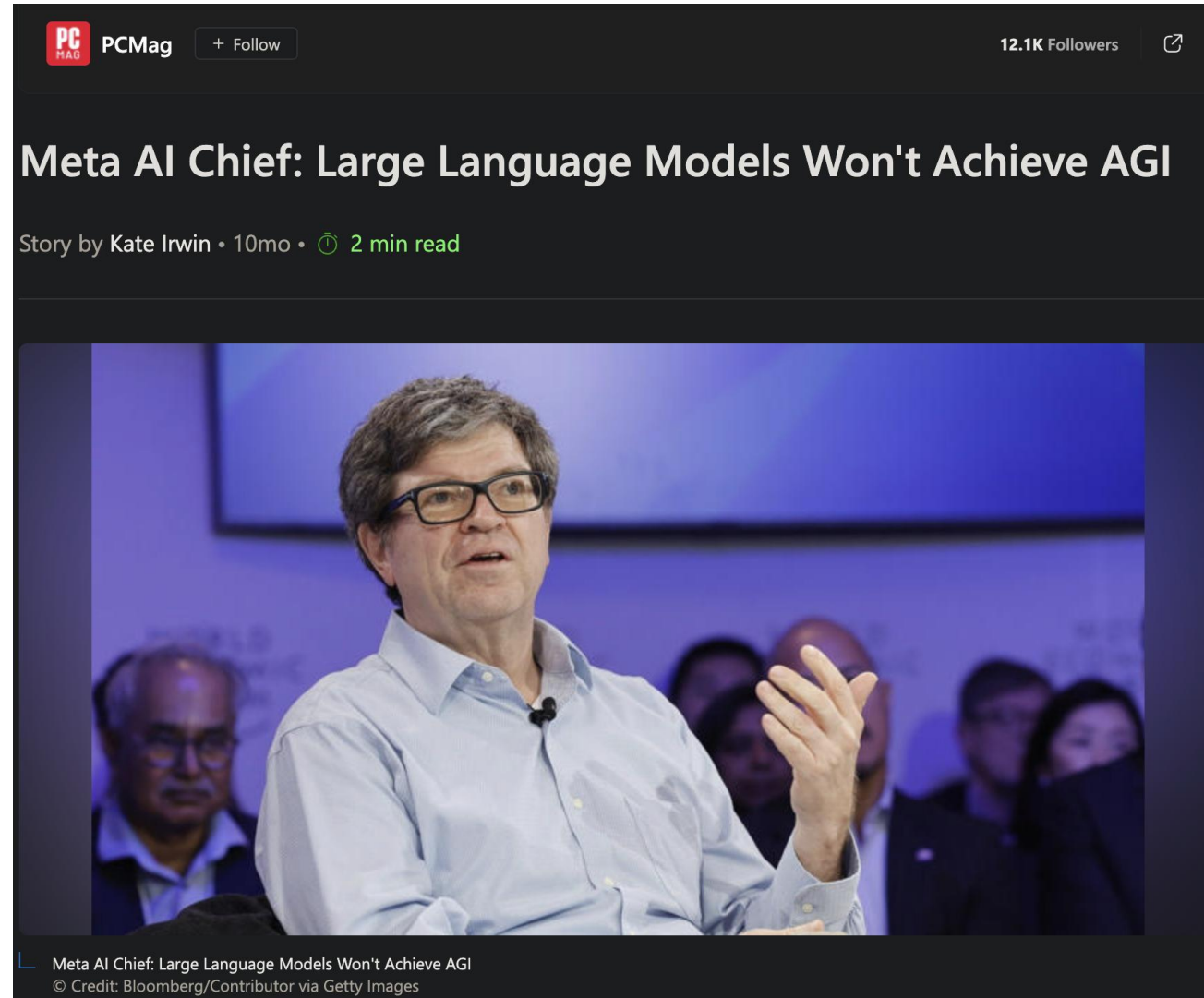


Question Answering



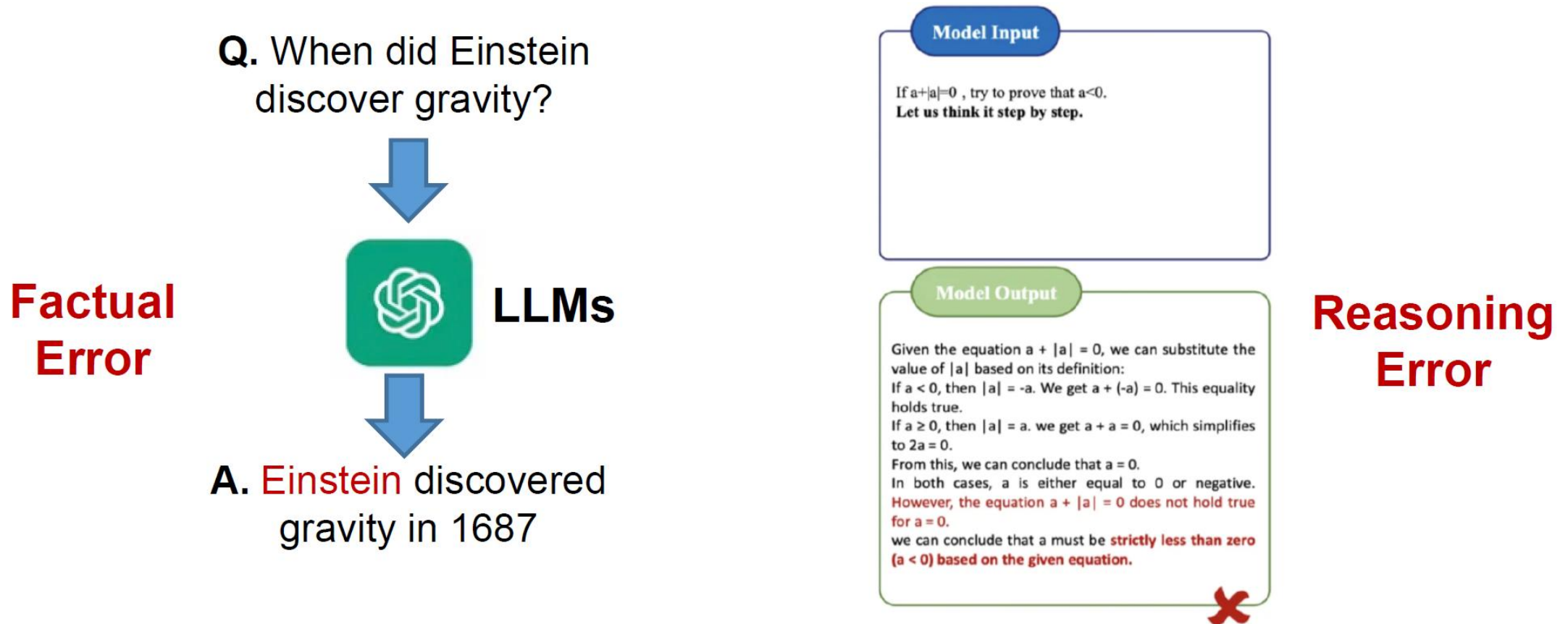
Math Solver

Can LLMs achieve AGI?



Limitations of LLMs

- ❑ LLMs suffer from hallucination problems during reasoning.



Hallucination impairs the trustworthiness of LLMs.

Limitations of LLMs

- ❑ LLMs limit in accessing up-to-date knowledge.

Apr 2


Mr. Trump added a **34 percent tariff** on imports from China, to take effect on April 9, on top of two earlier rounds of 10 percent tariffs he had already imposed.

Apr 8

Trump Threatens to Slap an Additional **50% Tariff on China**

By Alyssa Lukpat, Reporter

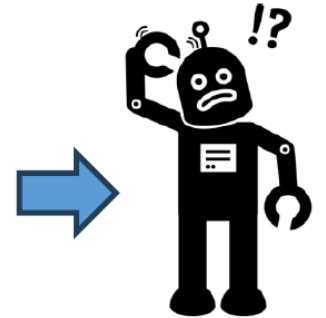


Donald J. Trump 
@realDonaldTrump

Apr 10

Based on the lack of respect that China has shown to the World's Markets, I am hereby raising the Tariff charged to China by the United States of America to **125%**, effective immediately. At some

Q. What is the current tariff on China?



LLMs

Limitations of LLMs

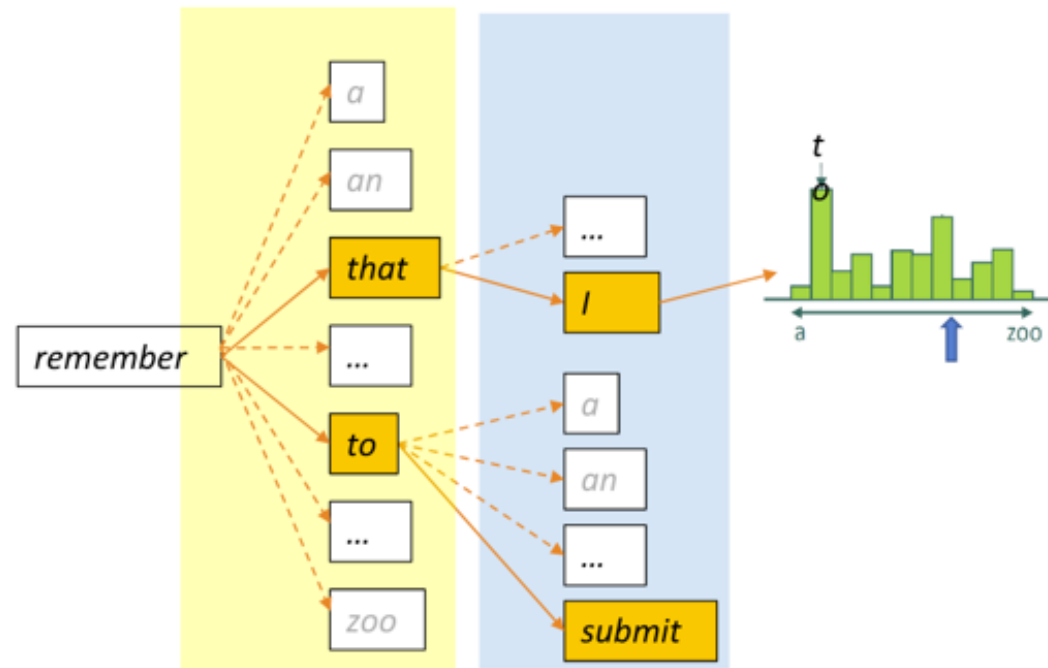
❑ LLMs lack **interpretability**.

- ✓ How to represent knowledge?
- ✓ Why make such a decision?



Limitations of LLMs

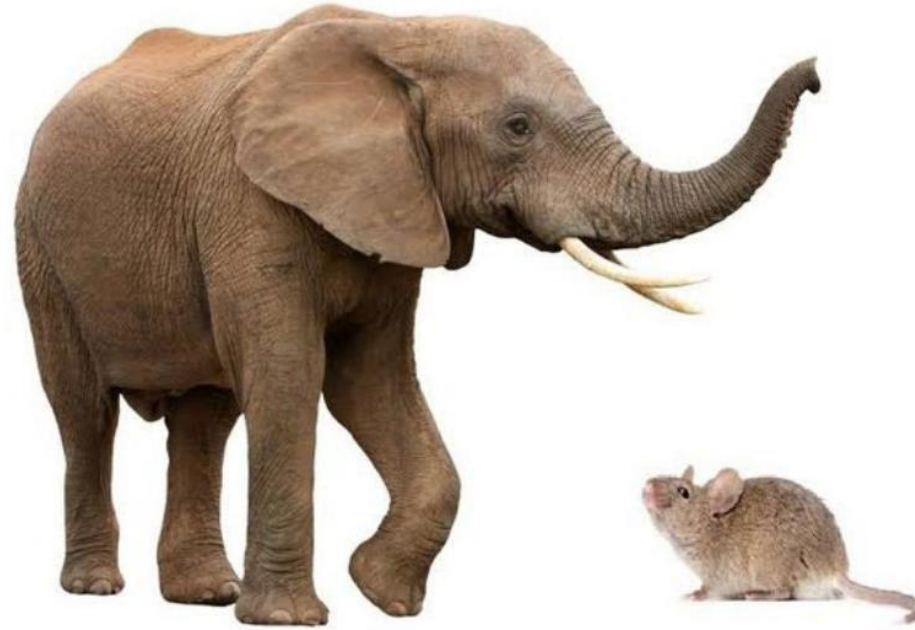
- ❑ LLMs are indecisive.
 - ✓ LLMs reason by probability.



Limitations of LLMs

❑ LLMs are **heavy**

- ✓ More data more parameters.
- ✓ Cannot generalize to a specific domain



Pretrain data

Task-specific data

CONTENTS

1. Today
2. Background
3. Large Language Models
- 4. Knowledge Graphs**
5. Graph-enhanced Training
6. Graph-enhanced Reasoning
7. LLM-based Recommendation

What is Knowledge?

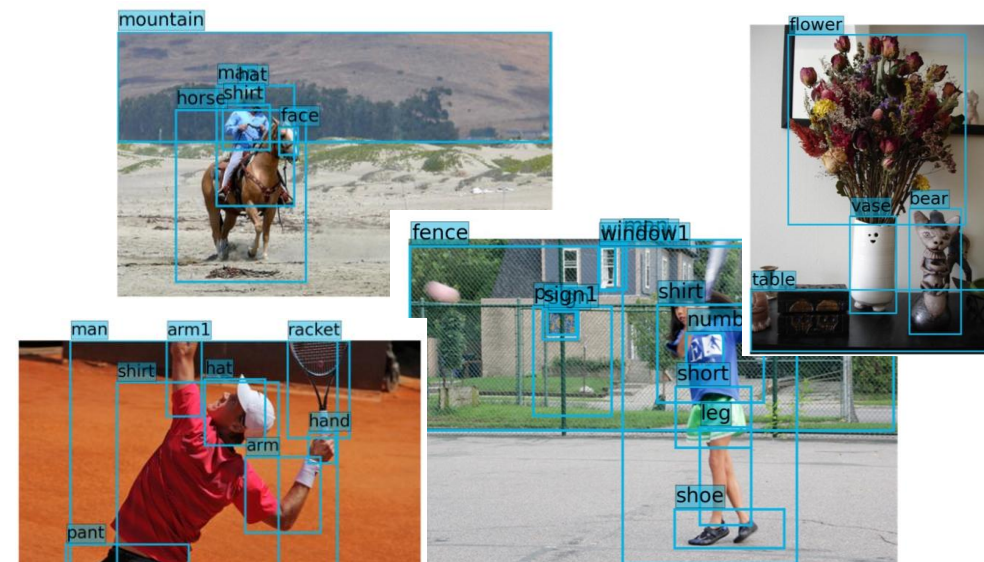
모든 데이터에는 human knowledge가 암묵적으로 담겨있음

ex) 텍스트: 단어 나열

- “아스피린은 두통을 완화한다” -> 개념(약물, 증상), 인과관계, 의학적 사실

ex) 이미지: 픽셀 배열

- "고양이가 소파 위에 있다" -> 객체(고양이, 소파), 관계(위에), 상식(고양이는 동물)



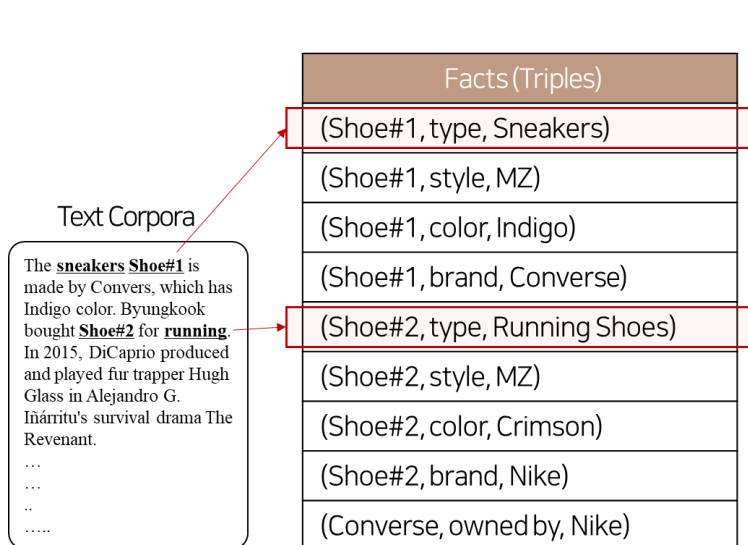
Text Data

Image Data

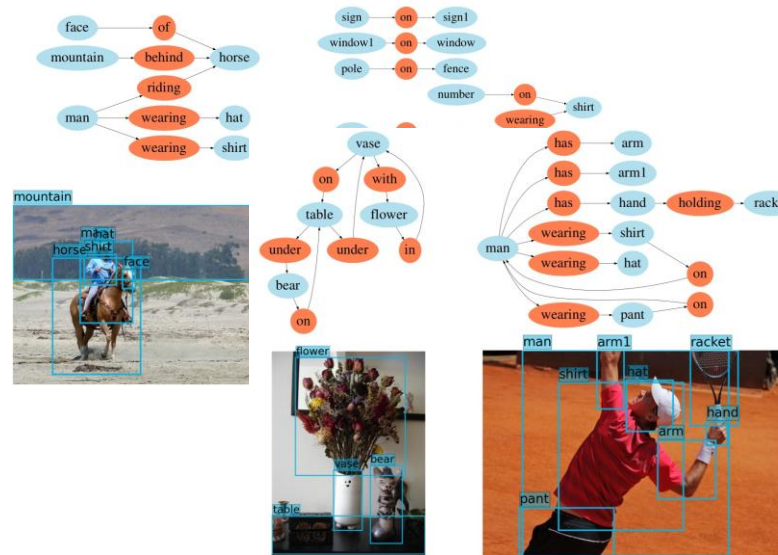
What is Knowledge Graph?

❑ Collection of facts about entities and semantic relations between entities

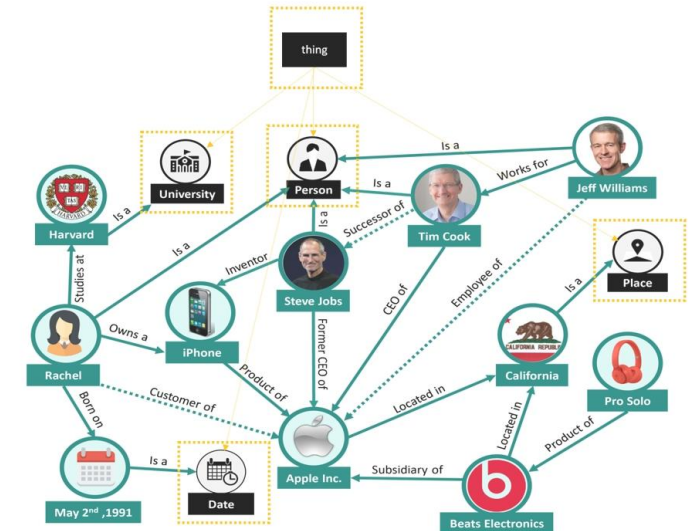
- ✓ Directed Labeled Multigraph: more generalized form than other graph forms
- ✓ 구조: Subject - Predicate - Object



Knowledge Graph from Text



Knowledge Graph from Images



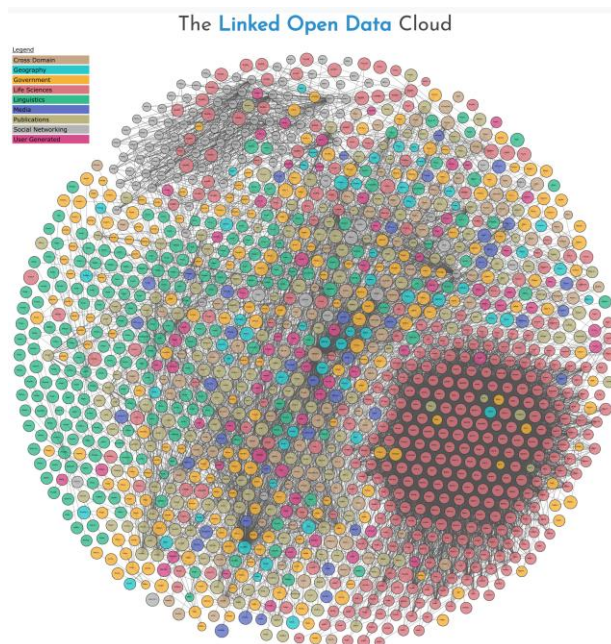
General Knowledge Graph

KGs store facts in a structural manner.

Knowledge Graph: Human Knowledge

❑ 지식그래프 (Knowledge Graph)

- ✓ 방대한 양의 지식 정보(데이터)와 데이터 간의 관계지식
- ✓ 복잡한 **인간의 지식**을 시각적이고 직관적인 방식으로 표현
 - 다양한 개념, 사실, 관계를 통합적으로 표현
 - 이를 통해 정보를 검색, 탐색 및 이해하기 쉽게 만들



Linked Open Data

Personalized Knowledge Graph

- 개인화된 지식 그래프
- 사용자 Context와 Contents에 대한 실시간 추론



Personal KG

사용자 컨텍스트 파악
맞춤형 추천 제공

Domain KG

도메인 특화 지식 정보 제공
전문 서비스 제공

Enterprise KG

비즈니스 데이터 검색
업무 생산성 향상 및 빠른 의사결정 도출

Knowledge Graph = Data and Ontology

- A structured collection of facts organized by semantic schemas
 - ✓ fact들의 구조화된 집합 (Data) + 의미론적 스키마로 조직화 (Ontology)

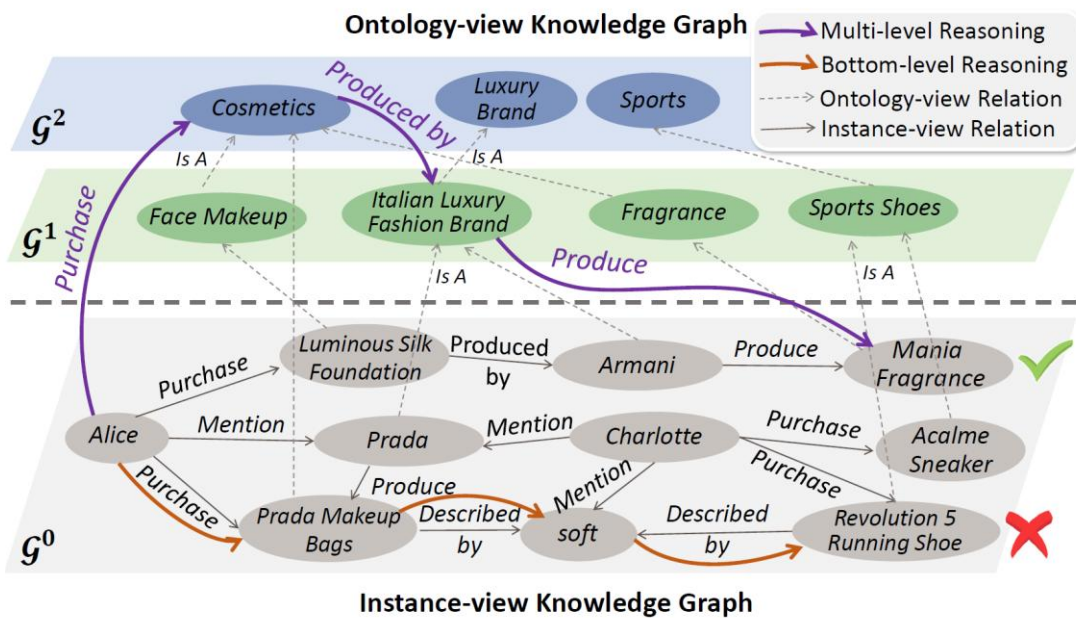
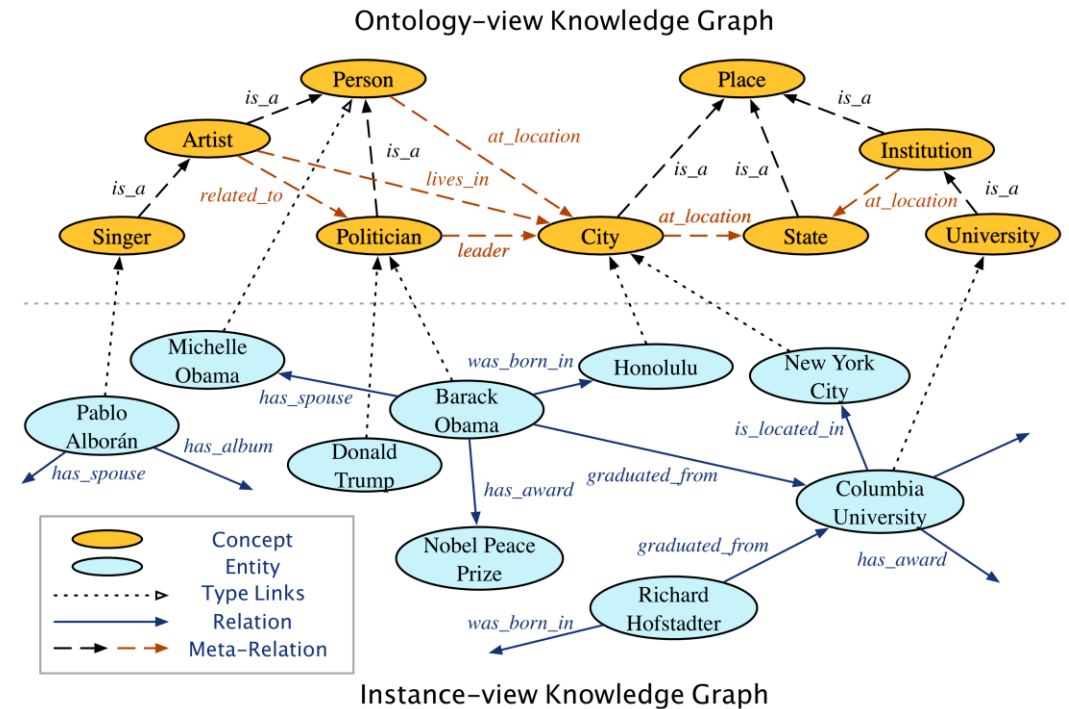


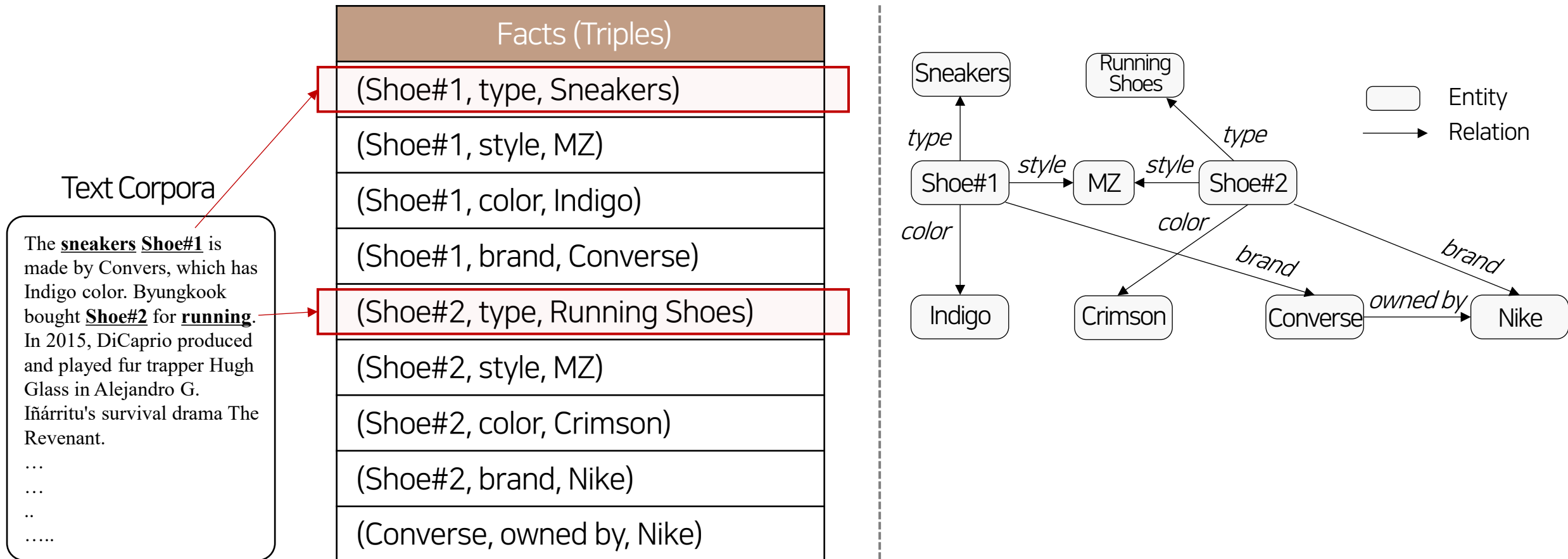
Figure 1: Multi-level reasoning over two-view KGs.



Knowledge Graph = Data and Ontology

❑ A structured collection of facts organized by semantic schemas

✓ fact들의 구조화된 집합 (Data) + 의미론적 스키마로 조직화 (Ontology)



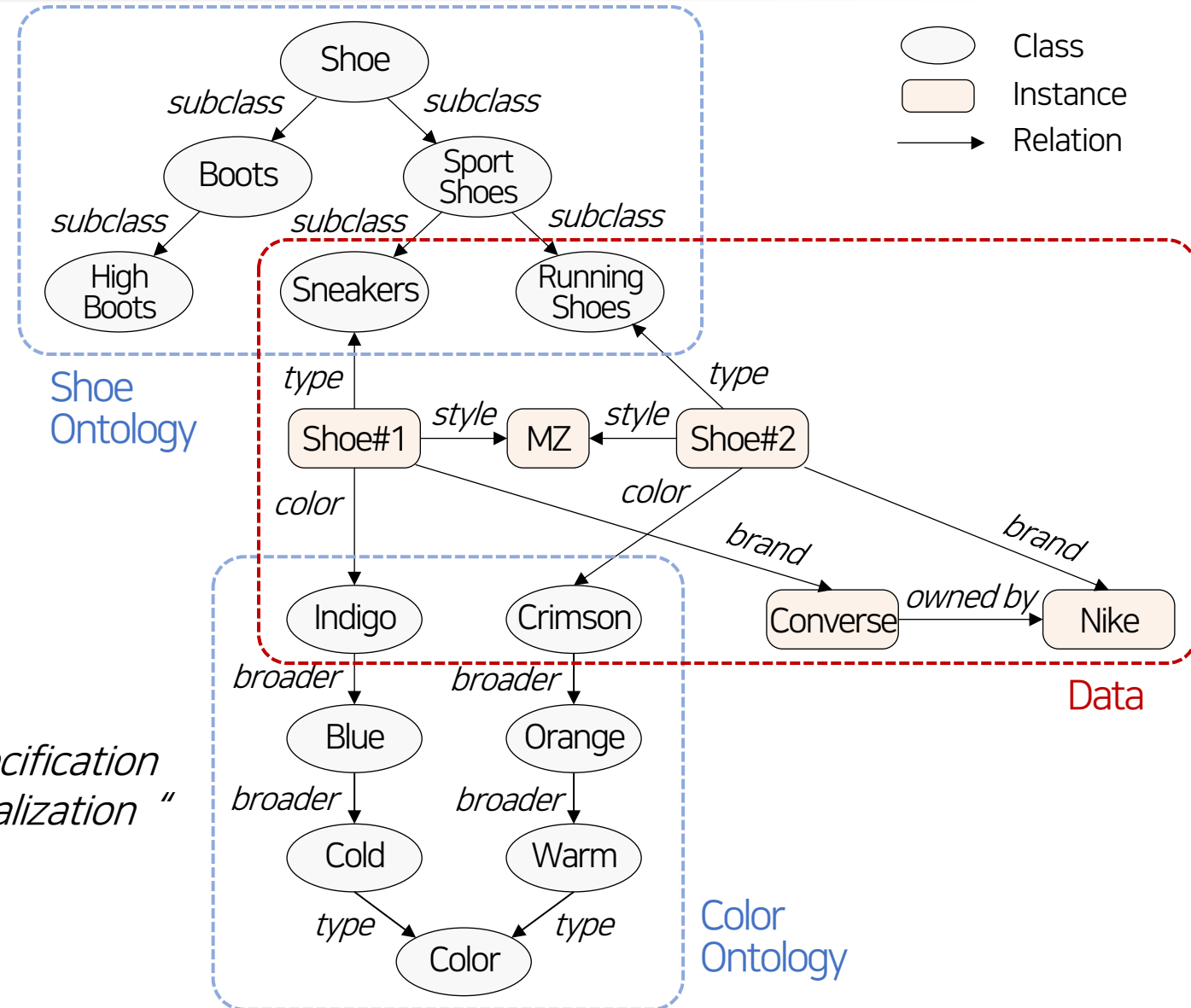
Knowledge Graph = Data and Ontology

✓ Data + Schema (Ontology)

- factual knowledge

✓ Reusable

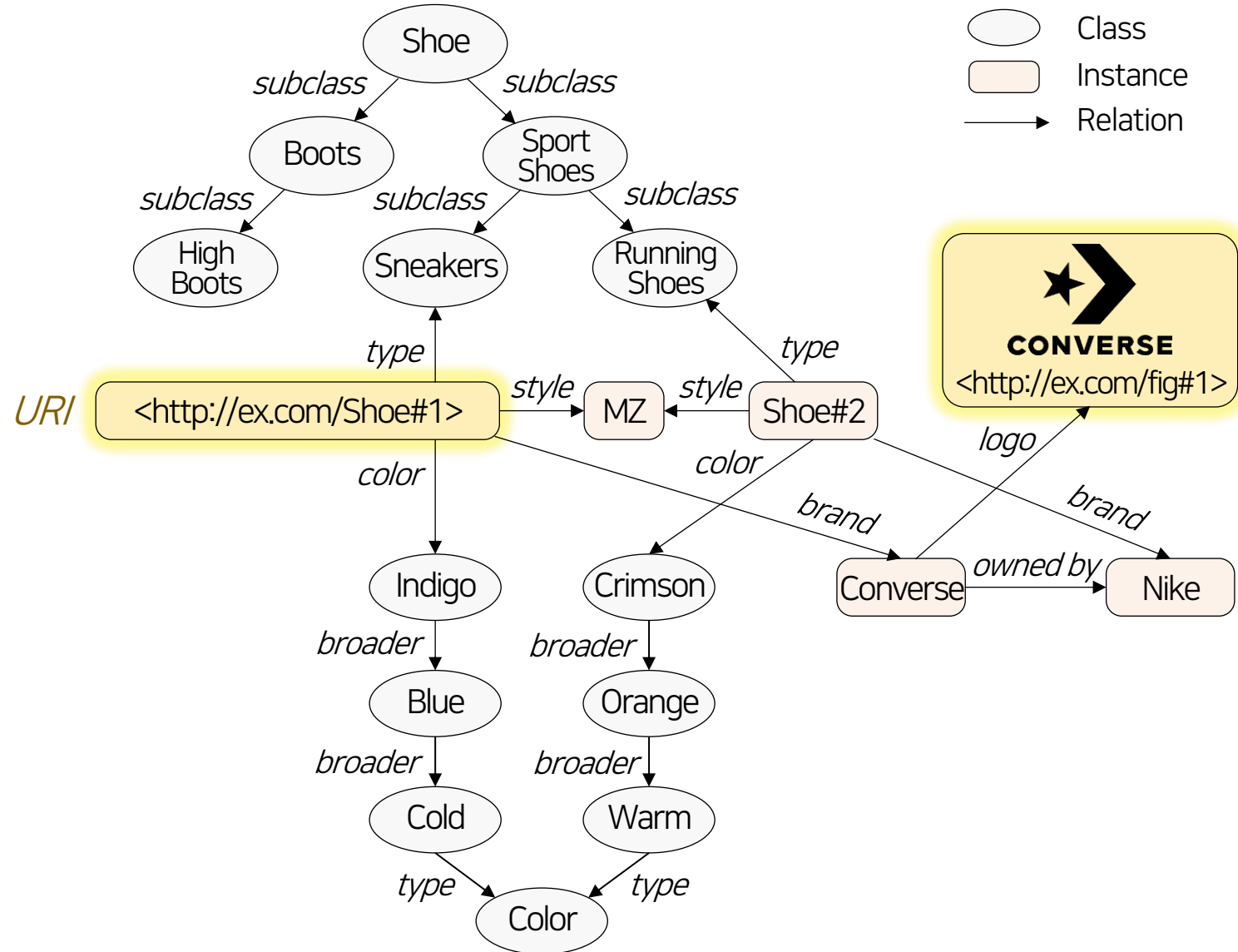
*" formal and explicit specification
of a shared conceptualization "*



Knowledge Graph = Data and Ontology

✓ Data + Schema (Ontology)

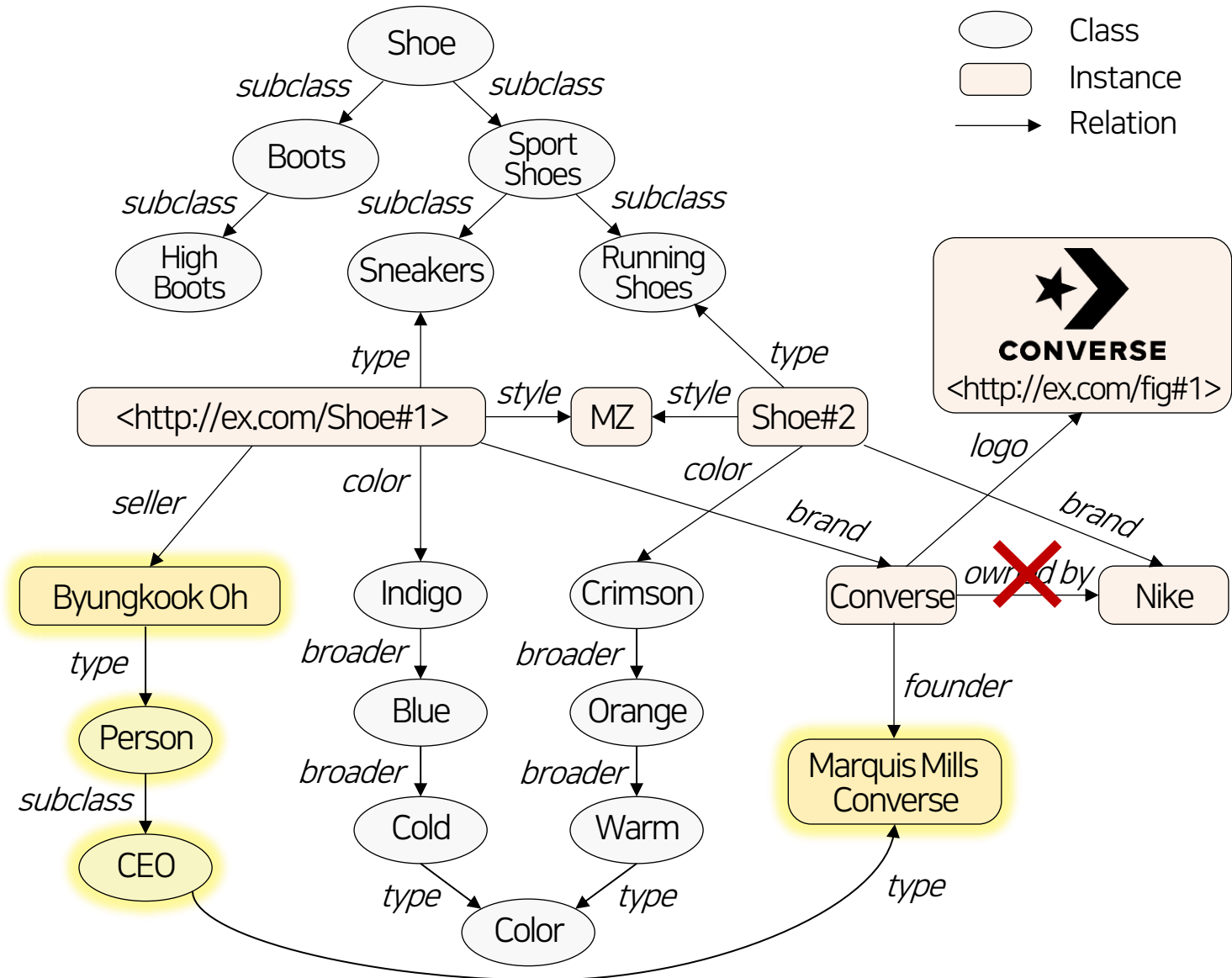
- factual knowledge
- ✓ Reusable
- ✓ Interoperable
- ✓ Machine Readable



Knowledge Graph = Data and Ontology

✓ Data + Schema (Ontology)

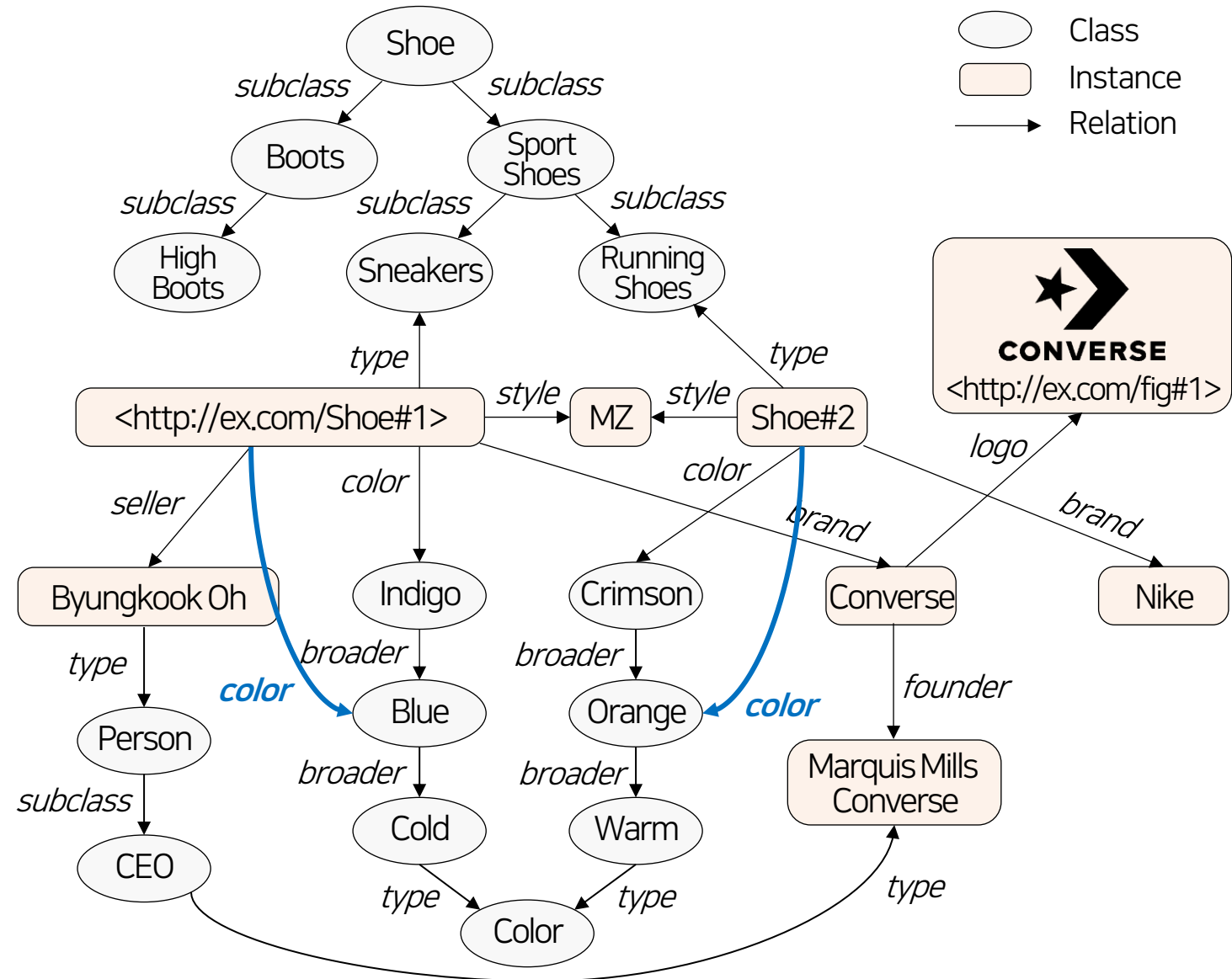
- factual knowledge
- ✓ Reusable
- ✓ Interoperable
- ✓ Machine Readable
- ✓ Flexible / Extensible



Knowledge Graph = Data and Ontology

✓ Data + Schema (Ontology)

- factual knowledge
- ✓ Reusable
- ✓ Interoperable
- ✓ Machine Readable
- ✓ Flexible / Extensible
- ✓ Inference (DL)
 - color + broader -> color



Knowledge Graph = Data and Ontology

✓ Data + Schema (Ontology)

- factual knowledge

✓ Reusable

✓ Interoperable

✓ Machine Readable

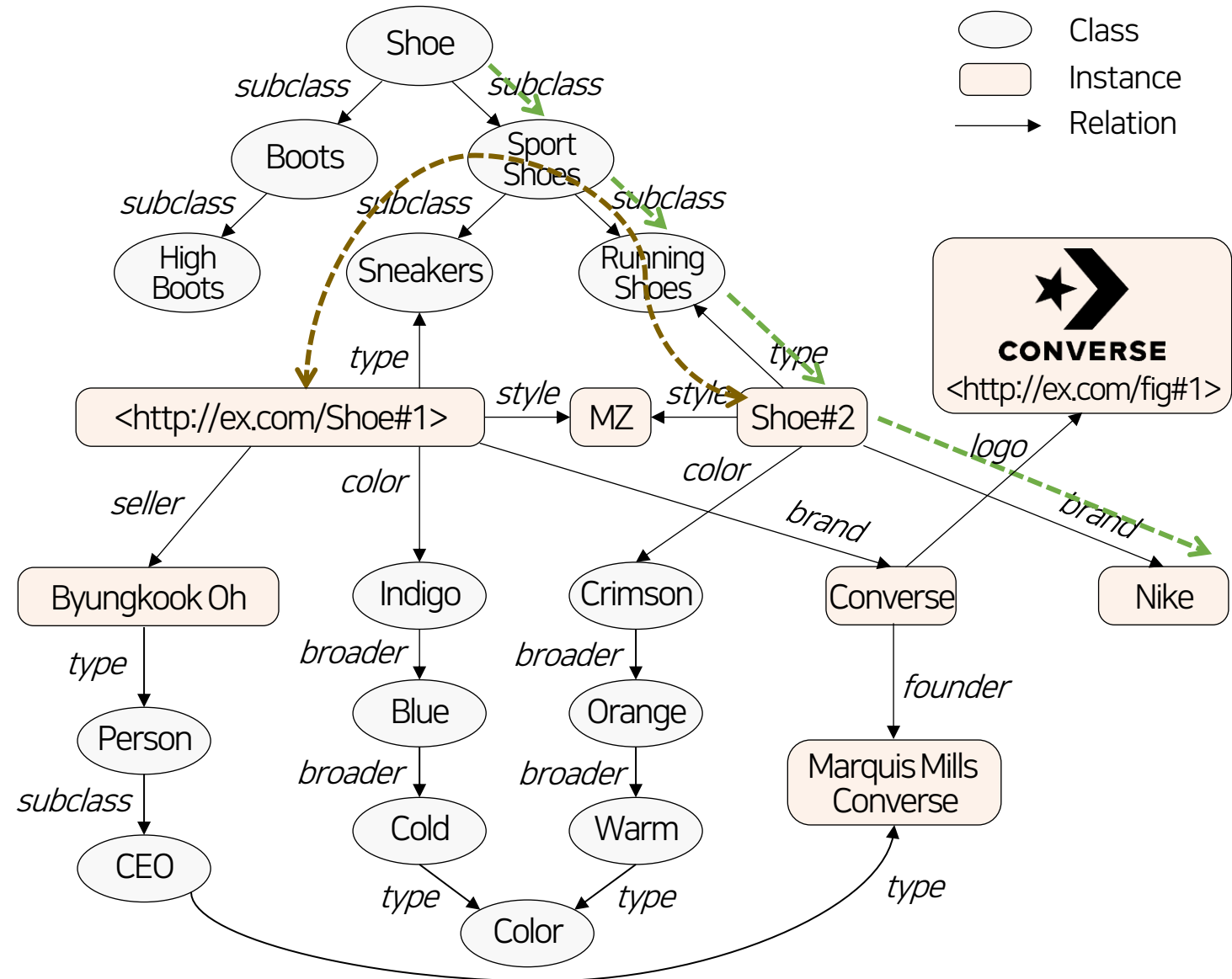
✓ Flexible / Extensible

✓ Inference (DL)

- color + broader -> color

✓ Other kinds of Queries

- Navigation
- Similarity & Locality (structure)



KGs help AGI

- ❑ KGs are transparent.

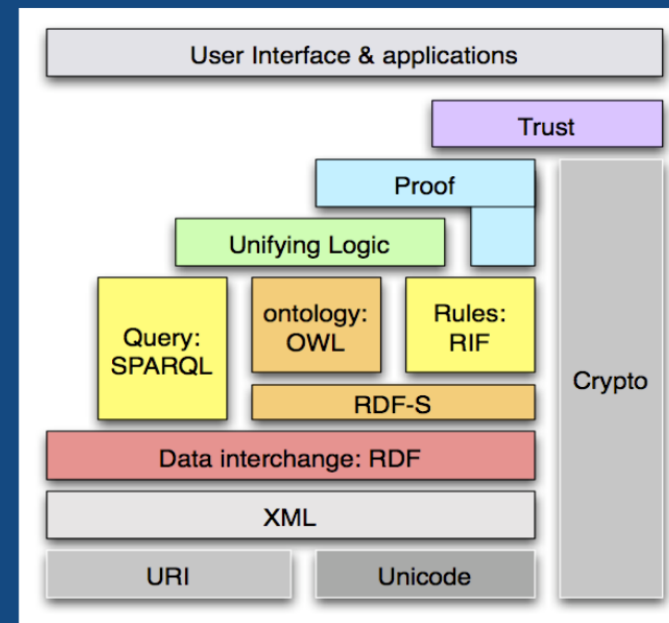
LLM is black-box

- How to represent knowledge?
- Why make such a decision?



KG is transparent

- Ontology and semantic definition
- Visible to users, e.g., nodes, edges
- Systematic store/exchange/update

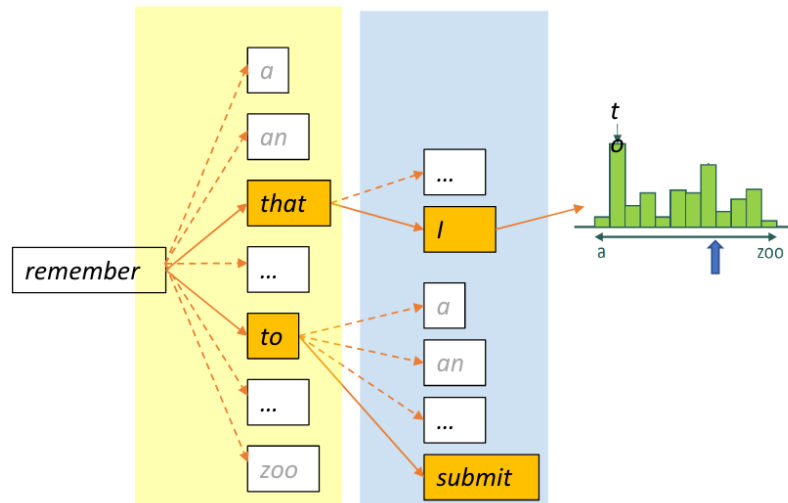


KGs help AGI

❑ KGs are adamant.

LLM is indecisive

- Easily swayed
- Anything with a probability



KG is adamant

- Mostly black and white facts
- Photographic memory

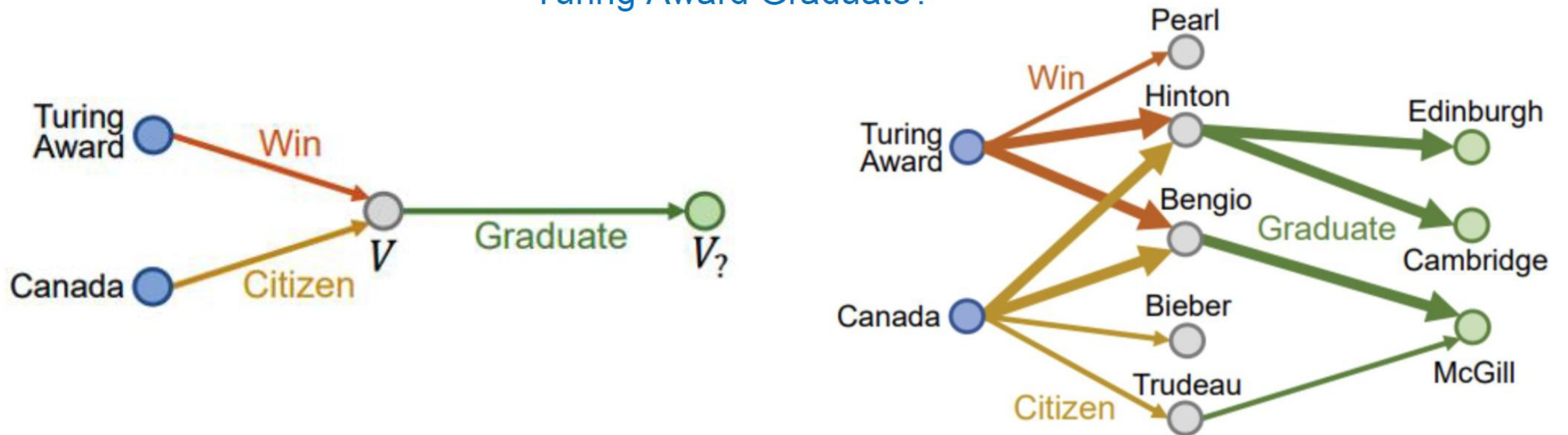


Capital	Singapore (city-state) 1°17'N 103°50'E
Official languages	English · Malay · Mandarin · Tamil
National language	Malay
Ethnic groups (2020)^[a]	74.3% Chinese 13.5% Malay 9.0% Indian 3.2% Others
Religion (2020)^[b]	31.1% Buddhism 20.0% No religion 18.9% Christianity 15.6% Islam 8.8% Taoism 5.0% Hinduism 0.6% Others
Demonym(s)	Singaporean
Government	Unitary dominant-party parliamentary republic
• President	Halimah Yacob
• Prime Minister	Lee Hsien Loong
Legislature	Parliament

KGs help AGI

- ❑ KGs power symbolic reasoning.

Q. Where did Canadian citizens with Turing Award Graduate?

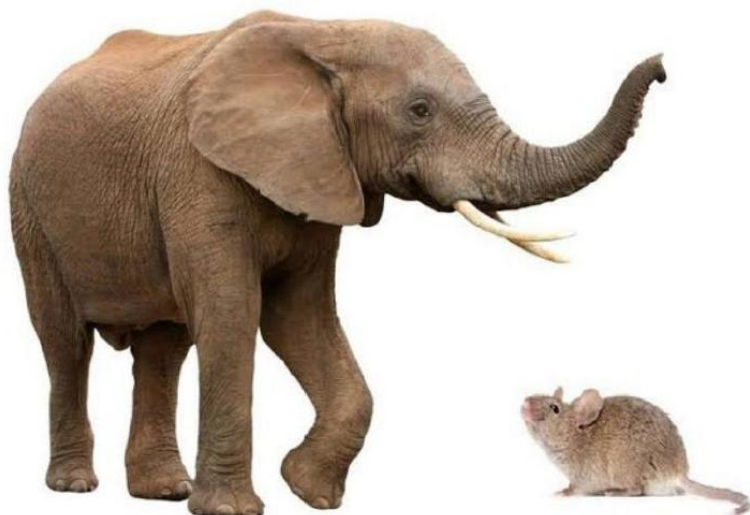


KGs help AGI

- ❑ KGs can provide domain-specific knowledge.

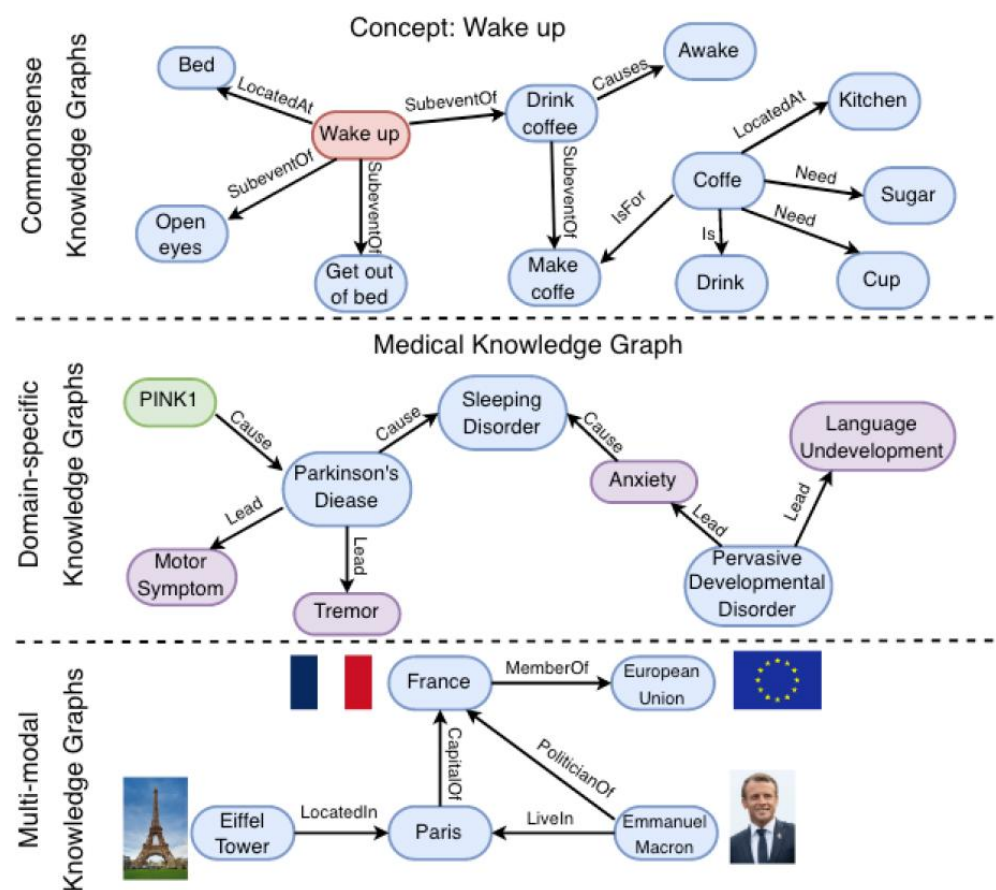
LLM is hungry

- More data more parameters
- Learn new knowledge inefficiently



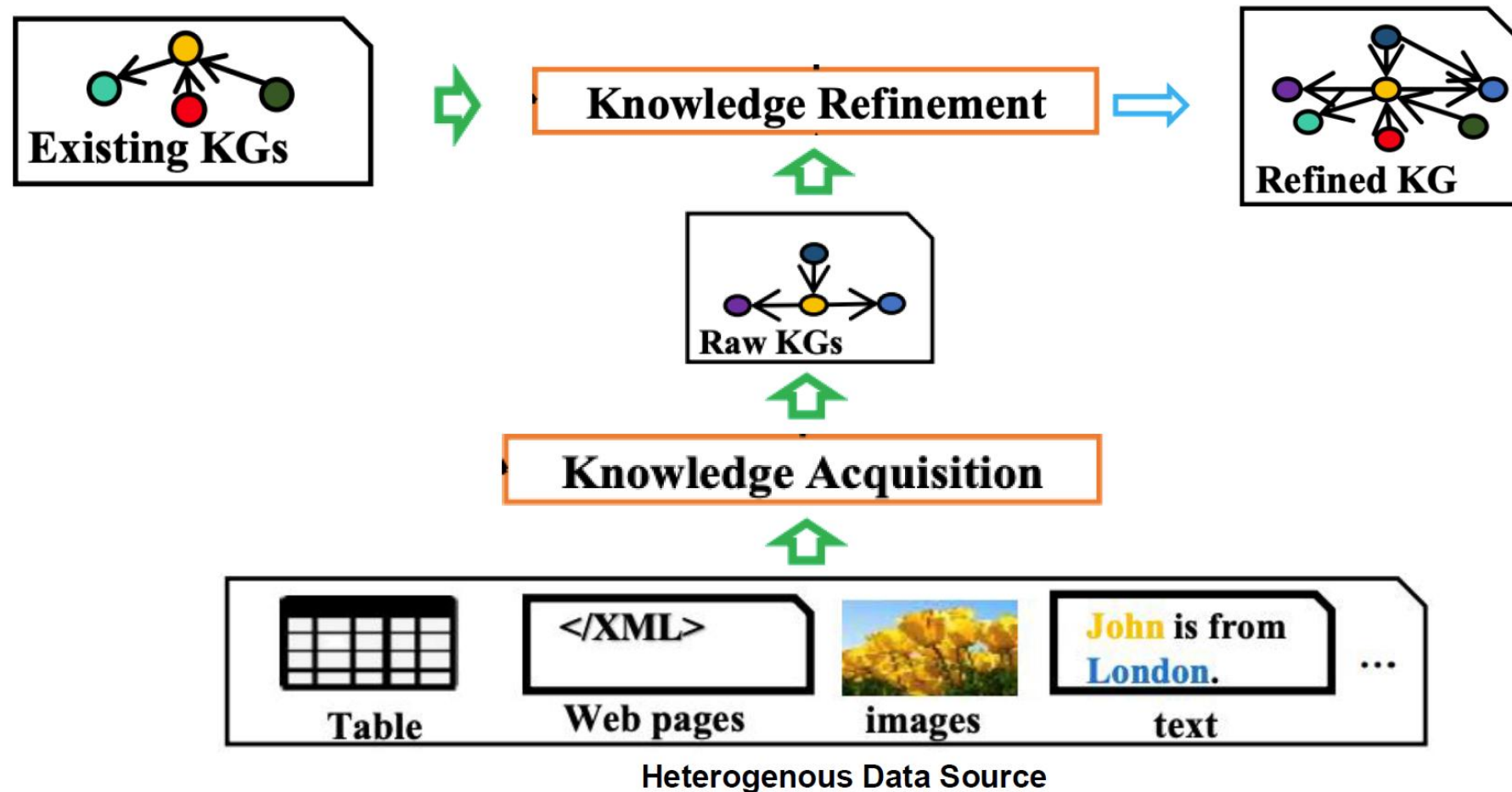
Pretrain data

Task-specific data



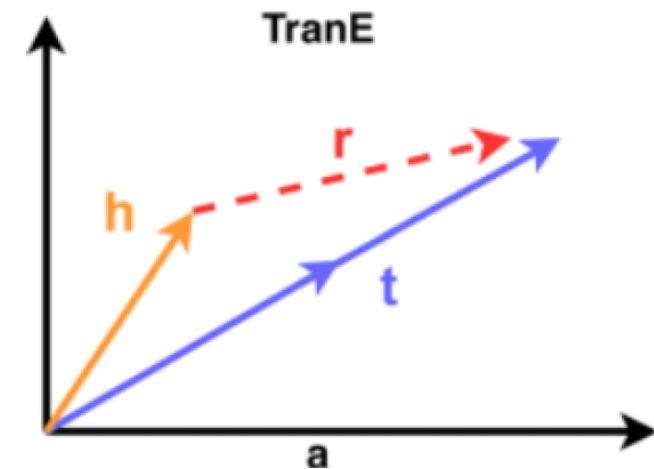
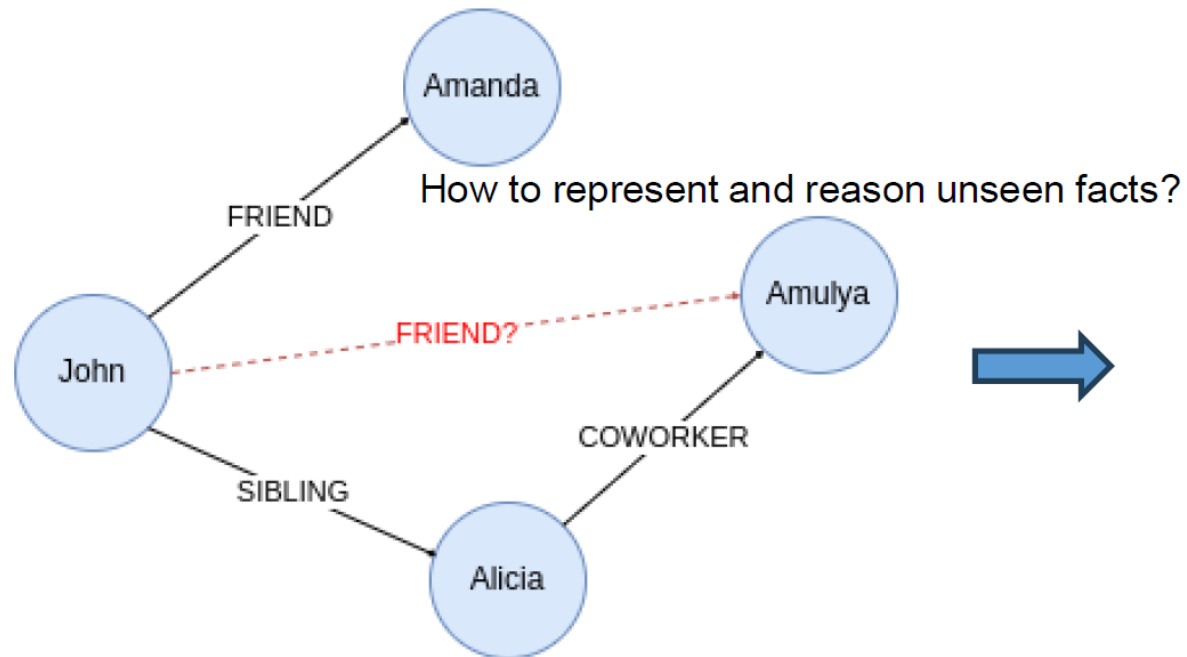
Limitations of KGs

- ❑ KGs are difficult to construct.



Limitations of KGs

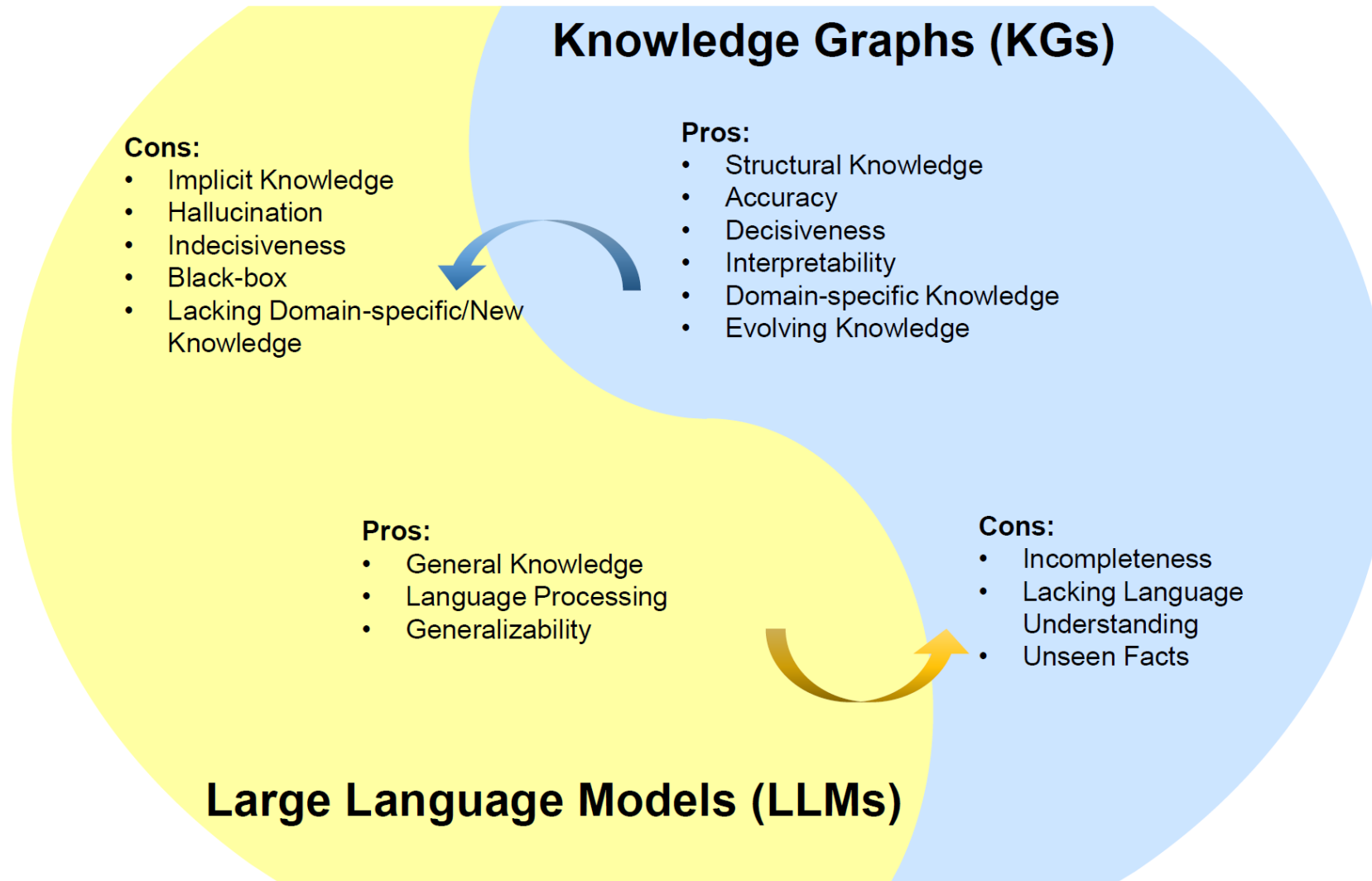
- ❑ KGs are **incomplete** and **noisy**.



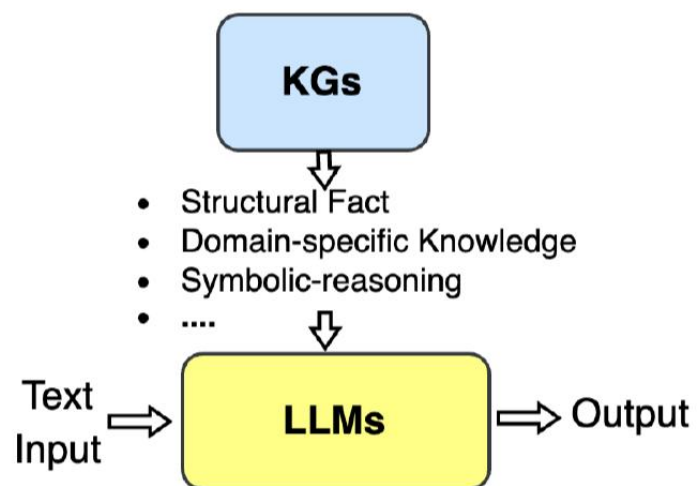
Embedding model

How to represent and reason unseen facts?

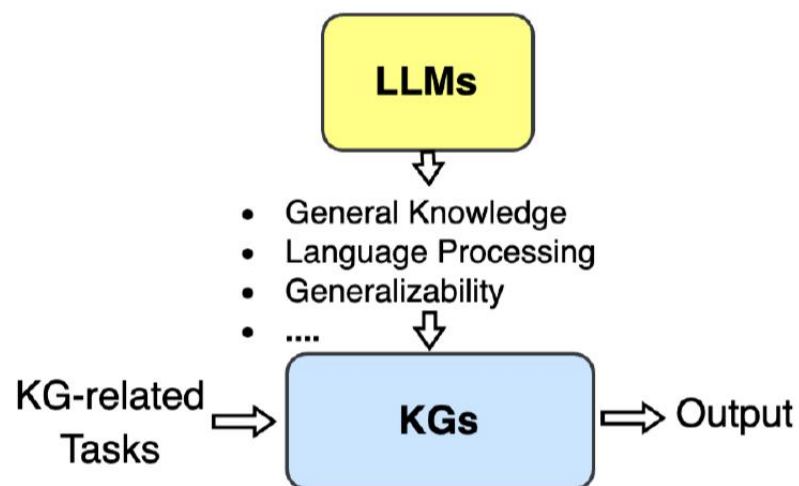
Synergy of LLMs and KGs towards AGI



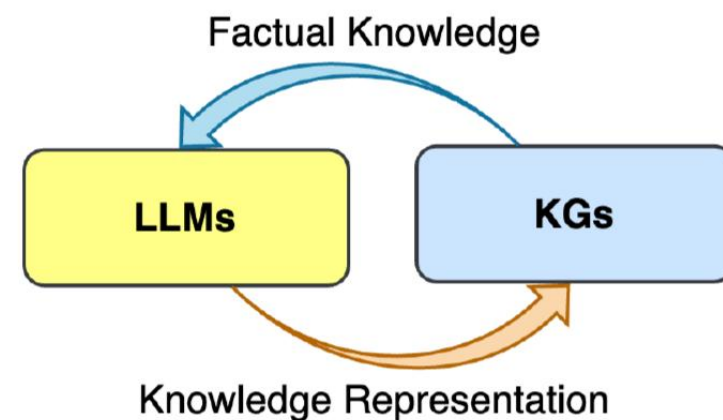
Roadmaps



a. KG-enhanced LLMs

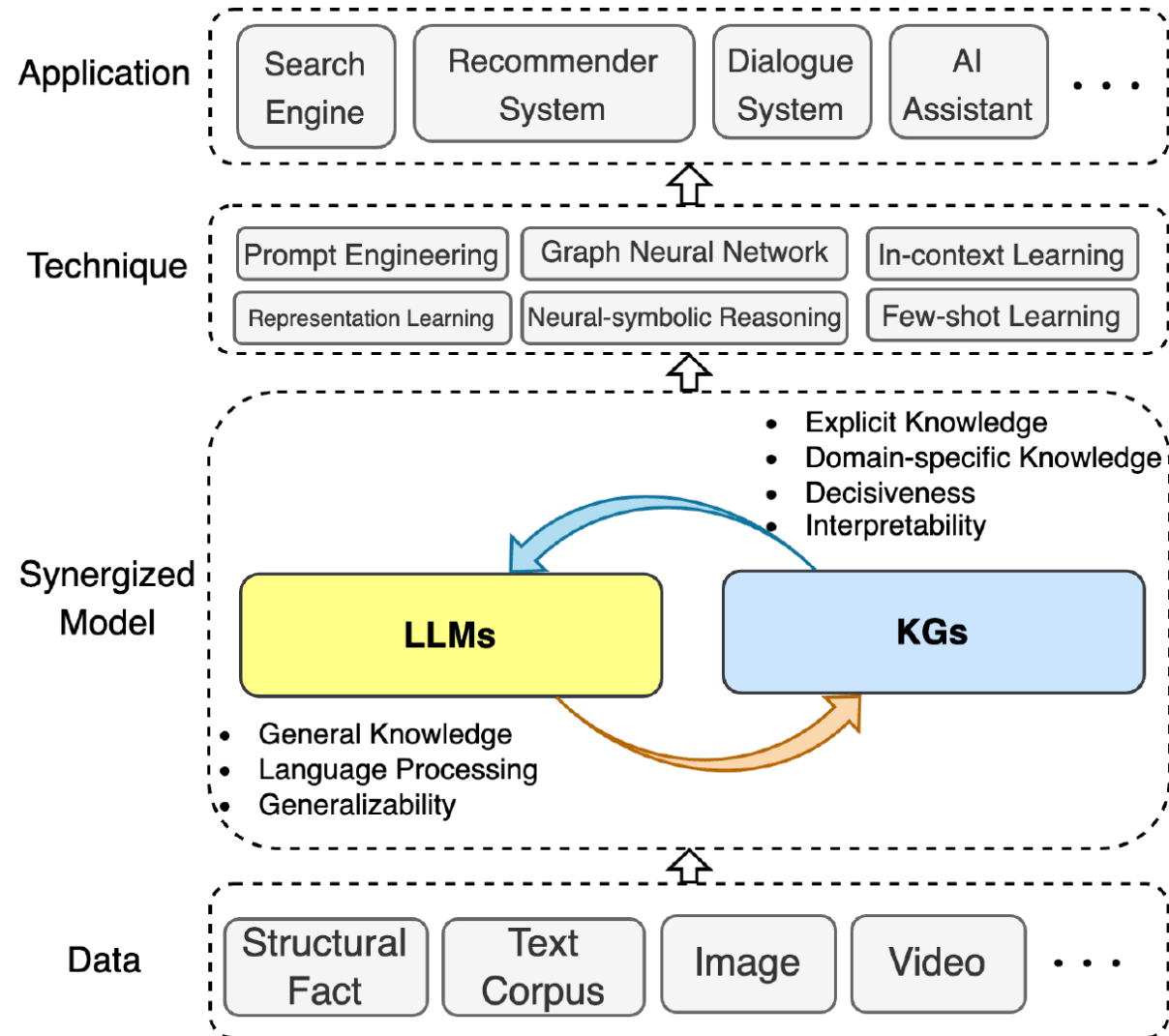


b. LLM-augmented KGs



c. Synergized LLMs + KGs

Roadmaps

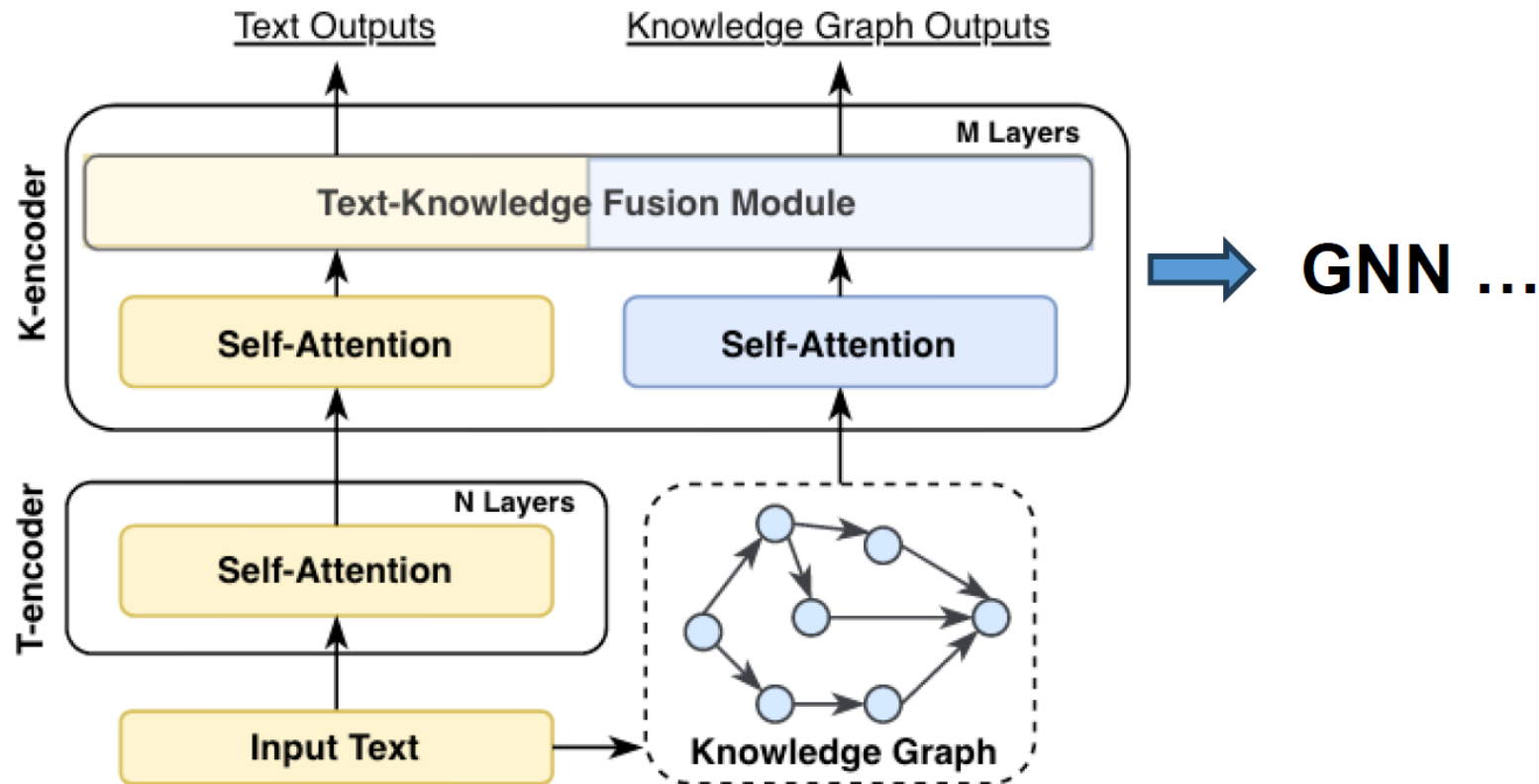


CONTENTS

1. Today
2. Background
3. Large Language Models
4. Knowledge Graphs
- 5. Graph-enhanced Training**
6. Graph-enhanced Reasoning
7. LLM-based Recommendation

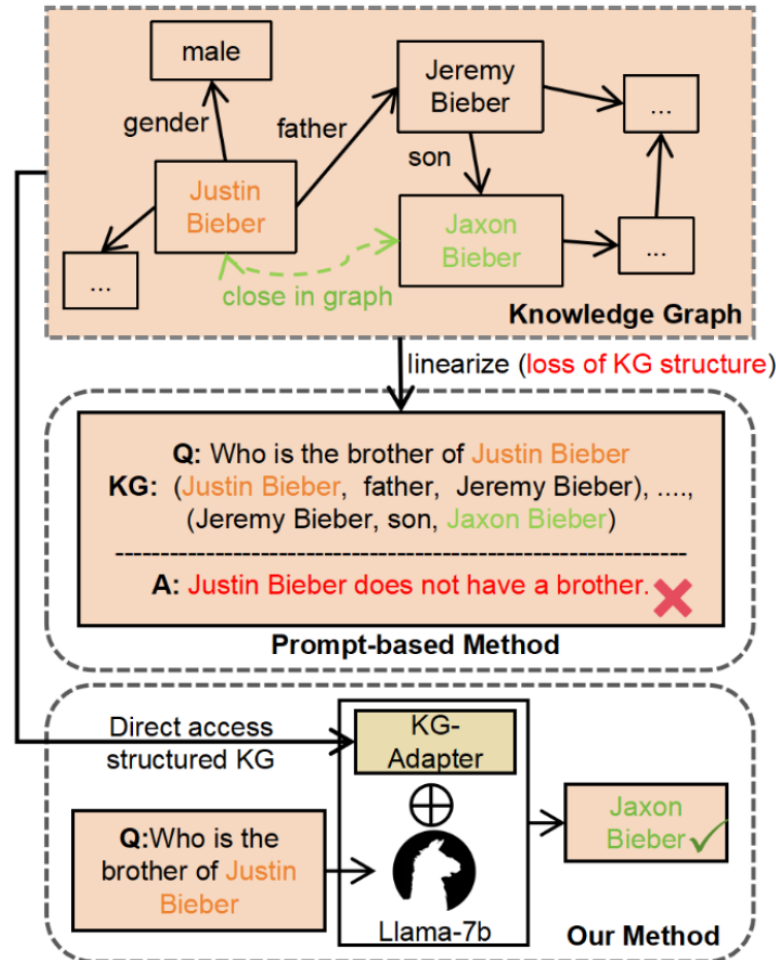
KG-enhanced LLM Training - Cokebert

- ❑ Integrating KGs by Additional Fusion Modules
 - ✓ Additional modules to better capture the structure knowledge of KGs.



KG-enhanced LLM Training - KG-adapter

❑ Integrating KGs by Additional Fusion Modules



KG-enhanced LLM Training - GraphGPT

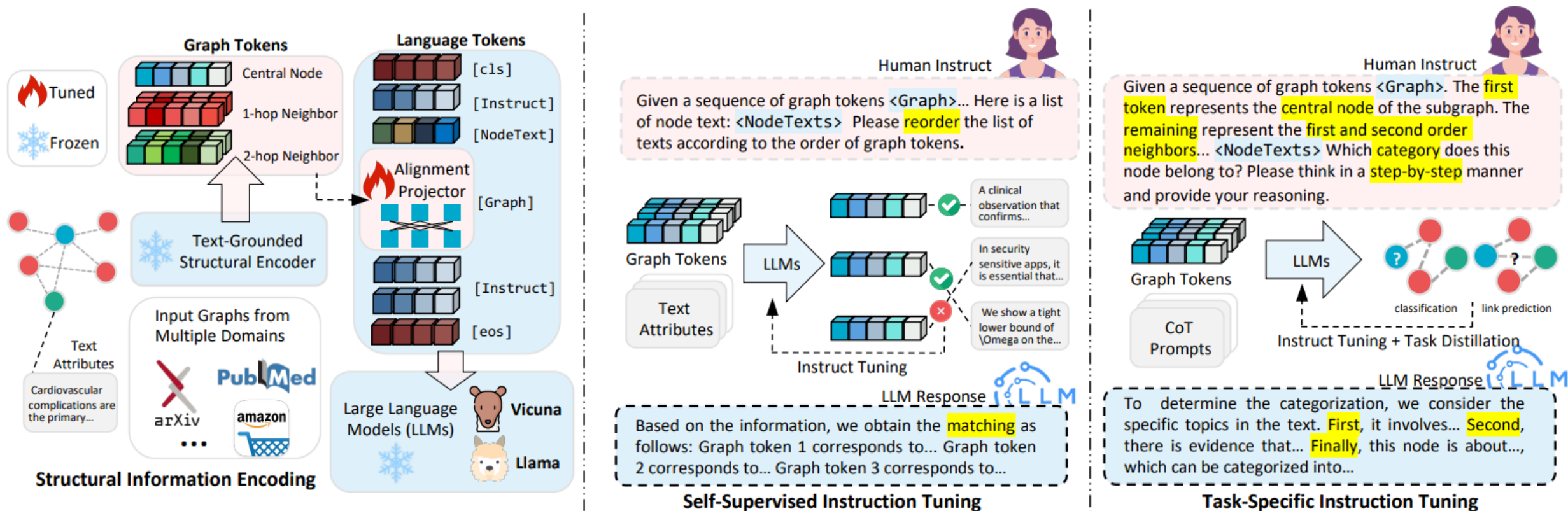


Figure 2: The overall architecture of our proposed GraphGPT with graph instruction tuning paradigm.

However,,,

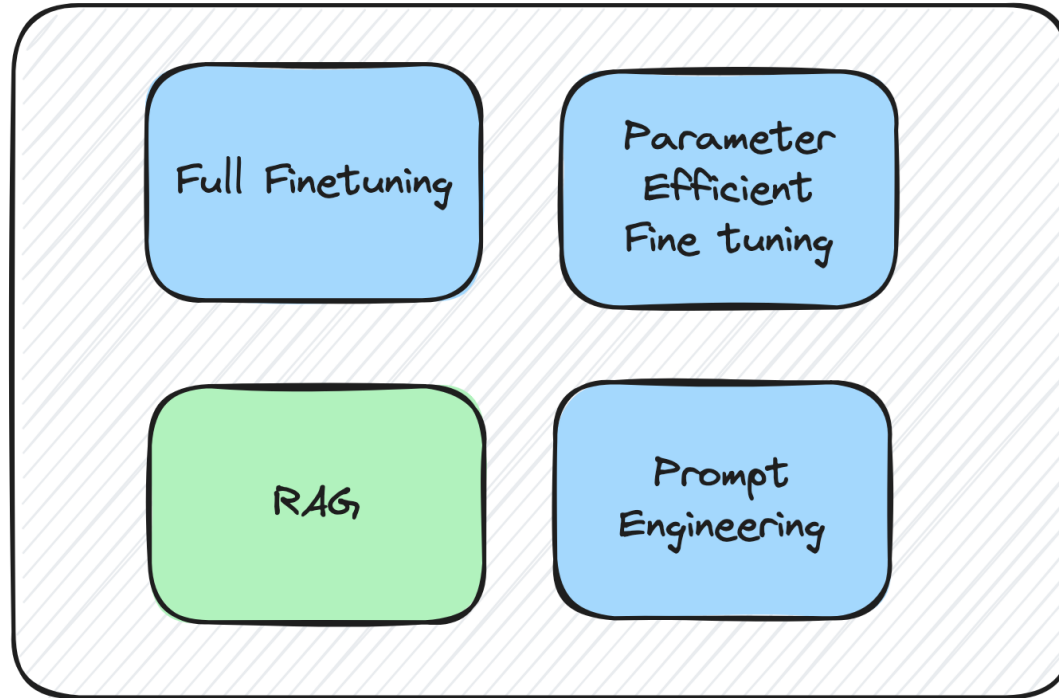
- ❑ KG-enhanced LLM training could fuse knowledge into LLMs.
- ❑ However, real-world knowledge is subject to change, and the pre-training approaches cannot update knowledge without retraining the model.
- ❑ KG-enhanced LLM Reasoning aims to separate the knowledge and text and inject the structural knowledge while LLM reasoning.

CONTENTS

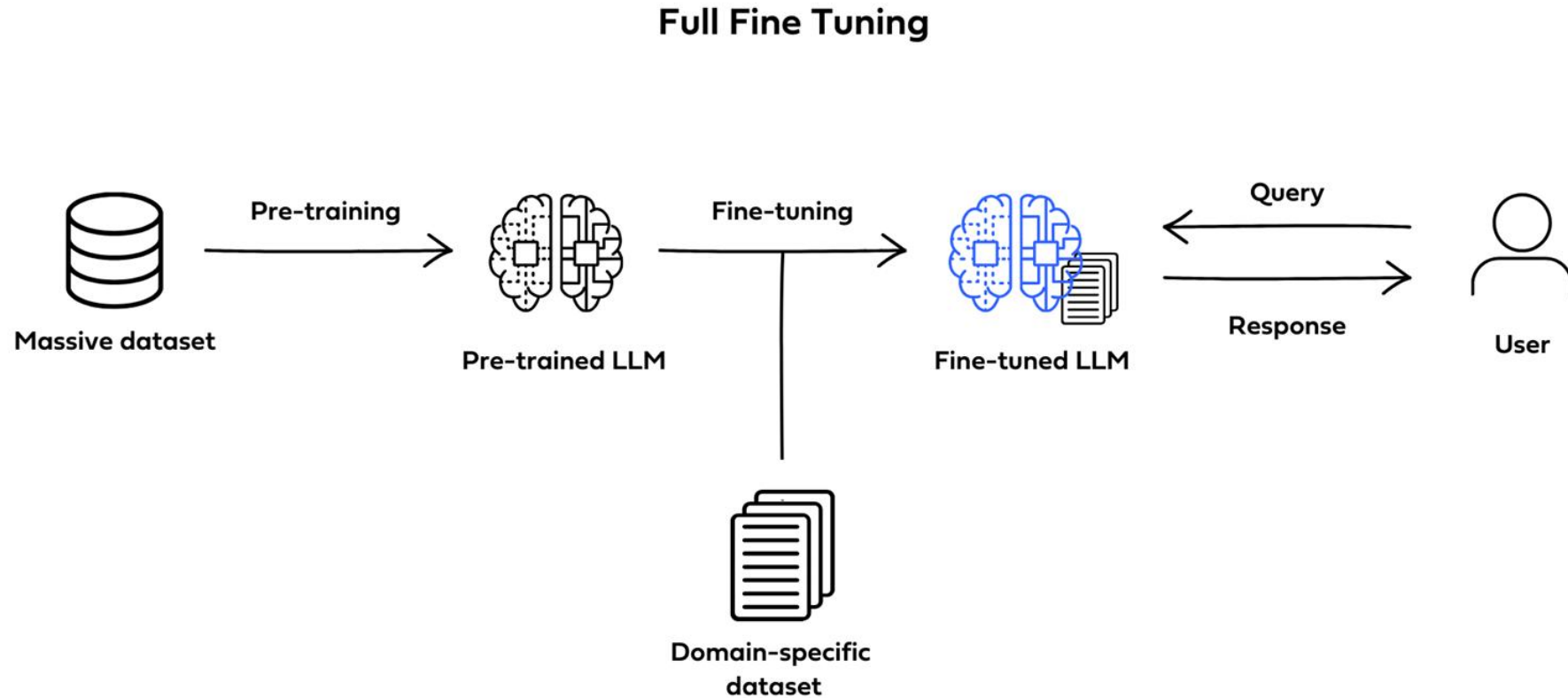
1. Today
2. Background
3. Large Language Models
4. Knowledge Graphs
5. Graph-enhanced Training
- 6. Graph-enhanced Reasoning**
7. LLM-based Recommendation

Retrieval Augmented Generation

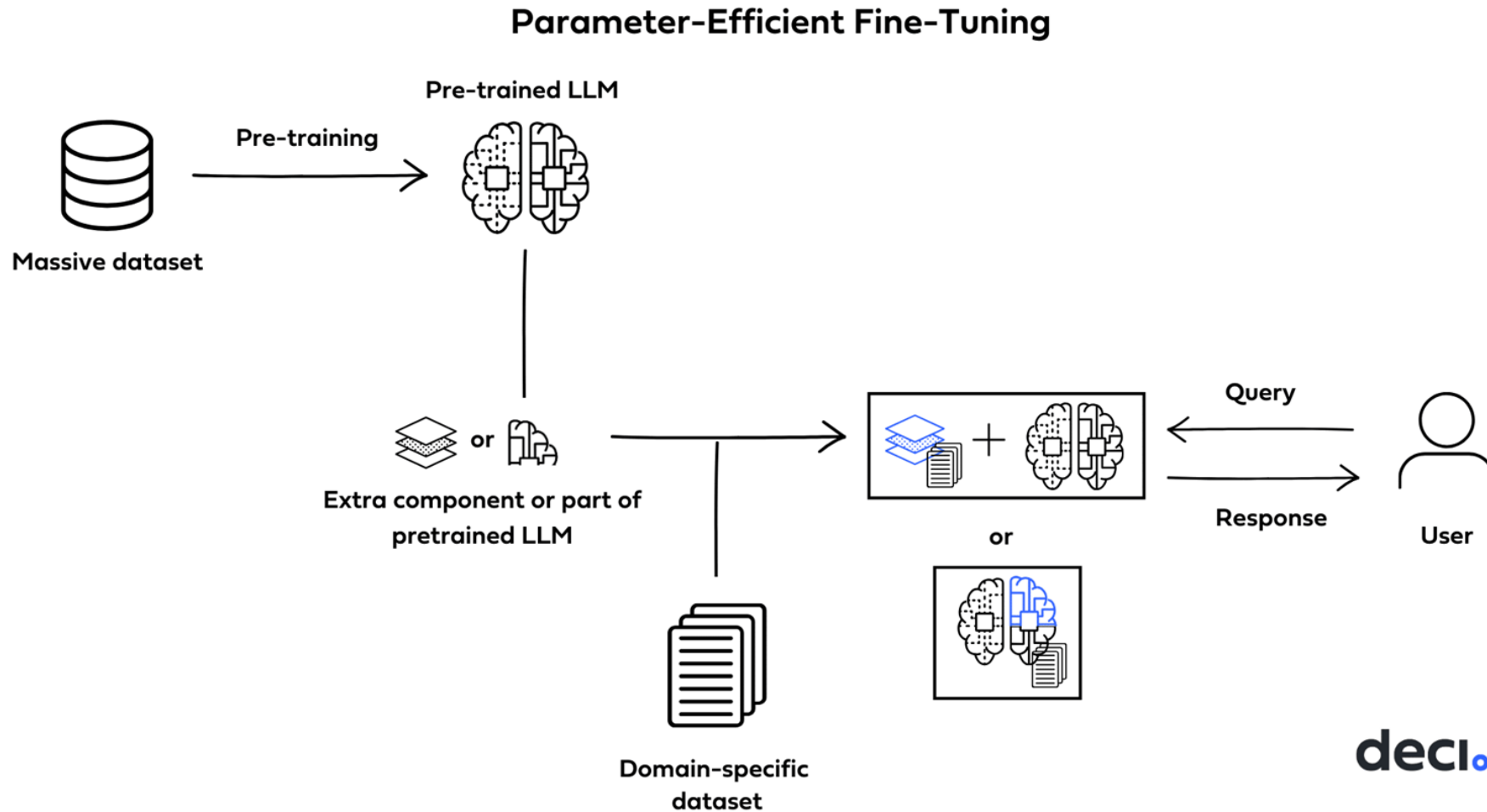
LLM Tuning method



Retrieval Augmented Generation

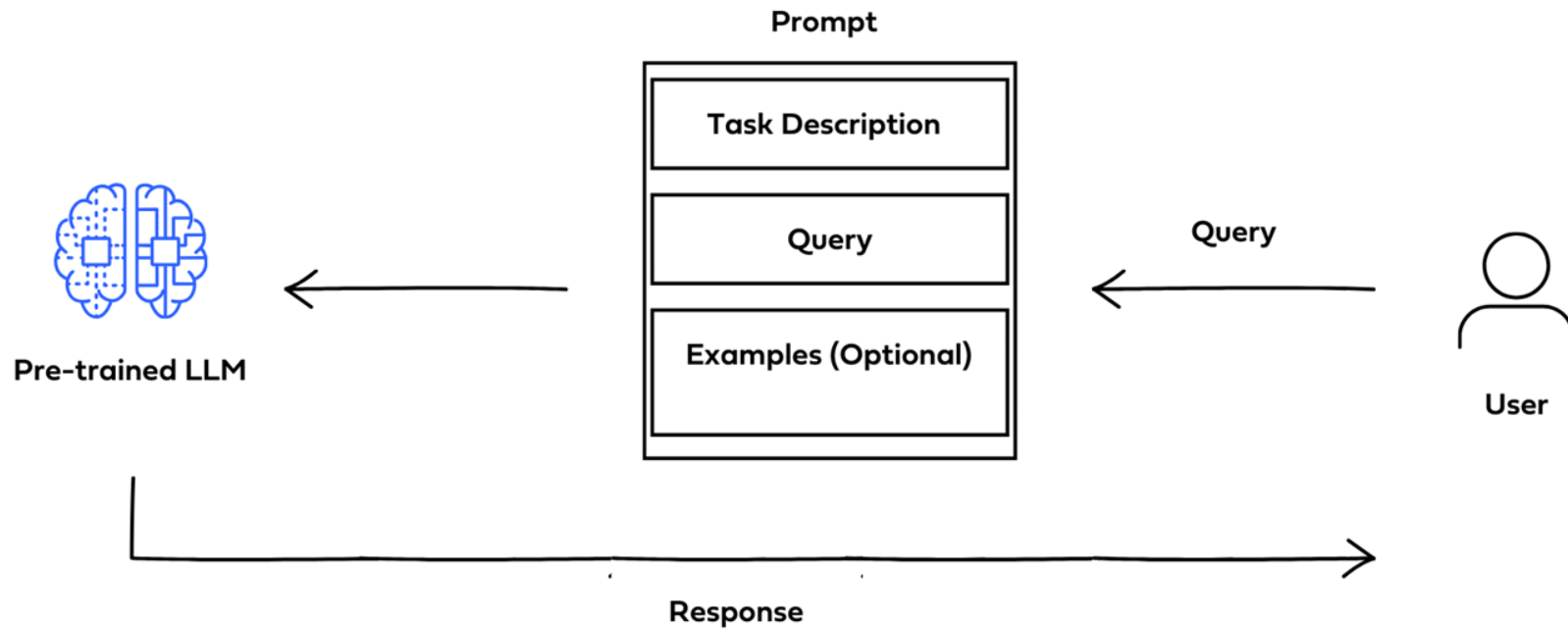


Retrieval Augmented Generation

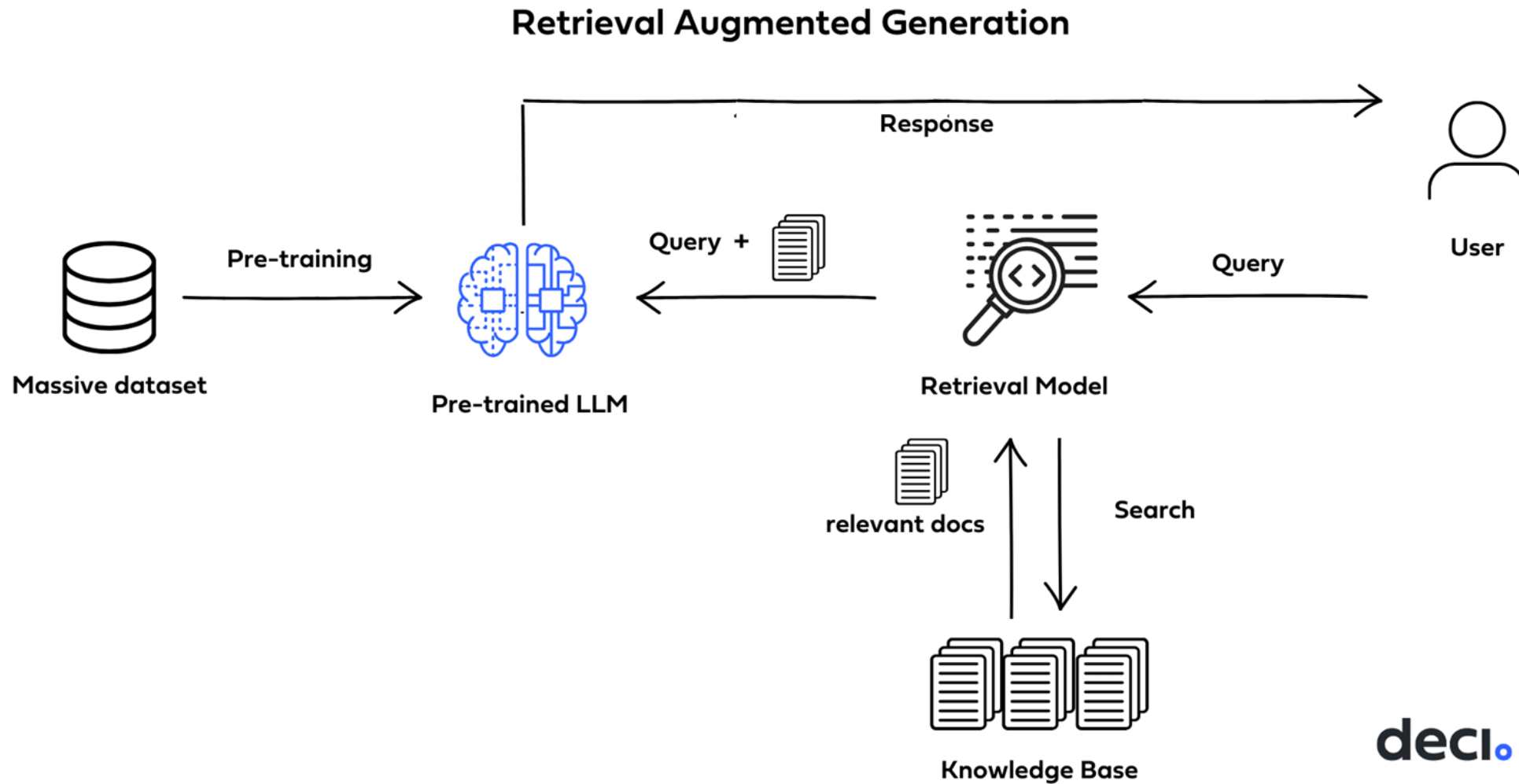


Retrieval Augmented Generation

Basic Prompt Engineering

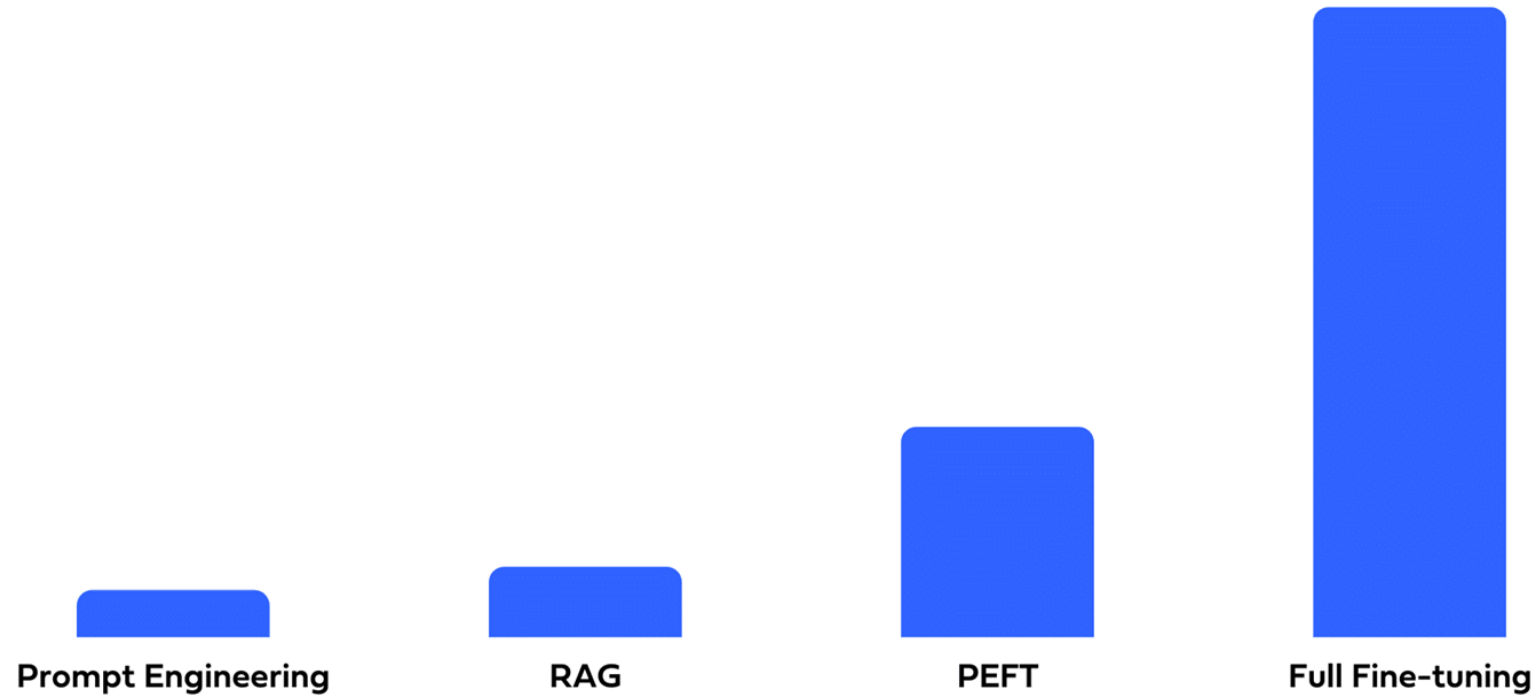


Retrieval Augmented Generation

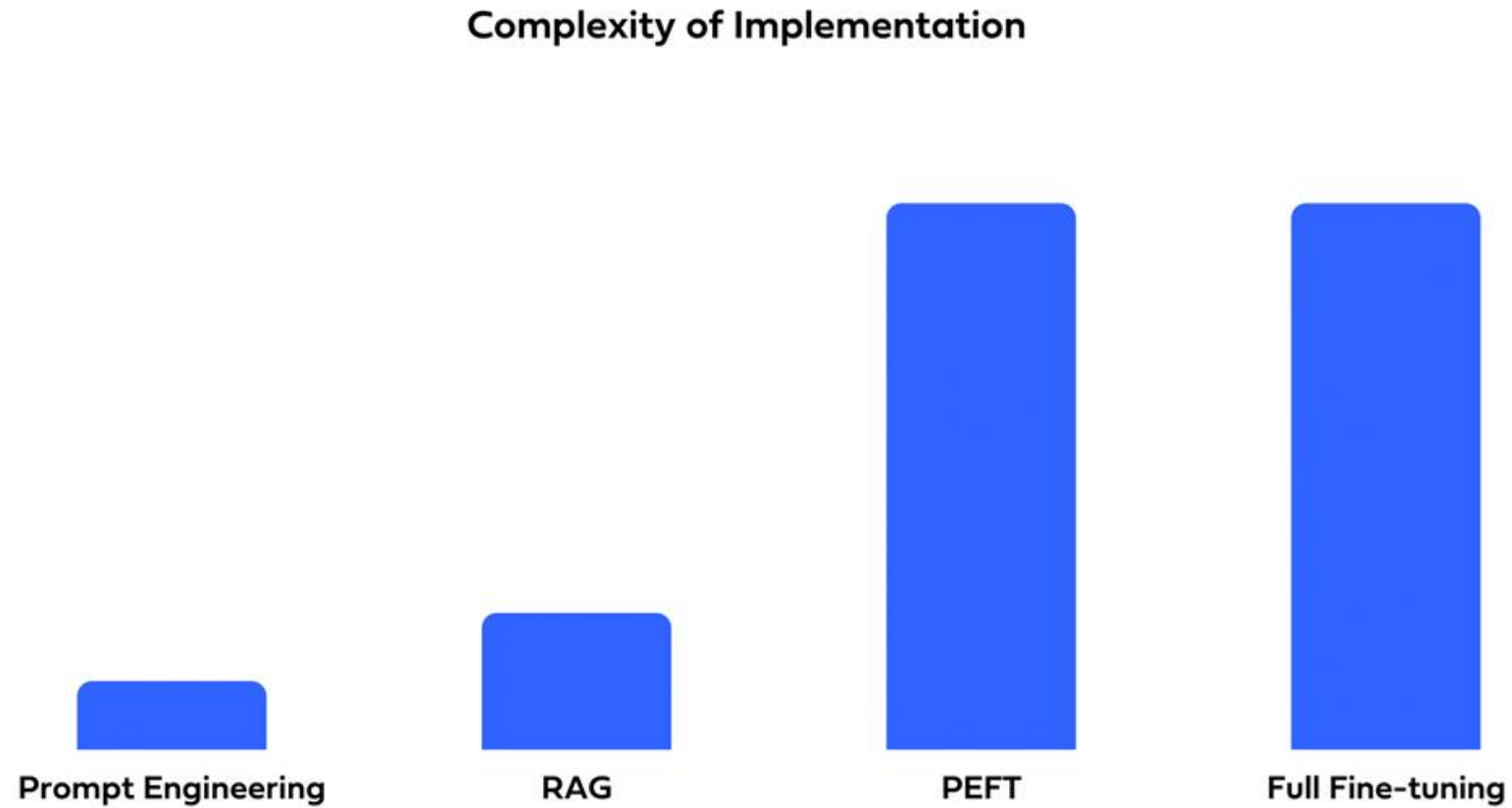


Retrieval Augmented Generation

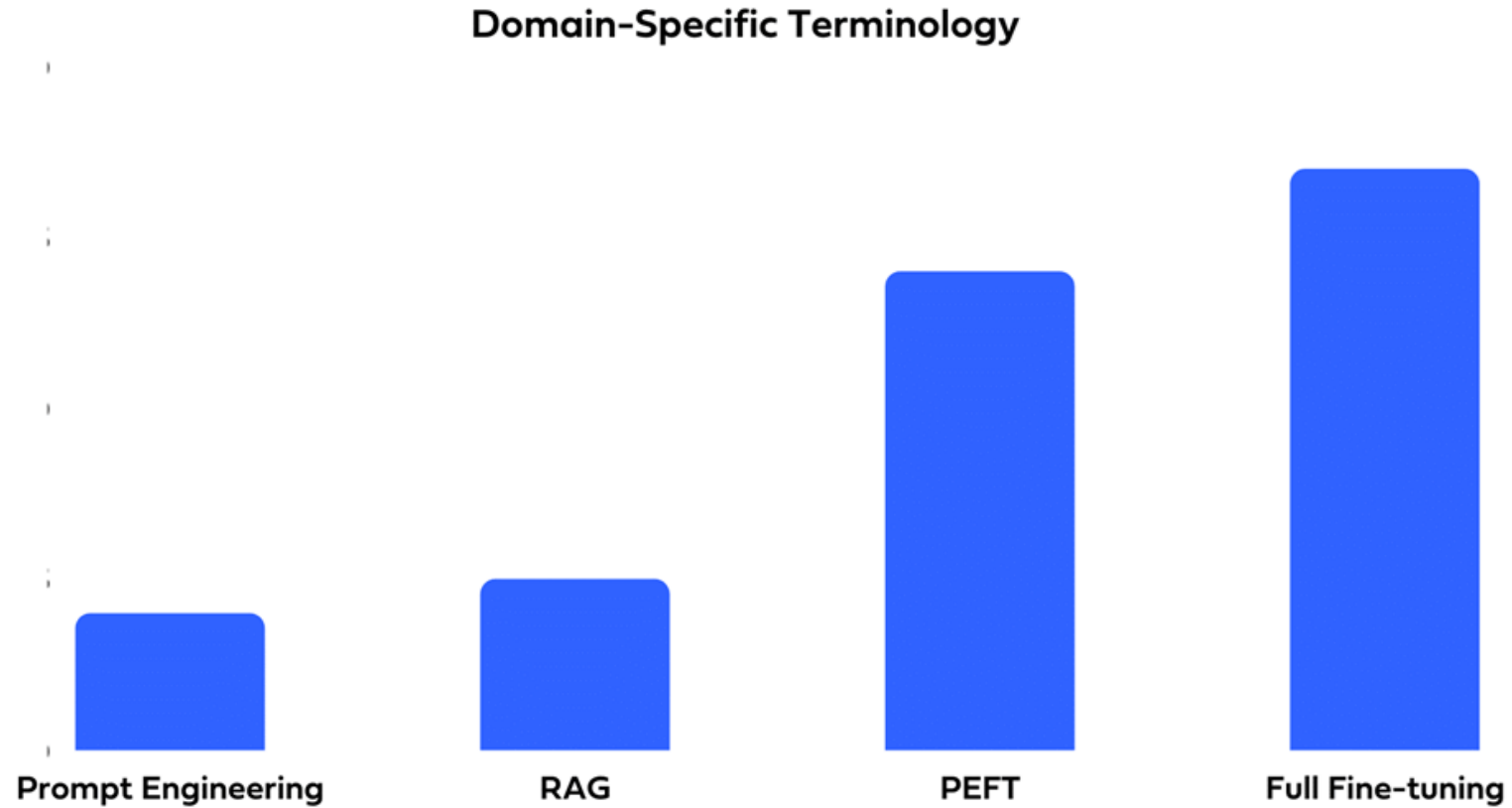
Cost of Implementation & Maintenance



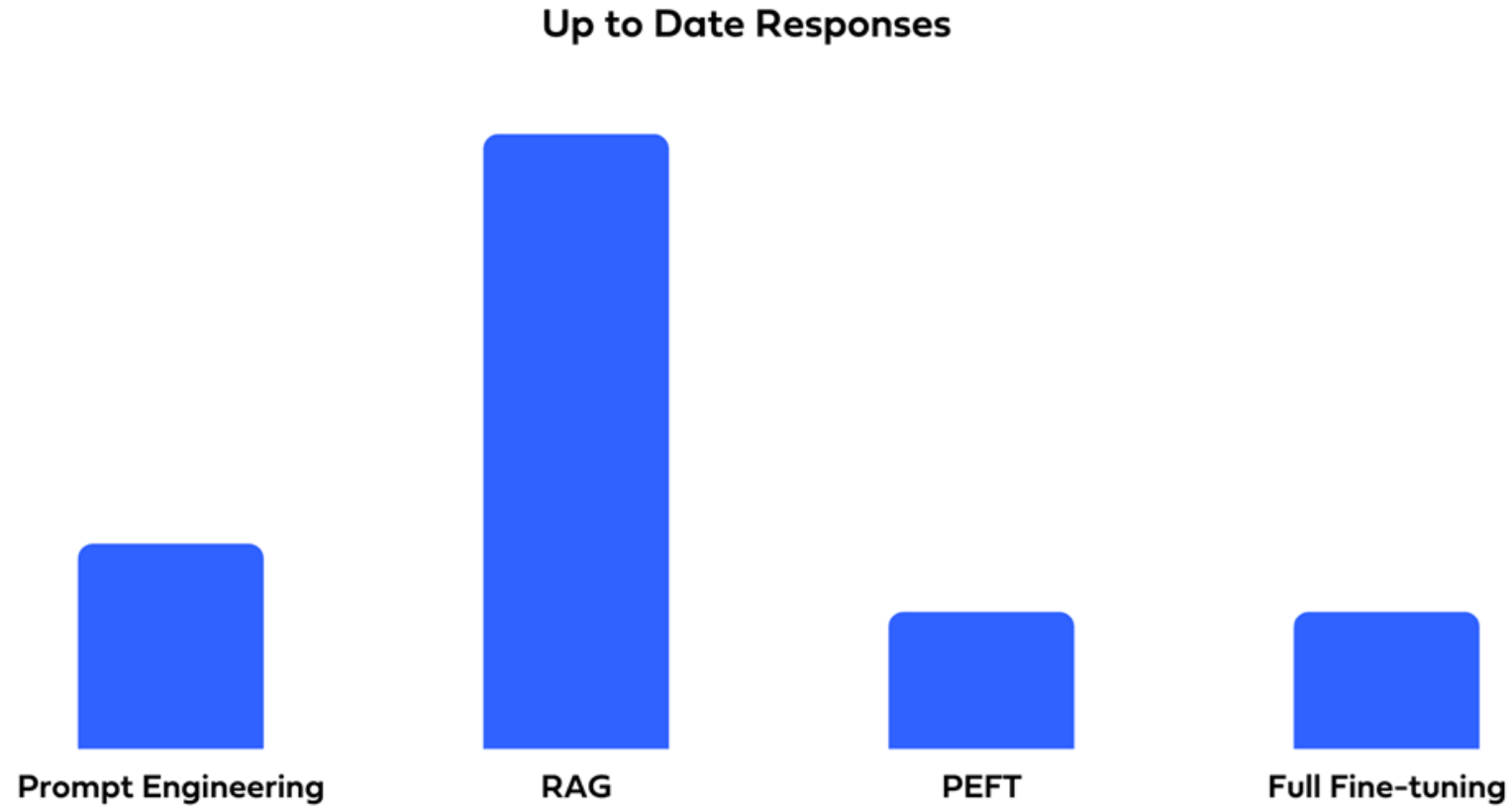
Retrieval Augmented Generation



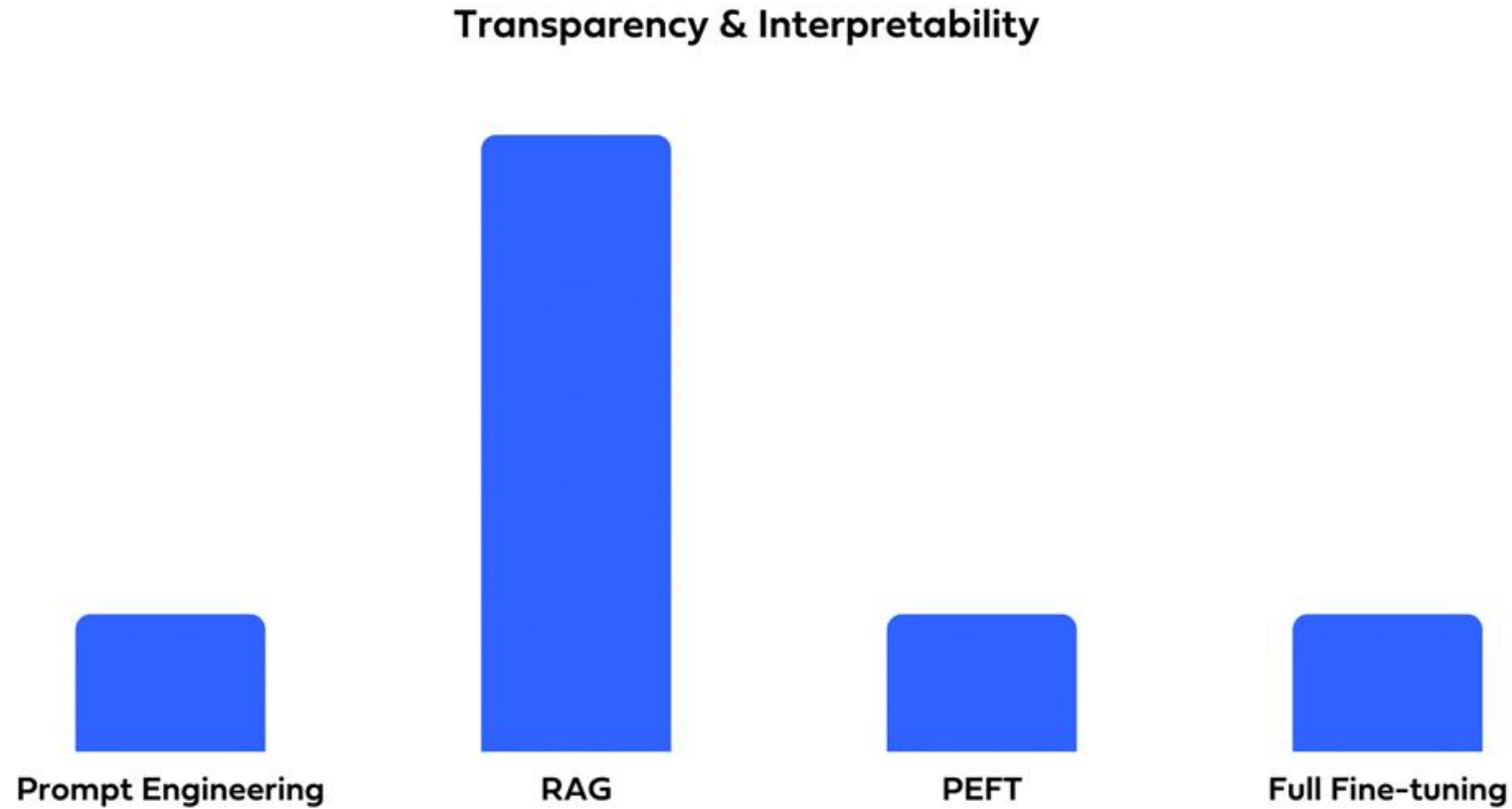
Retrieval Augmented Generation



Retrieval Augmented Generation



Retrieval Augmented Generation



Retrieval Augmented Generation

When to fine-tune

We may have some tasks where even the most cleverly designed prompts fall short. For example, even after significant prompt engineering, our system may still be a ways from returning reliable, high-quality output. If so, then it may be necessary to finetune a model for your specific task.

Successful examples include:

- Honeycomb's Natural Language Query Assistant: Initially, the "programming manual" was provided in the prompt together with n-shot examples for in-context learning. While this worked decently, fine-tuning the model led to better output on the syntax and rules of the domain-specific language.
- ReChat's Lucy: The LLM needed to generate responses in a very specific format that combined structured and unstructured data for the frontend to render correctly. Fine-tuning was essential to get it to work consistently.

도메인 specific할 때 fine-tuning
(ex, NL-to-SQL)

When to RAG

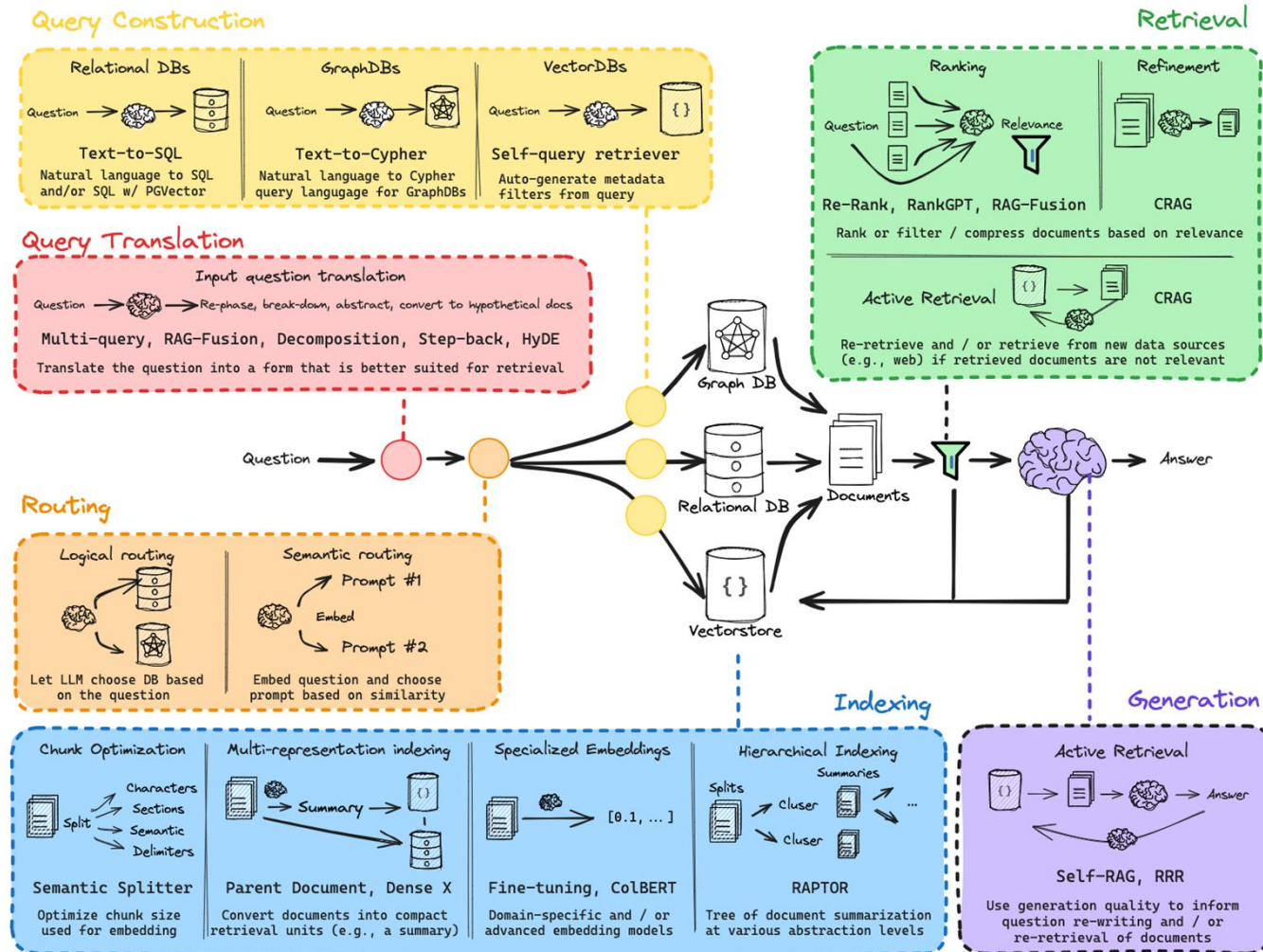
New Knowledge

Long-context(memory)

문서가 너무 길 때
새로운 정보가 중요할 때

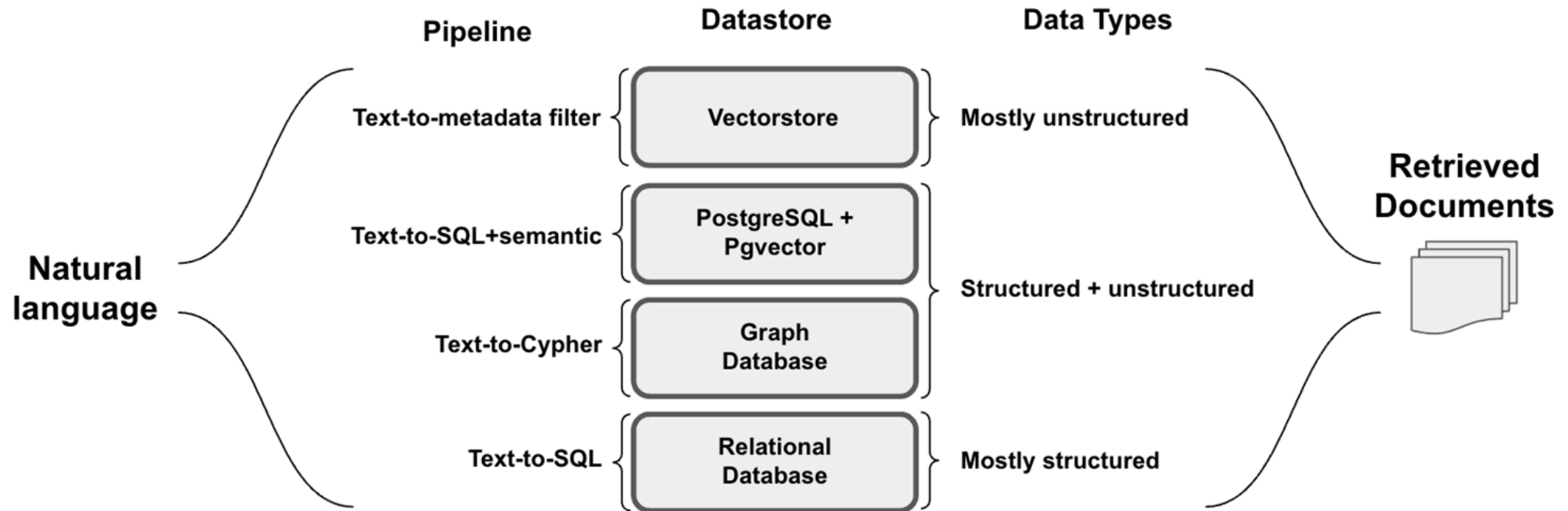
Retrieval Augmented Generation

□ 성공적인 RAG를 위한 다양한 전략들



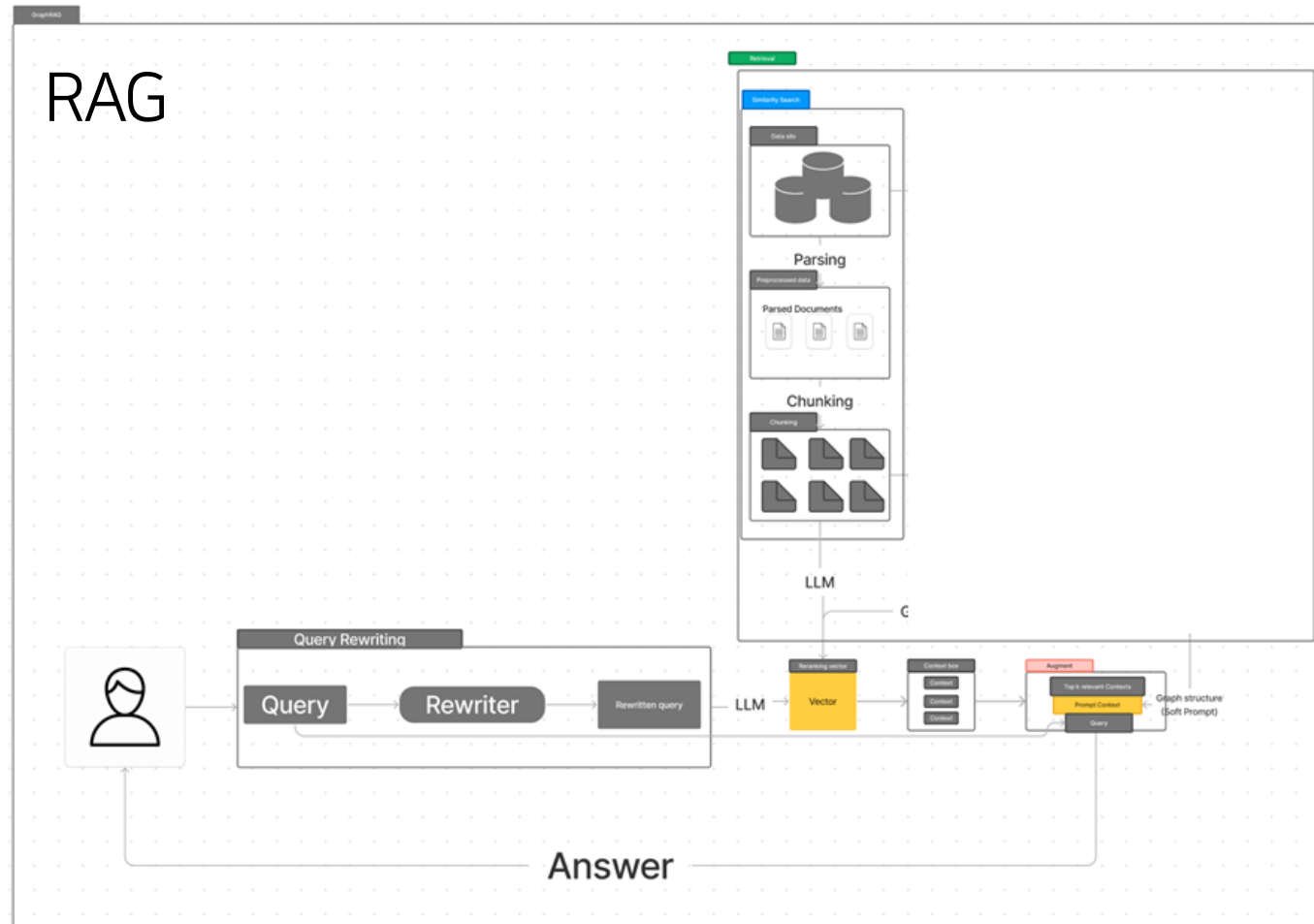
Retrieval Augmented Generation

□ 정형 / 비정형 데이터가 혼합된 상황



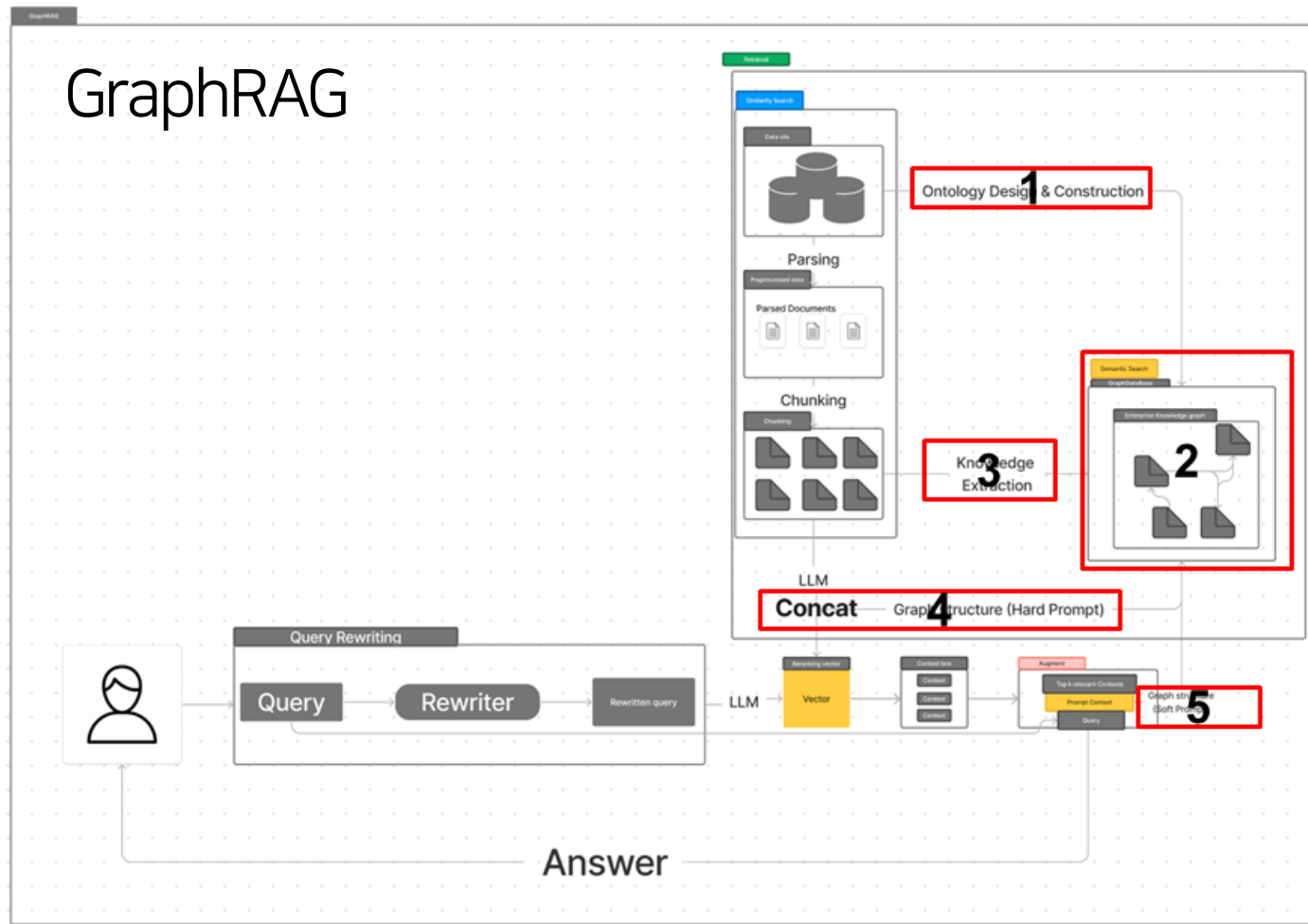
Retrieval Augmented Generation

❑ RAG to GraphRAG



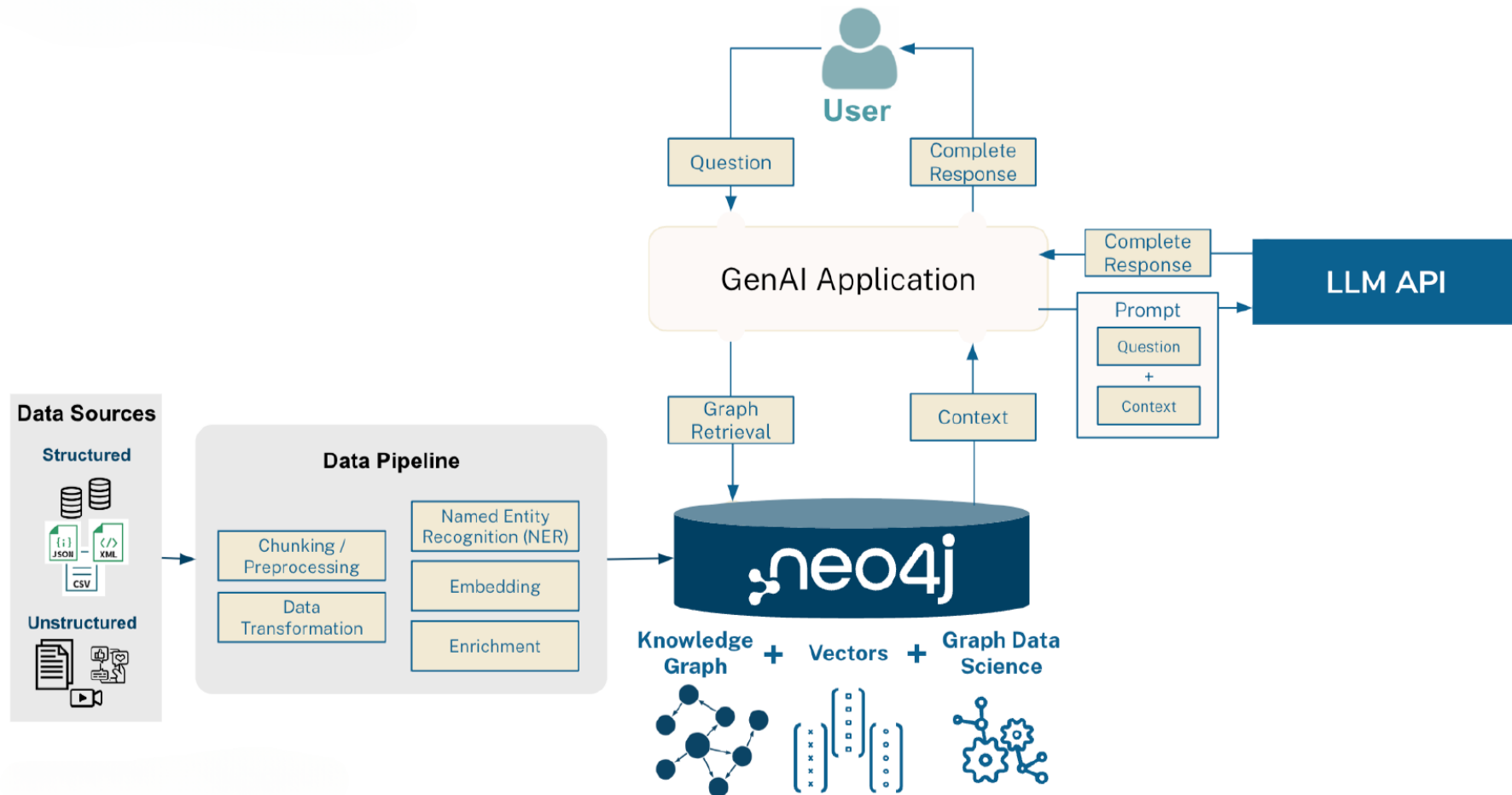
Graph-enhanced RAG from Documents

❑ RAG to GraphRAG



Graph-enhanced RAG from Documents

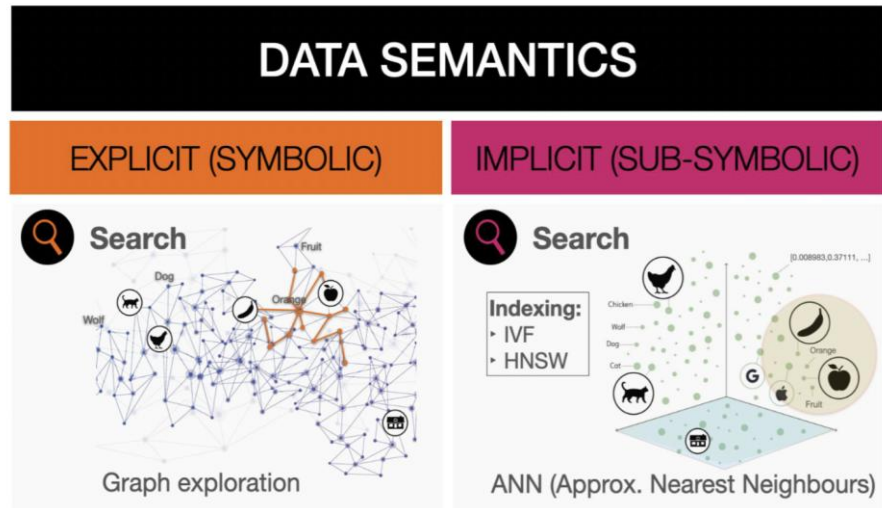
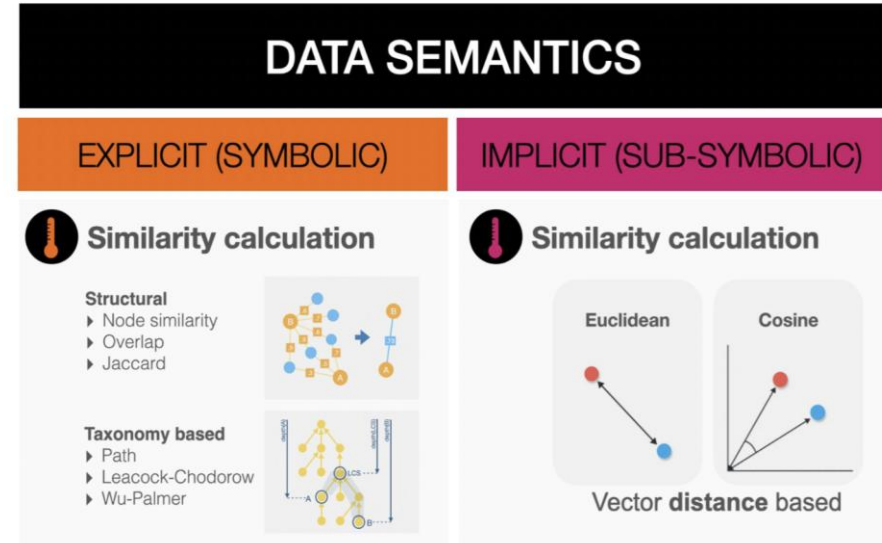
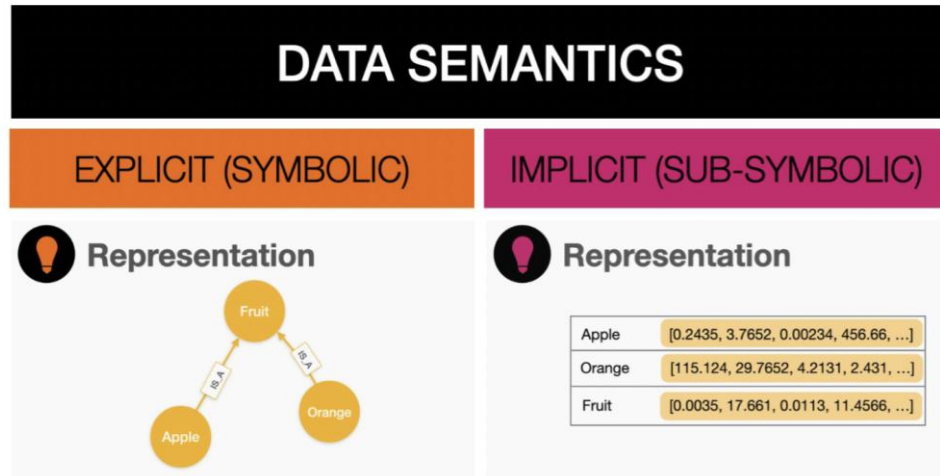
- ❑ KG-enhanced LLM retrieval augmented generation (GraphRAG).



Graph-enhanced RAG from Documents

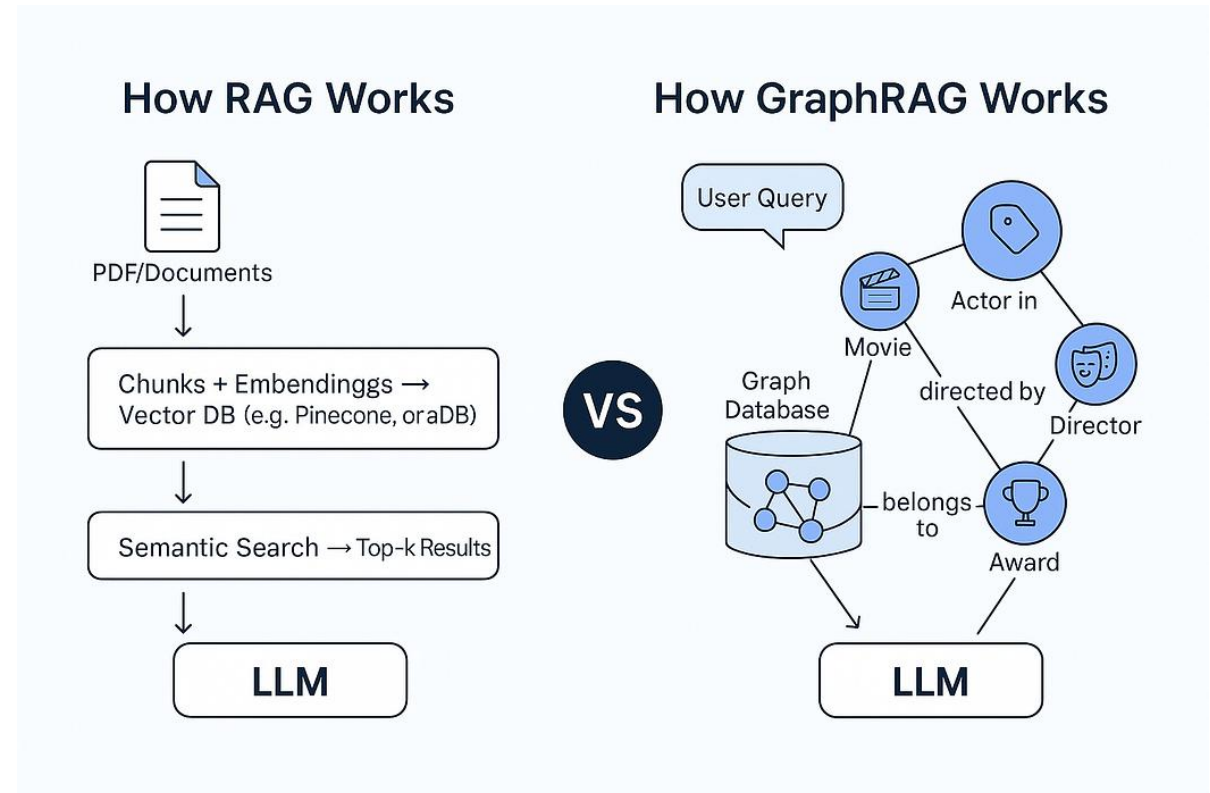
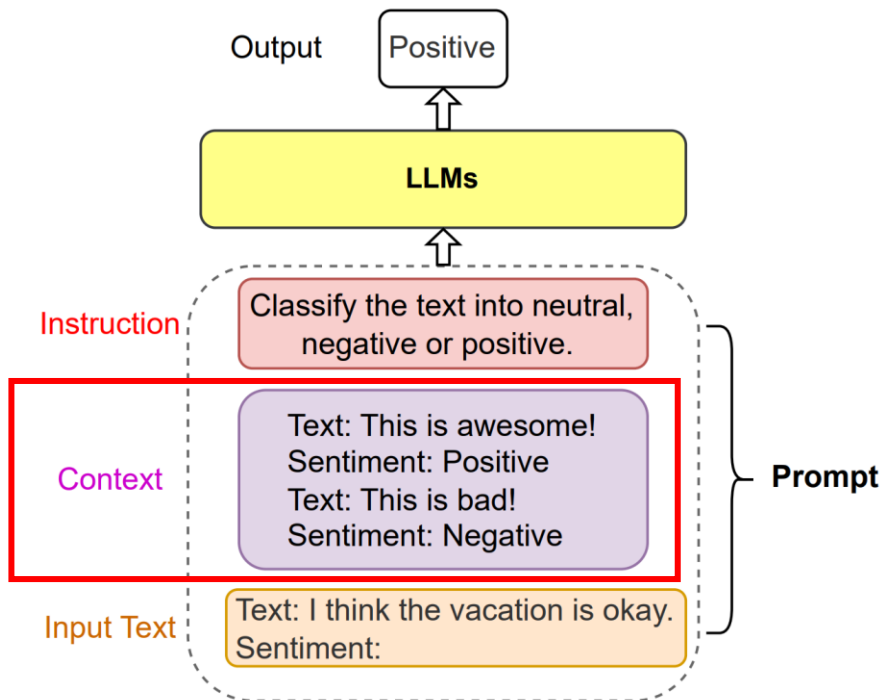
❑ RAG to GraphRAG 차이점

- ✓ 데이터 표현 방식
- ✓ 데이터 유사성 비교 및 연산
- ✓ 데이터 조회 방식



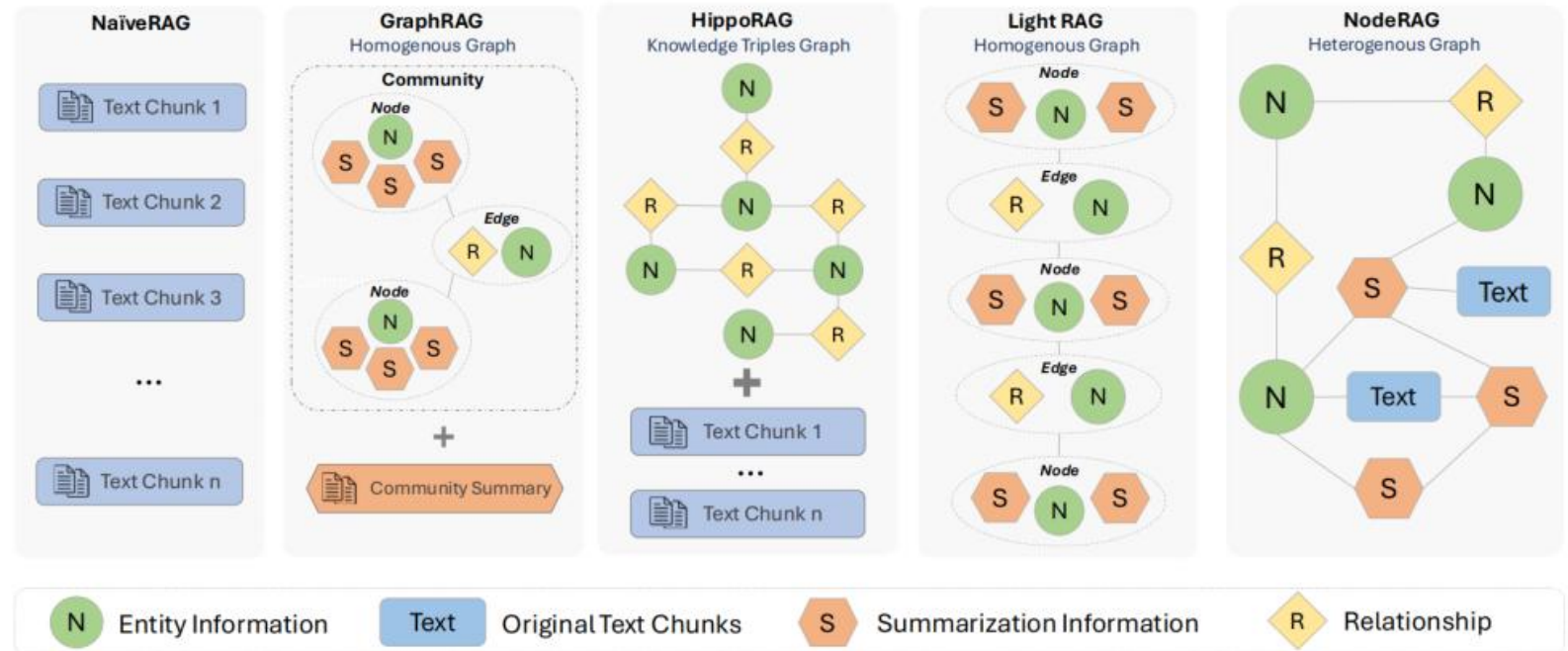
Graph-enhanced RAG from Documents

- ❑ 다양한 소스의 최신 지식을 실시간으로 검색·활용하여 생성모델의 환각을 줄이고 사실 검증 가능성을 높이는 것을 목표
 - ✓ 기존 방식: Query와 가장 유사한 chunk만을 개별적으로 검색하기 때문에 여러 문서에 흩어진 관련 정보들 사이의 연결성을 포착하지 못함



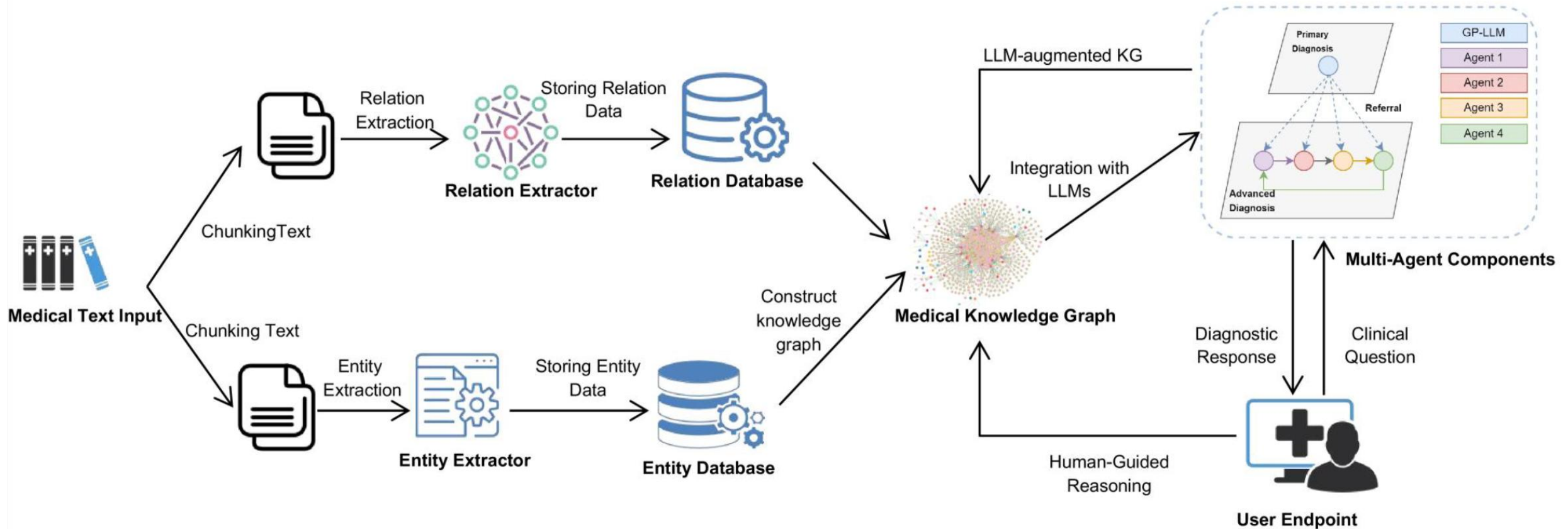
Graph-enhanced RAG from Documents

- ❑ 다양한 소스의 최신 지식을 실시간으로 검색·활용하여 생성모델의 환각을 줄이고 사실 검증 가능성을 높이는 것을 목표
 - ✓ 전역적 이해 (Global Understanding)
 - ✓ 다중 홉(Multi-hop) 추론
 - ✓ 맥락 연결
 - ✓ 중복 제거 및 일관성



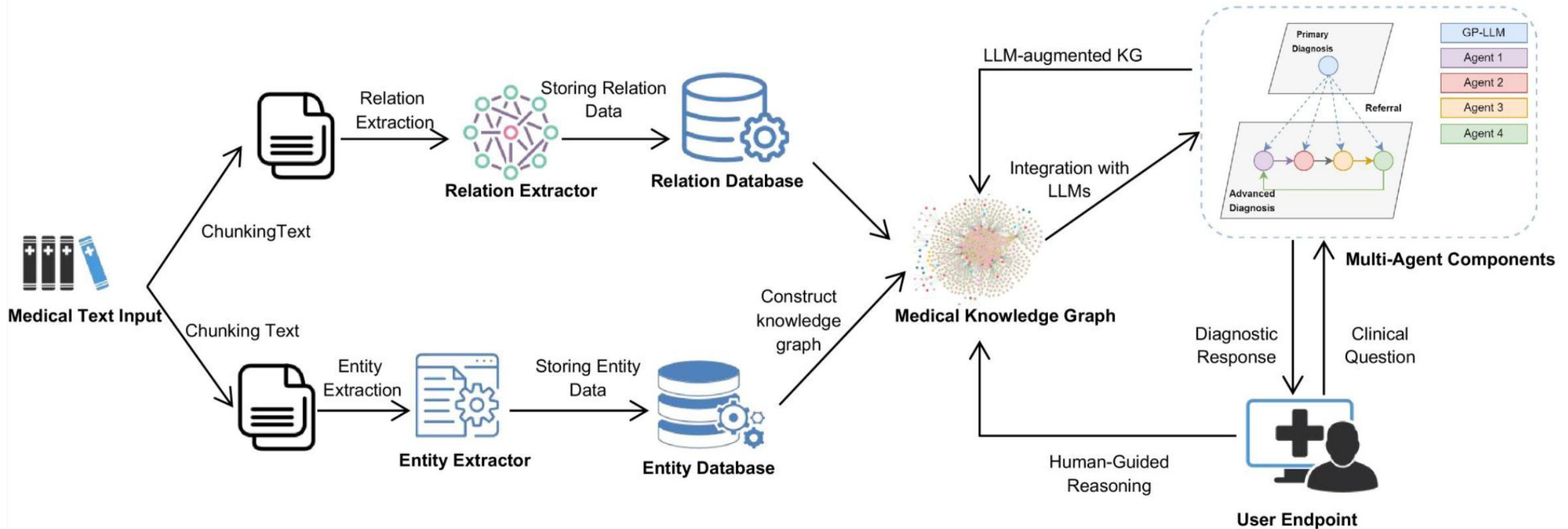
Graph-enhanced RAG from Documents

- ❑ KG+LLM for medical diagnosis.



Graph-enhanced RAG from Documents

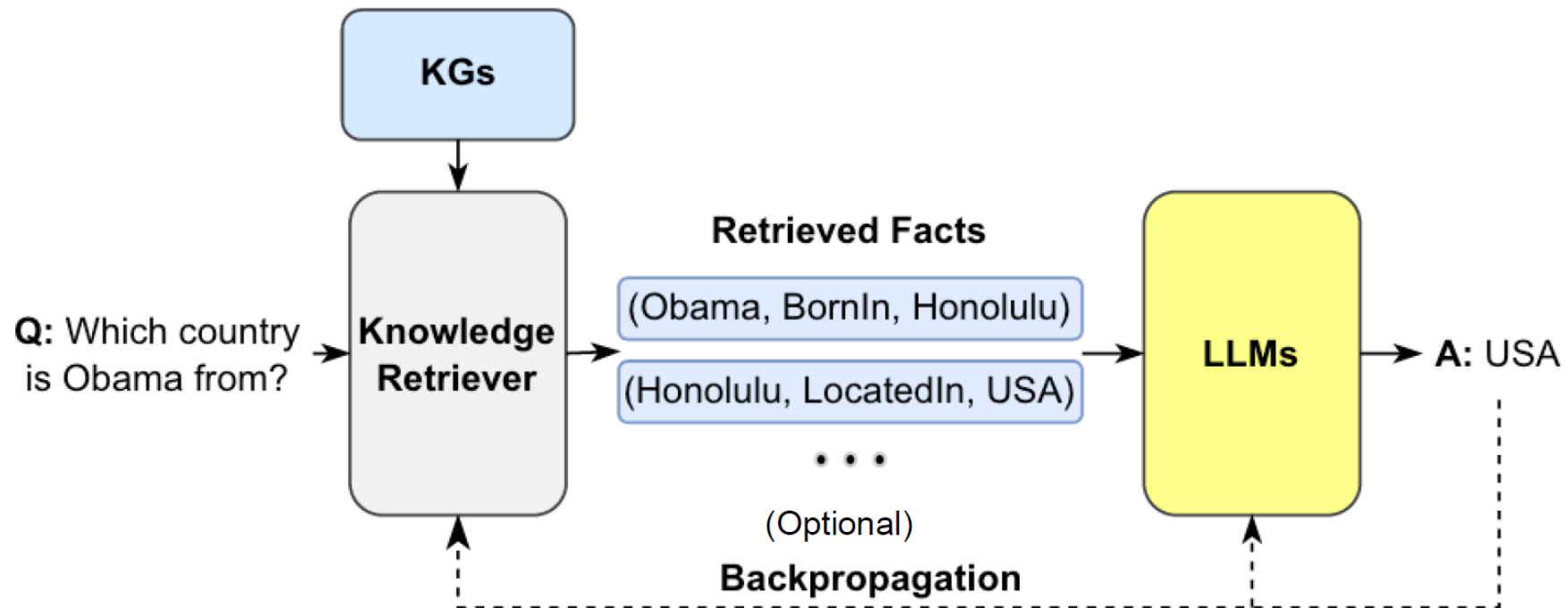
- ❑ KG+LLM for medical diagnosis.



KG-enhanced LLM Reasoning

- **Retrieval-augmented Knowledge Fusion**

- Retrieve-then-reasoning.
- Parameters-free.
- Can be applied to closed-source LLMs (e.g., ChatGPT).
- Widely used in **applications**.



KG-enhanced LLM Reasoning

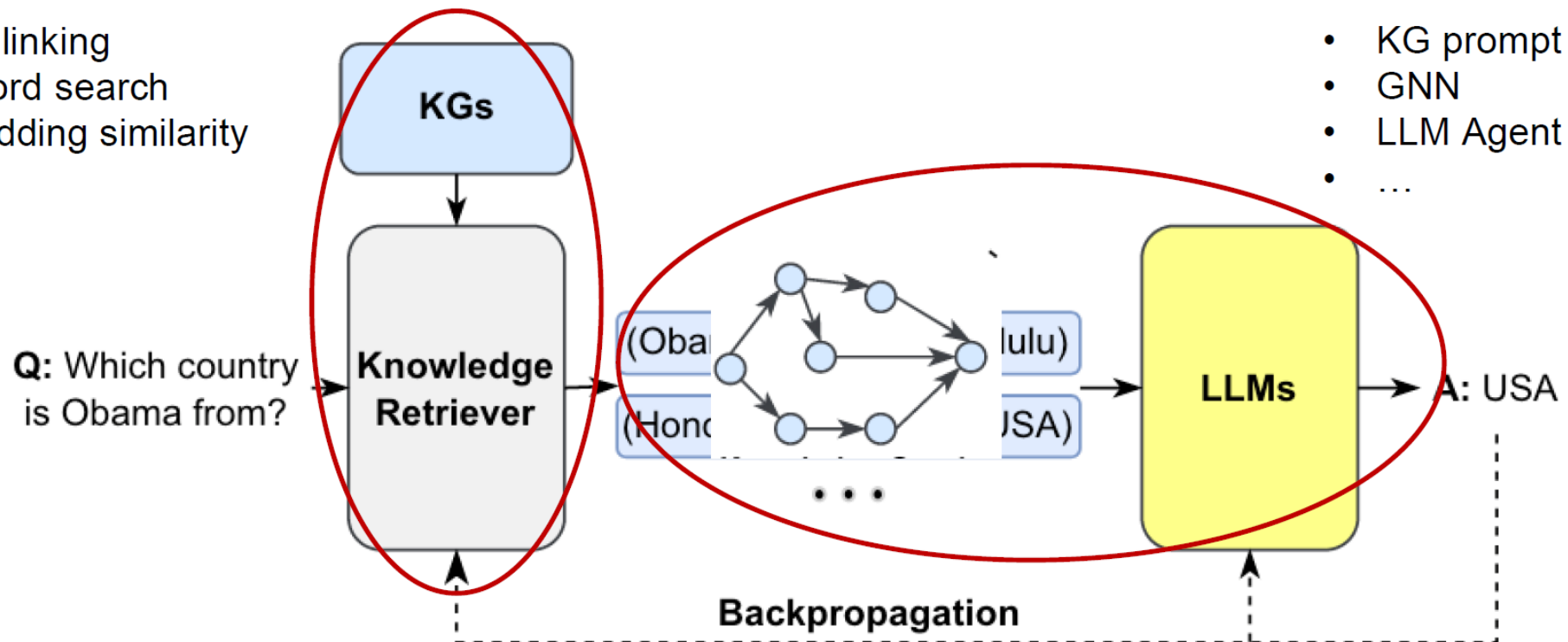
- **Retrieval-augmented Knowledge Fusion**
 - Techniques and challenges.

How to effectively retrieve on KGs?

- Entity linking
- Keyword search
- Embedding similarity
- ...

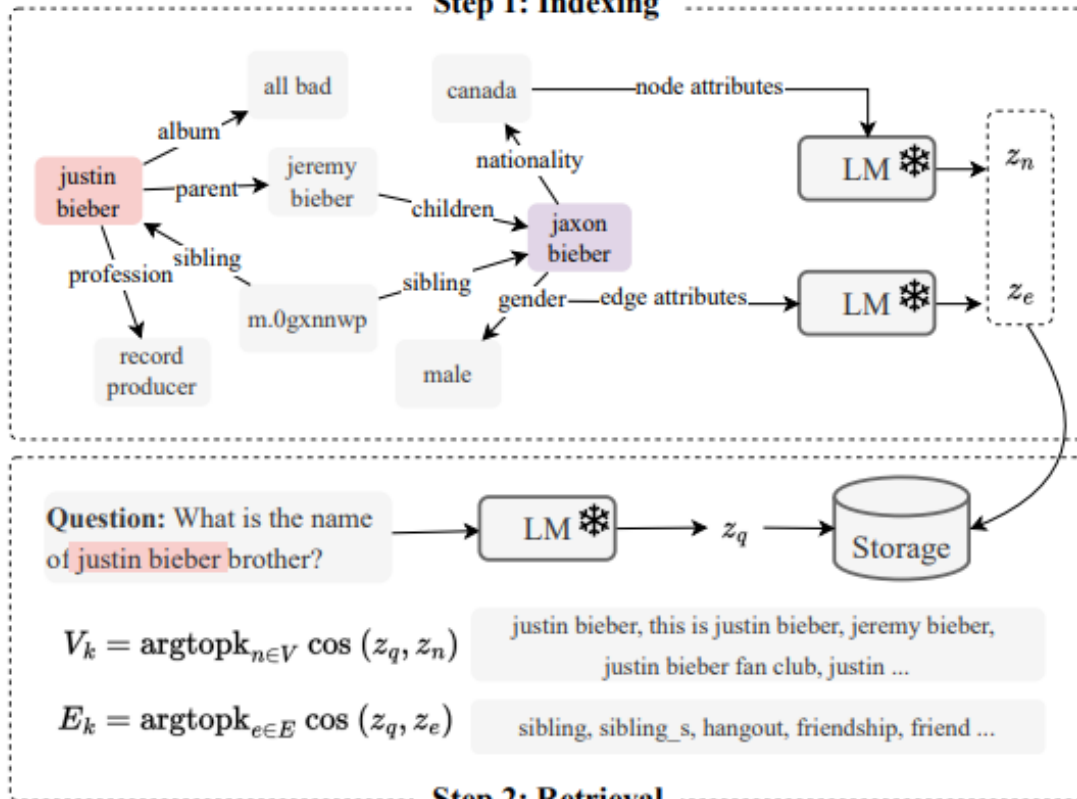
How to reason on retrieved KG structure?

- KG prompt
- GNN
- LLM Agent
- ...

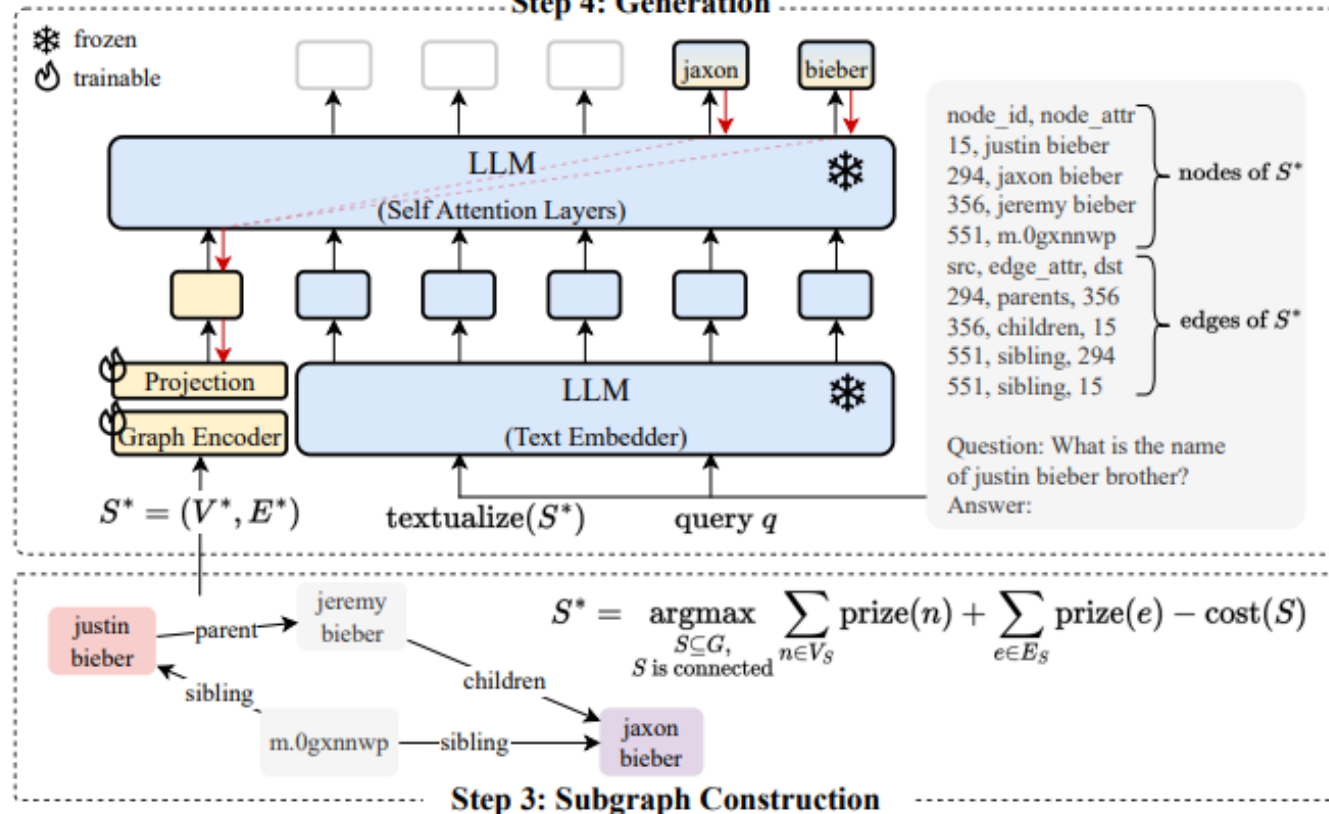


KG-enhanced LLM Reasoning – G-Retriever

Step 1: Indexing



Step 4: Generation

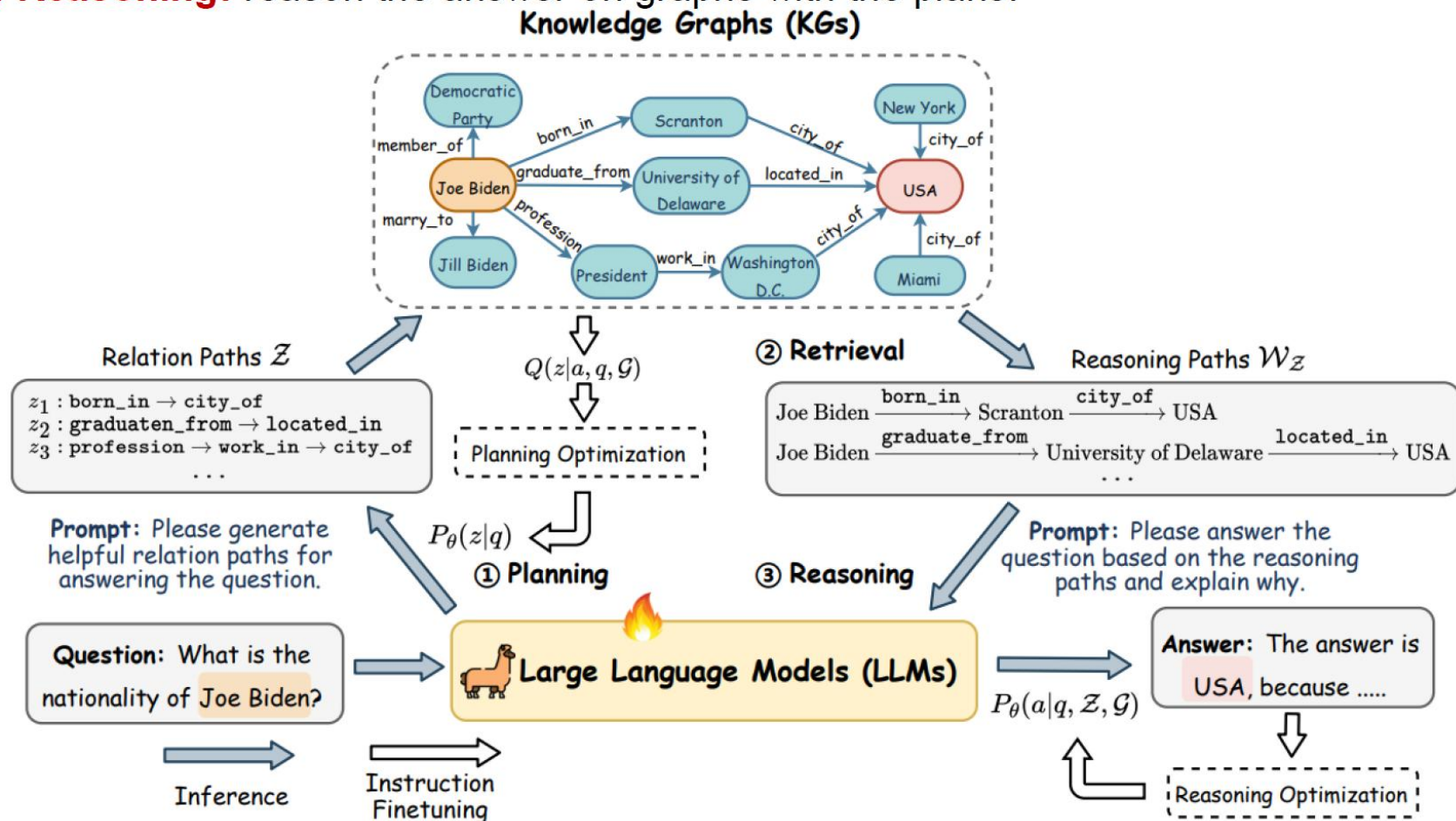


KG-enhanced LLM Reasoning – RoG

- **Planning-retrieval-reasoning.**

Planning: generate faithful relation paths as plans.

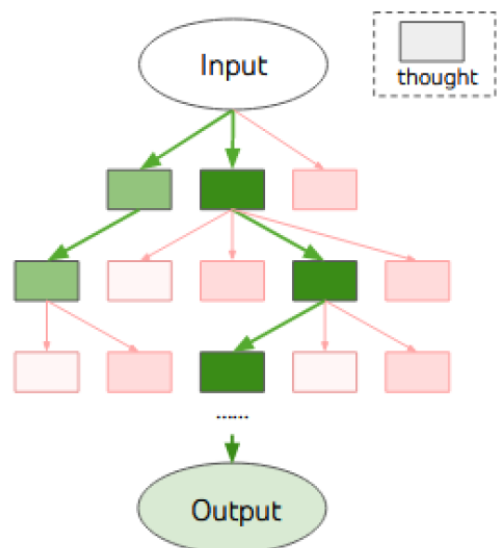
Retrieval-Reasoning: reason the answer on graphs with the plans.



KG-enhanced LLM Reasoning – GCR

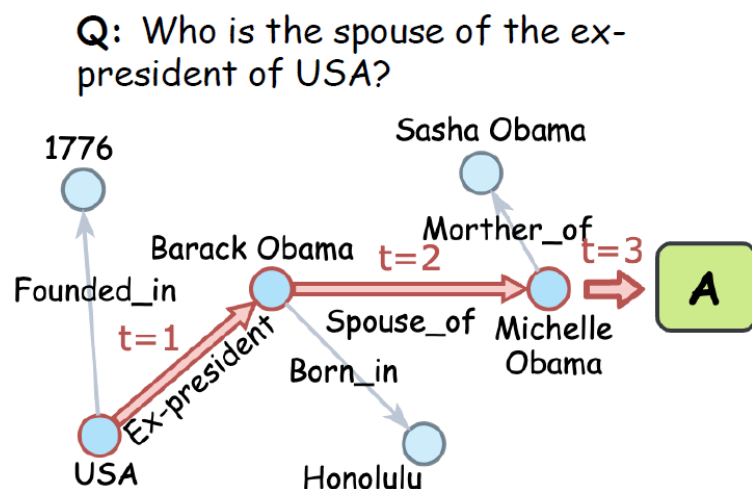
- **Graph-constrained Reasoning (GCR):**

- Incorporates KGs into the decoding process of LLMs to achieve KG-grounded faithful reasoning (**decoding on graphs**)

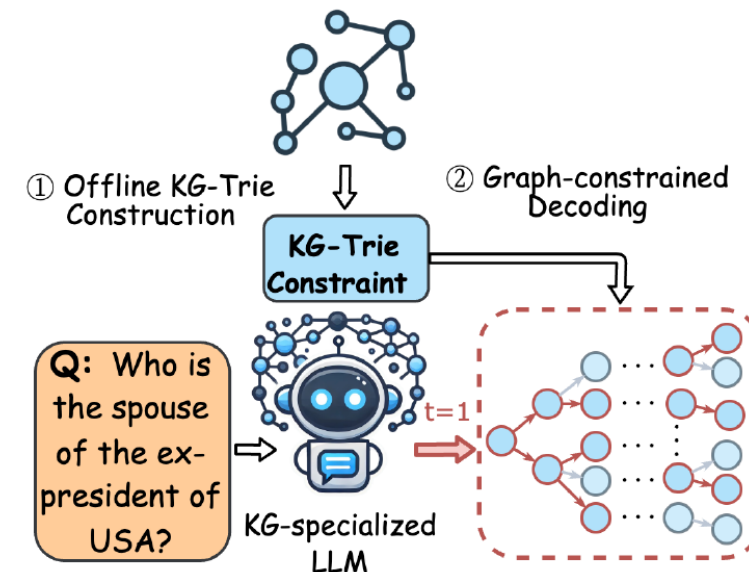


(d) Tree of Thoughts (ToT)

LLM reasoning



Graph Reasoning



Graph-constrained Reasoning

CONTENTS

1. Today
2. Background
3. Large Language Models
4. Knowledge Graphs
5. Graph-enhanced Training
6. Graph-enhanced Reasoning
- 7. LLM-based Recommendation**

Language Modeling vs. User Behavior Modeling

Language Modeling

- Dense world knowledge
- Text tokens (Ten thousands level)

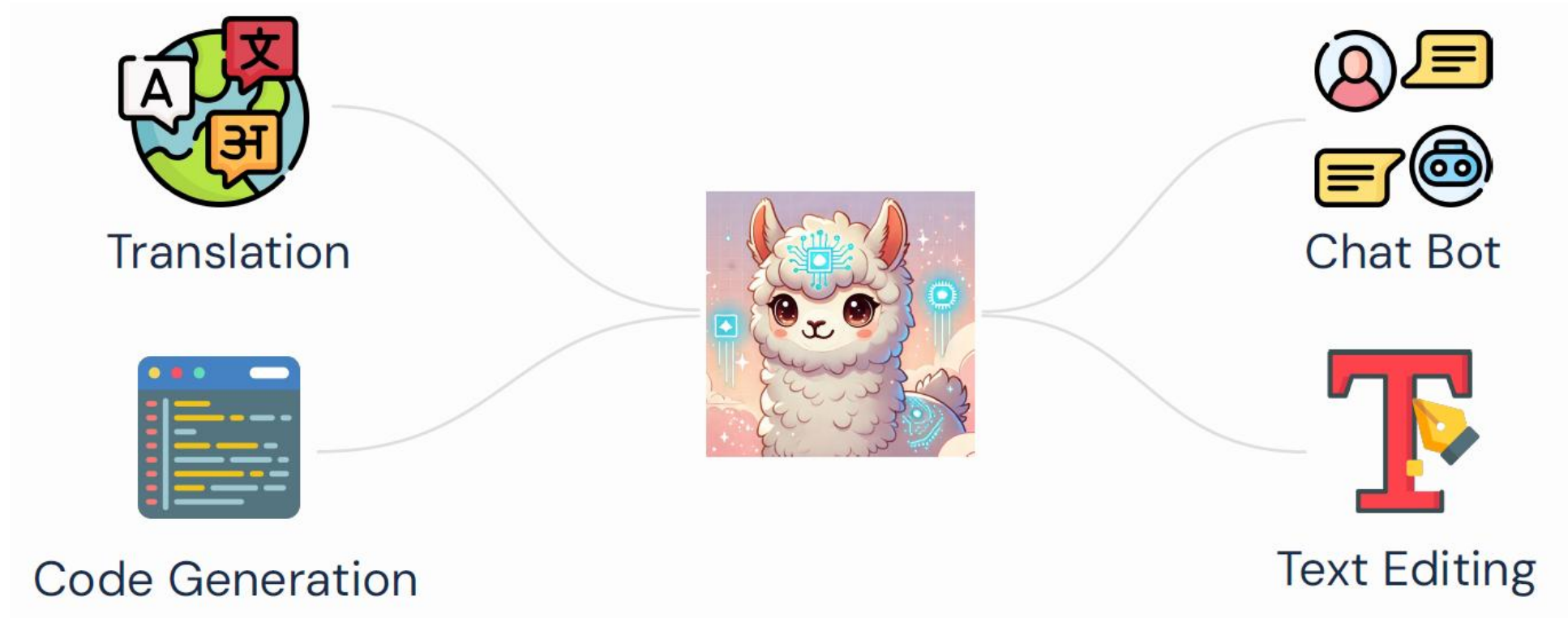


User Behavior Modeling

- Sparse user-item interactions
- Items (Billion to trillion level)

Large Language Models

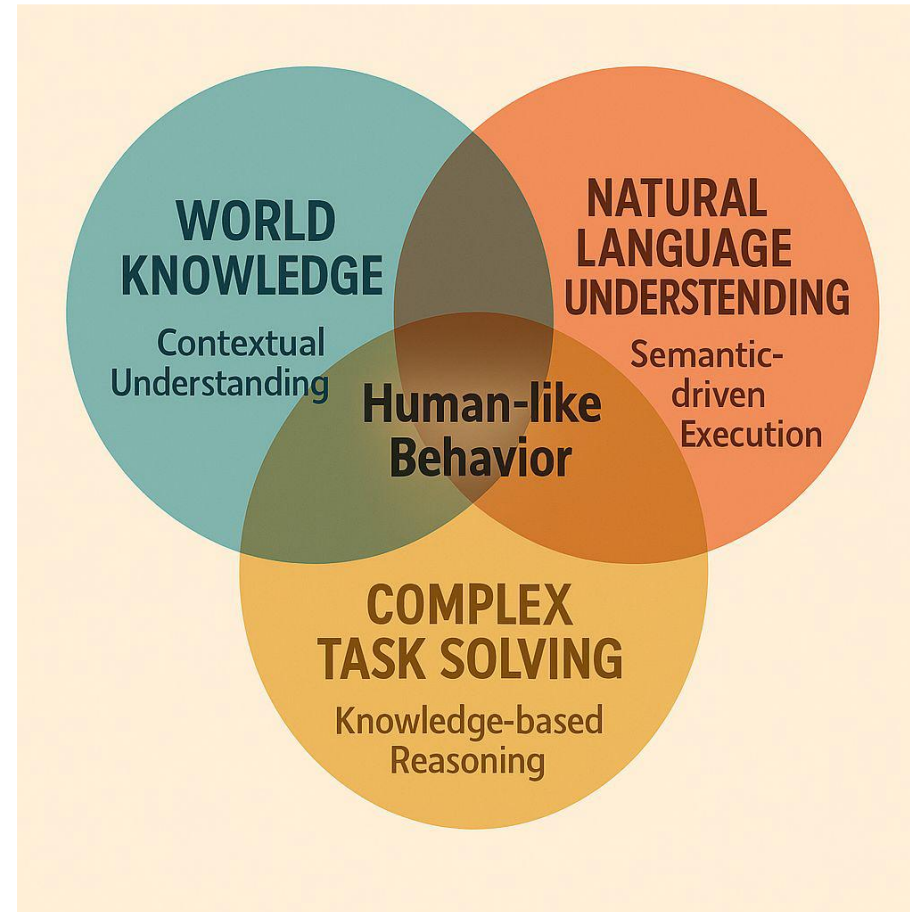
- ❑ LLMs are machine learning models that can perform a variety of natural language processing (NLP) tasks



Benefits of LLMs for Recommendation

❑ Key features of LLMs:

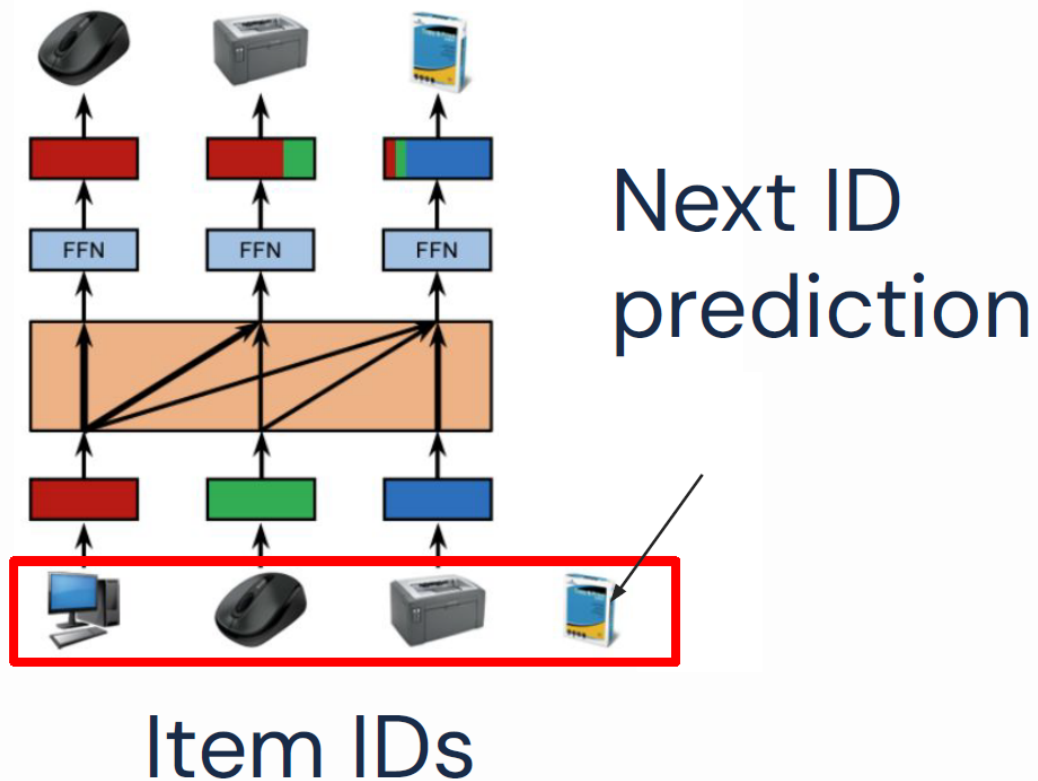
- ✓ World knowledge.
- ✓ Natural language understanding.
- ✓ Human-like behavior.



How can these features benefit recommender systems?

Benefits of LLMs for Recommendation

❑ (1) World knowledge

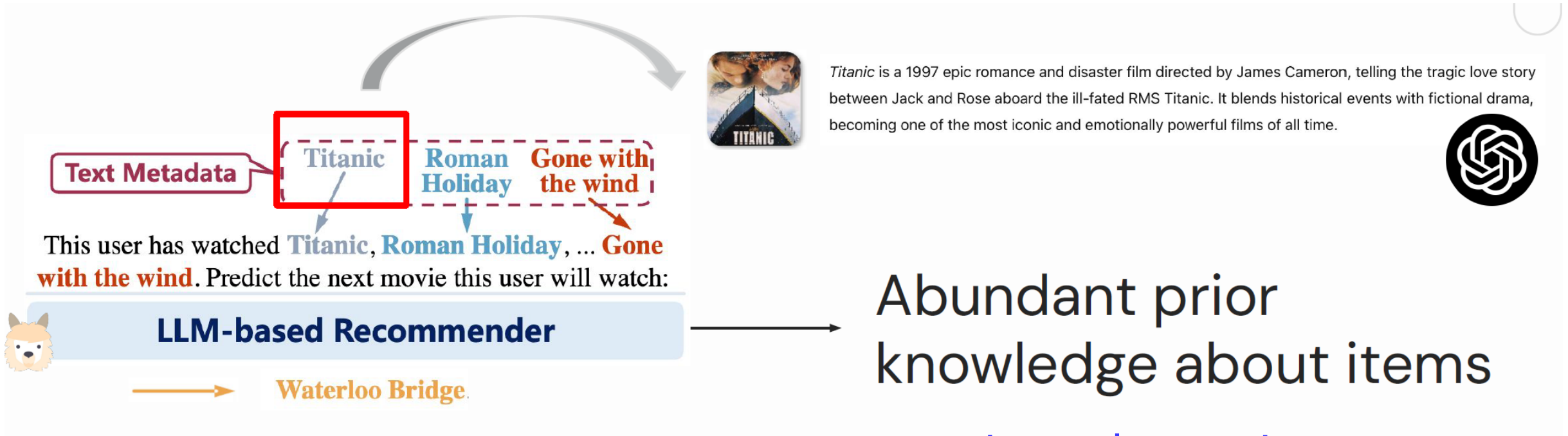


ID-based item modeling
lack semantic meanings

Example: SASRec [*ICDM'18*]

Benefits of LLMs for Recommendation

❑ (1) World knowledge



LLM as sequential recommender

Abundant prior knowledge about items

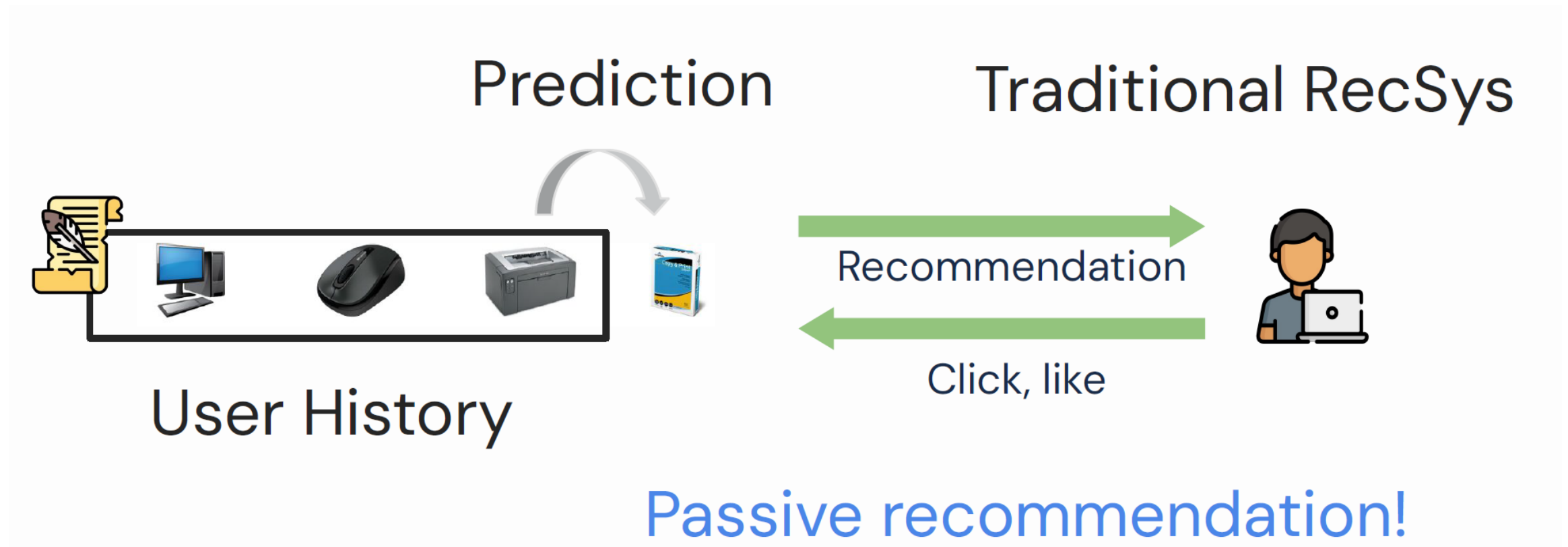
Lower data requirement

Cross-domain ability

Cold-start ability

Benefits of LLMs for Recommendation

- ❑ (2) Natural language understanding.



Benefits of LLMs for Recommendation

- ❑ (2) Natural language understanding.



*I would like to
recommend...*

Recommendation

Click, like

Conversation



*Some scientific
movies.*

LLM as conversational recommender

Interactive

User-friendly

More accurate



hello I'm open to any movie

Hi there. I would like to suggest some *comedies* you could watch, have you seen *The Wedding Singer (1998)*?

I have not seen it but I watched *American Pie 2 (2001)*. I just watched *Avengers: Infinity War (2018)* and I liked it.



Benefits of LLMs for Recommendation

❑ (3) Human-like behavior.



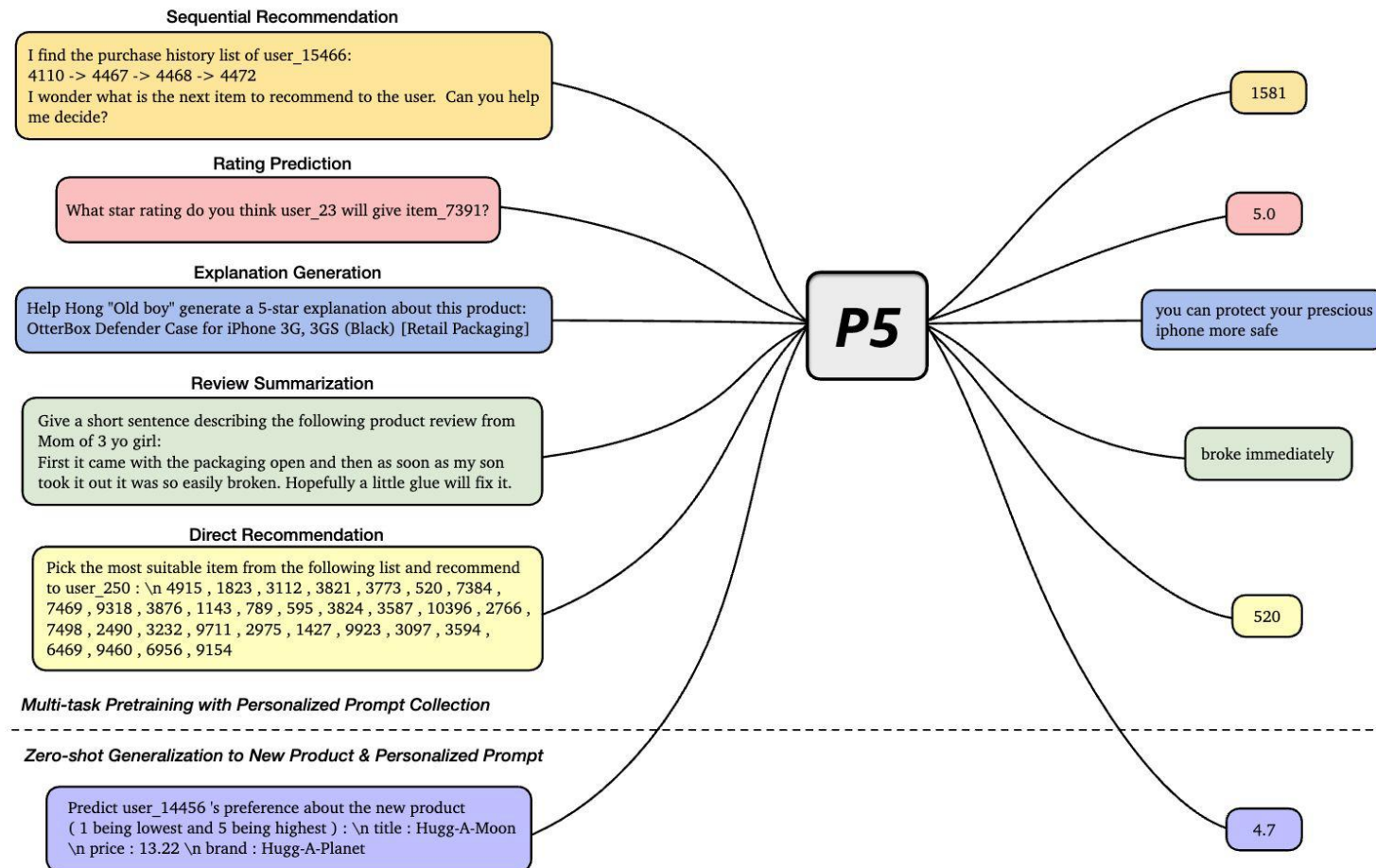
Generative Agents can
(mostly) simulate human
behaviors

- Cooperation
- Organization

Simulating user behaviors for evaluating recommenders.

LLM as Sequential Recommender

❑ Multi-task alignment (P5) -> general recommender



LLM as Sequential Recommender

- Training on different task prompts -> multiple recommendation abilities.

