

Lecture#1: Introduction to Knowledge Graphs and their Real-World Applications

Joyce Jiyoung Whang

Department of AI Computing, KAIST

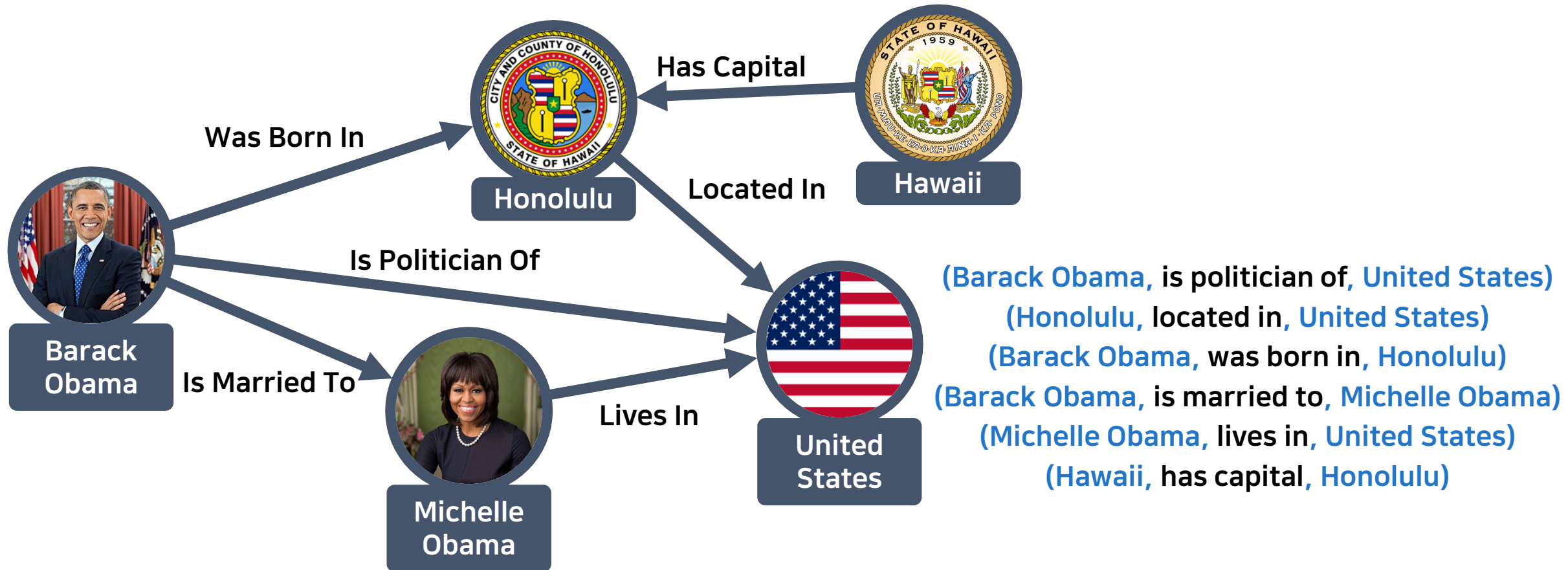
Knowledge Discovery with Knowledge Graphs: From Structural Embeddings to Generative Reasoning (KnowKG), KDD 2026 Tutorial

<https://bdi-lab.kaist.ac.kr>



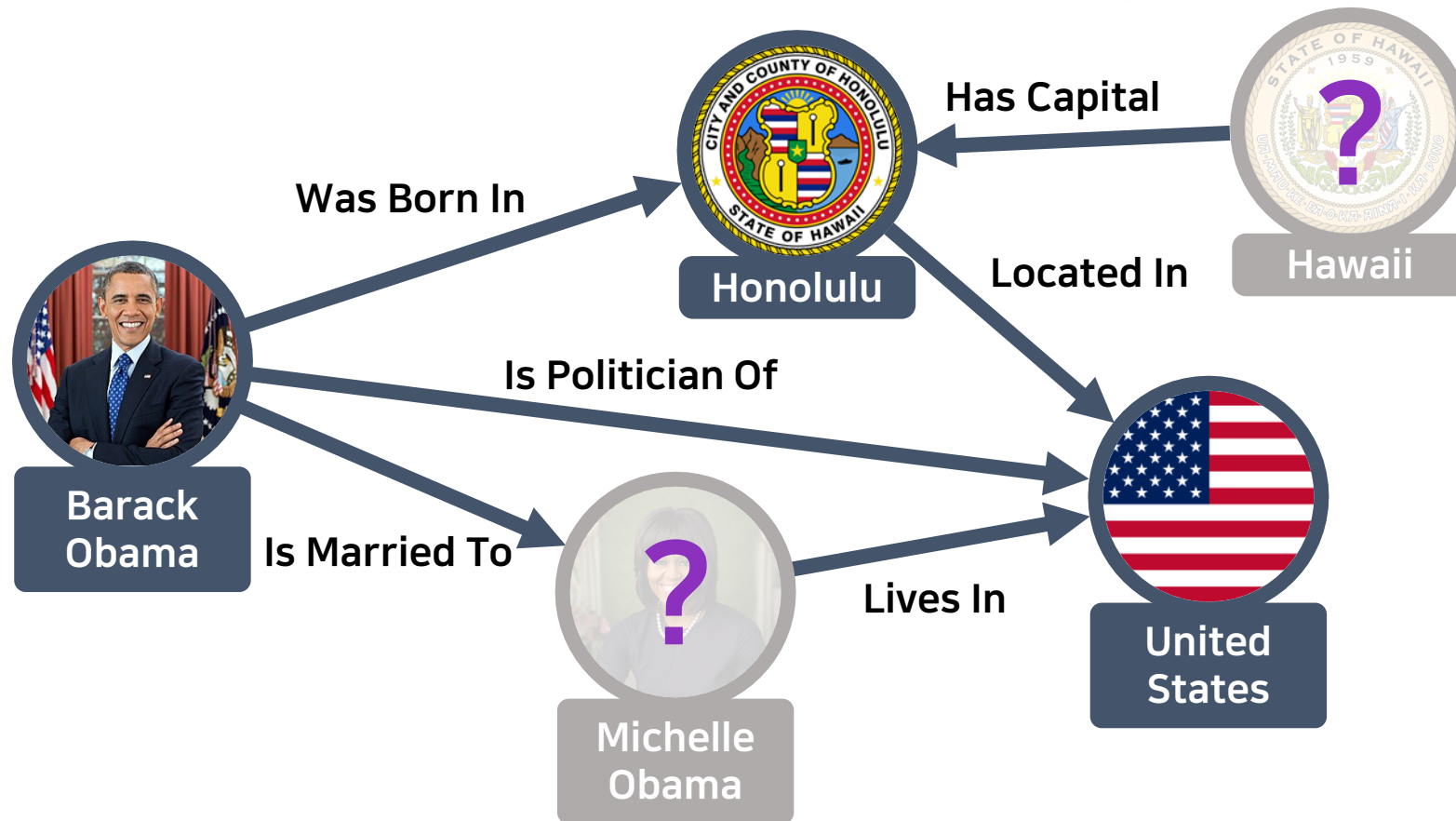
Knowledge Graphs

- Graphical Representation of Human Knowledge
 - Each fact is represented by a triplet (head entity, relation, tail entity)



01 Knowledge Graph Completion

- Given a head entity and a relation, predict the tail entity.
- Given a tail entity and a relation, predict the head entity.



01 Examples of Knowledge Graphs

- Google's search engine
 - An infobox next to the search results is created based on knowledge graphs

The image shows a Google search results page for the query 'kaist'. The search bar at the top shows 'kaist' with a search button. Below the search bar, there are tabs for 'All', 'Images', 'News', 'Maps', 'Videos', and 'More'. The search results show 'About 8,490,000 results (0.82 seconds)'. The first result is 'https://www.kaist.ac.kr' with the title 'KAIST'. Below the title, it says 'No information is available for this page.' and 'Learn why'. It also says 'You've visited this page many times. Last visit: 5/17/21'. There is a search bar for 'Results from kaist.ac.kr'. Below this, there are two columns of links: 'Korea Advanced Institute of ...' and 'Computer Science'. The 'People also ask' section has four questions: 'Is kaist a good university?', 'What is kaist known for?', 'Does kaist teach in English?', and 'Is kaist good for international students?'. At the bottom, there is a link to 'KAIST - Wikipedia' with the text 'KAIST (formally the Korea Advanced Institute of Science and Technology) is a national'. On the right side, there is a large blue-bordered box labeled 'Infobox' with an arrow pointing to it. The infobox contains a photo of the KAIST campus, a map, and text in Korean and English. The text includes the name 'KAIST 한국과학기술원 대덕캠퍼스', the address '291 Daehak-ro, Eoeun-dong, Yuseong-gu, Daejeon', and other details like 'Hours', 'Phone', 'Academic staff', 'Total enrollment', 'Postgraduates', and 'Founded'.

Google

kaist

About 8,490,000 results (0.82 seconds)

https://www.kaist.ac.kr

KAIST

No information is available for this page.
[Learn why](#)

You've visited this page many times. Last visit: 5/17/21

Results from kaist.ac.kr

Korea Advanced Institute of ...
KAIST Seals. KAIST stands for the
Korea Advanced Institute of ...

Computer Science
Faculty - Seminar Room -
Courses - Research Labs -
Education - ...

People also ask

Is kaist a good university?

What is kaist known for?

Does kaist teach in English?

Is kaist good for international students?

Feedback

https://en.wikipedia.org › wiki › KAIST

KAIST - Wikipedia
KAIST (formally the Korea Advanced Institute of Science and Technology) is a national

KAIST 한국과학기술원 대덕캠퍼스

See photos

Map data ©2021 SK Telecom

Website Directions Save Call

University in Daejeon

KAIST is a national research university located in Daedeok Innopolis, Daejeon, South Korea. KAIST was established by the Korean government in 1971 as the nation's first public, research-oriented science and engineering institution. KAIST is generally known as the most prestigious university in South Korea. [Wikipedia](#)

Address: 291 Daehak-ro, Eoeun-dong, Yuseong-gu, Daejeon

Hours: Open · Closes 6PM

Phone: 042-350-2114

Academic staff: 639 (2019)

Total enrollment: 10,249 (2013)

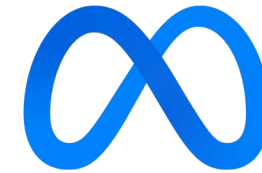
Postgraduates: 6,738

Founded: February 16, 1971

← Infobox

Examples of Knowledge Graphs

- **Google's** search engine
 - Knowledge graphs are embedded in the search engine
- **Microsoft's** knowledge mining API
 - Used for the Bing search engine, QnA pair mining, processing LinkedIn data
- **Meta's** heterogeneous graphs
 - Analyze connections between people, events, ideas, and news
- **Amazon's** product networks
 - Utilize relationships between users, products, and their metadata
- **IBM's** knowledge graphs
 - Provide a framework to develop internal knowledge graphs
- **Palantir's** ontology
 - Build business-level ontology for real-world organization management



01 Examples of Knowledge Graphs in NVIDIA

- Use cases of knowledge graphs across various domains 
- Healthcare: Enable advanced decision making by mapping medical knowledge into graphs
- Recommender Systems: Deliver personalized experiences by linking preferences with content
- Finance: Detect frauds by analyzing transaction graphs and identifying hidden relationships

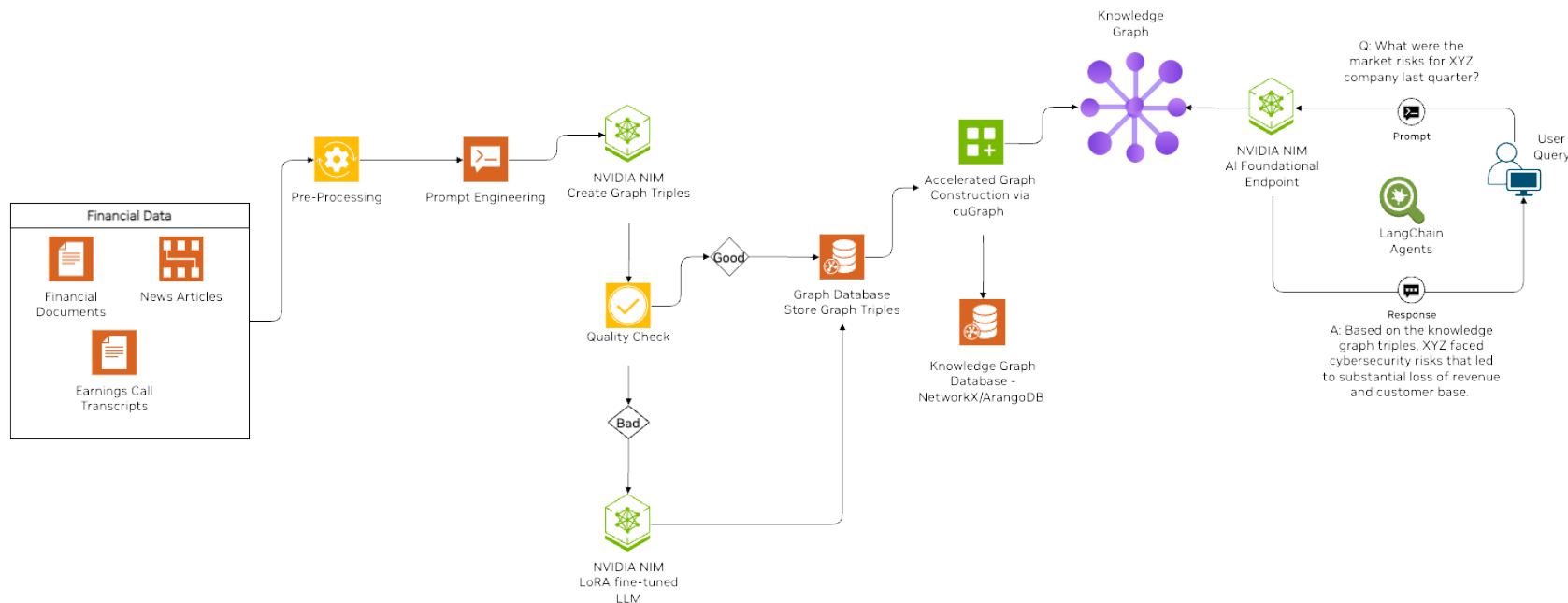


Image: <https://developer.nvidia.com/blog/insights-techniques-and-evaluation-for-llm-driven-knowledge-graphs/>

01 Examples of Knowledge Graphs in SAMSUNG

- Samsung Smartphones
 - Provide personalized user experiences using the Personal Data Engine powered by knowledge graphs

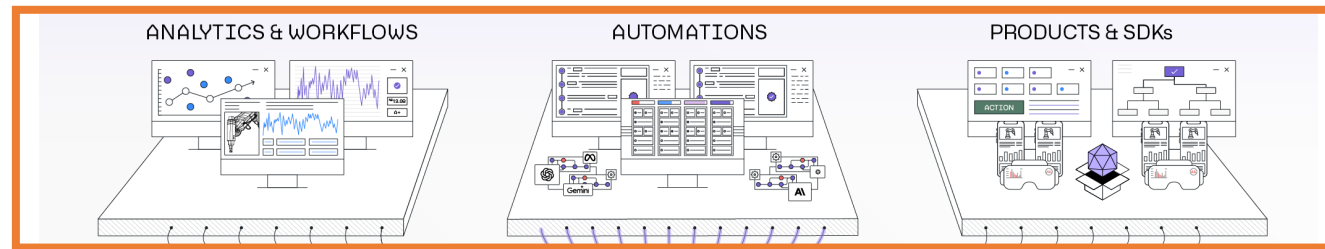


Image: <https://www.oxfordsemantic.tech/blog>

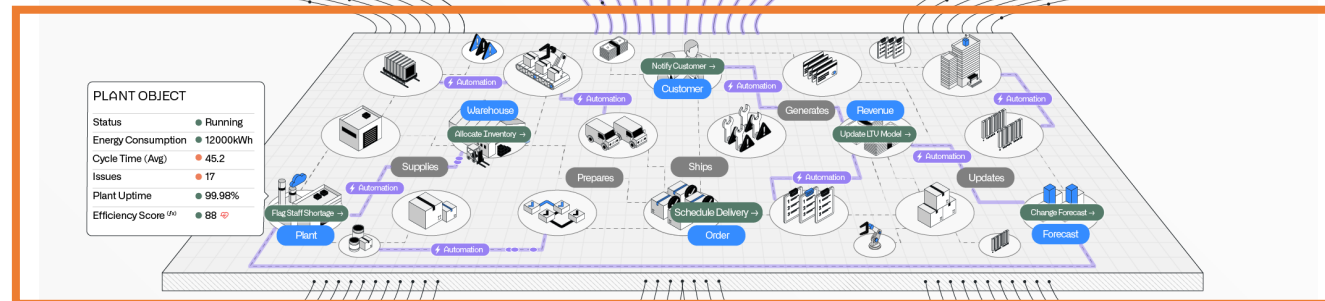
01 Examples of Knowledge Graphs in Palantir

- Construct ontology that integrates the data, logic, action, and security of the enterprise

Ontology Usage



Ontology



Information encoded in the Ontology

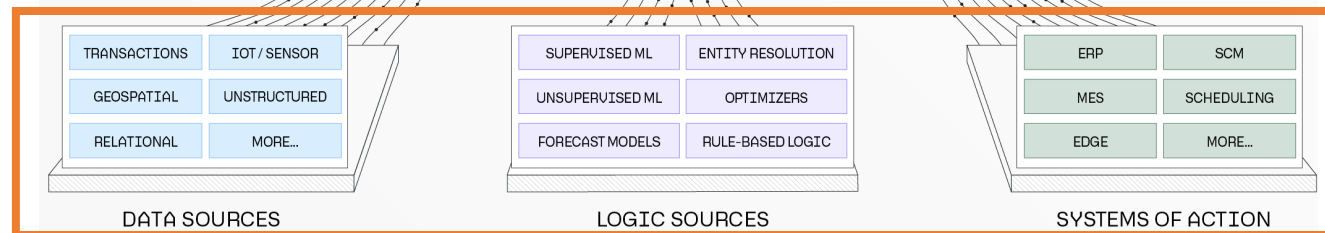


Image: <https://www.palantir.com/docs/foundry/architecture-center/ontology-system>

01 Examples of Knowledge Graphs in Palantir

- Palantir Ontology
 - Designed to represent the complex, interconnected decisions of an enterprise

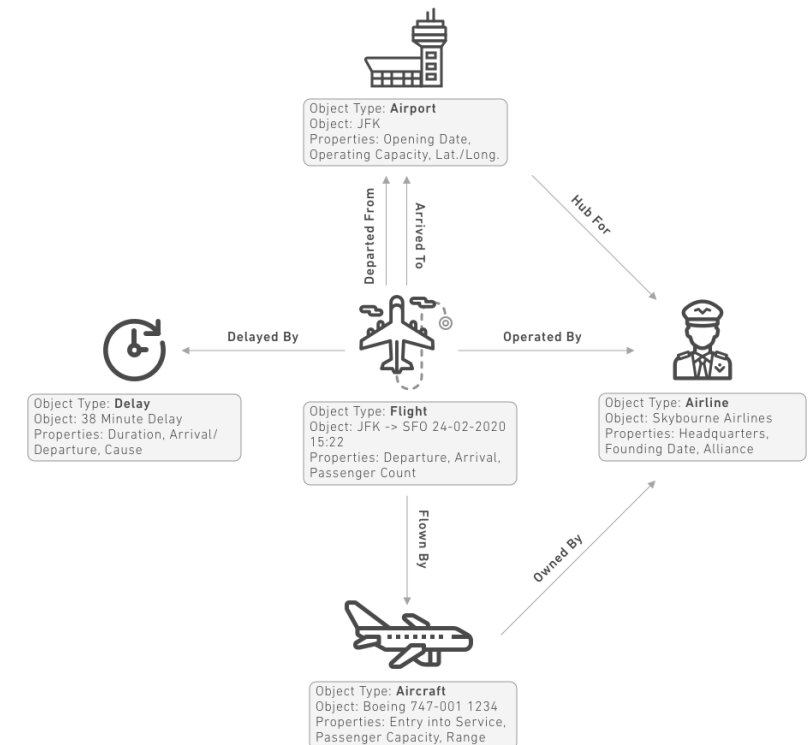
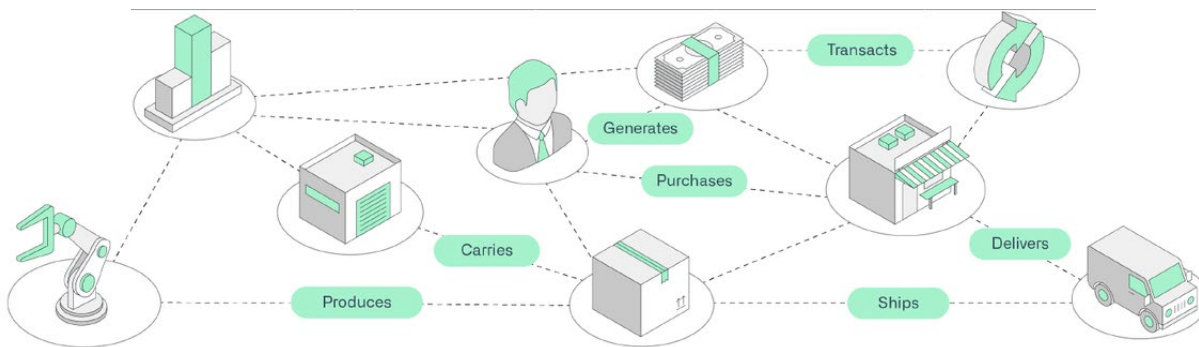
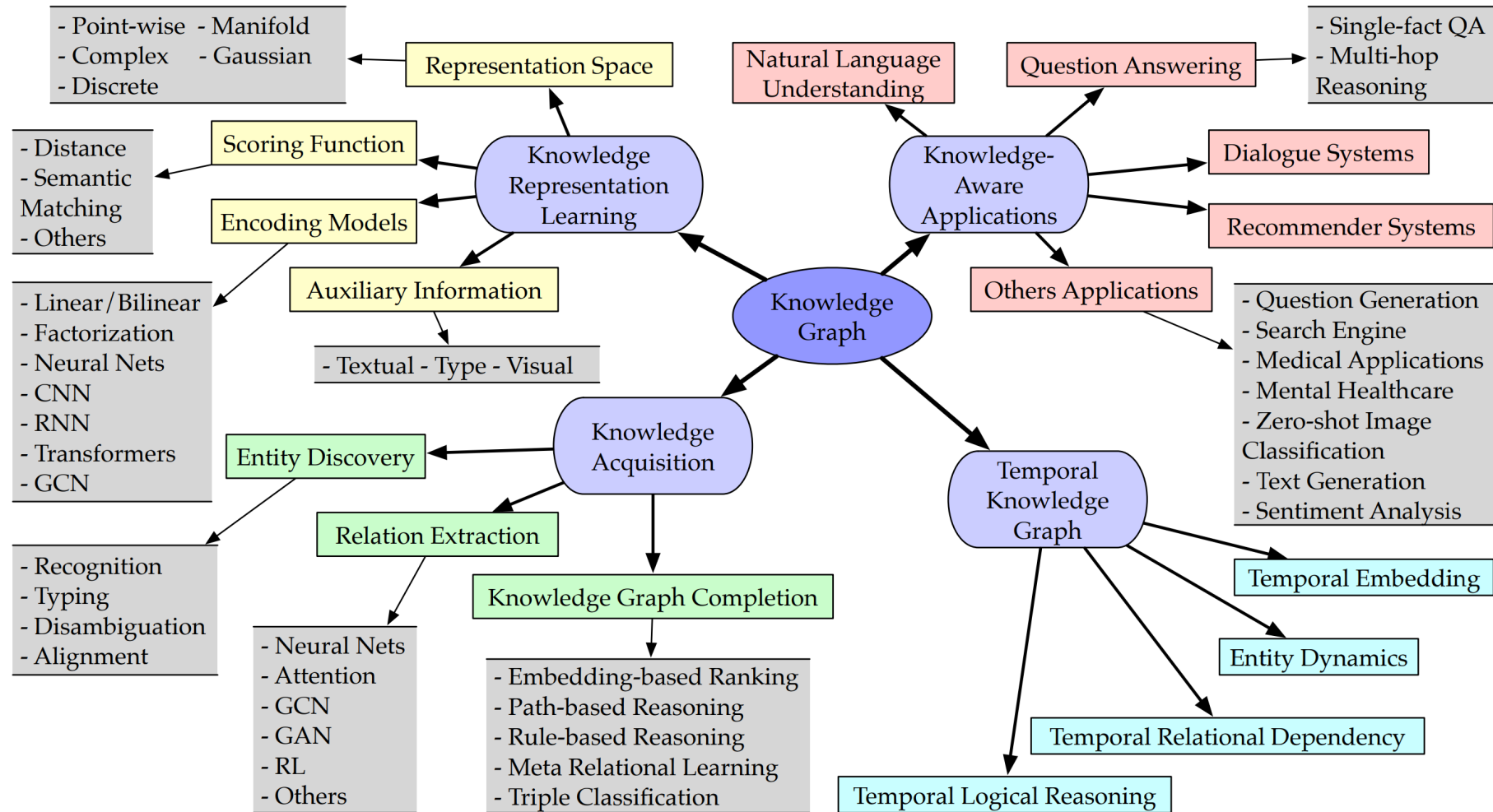


Image: <https://www.palantir.com/docs/foundry/ontology/core-concepts/?ref=community.heartcount.io>

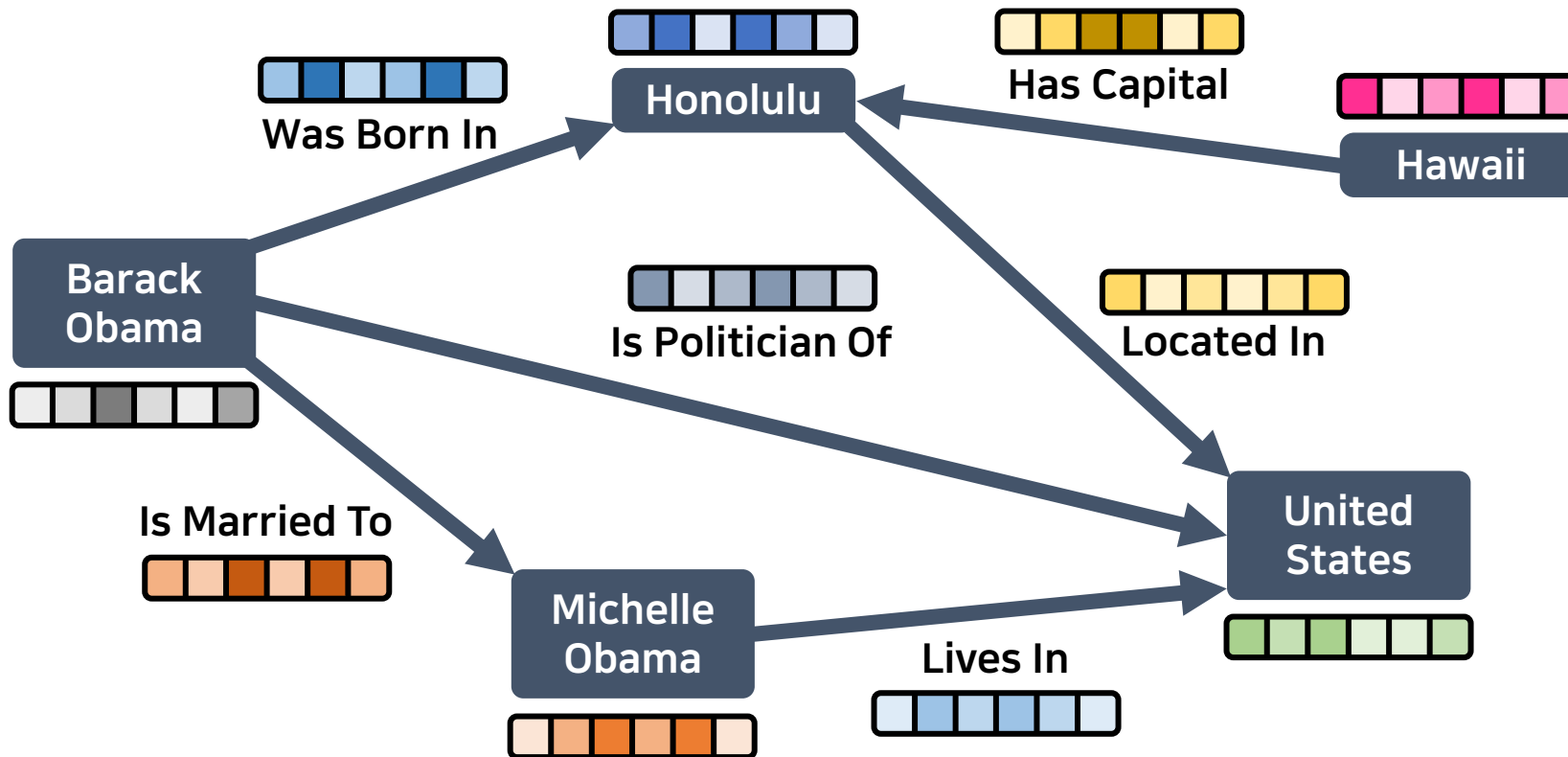
Research on Knowledge Graphs



01

Knowledge Graph Embedding

- Represent the entities and relations in a **low-dimensional feature space**.
- Simplify the manipulation while **preserving the inherent structure** of the graph.

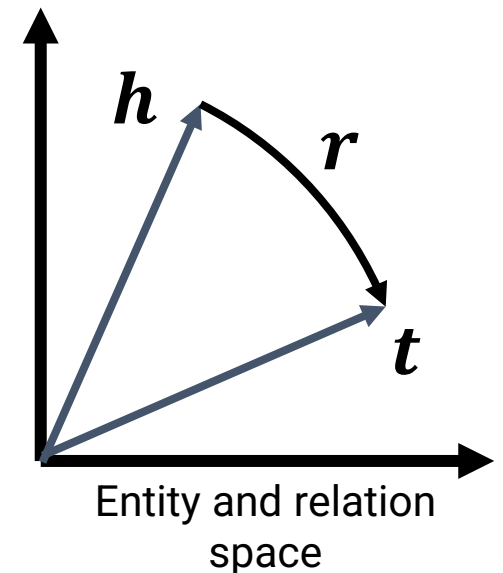
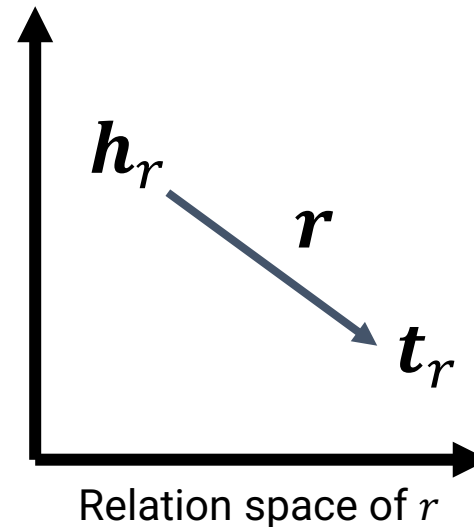
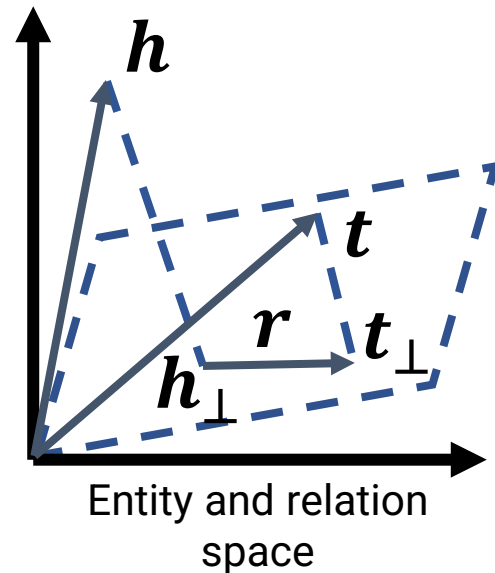
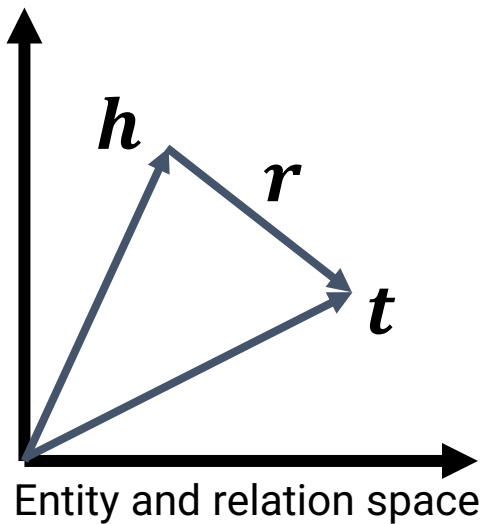


01

Knowledge Graph Embedding Models

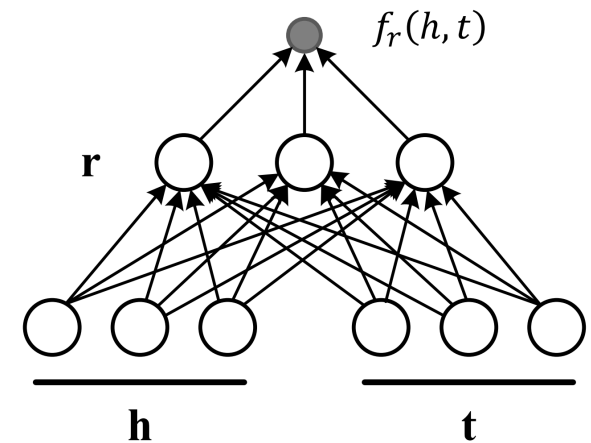
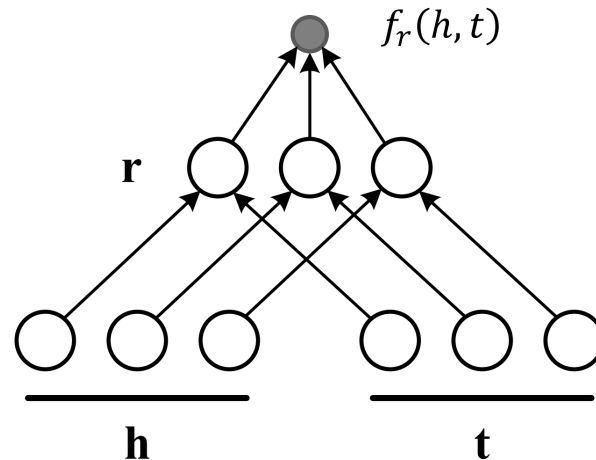
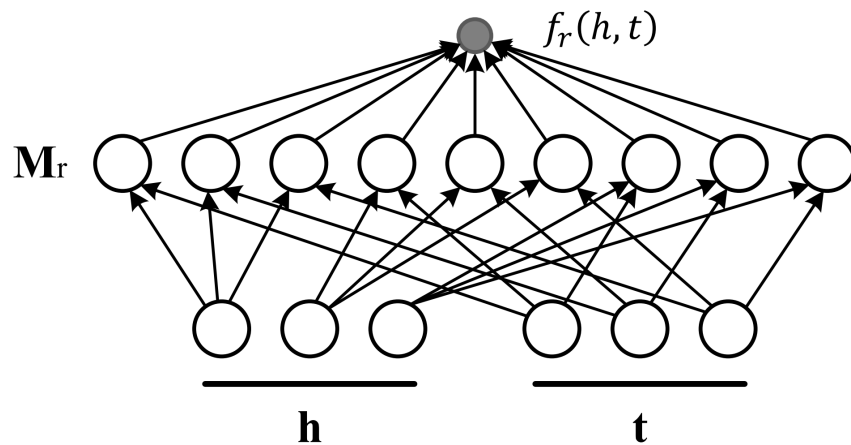
- Translational distance models

- Use distance-based scoring functions.
- Measure the plausibility of a fact as the distance between the two entities, usually after a translation carried out by the relation.
- Examples: TransE, TransH, TransR, RotatE



- Semantic matching models

- Exploit similarity-based scoring functions.
- Measure plausibility of facts by matching latent semantics of entities and relations embodied in their vector space representations.
- Examples: RESCAL, DistMult, HolE

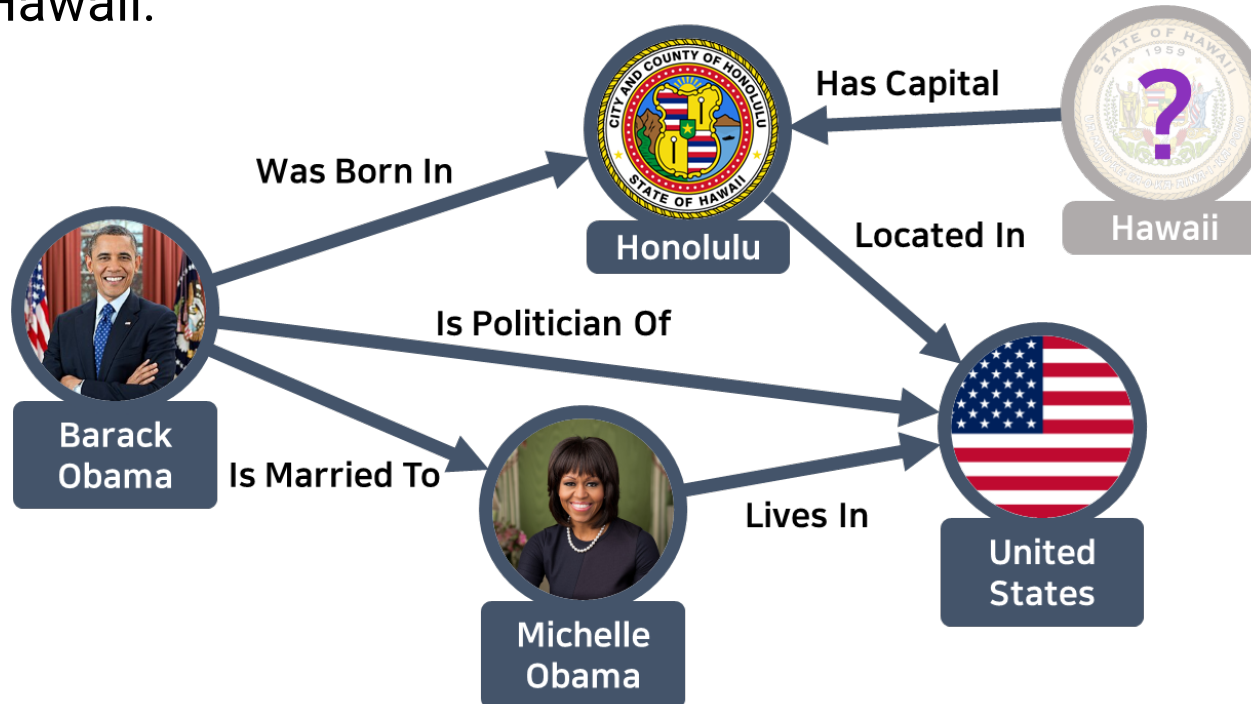


01 Training under Open World Assumption

- Knowledge graphs contain only true facts and non-observed facts can be either false or just missing.
- **Negative Sampling**: Generating negative training examples
 - Given a positive fact $(h, r, t) \in \mathcal{E}$, replace either h or t with a random entity
$$\mathcal{E}^- = \{(h', r, t) | h' \in \mathcal{V} \wedge h' \neq h \wedge (h, r, t) \in \mathcal{E}\} \cup \{(h, r, t') | t' \in \mathcal{V} \wedge t' \neq t \wedge (h, r, t) \in \mathcal{E}\}$$
 - False negative examples:
(JamesCameron, Gender, Male) $\rightarrow$ (AlfredHitchcock, Gender, Male)
 - To reduce such false-negative examples, set different probabilities for replacing the head and the tail.
 - Give more chance to replacing the head if the relation is 1-to-N and the tail if the relation is N-to-1

01 Link Prediction in Knowledge Graphs

- Assume that **we learned embeddings** using TransE.
- To solve (?, HasCapital, Honolulu), compute $f_r(h, t) = -\|h + r - t\|$.
 - **A ranked list**: Hawaii, United States, Barack Obama, Michelle Obama
 - Predict ? to be Hawaii.



- Record **ranks of correct answers** in the ordered lists.
 - **Mean Rank (MR)**
 - Compute the average of the ranks.
 - The lower the better
 - **Mean Reciprocal Rank (MRR)**
 - The average of reciprocal ranks.
 - The higher the better
 - **Hit@N**
 - The proportion of ranks no larger than N .
 - The higher the better

Evaluation of Link Prediction

[Problem 1] Answer: D

Rank	Entity
1	B
2	C
3	D
4	A
5	E
6	F
7	H
8	J
9	K
10	G
11	I
12	L
⋮	

[Problem 2] Answer: G

Rank	Entity
1	G
2	B
3	A
4	C
5	E
6	D
7	H
8	J
9	L
10	F
11	K
12	I
⋮	

[Problem 3] Answer: K

Rank	Entity
1	A
2	C
3	F
4	B
5	J
6	D
7	E
8	G
9	H
10	K
11	I
12	L
⋮	

[Problem 4] Answer: L

Rank	Entity
1	C
2	D
3	F
4	B
5	E
6	G
7	I
8	A
9	J
10	K
11	H
12	L
⋮	

$$MR=(3+1+10+12)/4=6.5, MRR=[1/3+1+1/10+1/12]/4=0.3792, \text{Hit}@10=3/4=0.75$$

03 References

- Some slides are made based on the following references.
 - Q. Wang et al., “Knowledge graph embedding: a survey of approaches and applications”, TKDE, 2017.
 - H. Cai et al., “A Comprehensive Survey of Graph Embedding: Problems, Techniques, and Applications”, TKDE, 2018.
 - S. Ji et al., “A Survey on knowledge graphs: representation, acquisition and applications”, TNNLS, 2021.