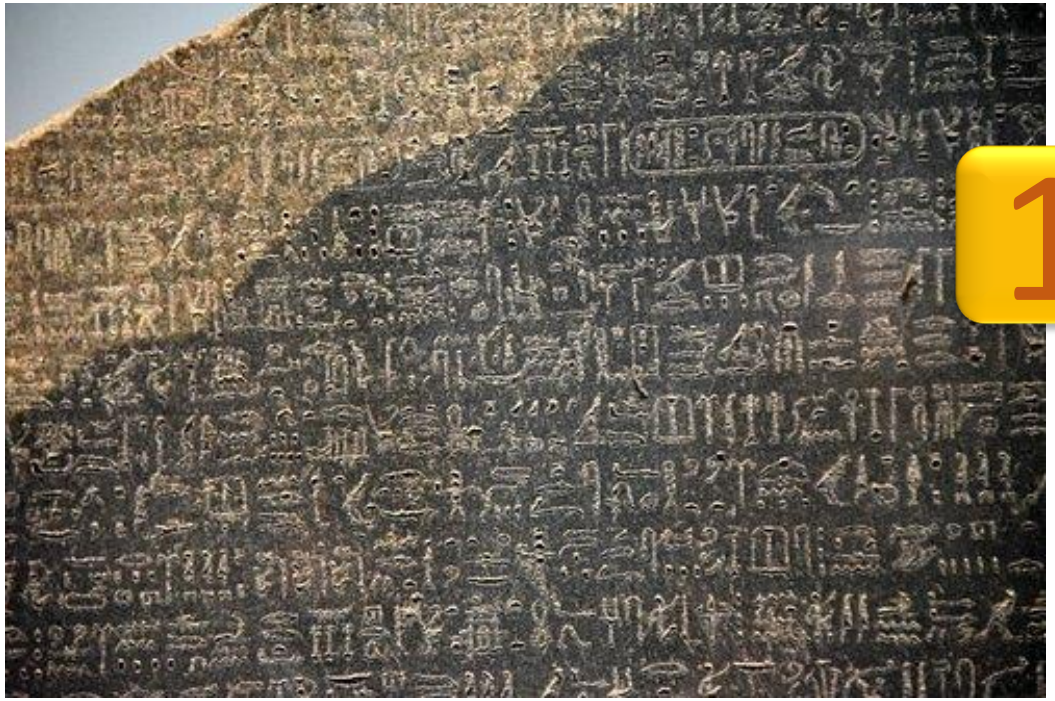
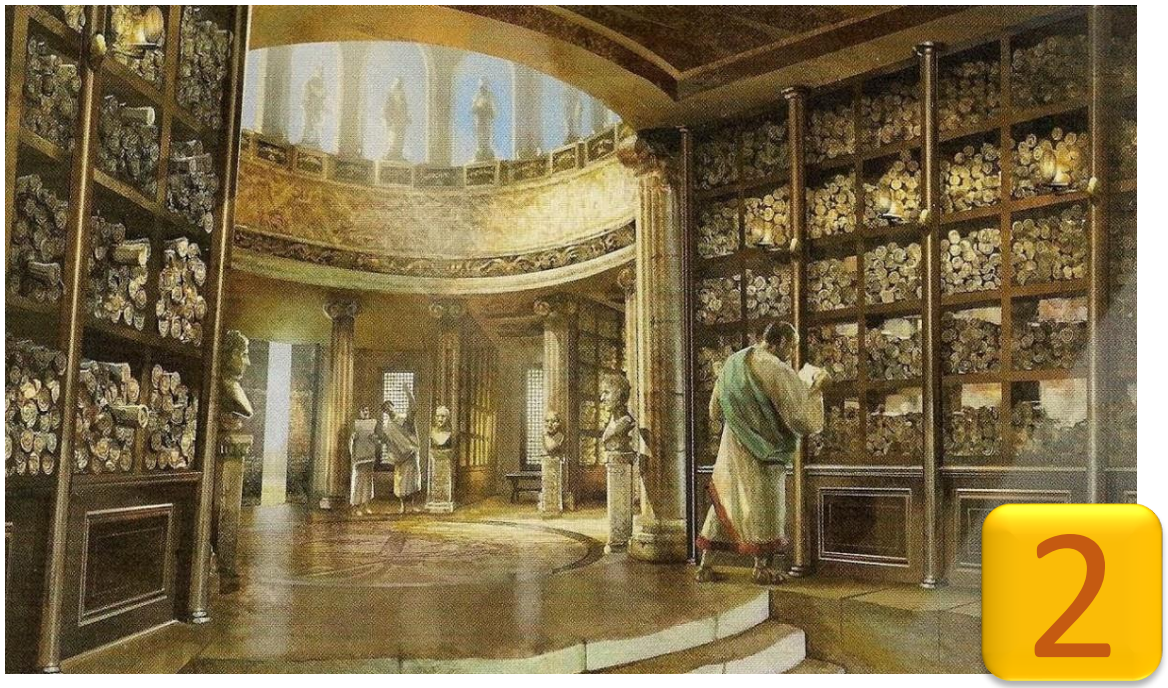
Knowledge and Retrieval

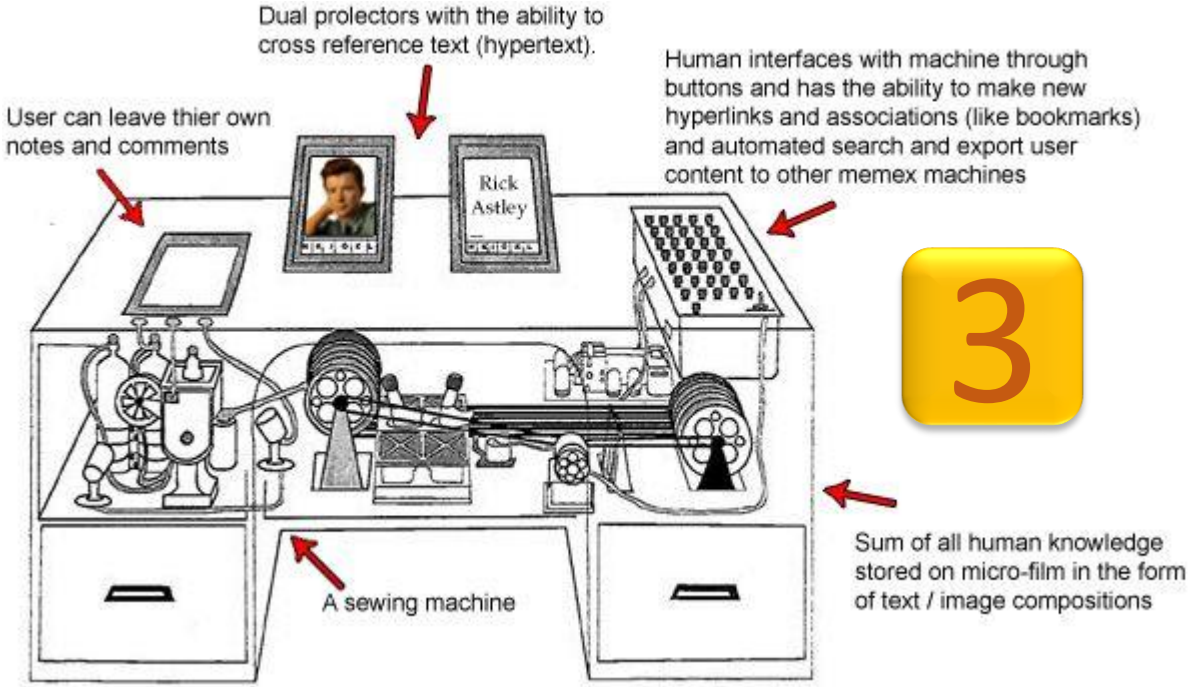
Knowledge Graph



1



2



3

Global WordNet Association

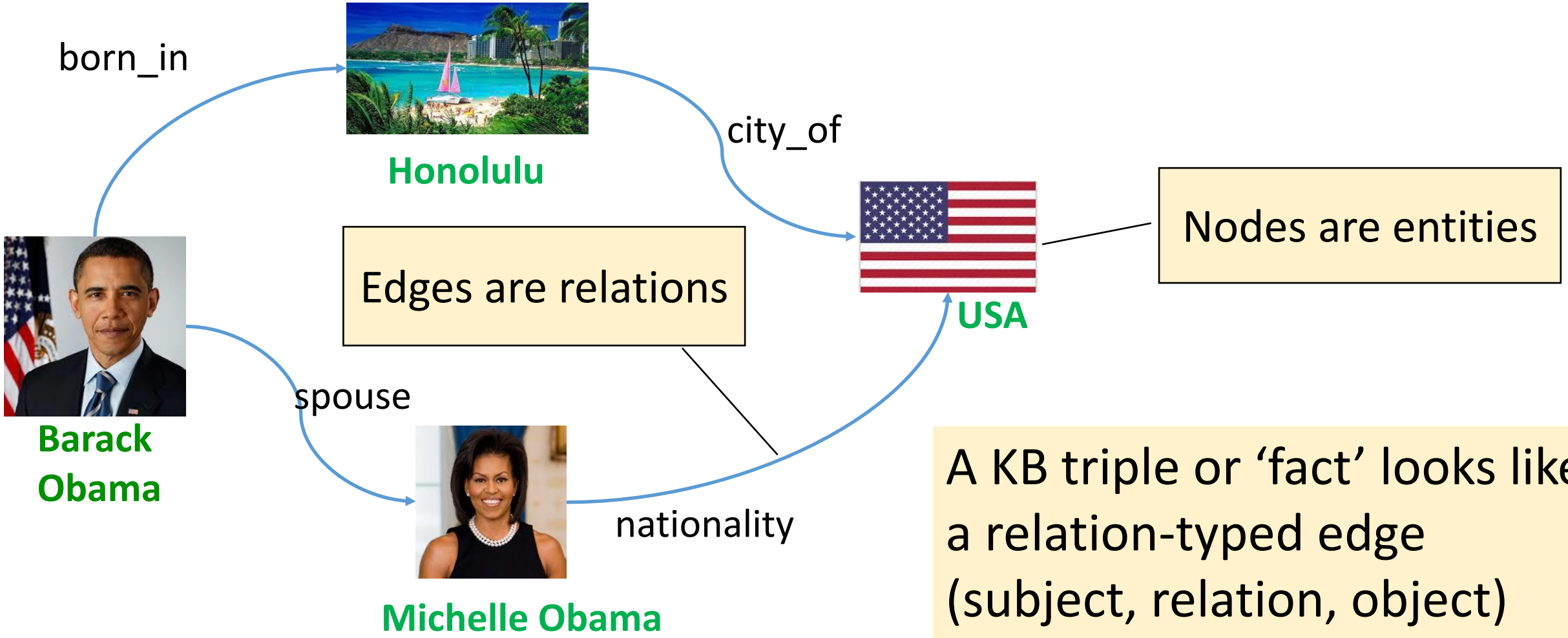
Cycorp

WIKIDATA

yago
select knowledge

Freebase™

Knowledge base (KB) / graph (KG)

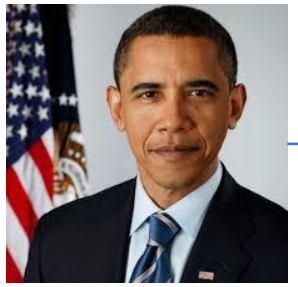


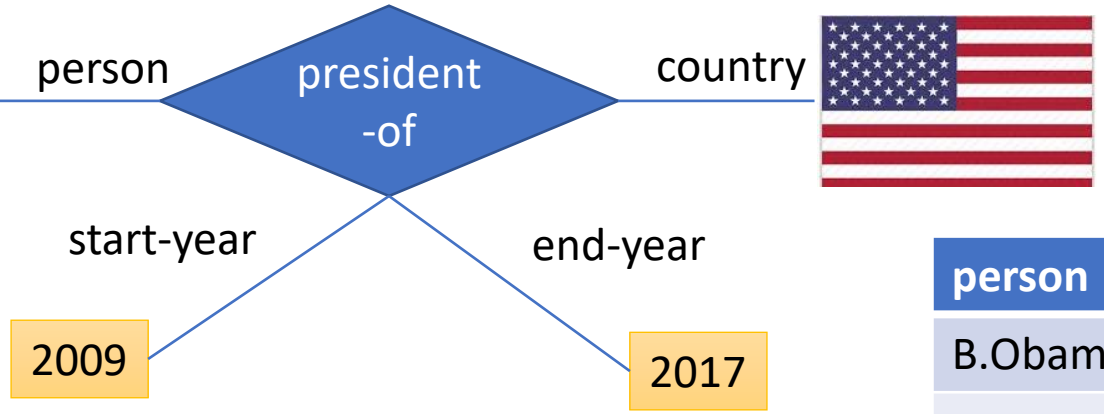
A KB triple or 'fact' looks like a relation-typed edge (subject, relation, object)

Knowledge base with entities and binary relations

Beyond binary relations

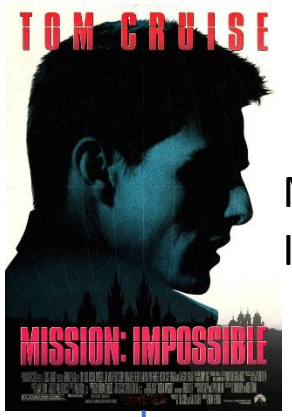
Barack Obama





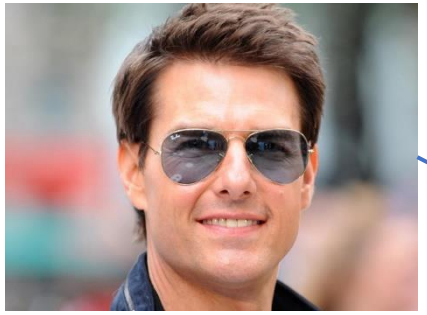
Table="president-of"

person	country	start-year	end-year
B.Obama	USA	2009	2017
J.Chirac	France	1995	2007
…	…	…	…

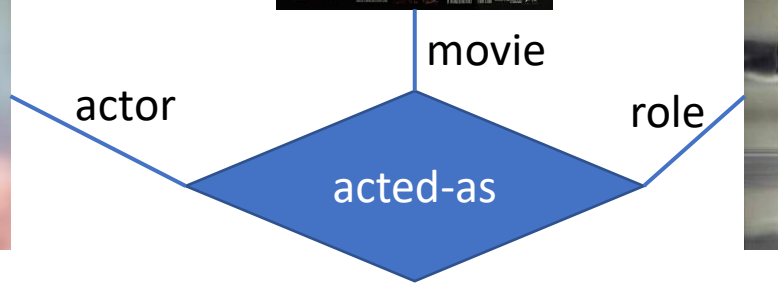


Mission: Impossible

Tom Cruise



Ethan Hunt



Table="acted-as"

actor	role	movie
Tom Hanks	Jim Lovell	Apollo 13
Tom Cruise	Ethan Hunt	Mission…
…	…	…

Wikidata: Entity ("topic") example

Name (can be ambiguous)

Unique canonical entity ID

Barack Obama (Q76)

44th president of the United States from 2009 to 2017

Barack Hussein Obama II | Barack Obama II | Barack Hussein Obama | Obama | Barak Obama | Barry Obama | President Obama | President Barack Obama | BHO | Barack | Barack H. Obama

▾ In more languages

Configure

Language	Label	Description	Also known as
English	Barack Obama	44th president of the United States from 2009 to 2017	Barack Hussein Obama II Barack Obama II Barack Hussein Obama Obama Barak Obama Barry Obama President Obama President Barack Obama BHO Barack Barack H. Obama
Hindi	बराक ओबामा	संयुक्त राज्य अमेरिका के 44वें राष्ट्रपति	बराक हुसैन ओबामा द्वितीय बराक ओबामा द्वितीय

A prominent description

Other English names or "aliases"

Aliases in other languages

Wikidata: Relations (“fact”) example

position held (P39)

Unique canonical relation ID

subject currently or formerly holds the object position or public office

political office held | political seat | public office | office held | position occupied | holds position | function | held position

▾ In more languages

Configure

Language	Label	Description	Also known as
English	position held	subject currently or formerly holds the object position or public office	political office held political seat public office office held position occupied holds position function held position
Hindi	पद पर आसीन	आइटम मौजूदा या पूर्व समय में इस राजनीतिक पद को धारण किए हुए है	
Bangla	ভারপ্রাপ্ত পদ	এক ব্যক্তি যে সকল পদ অধিগ্রহণ করেছেন	অলঙ্কৃত পদ

Relation aliases in multiple languages

position held

replaced by Donald Trump

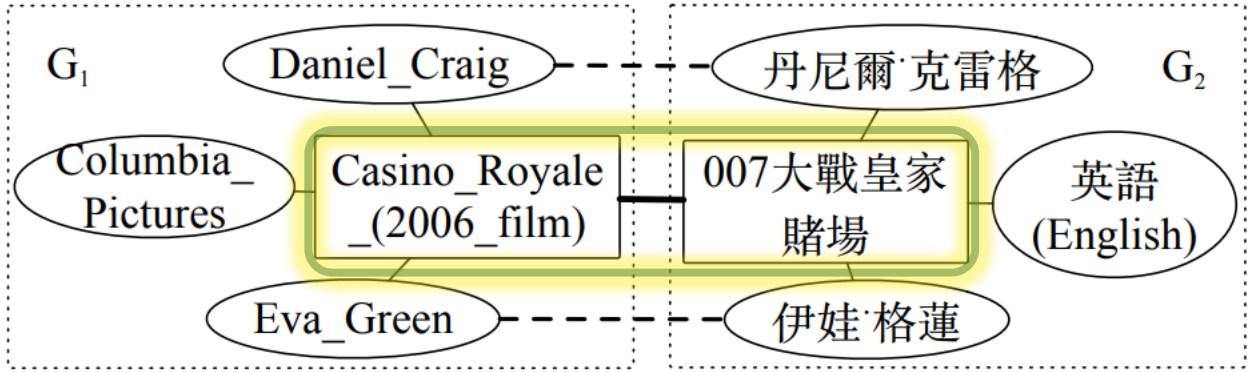
KGs: Useful but incomplete

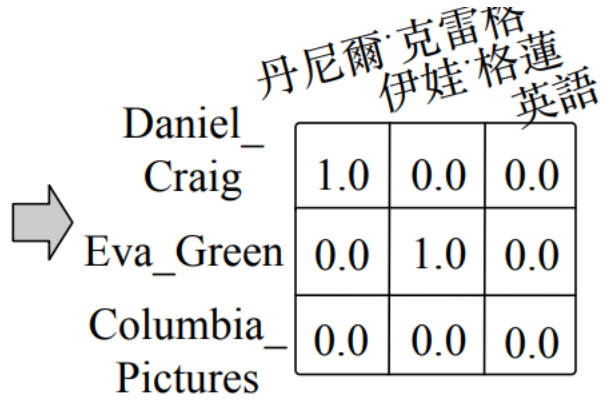
- Used for question answering, product search, dialog, information extraction
- Not possible to manually cover all human knowledge nor stay current
- Makes KG completion an important task

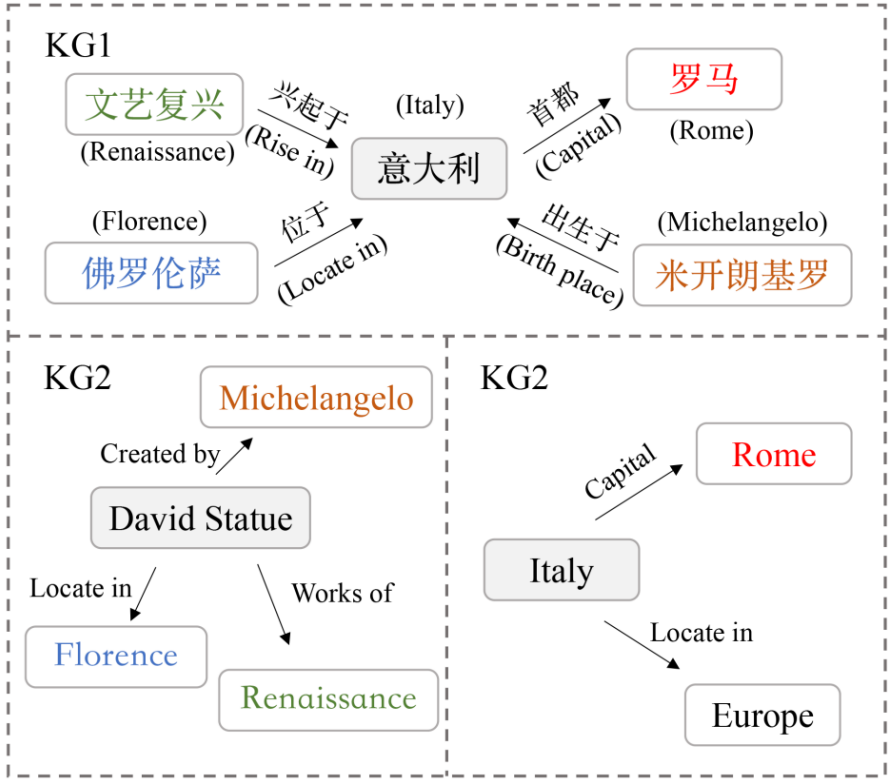
[Min et al 2013]

Freebase relations	% INcomplete
/people/person/parents	98.8
/people/person/places_lives	96.6
/people/person/place_of_birth	93.8
/people/person/employment_history	92.3
/people/person/nationality	78.5
/people/person/education	79.2

KG alignment



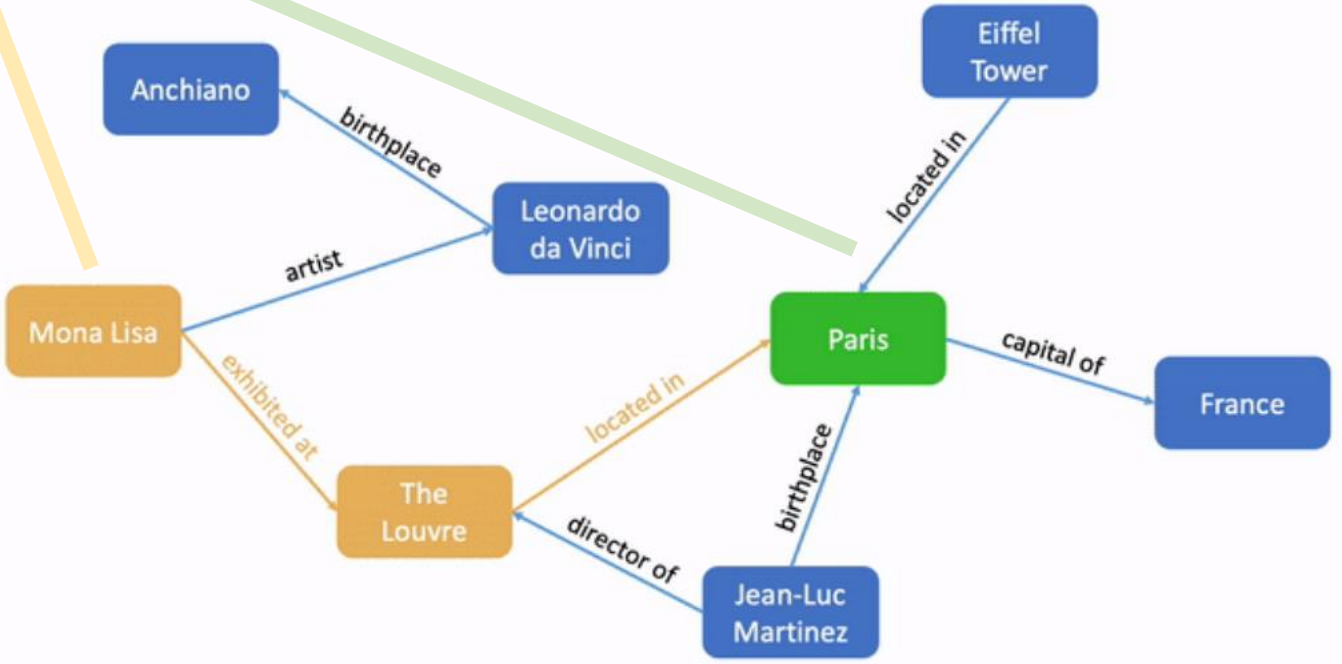




- Transliteration, translation give some clue for entity alignment
- Agreement of neighboring entities is also a strong signal
- Neighbors through what relation? May also be important
- Synergy between entity and relation alignment, with KG completion

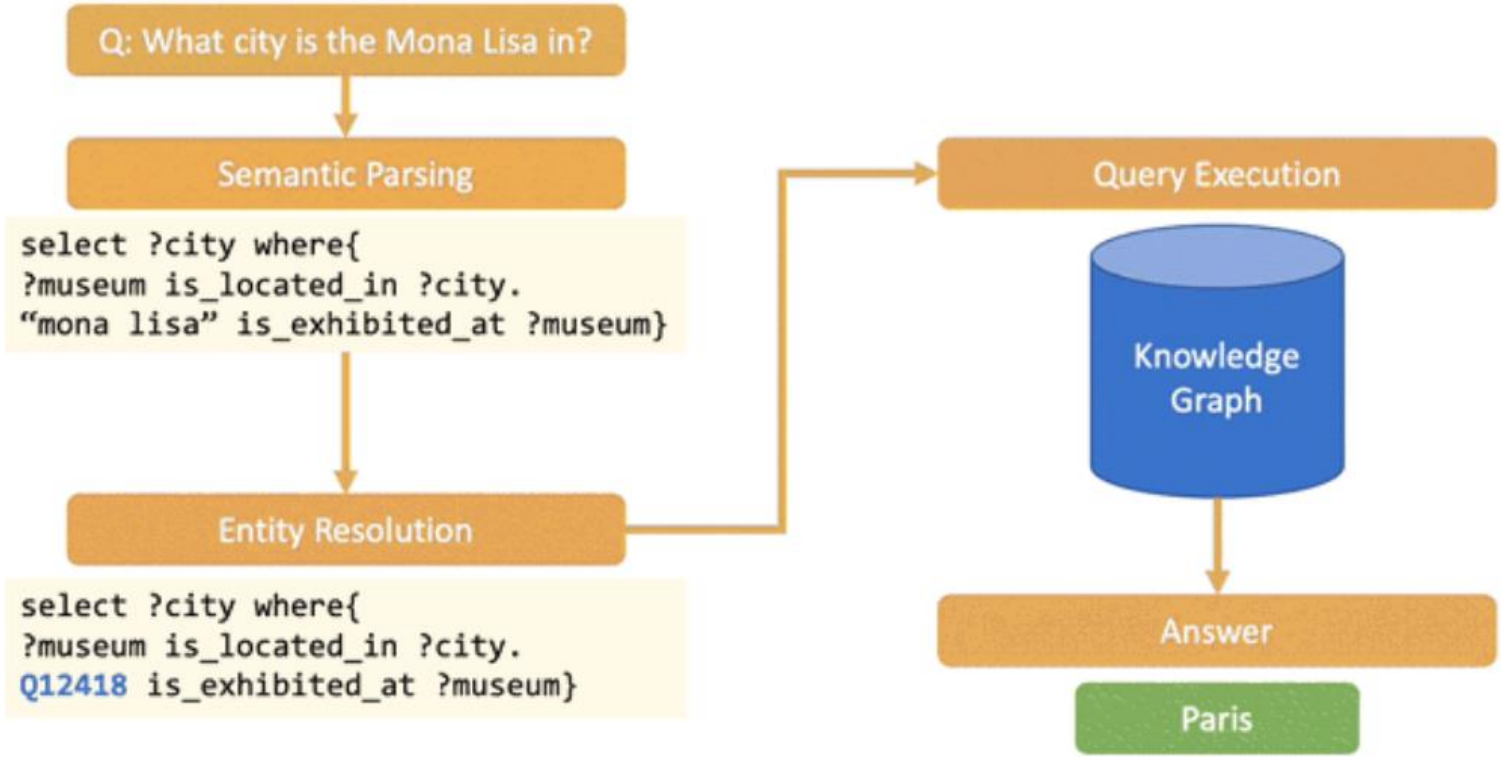
KG-based question answering (KGQA)

- What city is the Mona Lisa in?
- Disambiguate Mona Lisa
- Infer that exhibited-at followed by located-in implies Mona Lisa is in Paris
- Look up or infer that Paris is a city
- Approaches
  - Semantic interpretation
  - End-to-end trainable



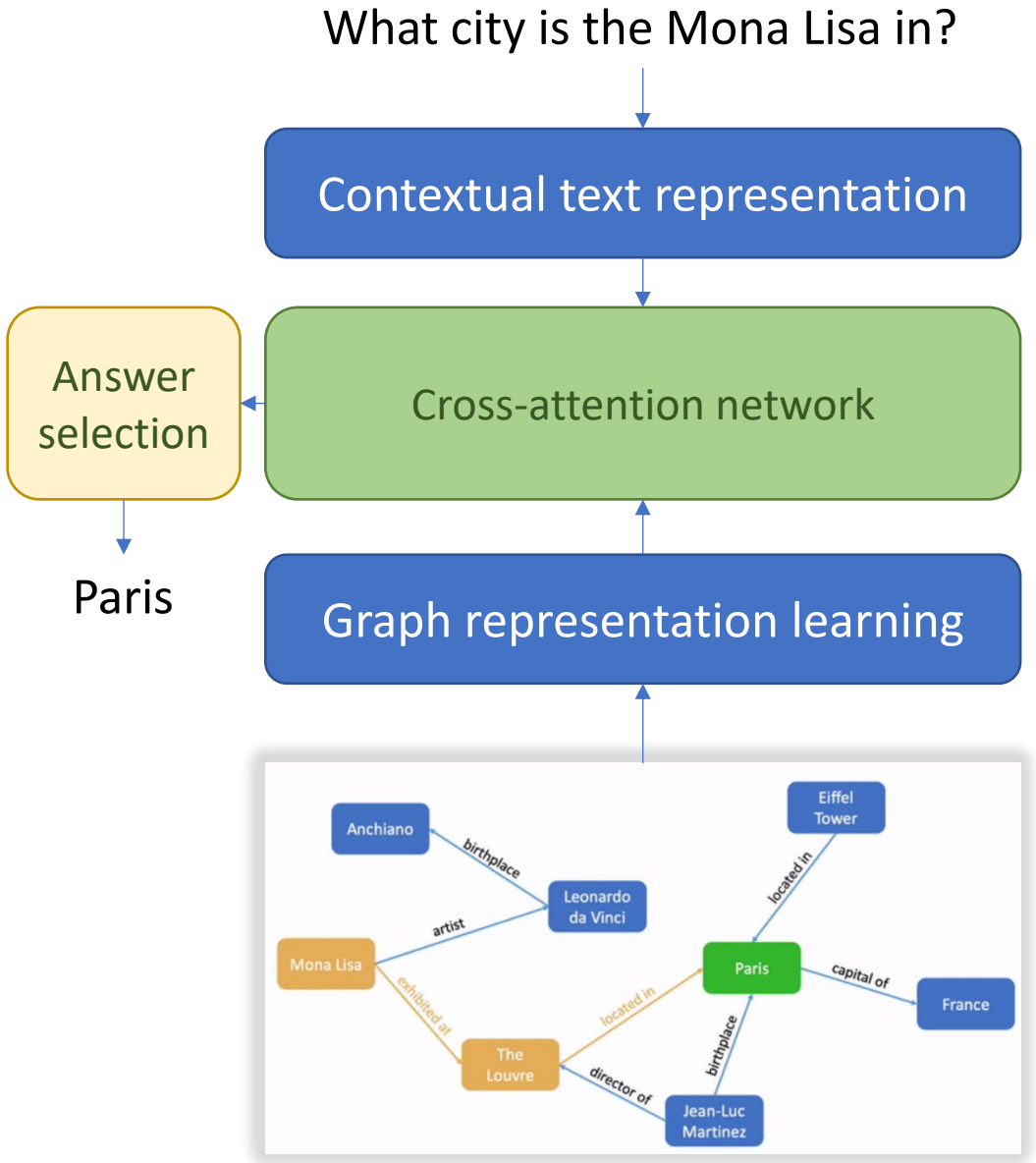
Semantic interpretation

- Translate natural language question to structured query cognizant of KG schema
- Execute structured query on KG
  - Interpretable
  - Tricky to implement
  - Not (easily) end-to-end trainable



“Differentiable KG”

- Find dense query embedding (e.g., using BERT)
- Find dense graph embedding (e.g., GCN, graph transformer)
- Cross attention layers establish correspondences between query and graph elements
- Answer selection identifies graph element(s) to transcribe to response
- 👍End-to-end trainable, no logical form needed
- 👎Not easily interpretable



Layout of this module

Module
Preliminaries (this module)
Fact scoring and sampling
Translation and rotation models
Factorization/multiplicative models
Embedding hierarchies
Embeddings temporal KGs
Multilingual KG alignment
Integration with LLMs, QA applications