



UCLA

Samueli  
Computer Science

# Graph Learning Session

## Incorporating Ontological Information in Knowledge Graph Learning and Applications

Junheng Hao

PhD Candidate, University of California, Los Angeles (UCLA)

Research Intern, Microsoft Search, Assistant and Intelligence Team (MSAI)

Aug 13, 2021



# Junheng Hao

Research Intern, MSAI (2020)

PhD Candidate, University of California Los Angeles (2017-)

Advisor: Wei Wang, Yizhou Sun

Website: [Jeff's Home \(haojunheng.com\)](http://haojunheng.com)

## Bio

- Currently 4th-year Ph.D. candidate at UCLA co-advised by Yizhou Sun and Wei Wang in UCLA Data Mining Group.
- My research interests include knowledge graph, graph representation learning, KG-empowered applications (NLP, Bioinformatics, recommender systems, etc.).
- Before joining UCLA, I graduated in 2017 from Department of Automation, Tsinghua University.

## Past Experiences

- PhD Research Intern, IBM, 2020
- Applied Science Intern, Amazon Product Graph, 2019
- Research Intern, NEC Labs America, 2018

# Today's Agenda

- Background: Knowledge graphs and representation learning
- **JOIE**: Joint learning on instance and ontology view on knowledge graphs
- Two JOIE-inspired applications: Bioinformatics (**Bio-JOIE**) and Recommender Systems (**P-Companion**)
- Summer intern project at MSAI: **DocGraph**

# Papers

- [Universal representation learning of knowledge bases by jointly embedding instances and ontological concepts](#) (KDD'19)
- [Bio-JOIE: Joint representation learning of biological knowledge bases](#) (ACM BCB'20, Best Student Paper)
- [P-Companion: A principled framework for diversified complementary product recommendation](#) (CIKM'20)

# Background: Knowledge Graphs and Representation Learning on KG

*What is a KG? What is the structure of a KG?*

# KG: Whenever you google...



Google

Q All News Images Videos Books More Settings Tools

About 73,600,000 results (0.89 seconds)

### Mike Bloomberg 2020 | Fighting for our future

[www.mikebloomberg.com/](http://www.mikebloomberg.com/)

You demand change, Mike will fight. Gun safety, education, healthcare, the environment. See how Mike's been successful at fighting President Trump, and how he'll fight for you.

Paid for by MIKE BLOOMBERG 2020 INC

**About Mike**  
From childhood to today  
Mike's life story

**Get Involved**  
Let's fight together  
Join us

### Top stories

Trump Bars Bloomberg News Journalists From Campaign Events  
The New York Times  
6 hours ago

Trump attacks 'Mini Mike Bloomberg' after campaign bars news outlet | TheHill  
TheHill  
1 hour ago

Trump attacks Bloomberg News after his campaign says it will deny press...  
Washington Post  
51 mins ago

→ More for mike bloomberg

### mike bloomberg on Twitter

<https://twitter.com/search/mike+bloomberg>

**Ronna McDaniel** (@GOPChairwoman)  
Media outlets should be independent and fair...

**Donald J. Trump** (@realDonaldTrump)  
Mini Mike Bloomberg has instructed his aides...

**Mike Bloomberg** (@MikeBloomberg)  
The NRA's latest effort to...

## Michael Bloomberg

CEO of Bloomberg L.P.

Michael Rubens Bloomberg is an American politician, businessman, and author. He is the co-founder, CEO, and majority owner of Bloomberg L.P. He was mayor of New York City from 2002 to 2013. On November 24, 2019 he announced his candidacy for the 2020 United States presidential election. [Wikipedia](#)

**Party:** Democratic Party Trending

**Born:** February 14, 1942 (age 77 years), Brighton, MA

**Height:** 5' 8"

**Net worth:** 54.6 billion USD (2019)

**Partner:** Diana Taylor (2000-)

**Children:** Georgina Bloomberg, Emma Bloomberg

### Profiles

[Twitter](#) [Facebook](#) [Instagram](#) [YouTube](#)

### People also search for

[View 15+ more](#)

Diana Taylor  
Partner

Andrew Cuomo  
Trending

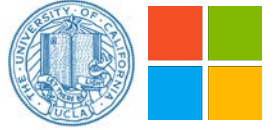
Georgina Bloomberg  
Daughter

Larry Ellison

Larry Page

*Facts from KG*

# Knowledge graphs (KGs) Are Everywhere



## General-purpose KGs



## Bio & Medical KGs



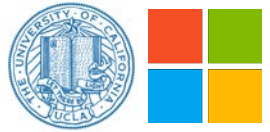
## Product Graphs & E-commerce



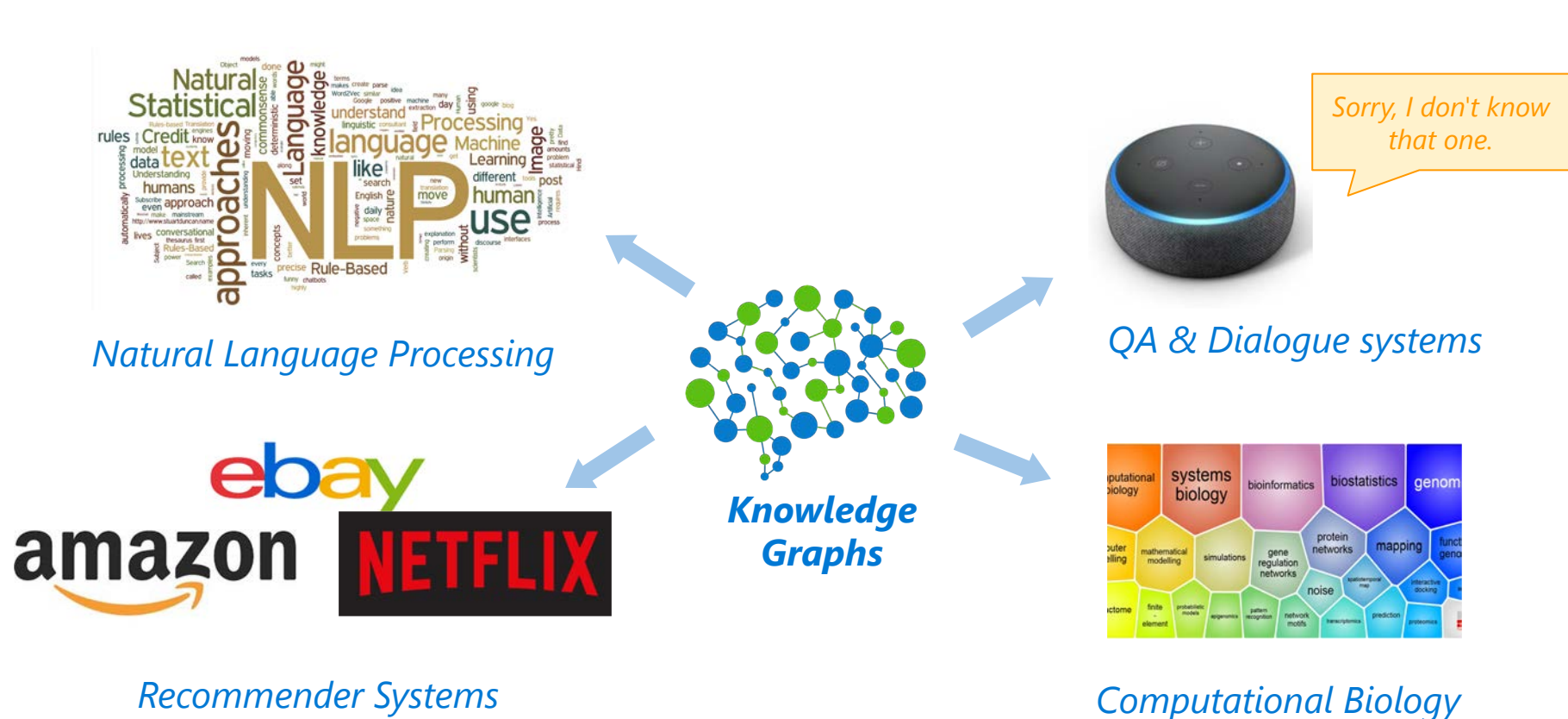
## Common-sense KGs & NLP



# Knowledge Graphs Are Important

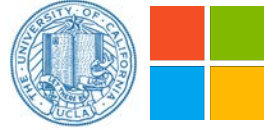


- Foundational to knowledge-driven AI systems
- Enable many downstream applications (NLP tasks, QA systems, etc.)



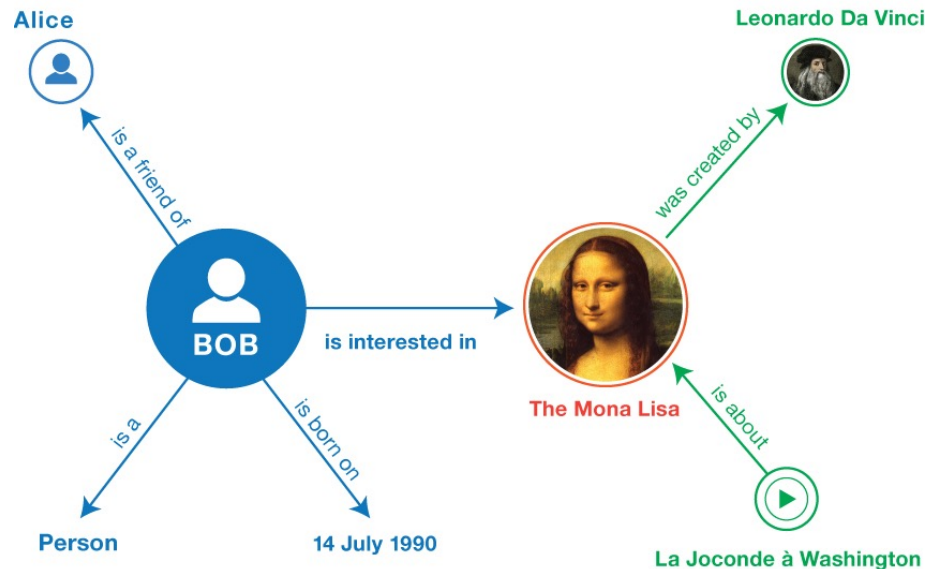


# How are KGs structured or formatted?



- **Triples (RDF)**

- Represented by: a node for the subject, a node for the object, and an arc/node for the predicate.
- Example: Semantic Web, medical ontologies, etc.



- **Label-property**

- Entity, labels, properties, qualifiers, etc.
- Example: Wikidata

label — **Douglas Adams** (Q42) — item identifier

description — English writer and humorist  
Douglas Noël Adams | Douglas Noel Adams — aliases  
▶ In more languages

property — **educated at** — value

end time	1974
academic major	English literature
academic degree	Bachelor of Arts
start time	1971

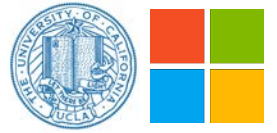
qualifiers

rank — ▼ 2 references

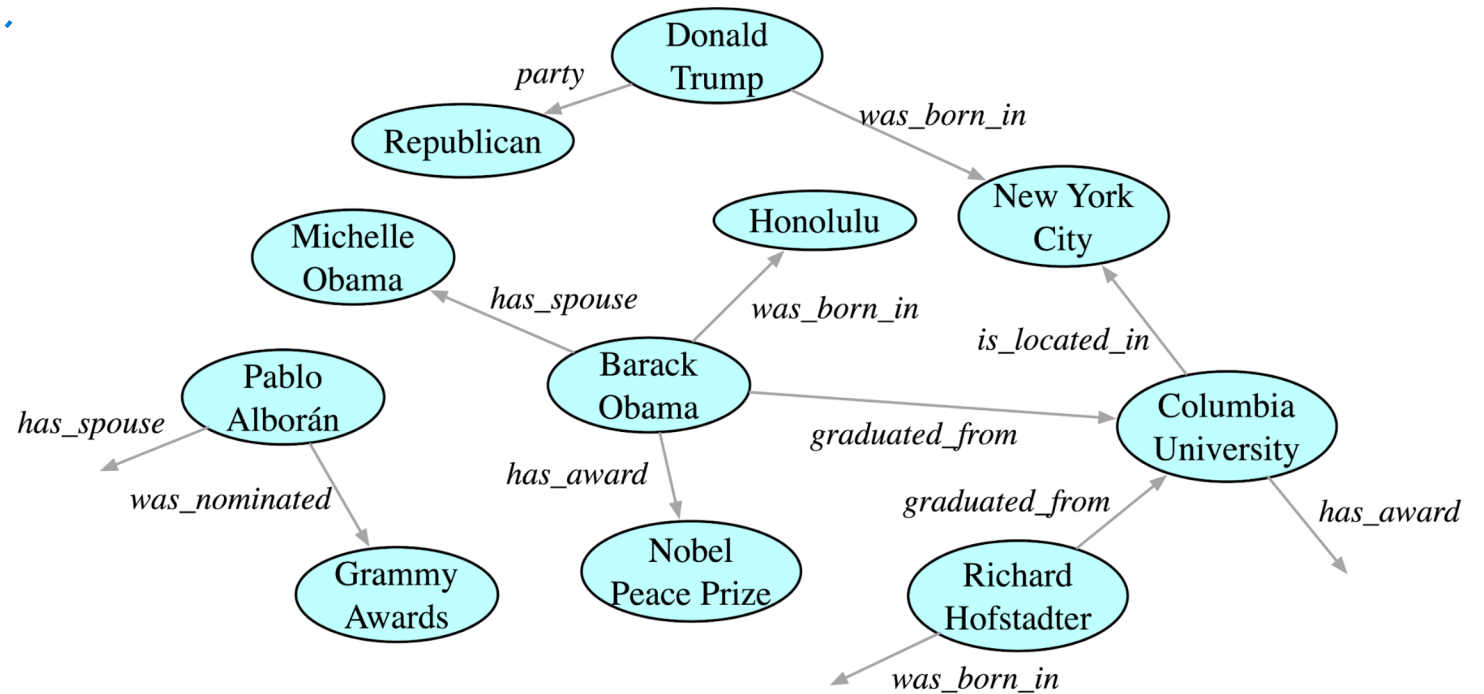
statement group	opened references
stated in	Encyclopædia Britannica Online
reference URL	http://www.nndb.com/people/731/000023662/
original language of work	English
retrieved	7 December 2013
publisher	NNDB
title	Douglas Adams (English)
	+ add reference

statement group	collapsed reference
Brentwood School	
end time	1970
start time	1959
	▶ 0 references
	+ add (statement)

# A KG Snapshot from YAGO, made with triples



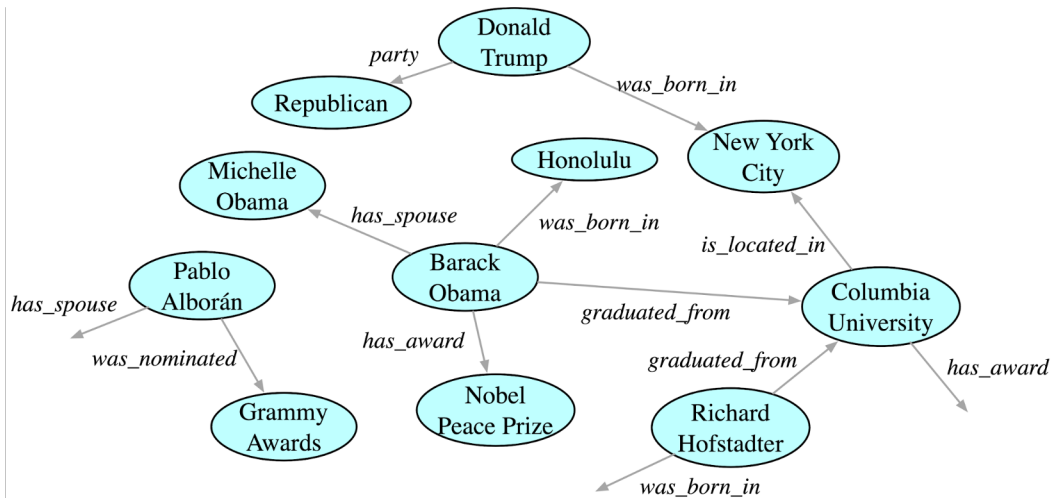
Example **UCLA** *Located In* *Los Angeles*



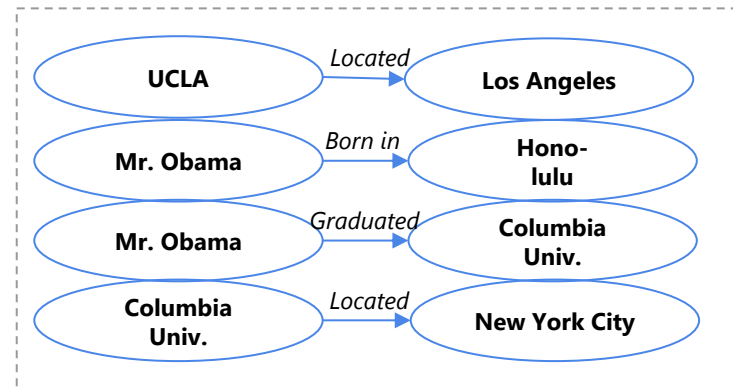
# KG Embeddings From Triples



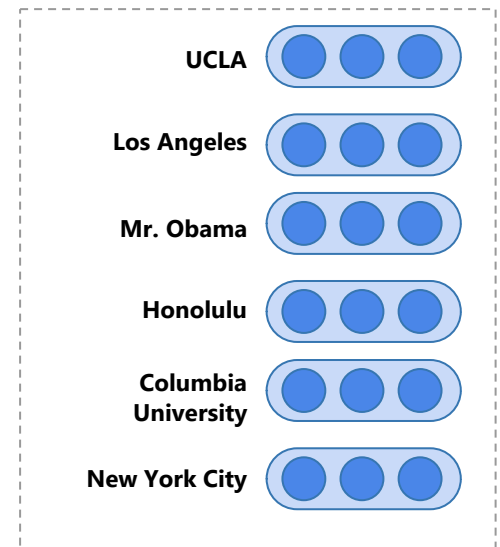
- Knowledge graph embeddings represent entities and relations as latent vectors or matrices and support effective relation learning and inference.
- **Input:** Relation facts (triples)
- **Output:** Embedding representations of objects and relations



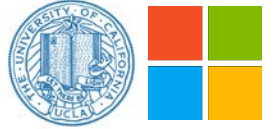
Input



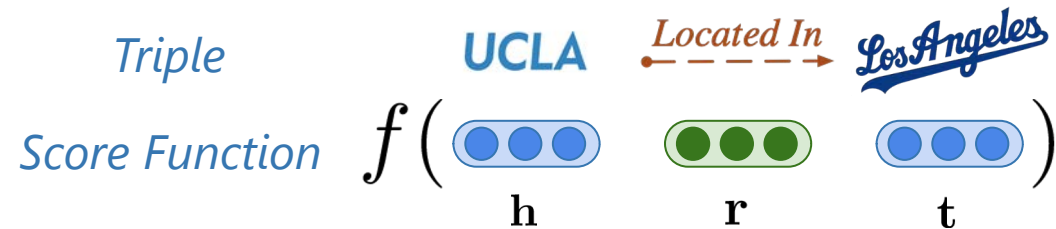
Output



# Learning KG Embeddings



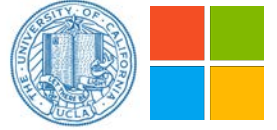
- Key of existing KG embedding methods: Triple score function



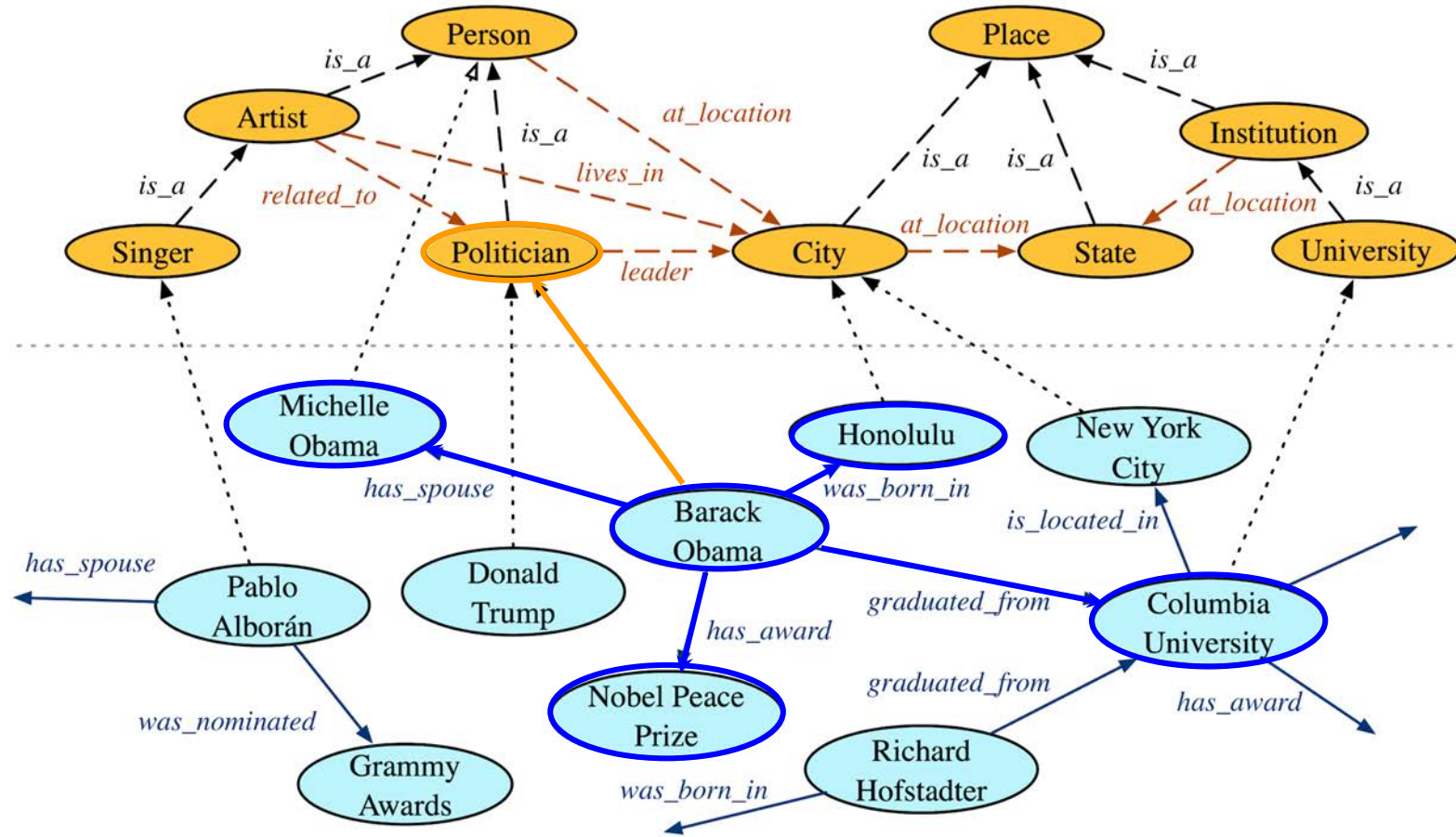
- Previous research employ various arithmetic methods to capture observed relations of entities in a single KG (for example, translational distance or similarity)

Model	Score Function	Embeddings
TransE (Bordes et al., 2013)	$-  \mathbf{h} + \mathbf{r} - \mathbf{t}  $	$\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{R}^k$
TransX	$-  g_{r,1}(\mathbf{h}) + \mathbf{r} - g_{r,2}(\mathbf{t})  $	$\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{R}^k$
DistMult (Yang et al., 2014)	$(\mathbf{h} \circ \mathbf{t}) \cdot \mathbf{r}$	$\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{R}^k$
HolE (Nickel et al., 2016)	$(\mathbf{h} \star \mathbf{t}) \cdot \mathbf{r}$	$\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{R}^k$
ComplEx (Trouillon et al., 2016)	$\text{Re}\langle \mathbf{r}, \mathbf{h}, \bar{\mathbf{t}} \rangle$	$\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{C}^k$
ConvE (Dettmers et al., 2017)	$\langle \sigma(\text{vec}(\sigma([\mathbf{r}, \mathbf{h}] * \Omega))\mathbf{W}), \mathbf{t} \rangle$	$\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{R}^k$
RotatE (Sun et al., 2019)	$-  \mathbf{h} \circ \mathbf{r} - \mathbf{t}  ^2$	$\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{C}^k,  r_i  = 1$

# A different view: How can we learn "Obama"?



Ontology-view Knowledge Graph

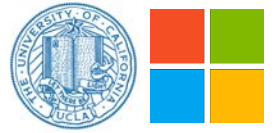


Instance-view Knowledge Graph

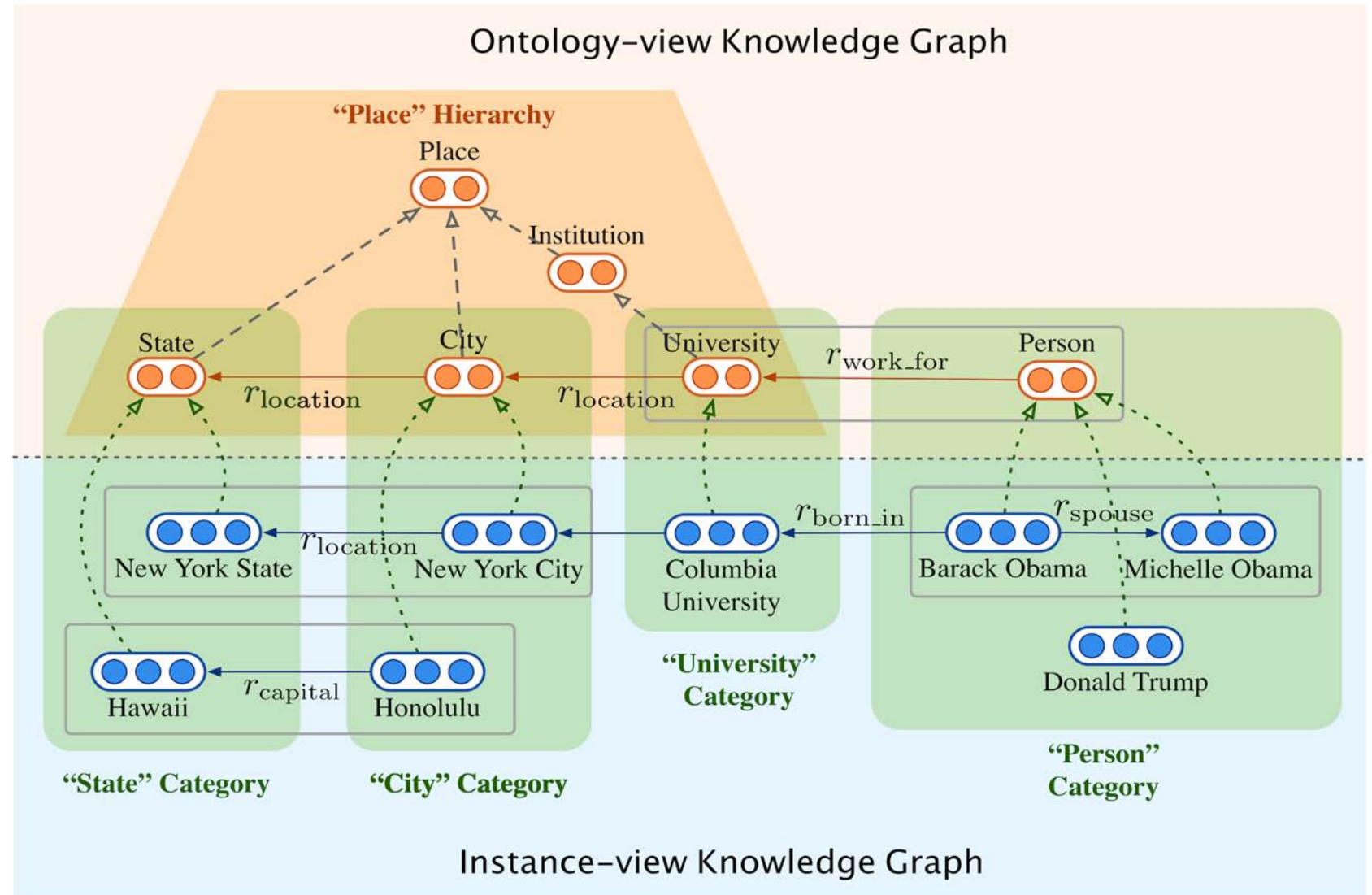
JOIE: Universal representation learning of knowledge bases by jointly embedding instances and ontological concepts

*How can we manage to jointly learn the instance and ontology?*

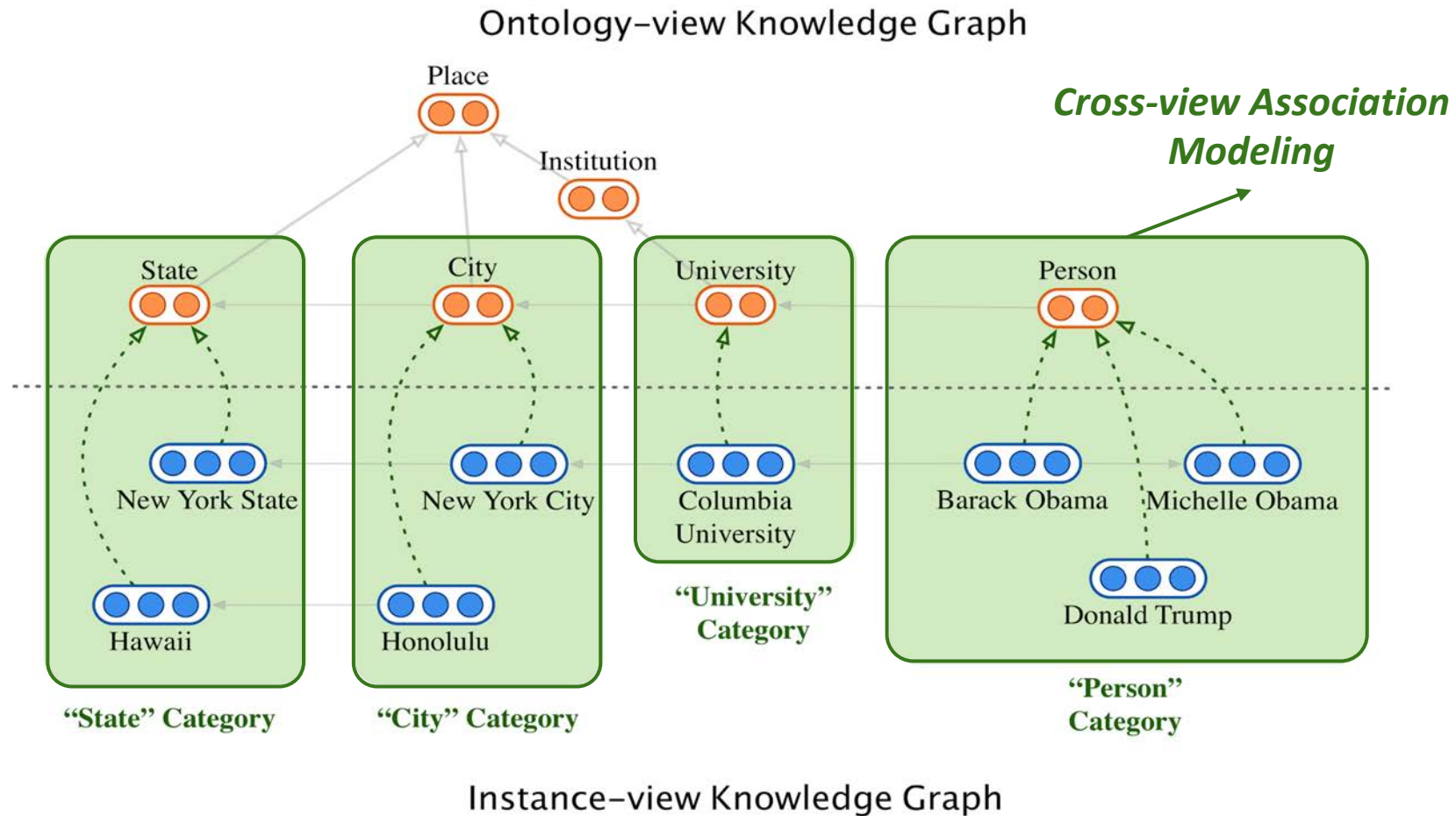
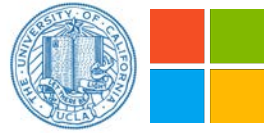
# JOIE: Learning on Instance & Ontology View



- Cross-view Association model
- Intra-view model



# JOIE: Cross-view Association





# JOIE: Cross-view Association Modeling



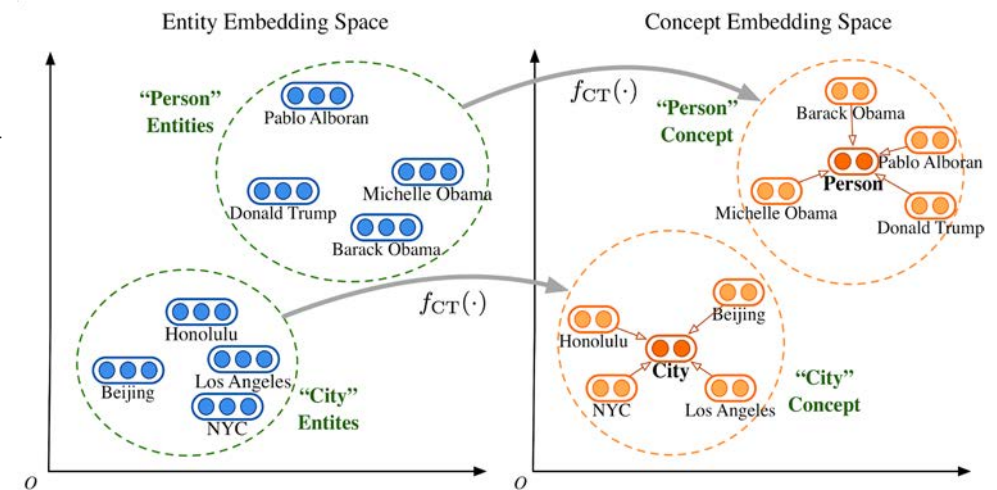
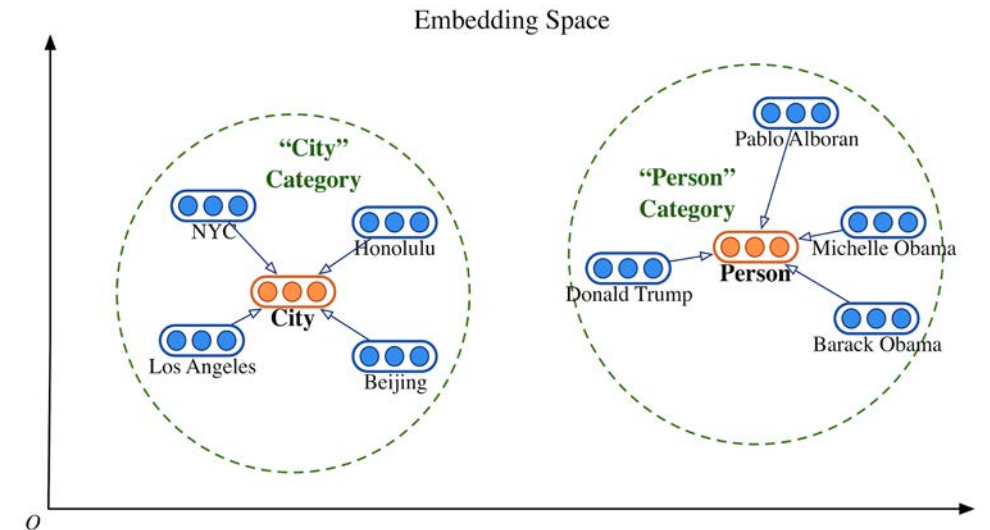
- **Goal:** capture associations between the entities  $\mathbf{e}$  and corresponding concepts  $\mathbf{c}$
- **Cross-view Grouping (CG)**

$$J_{\text{Cross}}^{\text{CG}} = \frac{1}{|\mathcal{S}|} \sum_{(e,c) \in \mathcal{S}} [\|\mathbf{c} - \mathbf{e}\|_2 - \gamma^{\text{CG}}]_+$$

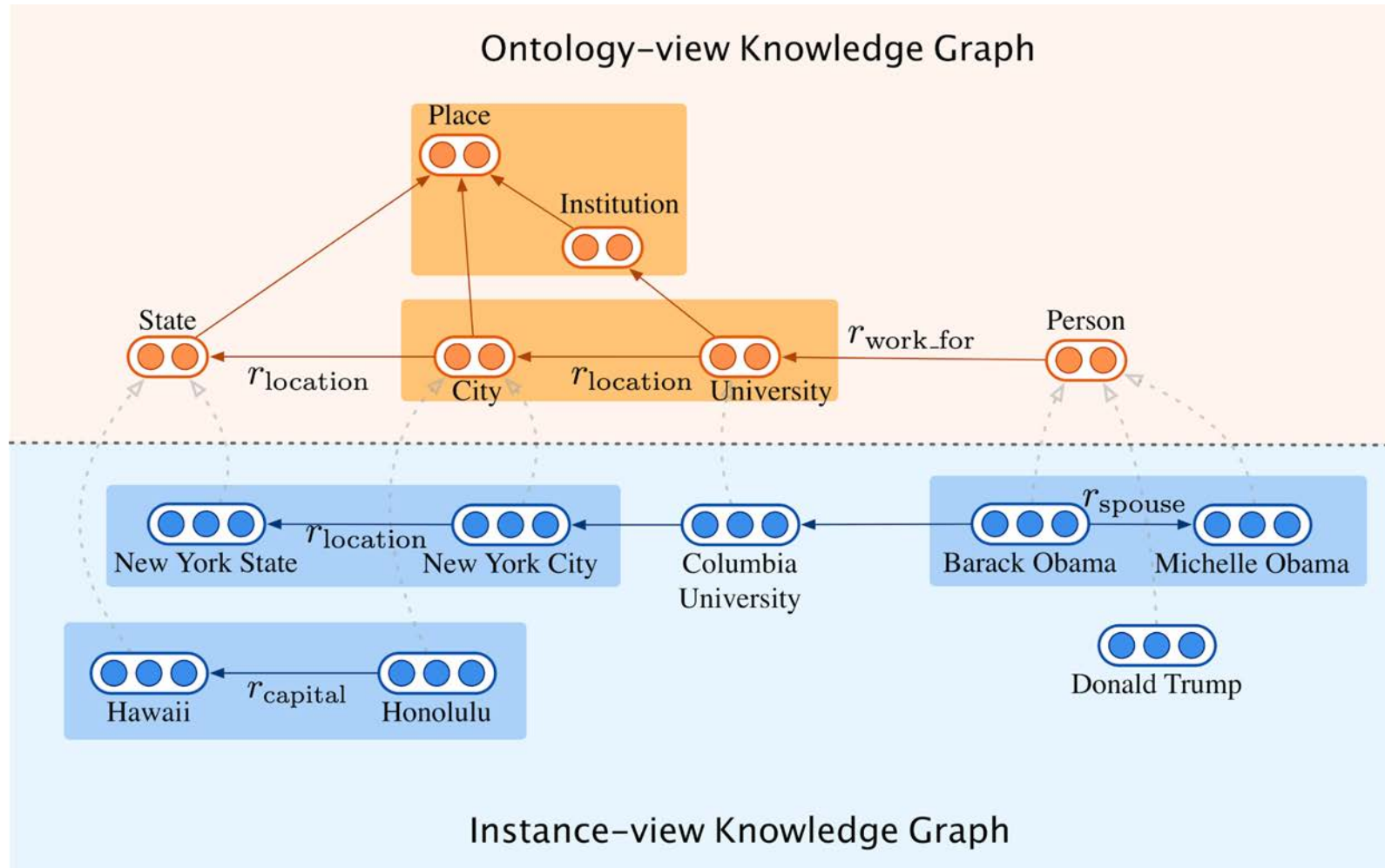
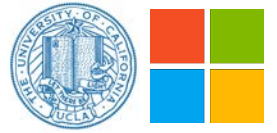
- **Cross-view Transformation (CT)**

$$f_{\text{CT}}(\mathbf{e}) = \sigma(\mathbf{W}_{\text{ct}} \cdot \mathbf{e} + \mathbf{b}_{\text{ct}})$$

$$J_{\text{Cross}}^{\text{CT}} = \frac{1}{|\mathcal{S}|} \sum_{\substack{(e,c) \in \mathcal{S} \\ \wedge (e,c') \notin \mathcal{S}}} [\gamma^{\text{CT}} + \|\mathbf{c} - f_{\text{CT}}(\mathbf{e})\|_2 - \|\mathbf{c}' - f_{\text{CT}}(\mathbf{e})\|_2]_+$$



# JOIE: Intra-view



# JOIE: Intra-view Model



- Goal: To embed the relational structures in the instance view of the KB
- Apply any KG embedding techniques on instance view
  - Three representatives: TransE, DistMult, and HolE

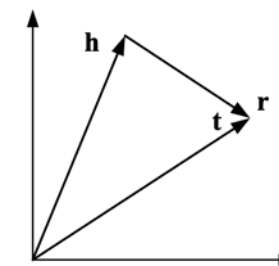
$$f_{\text{TransE}}(\mathbf{h}, \mathbf{r}, \mathbf{t}) = -\|\mathbf{h} + \mathbf{r} - \mathbf{t}\|_2$$

$$f_{\text{Mult}}(\mathbf{h}, \mathbf{r}, \mathbf{t}) = (\mathbf{h} \circ \mathbf{t}) \cdot \mathbf{r}$$

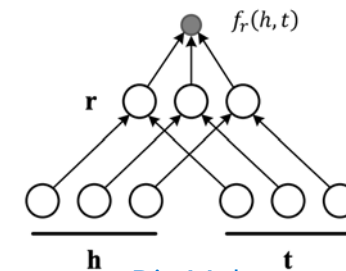
$$f_{\text{HolE}}(\mathbf{h}, \mathbf{r}, \mathbf{t}) = (\mathbf{h} \star \mathbf{t}) \cdot \mathbf{r}$$

- Training on contrastive margin loss

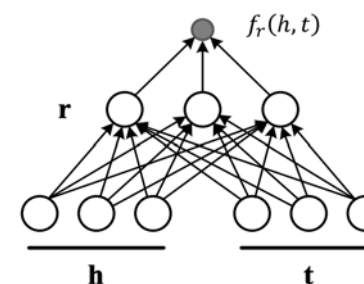
$$J_{\text{Intra}}^{\mathcal{G}} = \frac{1}{|\mathcal{G}|} \sum_{\substack{(h,r,t) \in \mathcal{G} \\ \wedge (h',r,t') \notin \mathcal{G}}} [\gamma^{\mathcal{G}} + f(\mathbf{h}', \mathbf{r}, \mathbf{t}') - f(\mathbf{h}, \mathbf{r}, \mathbf{t})]_+$$



TransE



DistMult



HolE

# JOIE: Joint Training & Model Summarization



- Two model components: Cross-view model and intra-view model
- Cross-view association model  $\Rightarrow J_{\text{Cross}}$ 
  - Categorical grouping (CG)
  - Categorical transformation (CT)
- Intra-view model  $\Rightarrow J_{\text{Intra}}$ 
  - Can apply any KG embedding on each view
  - Optional: Hierarchical-aware modeling on ontological view specifically for taxonomy meta relations
- Joint training on cross-view loss and intra-view loss

$$J = J_{\text{Intra}} + \omega \cdot J_{\text{Cross}}$$

# JOIE: Experiment Setup

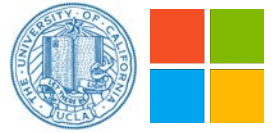


- Datasets: YAGO26K-906 (from YAGO) and DB11K-184 (from DBpedia), ontology-view leveraged from ConceptNet
- Tasks: *Triple completion* and *entity typing*
- Evaluation metrics
  - Triple completion: *MRR*, *Hit@K score* ( $K=1,3,10$ )
  - Entity typing: *Accuracy* (*Hit@1*), *Hit@3 Score*
- Baselines: TransE, DistMult, HoIE, TransC, MTransE



Dataset	Instance Graph $\mathcal{G}_I$			Ontology Graph $\mathcal{G}_O$			Type Links $\mathcal{S}$
	#Entities	#Relations	#Triples	#Concepts	#Meta-relations	#Triples	
YAGO26K-906	26,078	34	390,738	906	30	8,962	9,962
DB11K-174	111,762	305	863,643	174	20	763	99,748

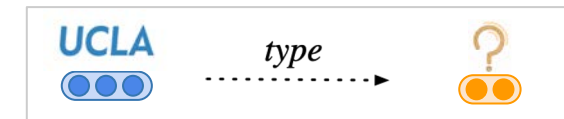
# JOIE: Results



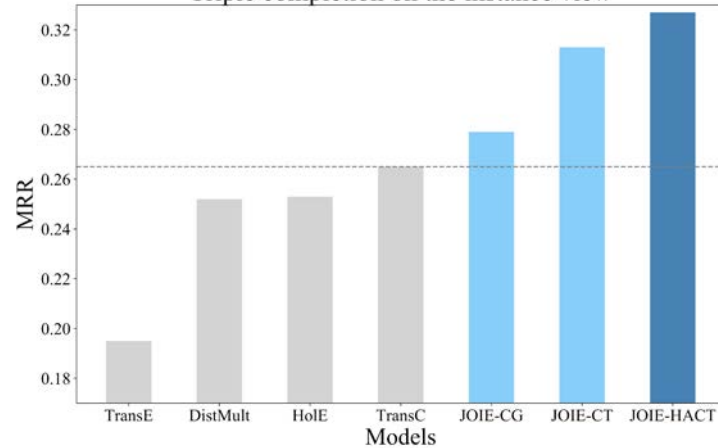
## Triple Completion



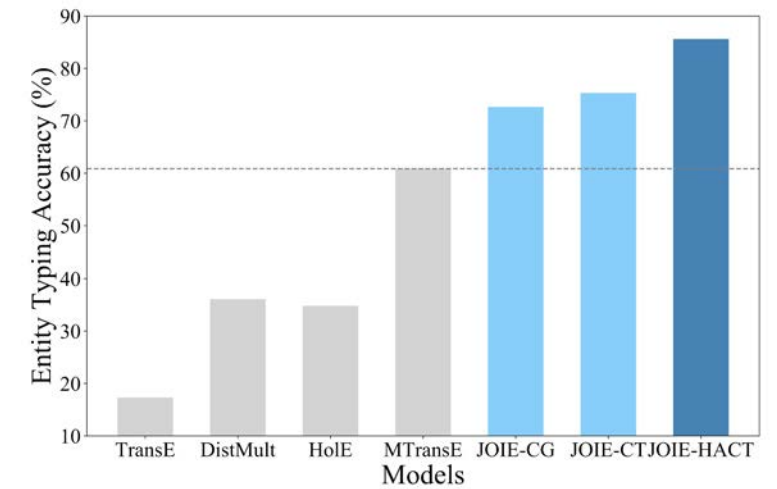
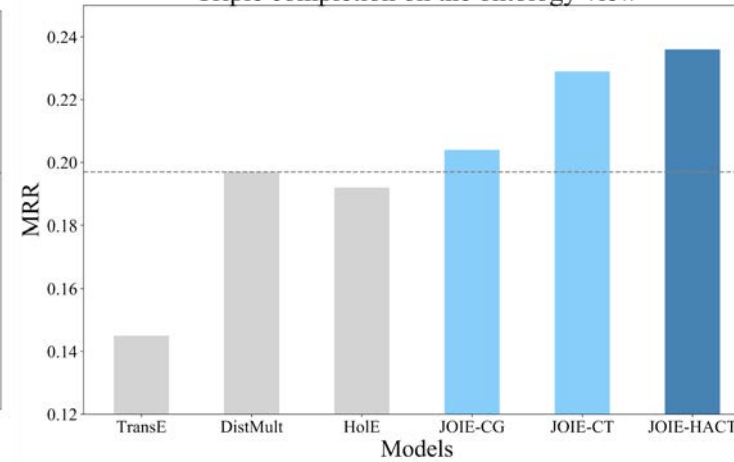
## Entity Typing



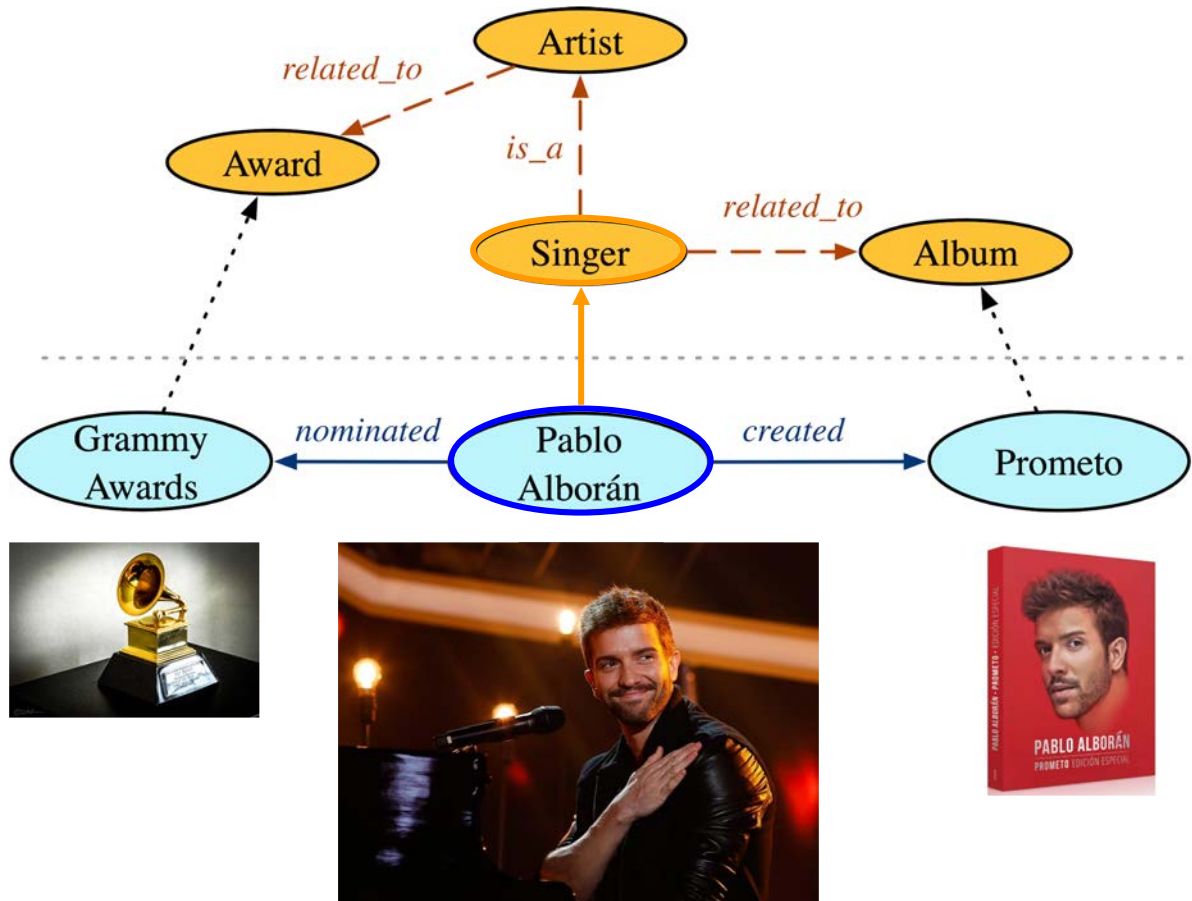
Triple completion on the instance view



Triple completion on the ontology view



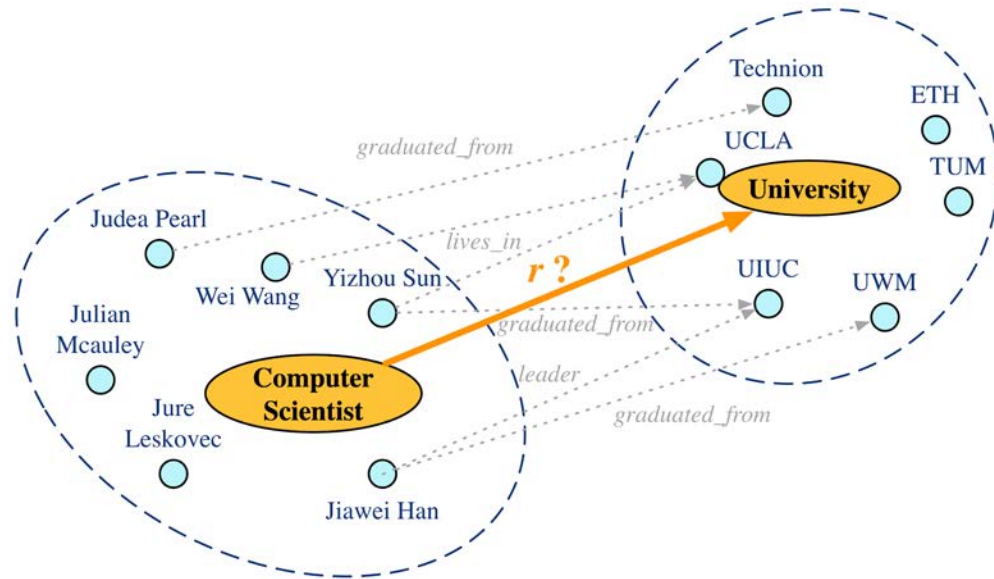
# Case Study: Long-tail Entity Typing



## Example of long-tail entity typing

Entity	Model	Top 3 Concept Prediction
Laurence Fishburne	DistMult MTransE JOIE	football team, club, team writer, <b>person</b> , artist <b>person</b> , artist, philosopher
Warangal City	DistMult MTransE JOIE	country, village, <b>city</b> administrative region, <b>city</b> , settlement <b>city</b> , town, country
Royal Victorian Order	DistMult MTransE JOIE	person, writer, administrative region election, award, <b>order</b> award, <b>order</b> , election

# Case Study: Ontology Population



## Examples of ontology population

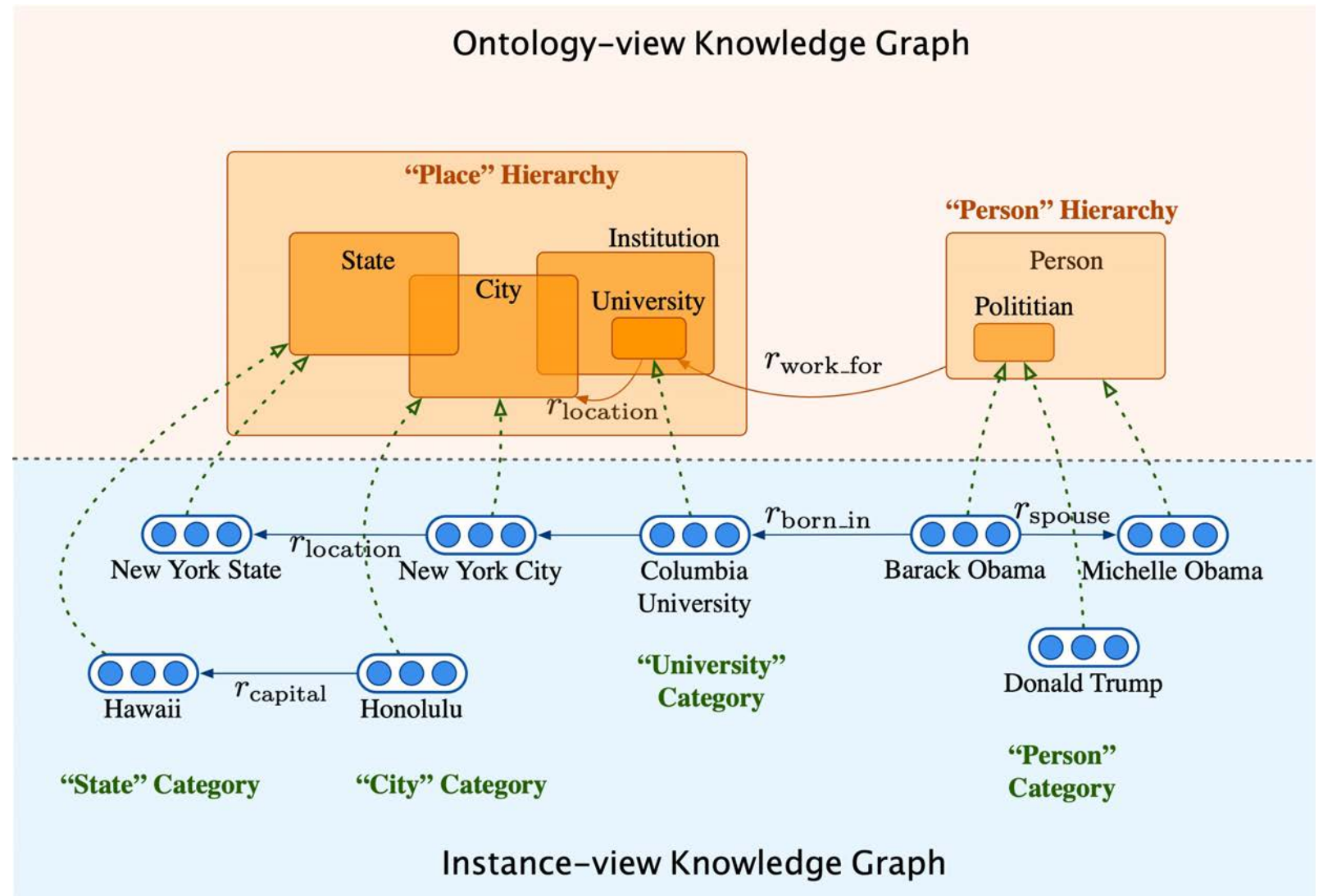
Query	Top 3 Populated Triples with distances
(scientist, ?r, university)	scientist, <i>graduated from</i> , university (0.499) scientist, <i>isLeaderOf</i> , university (1.082) scientist, <i>isKnownFor</i> , university (1.098)
(boxer, ?r, club)	boxer, <i>playsFor</i> , club (1.467) boxer, <i>isAffiliatedTo</i> , club (1.474) boxer, <i>worksAt</i> , club (1.479)
(scientist, ?r, scientist)	scientist, <i>doctoralAdvisor</i> , scientist (0.204) scientist, <i>doctoralStudent</i> , scientist (0.221) scientist, <i>relative</i> , scientist (0.228)



# Extension: JOIE + Ontological Box Embedding



Use box embedding to better capture the hierarchy in ConceptNet (common sense) ontology.

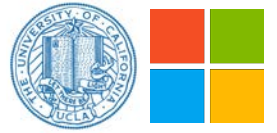


Application 1: KG in Bioinformatics

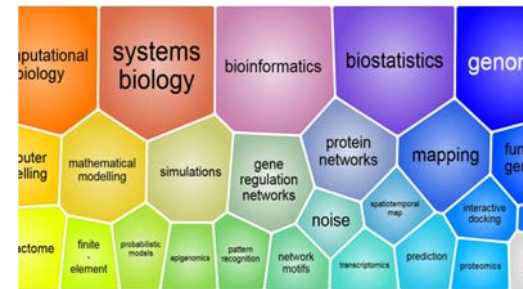
Bio-JOIE: Joint representation learning of  
biological knowledge bases

*A story of protein interaction networks and gene ontology.  
Multiple species, more views, more informational.*

# Application 1: Bio-JOIE

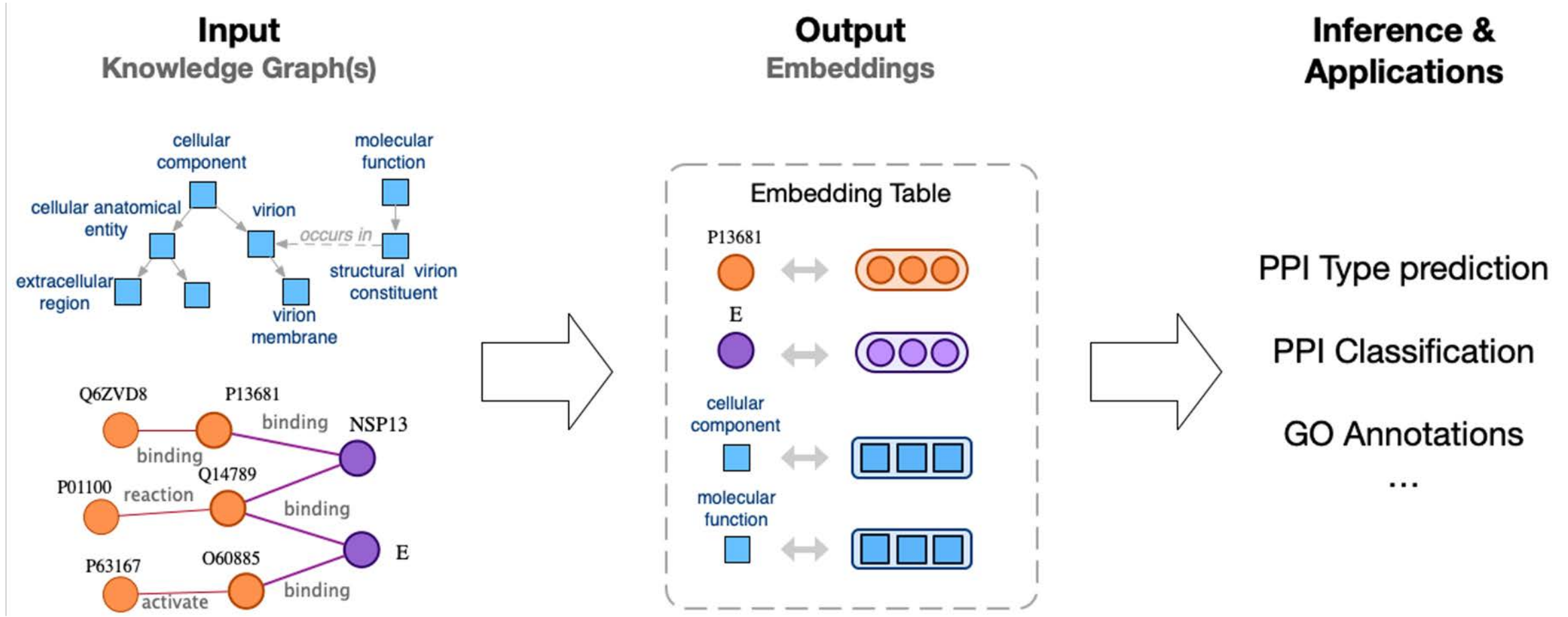
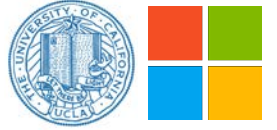


*Knowledge Graphs*

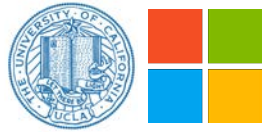


*Computational Biology & Bioinformatics*

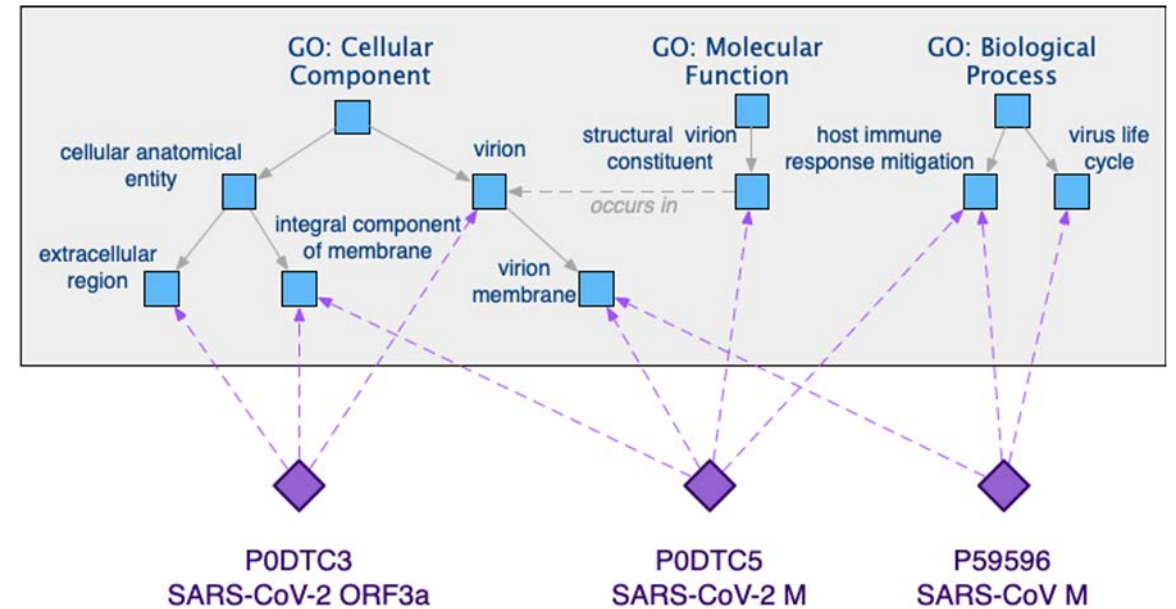
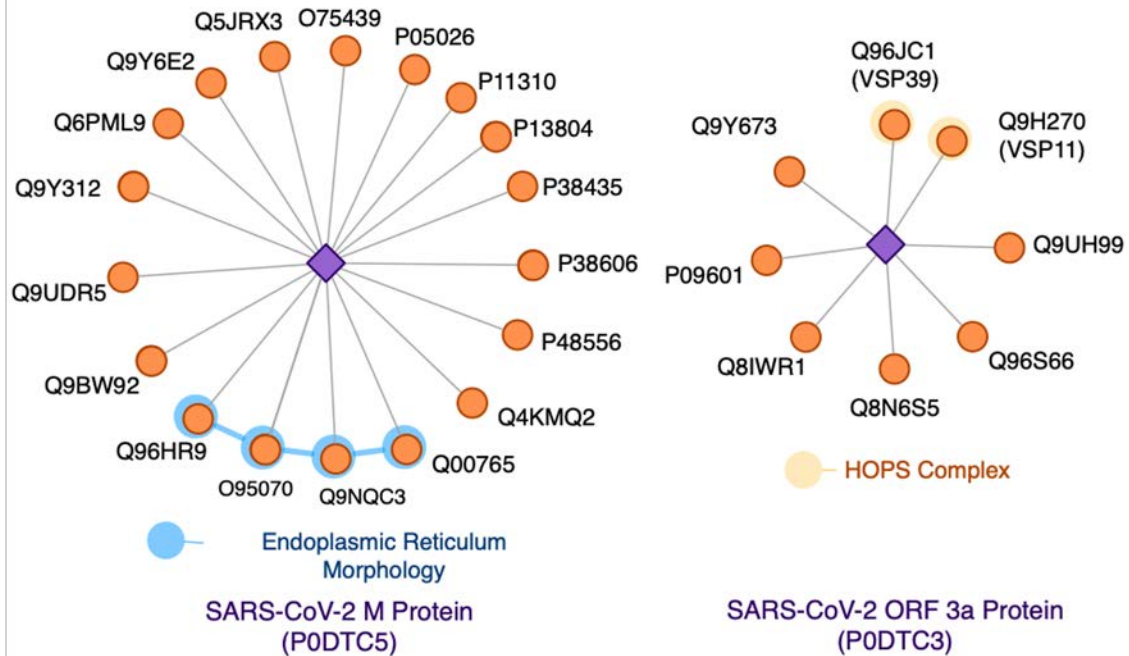
# KG Embedding for Medical Knowledge



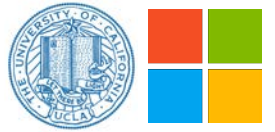
# Similar Ontology-Instance Views in Bioinformatics



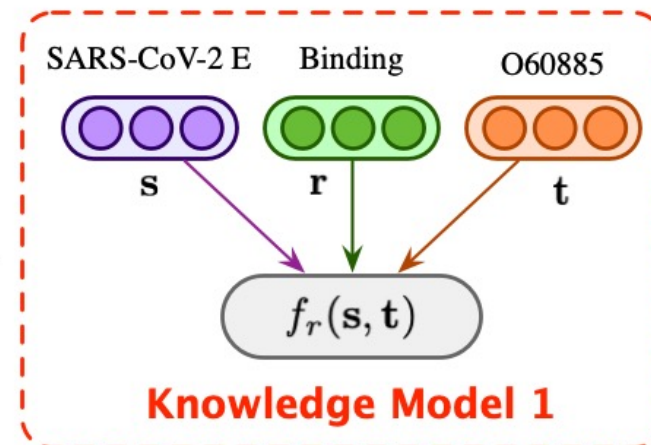
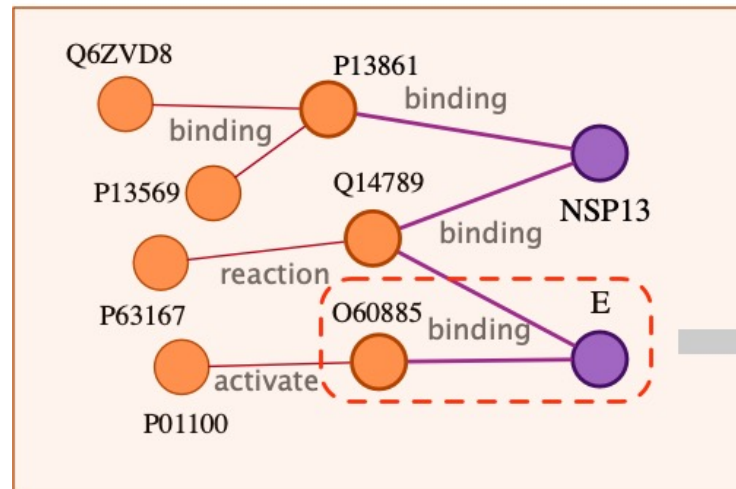
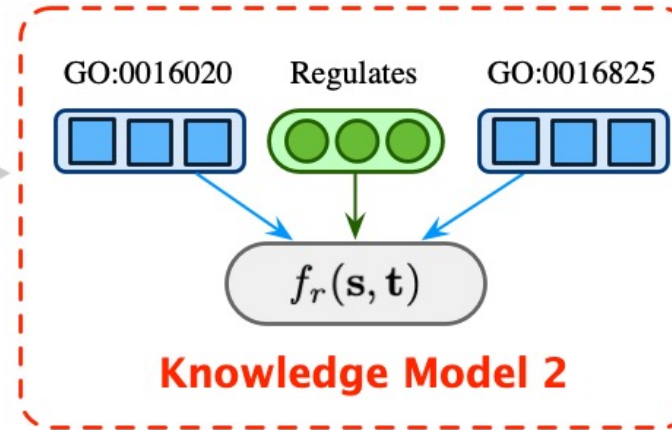
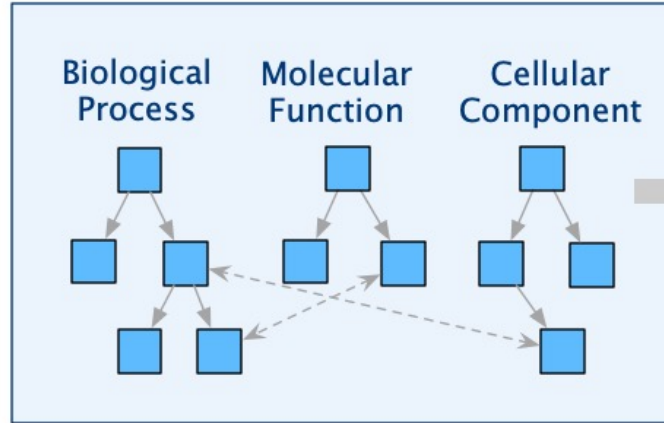
SARS-CoV-2 Human host interactions (Left) and SARS-CoV-2 Gene Ontology (GO) annotations (Right)



# Bio-JOIE: Extension from JOIE

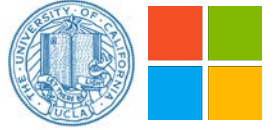


## Gene Ontology Domain

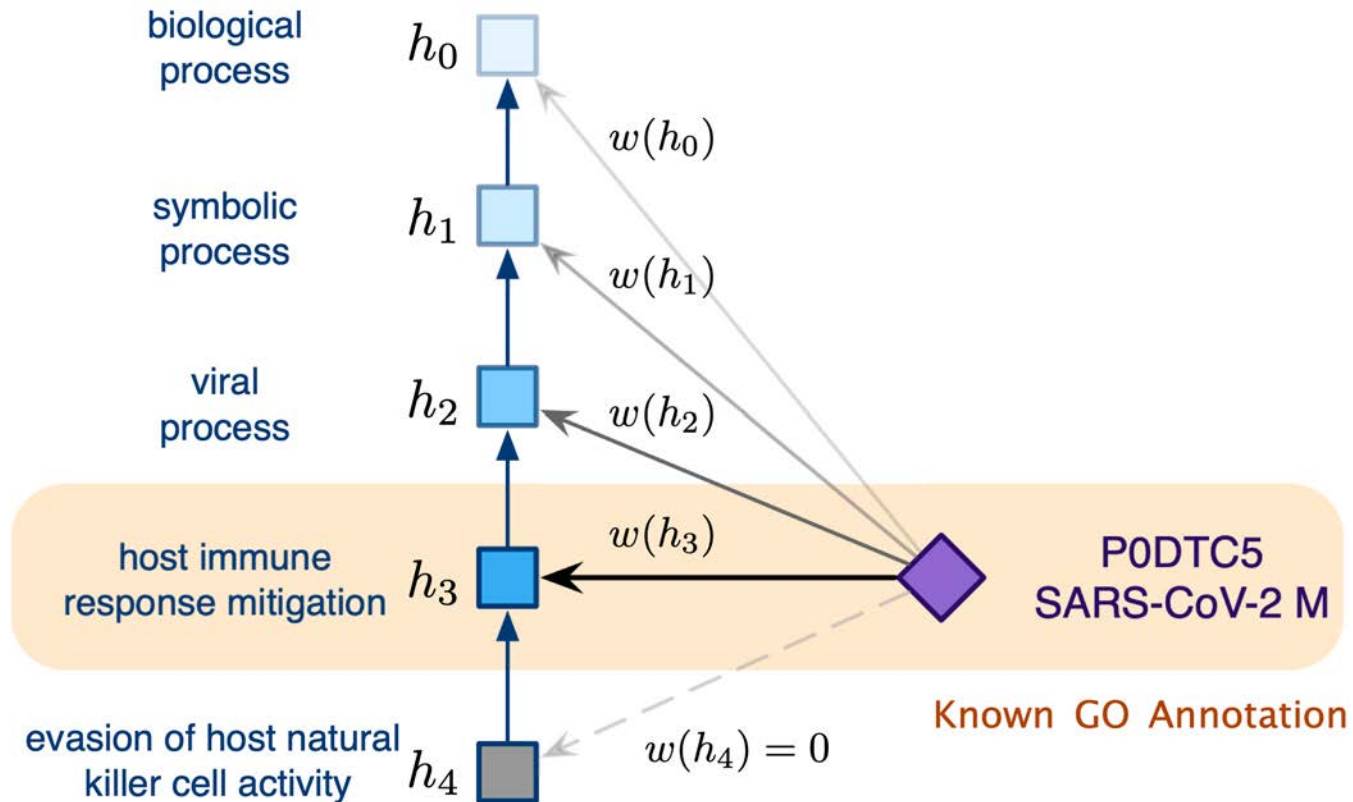


## Protein Interaction Domain

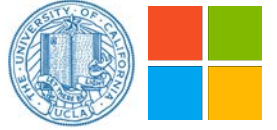
# New in Bio-JOIE: Weighted Alignment



**Intuition:** Assign higher weights to association of protein and a specific GO term compared to a general GO term, in terms of known GO annotations



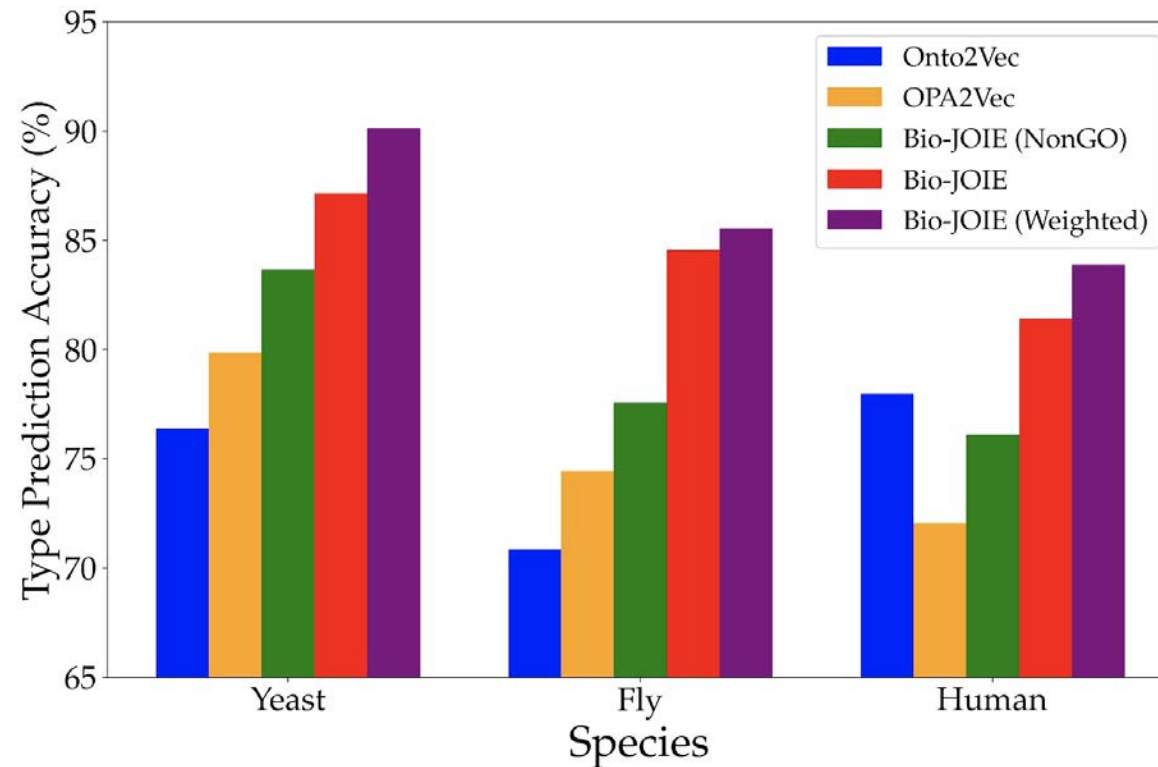
# Bio-JOIE: PPI Prediction



**Task:** Interaction type prediction given pairs of proteins

**Evaluation metric:** Prediction accuracy

**Baselines:** Onto2Vec (variants: Parent, Ancestor, Sum, Mean) , OPA2Vec, Bio-JOIE (NonGO)





# Bio-JOIE: PPI Prediction, different GO aspects

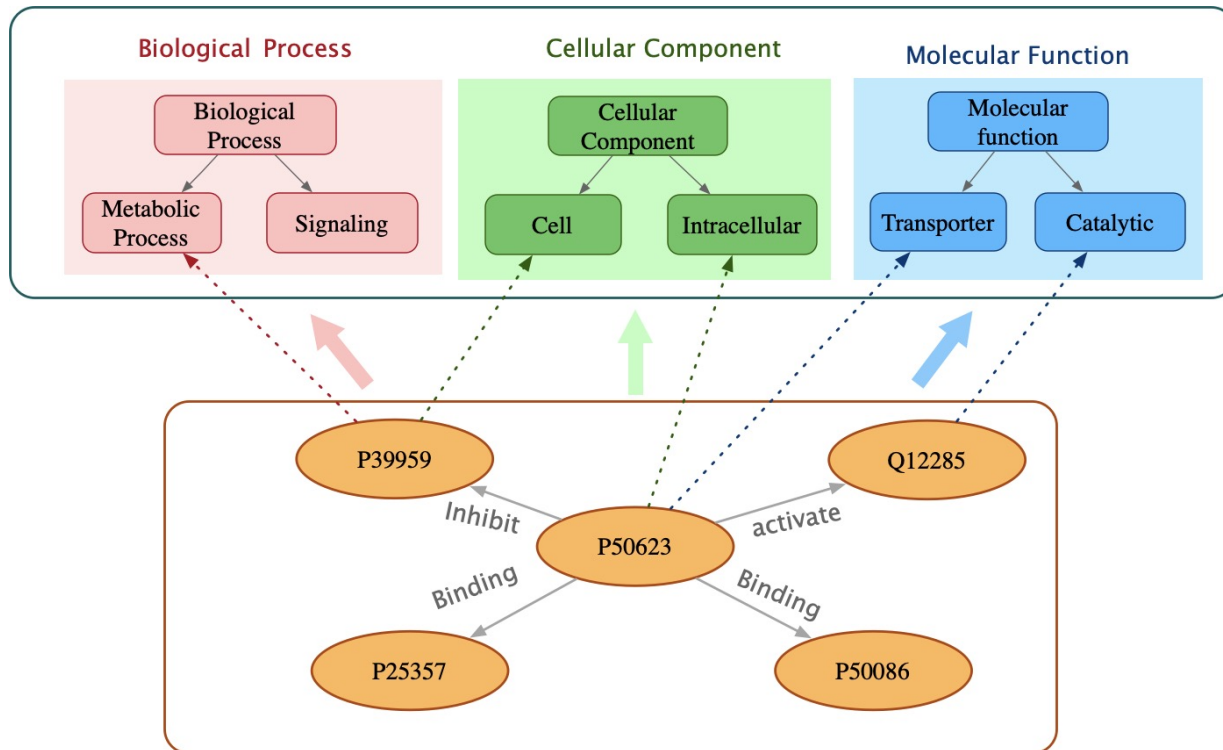
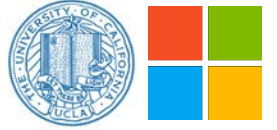


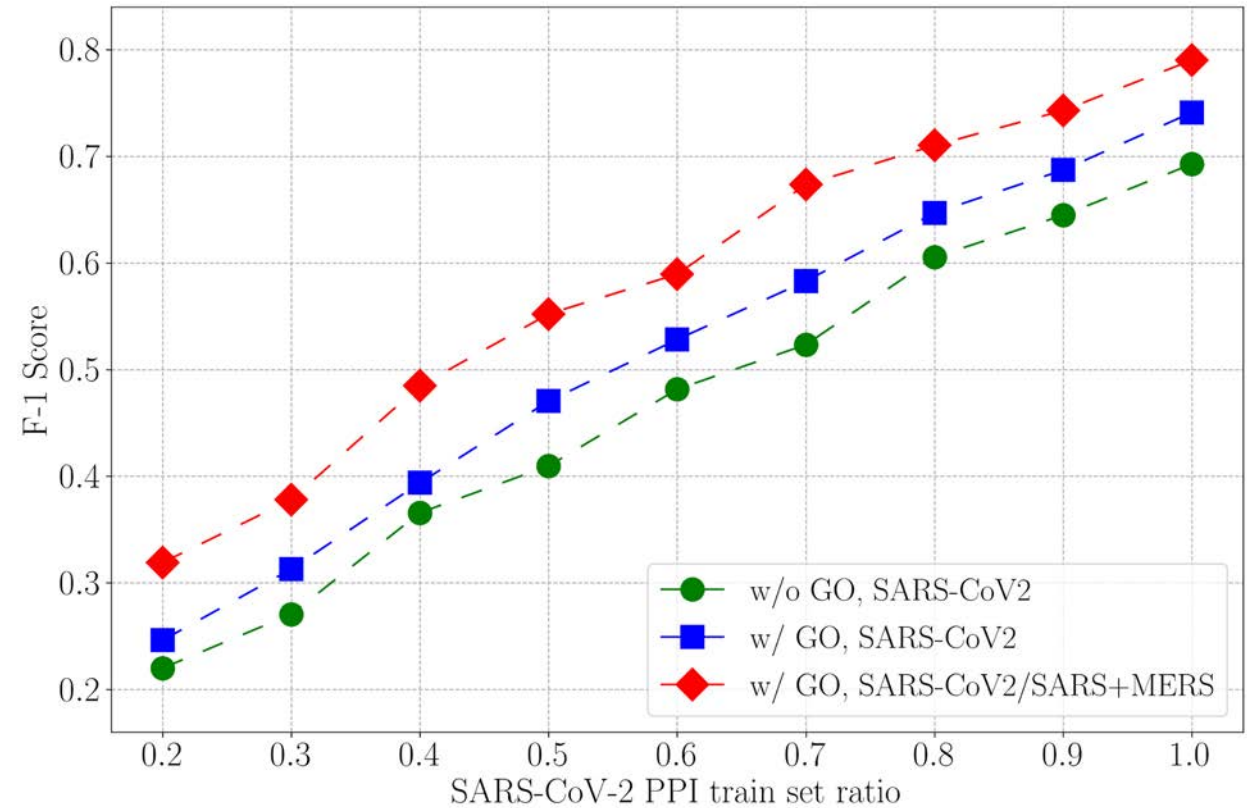
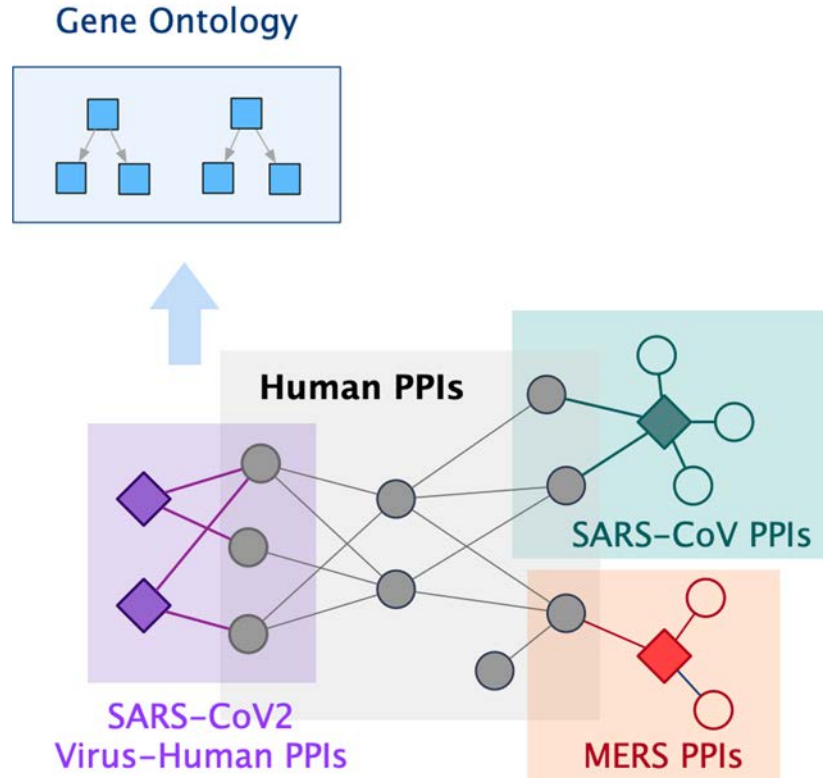
Table: Comparison of Bio-JOIE performance on combinations of three different aspects in GO.

#	Aspects	Yeast	Fly	Human
1	BP	0.8794	0.8402	0.8153
	CC	0.8499	0.8272	0.8054
	MF	0.8539	0.8386	0.8165
2	BP+CC	0.8717	0.8473	0.8271
	BP+MF	0.8673	0.8471	0.8163
	CC+MF	0.8569	0.8466	0.8170
3	AllGO	<b>0.9012</b>	<b>0.8555</b>	<b>0.8389</b>

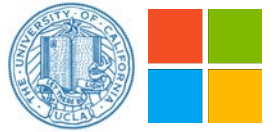
# Experiment: SARS-CoV-2 PPI Classification



**Task:** Virus-human PPI classification by embeddings learned from multiple gene ontology aspects and similar viruses

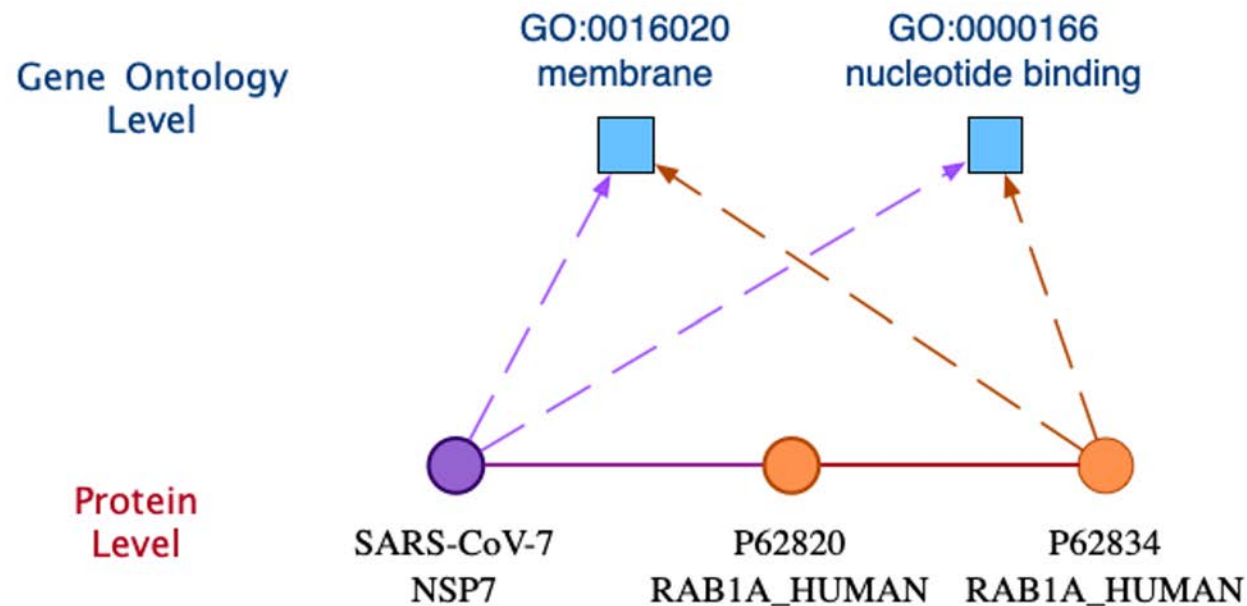


# Experiment: SARS-CoV-2 Target Prediction



SARS-CoV-2 Protein	Top Predicted Human Target Proteins
NSP7	P62834(0.685), P51148(0.879), P62070(0.418), P67870, O14578, Q8WTV0(0.854), P53618(0.350), Q9BS26, O94973, Q7Z7A1

Diving deep into the top-1 prediction:



## Application 2: KG in Recommendation

P-Companion: A principled framework for diversified complementary product recommendation

*How can we manage to jointly learn the instance and ontology?*

# Application 2: Recommender System

---



*Knowledge Graphs*



*Recommender Systems*

# Task: Complementary Recommendation



Home > Bedding > Blankets & Throws > Weighted Blankets

Quality Premium Kids' Weighted Blanket & Pillowcase Set | 12 lbs | 60"x80" | for Kids Between 110-140 lbs | Premium Glass Beads | Grey/Navy Blue  
by Quility  
★★★★★ 9,931 ratings  
#1 Best Seller in Kids' Quilt Sets

Price: \$99.70 ✓prime FREE C  
Get \$70 off instantly: Pay \$29.70 with the Amazon Prime Rewards Visa

Size: 60"x80" | 12lbs

36"x48"   05lbs	41"x60"   10lbs
48"x72"   12lbs	48"x72"   15lbs
60"x80"   15lbs	60"x80"   20lbs
86"x92"   15lbs	86"x92"   20lbs
86"x92"   30lbs	

Color: Grey Cotton Blanket + Navy Blue Pillowcase

- 100% Cotton
- 7-LAYERED PREMIUM BLANKET: The 7-layered blanket is designed to use the most advanced sewing techniques and the highest-quality materials to provide superior comfort and durability.

Roll over image to zoom in

**Added to Cart**  
Cart Subtotal (1 item): \$99.70  
View Cart Proceed to checkout

**Customers who bought this item also bought**

- Quality Premium Adult Removable Duvet Cover for Weighted Blanket | 60"x80" | Full Size Bed | 100% Cotton Cover Fabric | Blue  
★★★★★ 218  
\$31.92 ✓prime  
Add to Cart
- Quality Premium Adult Removable Duvet Cover for Weighted Blanket | 60"x80" | Full Size Bed | 100% Cotton Cover Fabric | Grey  
★★★★★ 218  
\$31.92 ✓prime  
Add to Cart
- Amazon.com Gift Card in a Greeting Card (Various Designs)  
★★★★★ 13,406  
\$10.00 - \$2,000.00  
Choose options

See More

*"How about just buying more? I want to go to the space." said J. Bezos*

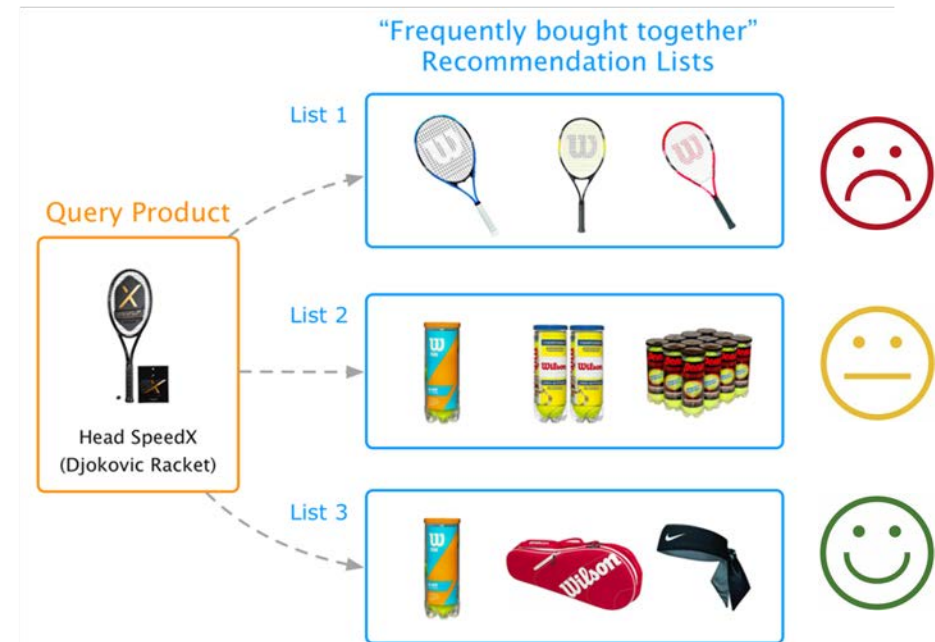
# Task: Complementary Recommendation



Think about one customer who plans to buy a tennis racket (e.g., Head SpeedX Djokovic racket).

What would you recommend for him to purchase together?

- List 1: three more tennis rackets? → **Sorry, we are not looking for substitutes!**
- List 2: three sets of tennis balls? → Hmm, not bad, but only need one is good enough. Can we do better?
- List 3: one tennis ball pack, one bag and one headband? → **Sound good this time!**



**Frequently bought together**

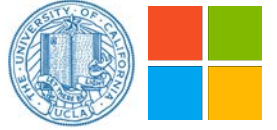
Total price: **\$92.97**

[Add all three to Cart](#)

[Add all three to List](#)

- ✓ **This item:** Cooling Shredded Memory Foam Bed Pillow for Sleeping- Adjustable to Thick Thin - Pillow for Side... **\$34.99**
- ✓ Beckham Hotel Collection Gel Pillow (2-Pack) - Luxury Plush Gel Pillow - Dust Mite Resistant... **\$34.99**
- ✓ Elegear Cooling Pillowcases for Night Sweats and Hot Flashes, Japanese Q-Max 0.4 Cooling Fiber... **\$22.99**

# Problem Definition



- Given the input as catalog features (including item type) and customers behavior data, for a query item  $i$ , we recommend a set of items  $S(i)$ , aiming at optimizing their co-purchase probability and recommendation diversity.



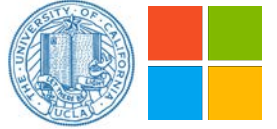
query item  $i$



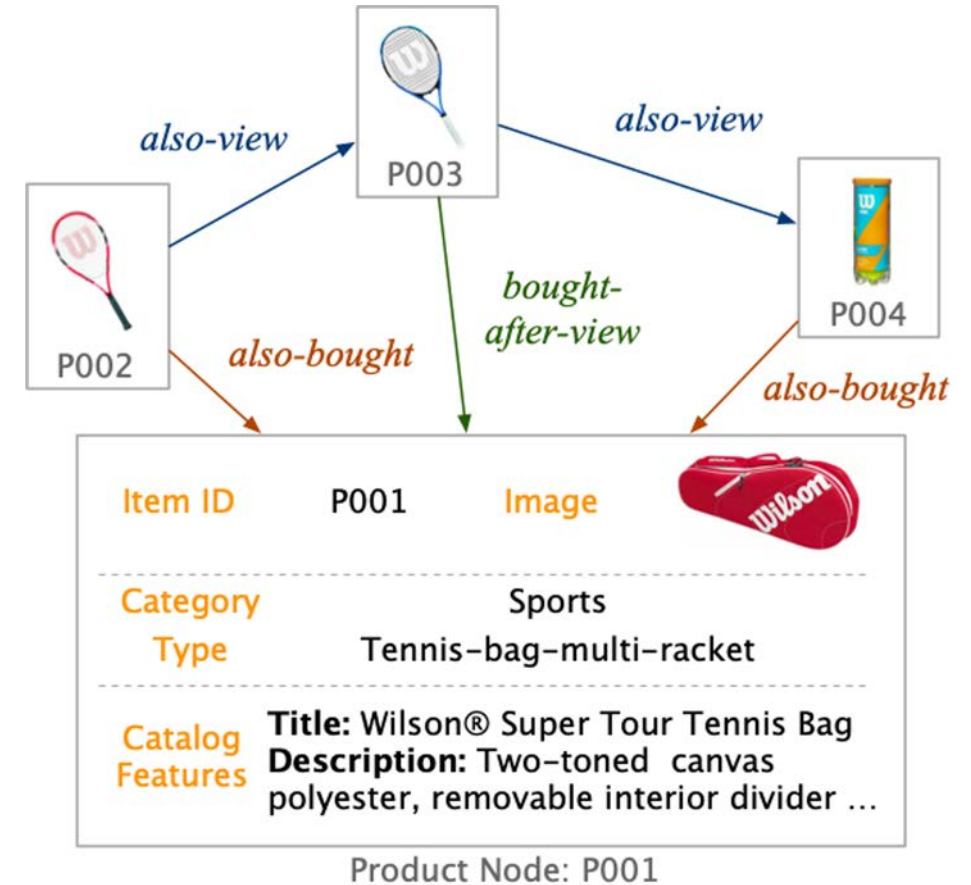
Recommendation set  $S(i)$



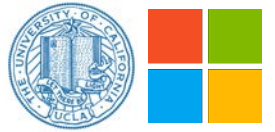
# "Behavior-based" Product Graphs (PG)



- Behavior based product graph → Attributed heterogeneous information networks (KG)
- **Node:** Product items with attributes (title, description, category, keywords)
- **Edges:** Customer browsing/purchase behaviors (such as **also-bought**, **also-view**, **bought-after-view**, as important indicators of **substitutes** or **complements**)
- Note that there are many alternative ways to construct product graphs, with different modeling goals.



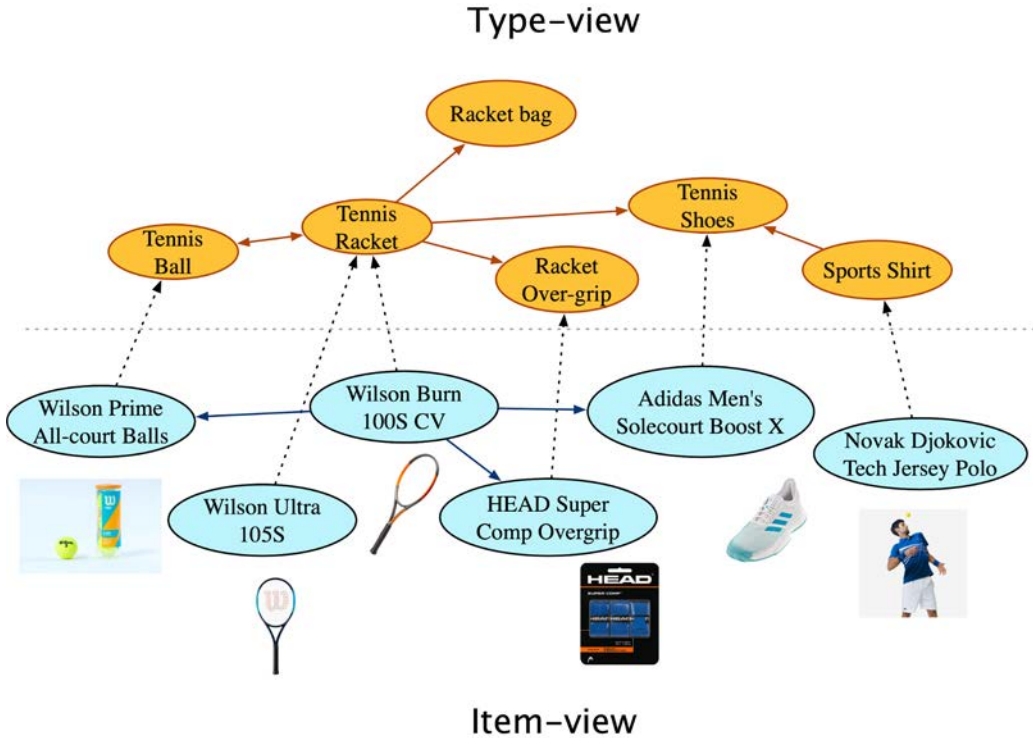
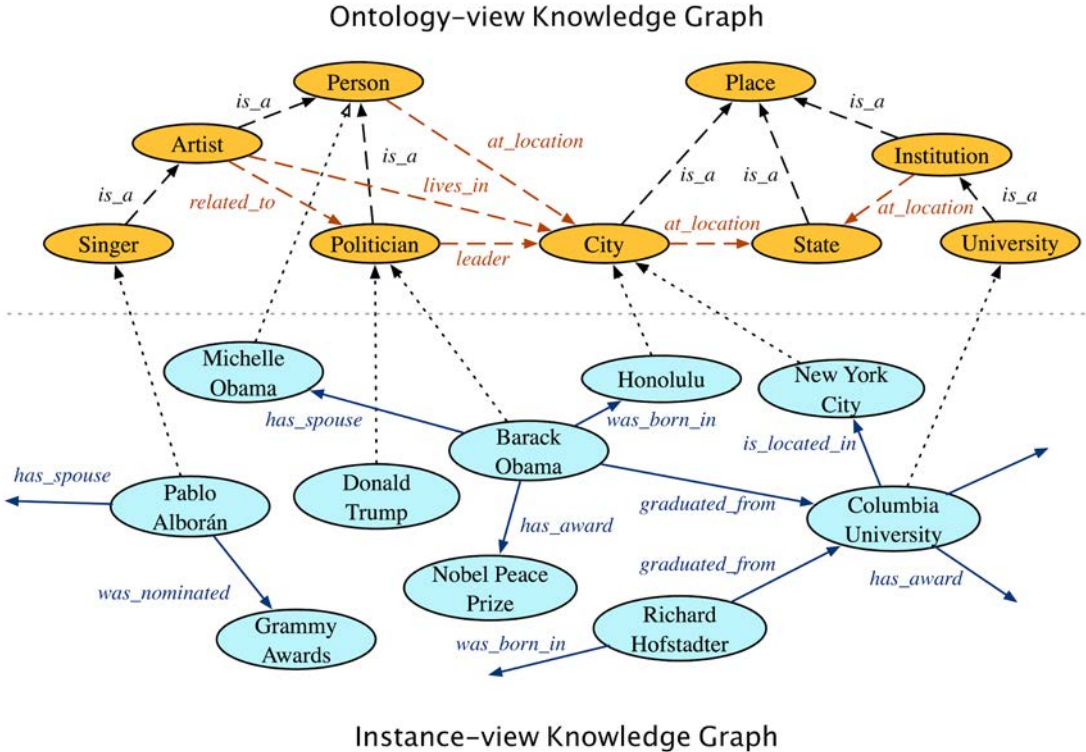
# \*Quick Comparison of KG and PG



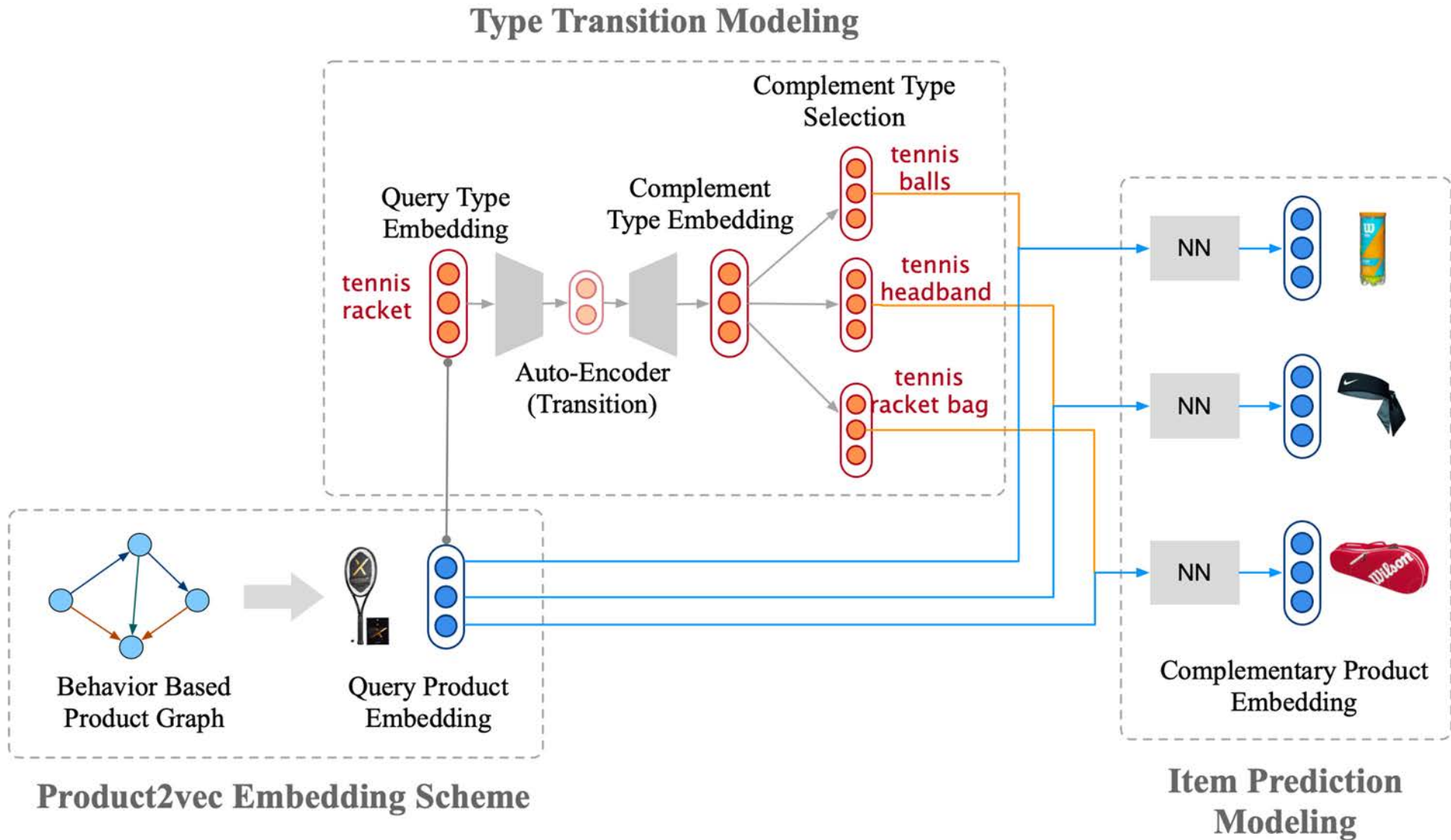
Comparison	Knowledge Graphs	Product Graphs
<i>Source</i>	Established facts	Product catalog, use-product interaction
<i>Quality</i>	Observed facts are well-established and plausible.	Much noisier
<i>Quantity of relations</i>	Typically, thousands of possible relations in real world, such as <code>born_in</code> , <code>director_of</code> , etc	A few relations defined from specified user behavior, such as <code>also_view</code> and <code>also_bought</code>
<i>Attributes</i>	Entity types, numerical features, descriptions, and many other additional features	
<i>Logic rules</i>	Available for logical inference and refinement.	Possibly a few rules. Similar products may have similar complements.
<i>Downstream tasks</i>	Knowledge completion, relation extraction, question answering, etc.	Recommendation, searching, personalization, etc.

# Connecting KG to PG

Product item to product type relation in PG is like entity-concept association in KG.



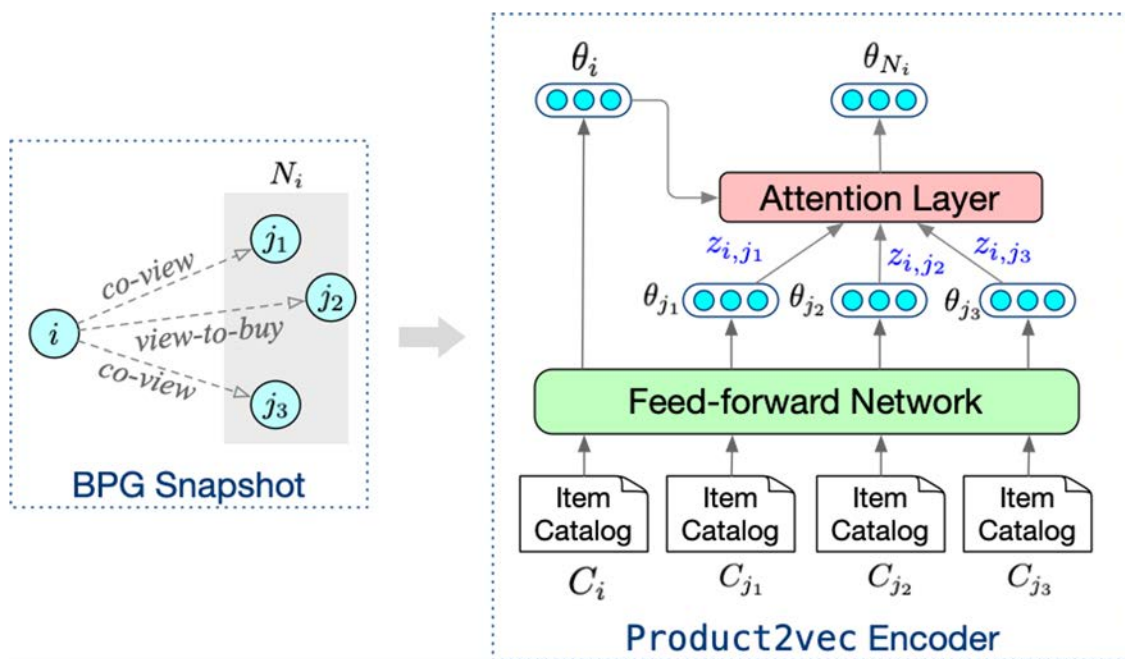
# Product Companion: Workflow



# Base Module: Product2vec



- GNN-based product representation learning framework
- FNN transforms the original textual features to latent embeddings and later aggregate the information from similar products selectively by the attention layer.



## FNN Model:

$$\theta_i = FFN(C_j) = \sigma \left( \sigma \left( C_i W^{(1)} + b^{(1)} \right) W^{(2)} + b^{(2)} \right) W^{(3)} + b^{(3)}$$

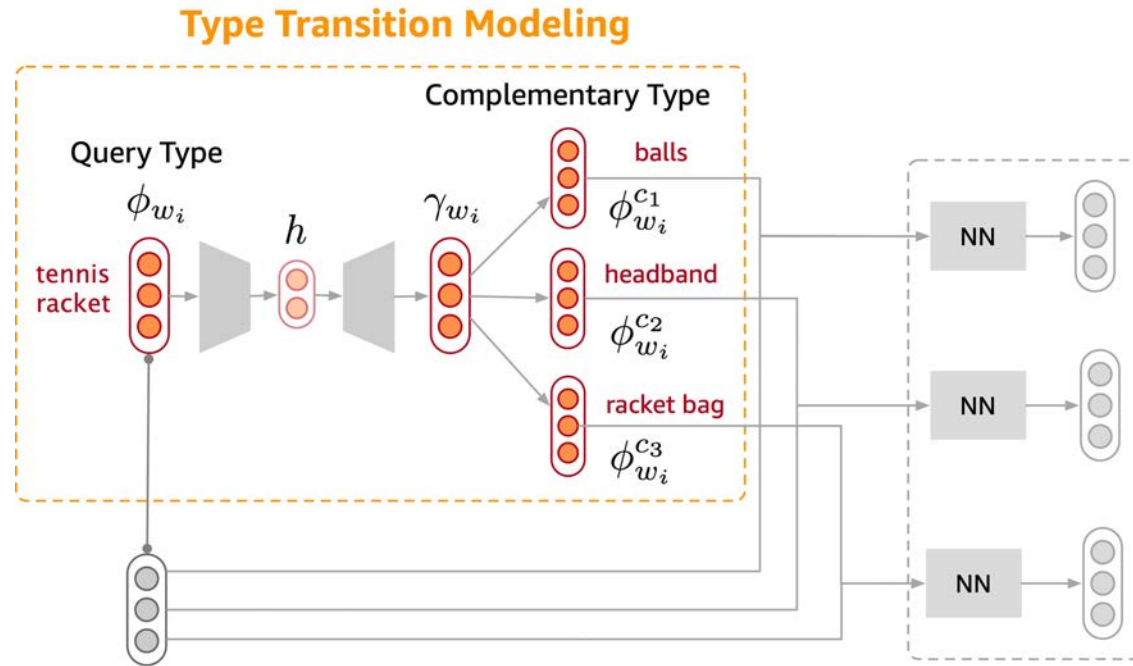
## Attention Weight:

$$z_{i,j} = \text{softmax}_j (\theta_i^T \theta_j) = \frac{\exp(\theta_i^T \theta_j)}{\sum_{j' \in N_i} \exp(\theta_i^T \theta_{j'})}$$

## Product2Vec training loss:

$$L = \sum_{i \in \mathcal{I}} \sum_{y \in \{\pm 1\}} \left\{ \max \left( \epsilon - y \cdot \left( \lambda - \|\theta_i - \theta_{N_i}\|_2^2 \right) \right) \right\}$$

# Module 2: Complementary Type Transition



Auto-encoder based type transition model:

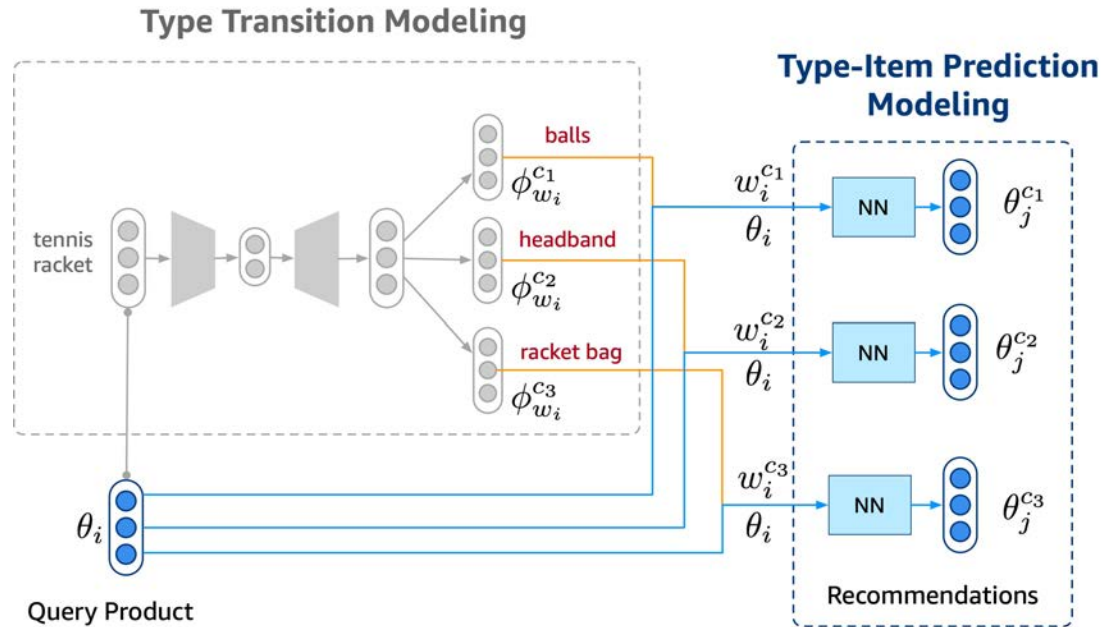
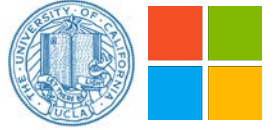
$$h = \text{Dropout} \left( \text{ReLU} \left( \phi_{w_i} W^{(4)} + b^{(4)} \right) \right)$$

$$\gamma_{w_i} = h W^{(5)} + b^{(5)}$$

Training loss:

$$\min \sum_{i,j \in \mathcal{T}} \left( \max \left\{ 0, \epsilon_w - y_{i,j} \left( \lambda_w - \|\gamma_{w_i} - \phi_{w_j}^c\|_2^2 \right) \right\} \right)$$

# Module 3: Complementary Item Prediction



Item prediction neural model:

$$\theta_i^{w_c} = \theta_i \odot (\phi_{w_c}^c W^{(6)} + b^{(6)}),$$

$$s.t., \|\phi_{w_c}^c - \gamma_{w_i}\|_2^2 \leq \beta$$

Training loss:

$$\min \sum_{i,j \in \mathcal{T}} \max \{0, \epsilon_i - y_{i,j} (\lambda_i - \|\theta_i^{w_c} - \theta_j\|_2^2)\}$$

# Joint Training and Inference

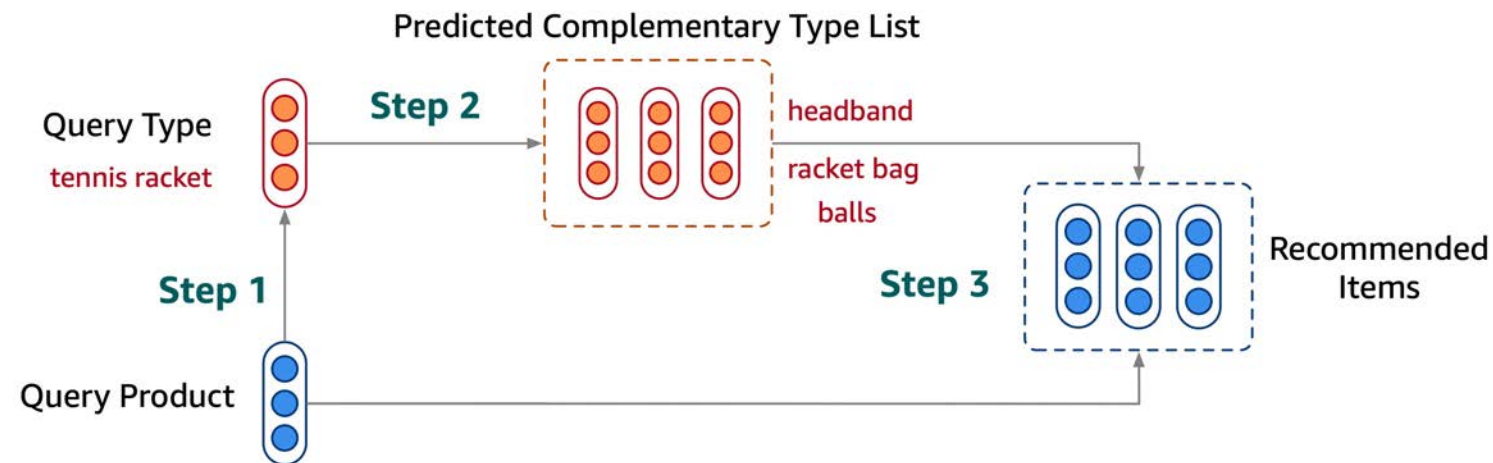


Joint training on type transition and item prediction:

$$\min \sum_{i,j \in \mathcal{T}} \alpha \left( \max \left\{ 0, \epsilon_i - y_{i,j} \left( \lambda_i - \|\theta_i^{w_j} - \theta_j\|_2^2 \right) \right\} \right) + (1 - \alpha) \left( \max \left\{ 0, \epsilon_w - y_{i,j} \left( \lambda_w - \|\gamma_{w_i} - \phi_{w_j}^c\|_2^2 \right) \right\} \right)$$

Item prediction loss Type transition loss

Inference stage:

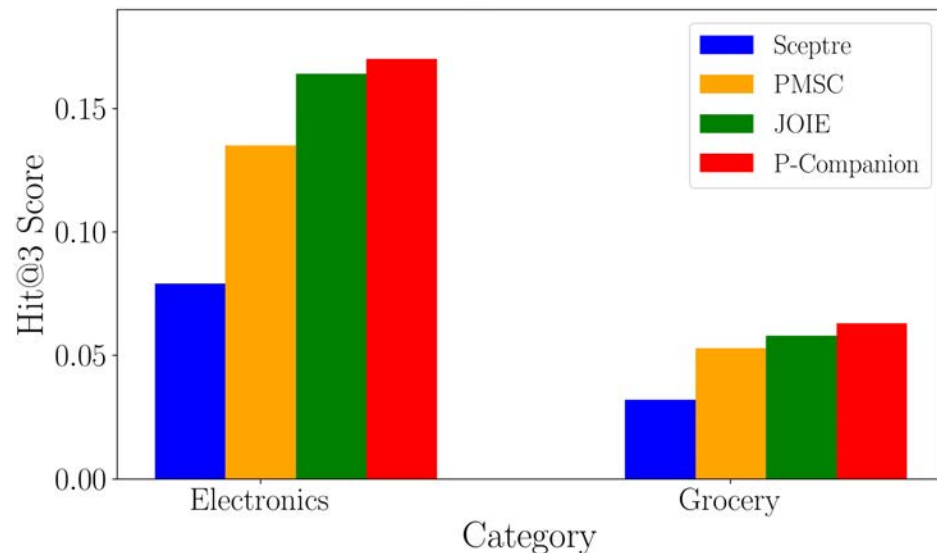




# Evaluation: From history purchase data

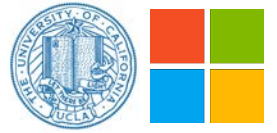


- Given a pair  $(i, j)$ , associated with type  $w_i$  and  $w_j$ , from co-purchase record as ground truth, we ask our model as well as all baselines to output recommendation list (with predicted complementary types), and consider the following:
  - whether item  $j$  is in the list.  $\rightarrow$  **Item level**
  - Whether type  $w_j$  is in the predicted types  $\rightarrow$  **Type level**
- Metric: Hit@K score (both item level and type level, if applicable)
- Baselines: Sceptre, PMSC, JOIE



Dataset		Electronics	Grocery
Model & Setting		Hit@60	Hit@60
Sceptre		0.124	0.085
PMSC		0.179	0.139
JOIE		0.200	0.155
P-Companion	1 type $\times$ 60 items	0.138	0.088
	3 types $\times$ 20 items	0.198	0.153
	5 types $\times$ 12 items	0.222	<b>0.189</b>
	6 types $\times$ 10 items	<b>0.227</b>	0.187

# Case Study: Type Transition Prediction



- Examples of Predicted Top-3 Complementary Type Predictions

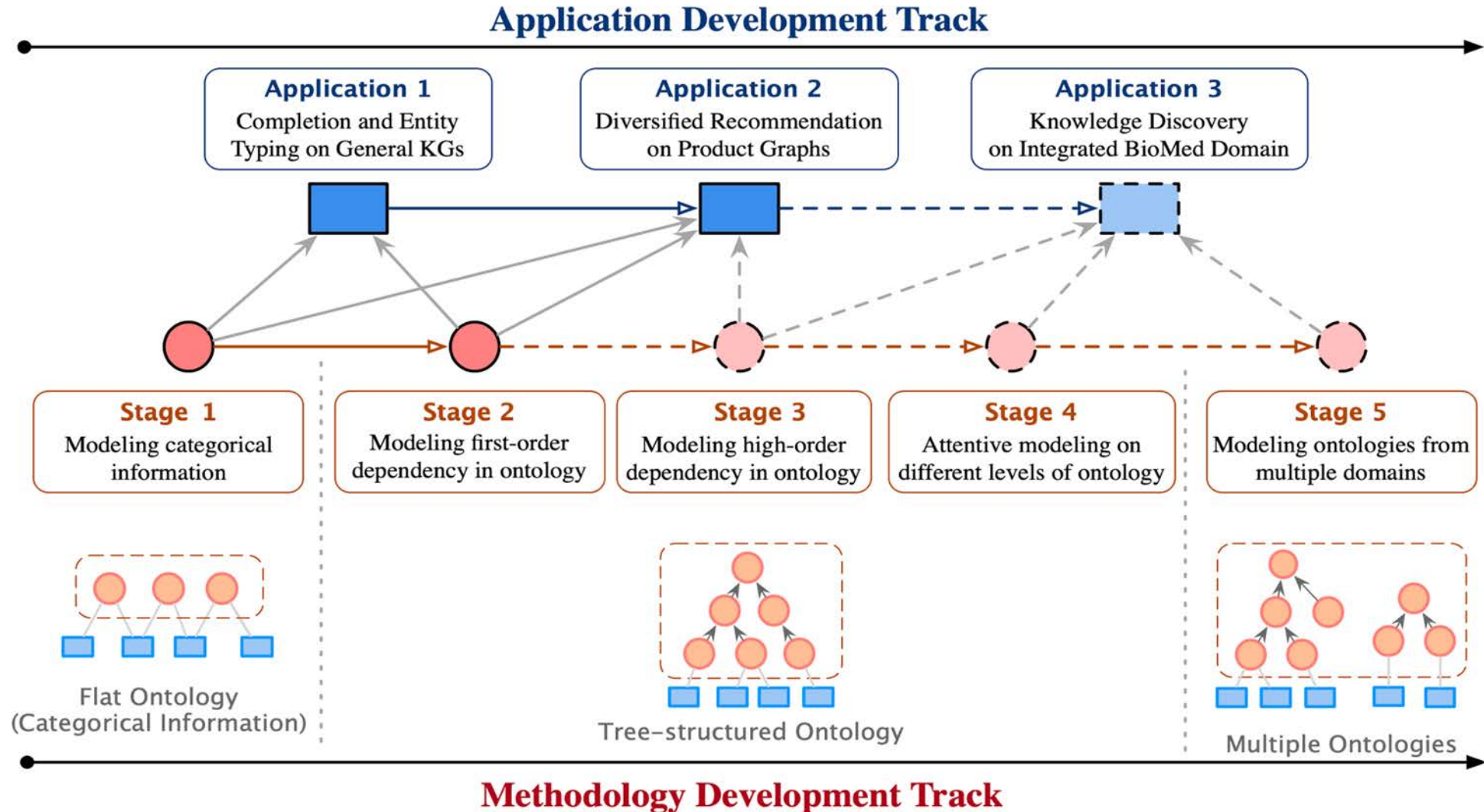
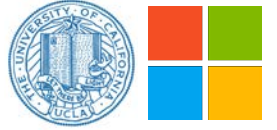
Query Type	Predicted Complementary Types
camera-power-adapter	(1) sec-digit-card (2) micro-sd-card (3) hdmi-cable
cell-phone-battery	(1) cell-phone-screen-protect (2) battery-charge-case (3) flip-cell-phone-carry-case
roast-coffee-bean	(1) fridge-coffee-cream (2) whole-bean (3) white-tea
fly-fish-line	(1) fluorocarbon-fish-line (2) surf-fish-rod (3) fly-fish-reel

# Case Study: Product Recommendation



Category	Query Item	Co-Purchase	Top-5 Recommendations from P-Companion
Electronics			    
Grocery			     
All-Group (Pet home)		None	    
All-Group (Fishing tools)		None	    

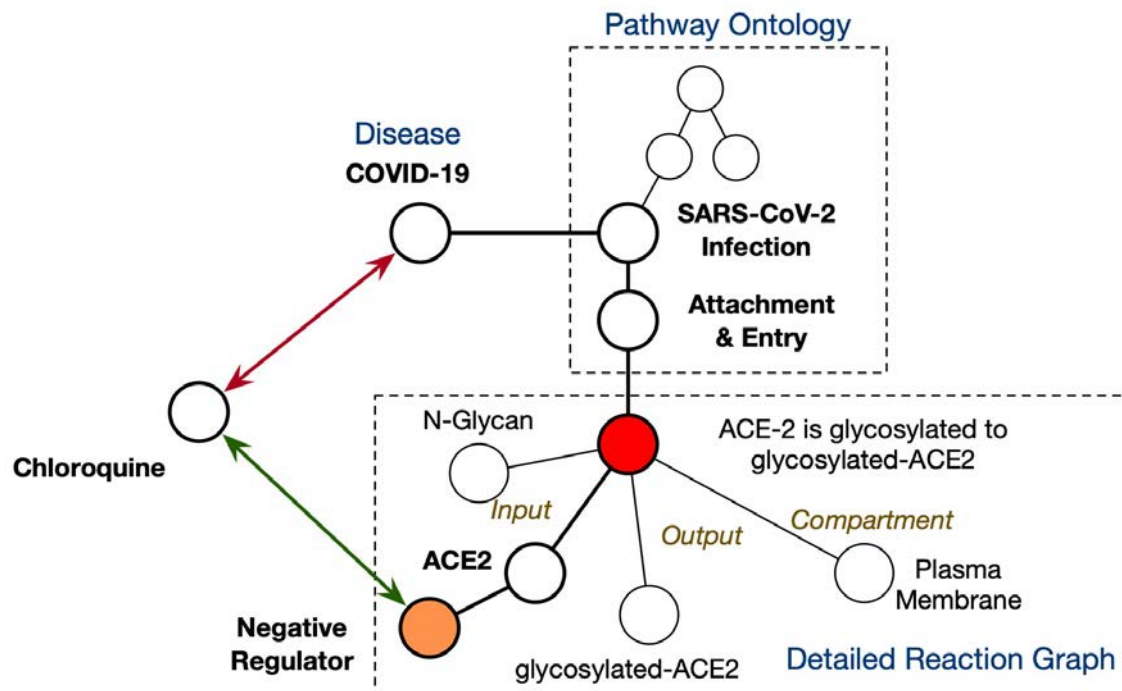
# Research Development Map (Ongoing)



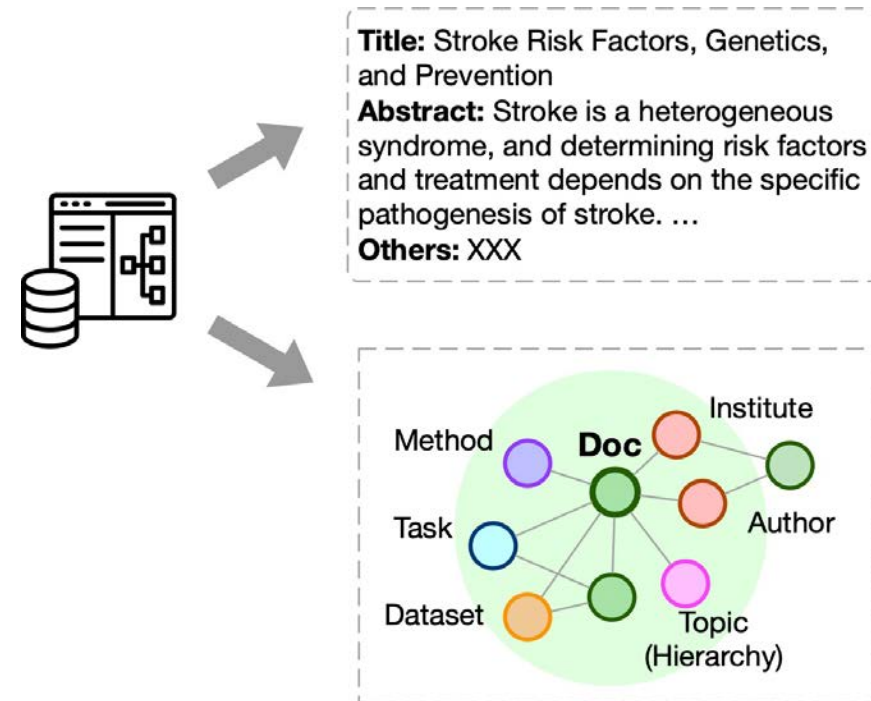
# KG Representation with Graph Attributes



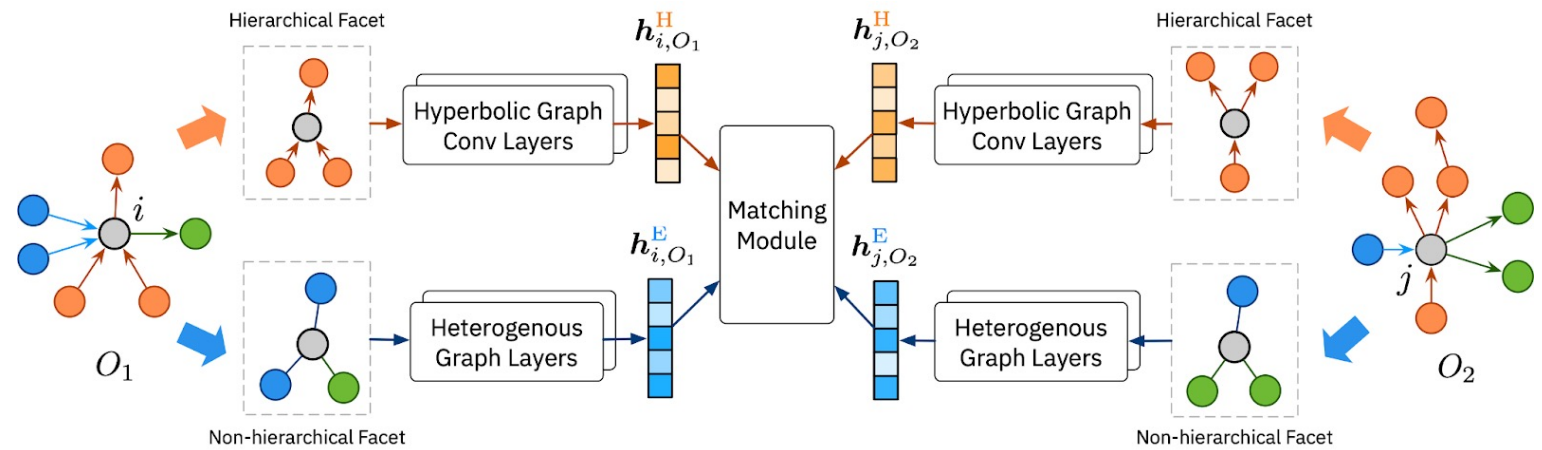
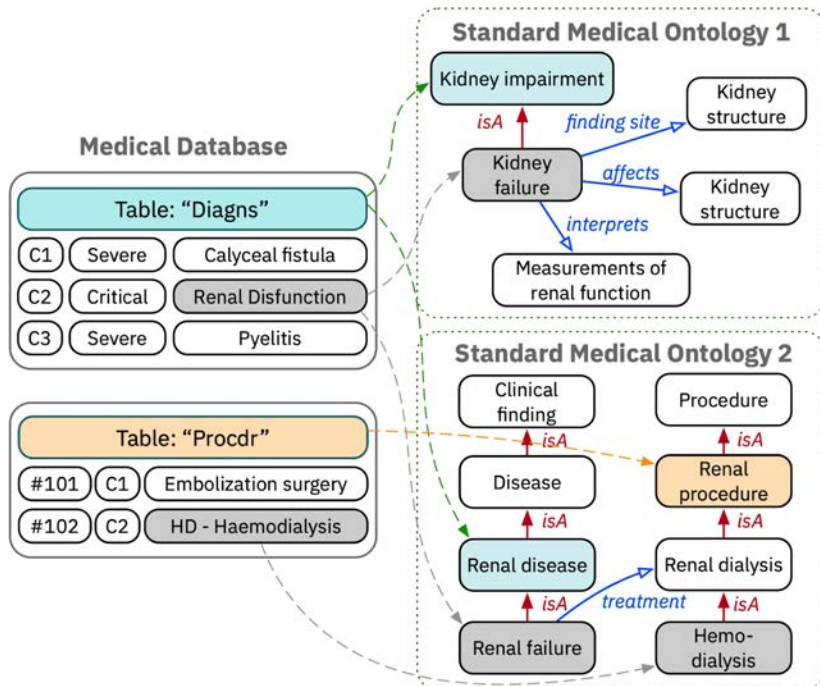
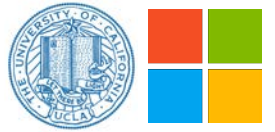
Example 1: Reaction Graph in Bio-KG



Example 2: Document Structured Knowledge



# Hybrid-GNN Data-Ontology Matching



# Summary

- Knowledge graphs often have ontological information, which is important for learning and inference but sparsely investigated.
- Joint learning on the instance and ontology views improves the KG embeddings. That is, incorporating ontologies in KGs is beneficial.
- Ontology-enhanced KG modeling can be applied in a wide selection of interdisciplinary applications, such as protein-protein interaction prediction in bioinformatics and diversified product recommendation in recommender systems.
- Graph neural networks have shown as a power tool on KG as relational data and graph-related downstream tasks, such as node classification, link prediction.

# Collaborators

Muhao Chen (USC ISI)

Chelsea J.-T. Ju (Amazon)

Wenchao Yu (NEC Labs)

Yizhou Sun (UCLA)

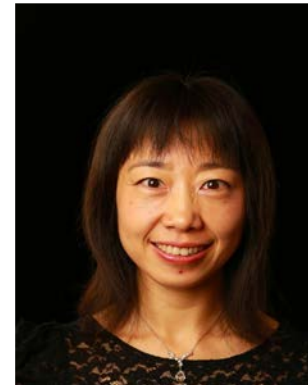
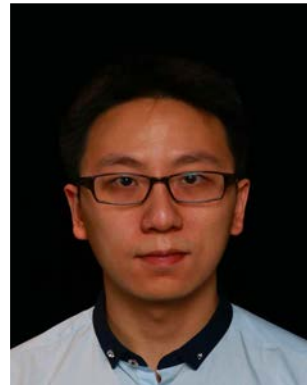
Carlo Zaniolo (UCLA)

Wei Wang (UCLA)

Tong Zhao (Amazon)

Luna Xin Dong (Facebook)

Christos Faloutsos (CMU,  
Amazon)







UCLA

Samueli  
Computer Science

Thank you!

Contact: [jhao@cs.ucla.edu](mailto:jhao@cs.ucla.edu), or [t-junhenghao@microsoft.edu](mailto:t-junhenghao@microsoft.edu)

Website: <http://www.haojunheng.com/>