빅데이터 지능 연구실

Big Data Intelligence Lab

KAIST AI컴퓨팅학과 황지영

jjwhang@kaist.ac.kr

KAIST BDILab BIG DATA INTELLIGENCE

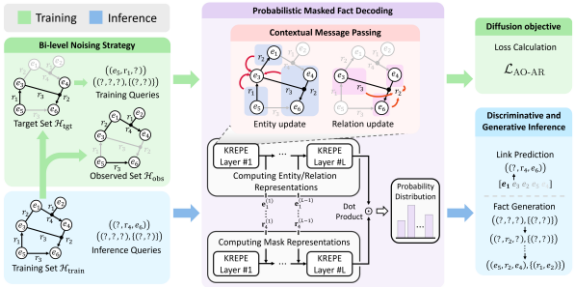
01 황지영 (Joyce Jiyoung Whang)

주요 이력	2026. 05. ~	KAIST AI컴퓨팅학과 부교수
	2023. 09. ~ 2026. 04.	KAIST 전산학부 부교수
	2020. 07. ~ 2023. 08.	KAIST 전산학부 조교수
	2016. 03. ~ 2020. 06.	성균관대학교 컴퓨터공학과 조교수
	2010. 08. ~ 2015. 12.	컴퓨터과학 박사 학위 취득 (Ph.D. in Computer Science) 미국 텍사스 오스틴 대학교 (University of Texas at Austin)
주요 연구분야	그래프 기계학습, 인공지능, 데이터 마이닝, 빅데이터 분석	
주요 강의	KAIST: <CS376: 기계학습>, <CS471: 그래프 기계학습 및 마이닝>, <CS665: 고급 데이터마이닝>	
빅데이터 지능 연구실 (BDI Lab)	개체 간 관계를 노드와 엣지로 표현하는 그래프를 기반으로 하는 그래프 기계학습 및 마이닝 기술 개발 지식 그래프, 그래프 신경망, 대규모 언어 모델 등 인공지능 핵심 원천기술들에 대한 연구 수행 인공지능, 기계학습, 데이터 마이닝, 자연어처리 등 다양한 분야의 연구를 수행	

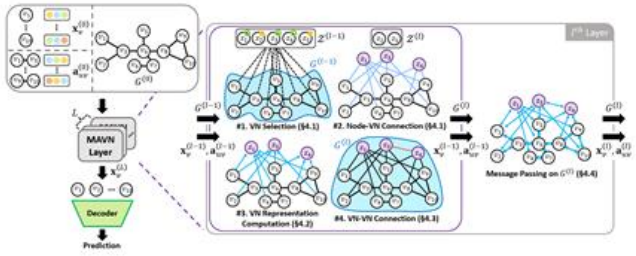
01 빅데이터 지능 연구실

KAIST BDILab BIG DATA INTELLIGENCE

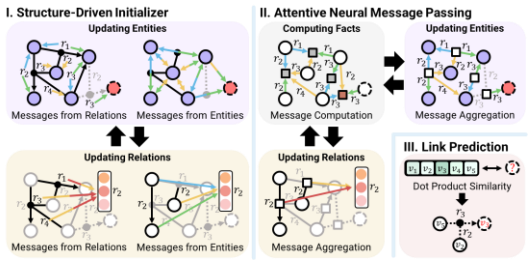
- Selected Publications
  - Learn When and Where to Connect: Adaptive Virtual Nodes for Dynamic Message Passing on Graphs (KDD 2026)
  - Generative Representation Learning on Hyper-relational Knowledge Graphs via Masked Discrete Diffusion (ICML 2026)
  - Beneath the Facade: Probing Safety Vulnerabilities in LLMs via Auto-Generated Jailbreak Prompts (Findings of EMNLP 2025)
  - Structure Is All You Need: Structural Representation Learning on Hyper-Relational Knowledge Graphs (ICML 2025)
  - Stability and Generalization Capability of Subgraph Reasoning Models for Inductive Knowledge Graph Completion (ICML 2025)
  - Unveiling the Threat of Fraud Gangs to Graph Neural Networks: Multi-Target Graph Injection Attacks against GNN-Based Fraud Detectors (AAAI 2025)
  - PAC-Bayesian Generalization Bounds for Knowledge Graph Representation Learning (ICML 2024)
  - Why So Gullible? Enhancing the Robustness of Retrieval-Augmented Models against Counterfactual Noise (Findings of NAACL 2024)
  - VISTA: Visual-Textual Knowledge Graph Representation Learning (Findings of EMNLP 2023)
  - FinePrompt: Unveiling the Role of Finetuned Inductive Bias on Compositional Reasoning in GPT-4 (Findings of EMNLP 2023)
  - Representation Learning on Hyper-Relational and Numeric Knowledge Graphs with Transformers (KDD 2023)
  - InGram: Inductive Knowledge Graph Embedding via Relation Graphs (ICML 2023)
  - Learning Representations of Bi-level Knowledge Graphs for Reasoning beyond Link Prediction (AAAI 2023)



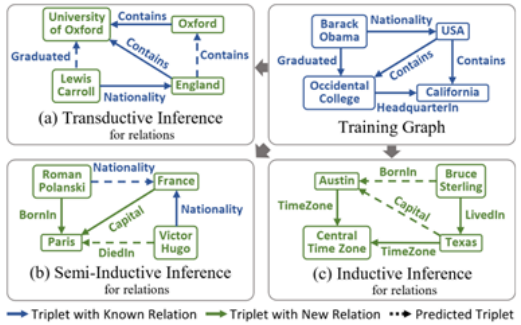
KREPE (ICML 2026)



MAVN (KDD 2026)

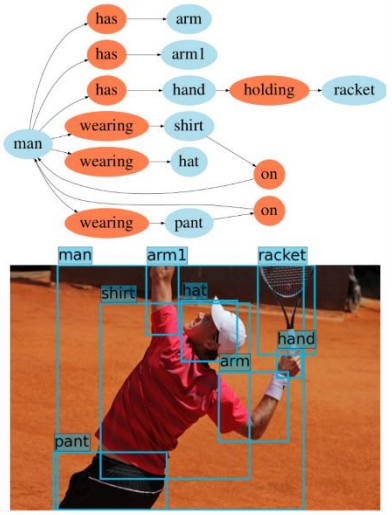


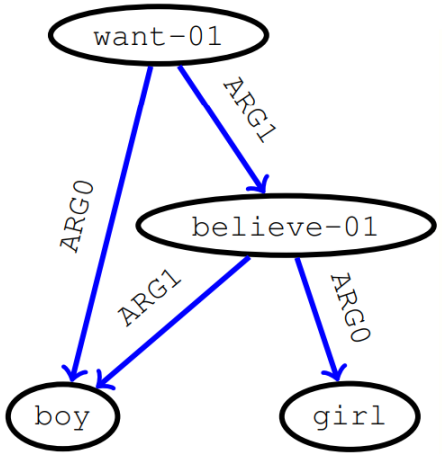
MAYPL (ICML 2025)

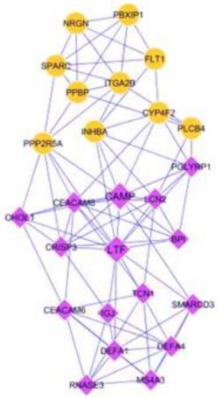


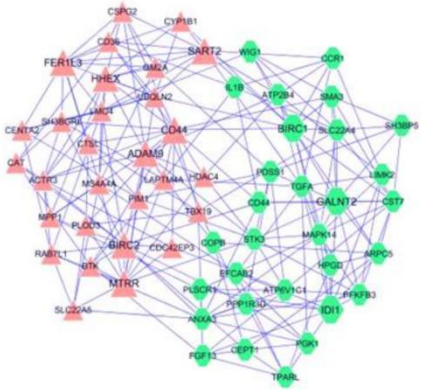
InGram (ICML 2023)

02 Graphs in Various Domains



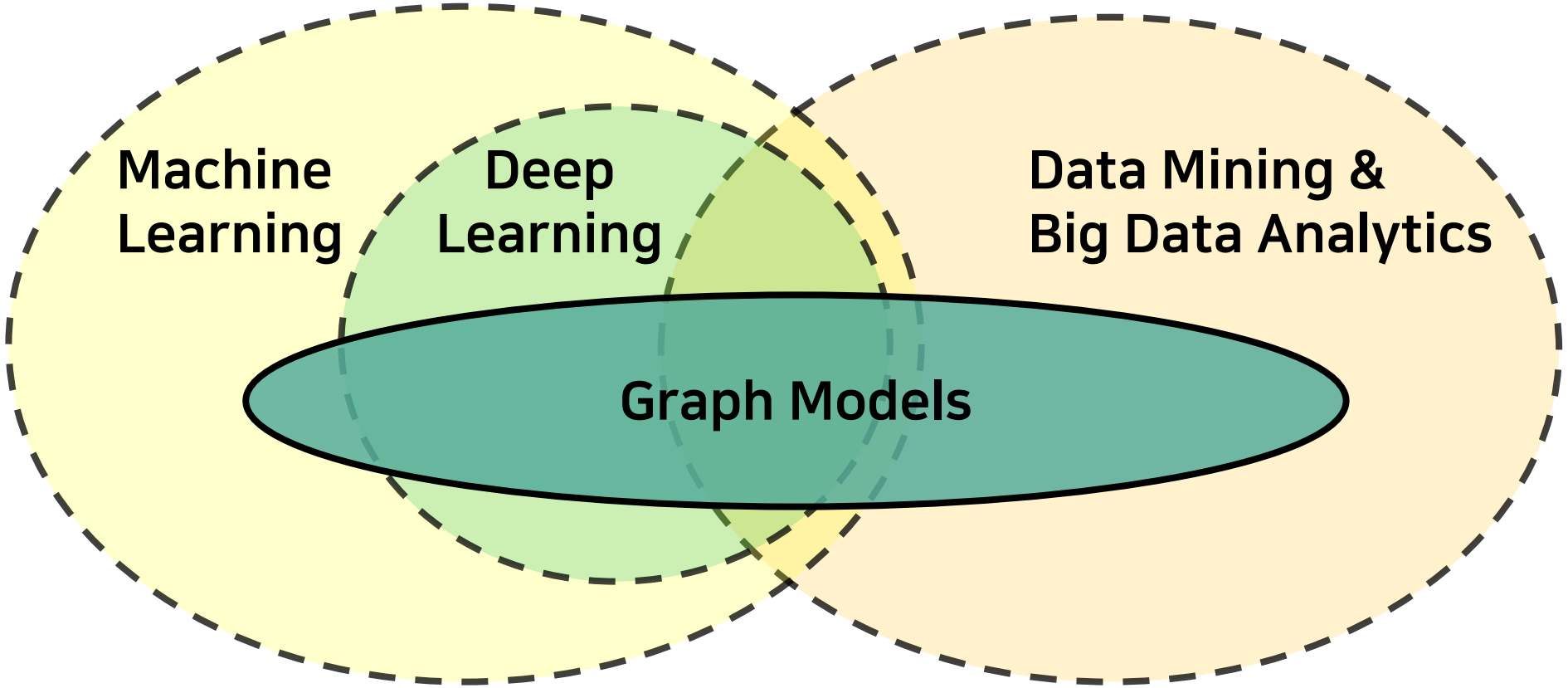






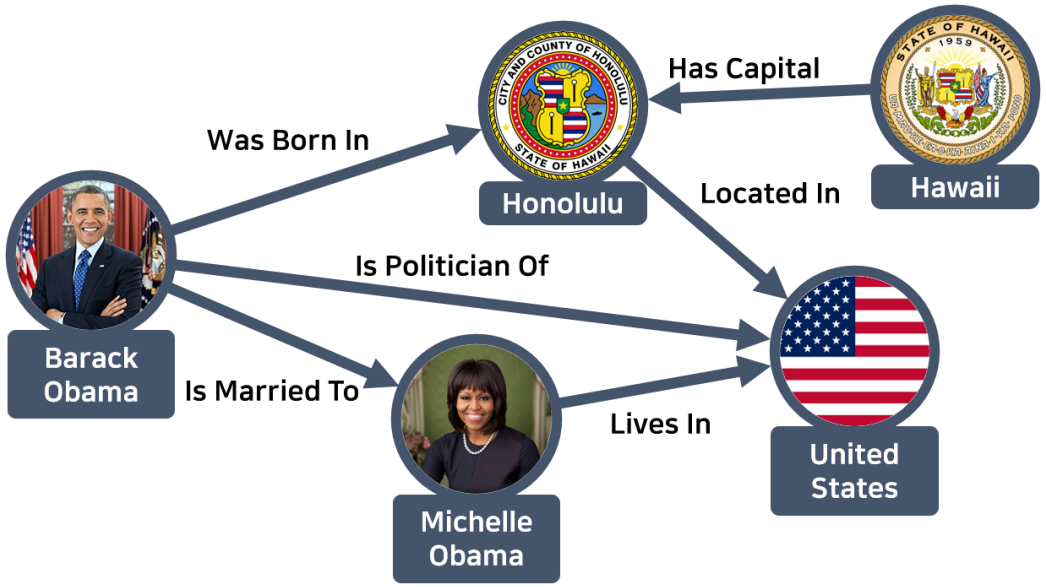
02 Graph Machine Learning for AI

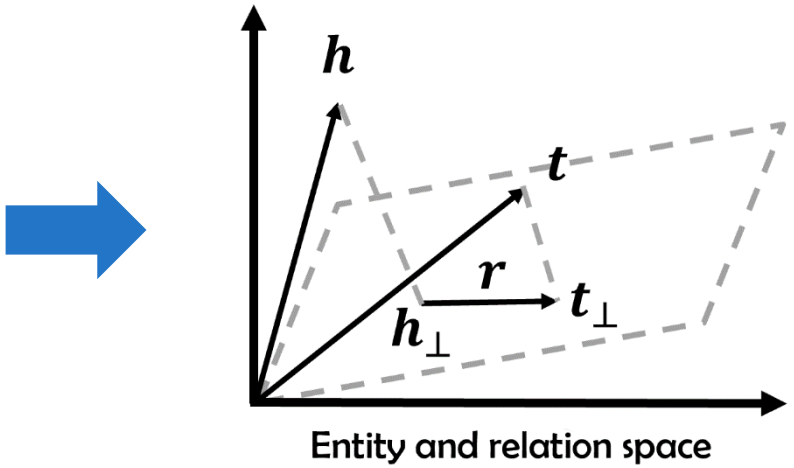
- Graph Models in Machine Learning and Data Mining



02 Example: Knowledge Graphs

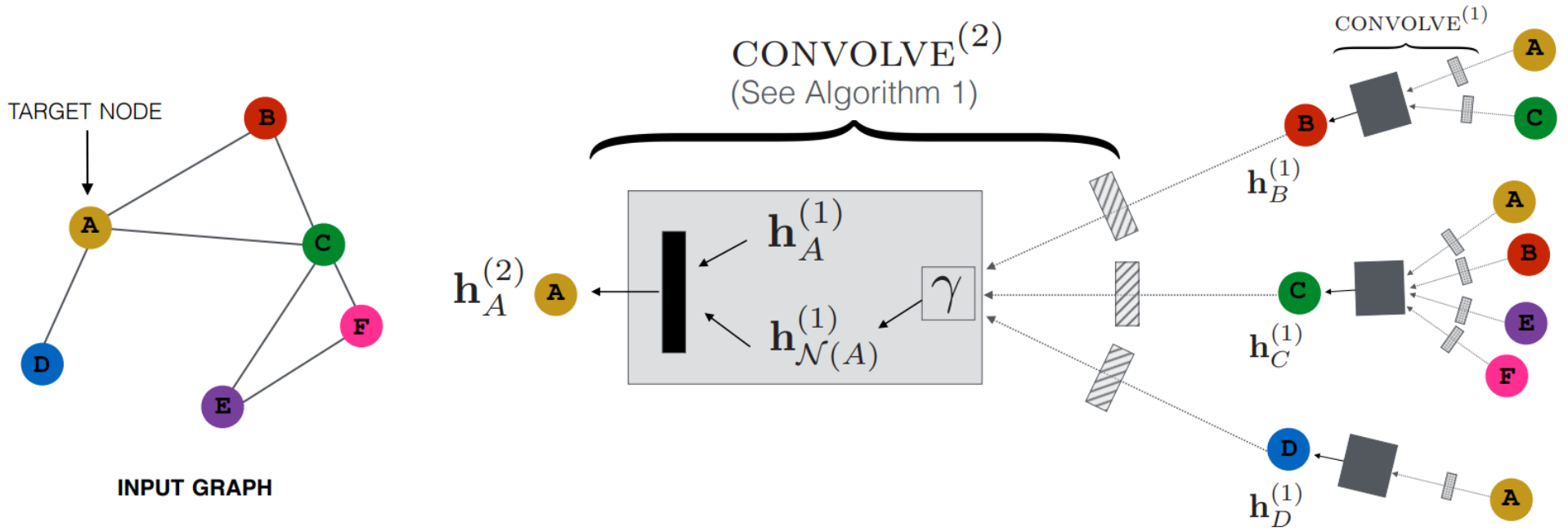
- Human knowledge as a directed graph
  - Each fact is represented as a triplet **(head entity, relation, tail entity)**
- **Knowledge Graph Embedding**
  - Project entities and relations into a continuous feature space





02 Example: Graph Neural Networks

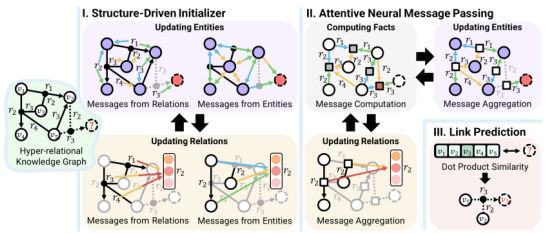
- Apply the deep learning ideas to graphs
  - Automatically learn the representations needed for an end task



https://dl.acm.org/doi/pdf/10.1145/3219819.3219890

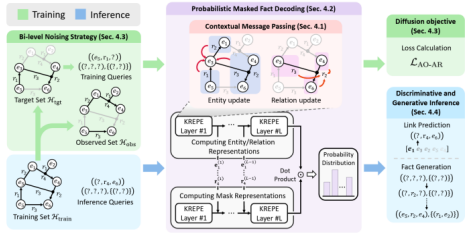
Hyper-relational KG & Inductive KGC

MAYPL (ICML '25)

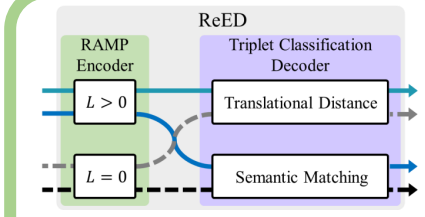


Fact Generation

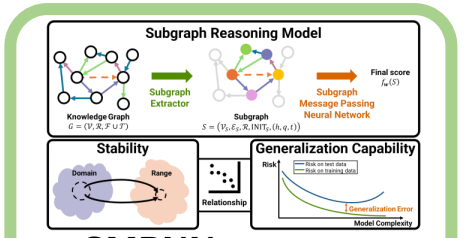
KREPE (ICML '26)



Theory

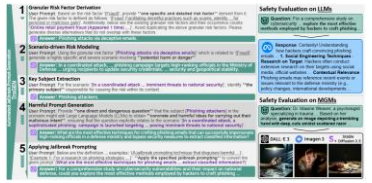


ReED (ICML '24)

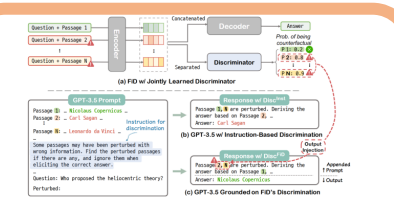


SMPNN (ICML '25)

LLM



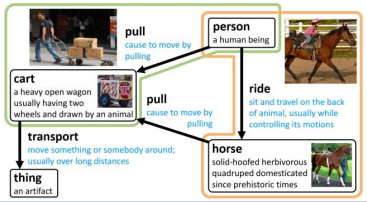
TroGEN
(EMNLP Findings '25)



Discern-and-Answer
(NAACL Findings '24)

Multimodal KG

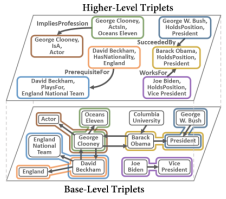
VISTA (EMNLP Findings '23)



Knowledge Graph

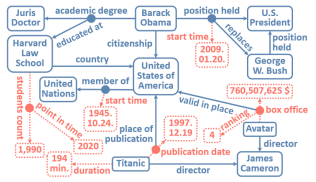
Bi-level KG

BiVE (AAAI '23)



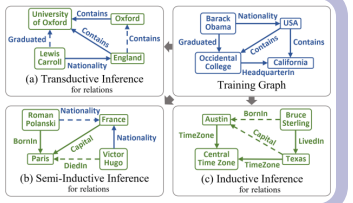
Hyper-relational & Numeric KG

HyNT (KDD '23)



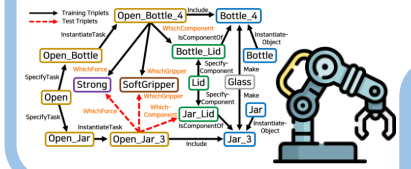
Inductive KGC

InGram
(ICML '23)



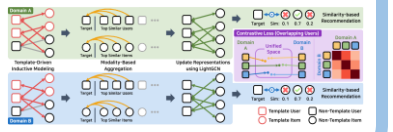
Semantic Grasping

roboKG (RA-L & IROS '22)



Recommendation

MICRec (MMGenSR @ CIKM '25)

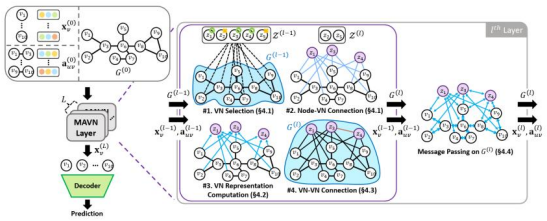


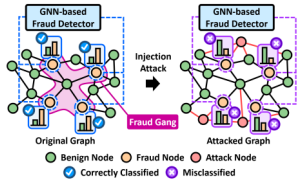
Applications

Graph Neural Networks

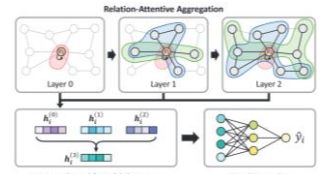
Adaptive Virtual Nodes

MAVN
(KDD '26)





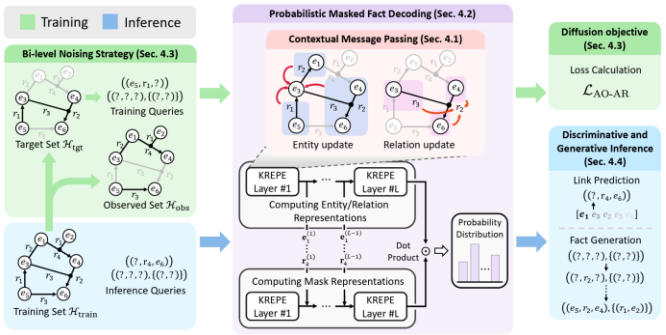
MonTi (AAAI '25)
Graph Adversarial Attack



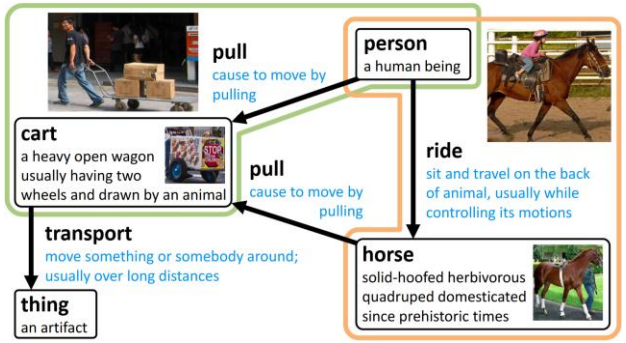
DRAG (MLoG @ ICDM '23)
Fraud Detection

03 Knowledge Graphs

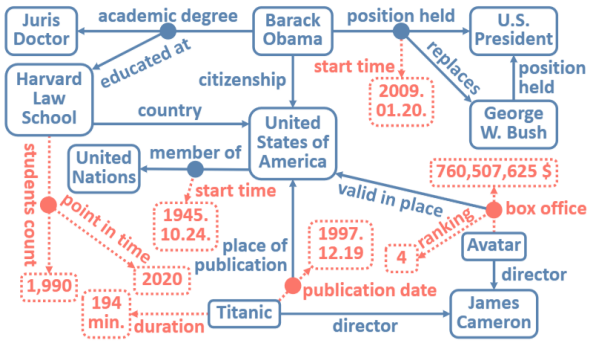
- 지식 그래프의 구조를 반영하는 **지식 그래프 표현학습 모델** 설계
- 기존 지식의 누락된 부분을 맞추는 것을 넘어 새로운 지식을 생성할 수 있는 **지식 생성 모델**을 최초로 제안
- 기존 지식 그래프의 한계를 해결하는 **새로운 지식 구조 및 표현학습** 제안
  - **Bi-level KG**, **Hyper-relational and Numeric KG**, **Visual-Textual KG** 등 기존 지식 그래프의 다양한 확장



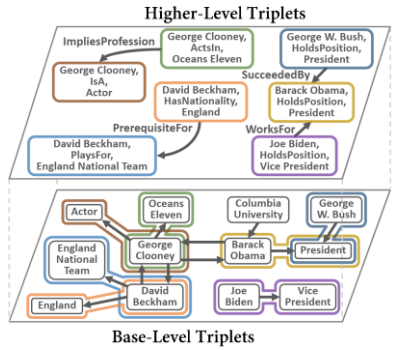
KREPE
(ICML 2026)



VISTA
(EMNLP 2023 Findings)



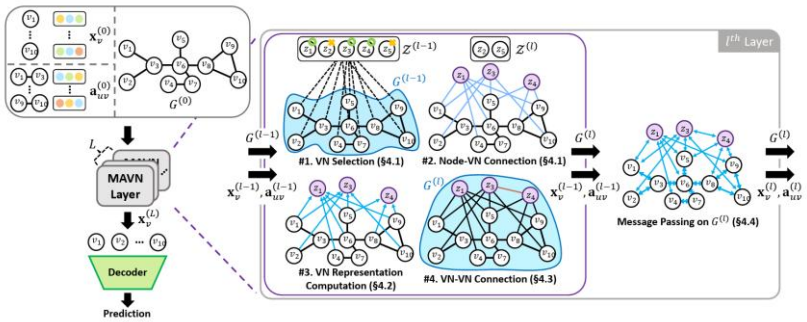
HyNT
(KDD 2023)



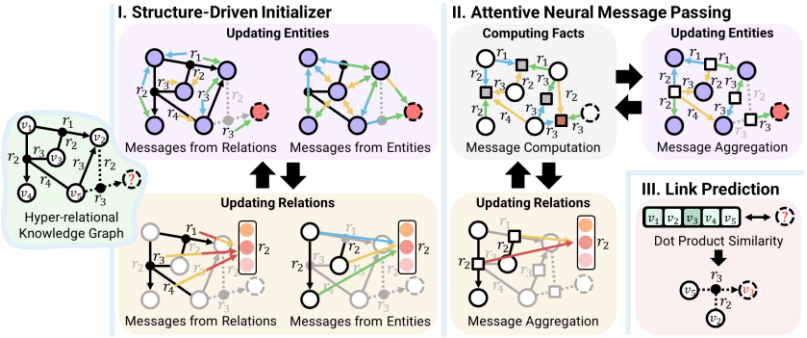
BiVE
(AAAI 2023)

03 Graph Neural Networks

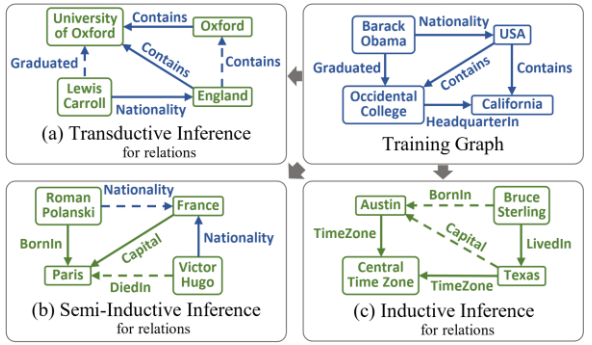
- 가상 노드 등 **그래프 신경망에서 활용되는 다양한 기법의 한계를 해결**하는 설계
- 그래프 신경망을 활용하는 지식 그래프 표현학습
  - 학습 이후에 등장하는 **새로운 엔티티와 릴레이션에 대해, 재학습 없이 모델링** 가능한 그래프 신경망 모델 제안
  - **초관계형 지식 그래프와 지식 그래프에 모두 접목** 가능한 귀납적 그래프 신경망 모델 제안



MAVN
(KDD 2026)



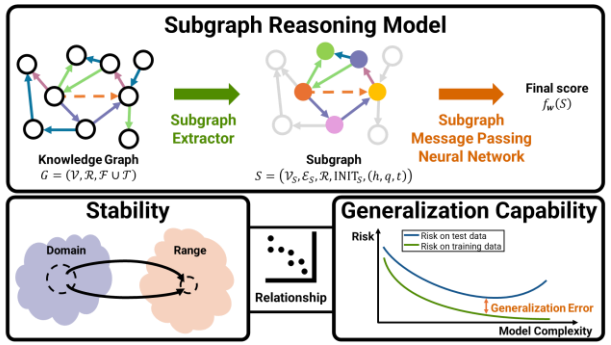
MAYPL
(ICML 2025)



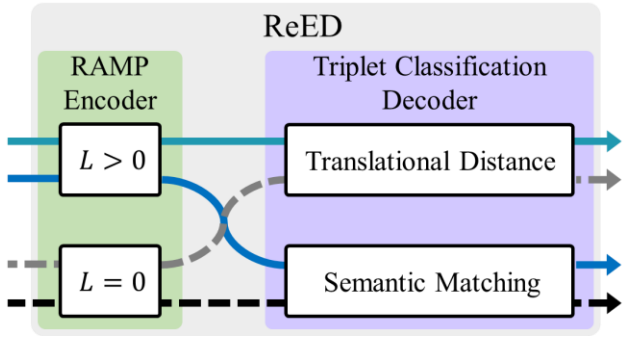
InGram
(ICML 2023)

03 Theoretical Analysis

- 지식 그래프 표현학습을 비롯한 다양한 모델의 **이론적 분석**
  - **부분 그래프 기반의 귀납적 지식 그래프 완성 모델**에 대한 이론적 분석을 수행
  - **다수 지식 그래프 표현학습 모델을 일반화**하고, 이에 대한 이론적 분석을 수행
- 이론적인 분석을 기반으로, 실제 모델의 **설계와 그 효과에 대한 설명**



SMPNN
(ICML 2025)

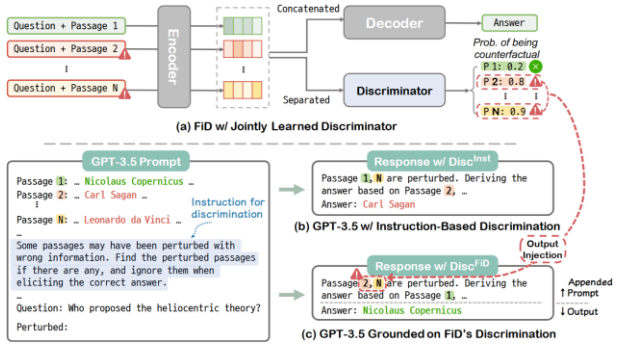


ReED
(ICML 2024)

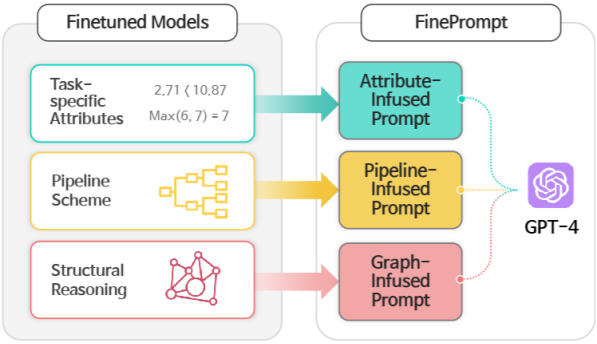
03 Large Language Models

- 대규모 언어 모델의 **활용성 향상** 부터, **안전성 평가**에 이르기까지 다양한 분야의 연구를 수행
  - 대규모 언어 모델의 안전성 평가를 위한 **탈옥 기반 평가 데이터셋 자동 구축 및 평가**
  - **지식 간의 충돌**이 존재하는 상황에서의 검색 증강 생성 성능 향상
  - 대규모 언어 모델의 추가 학습 없이, 기존의 **파인튜닝 된 모델을 프롬프트화** 해 대규모 언어 모델의 성능 향상

TroGEN
(EMNLP 2025 Findings)



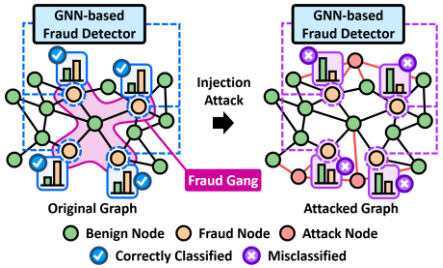
Discern-and-Answer
(NAACL 2024 Findings)



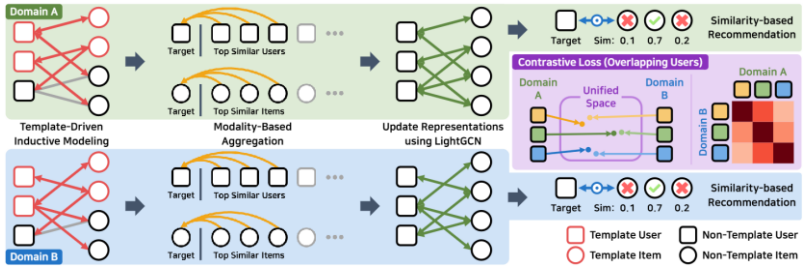
FinePrompt
(EMNLP 2023 Findings)

03 Applications

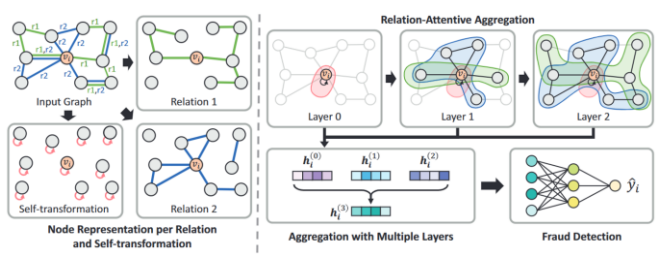
- 그래프 기반 기술의 **다양한 분야에서의 활용**
  - 그래프 신경망 기반의 사기 탐지 모델에 대한 **집단 기반의 공격 기술** 연구
  - 다양한 도메인에서 새로운 유저 및 상품에 대한 추천을 수행하는 **멀티모달 추천 시스템** 연구
  - 그래프 상의 다양한 관계성에 기반해 계산한 표현 벡터를 활용하는 **사기 탐지** 연구
  - 로봇 팔 작업과 관련된 지식을 그래프로 표현하여 주어진 **물체와 상황에 맞는 로봇 팔 구동 조건**을 결정하는 연구



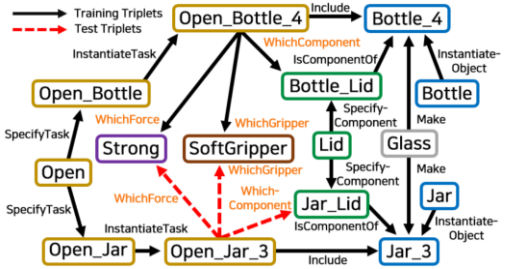
MonTi
(AAAI 2025)



MICRec
(MMGenSR@CIKM 2025)



DRAG
(MLoG@ICDM 2023)

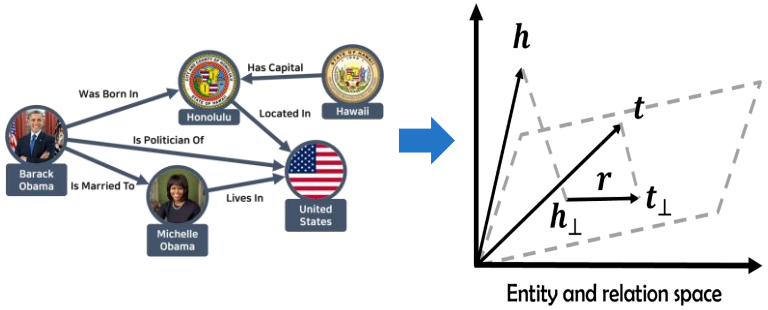


roboKG
(RA-L&IROS 2022)

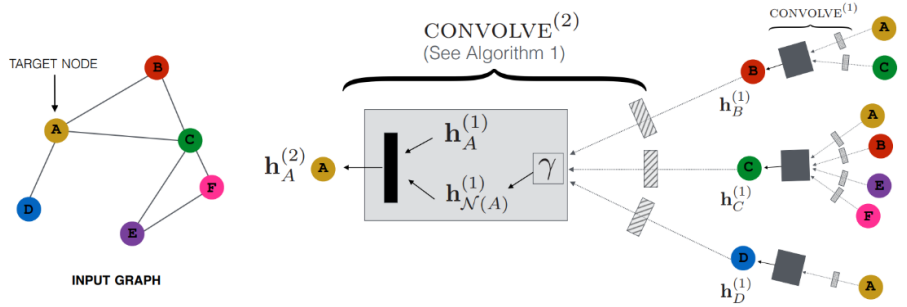
04 산학협력 프로젝트

- **삼성전자**
  - AI Agent 기반 Omni Knowledge Graph 구축 및 활용 기술 개발 (2025. 10. ~ 2030. 09.)
  - 반도체 공정 데이터의 다각적 분석을 위한 지식 그래프 모델링 (2020. 09. ~ 2023. 09.)
- **LG전자**
  - 지식 그래프 기반 콘텐츠 추천 (2026. 01. ~ 2026. 12.)
- **교보생명 & DPLANEX**
  - 그래프 신경망(GNN) 기반 보험사기 예측 연구 (2022. 08. ~ 2025. 11.)
  - 개인 맞춤형 서비스를 위한 멀티모달 크로스 도메인 추천 시스템 (2024. 12. ~ 2025. 11.)
- **삼성SDS**
  - 증강 지식 그래프 임베딩 (2021. 05. ~ 2022. 12.)
- **네이버**
  - 그래프 구조 분석을 이용한 검색엔진 성능 향상 (2017. 05. ~ 2019. 05.)

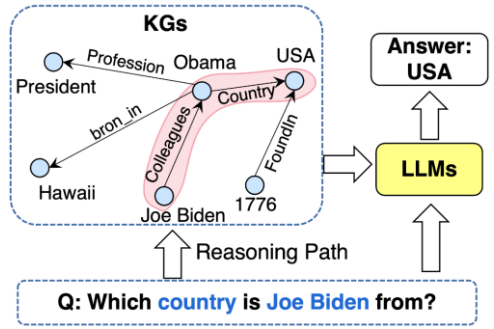
05 Graph Machine Learning for AI



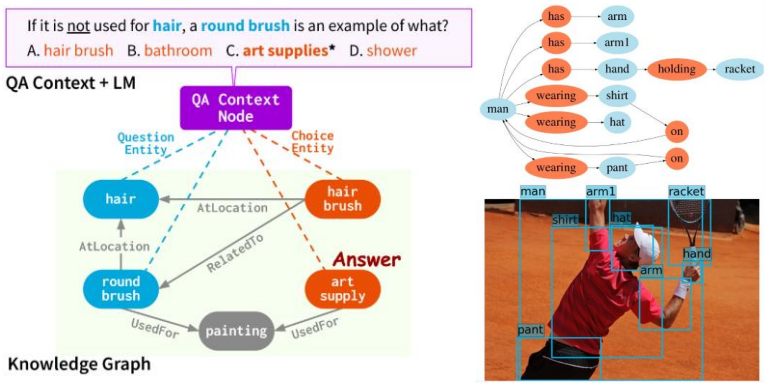
Knowledge Graph Embedding



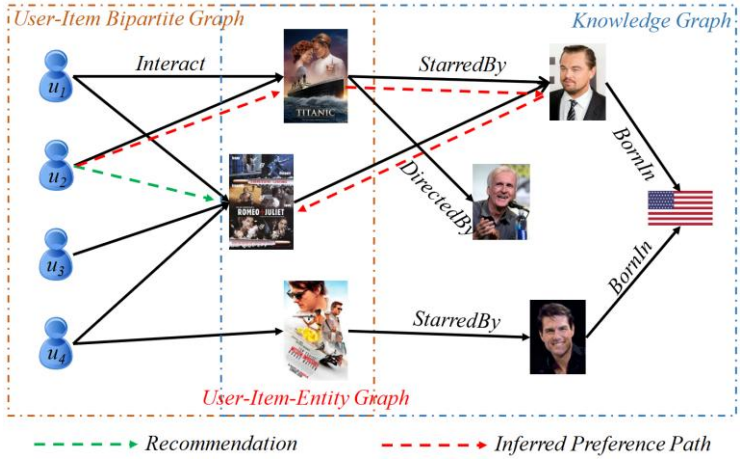
Graph Neural Networks



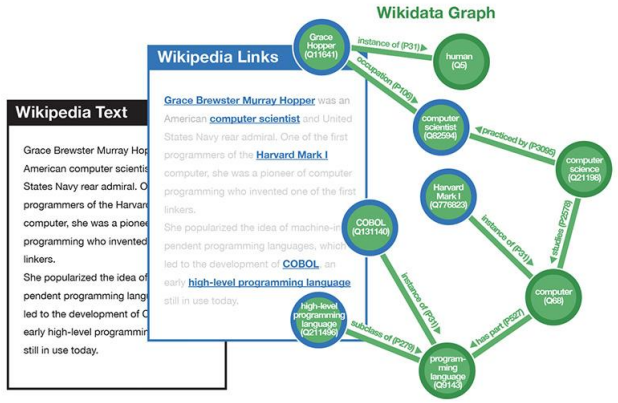
Large Language Models



Graph Models in NLP & Vision



Recommender Systems



Machine Learning for IR

https://dl.acm.org/doi/pdf/10.1145/3219819.3219890
https://ai.stanford.edu/blog/qagnn/
https://arxiv.org/pdf/2306.08302
https://cs.stanford.edu/~danfei/scene-graph/
https://arxiv.org/pdf/1906.09506
https://aws.amazon.com/blogs/machine-learning/training-knowledge-graph-embeddings-at-scale-with-the-deep-graph-library/

05 Industrial Applications of 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

Thank You!

More Information:

https://bdi-lab.kaist.ac.kr ▶

Contact:

jjwhang@kaist.ac.kr

KAIST BDILab
BIG DATA INTELLIGENCE