Knowledge-Guided IE

New Frontiers of Information Extraction (Part III)

Manling Li

Department of Computer Science

University of Illinois Urbana-Champaign

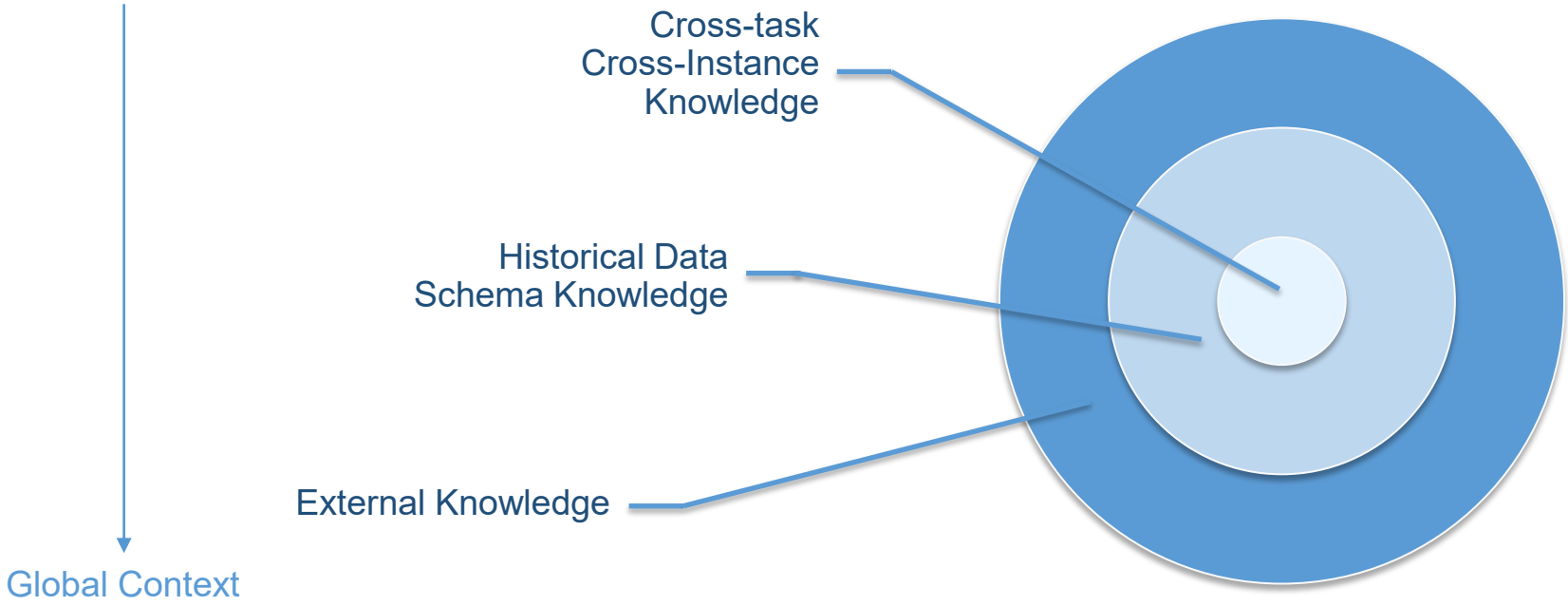
July 2022

NAACL Tutorials

New Frontiers of Information Extraction

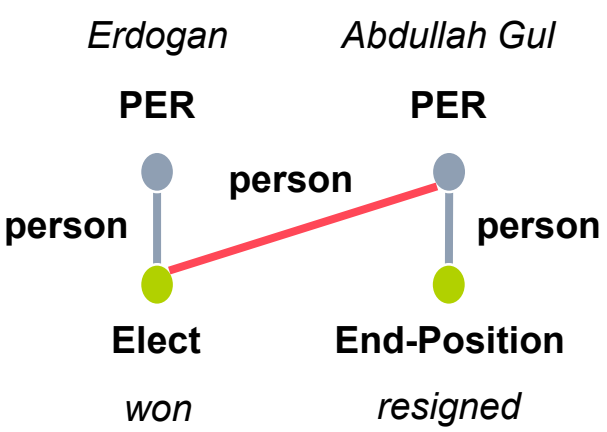
Knowledge-Enhanced IE

- What knowledge is useful for information extraction?
  - cross-task cross-instance knowledge such as the interactions between knowledge elements
  - schema knowledge induced from historical data
  - external knowledge such as commonsense knowledge

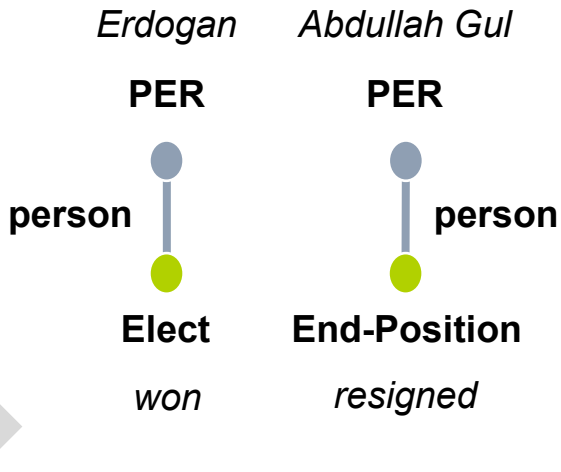


Cross-task Cross-instance Knowledge

- Existing neural models do not explicitly model cross-task and cross-instance interactions among knowledge elements
- Example: *Prime Minister **Abdullah Gul** resigned earlier Tuesday to make way for **Erdogan**, who won a parliamentary seat in by-elections Sunday.*



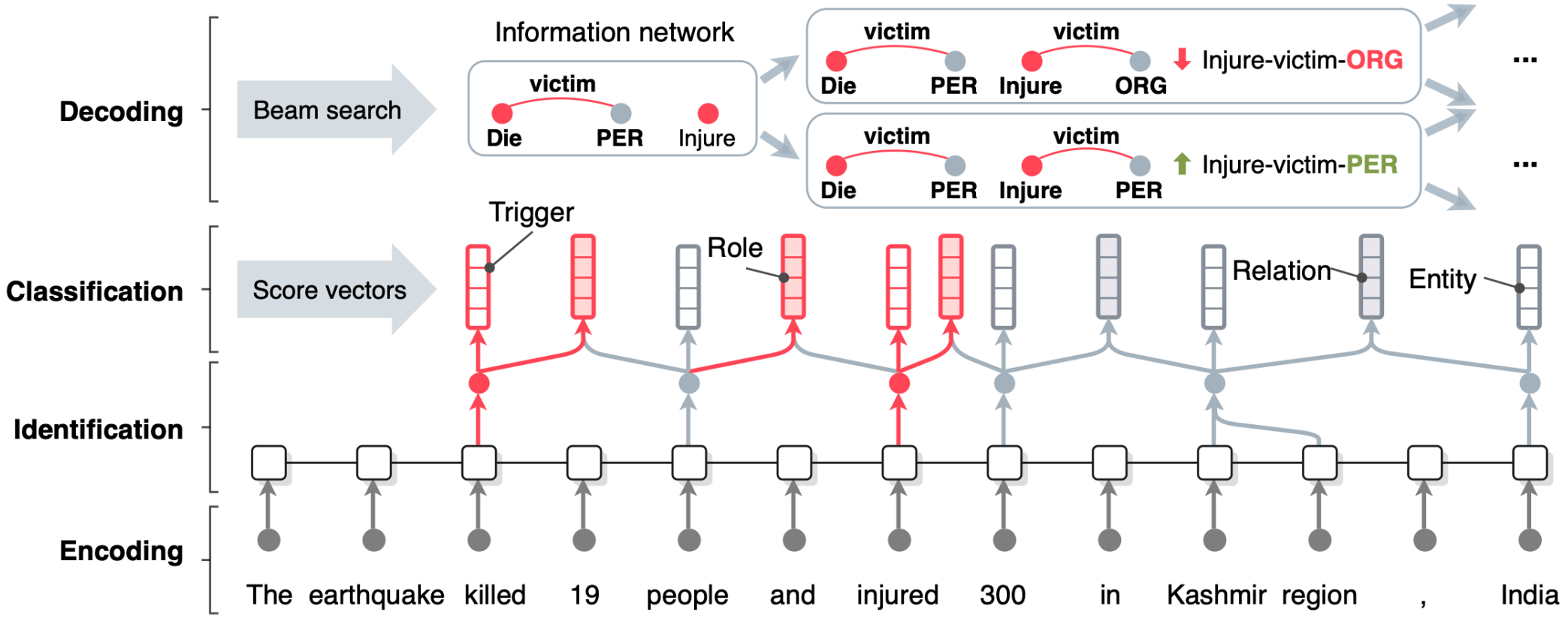
1. An **Elect** event usually has only one **Person** argument
2. An entity is unlikely to act as a **Person** argument for **End-Position** and **Elect** events at the same time



[Lin et al, 2020]

OneIE: Justify whether the entire graph makes sense

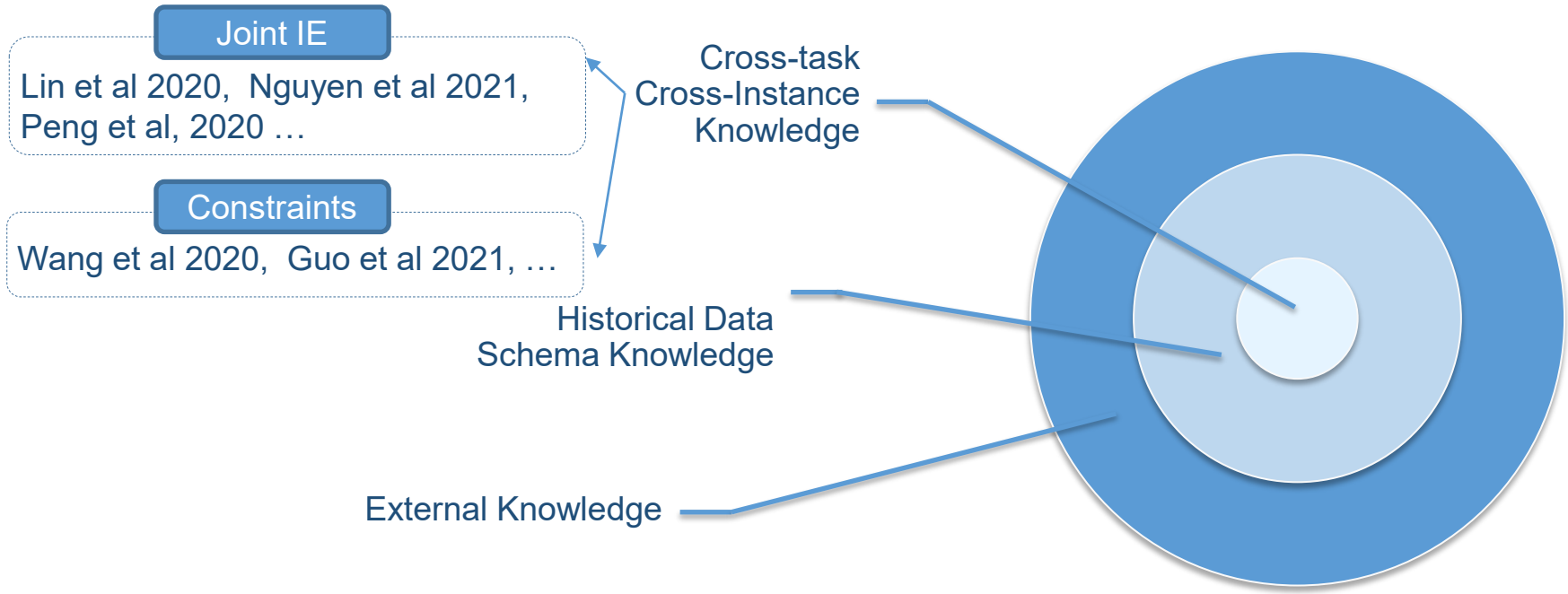
- OneIE framework extracts the information graph (nodes: entities and events, edges: relations and arguments) from a given sentence. (Lin et al., 2020)
- Main challenge for Joint IE: How to capture interactions between knowledge elements?



[Lin et al, 2020]

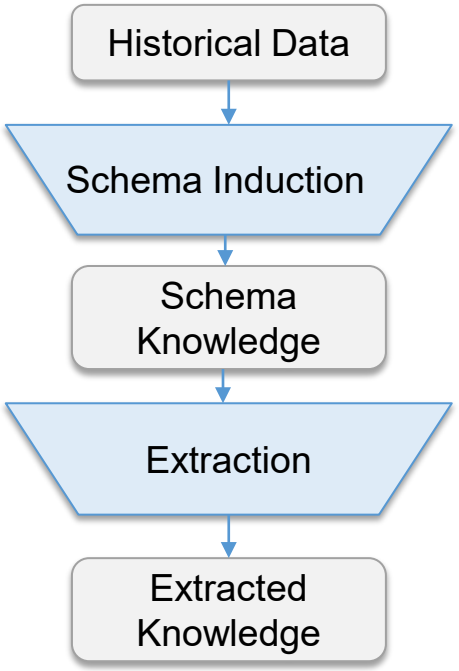
Knowledge-Enhanced IE

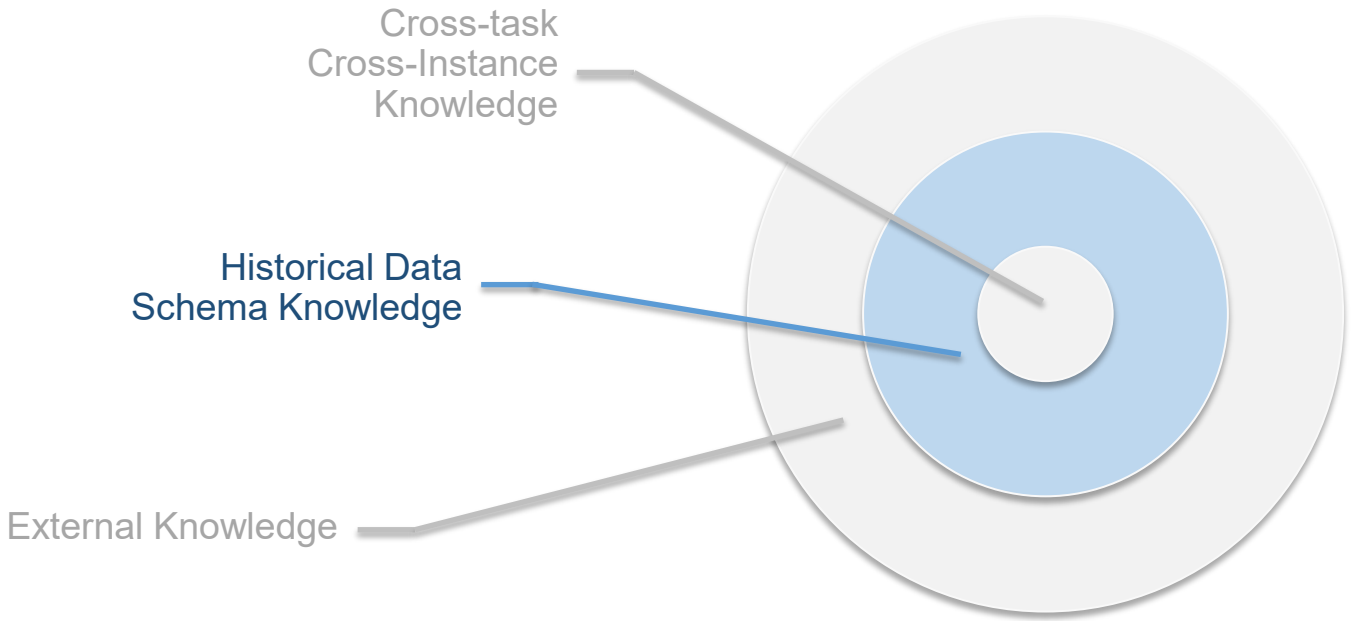
- What knowledge is useful for information extraction?
  - cross-task cross-instance knowledge such as the interactions between knowledge elements
  - schema knowledge induced from historical data
  - external knowledge such as commonsense knowledge



Knowledge-Enhanced IE

- What knowledge is useful for information extraction?
  - cross-task cross-instance knowledge such as the interactions between knowledge elements
  - schema knowledge induced from historical data
  - external knowledge such as commonsense knowledge, or domain knowledge





Schema Knowledge: Path Language Model

- A good schema consists of **salient** and **coherent** paths between them (Li et al, 2020).
  - Salience: recurring event-event connection patterns
  - Coherence: semantically coherent

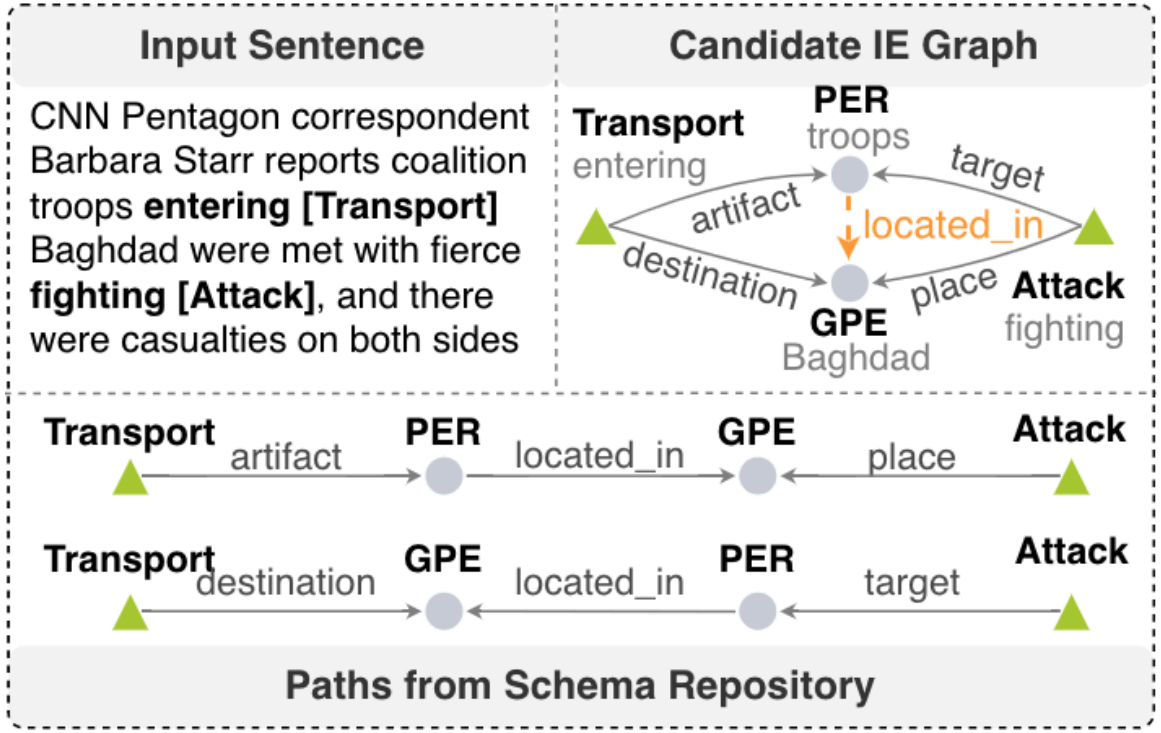
Criteria			Examples	Frequency
Single Path	Salience	High	Transport —agent→ GPE —affiliation→ PER ←attacker— Attack	31
		Low	Transport —destination→ GPE —affiliation→ PER ←attacker— Attack	2
	Semantic Coherence	High	Transport —origin→ FAC —part-whole→ LOC ←part-whole— GPE —affiliation→ PER ←attacker— Attack	9
		Low	Transport —agent→ GPE —affiliation→ PER —affiliation→ GPE ←resident— PER ←target— Attack	24
Multiple Paths	Semantic Consistency	High	Transport —destination→ GPE —place→ Attack Transport —artifact→ PER —located_in→ GPE —place→ Attack	20
		Low	Transport —destination→ GPE —place→ Attack Transport —origin→ GPE —place→ Attack	0

[Li et al, 2020]

to guide information extraction. We use the paths from

Schema-Guided Information Extraction

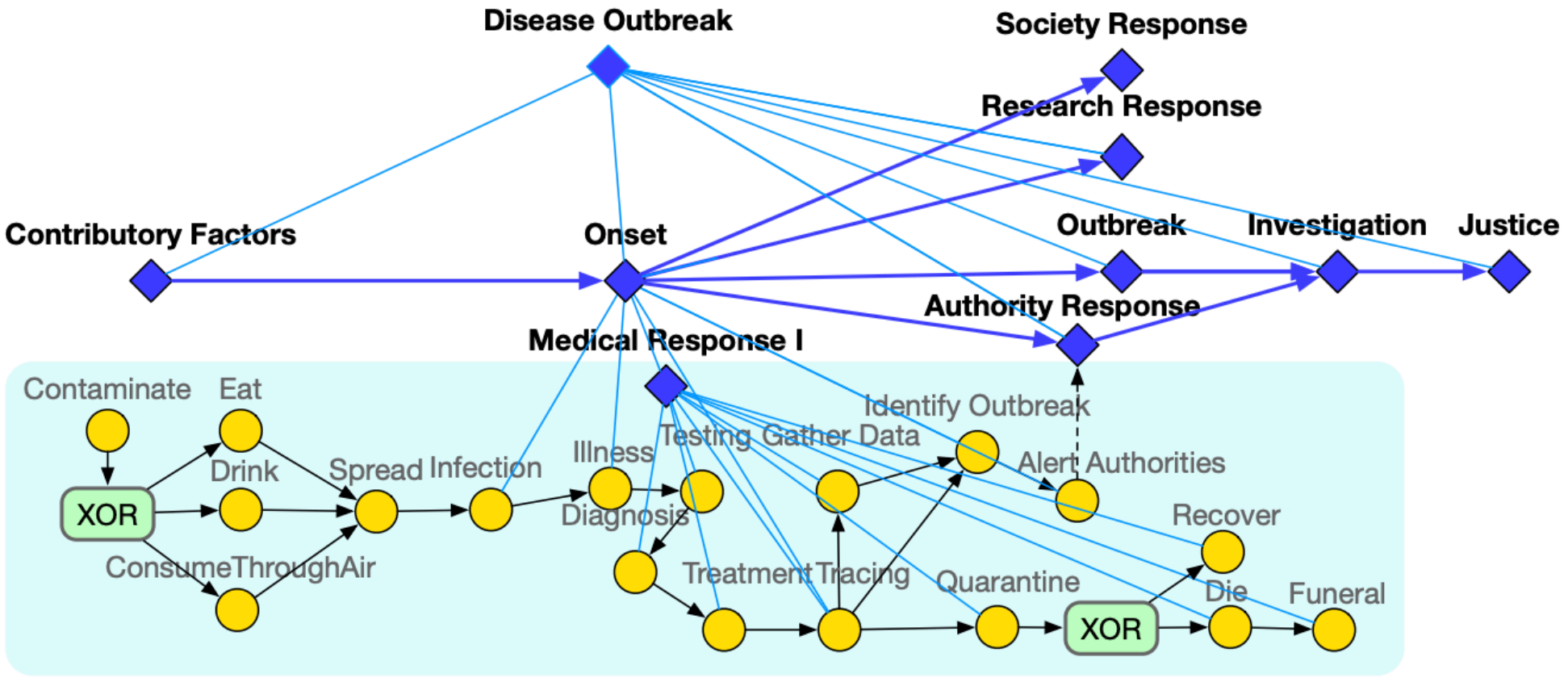
- Use the state-of-the-art IE system OneIE (Lin et al, 2020) to decode converts each input document into an IE graph
- Each path in the graph schema is encoded as a single global feature for scoring candidate IE graphs
- OneIE promotes candidate IE graphs containing paths matching schema graphs
- http://blender.cs.illinois.edu/software/oneie
- F-scores (%) on ACE2005 data [Lin et al., ACL2020]:



Dataset	Entity	Event Trigger Identification	Event Trigger Classification	Event Argument Identification	Event Argument Classification	Relation
Baseline	90.3	75.8	72.7	57.8	55.5	44.7
+PathLM	90.2	**76.0**	**73.4**	**59.0**	**56.6**	**60.9**

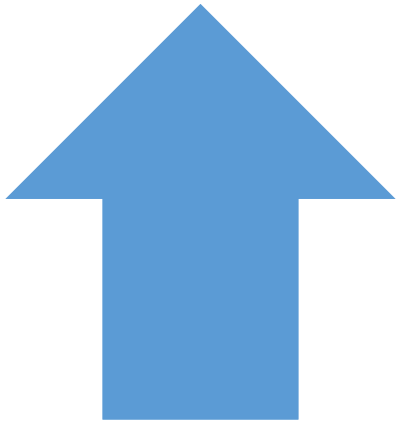
Schema Knowledge: Graph Schema

- Example schema of Disease Outbreak (Du et al, 2022)

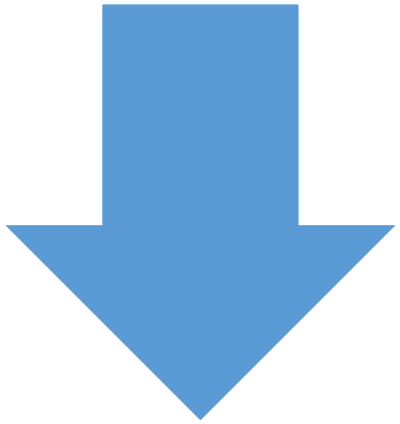


[Du et al, 2021]

Schema-Guided IE Challenges



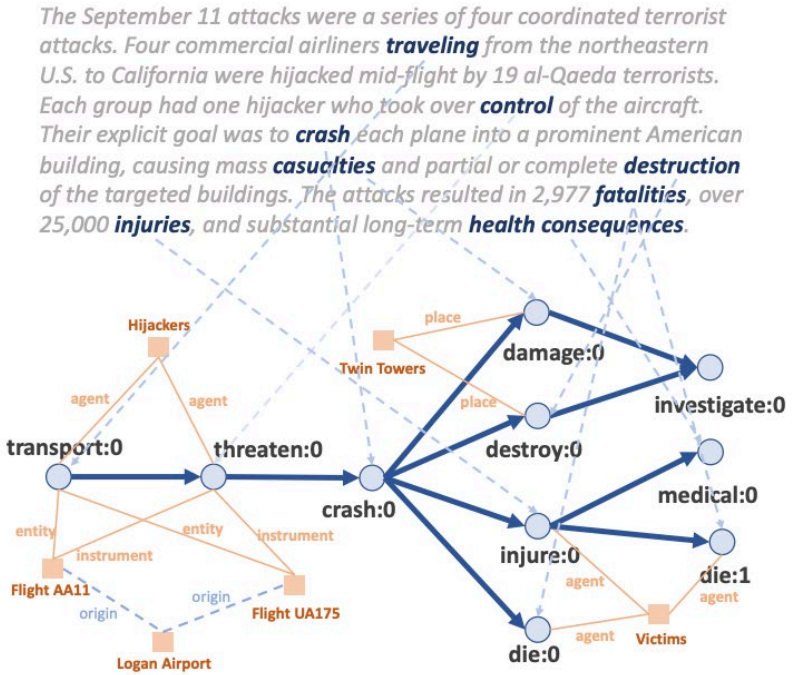
How to obtain schema knowledge?



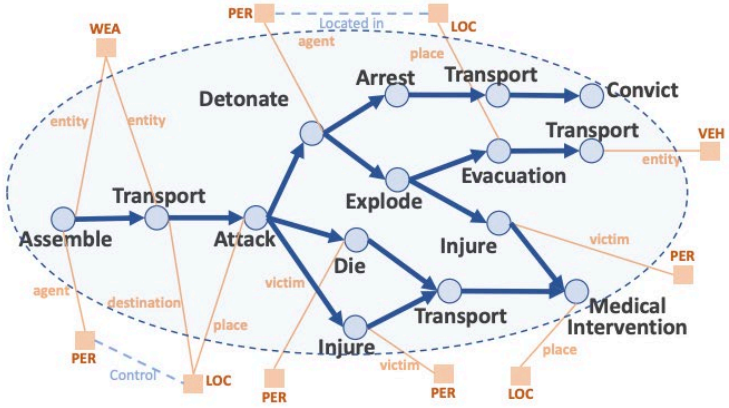
How to leverage schema knowledge in IE?

Graph Schema Induction

- Temporal Graph Schema (Li et al, 2021, Jin et al, 2022)
  - Goal:
    - Learn the frequent recurring patterns.
  - Idea:
    - Schemas are the hidden knowledge to control instance graph generation.
  - Design:
    - Graph Generation Model
  - Model as schemas:
    - Prediction ability
    - Condition on instance graph



(a) Event instance graph



(b) Event schema graph

[Li et al, 2021, Jin et al, 2022]

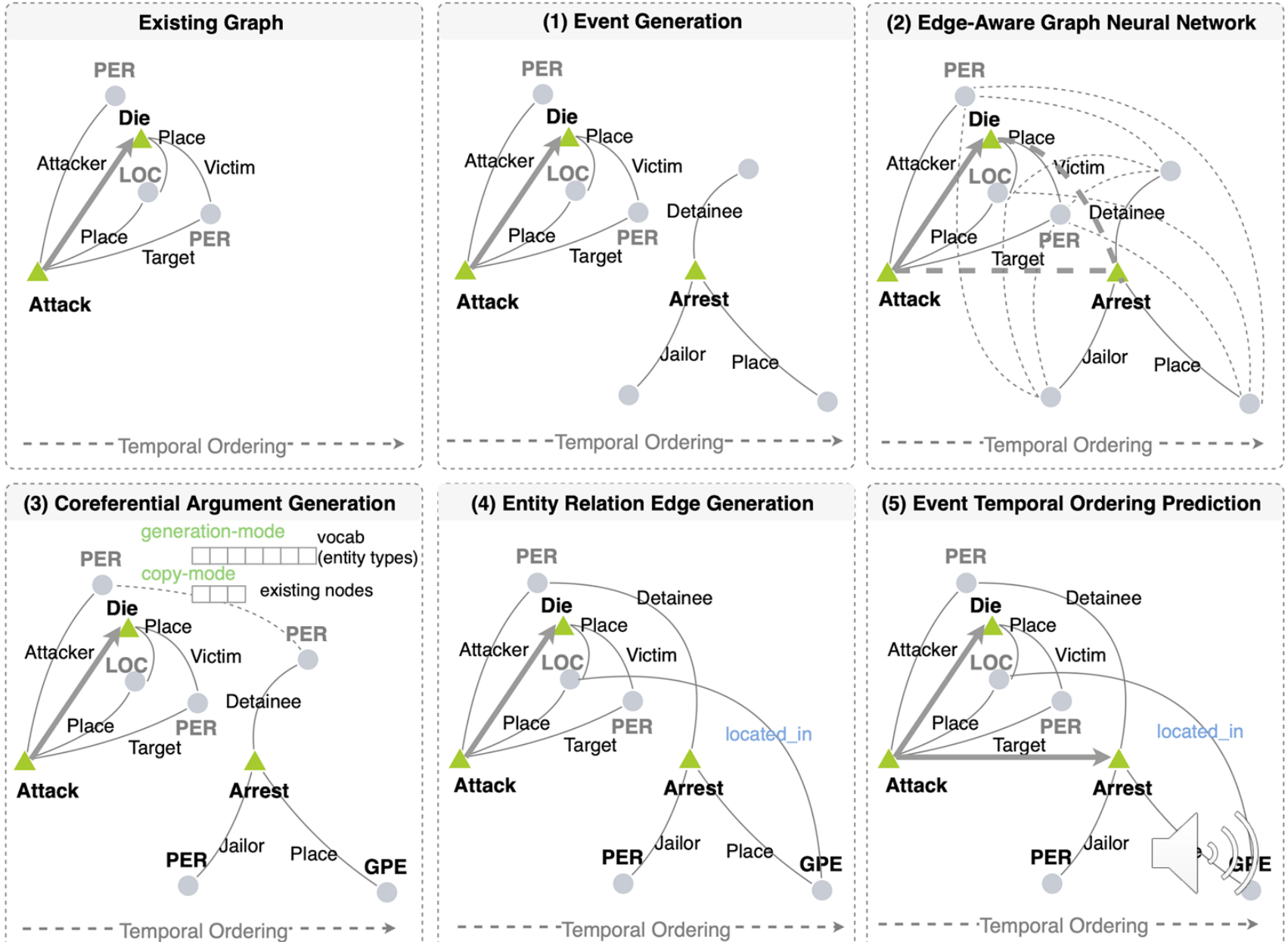
Generative Event Graph Model

Schema as graph generation

$$p(G) = \prod_{i=0}^{|\mathcal{E}|} p(G_i|G_{<i}).$$

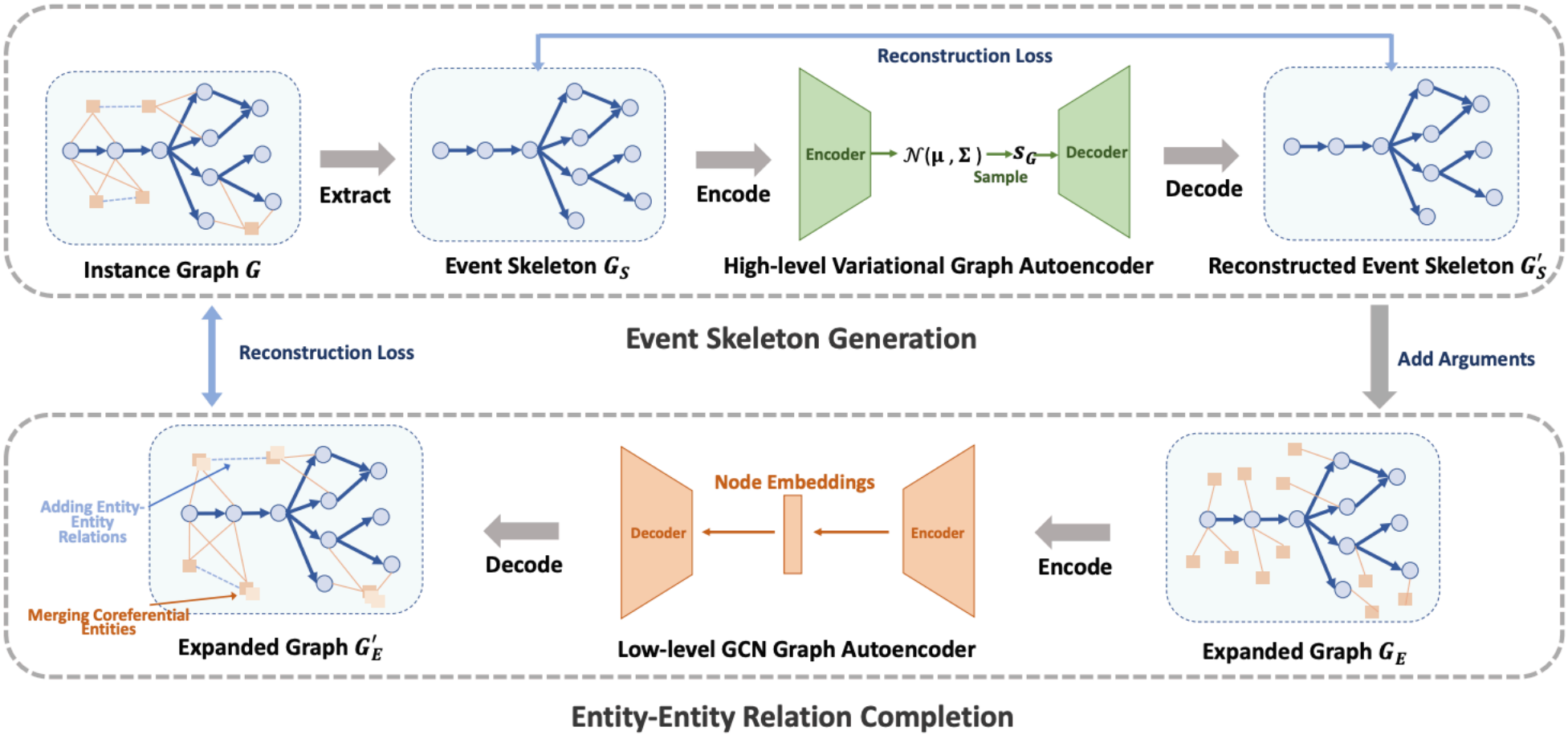
- Step 1.
  Event Node Generation
- Step 2.
  Message Passing
- Step 3.
  Argument Node Generation
- Step 4.
  Relation Edge Generation
- Step 5.
  Temporal Edge Generation

[Li et al, 2021]



Graph Schema Knowledge Acquisition

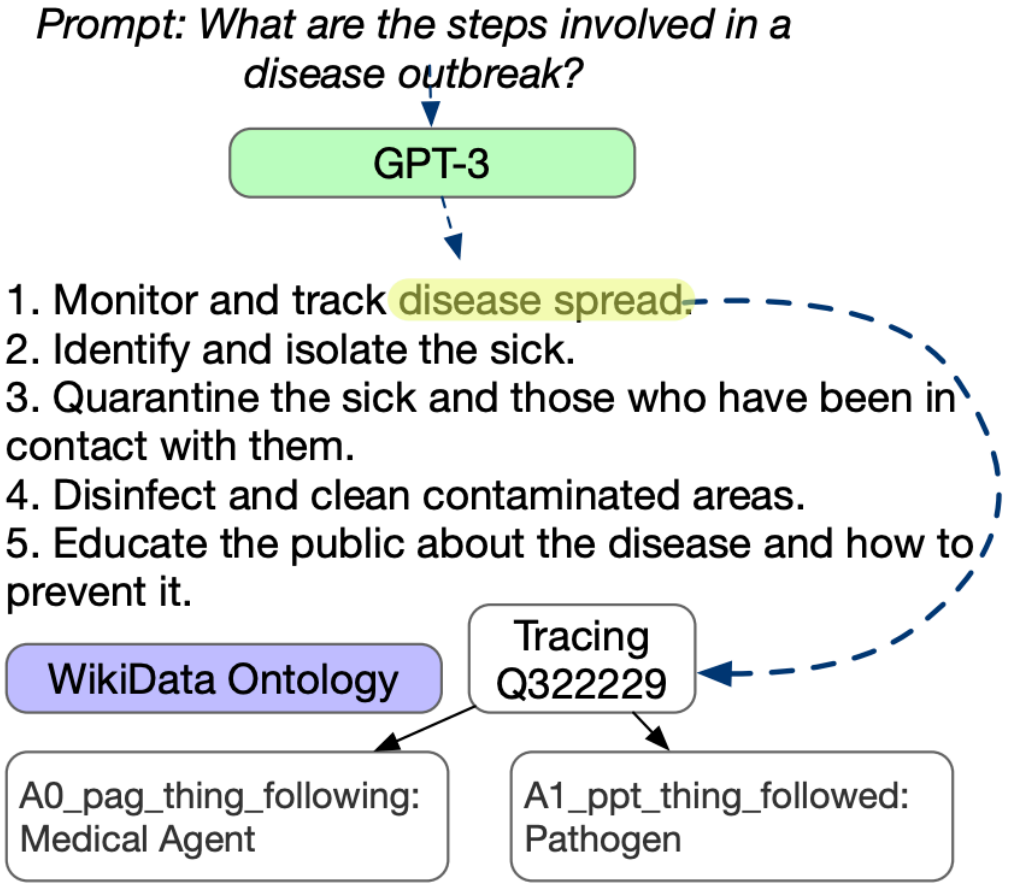
- Two-stage Graph Generation Model (Jin et al, 2022)



[Jin et al, 2022]

Schema Knowledge: Graph Schema Resources

- RESIN-11: A library of hierarchical schemas for 11 scenarios (Du et al, 2022)
  - capturing a wide coverage of newsworthy events.

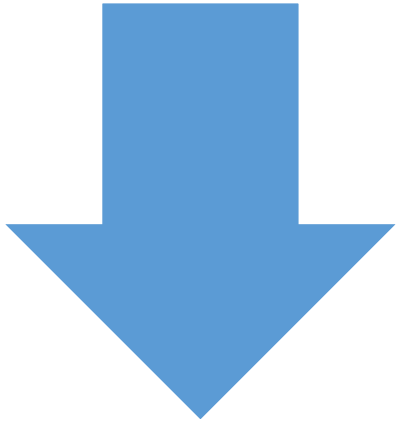


Scenario	# Episodes	# Events	# Ents	# Rels
Business Change	18	81	24	54
Civil Unrest	6	34	18	24
Disease Outbreak	19	102	27	93
Election	8	35	14	33
International Conflict	17	95	56	50
Kidnapping	9	66	15	56
Mass Shooting	8	37	13	31
Sports Events	4	17	14	19
Terrorist Attacks	8	36	11	26
Disaster/Manmade Disaster	8	38	10	29
Disaster/Natural Disaster	4	23	8	18
IED/General Attack	19	52	40	22
IED/General IED	10	48	18	39
IED/Drone Strikes	10	50	19	43
IED/Backpack IED	10	49	18	40
IED/Roadside IED	10	48	19	39
IED/Car IED	10	50	19	43

[Du et al, 2022]

Schema-Guided IE Challenges

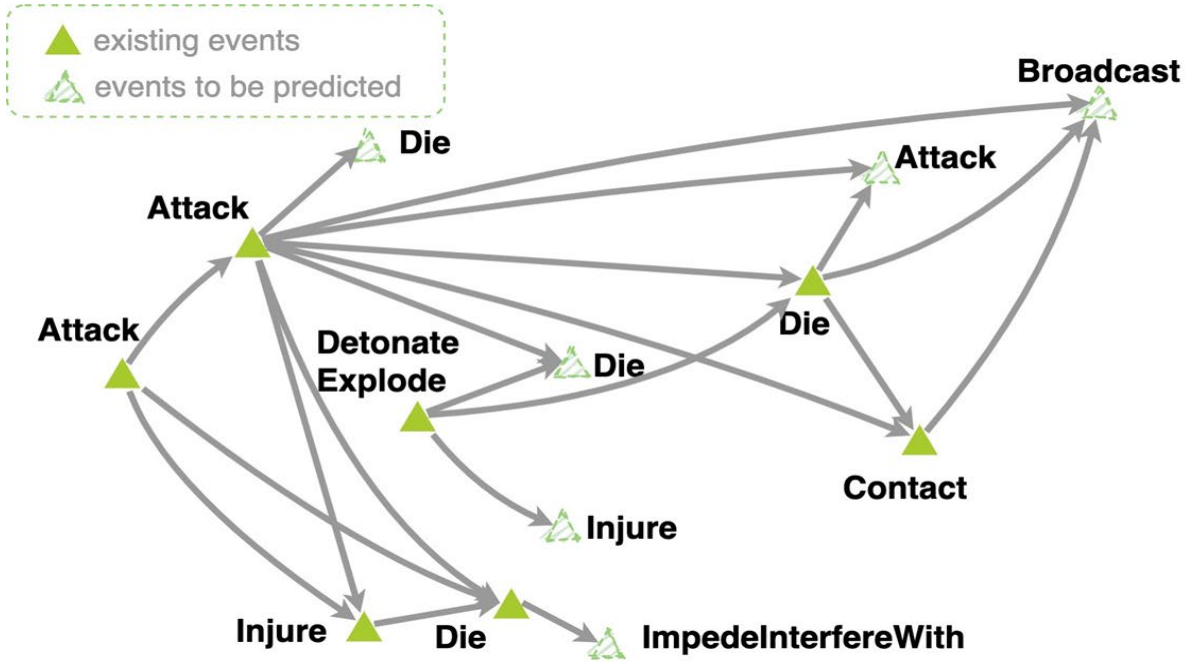
How to obtain schema knowledge?



How to leverage schema knowledge in IE?

“Model as Schema” for Event Prediction

- **Schema-guided Event Prediction:** The task aims to predict ending events of each graph.
  - Considering that there can be multiple ending events in one instance graph, we rank event type prediction scores and adopt MRR and HITS@1 as evaluation metrics.



Event Prediction

Human Schema	FireExplosion
	Die
	TrialHearing
	Transportation
	Sentence
	Broadcast
Graph Temporal Schema	Die
	Injure
	Attack
	Broadcast
	Arrest

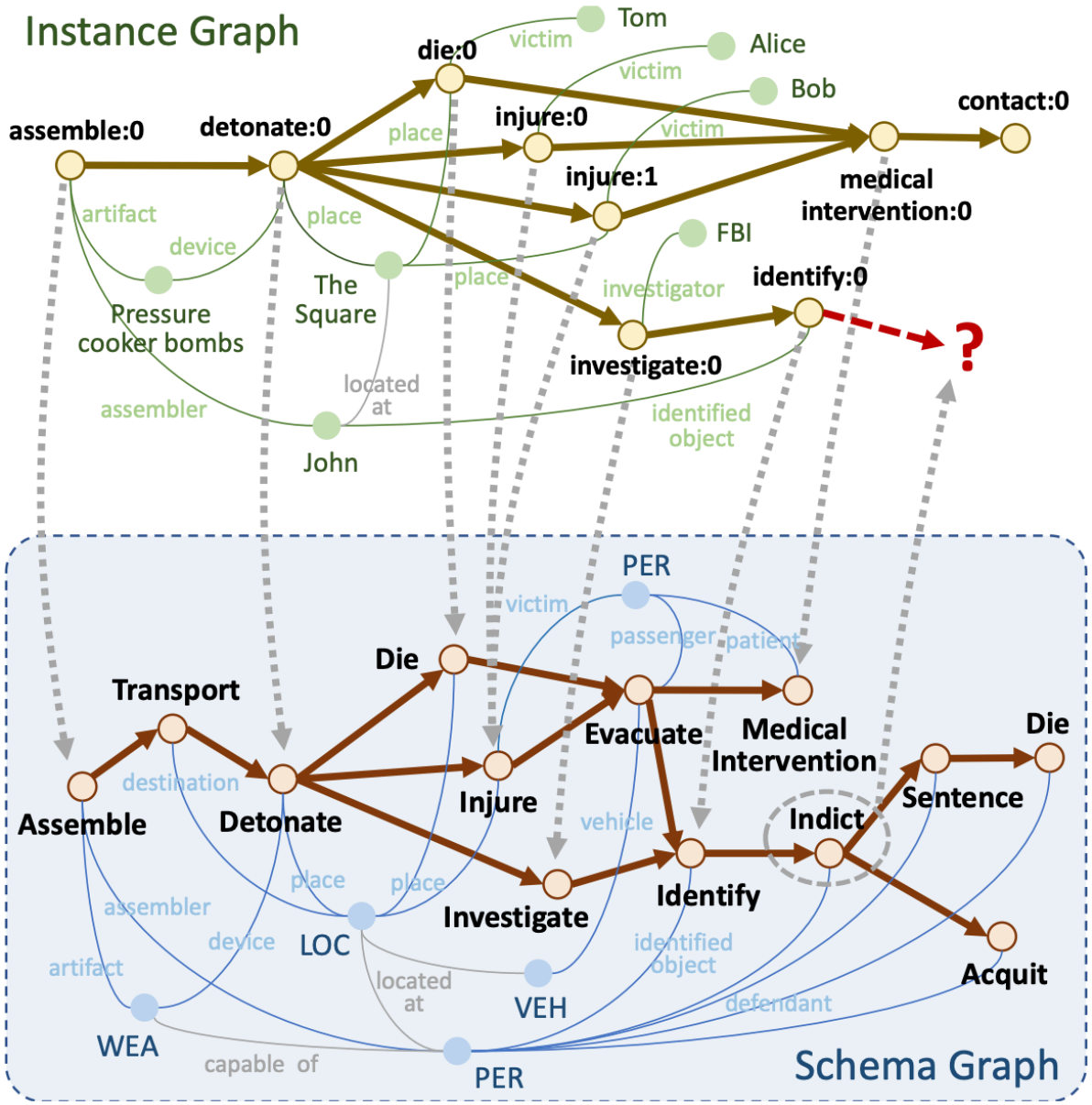
Dataset	Models	MRR	HITS@1
General	Human Schema	0.173	0.205
	Event Graph Model	**0.401**	**0.520**

Dataset	Models	MRR	HITS@1
IED	Human Schema	0.072	0.222
	Event Graph Model	**0.224**	**0.741**

[Li et al, 2021]

Schema Library Guided Event Graph Completion

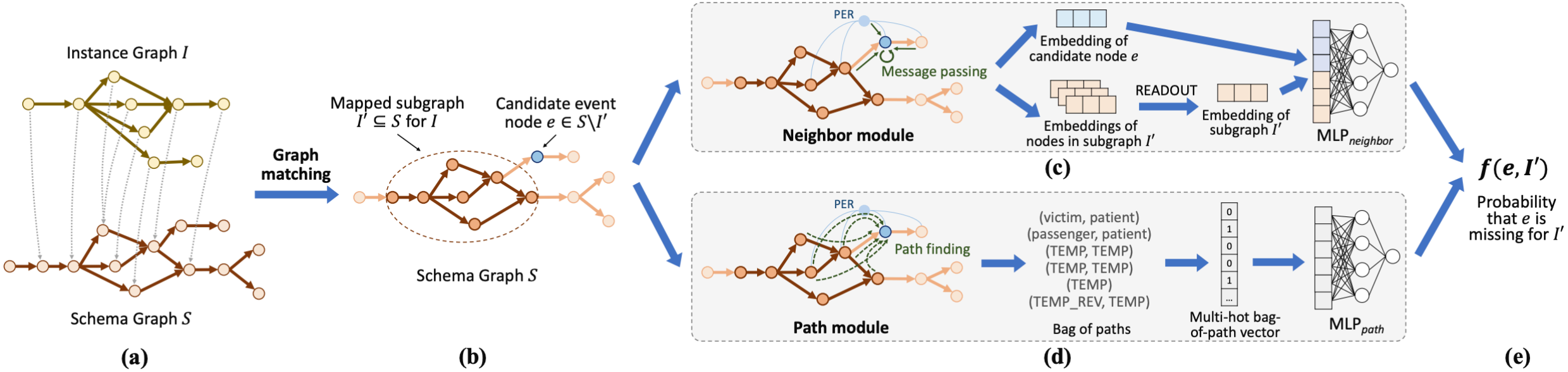
- Problem:
  - Missing event nodes
  - Missing edges
- Task: Event Graph Completion (Wang et al, 2022)
- Input
  - Incomplete event graph
  - Schema graph
- Output
  - Complete event graph



[Wang et al, 2022]

Schema Library Guided Event Graph Completion

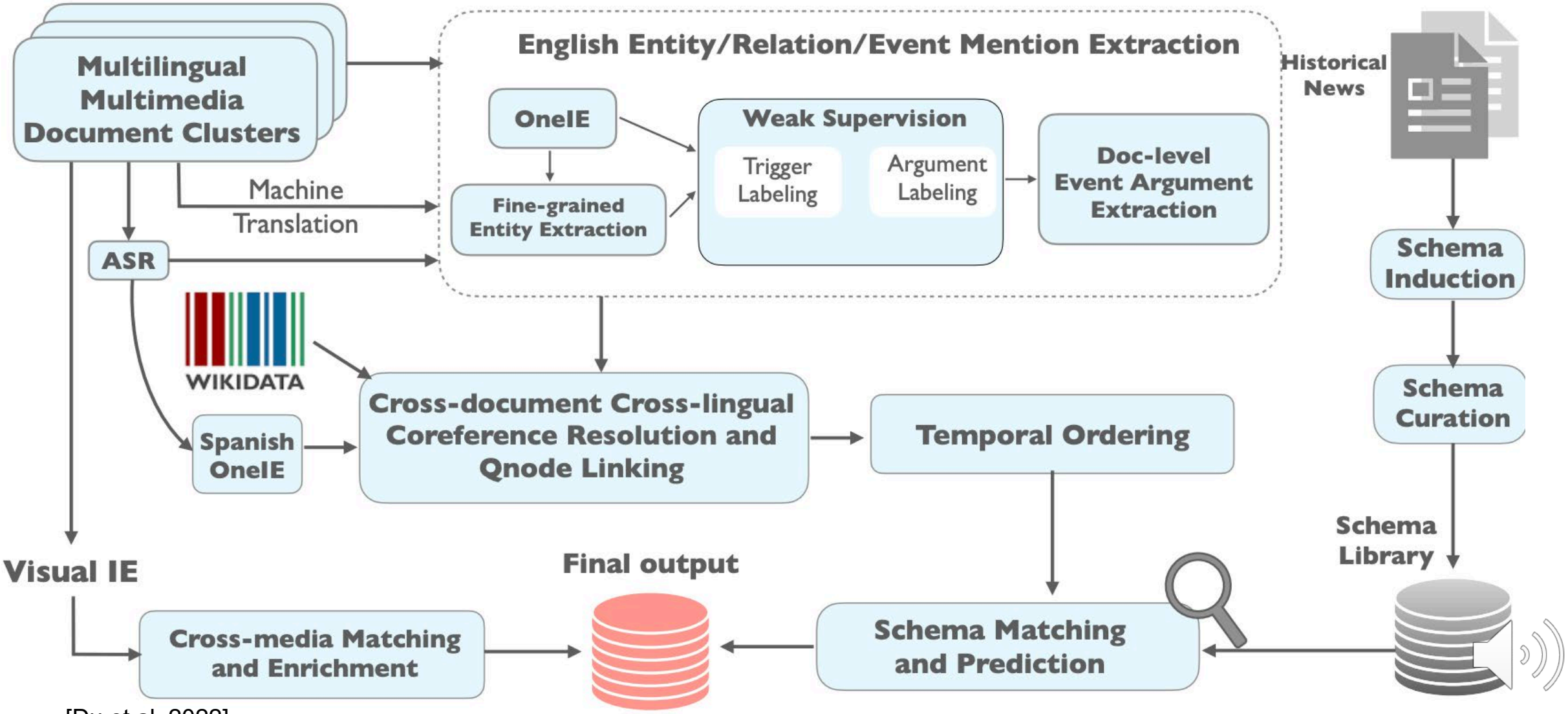
- Step 1: Matching between event graph and schema graph
- Step 2: Decide whether add a node according to the schema graph, based on:
  - neighbors of the candidate node and the matched subgraph
  - paths that connect the candidate node and the matched subgraph.



[Wang et al, 2022]

Schema-Guided Event Prediction: RESIN-11

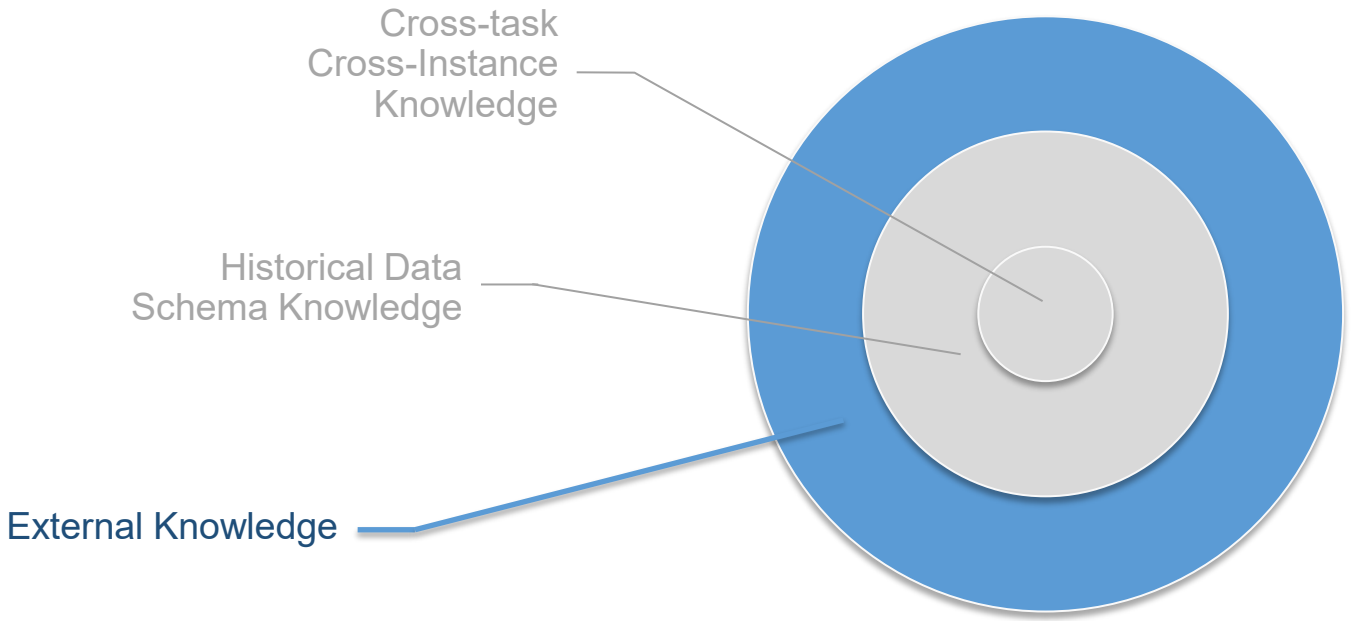
Dockerized system publicly available at Github: https://github.com/RESIN-KAIROS/RESIN-11



[Du et al, 2022]

Knowledge-Enhanced IE

- What knowledge is useful for information extraction?
  - cross-task cross-instance knowledge such as the interactions between knowledge elements
  - schema knowledge induced from historical data
  - external knowledge such as commonsense knowledge



Commonsense Knowledge Enhanced IE

- Human can construct a latent timeline about events' start and end times

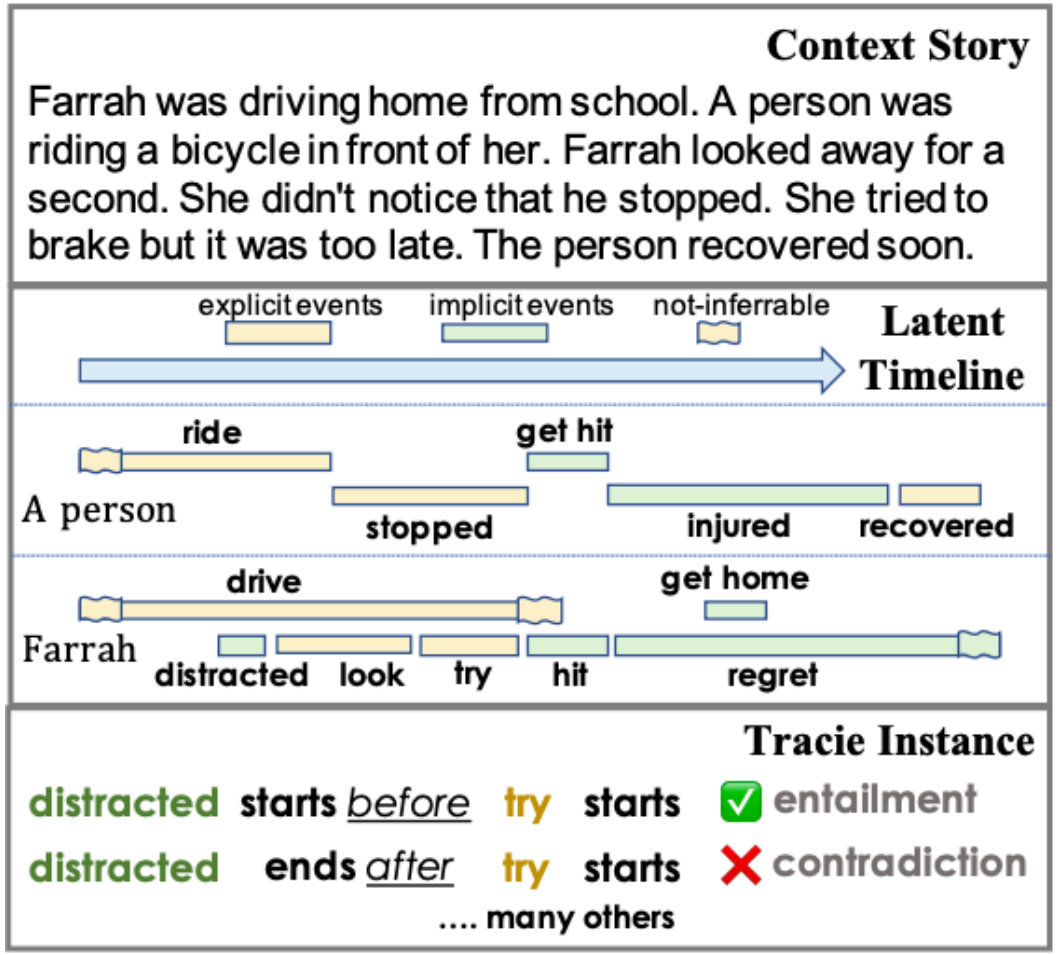
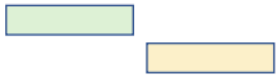
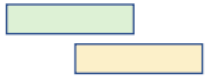
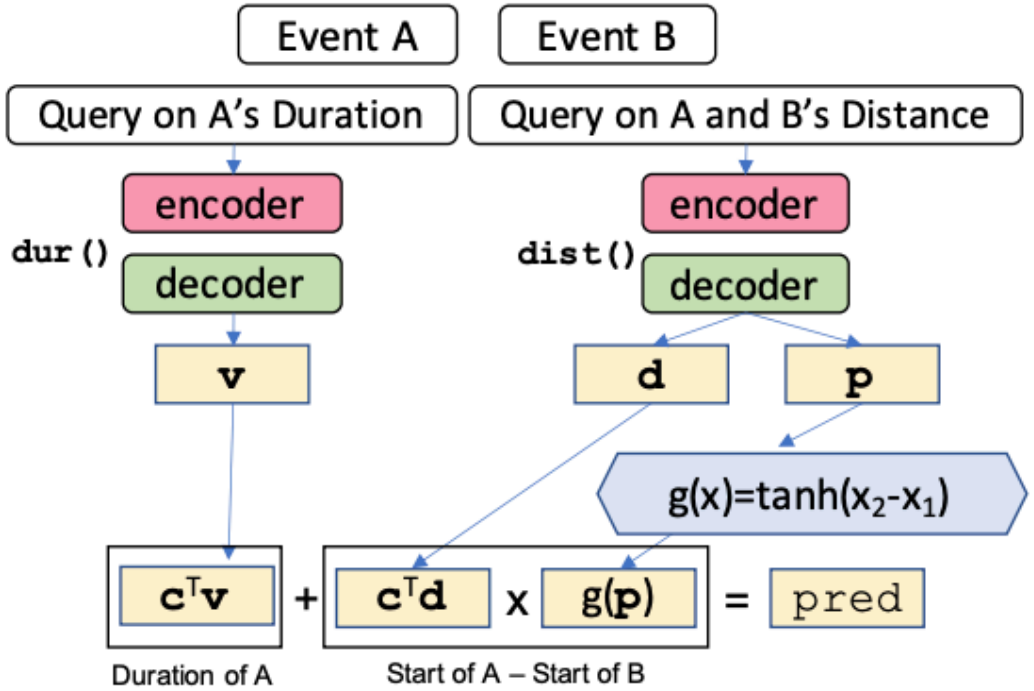


Illustration	Allen's Relation	Tracie's Relation
	Precedes, Meets	Starts Before Ends Before
	Overlaps, Finished-by, Contains, Starts, Equals, Started-by	Starts Before Ends After
	During, Finishes, Overlapped-by, Met-by, Preceded-by	Starts After Ends After

[Zhou et al, 2021]

Commonsense Knowledge Enhanced IE

- Key idea: Leveraging duration



text

I went to the park on January 1st. I was very hungry after some hiking. Luckily, I purchased a lot of food before I went to the park. I enjoyed the trip and wrote an online review about the trip on the 10th.

within-sentence

[I purchased food, I went to the park.]: **before**

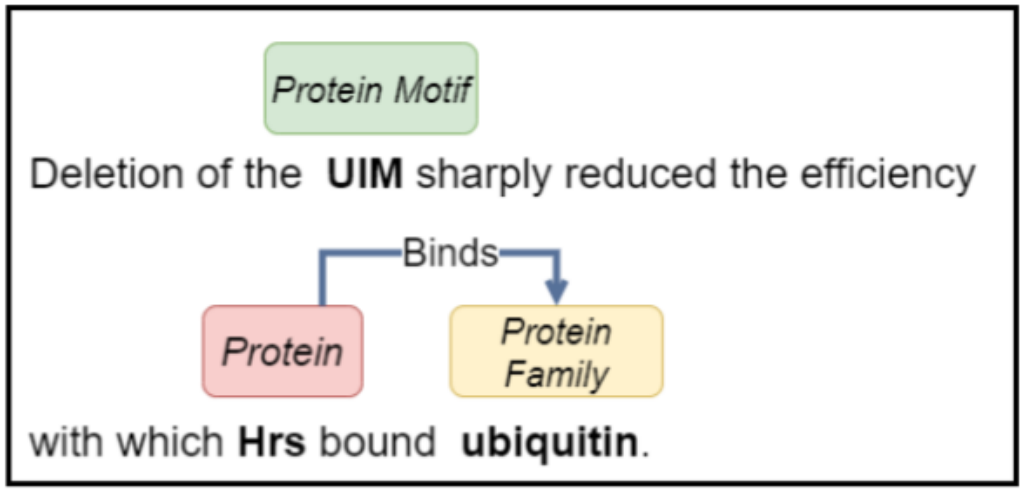
cross-sentence

[I went to the park, I wrote a review]: **before**, weeks

System	Start	End	All	Story
Majority	57.3	69.8	64.1	18.1
BiLSTM	53.7	63.5	59.1	10.9
Roberta-Large	78.5	78.3	78.4	26.1
T5-3B	79.4	77.4	78.3	26.9
BaseLM (T5-large)	75.5	75.4	75.4	22.6
BaseLM-MATRES	76.7	76.3	76.5	25.3
PTNTIME (ours)	81.4	77.5	79.3	31.0
SYMTIME (ours)	**82.1**	**79.4**	**80.6**	**32.0**
SYMTIME-ZEROSHOT	77.0	73.1	74.9	21.6

[Zhou et al, 2021]

External Knowledge Enhanced IE



UIM is an abbreviation of “Ubiquitin-Interacting Motif”.

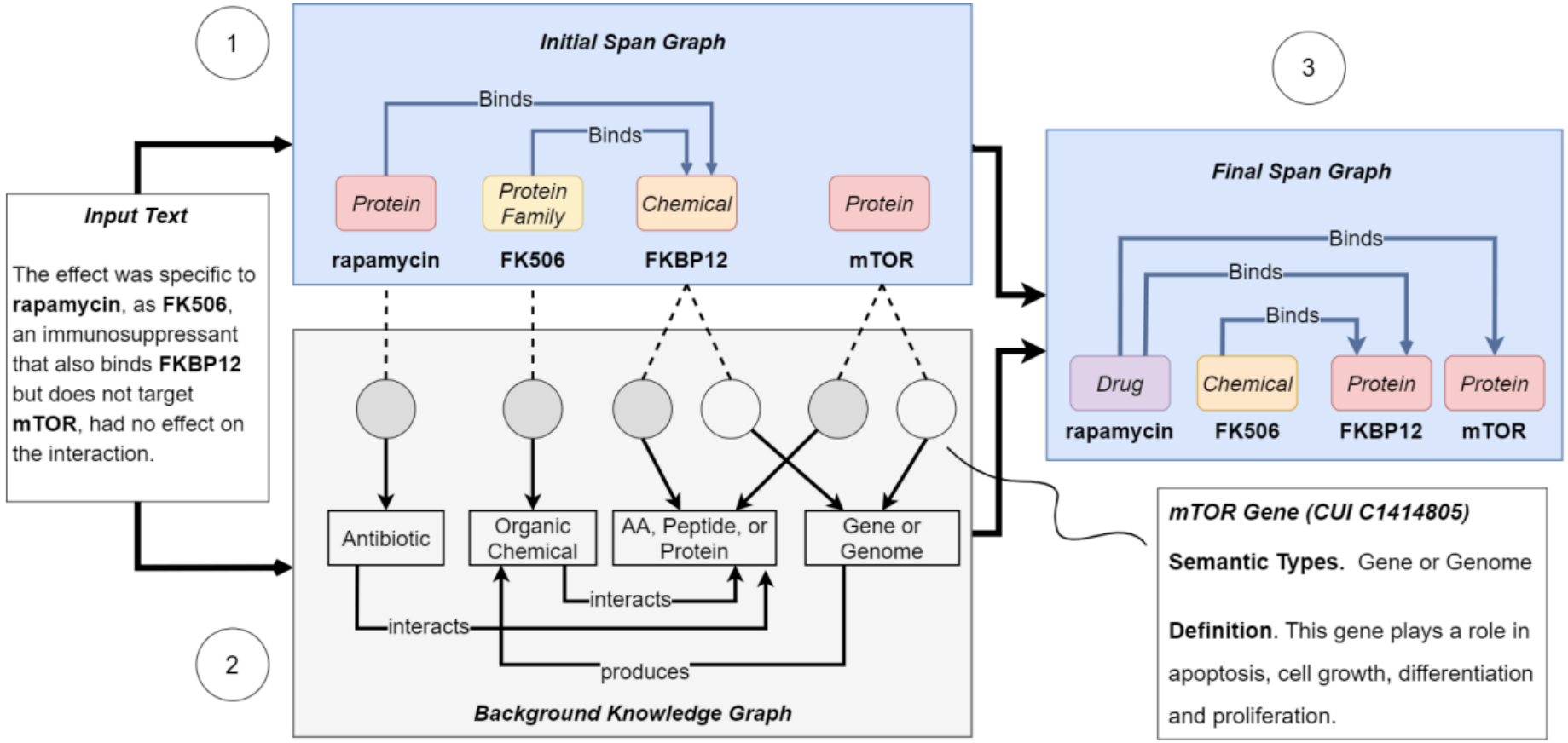
Our baseline SciBERT model incorrectly predicts the mention as a “DNA”.

- IE from biomedical text requires broad domain knowledge.
  - Many highly specialized terms, acronyms, and abbreviations.
- We introduce **KECI** (Knowledge-Enhanced Collective Inference), an end-to-end framework that utilizes external domain knowledge for joint entity and relation extraction.

[Lai et al, 2021]

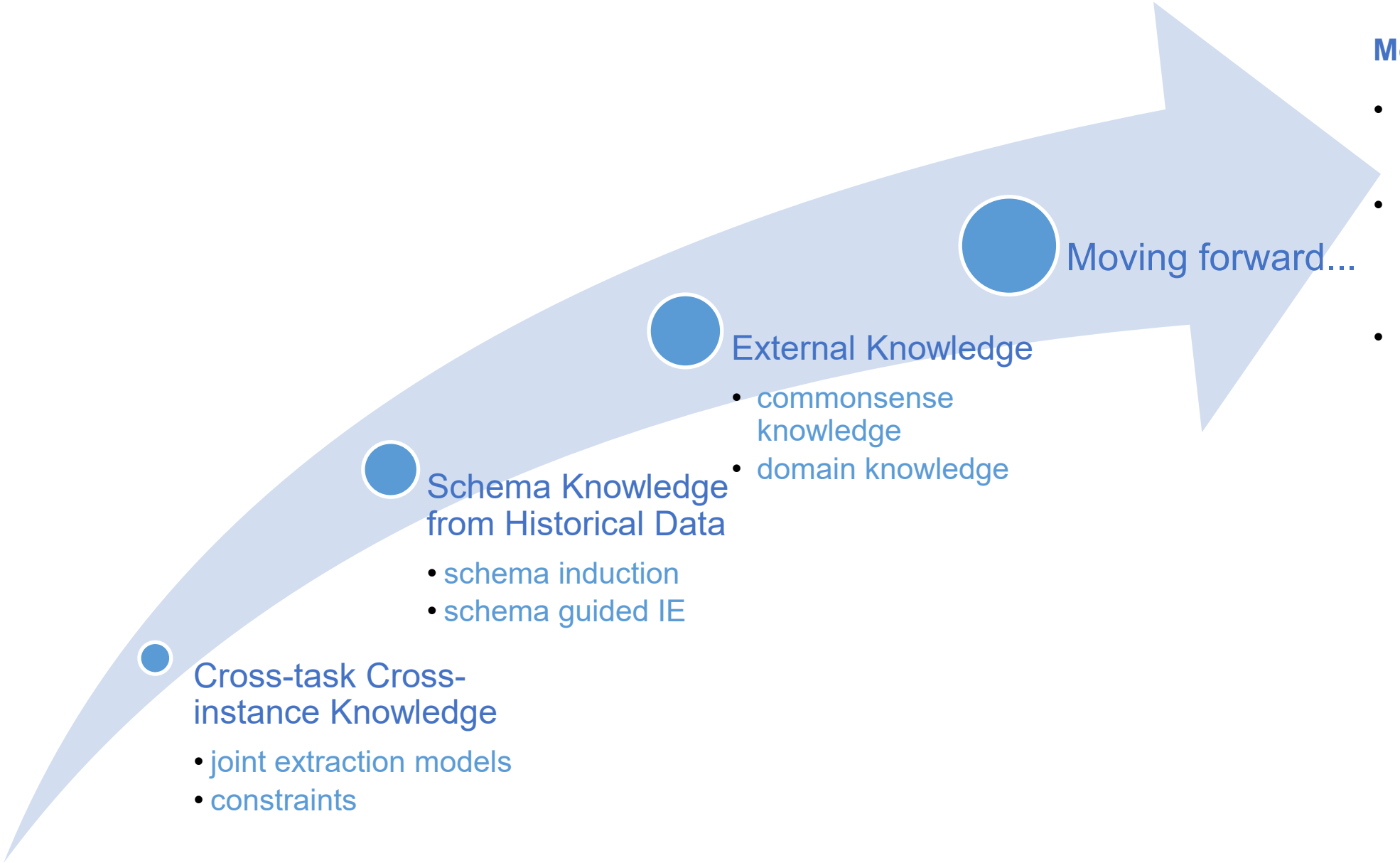
External Knowledge Enhanced IE

- Idea: linking to build a background knowledge graph containing all potentially relevant biomedical entities from an external KB.



[Lai et al, 2021]

Conclusions & Future Direction



Moving forward…

- How to induce schema knowledge?
- How to apply knowledge of different forms?
- How to capture wider and more global context, such as resolving corpus-level coreference during knowledge reasoning?

Thank You