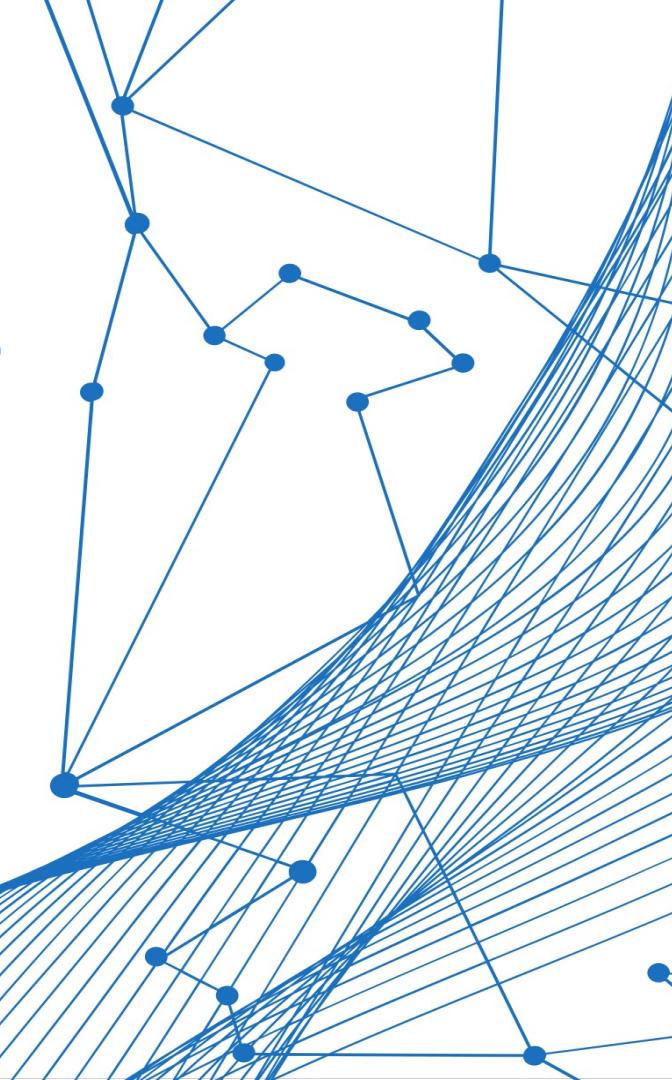


# Applications of Advanced/Deep Neural Networks in Knowledge Graphs Construction

Tek Raj Chhetri

08 April 2023

Faculty Development Program (Online) organised by  
Department of Computer Science & Engineering, Anil  
Neerukonda Institute of Technology and Sciences  
in association with Shodhguru Innovation &  
Research Labs



# About me

Tek Raj Chhetri

PhD Student | Instructor | University of Innsbruck

Research: Knowledge Graphs, Data Privacy (GDPR), Machine Learning,  
Predictive analytics, Data Sharing, Autonomous Vehicles, Cloud Computing

Web: <https://tekrajchhetri.com> | <https://www.cair-nepal.org>

# Demo source codes

[https://github.com/CAIRNepal/talks\\_tutorials\\_lectures/tree/main/fdd\\_india\\_2023\\_08\\_april](https://github.com/CAIRNepal/talks_tutorials_lectures/tree/main/fdd_india_2023_08_april)

# Outline

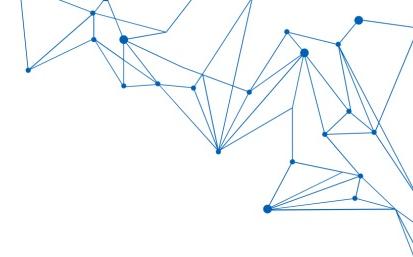
1. Introduction
2. Knowledge Graphs Construction
3. Advanced/Deep neural networks (DNN) for Knowledge Graphs Construction
4. Our Work
5. Conclusion

# 1.

## Introduction

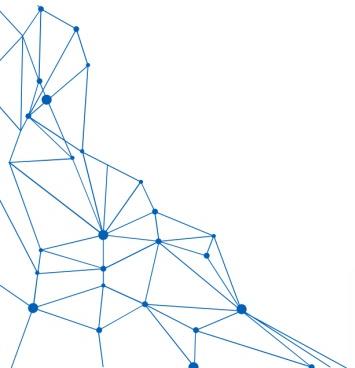
### *1.1. Knowledge Graphs*

# 1.1. Knowledge Graphs

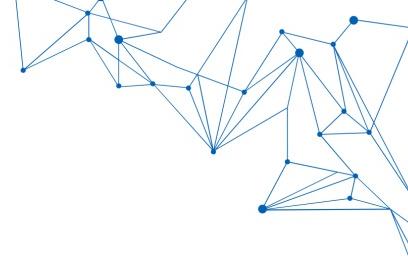


“A graph of data intended to accumulate and convey knowledge of the real world, whose nodes represent entities of interest and whose edges represent potentially different relations between these entities.”

-- Hogan et al., 2021

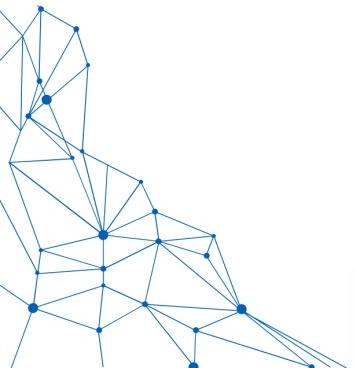


# 1.1. Knowledge Graphs

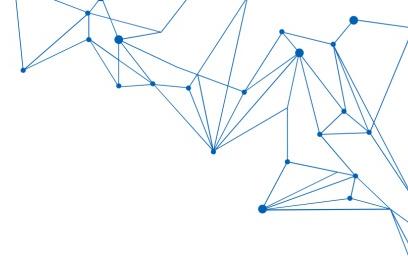


“A **graph** of data intended to accumulate and convey knowledge of the real world, whose nodes represent entities of interest and whose edges represent potentially different relations between these entities.”

-- Hogan et al., 2021

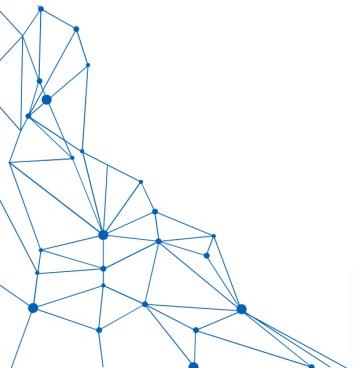


# 1.1. Knowledge Graphs

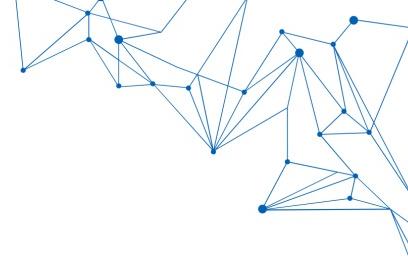


“A graph of data intended to accumulate and convey knowledge of the real world, whose nodes represent entities of interest and whose edges represent potentially different relations between these entities.”

-- Hogan et al., 2021

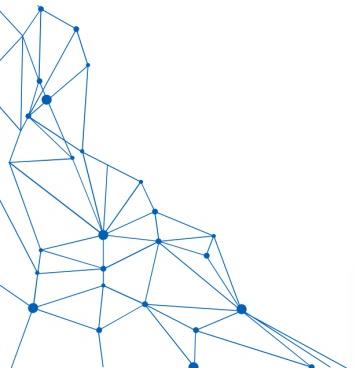


# 1.1. Knowledge Graphs

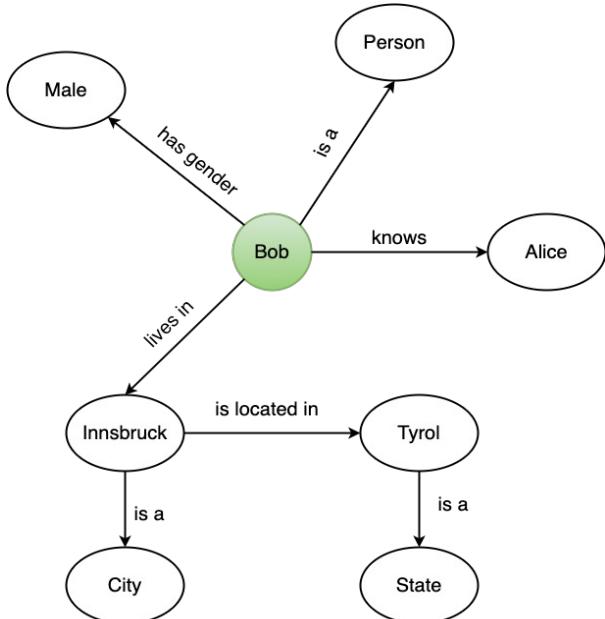
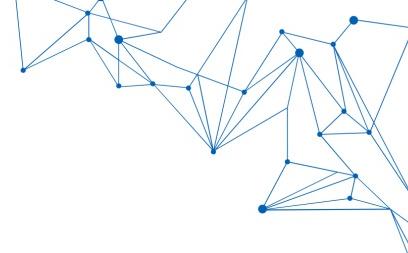


“A graph of data intended to accumulate and convey knowledge of the real world, whose nodes represent entities of interest and whose edges represent potentially different relations between these entities.”

-- Hogan et al., 2021



# 1.1. Knowledge Graphs



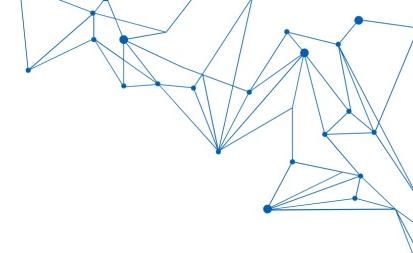
graph?

accumulate and convey  
knowledge?

nodes represent entities of interest  
and whose edges represent  
potentially different relations between  
these entities?

Figure 1: Example knowledge graph

# 1.1. Knowledge Graphs



Sources of knowledge graphs.

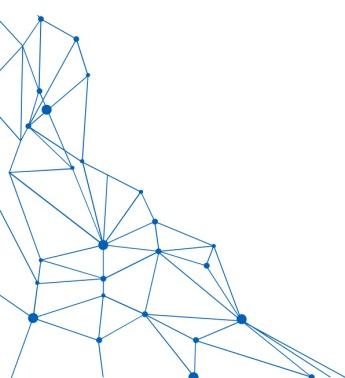
## Structured text

Tables, Wikipedia,  
Databases

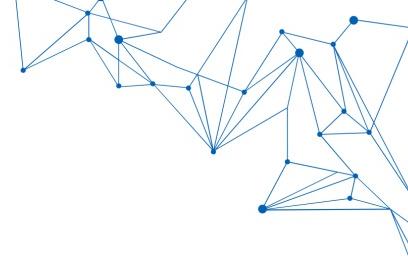
## Unstructured text

Internet, News

Images

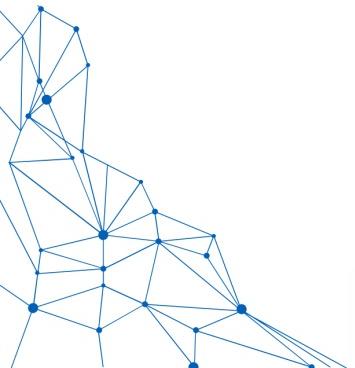


# 1.1. Knowledge Graphs

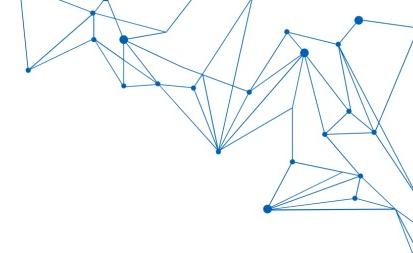


Examples of large knowledge graphs:

- Google knowledge graph: used for Google search.
- WordNet: large lexical database of English language.



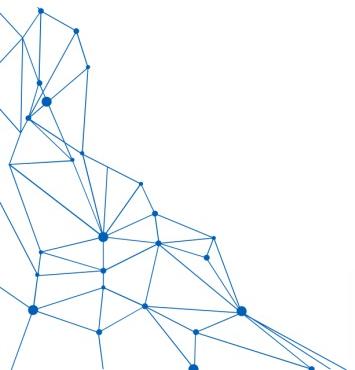
# 1.1. Knowledge Graphs



Where are knowledge graphs used?

- Healthcare, e.g., drug discovery
- Manufacturing, e.g., predictive maintenance
- Search

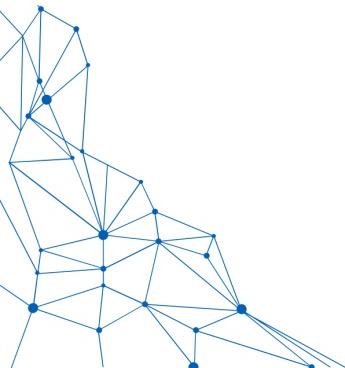
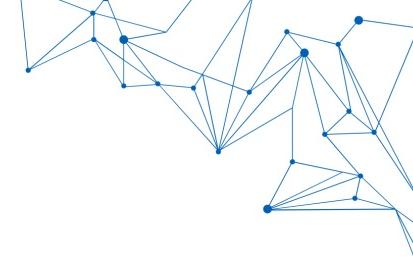
Who is using knowledge graphs?



# 1.

## Introduction

### *1.2. Knowledge Graphs Creation Methodology*



## 1.2. Knowledge Graphs Creation Methodology

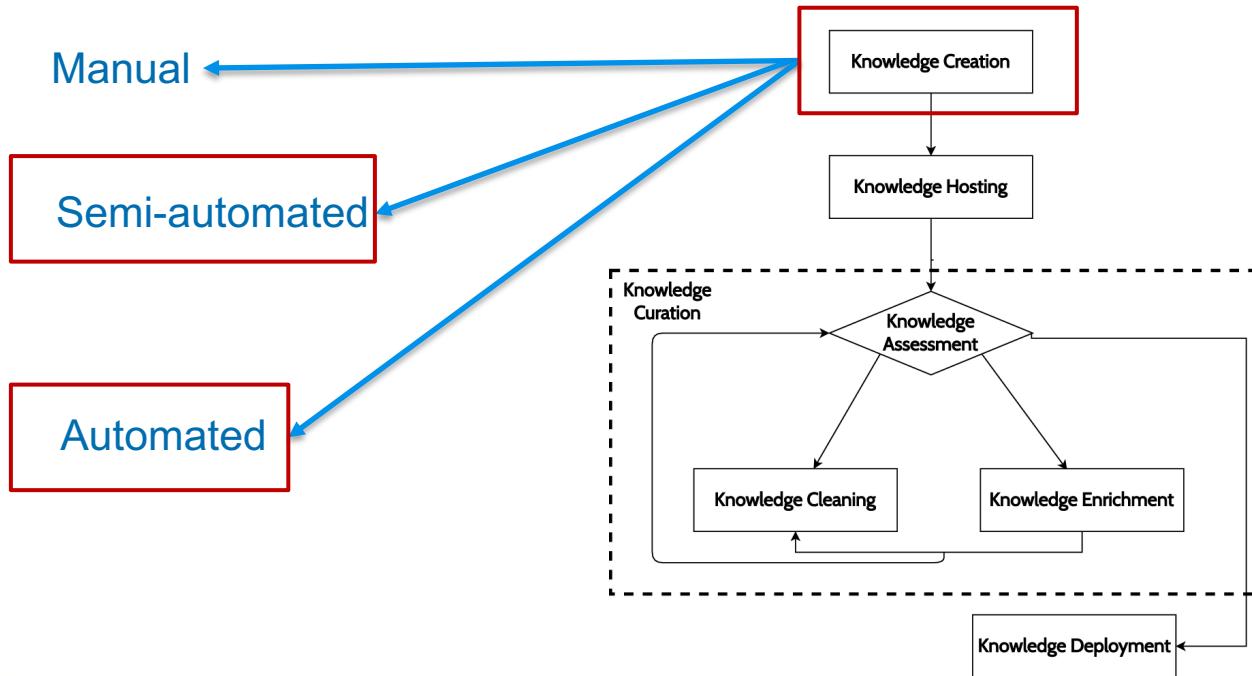


Figure 2: Methodology for building knowledge graphs (Simsek et al., 2022)

## 1.2. Knowledge Graphs Creation Methodology

Bottom-up: annotating the reference ontology, such as Schema.org.

Top-down: a large scale instance generation task that is specific to certain domain-specific patterns, e.g., annotating using NLP (Natural language processing).

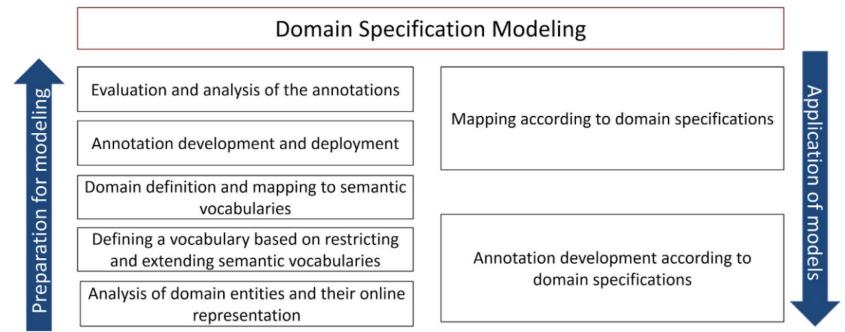
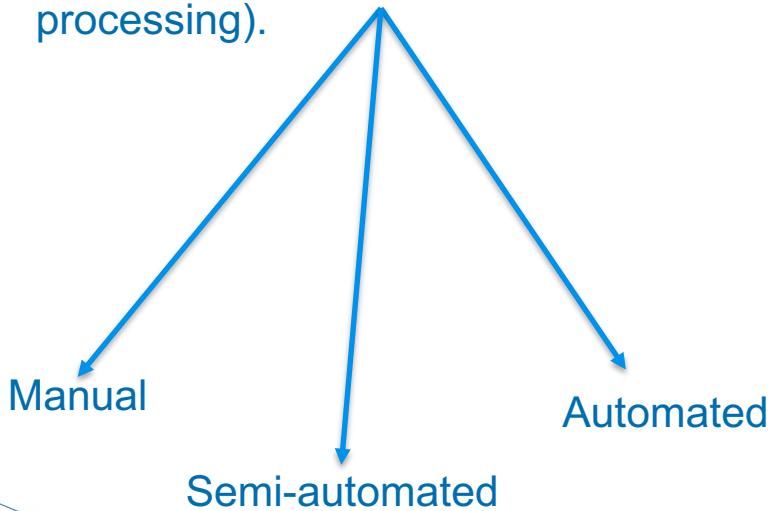


Figure 3: Knowledge creation approach  
(Simsek et al., 2022)

# 2.

## Knowledge Graphs Construction

### *2.1. Knowledge Graphs Construction Steps*

## 2.1. Knowledge Graphs Construction Steps

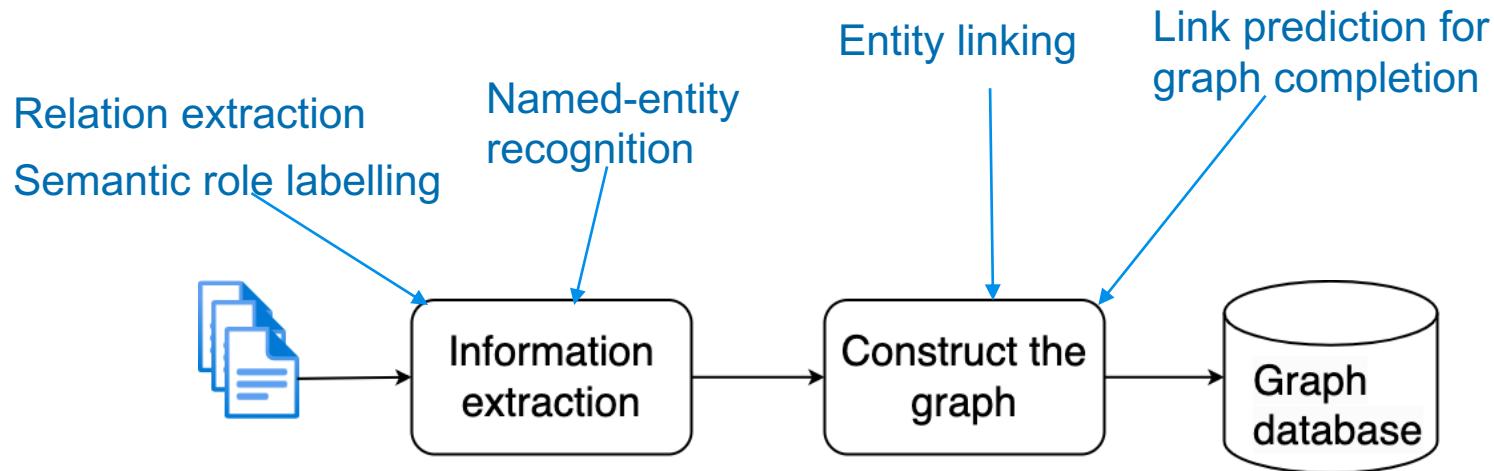


Figure 3: Steps for knowledge graph construction

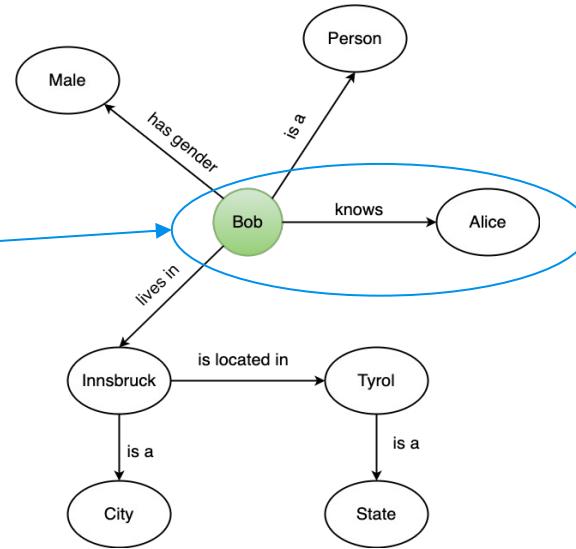
# 2.

## Knowledge Graphs Construction

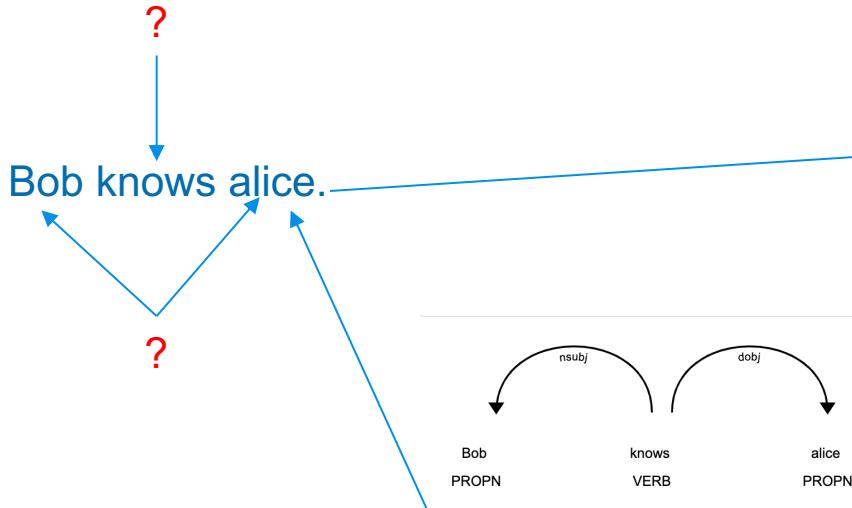
### *2.2. Linguistic Structure for Knowledge Graphs Construction*

## 2.2. Linguistic Structure for Knowledge Graphs Construction

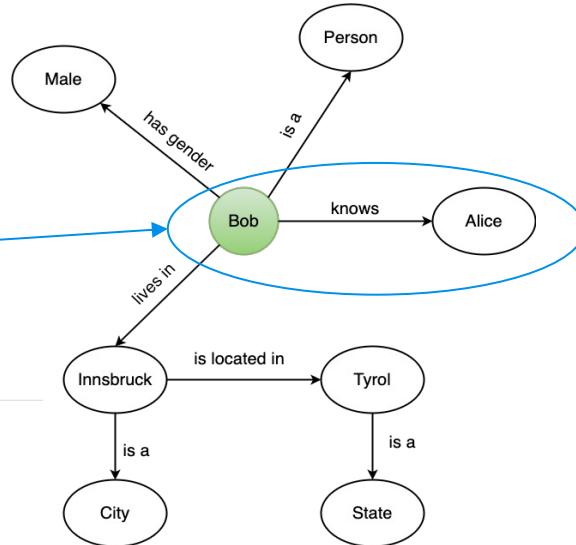
?  
↓  
Bob knows alice.  
?  
? ↗



## 2.2. Linguistic Structure for Knowledge Graphs Construction



```
: spacy.explain("PROPN"), spacy.explain("dobj")
: ('proper noun', 'direct object')
```



## 2.2. Linguistic Structure for Knowledge Graphs Construction

Can we leverage our knowledge about language to construct knowledge graphs?

## 2.2. Linguistic Structure for Knowledge Graphs Construction

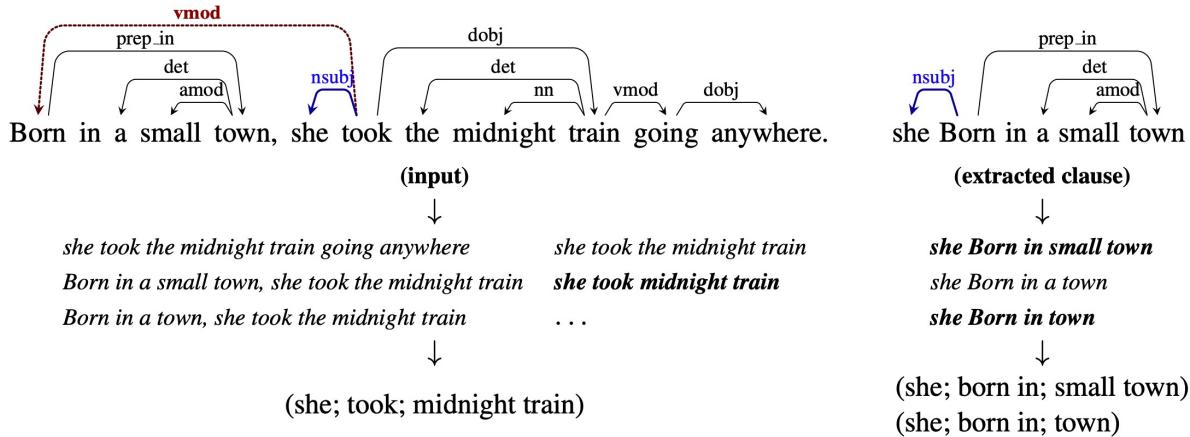
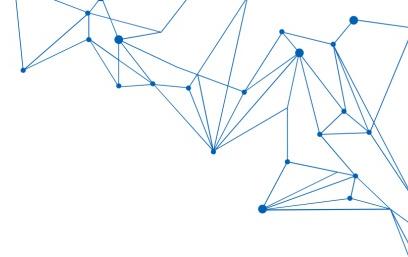


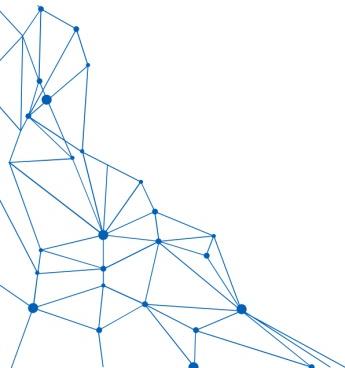
Figure 3: Illustration of OpenIE (Open Domain Information Extraction) approach (Angeli et al., 2015)

## 2.2. Linguistic Structure for Knowledge Graphs Construction



Demo

Notebook file: OpenIE-Demo



## 2.2. Linguistic Structure for Knowledge Graphs Construction

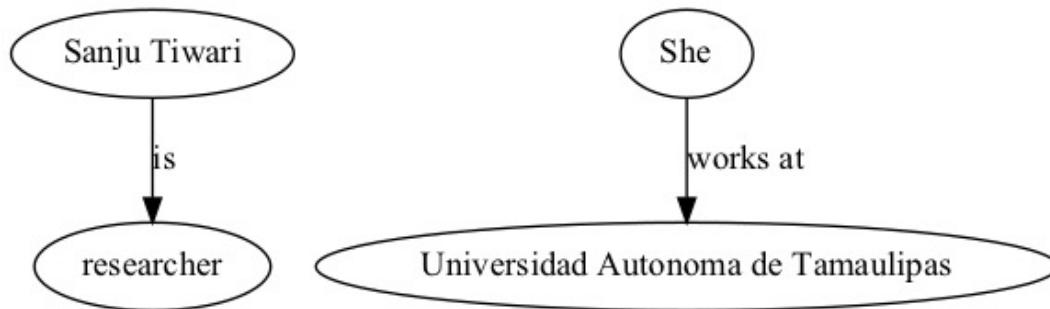


Figure 4: Generated knowledge graphs

## 2.2. Linguistic Structure for Knowledge Graphs Construction

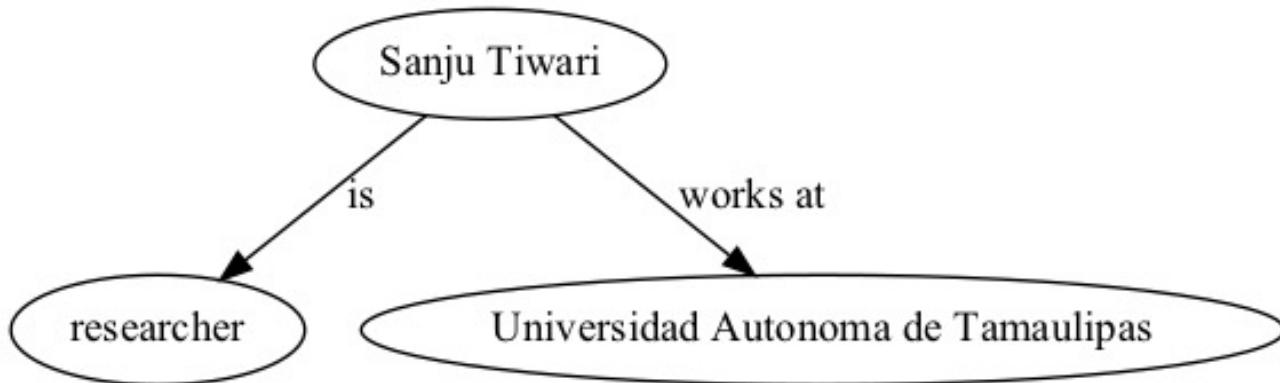


Figure 5: Generated knowledge graph

## 2.2. Linguistic Structure for Knowledge Graphs Construction

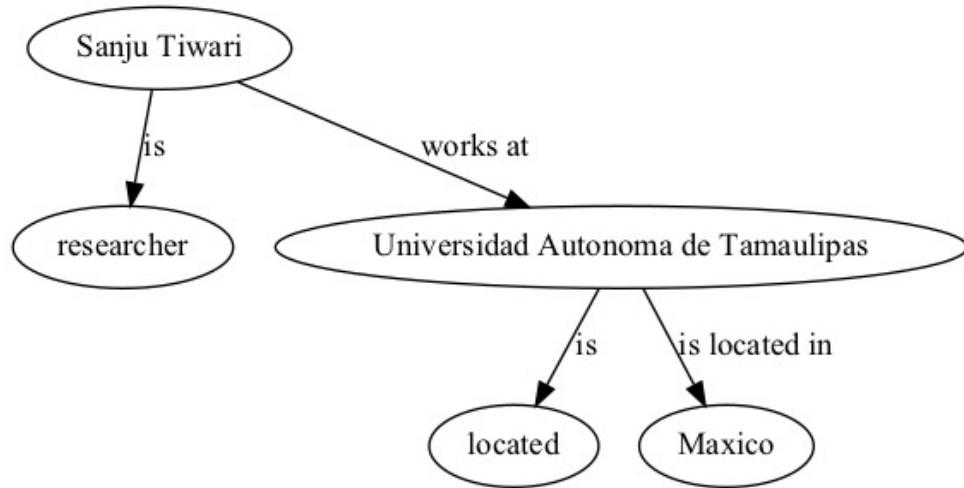


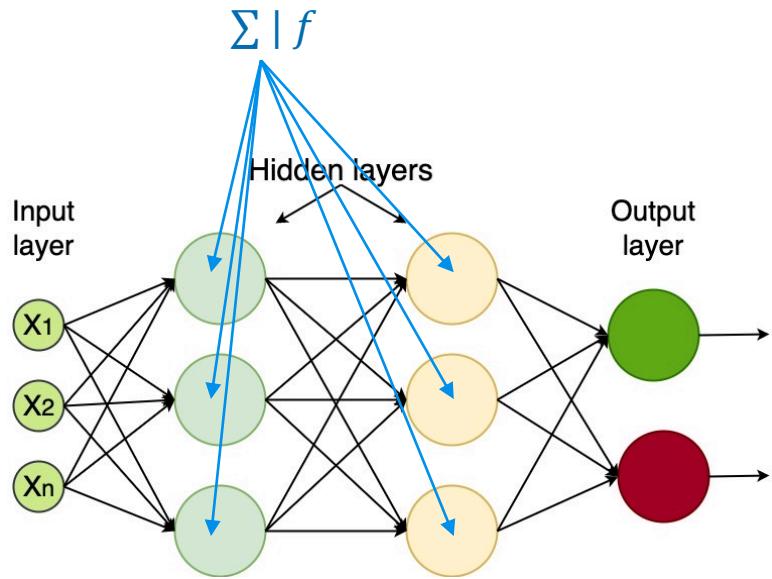
Figure 6: Generated knowledge graph

# 3.

## Advanced/Deep Neural Networks (DNN) for Knowledge Graphs Construction

### *3.1. Overview of Advanced/DNN*

## 3.1. Overview of Advanced/DNN



$$f(W * X + b)$$

$$\text{sigmoid} = \frac{1}{1 + e^{-x}}$$

Figure 7: Neural network

## 3.1. Overview of Advanced/DNN



Today, transformers are the state-of-the art in many tasks including language processing.

Two types of language modelling task:

1. Auto-regressive models: predict the future word given past or future, not both.  
Vienna is capital of \_\_\_\_\_. Given past, we want to predict future.  
\_\_\_\_ is in Austria. Given future, we want to predict past.  
E.g.: GPT family – used for natural language generation.
2. Auto-encoding models: learn the representation of sentence by predicting tokens given past and future.

He doesn't \_\_\_\_\_ homework, he will be punished.

E.g.: BERT – used for natural language understanding tasks like named entity recognition, sequence classification.



# 3.1. Overview of Advanced/DNN

The foundation of the modern transformers was introduced Vaswani et al. in 2017.

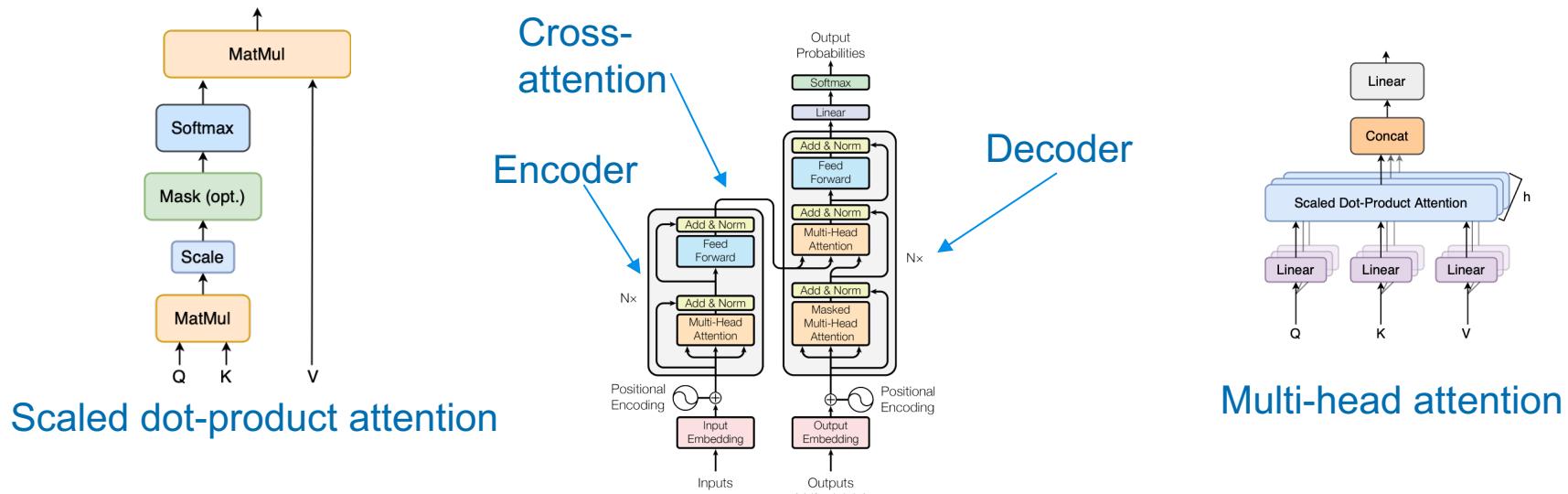
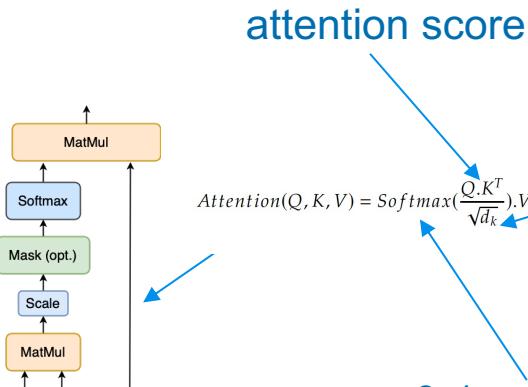


Figure 8: The Transformer - model architecture (Vaswani et al., 2017)

# 3.1. Overview of Advanced/DNN

$Q$  = query,  $K$  = key and  $V$  = view.  $Q$ ,  $K$  and  $V$  are matrices.



Scaled dot-product attention

make value smaller

0-1 range

$$\begin{aligned} \text{MultiHead}(Q', K', V') &= \text{Concat}(\text{Attention}(Q_1, K_1, V_1), \text{Attention}(Q_2, K_2, V_2), \\ &\dots, \text{Attention}(Q_n, K_n, V_n)).W^O \\ \text{head}_i &= \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \end{aligned}$$

32

Multi-head attention

## 3.1. Overview of Advanced/DNN

Sentence: Jane is good citizen.

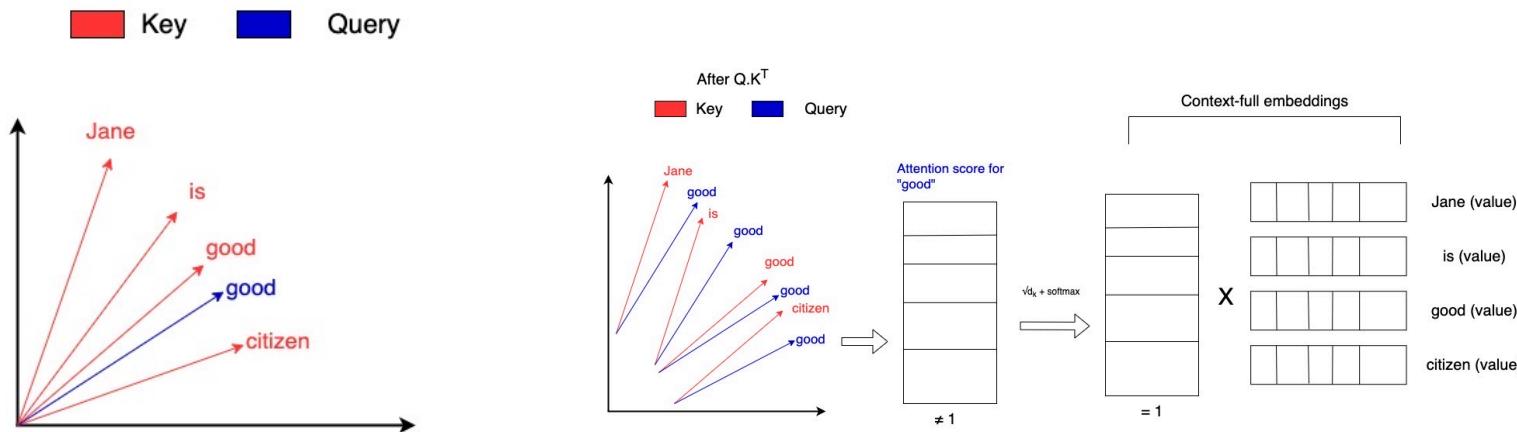


Figure 9: Scaled dot-product attention example

# 3.1. Overview of Advanced/DNN

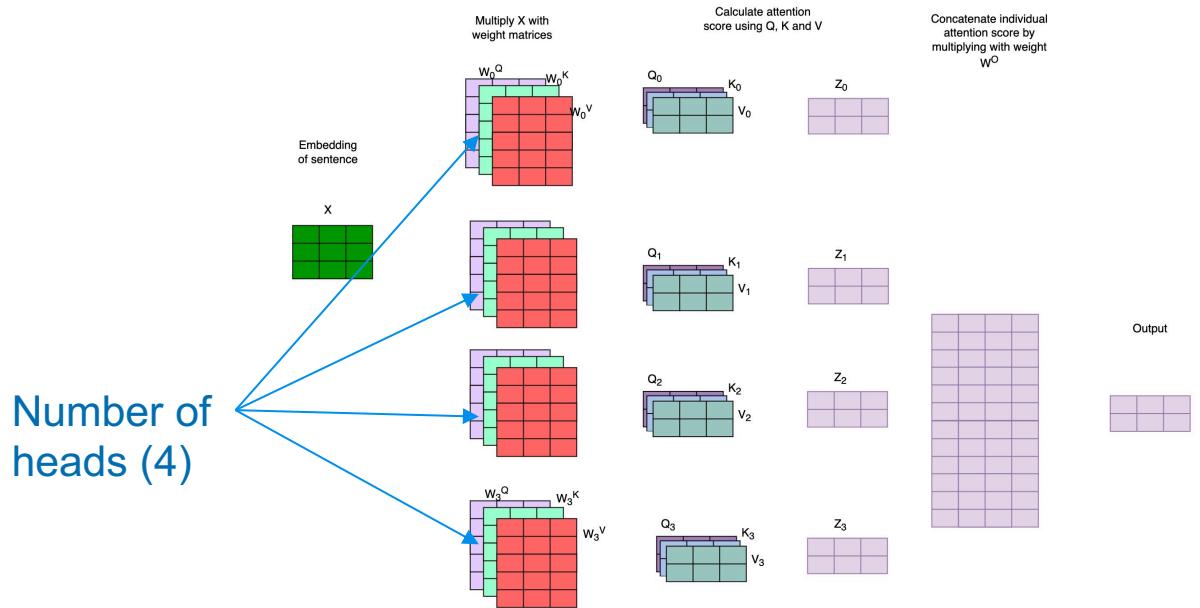


Figure 10: Multi-head attention example

## 3.1. Overview of Advanced/DNN

Input sentence: My friends told me about attention paper and I enjoyed reading it.

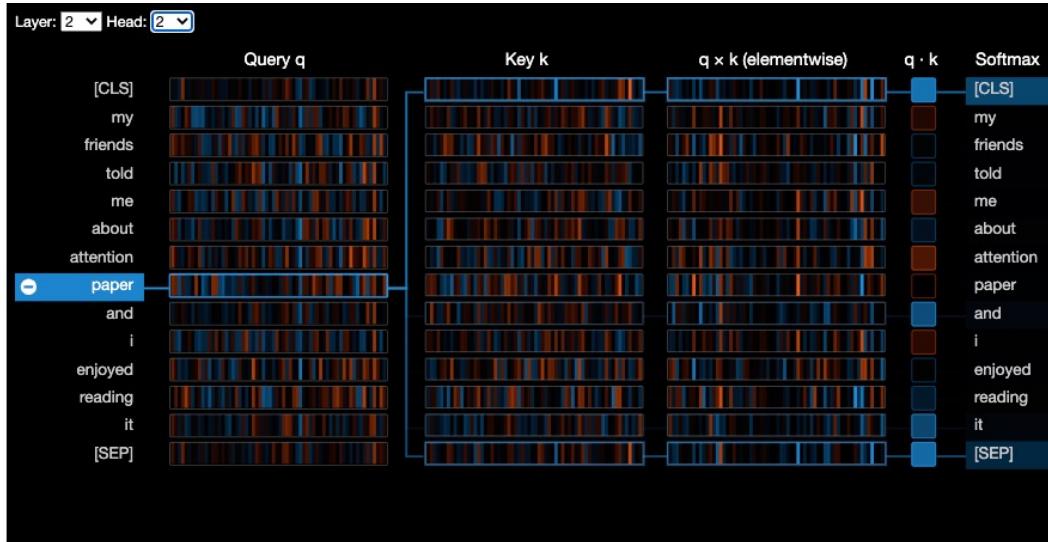


Figure 11: Visualisation of Bidirectional Encoder Representations from Transformers (BERT) neuron view

# 3.1. Overview of Advanced/DNN

Input sentence: My friends told me about attention paper and I enjoyed reading it.

	my	friends	told	me	about	attention	paper	and	i	enjoyed	reading	it
my	0.100934	0.069470	0.117299	0.120445	0.079473	0.046395	0.052692	0.073218	0.107774	0.114604	0.055917	0.061779
friends	0.038067	0.096476	0.117915	0.049149	0.067461	0.113744	0.137340	0.038570	0.027052	0.174089	0.107412	0.032725
told	0.279546	0.071563	0.141285	0.025596	0.064213	0.028561	0.044704	0.079137	0.071826	0.056287	0.106594	0.030688
me	0.080006	0.124975	0.148152	0.069366	0.097184	0.089837	0.042953	0.120903	0.067797	0.093817	0.055603	0.009408
about	0.114902	0.089267	0.121241	0.048999	0.095998	0.082544	0.055986	0.056864	0.039855	0.165398	0.061551	0.067397
attention	0.061192	0.103999	0.109819	0.031648	0.078422	0.134510	0.072197	0.046429	0.028284	0.172710	0.124742	0.036049
paper	0.086109	0.106978	0.110281	0.069261	0.056677	0.085790	0.087418	0.027939	0.062958	0.124162	0.144700	0.037727
and	0.179758	0.109041	0.072173	0.130123	0.047328	0.065694	0.038095	0.042239	0.094282	0.067172	0.058225	0.095871
i	0.045420	0.076360	0.073031	0.061148	0.060051	0.120220	0.117625	0.067718	0.073002	0.129334	0.099807	0.076283
enjoyed	0.232717	0.128013	0.066172	0.089593	0.033107	0.060076	0.098414	0.029566	0.065224	0.072466	0.095945	0.028707
reading	0.046663	0.017130	0.123428	0.036269	0.124542	0.064962	0.099627	0.072917	0.068067	0.119192	0.130926	0.096278
it	0.189348	0.042706	0.110883	0.047178	0.046489	0.172225	0.065286	0.032828	0.047367	0.148391	0.062988	0.034311

Table 1: Final attention score from BERT

# 3.1. Overview of Advanced/DNN

The colour indicates head, 12 for BERT because it has 12 heads.

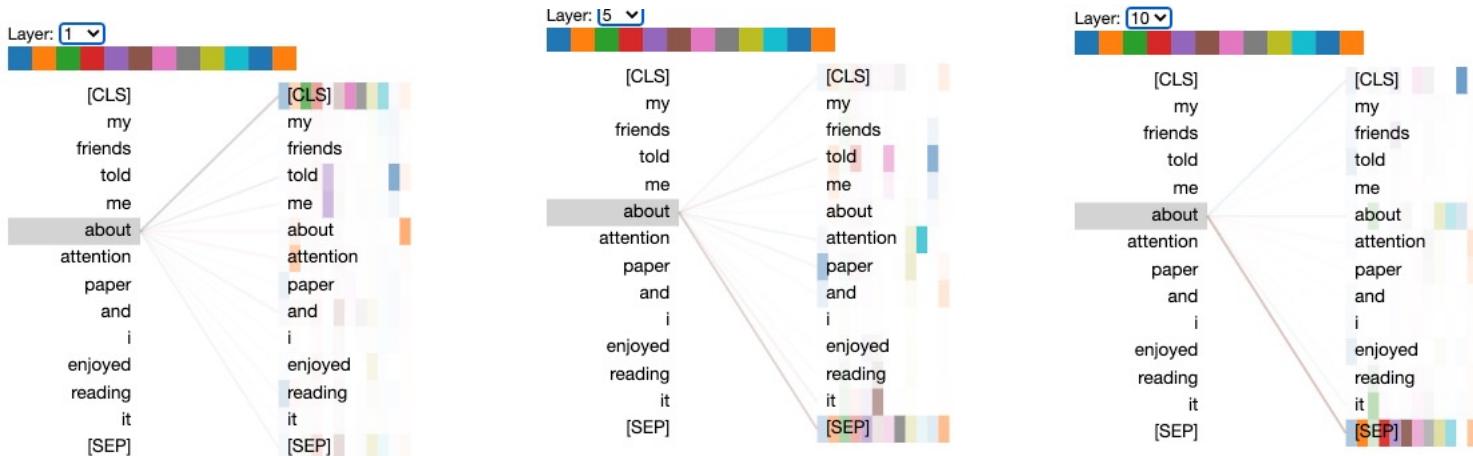


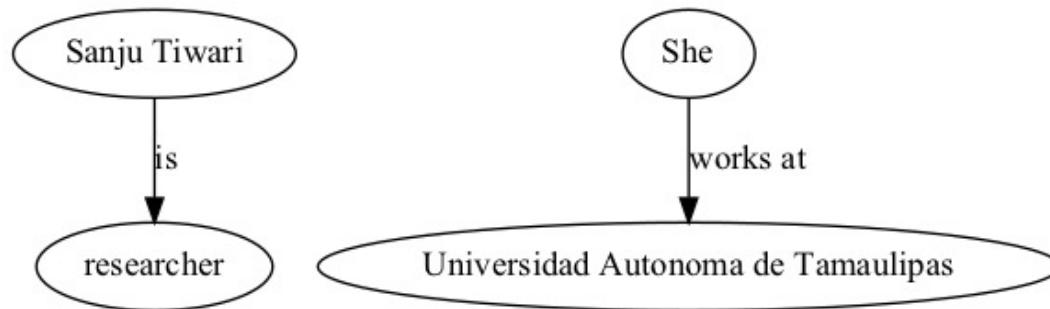
Figure 12: Head view visualisation of BERT

# 3.

## Advanced/Deep Neural Networks (DNN) for Knowledge Graphs Construction

### *3.2. Conference Resolution*

## 3.2. Conference Resolution



## 3.2. Conference Resolution

Coreference resolution is the task of clustering textual mentions that refer to the same discourse entity (Otmarzgin et al., 2022).

## 3.2. Conference Resolution

Jane voted for Obama because he is aligned with her democratic values, she said.

## 3.2. Conference Resolution

Jane voted for Obama because he is aligned with her democratic values, she said.

## 3.2. Conference Resolution

Jane voted for Obama because he is aligned with her democratic values, she said.

## 3.2. Conference Resolution

Jane stated that she voted for Obama because he shares democratic values that align with her.

## 3.2. Conference Resolution

Jane stated that **she** voted for Obama because he shares democratic values that align with **her**.

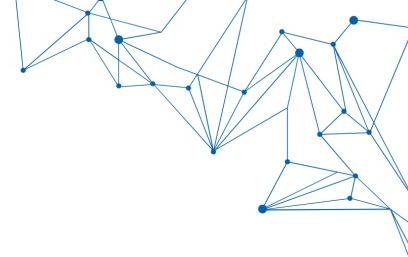
## 3.2. Conference Resolution

Jane stated that she voted for Obama because he shares democratic values that align with her.

## 3.2. Conference Resolution

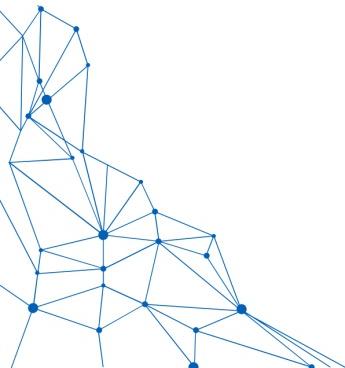
Jane stated that she voted for Obama because he shares democratic values that align with her.

## 3.2. Conference Resolution



Demo

Notebook file: Conference-resolution-demo



## 3.2. Conference Resolution

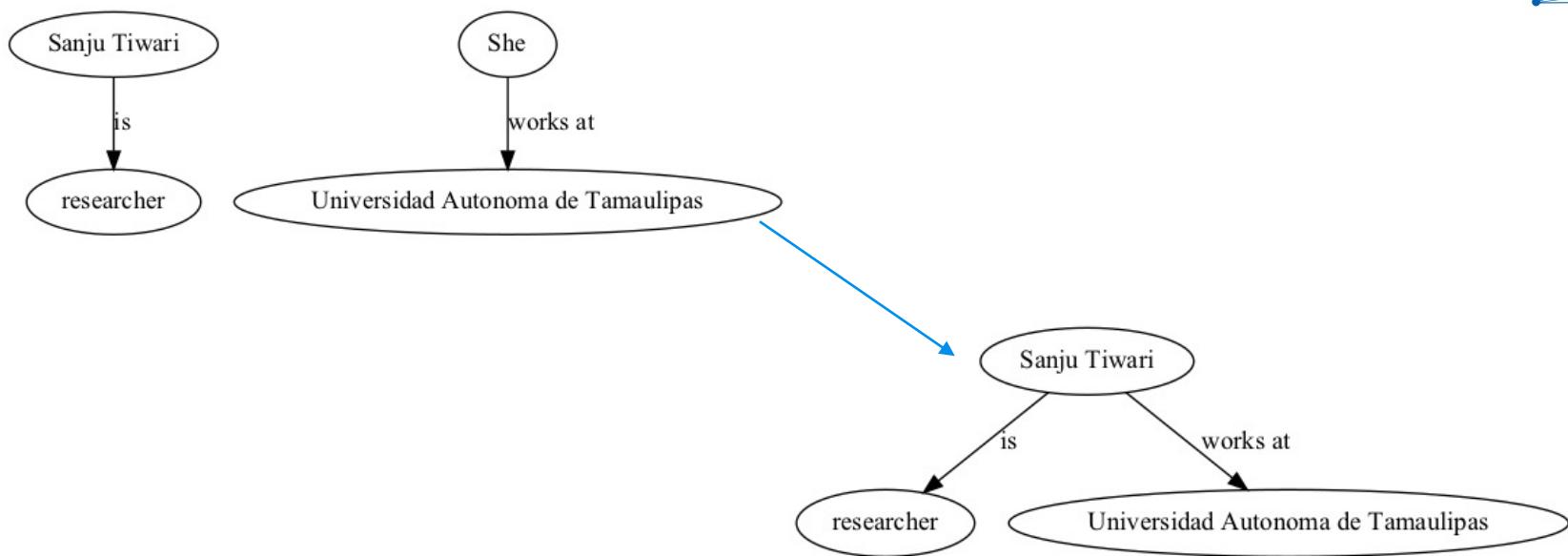
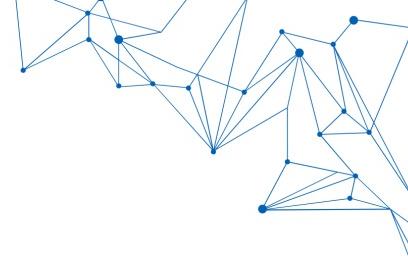


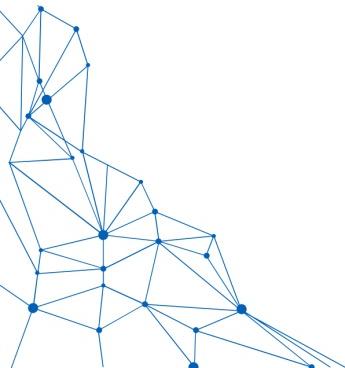
Figure 13: Conference resolution

## 3.2. Conference Resolution



Demo

Notebook file: Conference-resolution-demo



## 3.2. Conference Resolution

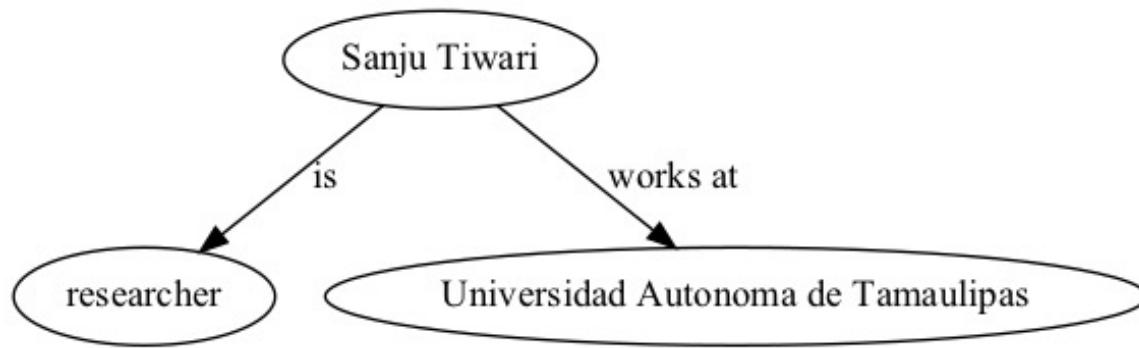
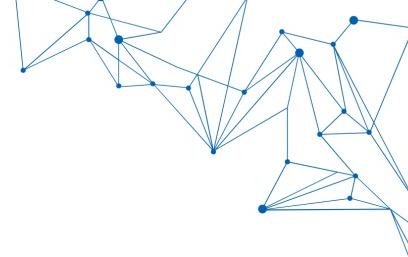
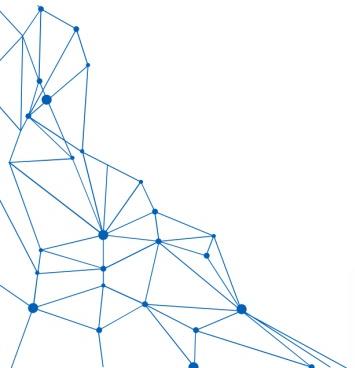


Figure 14: Result knowledge graph after conference resolution

## 3.2. Conference Resolution



Let's create some large knowledge graphs.



## 3.2. Conference Resolution

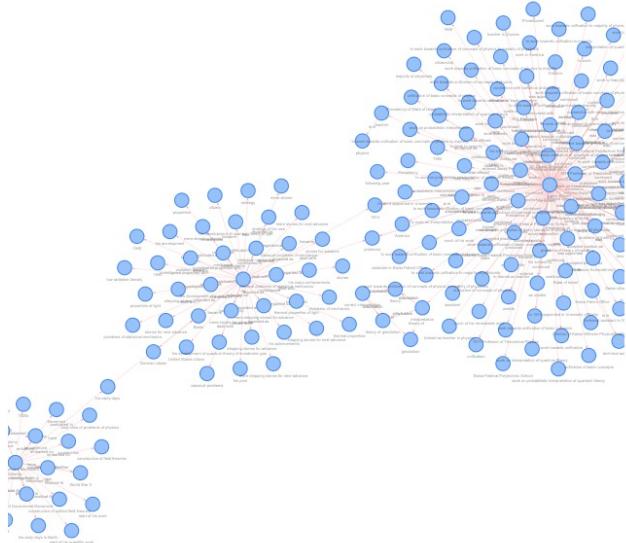


Figure 15: Knowledge graph about Albert Einstein created from raw text (without conference resolution)

## 3.2. Conference Resolution

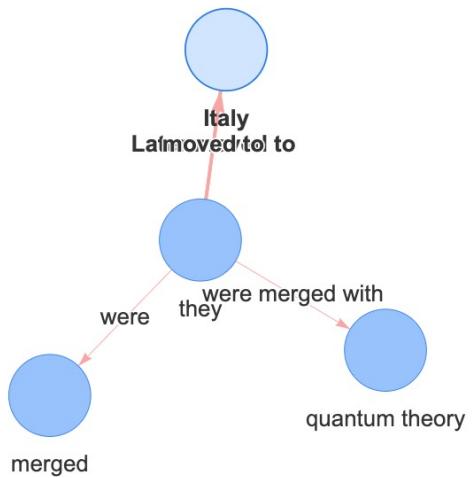


Figure 16: Knowledge graph about Albert Einstein created from raw text (without conference resolution)

## 3.2. Conference Resolution

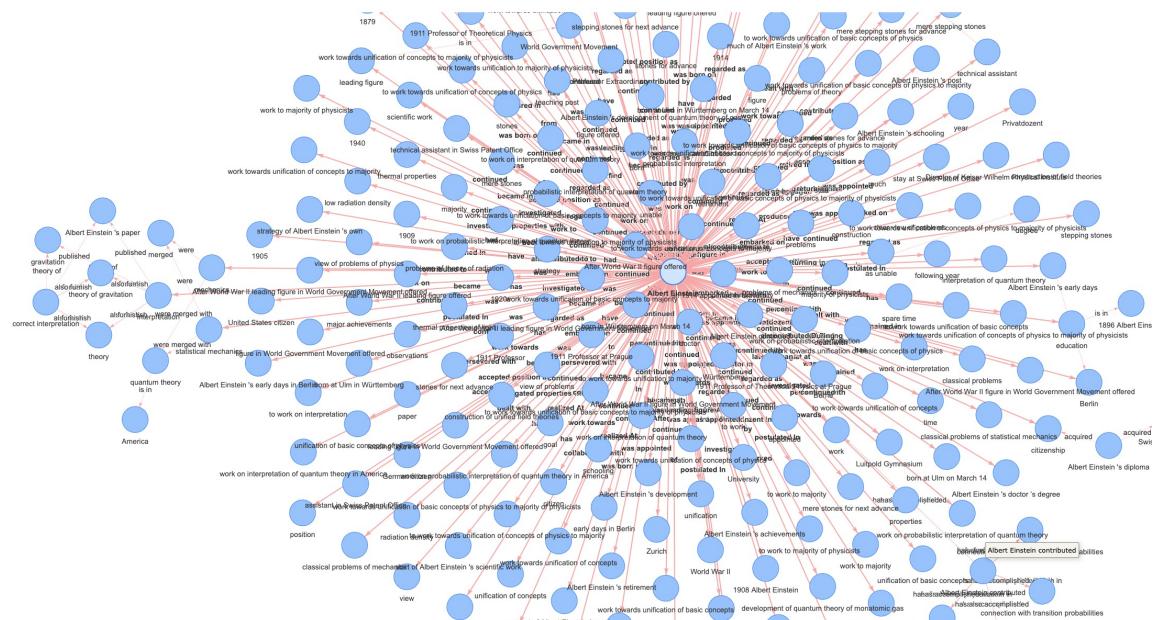


Figure 17: Knowledge graph about Albert Einstein created from raw text (with conference resolution)

# 3.

## Advanced/Deep Neural Networks (DNN) for Knowledge Graphs Construction

### *3.3. Relation Extraction & Entity Linking*

### 3.3. Relation Extraction & Entity Linking

Entity linking is the task of identifying and mapping the entities to the correct entries.

### 3.3. Relation Extraction & Entity Linking

Albert Einstein was born at Ulm, in Württemberg, Germany. Six weeks later Einstein family moved to Munich.

### 3.3. Relation Extraction & Entity Linking

Albert Einstein was born at Ulm, in Württemberg, Germany. Six weeks later Einstein family moved to Munich.

### 3.3. Relation Extraction & Entity Linking

	subject	relation	object	
0	Albert Einstein	born_in	Ulm	?
1	Albert Einstein	born_in	Württemberg	
2	Albert Einstein	born_in	Germany	
3	Einstein	lived_in	Munich	

Figure 18: Knowledge graph triples from text

### 3.3. Relation Extraction & Entity Linking

Albert Einstein was born at Ulm, in Württemberg, Germany. Six weeks later Einstein family moved to Munich.

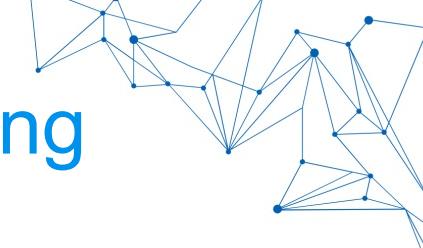
Both refers to the same person, Albert Einstein

### 3.3. Relation Extraction & Entity Linking

	subject	relation	object
0	Albert Einstein	born_in	Ulm
1	Albert Einstein	born_in	Württemberg
2	Albert Einstein	born_in	Germany
3	Albert Einstein	lived_in	Munich

Figure 19: Knowledge graph triples from text

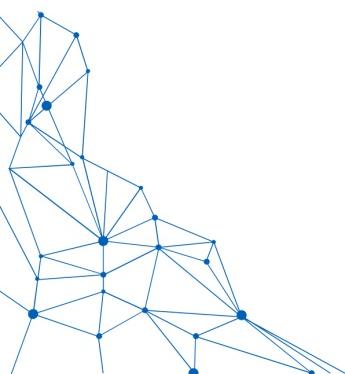
### 3.3. Relation Extraction & Entity Linking



Relation extraction refers to the extraction of the semantic relationship between entities.

It involves two tasks:

- Identifying (or detecting) the relevant entities pair of interest.
- Detecting and extracting the identified relations.



### 3.3. Relation Extraction & Entity Linking

	subject	relation	object
0	Albert Einstein	born_in	Ulm
1	Albert Einstein	born_in	Württemberg
2	Albert Einstein	born_in	Germany
3	Albert Einstein	lived_in	Munich

Figure 20: Knowledge graph triples from text

## 3.3. Relation Extraction & Entity Linking

Demo

Notebook file: `relation_extraction_entity_linking`

### 3.3. Relation Extraction & Entity Linking

REBEL: Relation Extraction By End-to-end Language generation (Cabot et al., 2021)

- An autoregressive approach that frames relation extraction as a seq2seq task.
- Pretrained on Encoder-Decoder Transformer (BART).
- Translates the raw input sentences into a set of triples that explicitly refer to those relations, and triples are expressed as a sequence of tokens.
  - Introduced new tokens, <triplet>, <subj> and <obj> for linerisation.

### 3.3. Relation Extraction & Entity Linking

<triplet> marks start of triplet with new head entity.

<subj> marks end of head entity and start of tail entity.

<obj> marks end of tail entity and start of relation.

“This Must Be the Place” is a song by new wave band Talking Heads, released in November 1983 as the second single from its fifth album “Speaking in Tongues”

(This Must Be the Place, performer, Talking Heads)  
(Talking Heads, genre, new wave)  
(This Must Be the Place, part of, Speaking in Tongues)  
(Speaking in Tongues, performer, Talking Heads)

<triplet> This Must Be the Place  
<subj> Talking Heads <obj> performer  
<subj> Speaking in Tongues <obj> part of  
<triplet> Talking Heads <subj> new  
wave <obj> genre <triplet> Speaking in  
Tongues <subj> Talking Heads <obj>  
performer

Figure 21: Triplet linearisation process for REBEL (Cabot et al., 2021)

### 3.3. Relation Extraction & Entity Linking

The screenshot shows the Hugging Face Model Card for the model `Babelscape/rebel-large`. The card includes the following information:

- Model Type:** Text2Text Generation, PyTorch, Safetensors, Transformers
- Dataset:** Babelscape/rebel-dataset
- Languages:** English
- License:** cc-by-nc-sa-4.0
- Model card:** Selected tab.
- Files and versions:** Available.
- Community:** 5 members.
- Performance Metrics (Ranked #2):** Relation Extraction on NYT (using additional training data).
- Performance Metrics (State of the Art):** Relation Extraction on CoNLL04 (using additional training data).
- Performance Metrics (State of the Art):** Joint Entity and Relation Extraction on DocRED (using additional training data).
- Performance Metrics (Ranked #5):** Relation Extraction on Adverse Drug Events (ADE) Corpus (using additional training data).
- Performance Metrics (Ranked #4):** Relation Extraction on Re-TACRED.

## 3.3. Relation Extraction & Entity Linking

Demo

Notebook file: Transformer\_example\_KG\_generation

### 3.3. Relation Extraction & Entity Linking

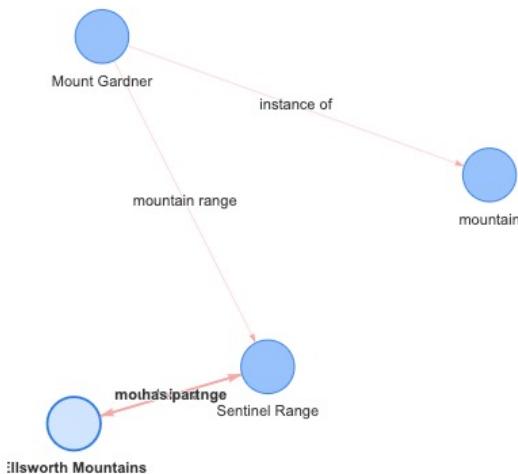


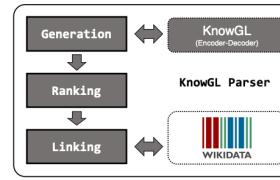
Figure 22: Generated knowledge graph

### 3.3. Relation Extraction & Entity Linking

KnowGL (Rossiello et al., 2023)

- Like REBEL, KnowGL also uses BART.
- Convert text into structured data as a set of ABox assertions compliant with the TBox of a given knowledge graph.

For the semantic web to function, computers must have access to structured collections of information and sets of inference rules.



```
[{"subject": {"mention": "semantic web", "entity_label": "Semantic Web", "type_label": "academic discipline", "entity_link": "Q54837", "type_link": "Q1862829"}, "relation": {"label": "uses", "link": "Property:P2283"}, "object": {"mention": "inference rules", "entity_label": "inference", "type_label": "process", "entity_link": "Q488386", "type_link": "Q619671"}, "score": -0.98}]
```

Figure 23: KnowGL Parser Framework (Rossiello et al., 2023).

# 3.

## Advanced/Deep Neural Networks (DNN) for Knowledge Graphs Construction

### *3.4. Knowledge Graph Completion*

## 3.4. Knowledge Graph Completion

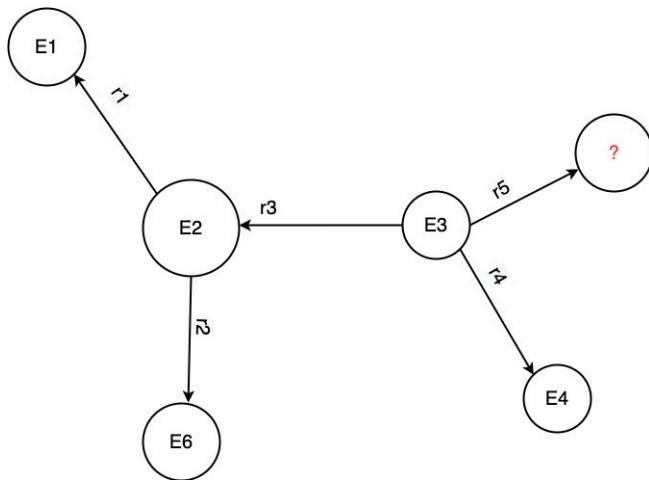


Figure 24: Knowledge graph with missing entity

## 3.4. Knowledge Graph Completion

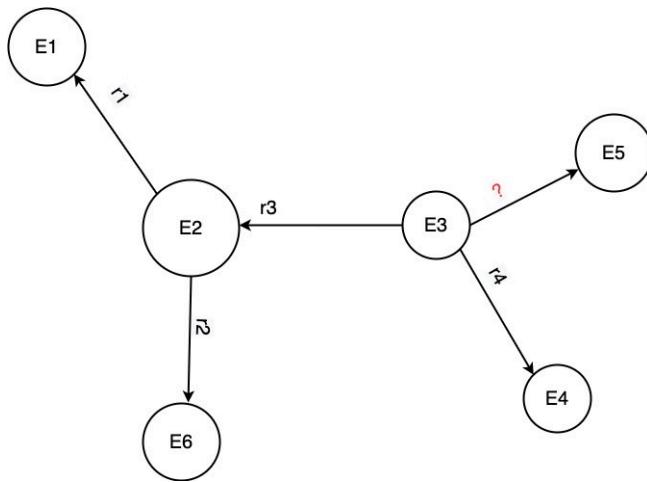


Figure 25: Knowledge graph with missing relation

# 4.

## Our work

### *4.1. Knowledge Graphs for Manufacturing Process Optimisation*

# 4.1. Knowledge Graphs for Manufacturing Process Optimisation



Image: Reuters.com

# 4.1. Knowledge Graphs for Manufacturing Process Optimisation

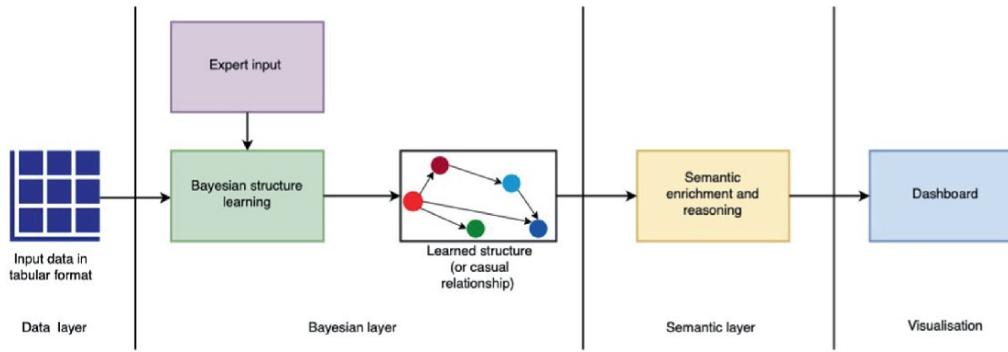


Figure 26: Methodology (Chhetri et al., 2023).

# 4.1. Knowledge Graphs for Manufacturing Process Optimisation

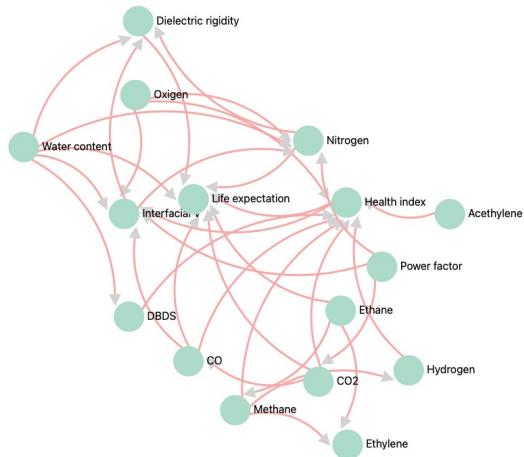


Figure 27: Directed acyclic graph (DAG)

# 4.1. Knowledge Graphs for Manufacturing Process Optimisation

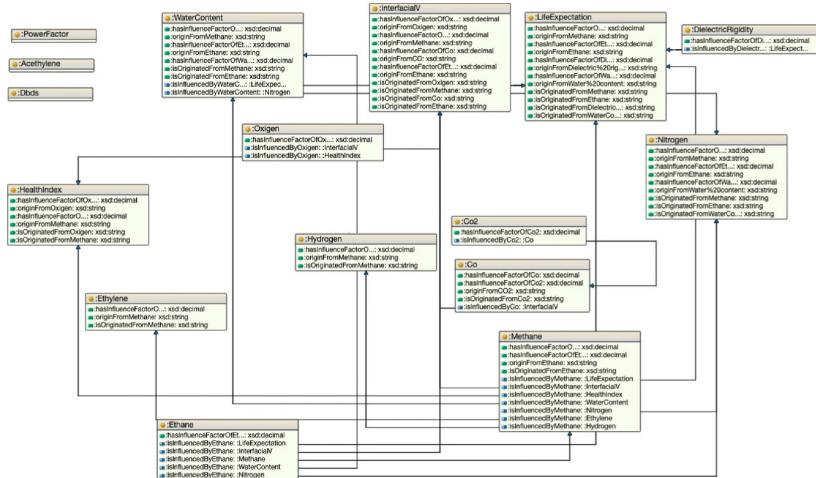


Figure 28: Automatically generated OWL ontology based on DAG learned from Bayesian structure learning.

# 4.1. Knowledge Graphs for Manufacturing Process Optimisation

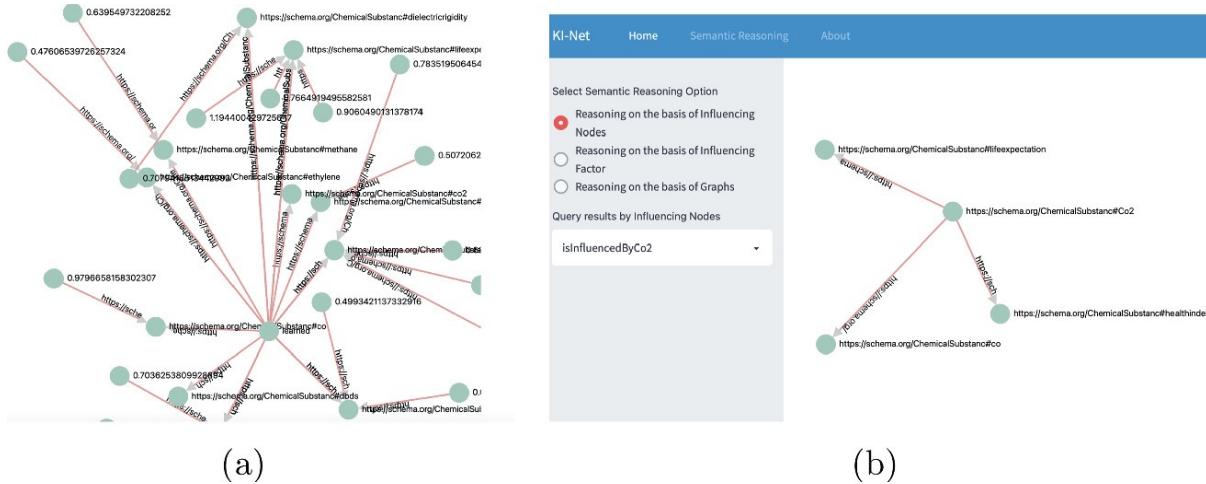
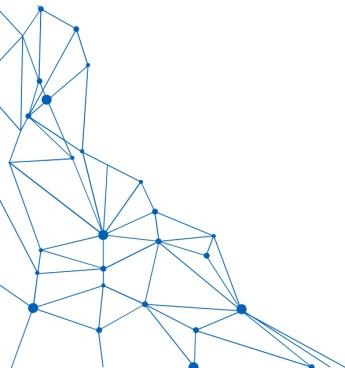
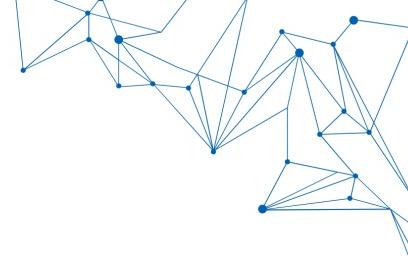


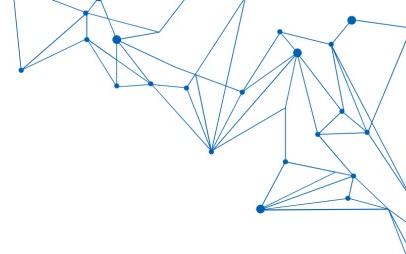
Figure 29: (a) Visualisation of semantic reasoning results according to influencing factor (or weights). (b) Visualisation of the outcomes of semantic reasoning according to their influencing nodes (or class) (Chhetri et al., 2023).

# 4.1. Knowledge Graphs for Manufacturing Process Optimisation





# 5. Conclusion



# 5. Conclusion

- Advanced/DNN have made it easy to construct knowledge graphs from text.
- There is no one technique that works for all domains.
- No matter how good your DNN models are, generated knowledge graphs are prone to errors. Consider evaluating the knowledge graphs using the quality assessment dimensions defined by Huaman & Fensel, 2021.



# 5. Conclusion

Open challenges (Groth et al., 2023):

- n to M relations issue.
- Knowledge extraction – what types?
- Trust, bias issue in language models.

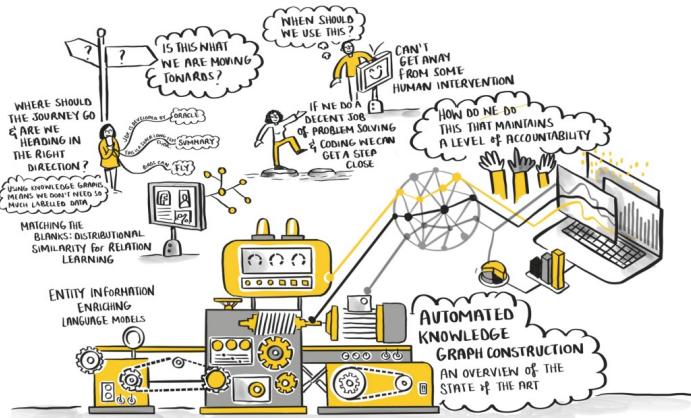
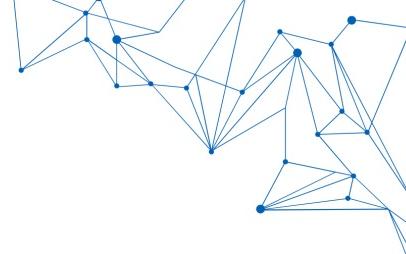
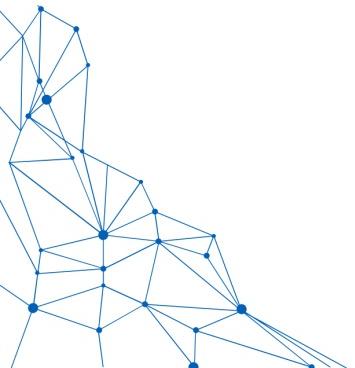


Figure 27: Automated Knowledge Graph Construction (Groth et al., 2023)



# Thank you!



# References

- Hogan, A., Blomqvist, E., Cochez, M., d'Amato, C., Melo, G.D., Gutierrez, C., Kirrane, S., Gayo, J.E.L., Navigli, R., Neumaier, S. and Ngomo, A.C.N., 2021. Knowledge graphs. ACM Computing Surveys (CSUR), 54(4), pp.1-37.
- Simsek, U., Kärle, E., Angele, K., Huaman, E., Opdenplatz, J., Sommer, D., Umbrich, J. and Fensel, D., 2022. A Knowledge Graph Perspective on Knowledge Engineering. SN Computer Science, 4(1), p.16.
- Angeli, G., Premkumar, M.J.J. and Manning, C.D., 2015, July. Leveraging linguistic structure for open domain information extraction. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers) (pp. 344-354).
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and Polosukhin, I., 2017. Attention is all you need. Advances in neural information processing systems, 30.

# References

- Otmazgin, S., Cattan, A. and Goldberg, Y., 2022. Lingmess: Linguistically informed multi expert scorers for coreference resolution. arXiv preprint arXiv:2205.12644.
- Cabot, P.L.H. and Navigli, R., 2021, November. REBEL: Relation extraction by end-to-end language generation. In Findings of the Association for Computational Linguistics: EMNLP 2021 (pp. 2370-2381).
- Rossiello, G., Chowdhury, F., Mihindukulasooriya, N., Cornec, O. and Gliozzo, A., 2023. KnowGL: Knowledge Generation and Linking from Text. AAAI.
- Huaman, E. and Fensel, D., 2021, December. Knowledge graph curation: A practical framework. In The 10th International Joint Conference on Knowledge Graphs (pp. 166-171).
- Chhetri, T.R., Aghaei, S., Fensel, A., Göhner, U., Gül-Ficici, S. and Martinez-Gil, J., 2023, February. Optimising Manufacturing Process with Bayesian Structure Learning and Knowledge Graphs. In Computer Aided Systems Theory—EUROCAST 2022: 18th International Conference, Las Palmas de Gran Canaria, Spain, February 20–25, 2022, Revised Selected Papers (pp. 594-602). Cham: Springer Nature Switzerland.
- Groth, P., Simperl, E., van Erp, M. and Vrandečić, D., 2023. ‘Knowledge Graphs and their Role in the Knowledge Engineering of the 21st Century (Dagstuhl Seminar 22372)’, Dagstuhl Reports. Edited by P. Groth et al. Dagstuhl, Germany: Schloss Dagstuhl -- Leibniz-Zentrum für Informatik, 12(9), pp. 60–120. doi: 10.4230/DagRep.12.9.60.