Information Extraction Topics WiSe 2023/24

Kathy Hämmerl
haemmerl@cis.lmu.de

CIS, LMU

08 November 2023

Table of Contents

Table of Contents

Fact Extraction and Verification

- Misinformation and disinformation are widespread problems

Fact Extraction and Verification

- Misinformation and disinformation are widespread problems
- Fact Extraction and Verification is the task of *finding claims in text* and *checking for evidence for or against* the claim(s)

Fact Extraction and Verification

- Misinformation and disinformation are widespread problems
- Fact Extraction and Verification is the task of *finding claims in text* and *checking for evidence for or against* the claim(s)
- "FEVER" dataset and workshop focusing especially on the *verification* part by finding relevant evidence

Fact Extraction and Verification

- James Thorne, Andreas Vlachos, Christos Christodoulopoulos, and Arpit Mittal. 2018. FEVER: A large-scale dataset for Fact Extraction and VERification. In *NAACL-HLT 2018*, pages 809–819, New Orleans, Louisiana. ACL.
- Andreas Hanselowski, Hao Zhang, Zile Li, Daniil Sorokin, Benjamin Schiller, Claudia Schulz, and Iryna Gurevych. 2018. UKP-Athene: Multi-Sentence Textual Entailment for Claim Verification. In *FEVER 2018*, pages 103–108, Brussels, Belgium. ACL.
- Daniel Guzman Olivares, Lara Quijano, and Federico Liberatore. 2023. Enhancing Information Retrieval in Fact Extraction and Verification. In *FEVER 2023*, pages 38–48, Dubrovnik, Croatia. ACL.

Table of Contents

Canonicalisation of Open Knowledge Bases

- Open knowledge bases can be extracted from web text, e.g., as triples of (noun phrase, verb phrase, noun phrase)

Canonicalisation of Open Knowledge Bases

- Open knowledge bases can be extracted from web text, e.g., as triples of (noun phrase, verb phrase, noun phrase)
- Multiple NPs can mean the same entity; multiple VPs can mean the same relation

Canonicalisation of Open Knowledge Bases

- Open knowledge bases can be extracted from web text, e.g., as triples of (noun phrase, verb phrase, noun phrase)
- Multiple NPs can mean the same entity; multiple VPs can mean the same relation
- Without addressing this, the database becomes redundant, very hard to search

Canonicalisation of Open Knowledge Bases

- Open knowledge bases can be extracted from web text, e.g., as triples of (noun phrase, verb phrase, noun phrase)
- Multiple NPs can mean the same entity; multiple VPs can mean the same relation
- Without addressing this, the database becomes redundant, very hard to search
- Canonicalisation: Recognise synonyms, cluster them together

Canonicalisation of Open Knowledge Bases

- Open knowledge bases can be extracted from web text, e.g., as triples of (noun phrase, verb phrase, noun phrase)
- Multiple NPs can mean the same entity; multiple VPs can mean the same relation
- Without addressing this, the database becomes redundant, very hard to search
- Canonicalisation: Recognise synonyms, cluster them together

Example

"Obama was born in Honolulu" == "Barack Obama's birthplace is Honolulu"

Canonicalisation of Open Knowledge Bases

- Luis Galárraga, Geremy Heitz, Kevin Murphy, and Fabian M. Suchanek. 2014. Canonicalizing Open Knowledge Bases. In *CIKM '14*. ACM.
- Shikhar Vashishth, Prince Jain, and Partha Talukdar. 2018. CESI: Canonicalizing Open Knowledge Bases using Embeddings and Side Information. In *WWW '18*.
- Sarthak Dash, Gaetano Rossiello, Nandana Mihindukulasooriya, Sugato Bagchi, and Alfio Gliozzo. 2021. Open Knowledge Graphs Canonicalization using Variational Autoencoders. In *EMNLP 2021*, pages 10379–10394, Online and Punta Cana, Dominican Republic. ACL.

Table of Contents

Knowledge Graph Reasoning

- Knowledge bases are most useful if we can apply them to, e.g., question answering

Knowledge Graph Reasoning

- Knowledge bases are most useful if we can apply them to, e.g., question answering
- That is, if we can *extract more information* from the relations that we already have

Knowledge Graph Reasoning

- Knowledge bases are most useful if we can apply them to, e.g., question answering
- That is, if we can *extract more information* from the relations that we already have
- This is *related* to the canonicalisation problem

Knowledge Graph Reasoning

- Knowledge bases are most useful if we can apply them to, e.g., question answering
- That is, if we can *extract more information* from the relations that we already have
- This is *related* to the canonicalisation problem
- This topic is a bit more complex—try not to get too lost in technical detail but focus more on goals, successes and limitations

Knowledge Graph Reasoning

- Bordes, Antoine, Nicolas Usunier, Alberto García-Durán, Jason Weston and Oksana Yakhnenko. 2013. Translating Embeddings for Modeling Multi-relational Data. In *NeurIPS 2013*.
- Shi, Baoxu and Tim Weninger. Open-World Knowledge Graph Completion. 2017. In *AAAI 2017*.
- Rajarshi Das, Ameya Godbole, Ankita Naik, Elliot Tower, Manzil Zaheer, Hannaneh Hajishirzi, Robin Jia, Andrew Mccallum. 2022. Knowledge Base Question Answering by Case-based Reasoning over Subgraphs. In *ICML 2022*.