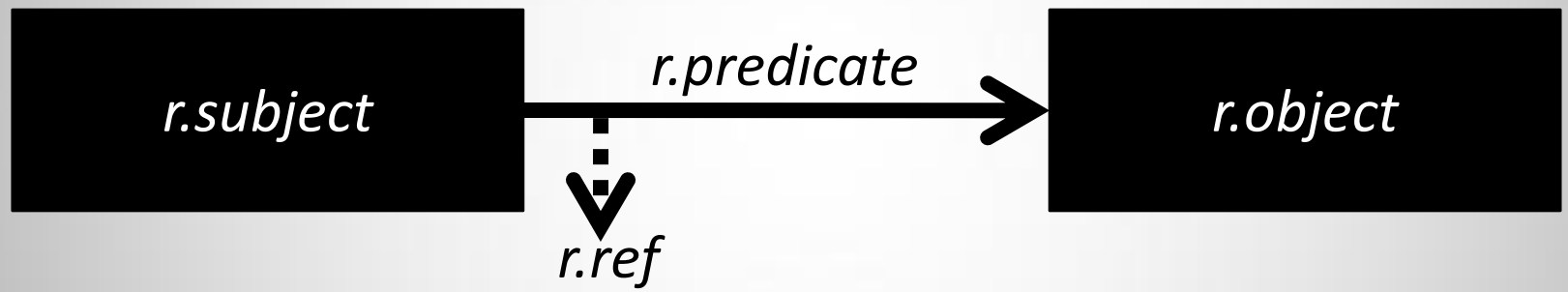
جامعة صفاقس
University of Sfax

REFB: ADDING REFERENCE SUPPORT TO BIOMEDICAL RELATIONS IN WIKIDATA

Houcemeddine Turki
University of Sfax, Tunisia

INTRODUCTION

- Biomedical relations: Biomedical statements where the subject and object are Wikidata items



- Over 25000 Biomedical relations missing references in Wikidata
- Most scientists work on screening research publications to automatically retrieve biomedical relations. Less scientists think about finding research publications that support a given biomedical relation.

جامعة صفاقس
University of Sfax

PUBMED CENTRAL (PMC)

- More than 6.4 million biomedical publications: *The sum of all human biomedical knowledge*
- Bibliographic metadata as well as full texts
- Involves a search engine that sorts publications according to their relevance
- Can be automatically processed using the Entrez API

جامعة صفاقس
University of Sfax

ALGORITHM

1	Extract all unsupported biomedical relations (*bior*) in Wikidata
2	**For** *r* in *bior* **do**
3	**begin**
4	Find the three most relevant research publications (*reflist*) supporting *r* in PMC
5	**For** *pub* in *reflist* **do**
6	**begin**
7	**if** (*r.subject* in *pub.abstract*) and (*r.object* in *pub.abstract*) **then**
8	**begin**
9	**if** *pub* is not fully or partially retracted **then** *r.ref* = *pub*
10	**end**
11	**end**
12	**end**

جامعة صفاقس
University of Sfax

RESULTS

- 98.6% Accuracy
- Screens ten biomedical relations every minute
- Adds references to two biomedical relations every minute

جامعة صفاقس
University of Sfax

CONCLUSION

- Promising results
- To be customized allowing the use of multiple queries to identify unsupported relations
- Assigning *Preferred Rank* to the most relevant reference
- Can be enhanced by using interesting Python libraries like Wikidata Integrator
- Can be generalized to support *Reference needed* Wikipedia sentences

جامعة صفاقس
University of Sfax

TO CITE

- Turki, H., et al. (2020). Using logical constraints to validate information in collaborative knowledge graphs: a study of COVID-19 on Wikidata. *Zenodo*. https://zenodo.org/record/4008359
- *MisinfoCon Post*: https://misinfocon.com/refdata-adding-trustworthiness-to-wikidata-d3cc68c21a6f
- *GitHub Repository*: https://github.com/Data-Engineering-and-Semantics/refb
- *Bot contributions*: https://www.wikidata.org/wiki/Special:Contributions/RefB_(WikiCred)

جامعة صفاقس
University of Sfax