

Open knowledge graph on clinical trials

Sandeep Kunkunuru^{*1,2}

¹IIPH Hyderabad

²VaidhyaMegha Private Limited

^{*} Correspondence: Sandeep Kunkunuru <hi18sk@iiphh.org>

Abstract

Using any clinical trial id from across the globe find the associated diseases, interventions, research articles and genes. Also discover relationships between various medical topics through co-occurrences in articles. Query the graph using SparQL from cli or GraphQL using any API client tool ex: Postman or curl

Introduction

Sources

- WHO's ICTRP
 - Registries covered in ICTRP include :
- AACT Clinicaltrials.gov
- NLM MeSH
- NLM MRCOC
- NLM PubMed
- NLM PheGenI

Feature list

- Using GraphQL API knowledge graph can be queried using any API client tool ex: curl or Postman.
- Graph includes trials from across the globe. Data is sourced from WHO's ICTRP and clinicaltrials.gov
- Links from trial to MeSH vocabulary are added for conditions and interventions employed in the trial.
- Links from trial to PubMed articles are added. PubMed's experts curate this metadata information for each article.
- Added MRCOC to the graph for the selected articles linked to clinical trials.

- Added PheGenI links i.e. links from phenotype to genotype as links between MeSH DUI and GeneID.
- Added SparQL query execution feature. Adding CLI mode. Adding a count SparQL query for demo.
- 5 co-existing bi-partite graphs together comprise this knowledge graph. Bi-partite graphs are between
 - trial-> condition
 - trial-> intervention
 - trial-> articles
 - article-> MeSH DUIs
 - gene id-> MeSH DUIs

Trial to condition

Trial to intervention

Trial to articles

Article to MeSH DUIs

Gene id to MeSH DUIs

Demonstration

Querying knowledge graph using SparQL

```
$ java -jar -Xms4096M -Xmx8144M \
    target/vaidhyamegha-knowledge-graphs-1.0-SNAPSHOT-jar-with-dependencies.jar \
    -m cli -q src/main/sparql/1_count_of_records.rq
...
Results:
-----
4766048~http://www.w3.org/2001/XMLSchema#integer
```

Querying knowledge graph using GraphQL (via Hyper-GraphQL)

Start server

```
java -Dorg.slf4j.simpleLogger.defaultLogLevel=debug -jar lib/hypergraphql-3.0.1-exe.jar \
    --config src/main/resources/hql-config.json
```

Start client

In a separate terminal execute GraphQL query using curl (alternatively use Postman)

```
$ curl --location --request POST 'http://localhost:8080/graphql' \
  --header 'Accept: application/ntriples' \
  --header 'Accept-Language: en-GB,en-US;q=0.9,en;q=0.8,kn;q=0.7' \
  --header 'Content-Type: application/json' \
  --data-raw \
  '{"query": "{\n trial_GET(limit: 30, offset: 1) {\n label\n }\n\n}", "variables": {}}'

<https://www.who.int/clinical-trials-registry-platform/EUCTR2007-006072-11-SE>
  <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
  <https://vaidhyamegha.com/open_kg/ct> .
<https://www.who.int/clinical-trials-registry-platform/EUCTR2007-006072-11-SE>
  <http://www.w3.org/2000/01/rdf-schema#label>
  "EUCTR2007-006072-11-SE"^^<http://www.w3.org/2001/XMLSchema#string> .
<https://clinicaltrials.gov/ct2/show/NCT02954757>
  <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
  <https://vaidhyamegha.com/open_kg/ct> .
<https://clinicaltrials.gov/ct2/show/NCT02954757>
  <http://www.w3.org/2000/01/rdf-schema#label>
  "NCT02954757"^^<http://www.w3.org/2001/XMLSchema#string> .
<https://www.who.int/clinical-trials-registry-platform/EUCTR2014-005525-13-FI>
  <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
  <https://vaidhyamegha.com/open_kg/ct> .
<https://www.who.int/clinical-trials-registry-platform/EUCTR2014-005525-13-FI>
  <http://www.w3.org/2000/01/rdf-schema#label>
  "EUCTR2014-005525-13-FI"^^<http://www.w3.org/2001/XMLSchema#string> .
<https://clinicaltrials.gov/ct2/show/NCT02721914>
  <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
  <https://vaidhyamegha.com/open_kg/ct> .
<https://clinicaltrials.gov/ct2/show/NCT02721914>
  <http://www.w3.org/2000/01/rdf-schema#label>
  "NCT02721914"^^<http://www.w3.org/2001/XMLSchema#string> .
...
<http://hypergraphql.org/query> <http://hypergraphql.org/query/trial_GET>
  <https://www.who.int/clinical-trials-registry-platform/EUCTR2016-002461-66-IT> .
<http://hypergraphql.org/query> <http://hypergraphql.org/query/trial_GET>
  <https://www.who.int/clinical-trials-registry-platform/CTRI/2020/08/027368> .
<http://hypergraphql.org/query> <http://hypergraphql.org/query/trial_GET>
  <https://www.who.int/clinical-trials-registry-platform/EUCTR2013-001294-24-DE> .
```

Further reading

For more information please read

Carroll et al. (2004). Grobe (2009).

Acknowledgements

References

- Carroll, J. J., Dickinson, I., Dollin, C., Reynolds, D., Seaborne, A., & Wilkinson, K. (2004). Jena: Implementing the semantic web recommendations. *Proceedings of the 13th International World Wide Web Conference on Alternate Track Papers & Posters*, 74–83. <https://doi.org/10.1145/1013367.1013381>
- Grobe, M. (2009). RDF, jena, sparql and the 'semantic web'. *Proceedings of the 37th Annual Acm Siguccs Fall Conference: Communication and Collaboration*, 131–138. <https://doi.org/10.1145/1629501.1629525>