6.11

DbPedia ausflug:

https://dbpedia.org/sparql

11.10.2023

<https://drive.google.com/file/d/1jzH1KhSn6UxxhayoZV5rCXQF1nsCI5a6/view> von marta sabtou sehr interessant

<https://www.wisecube.ai/blog/knowledge-graphs-rdf-or-property-graphs-which-one-should-you-pick/> -> eigentlich property graph für evolving kg

<https://ceur-ws.org/Vol-2456/paper36.pdf> -> paper impact from evolving knowledge graphs

<https://alenaschmickl.medium.com/5-steps-to-find-inconsistencies-in-evolving-knowledge-graphs-6f3f88c0ab7b> get changes in kg

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8126660/> correctness after updated to kg

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10495951/> kg augmentation

<https://www.cs.sjtu.edu.cn/~fu-ly/paper/EvolvingKG.pdf> evolving kg

<https://ieeexplore.ieee.org/document/8737547> evolvekg

<https://arxiv.org/pdf/2112.05003.pdf> wikidata diff

<https://www.sciencedirect.com/science/article/pii/S1570826823000252> iwas

[file:///C:/Users/evapu/Downloadshttps://fileproinfo.com/tools/comparison/ttl/data\_versioning\_for\_graph\_databases.pdf](file:///C:/Users/evapu/Downloads/data_versioning_for_graph_databases.pdf) version control for db

<https://repository.tudelft.nl/islandora/object/uuid%3A63aeab75-64a5-4b59-9cb0-241b603bd00d> versioncontrol kg <file:///C:/Users/evapu/Downloads/MSc_thesis_Lisa_Meijer.pdf>

<https://www.w3.org/2001/sw/wiki/How_to_diff_RDF>

<https://rdflib.readthedocs.io/en/stable/apidocs/rdflib.html#module-rdflib.compare>

jena verwenden

Data Wrangling als Überbegriff für Daten sammeln

<https://dl.acm.org/doi/pdf/10.1145/3487553.3524253> paper von hose, warum qse gebaut wurde

Knowledge graph vorlesung: use cases: bank of Italy, versteckte verbindungen aufdecken. Für logische reasoning, kg embeddings, graph neuronal networks; record linkeage mit human intervention (bei kg creation)

KG-07-06: Beispiele, wann ein graph geändert werden muss. Wann das Schema geändert werden muss, was beachtet werden muss…. Verändert auch das abgeleitete wissen, wenn sich die daten ändern

KG-07-07: Schema evolution: wie ein schema geändert werden kann. Kann sein, dass eine property gelöscht oder umgeändert wird (was macht man dann: report drucken, ganzen graph ändern). Oder die klassenhierarchie verändert wird. Superklassen müssen dann neu zugeordnet werden, subklassen…

KG-07-08: View maintenance (materialized views) und truth maintenance (update ergebnis, wenn sich darunter liegende daten ändern) eher unwichtig

KG-09-03: Graphdatenbanken Temporal: Als Attribute, Properties, Nodes oder Edges. Versioned Rooted Tree, hat dann edges wie :previous drinnen (<https://docs.google.com/presentation/d/1lCgosyT8VOszY1GAqaGqu7pB-Cz7J4VA_ExIJuw_XwM/edit#slide=id.g4f4671db3_0375> ). Oder in den edges @versionsnummer. Graph Versioning als Keyword  
Oder time tree model (<https://graphaware.com/blog/framework/from-graphaware-framework-to-graphaware-hume.html>), sieht aber nicht so interessant aus

KG-Vorlesung: Use Cases für Knowledge graphs: AI, Neurosymbolic AI, Explainable AI

Ideen aus Thema 1:

The following extensions can be made to our existing QSE5 algorithm:

- Extend QSE's scalability and adaptability by transforming it into an anytime algorithm,

enabling on-the-fly shape extraction in such a way that users can interrupt the

extraction at any time and QSE will still output useful shapes.

- Support for mining additional types of SHACL constraints (e.g., recursive constraints)

The following extensions can be made to our existing SHACTOR tool6

- Grouping related shapes to help users understand non-obvious similarities.

- Support of an alternative shape language: SheX7

- Support of additional knowledge graph formats, e.g., Turtle8 (currently support of NTriples(

NT)9-formatted RDF files and GraphDB Triplestore via SPARQL)

- Visualization of the extracted shapes along with the corresponding data constraints

- Interactive identification of low-quality (erroneous/spurious) data, enabling users to

comprehend the data quality within the knowledge graph.

- Using the tool to analyze several real-world knowledge graphs, e.g., WikiData

3.10

Vergleich

Themen von thema 1

Kann mich mit dorktor zusammenreden, um shactor zu erweitern

Webtool

Exposee schreiben mit genauer definition

Grundvariante funktioniert immer, samples für große graphen

Geht schon mit shapes vergleichen

Diff von graphen -> wie sich das auf shapes auswirkt, zweite variante

Selbst ausprobieren, knowledge graph basteln

Laufzeit messen, welche variatne schneller ist

Changesets graphen diffen

Trotzdem shactor parameter verwenden?

https://relweb.cs.aau.dk/rdfev/

<https://content.iospress.com/articles/semantic-web/sw210434>

Ja, parametern verwenden. Es kann sein,d ass ein parameter den wert unterschreitet. Kann sein,d ass ergebnis hinuasfällt. Recherchieren, warum shapes hinausfällt.

Neue website, sachen kopieren

Eine seite exposse, titel name, referenzen

Man kann fragen auch ändern, wenn es sich währenddessen anders besser ergibt

Todo exposee: themen von 1, webtool, title, refernezen, shapes vergleich, graph vergleich, laufzeiten messen, parameter, warum shapes herausfallen

Ist der vergleich nicht nur ein string vergleich? Shacl forms sind ja nur text

QSE Algorithm noch nicht angsehen -> kann noch nicht einschätzen, wie viel ich da einbauen kann. If time allows… wie zu verstehen?

Trotzdem shactor parameter verwenden?

Also ich baue nicht auf Shactor tool auf, aber kann vll code teile davon wiederverwenden? (z.B. Daten einlesen, Visualisierung etc.) ja

Werde ich website bauen?

Gibt es nicht schon Lösungen, die mit versionen von graphen arbeiten?

Nächste schritte? Proposal? Literatur?

Teams chatten?

Fixe termine? Oder fixe zeitspanne, an der sie zeit hat?

22.9

RunConfiguration QSE -> es muss der Pfad von den config properties angegeben werden

GraphDb ausprobiert. Mit Lubm mini ausprobiert C:\Users\evapu\Documents\Wien\TU\WINF\MA\shactor\qse\src\main\resources

In intellj gestartet, pfad zu lubm mini geändert

Main file einfach gerunt, website öffnet sich von alleine. Andere optionen zum graph einlesen haben nicht funktioniert.

Qse algorithmus noch nicht probiert.

nt ist format zum speichern

Graphdb kostenloses programm + website zum ansehen von graphdaten

Sparql abfragesprache, vll noch mehr einlesen und übungsbeispiele machen mit graphdb

Description Framework (RDF). Cypher is the corresponding language for the data represented in property graphs.

Shacl hat shapes (property oder node shapes), die qualitätsregeln vorgeben. Property shapes e.g. name darf nur ein string sein, node: muss name verbindung haben)

Warum ist alles so alt?

8.9. todo

Andere professoren research

Zurückschreiben

Sonst thema provenance 1

Datenbanksysteme Vertiefung

Business Intelligence

Genaues Thema?

Zeitplan: 2 Semester

Zusammanarbeit mit AWS? Noch nicht

Wo data engineering?

Amazon neptune

Azure Cosmos DB for Apache Gremlin

Apache tinkerpop (grundlage)

Tigergraph

Graphvisualisierung manchmal nur draufgesetzt auf nosql

Cypher

Cayley

Betrugsfälle, neo4j, viele verbindungen, kein festes schema, für bessere kis

Papers: Aufgabenunterteilung Suchanfragen

datenextraktion ozeanographische Daten

sparql

cypher

What is the difference between SPARQL and Cypher Query Language? SPARQL is the query language for accessing data in the Resource Description Framework (RDF). Cypher is the corresponding language for the data represented in property graphs.

theorie praxis?

Id für metadaten

Adfs star star nesting

Optimierung aufteilung

Spm semi rings

1. Thema anfrage quelle finden, möglichst effizient how prominänzprovenance

Nur eine instanz, mehrere graphen verteilt?

Ecosystem für graph daten banken -> nicht so einfach für thmea

Shacl validation

Shape ist validation

Dpedia

Doktorant tool existiert

Clustern nach ähnlichkeit validations

1. Graph dynamisch, neue version, vergleichen

Provenance manipulation fehler

Rdf datenmodell -> für knoledge graphs

Sharctor

Rdf und property grpah (neo4j)

Nur solange es kein anderer macht