Ontologies, Knowledge Graphs, and all that

Pascal Hitzler
Data Semantics Laboratory (DaSe Lab)
Kansas State University

http://www.daselab.org
Knowledge Graphs

Knowledge Graphs (and their schemas) are made to enable easier

- data sharing
- data discovery
- data integration
- data reuse
Laura Kelly is an American politician serving as the 48th governor of Kansas since 2019. A member of the Democratic Party, she represented the 18th district in the Kansas Senate from 2005 to 2019. Kelly ran for governor in the 2018 election and defeated the Republican nominee, Kansas Secretary of State Kris Kobach. Wikipedia

Born: January 24, 1950 (age 69 years), New York, NY
Spouse: Ted Daughety
Party: Democratic Party
Office: Governor of Kansas since 2019
Education: Indiana University, Bradley University, Indiana University Bloomington
Children: Kathleen Daughety, Molly Daughety
Knowledge Graphs

Laura Kelly

hasEducation

Indiana University

hasPresident

Michael McRobbie

hasStudents

110,436

hasEducation

University of Queensland

01/24/1950
A good schema is critical for ease of reuse
This is not a good Knowledge Graph!

Laura Kelly hasEducation Indiana University

hasBirthDate 01/24/1950

hasStudents 110,436

Michael McRobbie hasPresident University of Queensland

hasEducation

hasStudents

hasPresident
Semantic Web Pre-History: before 2001

- Cyc, gene ontology, and others
- semantic networks
- knowledge representation in Artificial Intelligence
- the world wide web
- data integration
- database schemas
Ontologies: 2001-2007

• “ontologies will solve all your data management problems”

• Fueled by large-scale EU funding in FP6.
• Many ontologies came out of this time
• W3C standards RDF/S, OWL, SPARQL

• The promise was that ontologies will be heavily re-used, but this didn’t happen.

Arguably:
• If often seemed easier to make a new ontology from scratch than to try understand an existing one and adapt it to your requirements.
Problems with Ontologies

- Large, complex, with little or no internal structure
- Insufficient documentation.
- Non-obvious design choices, unexplained.
- Laden with ambiguity.

Very difficult to understand what’s going on inside.
Problems with Ontologies

And this is just looking at the class hierarchy!
Problems with Ontologies

- Designed for single use case.
- Granularity of representation highly varying.
- Large and monolithic, hard to assess what any change will entail.

→ Very difficult to adapt to a new setting.
W3C Standards

RDF 1.1 Concepts and Abstract Syntax

W3C Recommendation 25 February 2014

This version:
http://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/

Latest published version:
http://www.w3.org/TR/rdf11-concepts/

Previous version:
http://www.w3.org/TR/2014/PR-rdf11-concepts-20140109/

Previous Recommendation:
http://www.w3.org/TR/rdf-concepts

Editors:
Richard Cyganiak, DERI, NUI Galway
David Wood, 3 Round Stones
Markus Lanthaler, Graz University of Technology

Both established 2004 as versions 1.0.

OWL 2 Web Ontology Language Primer (Second Edition)

W3C Recommendation 11 December 2012

This version:
http://www.w3.org/TR/2012/REC-owl2-primer-20121211/

Latest version (series 2):
http://www.w3.org/TR/owl2-primer/

Latest Recommendation:
http://www.w3.org/TR/owl-primer

Previous version:
http://www.w3.org/TR/2012/PER-owl2-primer-20121018/

Editors:
Pascal Hitzler, Wright State University
Markus Krötzsch, University of Oxford
Bijan Parsia, University of Manchester
Peter F. Patel-Schneider, Nuance Communications
Sebastian Rudolph, FZI Research Center for Information
Linked Data: 2007-2013

- “linked data will solve all your data management problems”

- Convert your data to RDF, link it to Dbpedia, and put it on the Web.

- A 2015 count: “more than 37 billion triples from over 650,000 data documents”

Figure 1: Number of RDF graphs in the Linked Open Data Cloud over time
Problems with Linked Data

Geoindexed Linked Data – courtesy of Krzysztof Janowicz, 2012
http://stko.geog.ucsb.edu/location_linked_data
Problems with Linked Data

“Nancy Pelosi voted for the Affordable Care Act.”
Knowledge Graphs: since 2013

• Term originating from the Google Knowledge Graph, launched in 2012.

• Essentially, still (RDF) graphs. But shift in emphasis:
  – industrial adoption of “their own” knowledge graph
  – openness is no longer a prominent aspect
  – more central control
  – de-emphasis of external links
  – more re-introduced awareness of schema/ontologies
Sustainable Knowledge Graph Design

Some key aspects of our own approach:

1. Modular approach
2. Careful schema design
3. Reuse components, not ontologies
Modular approach

- Divide and conquer complexity
- Modules resonate with human expert conceptualizations
- De-emphasize class hierarchies in favor of modules
- Modifications remain local

For sustainable data reuse.
Design interconnected modules
From Patterns to Modules
Careful Schema Design

- Involve group of domain experts
- Relate, not define
- Design general purpose patterns

For sustainable data reuse
Reuse components, not ontologies

• Ontology Design Patterns: Reuseable solutions to recurring modeling problems
• Use as templates: Adjust to scope at hand
• Development and use of pattern libraries

For rapid deployment of high quality schema.
Enslaved
Peoples of the Historic Slave Trade

Building a Linked Open Data Platform for the study and exploration of the historical slave trade.

Learn More
Grain Traceability

Data model is central for lowering cost in all parts of the pipeline.

- Tracing along splits and merge.
- Elevators as black boxes.
- Containers may carry contaminants
- ...

Figure acknowledgement: NIST / Evan Wallace
NSF KnowWhereGraph

- spatial and temporal aspects of knowledge graphs
- applications e.g. in disaster relief, soil health
- tool development with goal of industrial dissemination

- Part of the NSF Convergence Accelerator Program in Track A “Open Knowledge Networks”
  - 2019-2020 Phase 1 ($1M)
  - 2020-2022 Phase 2 ($5M)
COModIDE

- Protege plug-in

- supports our Modular Ontology Modeling process

(Cogan Shimizu, lead developer)
Thanks!


Guus Schreiber, Yves Raimond, RDF 1.1 Primer. W3C working Group Note 24 June 2014


https://lod-cloud.net
References


References


Cogan Shimizu, Karl Hammar, Pascal Hitzler, Modular Ontology Modeling. Under review. http://www.semantic-web-journal.net/content/modular-ontology-modeling
Thanks!