PADERBORN UNIVERSITY

Concept Learning in Description Logics

Motivation

Axel Ngonga

September 03rd, 2023

Section 1

Motivation

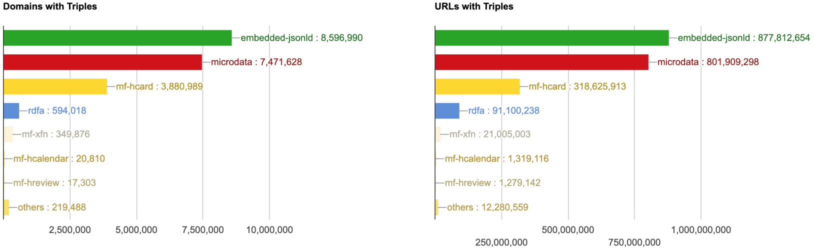
Motivation

Data Web

- Most popular application of knowledge graphs (KGs) is Web search
- Approx. 30% of search queries at Google answered by the Google Knowledge Graph
- Further applications in finance, health, e-commerce, and Industry 4.0
- Predominant portion of open KGs are in RDF

Motivation

Data Web

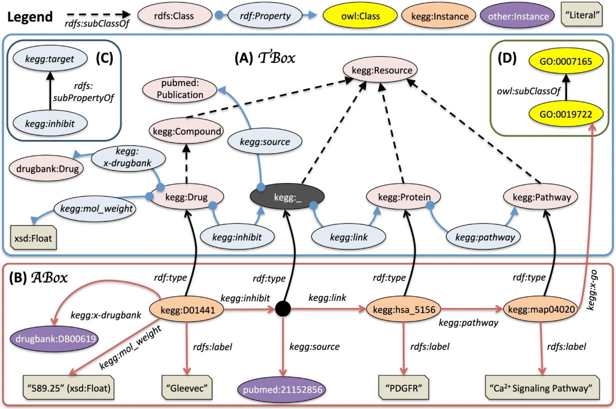


- RDF knowledge bases are now first-class citizens of the Web
- Approx. 50% of websites contain RDF[1]
- 2+ billion URLs contain RDF statements
- Ca. 100 billion statements in Linked Open Data

[1] See http://webdatacommons.org/structureddata/#results-2022-1

Motivation

Description Logics



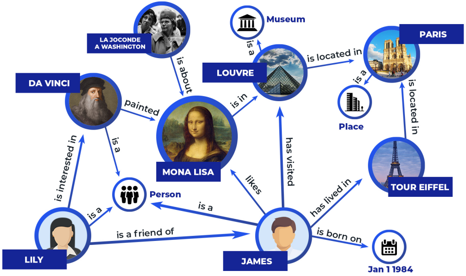
- Terminology of RDF datasets in description logics
- Popular DLs include $\mathcal{ELH}$ (e.g., for biomedical domain), $\mathcal{ALC}$ (e.g., for ML-driven applications), and $\mathcal{SROIQ}$ (e.g., on the Web)

Section 2

Basic Setting

Basic Setting

Learning Problem



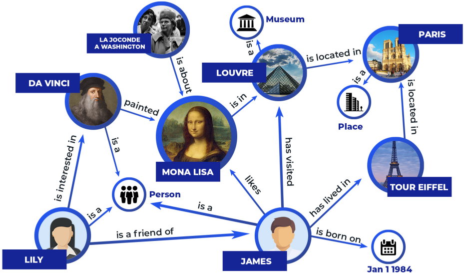
[2]

- Given
  - Knowledge base G (often called background knowledge)

[2]Source: https://bit.ly/3sxCj6e

Basic Setting

Learning Problem



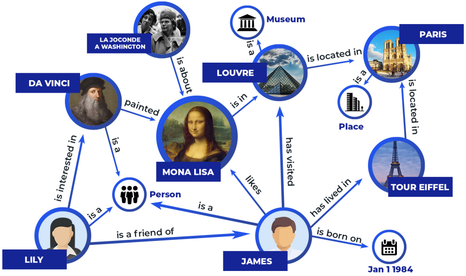
2

- Given
  - Knowledge base G (often called background knowledge)
  - Set of positive examples, e.g., $E^+ = \{Louvre, TourEiffel\}$

[2] Source: https://bit.ly/3sxCj6e

Basic Setting

Learning Problem



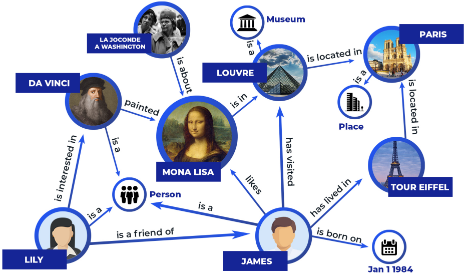
[2]

- Given
  - Knowledge base G (often called background knowledge)
  - Set of positive examples, e.g., $E^+ = \{Louvre, TourEiffel\}$
  - Set of negative examples, e.g., $E^- = \{Lily, James\}$

[2] Source: https://bit.ly/3sxCj6e

Basic Setting

Learning Problem



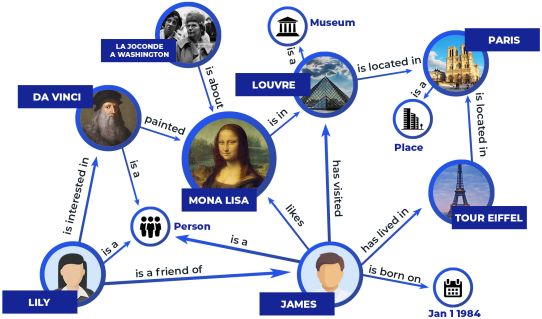
[2]

- Given
  - Knowledge base G (often called background knowledge)
  - Set of positive examples, e.g., $E^+ = \{Louvre, TourEiffel\}$
  - Set of negative examples, e.g., $E^- = \{Lily, James\}$
- Goal: Find concept H that "describes" E^+ and "does not describe" E^-, e.g., $H = \exists\ isLocatedIn.Place$

[2]Source: https://bit.ly/3sxCj6e

Basic Setting

Symbolic Approaches



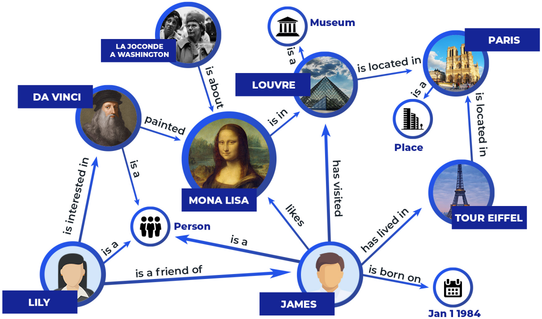
[3]

- Often based on refinement operators

[3] Source: https://bit.ly/3sxCj6e

Basic Setting

Symbolic Approaches



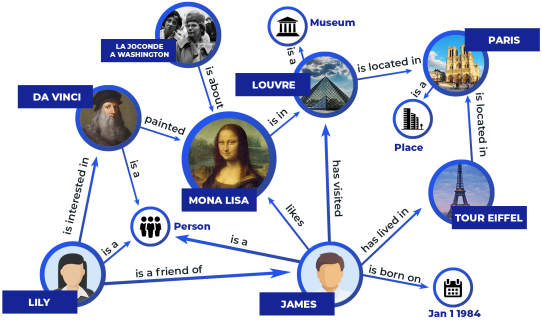
3

- ► Often based on refinement operators
- + Explainable, exploits background knowledge
- − Slow :-(

[3]Source: https://bit.ly/3sxCj6e

Basic Setting

Neural Approaches



[4]

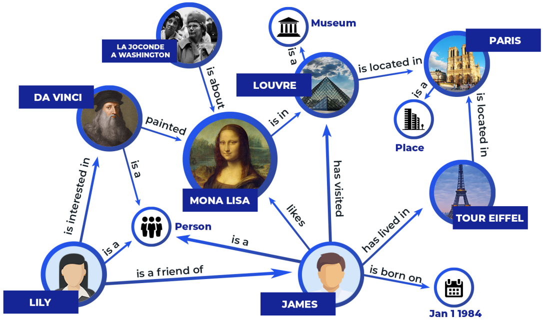
- **Deep Learning**: $e(v_i) := \varphi\left(\bigoplus_{(v_i,p_k,v_j)\in G} e(p_k, v_j), e(v_i)\right)$

---

[4] Source: `https://bit.ly/3sxCj6e`

Basic Setting

Neural Approaches



[4]

- Deep Learning: $e(v_i) := \varphi\left(\bigoplus_{(v_i,p_k,v_j)\in G} e(p_k, v_j), e(v_i)\right)$

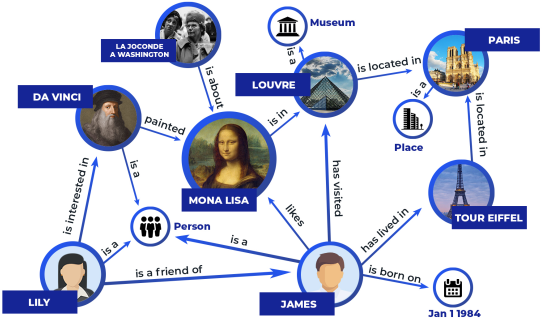
+ Time-efficient

− Unintelligible, does not exploits background knowledge

[4]Source: https://bit.ly/3sxCj6e

Basic Setting

Neuro-Symbolic Approaches



[5]

- Exploit time efficiency of neural approaches
- Keep explainability of symbolic approaches

[5] Source: https://bit.ly/3sxCj6e