

Open Information Extraction

Dominik Both, Tonio Weidler

Proseminar *Text Mining*
Andrea Zielinski

Institut für Computerlinguistik, Universität Heidelberg, 15.07.2016

Strukturierung

- 1 Einleitung
- 2 OIE - Grundlagen
- 3 Anwendungsbeispiel: LODifier
- 4 LODifier und andere Systeme im Kontext
- 5 Fazit und Ausblick

Einleitung



OIE - Grundlagen



OIE - Grundlagen

Motivation und Hürden



OIE - Grundlagen

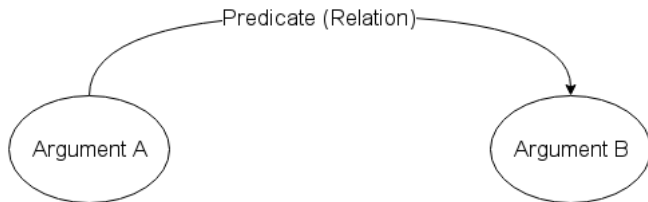
Verfahren



OIE - Grundlagen

Datenrepräsentationsformen

Standardpattern



Argument A geht eine gerichtete **Relation** mit **Argument B** ein.

Unnormierte Annotation

(argument_a, predicate_x, argument_b)
(argument_a, predicate_y, argument_c)
(argument_a, predicate_y, argument_d)

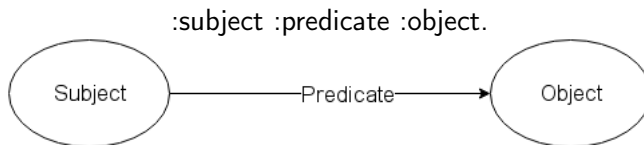
Probleme

- redundant
- unnormiert
- kann nur zweistellige Relationen ausdrücken

RDF und Linked Data

Resource Description Framework

Modelliert Aussagen (triples) über **Subjekte**, **Objekte** und **Prädikate**, in denen das Objekt das Subjekt anhand einer Relation (dem Prädikat) genauer beschreibt. Dabei entsteht ein gerichteter Graph.



RDF Konzepte und Notationen

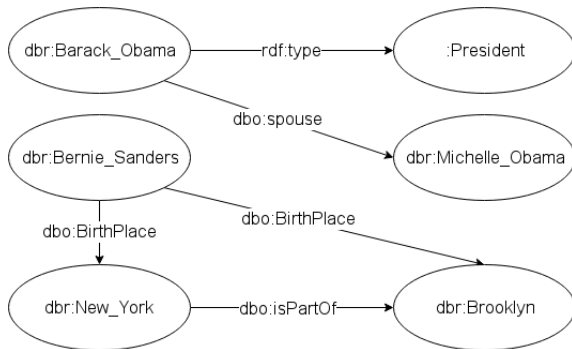
- **URIs**
macht Ressourcen (S, R, O) eindeutig identifizierbar und
verweist auf zusätzliche Informationen
- **Interpretierbar** erlaubt Schlussfolgerungen nach definierten
Regeln
- **Turtle Syntax** erlaubt verkürzte Schreibweisen

RDF Syntax

```
dbr:Barack_Obama a foaf:person, :President;
    dbo:spouse dbr:Michelle_Obama.
dbr:Bernie_Sanders dbo:birthPlace dbr:New_York,
    dbr:Brooklyn;
dbr:Brooklyn dbo:isPartOf dbr:New_York
```

Datenrepräsentationsformen

... als Graph



Anwendungsbeispiel: LODifier

Anwendungsbeispiel: LODifier

Struktur

Anwendungsbeispiel: LODifier

Preprocessing

Anwendungsbeispiel: LODifier

RDF Construction

Anwendungsbeispiel: LODifier

Bewertung

LODifier und andere Systeme im Kontext

LODifier und andere Systeme im Kontext

Bewertung der Ansätze

Fazit und Ausblick



