

Open Information Extraction

Dominik Both, Tonio Weidler

Proseminar *Text Mining*
Andrea Zielinski

Institut für Computerlinguistik, Universität Heidelberg, 15.07.2016

Strukturierung

- 1 Introduction
- 2 OIE - Principles
- 3 Example: LODifier
- 4 OIE Systems in Context
- 5 Conclusion

Introduction

OIE - Principles

OIE - Principles

Motivation

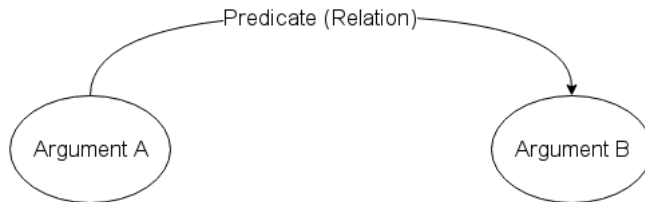
OIE - Principles

Methods

OIE - Principles

Data Representation

Standard Patterns



Argument A is in a directed relation to **Argument B**.

Unnormalized Annotation

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

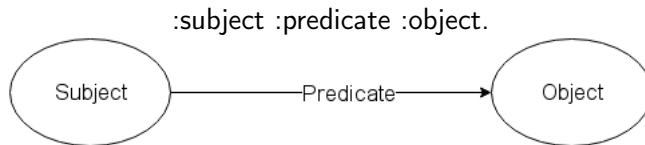
Problems

- redundant
- unnormalized
- can only produce binary predicates

RDF and Linked Data

Resource Description Framework

Models propositions by constructing *triples* including **Subjects**, **Objects** and **Predicates**
Generates a directed graph



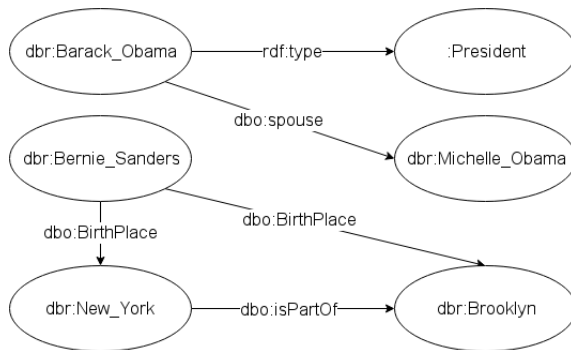
RDF Concepts and Notation

- **URIs**
identifies ressources (S, R, O) distinctivly and references further informations (triples)
- **Conclusions**
allows to draw conclusions using rules
- **Turtle**
allows syntax abbreviations

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
```

... als Graph



Example: LODifier

Example: LODifier

Architecture

Example: LODifier

Preprocessing

Example: LODifier

RDF Construction

Example: LODifier

Conclusions

OIE Systems in Context

OIE Systems in Context

Comparison



OIE Systems in Context

Evaluating the Approaches

Conclusion



Conclusion

Problems and Obstacles



Conclusion Future Opportunities