Master thesis presentation: Mining adverse events from healthcare data

Dries Van Daele

KU Leuven, DTAI

6 september, 2013

Problem description

Problem: voluntary reporting records a fraction of the adverse events

manual detection

- is costly
- is limited in scope
- is driven by intuition

data mining

- treats all patient data uniformly
- can reason over all relevant data
- enables automation
- makes biases explicit

Problem description

Task: Apply data mining on a given database to perform the detection of adverse events

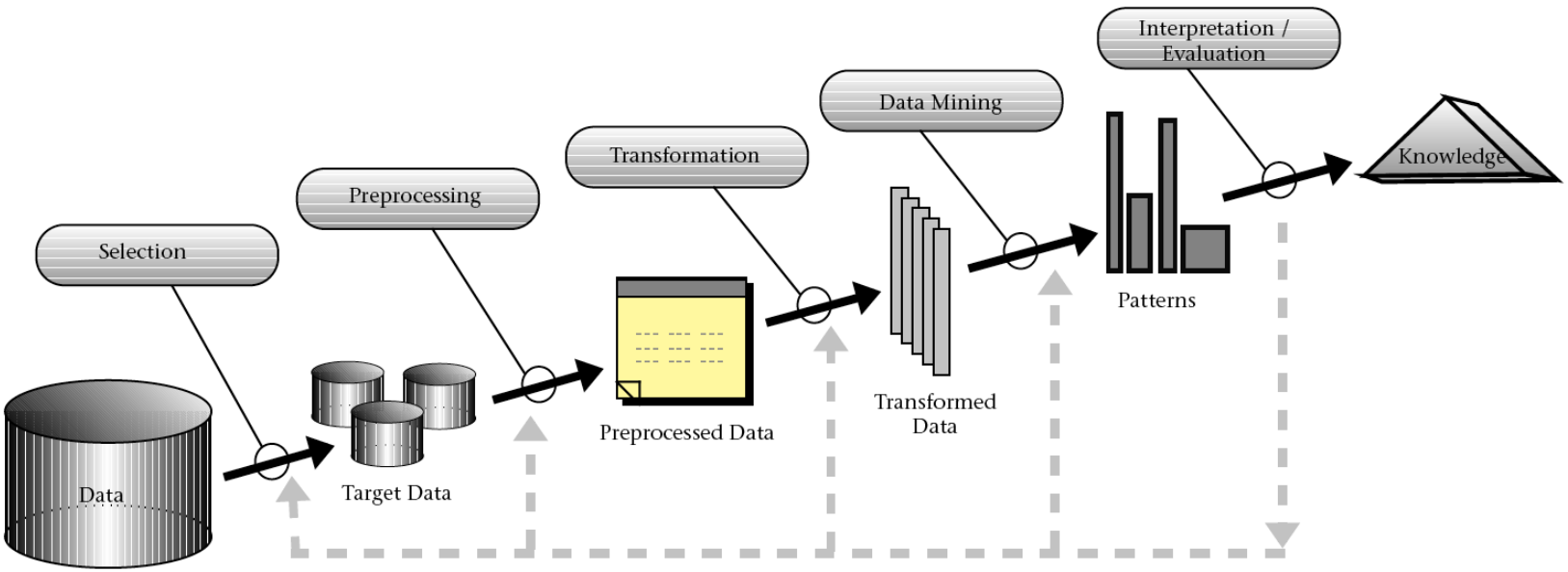
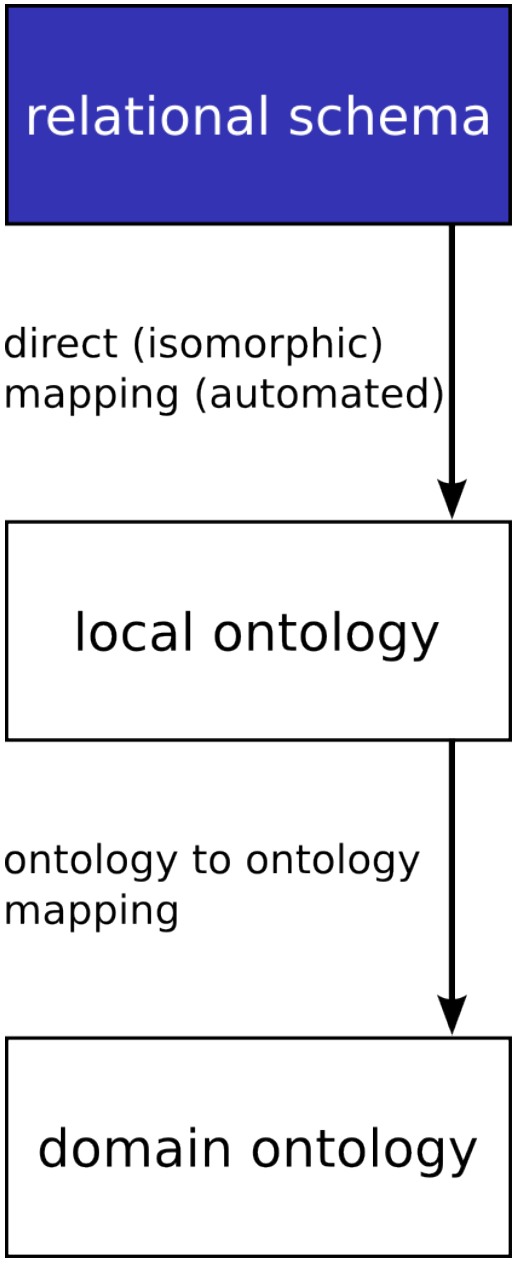


Figure: http://www.kdnuggets.com/gpspubs/aimag-kdd-overview-1996-Fayyad.pdf

Data Preparation

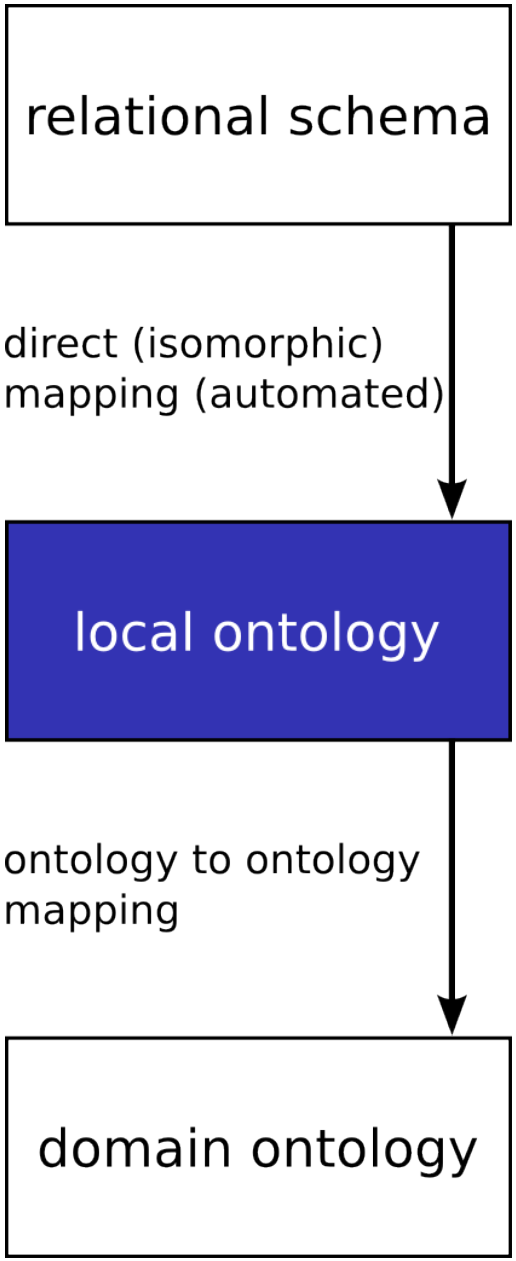


patient_id	birth_date	gender
patient_1	12-AUG-1956	M
...	...	...

medical_case_id	patient_id	admission_date	discharge_date
medical_case_100	patient_1	17-JAN-2009	19-JAN-2009
medical_case_101	patient_1	03-SEP-2009	27-SEP-2009
...	...	...	...

diagnose_id	medical_case_id	ICD_code
diagnose_1	medical_case_101	F48
diagnose_2	medical_case_101	T85.4
...	...	...

Data Preparation




```
@prefix diagnose: <http://www.example.org/ddo/diagnose#>.
@prefix medical-case: <http://www.example.org/ddo/medical-case#>
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>.
@prefix xsd: <http://www.w3.org/2001/XMLSchema#>.

diagnose:Diagnose a rdfs:Class.
diagnose:diagnose-id a rdf:Property ;
                     rdfs:domain diagnose:Diagnose ;
                     rdfs:range xsd:Literal.
diagnose:medical-case-id a rdf:Property ;
                         rdfs:domain diagnose:Diagnose ;
                         rdfs:range medical-case:Medical-Case.
diagnose:icd-code a rdf:Property ;
                  rdfs:domain diagnose:Diagnose ;
                  rdfs:range xsd:Literal

...
```

Data Preparation

relational schema

direct (isomorphic) mapping (automated)

local ontology

ontology to ontology mapping

domain ontology

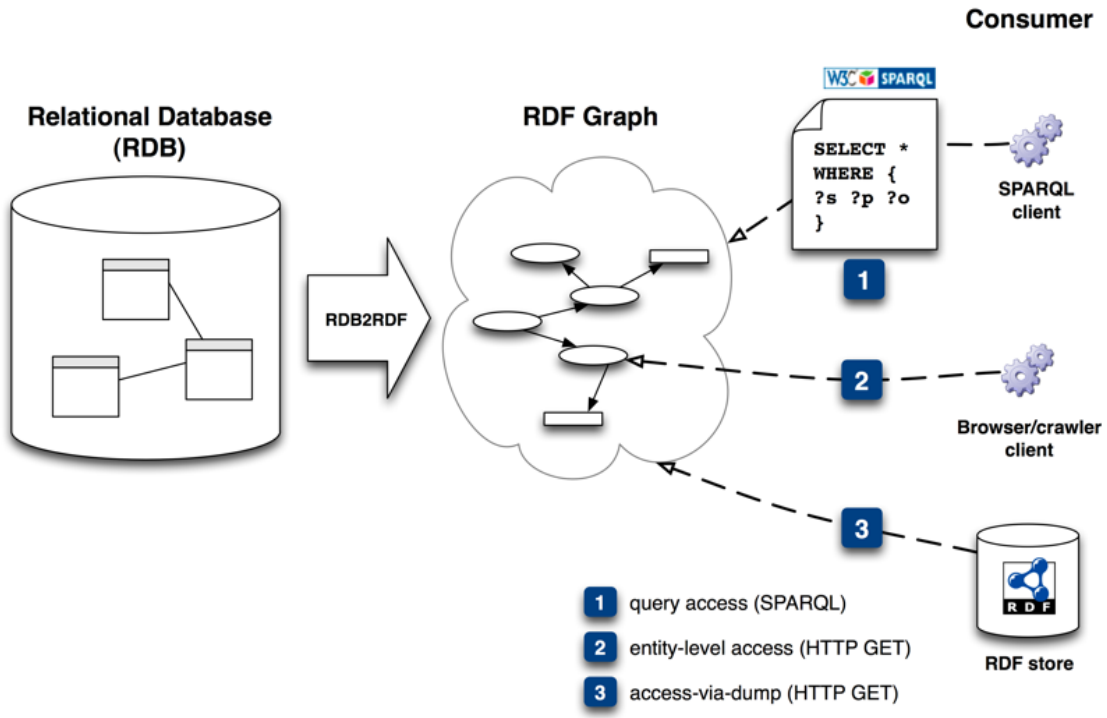
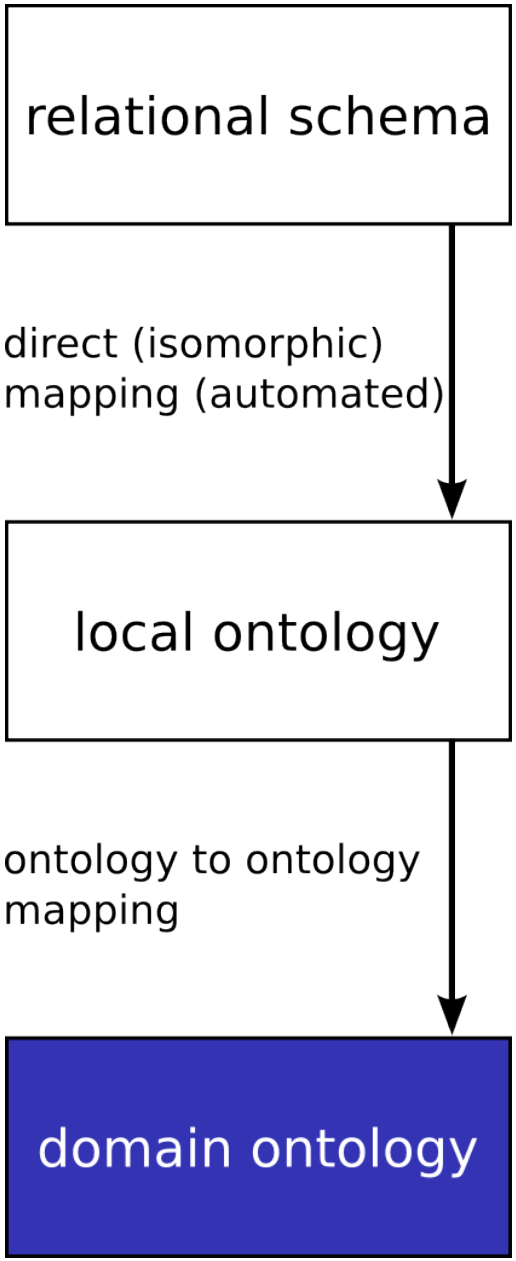


Figure:
http://www.w3.org/2001/sw/rdb2rdf/use-cases/

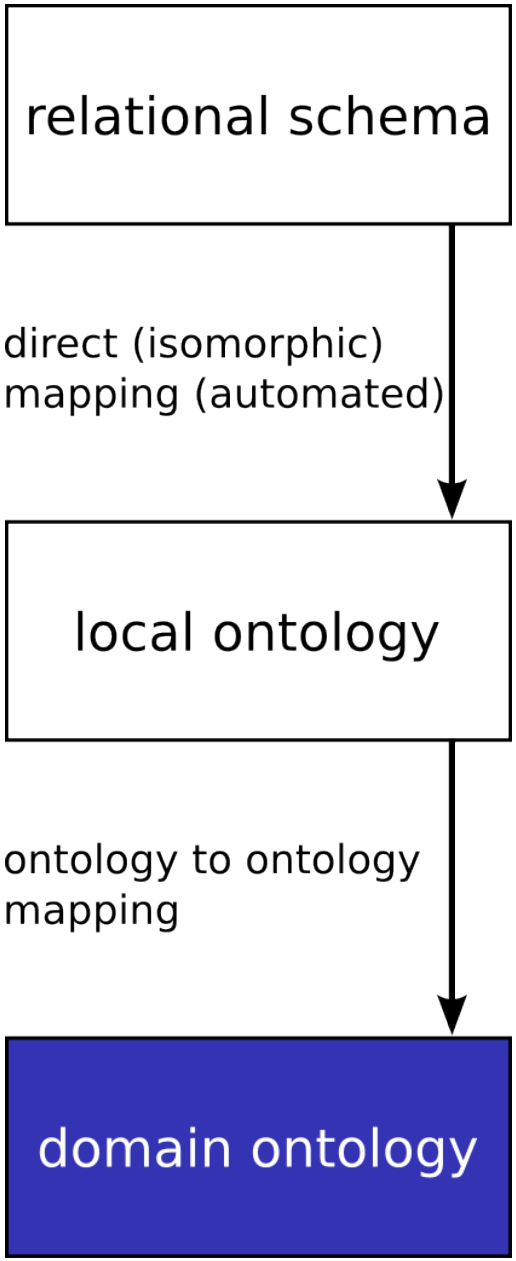
Data Preparation




```
@prefix patient: <http://www.example.org/ddo/patient#>.
@prefix heca: <http://www.agfa.com/w3c/2009/healthCare#>.
{
    ?patient a patient:Patient.
} => {
    ?patient a heca:Patient.
}.

...
```

Data Preparation




```
@prefix orgStructure: <http://www.example.org/ddo/orgStructure#>.
@prefix space: <http://eulersharp.sourceforge.net/2003/03swap/space#>
{
    _:structure
        orgStructure:structID ?s ;
        orgStructure:innerStructID ?is .
} => {
    ?is space:containedBy ?s .
}.

{
    ?startNode space:containedBy ?middleNode .
    ?middleNode space:containedBy ?endNode .
} => {
    ?startNode space:containedBy ?endNode .
}.
```

Problem approach

split problem up AE's based on same trigger

Data mining approach characteristics

- support relational input data
- perform induction
- incorporate reverse machine learning
- probability

Data mining approach characteristics

example patient record timeline

Conclusions

some stuff