# Ontology-based Data Access: Theory and Practice

## **Guohui Xiao**

Roman Kontchakov

KRDB Research Centre

re Department of Computer Science & Inf. Systems

Free University of Bozen-Bolzano

Birkbeck, University of London

http://ontop.inf.unibz.it/ijcai-2018-tutorial

## Motivation: Geologists at Statoil (Equinor)

900 geologists & geophysicists in Statoil Exploration develop stratigraphic models of unexplored areas on the basis of data acquired from previous operations at nearby locations

# Motivation: Geologists at Statoil (Equinor)

900 geologists & geophysicists in Statoil Exploration develop stratigraphic models of unexplored areas on the basis of data acquired from previous operations at nearby locations

G&G's information need (009)

In my area of interest, return all pressure data tagged with key stratigraphy information with understandable QC attributes (and suitable for further filtering).

LITHOSTRATIGRAPHIC CHART

## Motivation: Geologists at Statoil (Equinor)

900 geologists & geophysicists in Statoil Exploration develop stratigraphic models of unexplored areas on the basis of data acquired from previous operations at nearby locations

G&G's information need (009)

In my area of interest, return all pressure data tagged with key stratigraphy information with understandable QC attributes (and suitable for further filtering).

Slegge database contains 1545 tables and 1727 views

table WELLBORE has 38 columns

formation pressure: a join of PTY\_PRESSURE, ACTIVITY, ACTIVITY\_CLASS and WELLBORE stratigraphic information: a join of PICKED\_STRATIGRAPHIC\_ZONES, PTY\_LOCATION\_1D, PTY\_PRESSURE, ACTIVITY, ACTIVITY\_CLASS and WELLBORE

data gathering is a huge problem in industry:
search for data and quality assessment,
e.g., in oil&gas takes 30–70% of engineers' time

(Crompton, 2008)

data gathering is a huge problem in industry:

search for data and quality assessment,

e.g., in oil&gas takes 30-70% of engineers' time

(Crompton, 2008)

# solution 1

use predefined SQL queries of the in-house system to retrieve information about

- (a) pressure measurements,
- (b) lithostratigraphy of wellbores, etc.

and integrate the results using a spreadsheet

data gathering is a huge problem in industry:

search for data and quality assessment,

e.g., in oil&gas takes 30–70% of engineers' time

(Crompton, 2008)

# solution 1

use predefined SQL queries of the in-house system to retrieve information about

- (a) pressure measurements,
- (b) lithostratigraphy of wellbores, etc.

and integrate the results using a spreadsheet

time-consuming, error-prone (e.g., different units of measurement), difficult to repeat

# solution 2

ask an IT expert to translate the information need into SQL

# solution 2

## ask an IT expert to translate the information need into SQL

```
SELECT.
   WELLBORE.IDENTIFIER.
   PTY PRESSURE.PTY_PRESSURE_S.
   STRATIGRAPHIC 7ONE.STRAT COLUMN IDENTIFIER.
   STRATIGRAPHIC ZONE.STRAT UNIT IDENTIFIER
FROM WELLBORE.
   PTY PRESSURE.
   ACTIVITY FP DEPTH DATA
       LEFT JOIN (PTY LOCATION 1D FP DEPTH PT1 LOC
           INNER JOIN PICKED STRATIGRAPHIC ZONES ZS
               ON ZS.STRAT.ZONE.ENTRY.MD <= FP_DEPTH_PT1_LOC.DATA_VALUE_1_O AND
                  ZS.STRAT_ZONE_EXIT_MD >= FP_DEPTH_PT1_LOC.DATA_VALUE_1_O AND
                  ZS.STRAT ZONE DEPTH UOM = FP DEPTH PT1 LOC.DATA VALUE 1 OU
           INNER JOIN STRATIGRAPHIC 70NF
               ON ZS.WELLBORE = STRATIGRAPHIC ZONE.WELLBORE AND
                  ZS.STRAT_COLUMN_IDENTIFIER = STRATIGRAPHIC_ZONE.STRAT_COLUMN_IDENTIFIER AND
                  7S.STRAT INTERP VERSION = STRATIGRAPHIC 7ONE.STRAT INTERP VERSION AND
                  ZS.STRAT ZONE IDENTIFIER = STRATIGRAPHIC ZONE.STRAT ZONE IDENTIFIER)
           ON FP DEPTH DATA FACILITY S = 7S.WELLBORE AND
              FP DEPTH DATA ACTIVITY S = FP DEPTH PT1 LOC ACTIVITY S.
   ACTIVITY CLASS FORM PRESSURE CLASS
WHERE WELLBORE WELLBORE S = FP DEPTH DATA FACILITY S AND
   FP DEPTH DATA.ACTIVITY S = PTY PRESSURE.ACTIVITY S AND
   FP_DEPTH_DATA.KIND.S = FORM_PRESSURE_CLASS.ACTIVITY_CLASS.S AND
   WELLBORE, REF EXISTENCE KIND = 'actual' AND
   FORM_PRESSURE_CLASS.NAME = 'formation pressure depth data'
```

# solution 2

## ask an IT expert to translate the information need into SQL

**SELECT** 

WELLBORE.IDENTIFIER.

PTY\_PRESSURE.PTY\_PRESSURE\_S.

STRATIGRAPHIC\_ZONE.STRAT\_COLUMN\_IDENTIFIER,

STRATIGRAPHIC\_ZONE.STRAT\_UNIT\_IDENTIFIER

FROM WELLBORE.

PTY\_PRESSURE,

ACTIVITY FP\_DEPTH\_DATA

LEFT JOIN (PTY\_LOCATION\_1D FP\_DEPTH\_PT1\_LOC

INNER JOIN PICKED\_STRATIGRAPHIC\_ZONES ZS

ON ZS.STRAT\_ZONE\_ENTRY\_MD <= FP\_DEPTH\_PT1\_LOC.DATA\_VALUE\_1\_O AND

ZS.STRAT.ZONE\_EXIT\_MD >= FP\_DEPTH\_PT1\_LOC.DATA\_VALUE\_1\_O AND ZS.STRAT.ZONE\_DEPTH\_UOM = FP\_DEPTH\_PT1\_LOC.DATA\_VALUE\_1\_OU

INNER JOIN STRATIGRAPHIC 70NF

ON ZS.WELLBORE = STRATIGRAPHIC ZONE.WELLBORE AND

ZS.STRAT\_COLUMN\_IDENTIFIER = STRATIGRAPHIC\_ZONE.STRAT\_COLUMN\_IDENTIFIER AND

ZS.STRAT\_INTERP\_VERSION = STRATIGRAPHIC\_ZONE.STRAT\_INTERP\_VERSION AND

ZS.STRAT\_ZONE\_IDENTIFIER = STRATIGRAPHIC\_ZONE.STRAT\_ZONE\_IDENTIFIER)

ON FP\_DEPTH\_DATA.FACILITY\_S = ZS.WELLBORE AND

FP\_DEPTH\_DATA.ACTIVITY\_S = FP\_DEPTH\_PT1\_LOC.ACTIVITY\_S,

ACTIVITY\_CLASS FORM\_PRESSURE\_CLASS

WHERE WELLBORE.WELLBORE\_S = FP\_DEPTH\_DATA.FACILITY\_S AND FP\_DEPTH\_DATA.ACTIVITY\_S = PTY\_PRESSURE.ACTIVITY\_S AND FP\_DEPTH\_DATA.KIND\_S = FORM\_PRESSURE\_CLASS.ACTIVITY\_CLA

WELLBORE.REF\_EXISTENCE\_KIND = 'actual' AND

FORM\_PRESSURE\_CLASS.NAME = 'formation pressure depth da...

# solution 2

ask an IT expert to translate the information need into SQL

**SELECT** 

WELLBORE.IDENTIFIER, PTY\_PRESSURE.PTY\_PRESSURE\_S,

STDATICDADHIC 70NE STDAT COLLIMN IDENTIFIED

## knowledge of the geological domain and database structure

(1545 tables and 1727 views, magic values, extensive denormalisation)

ON ZS.STRAT\_ZONE\_ENTRY\_MD <= FP\_DEPTH\_PT1\_LOC.DATA\_VALUE\_1\_O AND ZS.STRAT\_ZONE\_EXIT\_MD >= FP\_DEPTH\_PT1\_LOC.DATA\_VALUE\_1\_O AND ZS.STRAT\_ZONE\_DEPTH\_UOM = FP\_DEPTH\_PT1\_LOC.DATA\_VALUE\_1\_OU

INNER JOIN STRATIGRAPHIC\_ZONE

ON ZS.WELLBORE = STRATIGRAPHIC\_ZONE.WELLBORE AND

ZS.STRAT\_COLUMN\_IDENTIFIER = STRATIGRAPHIC\_ZONE.STRAT\_COLUMN\_IDENTIFIER AND ZS.STRAT\_INTERP\_VERSION = STRATIGRAPHIC\_ZONE.STRAT\_INTERP\_VERSION AND ZS.STRAT\_ZONE IDENTIFIER = STRATIGRAPHIC\_ZONE.STRAT\_ZONE IDENTIFIER)

ON FP\_DEPTH\_DATA.FACILITY\_S = ZS.WELLBORE AND

FP\_DEPTH\_DATA.ACTIVITY\_S = FP\_DEPTH\_PT1\_LOC.ACTIVITY\_S,

ACTIVITY\_CLASS FORM\_PRESSURE\_CLASS

WHERE WELLBORE.WELLBORE.S = FP\_DEPTH\_DATA.FACILITY\_S AND FP\_DEPTH\_DATA.ACTIVITY\_S = PTY\_PRESSURE.ACTIVITY\_S AND FP\_DEPTH\_DATA.KIND\_S = FORM\_PRESSURE.CLASS.ACTIVITY\_CLA WELLBORE.REF\_EXISTENCE\_KIND = 'actual' AND FORM\_PRESSURE\_CLASS.NAME = 'formation pressure depth data.

# solution 2

ask an IT expert to translate the information need into SQL

**SELECT** 

WELLBORE.IDENTIFIER,
PTY\_PRESSURE.PTY\_PRESSURE\_S,
STDATIGDADHIC 7ONE STDAT COLLIMN IDENTIFIED

knowledge of the geological domain and database structure

(1545 tables and 1727 views, magic values, extensive denormalisation)

ON ZS.STRAT.ZONE\_ENTRY\_MD <= FP\_DEPTH\_PT1\_LOC.DATA\_VALUE\_1\_O AND ZS.STRAT.ZONE\_EXIT\_MD >= FP\_DEPTH\_PT1\_LOC.DATA\_VALUE\_1\_O AND ZS.STRAT\_ZONE\_EXIT\_MO == FP\_DEPTH\_PT1\_DATA\_VALUE\_1\_O AND ZS.STRAT\_VALUE\_1\_O AND ZS.STRAT\_ZS.ST

# encode the domain knowledge in an ontology

ZS.STRAT\_COLOIVIN\_IDENTIFIER = STRATIGRAPHIC\_ZONE.STRAT\_INTERP\_VERSION AND
ZS.STRAT\_INTERP\_VERSION = STRATIGRAPHIC\_ZONE.STRAT\_INTERP\_VERSION AND
ZS.STRAT\_ZONE\_IDENTIFIER = STRATIGRAPHIC\_ZONE.STRAT\_ZONE\_IDENTIFIER)
ON EP\_DEPTH\_DATA\_FACILITY\_S = 7S.WELLBORE\_AND

FP\_DEPTH\_DATA.ACTIVITY\_S = FP\_DEPTH\_PT1\_LOC.ACTIVITY\_S,

ACTIVITY CLASS FORM PRESSURE CLASS

WHERE WELLBORE.WELLBORE.S = FP\_DEPTH\_DATA.FACILITY\_S AND FP\_DEPTH\_DATA.ACTIVITY\_S = PTY\_PRESSURE.ACTIVITY\_S AND FP\_DEPTH\_DATA.KIND\_S = FORM\_PRESSURE.CLASS.ACTIVITY\_CLA WELLBORE.REF\_EXISTENCE\_KIND = 'actual' AND FORM\_PRESSURE\_CLASS.NAME = 'formation pressure depth data.

# solution 2

ask an IT expert to translate the information need into SQL

**SELECT** 

WELLBORE.IDENTIFIER, PTY\_PRESSURE.PTY\_PRESSURE\_S,

STDATICDADHIC 70NE STDAT COLLIMN IDENTIFIED

knowledge of the **geological domain** and **database structure** 

(1545 tables and 1727 views, magic values, extensive denormalisation)

ON ZS.STRAT.ZONE\_ENTRY\_MD <= FP\_DEPTH\_PT1\_LOC.DATA\_VALUE\_1\_O AND ZS.STRAT.ZONE\_EXIT\_MD >= FP\_DEPTH\_PT1\_LOC.DATA\_VALUE\_1\_O AND ZS.STRAT\_ZONE\_EXIT\_MO == FP\_DEPTH\_PT1\_DATA\_VALUE\_1\_O AND ZS.STRAT\_VALUE\_1\_O AND ZS.STRAT\_ZS.ST

encode the domain knowledge in an ontology

and database structure in mappings

ON

FP\_DEPTH\_DATA.ACTIVITY\_S = FP\_DEPTH\_PT1\_LOC.ACTIVITY\_S,

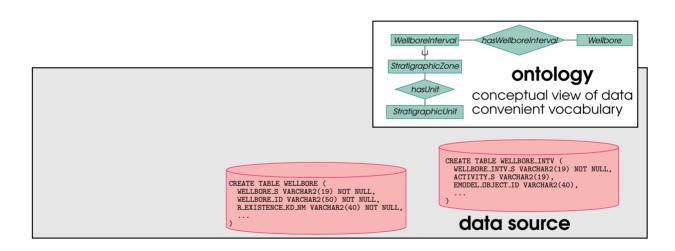
ACTIVITY CLASS FORM PRESSURE CLASS

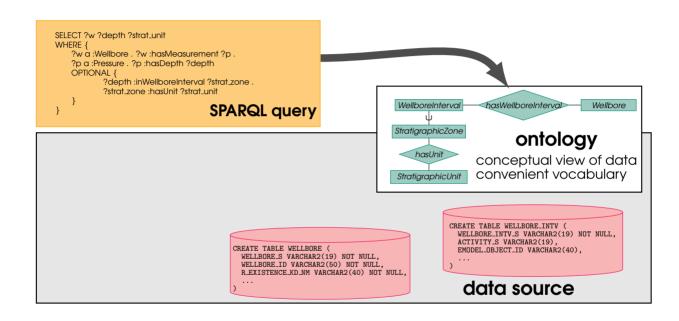
WHERE WELLBORE.WELLBORE.S = FP\_DEPTH\_DATA.FACILITY\_S AND FP\_DEPTH\_DATA.ACTIVITY\_S = PTY\_PRESSURE.ACTIVITY\_S AND FP\_DEPTH\_DATA.KIND\_S = FORM\_PRESSURE.CLASS.ACTIVITY\_CLA WELLBORE.REF\_EXISTENCE\_KIND = 'actual' AND FORM\_PRESSURE\_CLASS.NAME = 'formation pressure depth data.

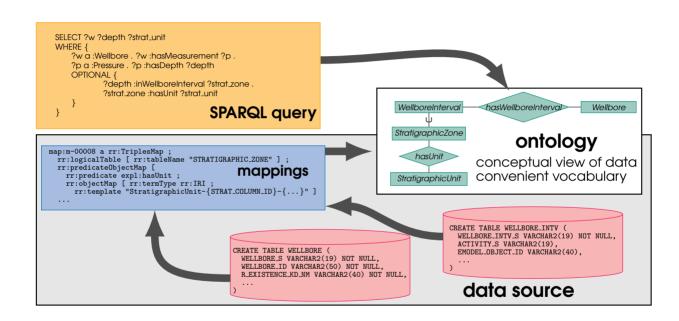
```
CREATE TABLE WELLBORE (
WELLBORE S VARCHAR2(19) NOT NULL,
WELLBORE ID VARCHAR2(50) NOT NULL,
R.EXISTENCE KD.NM VARCHAR2(40) NOT NULL,

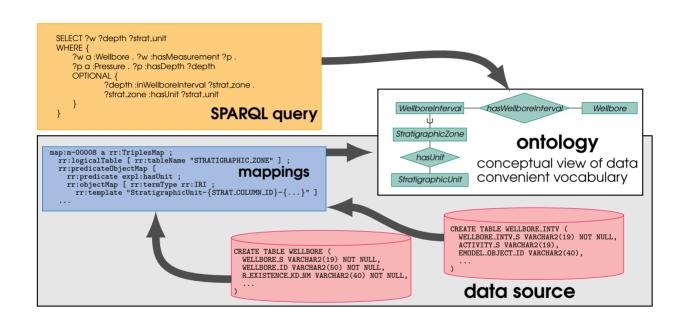
...

data source
```









reduced time for translating information needs into (SPARQL) queries

days  $\longrightarrow$  minutes

#### **Tutorial Plan**

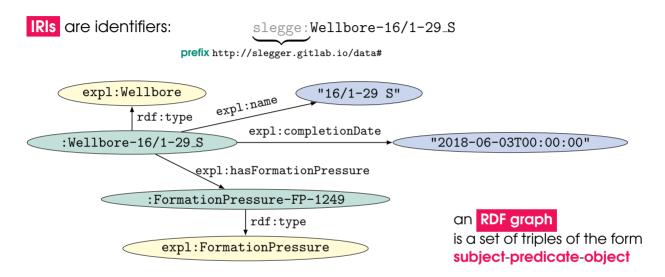
- Semantic Web standards
  - RDF (Data Model)
  - OWL 2 QL (Ontology Language)
  - SPARQL (Query Language)
  - R2RML (Mapping Language)
- Tutorial with Ontop
- coffee break
- Using OBDA in Practice
- Basics: Query Rewriting and Optimisation
- Extending the Foundations
  - Approximating Expressive Ontologies
  - Dealing with Identity: SameAs
  - Ontology-Mediated Query Answering and Circuit Complexity
- Recent Advances / Challenges

# RDF: Resource Description Framework (RDF in 2004, RDF 1.1 in 2014)

IRIs are identifiers: slegge: Wellbore-16/1-29\_S

prefix http://slegger.gitlab.io/data#

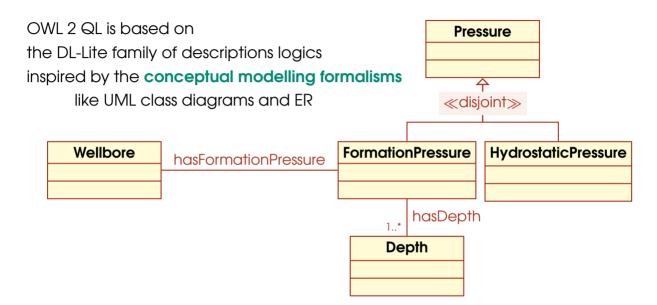
## RDF: Resource Description Framework (RDF in 2004, RDF 1.1 in 2014)

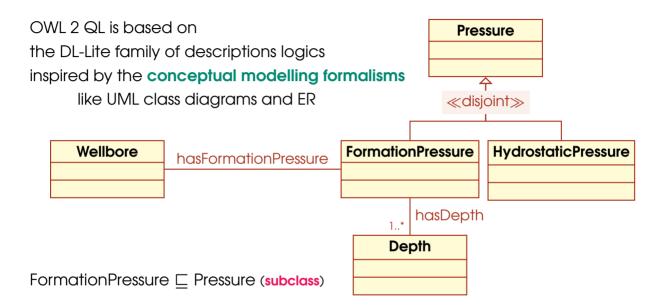


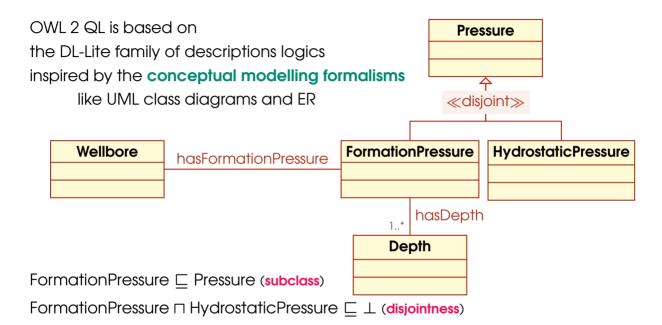
## RDF: Resource Description Framework (RDF in 2004, RDF 1.1 in 2014)

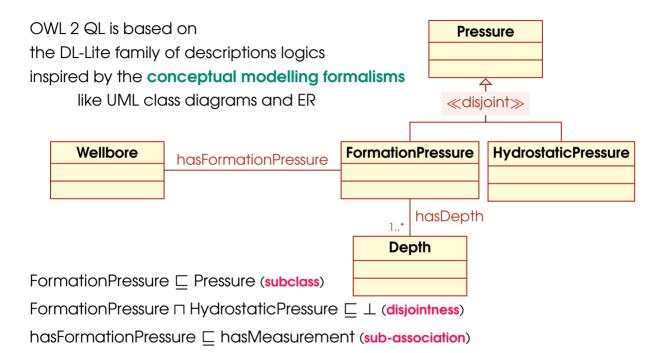
**IRIs** are identifiers: slegge: Wellbore-16/1-29\_S prefix http://slegger.gitlab.io/data# expl:Wellbore "16/1-29 S" expl:name rdf:type expl:completionDate :Wellbore-16/1-29\_S "2018-06-03T00:00:00" expl:hasFormationPressure :FormationPressure-FP-1249 an RDF graph rdf:type is a set of triples of the form expl:FormationPressure subject-predicate-object @prefix : <http://slegger.gitlab.io/data#> . @prefix expl: <a href="http://slegger.gitlab.io/slegge-obda/ontology/subsurface-exploration#"> .</a>. **Turtle syntax** :Wellbore-16/1-29\_S rdf:type expl:Wellbore . :Wellbore-16/1-29\_S expl:name "16/1-29 S" . :Wellbore-16/1-29\_S expl:completionDate "2018-06-03T00:00:00"^xsd:dateTime . :Wellbore-16/1-29\_S expl:hasFormationPressure :FormationPressure-FP-1249 . :FormationPressure-FP-1249 rdf:type expl:FormationPressure .

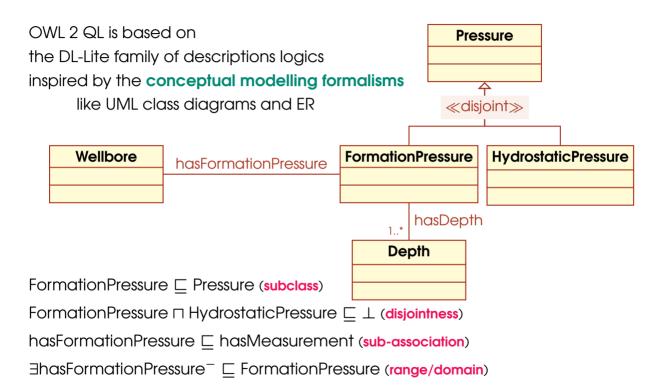
OWL 2 QL is based on the DL-Lite family of descriptions logics inspired by the **conceptual modelling formalisms** like UML class diagrams and ER

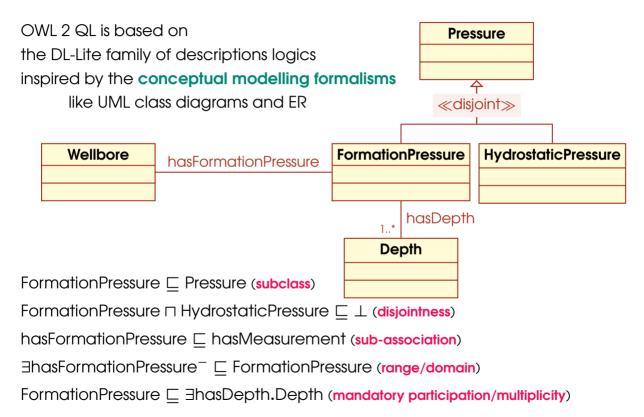












DLs are fragments of First-Order Logic:

classes are unary predicates / properties are binary predicates

DLs are fragments of First-Order Logic:

classes are unary predicates / properties are binary predicates

FormationPressure 

☐ Pressure

DLs are fragments of First-Order Logic:

classes are unary predicates / properties are binary predicates

FormationPressure 

□ Pressure

 $\forall x \ (FormationPressure(x) \rightarrow Pressure(x))$ 

DLs are fragments of First-Order Logic:

classes are unary predicates / properties are binary predicates

FormationPressure 

□ Pressure

 $\forall x \ (\text{FormationPressure}(x) \rightarrow \text{Pressure}(x))$ 

FormationPressure  $\Box$  HydrostaticPressure  $\sqsubseteq$   $\bot$ 

DLs are fragments of First-Order Logic:

classes are unary predicates / properties are binary predicates

FormationPressure 

□ Pressure

 $\forall x \, (\text{FormationPressure}(x) \rightarrow \text{Pressure}(x))$ 

FormationPressure  $\Box$  HydrostaticPressure  $\sqsubseteq$   $\bot$ 

 $\forall x \ (\text{FormationPressure}(x) \land \text{HydrostaticPressure}(x) \rightarrow \bot)$ 

DLs are fragments of First-Order Logic:

classes are unary predicates / properties are binary predicates

FormationPressure 

□ Pressure

 $\forall x \, (\text{FormationPressure}(x) \rightarrow \text{Pressure}(x))$ 

FormationPressure  $\Box$  HydrostaticPressure  $\sqsubseteq$   $\bot$ 

 $\forall x \ ( Formation Pressure(x) \land Hydrostatic Pressure(x) \rightarrow \bot )$ 

hasFormationPressure 

□ hasMeasurement

DLs are fragments of First-Order Logic:

classes are unary predicates / properties are binary predicates

FormationPressure 

□ Pressure

 $\forall x \, (\text{FormationPressure}(x) \rightarrow \text{Pressure}(x))$ 

FormationPressure  $\Box$  HydrostaticPressure  $\sqsubseteq$   $\bot$ 

 $\forall x \ (\text{FormationPressure}(x) \land \text{HydrostaticPressure}(x) \rightarrow \bot)$ 

hasFormationPressure 

□ hasMeasurement

 $\forall xy \ (\mathsf{hasFormationPressure}(x,y) \to \mathsf{hasMeasurement}(x,y))$ 

DLs are fragments of First-Order Logic:

classes are unary predicates / properties are binary predicates

FormationPressure 

☐ Pressure

 $\forall x \, (Formation Pressure(x) \rightarrow Pressure(x))$ 

FormationPressure  $\Box$  HydrostaticPressure  $\sqsubseteq$   $\bot$ 

 $\forall x \ ( Formation Pressure(x) \land Hydrostatic Pressure(x) \rightarrow \bot )$ 

hasFormationPressure 

□ hasMeasurement

 $\forall xy \ (\text{hasFormationPressure}(x,y) \rightarrow \text{hasMeasurement}(x,y))$ 

 $\exists$ hasFormationPressure $^ \sqsubseteq$  FormationPressure

DLs are fragments of First-Order Logic:

classes are unary predicates / properties are binary predicates

FormationPressure □ Pressure

 $\forall x \ (Formation Pressure(x) \rightarrow Pressure(x))$ 

FormationPressure □ HydrostaticPressure □ ⊥

 $\forall x \ (\text{FormationPressure}(x) \land \text{HydrostaticPressure}(x) \rightarrow \bot)$ 

hasFormationPressure 

□ hasMeasurement

 $\forall xy \ (\mathsf{hasFormationPressure}(x,y) \to \mathsf{hasMeasurement}(x,y))$ 

∃hasFormationPressure 

FormationPressure

 $\forall xy \ (\mathsf{hasFormationPressure}(x,y) \to \mathsf{FormationPressure}(y))$ 

DLs are fragments of First-Order Logic:

classes are unary predicates / properties are binary predicates

FormationPressure 

☐ Pressure

 $\forall x \ (Formation Pressure(x) \rightarrow Pressure(x))$ 

FormationPressure □ HydrostaticPressure □ ⊥

 $\forall x \ (\text{FormationPressure}(x) \land \text{HydrostaticPressure}(x) \rightarrow \bot)$ 

hasFormationPressure 

□ hasMeasurement

 $\forall xy \ (\mathsf{hasFormationPressure}(x,y) \to \mathsf{hasMeasurement}(x,y))$ 

∃hasFormationPressure 

FormationPressure

 $\forall xy \ (\mathsf{hasFormationPressure}(x,y) \to \mathsf{FormationPressure}(y))$ 

FormationPressure 

∃hasDepth.Depth

## DLs are fragments of First-Order Logic:

classes are unary predicates / properties are binary predicates

FormationPressure 

☐ Pressure

 $\forall x \ (Formation Pressure(x) \rightarrow Pressure(x))$ 

FormationPressure □ HydrostaticPressure □ ⊥

 $\forall x \ (\text{FormationPressure}(x) \land \text{HydrostaticPressure}(x) \rightarrow \bot)$ 

hasFormationPressure 

□ hasMeasurement

 $\forall xy \ (\mathsf{hasFormationPressure}(x,y) \to \mathsf{hasMeasurement}(x,y))$ 

∃hasFormationPressure 

FormationPressure

 $\forall xy \ (\mathsf{hasFormationPressure}(x,y) o \mathsf{FormationPressure}(y))$ 

FormationPressure 

∃hasDepth.Depth

 $\forall x \, (\text{FormationPressure}(x) \rightarrow \exists y \, (\text{hasDepth}(x,y) \land \text{Depth}(y)))$ 

```
DLs are fragments of First-Order Logic:
```

classes are unary predicates / properties are binary predicates

FormationPressure □ Pressure

 $\forall x \ (Formation Pressure(x) \rightarrow Pressure(x))$ 

FormationPressure □ HydrostaticPressure □ ⊥

 $\forall x \ (\text{FormationPressure}(x) \land \text{HydrostaticPressure}(x) \rightarrow \bot)$ 

hasFormationPressure 

□ hasMeasurement

 $\forall xy \ (\mathsf{hasFormationPressure}(x,y) \to \mathsf{hasMeasurement}(x,y))$ 

∃hasFormationPressure 

FormationPressure

 $\forall xy \ (\mathsf{hasFormationPressure}(x,y) \to \mathsf{FormationPressure}(y))$ 

FormationPressure 

∃hasDepth.Depth

 $\forall x \, (\text{FormationPressure}(x) \rightarrow \exists y \, (\text{hasDepth}(x,y) \land \text{Depth}(y)))$ 

## tuple-generating dependencies / existential rules

Web Ontology Language (OWL) was standardised in 2004; OWL 2 in 2012

OWL 2 QL is one of the three profiles of OWL 2

(fragments tailored for specific applications)

#### OWL 2 QL contains

subclass axioms, where

existential quantification (ObjectSomeValuesFrom) is limited to owl:Thing

Web Ontology Language (OWL) was standardised in 2004; OWL 2 in 2012

OWL 2 QL is one of the three profiles of OWL 2

(fragments tailored for specific applications)

#### OWL 2 QL contains

- subclass axioms, where
   existential quantification (ObjectSomeValuesFrom) is limited to owl:Thing
- inverse object properties (InverseObjectProperties)

Web Ontology Language (OWL) was standardised in 2004; OWL 2 in 2012

OWL 2 QL is one of the three profiles of OWL 2

(fragments tailored for specific applications)

#### OWL 2 QL contains

- subclass axioms, where
   existential quantification (ObjectSomeValuesFrom) is limited to owl:Thing
- inverse object properties (InverseObjectProperties)
- property domain axioms (ObjectPropertyDomain and DataPropertyDomain)
- property range axioms (ObjectPropertyRange and DataPropertyRange)

Web Ontology Language (OWL) was standardised in 2004; OWL 2 in 2012

OWL 2 QL is one of the three profiles of OWL 2

(fragments tailored for specific applications)

#### OWL 2 QL contains

- subclass axioms, where
   existential quantification (ObjectSomeValuesFrom) is limited to owl:Thing
- inverse object properties (InverseObjectProperties)
- property domain axioms (ObjectPropertyDomain and DataPropertyDomain)
- property range axioms (ObjectPropertyRange and DataPropertyRange)
- property inclusions (SubObjectPropertyOf but no property chains)

Web Ontology Language (OWL) was standardised in 2004; OWL 2 in 2012

OWL 2 QL is one of the three profiles of OWL 2

(fragments tailored for specific applications)

#### OWL 2 QL contains

- subclass axioms, where
   existential quantification (ObjectSomeValuesFrom) is limited to owl:Thing
- inverse object properties (InverseObjectProperties)
- property domain axioms (ObjectPropertyDomain and DataPropertyDomain)
- property range axioms (ObjectPropertyRange and DataPropertyRange)
- property inclusions (SubObjectPropertyOf but no property chains)

#### does not contain

- universal quantification (ObjectAllValuesFrom, DataAllValuesFrom)
- enumeration of individuals and literals (ObjectOneOf, DataOneOf)
- disjunction (ObjectUnionOf, DisjointUnion and DataUnionOf)
- individual equality assertions (sameAs) (SameIndividual)

SPARQL is the query language for RDF

```
SELECT ?w ?d
WHERE {
    ?w a expl:Wellbore .
    ?w expl:hasMeasurement ?p .
    ?p a expl:Pressure .
    ?p expl:hasDepth ?d
}
```

SPARQL is the guery language for RDF

```
:Wellbore-16/1-29_S a expl:Wellbore .
:Wellbore-30/8-5 a expl:Wellbore .
:Wellbore-16/1-29_S expl:hasMeasurement :FormationPressure-FP-1249 .
:FormationPressure-FP-1249 a expl:Pressure .
:FormationPressure-FP-1249 expl:hasDepth :TVD-FP-1249 .
:Wellbore-16/1-29_S expl:hasMeasurement :FormationPressure-FP-1377 .
:FormationPressure-FP-1377 a expl:Pressure .
:FormationPressure-FP-1377 expl:hasDepth :TVD-FP-1377 .
```

SPARQL is the guery language for RDF

```
variables to appear in the query results
 SELECT ?w ?d
 WHERE {
      ?w a expl:Wellbore .
                                          triple patterns separated by.
      ?w expl:hasMeasurement ?p .
                                          a abbreviates rdf:type
      ?p a expl:Pressure .
      ?p expl:hasDepth ?d
          `find all depths of pressure measurements for all wellbores'
                                  :Wellbore-16/1-29_S a expl:Wellbore .
                                  :Wellbore-30/8-5 a expl:Wellbore .
                                  :Wellbore-16/1-29_S expl:hasMeasurement :FormationPressure-FP-1249 .
                                  :FormationPressure-FP-1249 a expl:Pressure .
                                  :FormationPressure-FP-1249 expl:hasDepth :TVD-FP-1249 .
                                  :Wellbore-16/1-29_S expl:hasMeasurement :FormationPressure-FP-1377 .
                                  :FormationPressure-FP-1377 a expl:Pressure .
                                  :FormationPressure-FP-1377 expl:hasDepth :TVD-FP-1377 .
answer:
                     ?d
  :Wellbore-16/1-29_S
                     :TVD-FP-1249
  :Wellbore-16/1-29_S
                     :TVD-FP-1377
```

SPARQL is the guery language for RDF

IJCAI 2018, Stockholm, 14,07,2018

```
variables to appear in the query results
 SELECT ?w ?d
 WHERE {
      ?w a expl:Wellbore .
                                          triple patterns separated by.
      ?w expl:hasMeasurement ?p .
                                          a abbreviates rdf:type
      ?p a expl:Pressure .
      ?p expl:hasDepth ?d
          `find all depths of pressure measurements for all wellbores'
                                 :Wellbore-16/1-29_S a expl:Wellbore .
                                 :Wellbore-30/8-5 a expl:Wellbore .
                                 :Wellbore-16/1-29_S expl:hasMeasurement :FormationPressure-FP-1249 .
                                 :FormationPressure-FP-1249 a expl:Pressure .
                                 :FormationPressure-FP-1249 expl:hasDepth :TVD-FP-1249 .
                                 :Wellbore-16/1-29_S expl:hasMeasurement :FormationPressure-FP-1377 .
                                 :FormationPressure-FP-1377 a expl:Pressure .
                                 :FormationPressure-FP-1377 expl:hasDepth :TVD-FP-1377 .
answer:
                     ?d
  :Wellbore-16/1-29_S
                     :TVD-FP-1249
  :Wellbore-16/1-29_S
                     :TVD-FP-1377
                                              expl:Wellbore
                                                                            expl:Pressure
                                 rdf:type
                                                               rdf:type
      graph matching
                                                                            expl:hasDepth
                                               expl:hasMeasurement
```

```
SELECT ?w
WHERE {
          { ?w a expl:Wellbore }
          UNION
          { ?w a expl:ExplorationWellbore }
}
```

:Wellbore-16/1-29\_S :Wellbore-1/2-U-3

SELECT with Basic Graph Patterns = Conjunctive Queries (CQs)

 $\exists p \ [\mathsf{Wellbore}(w) \land \mathsf{hasMeasurement}(w,p) \land \mathsf{Pressure}(p) \land \mathsf{hasDepth}(p,d)]$ 

SELECT with Basic Graph Patterns + UNION ≈ Unions of Conjunctive Queries (UCQs)

ExplorationWellbore $(w) \vee Wellbore(w)$ 

SELECT with Basic Graph Patterns = Conjunctive Queries (CQs)

 $\exists p \ [\mathsf{Wellbore}(w) \land \mathsf{hasMeasurement}(w,p) \land \mathsf{Pressure}(p) \land \mathsf{hasDepth}(p,d)]$ 

SELECT with Basic Graph Patterns + UNION ≈ Unions of Conjunctive Queries (UCQs)

 $\mathsf{ExplorationWellbore}(w) \lor \mathsf{Wellbore}(w)$ 

observe that expl:ExplorationWellbore is naturally a subclass of expl:Wellbore

why do we need the UNION above then?

## **Entailment Regime**

OWL 2 Direct Semantics Entailment Regime is part of SPARQL 1.1 recommendation it allows to take account of the **ontology** (subclass & subproperty axioms, etc.) when **answering** SPARQL queries

## **Entailment Regime**

OWL 2 Direct Semantics Entailment Regime is part of SPARQL 1.1 recommendation it allows to take account of the **ontology** (subclass & subproperty axioms, etc.) when **answering** SPARQL queries

if  $expl:ExplorationWellbore \sqsubseteq expl:Wellbore$  , then

```
SELECT DISTINCT ?w
WHERE {
    ?w a expl:Wellbore
}
```

is equivalent (under the entailment regime) to

```
SELECT DISTINCT ?w
WHERE {
          { ?w a expl:Wellbore }
          UNION
          { ?w a expl:ExplorationWellbore }
}
```

## **Entailment Regime**

OWL 2 Direct Semantics Entailment Regime is part of SPARQL 1.1 recommendation it allows to take account of the **ontology** (subclass & subproperty axioms, etc.) when **answering** SPARQL queries

if  $\mathsf{expl}\mathsf{:}\mathsf{Exploration}\mathsf{Wellbore} \sqsubseteq \mathsf{expl}\mathsf{:}\mathsf{Wellbore}$  , then

```
SELECT DISTINCT ?w
WHERE {
    ?w a expl:Wellbore
}
```

is equivalent (under the entailment regime) to

```
SELECT DISTINCT ?w
WHERE {
      { ?w a expl:Wellbore }
      UNION
      { ?w a expl:ExplorationWellbore }
}
```

more precisely, when evaluating basic graph patterns (BGPs), graph matching is replaced by 'entailed by the ontology'

the semantics of all other constructs of SPARQL is the same

```
SELECT ?w ?d ?u
WHERE {
     ?w a expl:Wellbore . ?w expl:hasMeasurement ?p .
    ?p a expl:Pressure . ?p expl:hasDepth ?d
                                                             if the optional part does not match,
    OPTIONAL {
                                                             it creates no bindings
          ?d expl:inWellboreInterval ?z .
                                                             but does not eliminate the solution
          ?z expl:hasUnit ?u
                                 :Wellbore-16/1-29_S a expl:Wellbore .
                                 :Wellbore-30/8-5 a expl:Wellbore .
                                 :Wellbore-16/1-29_S expl:hasMeasurement :FormationPressure-FP-1249 .
                                 :FormationPressure-FP-1249 a expl:Pressure .
                                 :FormationPressure-FP-1249 expl:hasDepth :TVD-FP-1249 .
                                 :Wellbore-16/1-29_S expl:hasMeasurement :FormationPressure-FP-1377 .
                                 :FormationPressure-FP-1377 a expl:Pressure .
                                 :FormationPressure-FP-1377 expl:hasDepth :TVD-FP-1377 .
                                 :TVD-FP-1377 expl:inWellboreInterval :SZ-4 .
                                 :SZ-4 expl:hasUnit :Stratigraphic-Unit-ROGALAND .
```

```
SELECT ?w ?d ?u
WHERE {
     ?w a expl:Wellbore . ?w expl:hasMeasurement ?p .
     ?p a expl:Pressure . ?p expl:hasDepth ?d
                                                             if the optional part does not match,
    OPTIONAL {
                                                             it creates no bindings
          ?d expl:inWellboreInterval ?z .
                                                             but does not eliminate the solution
          ?z expl:hasUnit ?u
                                 :Wellbore-16/1-29_S a expl:Wellbore .
                                 :Wellbore-30/8-5 a expl:Wellbore .
                                 :Wellbore-16/1-29_S expl:hasMeasurement :FormationPressure-FP-1249 .
                                 :FormationPressure-FP-1249 a expl:Pressure .
                                 :FormationPressure-FP-1249 expl:hasDepth :TVD-FP-1249 .
                                 :Wellbore-16/1-29_S expl:hasMeasurement :FormationPressure-FP-1377 .
                                 :FormationPressure-FP-1377 a expl:Pressure .
                                 :FormationPressure-FP-1377 expl:hasDepth :TVD-FP-1377 .
                                 :TVD-FP-1377 expl:inWellboreInterval :SZ-4 .
                                 :SZ-4 expl:hasUnit :Stratigraphic-Unit-ROGALAND .
```

#### answer:

```
SELECT ?w ?d ?u
WHERE {
     ?w a expl:Wellbore . ?w expl:hasMeasurement ?p .
    ?p a expl:Pressure . ?p expl:hasDepth ?d
                                                             if the optional part does not match,
    OPTIONAL {
                                                             it creates no bindings
          ?d expl:inWellboreInterval ?z .
                                                             but does not eliminate the solution
          ?z expl:hasUnit ?u
                                 :Wellbore-16/1-29_S a expl:Wellbore .
                                 :Wellbore-30/8-5 a expl:Wellbore .
                                 :Wellbore-16/1-29_S expl:hasMeasurement :FormationPressure-FP-1249 .
                                 :FormationPressure-FP-1249 a expl:Pressure .
                                 :FormationPressure-FP-1249 expl:hasDepth :TVD-FP-1249 .
                                 :Wellbore-16/1-29_S expl:hasMeasurement :FormationPressure-FP-1377 .
                                 :FormationPressure-FP-1377 a expl:Pressure .
                                 :FormationPressure-FP-1377 expl:hasDepth :TVD-FP-1377 .
                                 :TVD-FP-1377 expl:inWellboreInterval :SZ-4 .
                                 :SZ-4 expl:hasUnit :Stratigraphic-Unit-ROGALAND .
```

#### answer:

?₩	?d	?u
:Wellbore-16/1-29_S	:TVD-FP-1249	_
:Wellbore-16/1-29_S	:TVD-FP-1377	:Stratigraphic-Unit-ROGALAND

#### NB: similar to LEFT JOIN in SQL

## Remarks on SPARQL 1.1

we have seen the following features of SPARQL:

- Basic Graph Patterns
- UNION
- OPTIONAL

## Remarks on SPARQL 1.1

we have seen the following features of SPARQL:

- Basic Graph Patterns
- UNION
- OPTIONAL

## SPARQL 1.1 has many additional features:

- complex FILTER conditions
- GROUP BY, to express aggregations and support aggregation operators
- MINUS, to remove possible solutions
- FILTER NOT EXISTS, to test for the absence of a pattern
- property paths (regular expressions)
- solution modifiers (LIMIT, ORDER BY)
- CONSTRUCT / ASK / DESCRIBE queries
- ...

 **source query** (in SQL for relational databases)

target query (atoms in the ontology vocabulary)

 source query
(in SQL for relational databases)
target query
(atoms in the ontology vocabulary)

the iri function constructs IRIs by concatenating any number of strings

#### WELLBORE table

IDENTIFIER	REF_EXISTENCE_KIND	
16/1-29_S	actual	
30/8-5	actual	
33/10-12	planned	

```
source query
(in SQL for relational databases)
target query
(atoms in the ontology vocabulary)
```

the iri function constructs IRIs by concatenating any number of strings

#### WELLBORE table

IDENTIFIER	REF_EXISTENCE_KIND	
16/1-29_S	actual	
30/8-5	actual	
33/10-12	planned	

```
result: Wellbore(Wellbore-16/1-29_S)
Wellbore(Wellbore-30/8-5)

↑ IRIs
```

```
a language for expressing
map:m-00001
   a rr:TriplesMap ;
                                                             customised mappings
   rr:logicalTable
                                                             from relational databases
       a rr:R2RMLView :
       rr:sqlQuery " SELECT * FROM WELLBORE WHERE REF_EXISTENCE_KIND = 'actual' "
                                                             to RDF datasets
   rr:predicateObjectMap
                                                             (W3C recommendation 2012)
       a rr:PredicateObjectMap;
       rr:objectMap [
           a rr:ObjectMap, rr:TermMap;
           rr:template "http://slegger.gitlab.io/data#TotalCoreLength-{IDENTIFIER}";
           rr:termType rr:IRI
       1;
       rr:predicate expl:hasTotalCoreLength
   rr:subjectMap
       a rr:SubjectMap, rr:TermMap;
       rr:class expl:Wellbore ;
       rr:template "http://slegger.gitlab.io/data#Wellbore-{IDENTIFIER} ";
       rr:termType rr:IRI
   ] .
```

```
a language for expressing
map:m-00001
   a rr:TriplesMap;
                                                             customised mappings
   rr:logicalTable
                                                             from relational databases
       a rr:R2RMLView :
       rr:sqlQuery " SELECT * FROM WELLBORE WHERE REF_EXISTENCE_KIND = 'actual' "
                                                             to RDF datasets
   rr:predicateObjectMap
                                                             (W3C recommendation 2012)
       a rr:PredicateObjectMap;
       rr:objectMap [
           a rr:ObjectMap, rr:TermMap;
           rr:template "http://slegger.gitlab.io/data#TotalCoreLength-{IDENTIFIER}" ;
           rr:termType rr:IRI
       1;
       rr:predicate expl:hasTotalCoreLength
   rr:subjectMap
       a rr:SubjectMap, rr:TermMap;
       rr:class expl:Wellbore ;
       rr:template "http://slegger.gitlab.io/data#Wellbore-{IDENTIFIER} ";
       rr:termType rr:IRI
```

produces RDF graph

```
:Wellbore-16/1-29_S a expl:Wellbore . :Wellbore-30/8-5 a expl:Wellbore .
```

```
a language for expressing
 map:m-00001
    a rr:TriplesMap;
                                                            customised mappings
    rr:logicalTable
                                                            from relational databases
        a rr:R2RMLView :
        rr:sqlQuery " SELECT * FROM WELLBORE WHERE REF_EXISTENCE_KIND = 'actual' "
                                                            to RDF datasets
    rr:predicateObjectMap
                                                            (W3C recommendation 2012)
        a rr:PredicateObjectMap;
        rr:objectMap [
            a rr:ObjectMap, rr:TermMap;
            rr:template "http://slegger.gitlab.io/data#TotalCoreLength-{IDENTIFIER}" ;
            rr:termType rr:IRI
        1;
        rr:predicate expl:hasTotalCoreLength
    rr:subjectMap
        a rr:SubjectMap, rr:TermMap;
        rr:class expl:Wellbore ;
        rr:template "http://slegger.gitlab.io/data#Wellbore-{IDENTIFIER} ";
        rr:termType rr:IRI
produces RDF graph
                            :Wellbore-16/1-29_S a expl:Wellbore .
                            :Wellbore-30/8-5 a expl:Wellbore .
```

NB: is this a GAV (Global-As-View) mapping?

```
a language for expressing
 map:m-00001
     a rr:TriplesMap;
                                                             customised mappings
     rr:logicalTable
                                                             from relational databases
        a rr:R2RMLView :
        rr:sqlQuery " SELECT * FROM WELLBORE WHERE REF_EXISTENCE_KIND = 'actual' "
                                                             to RDF datasets
     rr:predicateObjectMap
                                                             (W3C recommendation 2012)
        a rr:PredicateObjectMap;
        rr:objectMap [
            a rr:ObjectMap, rr:TermMap;
            rr:template "http://slegger.gitlab.io/data#TotalCoreLength-{IDENTIFIER}" ;
            rr:termType rr:IRI
        1;
        rr:predicate expl:hasTotalCoreLength
     rr:subjectMap
        a rr:SubjectMap, rr:TermMap;
         rr:class expl:Wellbore ;
        rr:template "http://slegger.gitlab.io/data#Wellbore-{IDENTIFIER} ";
        rr:termType rr:IRI
produces RDF graph
                            :Wellbore-16/1-29_S a expl:Wellbore .
```

:Wellbore-30/8-5 a expl:Wellbore .

NB: is this a GAV (Global-As-View) mapping?

not quite — the IRI function can simulate GLAV (more in part 2...)

## References

- G. Xiao, D. Calvanese, R. Kontchakov, D. Lembo, A. Poggi, R. Rosati & M. Za-kharyaschev. "Ontology-Based Data Access: A Survey". In: Proc. of the 27th Int. Joint Conf. on Artificial Intelligence, IJCAI-ECAI 2018, 5511–5519. ijcai.org, 2018.
- 2. D. Calvanese, G. De Giacomo, D. Lembo, M. Lenzerini & R. Rosati. "Tractable reasoning and efficient query answering in description logics: The DL-Lite family." In: JAR, 39(3):385–429, 2007.
- 3. "RDF 1.1 Concepts and Abstract Syntax" R. Cyganiak, D. Wood, M. Lanthaler, editors. W3C, 2014. https://www.w3.org/TR/rdf11-concepts
- 4. "OWL 2 Web Ontology Language Profiles (Second Edition)" B. Motik, B. Cuenca Grau, I. Horrocks, Z. Wu, A. Fokoue, C. Lutz, editors. W3C, 2012. https://www.w3.org/TR/owl2-profiles
- 5. "SPARQL 1.1 Query Language" S. Harris & A. Seaborne, editors. W3C, 2013. https://www.w3.org/TR/sparql11-query
- "R2RML: RDB to RDF Mapping Language" S. Das, S. Sundara, R. Cyganiak, editors. W3C, 2012.
  - https://www.w3.org/TR/r2rml