



**KESW 2015**

INTERNATIONAL CONFERENCE ON KNOWLEDGE  
ENGINEERING AND SEMANTIC WEB  
SEP 30 - OCT 2, 2015, MOSCOW, RUSSIA



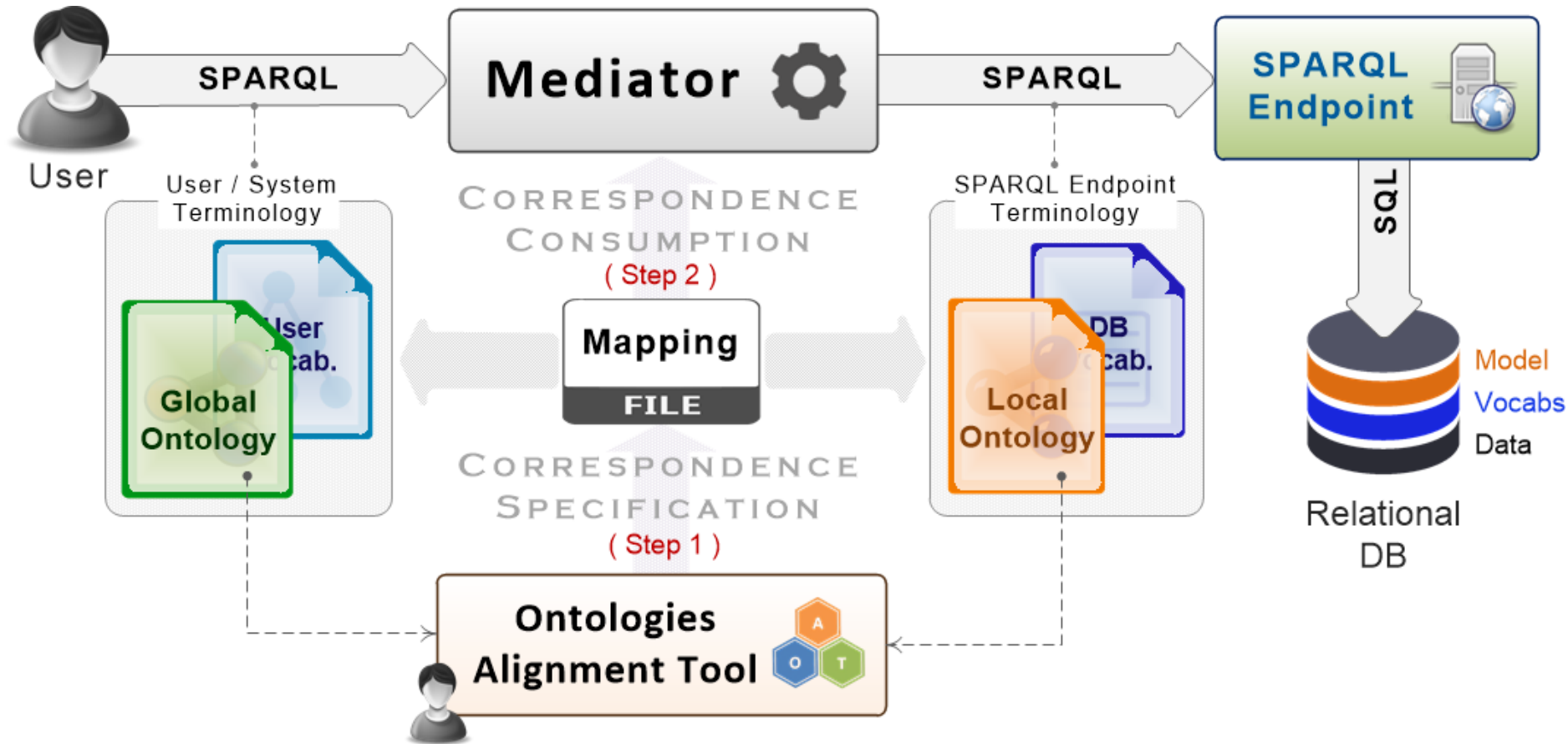
# An Advanced Query and Result Rewriting Mechanism for Information Retrieval Purposes from RDF Datasources

Efthymios Chondrogiannis  
ICCS / NTUA  
Greece

# Introduction

- Relational Databases
  - Widely Used in Software Industry
  - Access via **SQL** queries
- SPARQL Endpoints (e.g., D2R server)
  - Enable users to treat a RDB as a Virtual RDF Graph
  - Access via **SPARQL** queries
    - Close to the data structure (DB schema) and
    - DB vocabularies (i.e., table-field values)
- Semantic Web Applications
  - Need access to existing (real or virtual) RDF datasources
  - Use different **Model** and **Vocabularies**

# Approach Followed



# Correspondence Specification – OAT

The screenshot displays the Ontologies Alignment Tool (OAT) interface. The top bar includes the title "Global EHR Ontology" on the left, the tool name "Ontologies Alignment Tool" in the center, and "HE1-EHR-O" on the right. The main workspace is divided into three panels. The left panel shows a tree view of the "Global EHR Ontology" with categories like "Classes", "ObjProp", "DatProp", and "Indiv". The middle panel, titled "Ontological Elements", shows the definition of the "Observation" class, including its description, a link to the HL7 RIM class, and a list of super classes. The right panel shows a tree view of the "HE1-EHR-O" ontology. An "Element Definition" dialog box is open in the center, displaying the definition of the "Amount" class and its super classes. The dialog box has an "Ok" button.

**Global EHR Ontology**

**Ontologies Alignment Tool**

**HE1-EHR-O**

**Classes** | **ObjProp** | **DatProp** | **Indiv**

**Observation**

An act that is intended to result in new information about a subject. The main difference between Observations and other Acts is that Observations have a value attribute. The code attribute of Observation and the value attribute of Observation must be considered in combination to determine the semantics of the observation.

Definition retrieved from HL7 RIM class Observation

**Super Classes (4)**

- **Thing**
- **Date Performed** **max 1** **dateTime**
- **Examination Code** **exactly 1** **Examination Code**
- **Examination Outcome** **max 1** ( **Amount** or **Ex** )

**Quick Elements Mapping Bar** **C** **:**

This class is being used for capturing all the import Performed.

**Super Classes (6)**

- **Data**
- **Hemat Exam Data Date Recorded** **exactly 1**
- **Hemat Exam Data ID** **exactly 1**
- **Hemat Exam Data Results** **exactly 1** **float**
- **Hematological Examination Data Belongs To** **exactly 1**
- **Hematological Examination HE Code** **exactly 1**

**Element Definition**

An amount is defined by a value followed by a unit or unit expression.

**Super Classes**

- **Complex Data Type**
- **Amount Unit** **exactly 1** **Unit Code**
- **Amount Value** **exactly 1** **float**

**Ok**

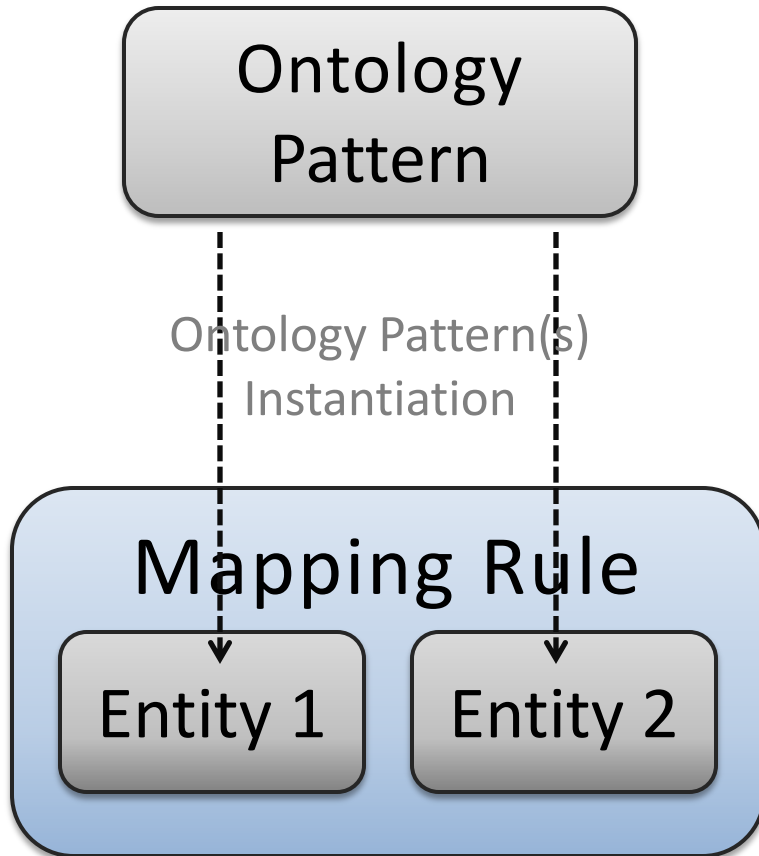
**Classes** | **ObjProp** | **DatProp** | **Indiv**

**HE1-EHR-O**

**Thing**

- **Data**
- **Patient Clinical Characteristics**
- **Diagnosis Data**
- **Drug Data**
- **ECG Data**
- **Hematological Examination Data**
- **Hormone Data**
- **Vocabulary**
- **ATC**
- **Boolean Values**
- **Country Values**
- **ICD-9-CM**
- **Ethnicity Values**
- **Hematological Examination Values**
- **Hormone Values**
- **Income Class Value**
- **Sex Values**
- **Marital Status**

# Ontology Patterns & Mapping Rules



## Ontology Pattern

- Uniquely determines an Entity of a Mapping Rule
- Specifies either an existing Ontological Element or a new one

## Mapping Rule

- Specify the correspondence among the ontology terms along with the changes that should take place in their values (if necessary)

# Ontology Pattern(s) Instantiation – Example

**Entity 1:**

**Relation Domain Restriction**

Relation:

Select Relation Pattern

Domain-Class:

Select Class Pattern

**Entity 2:**

Specify...

**Transformation:**

**Relation:**

Relation of Entity 1 towards Entity 2:

**Direction:**

**Comments:**

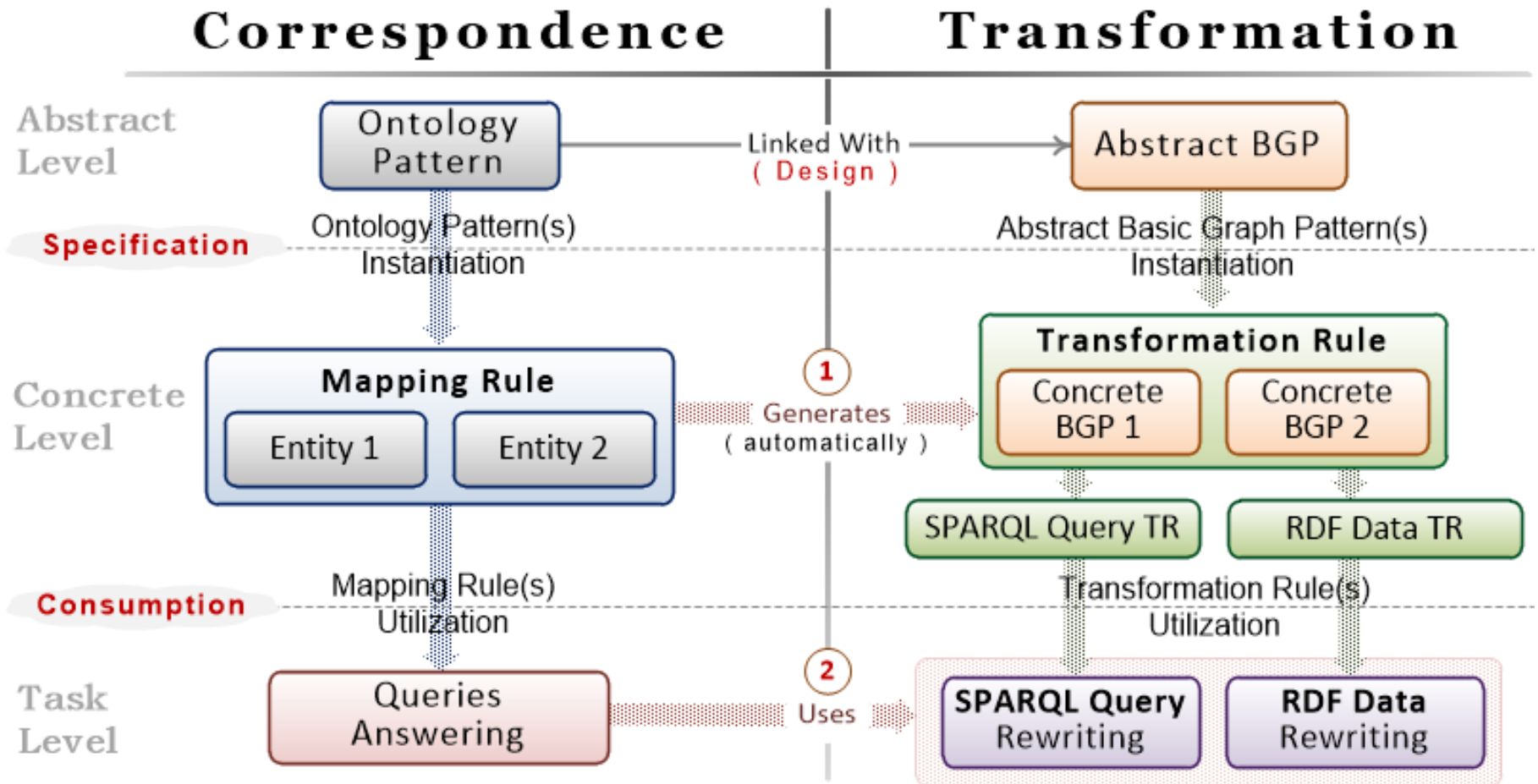
Relation Pattern Selection

- Simple Relation Pattern**
- Inverse Relation Pattern
- Relation Domain Restriction Pattern
- Relation Range Restriction Pattern
- Relation Path Pattern
- Relation Domain-Range Restriction Pattern

# Correspondence Consumption – Mediator

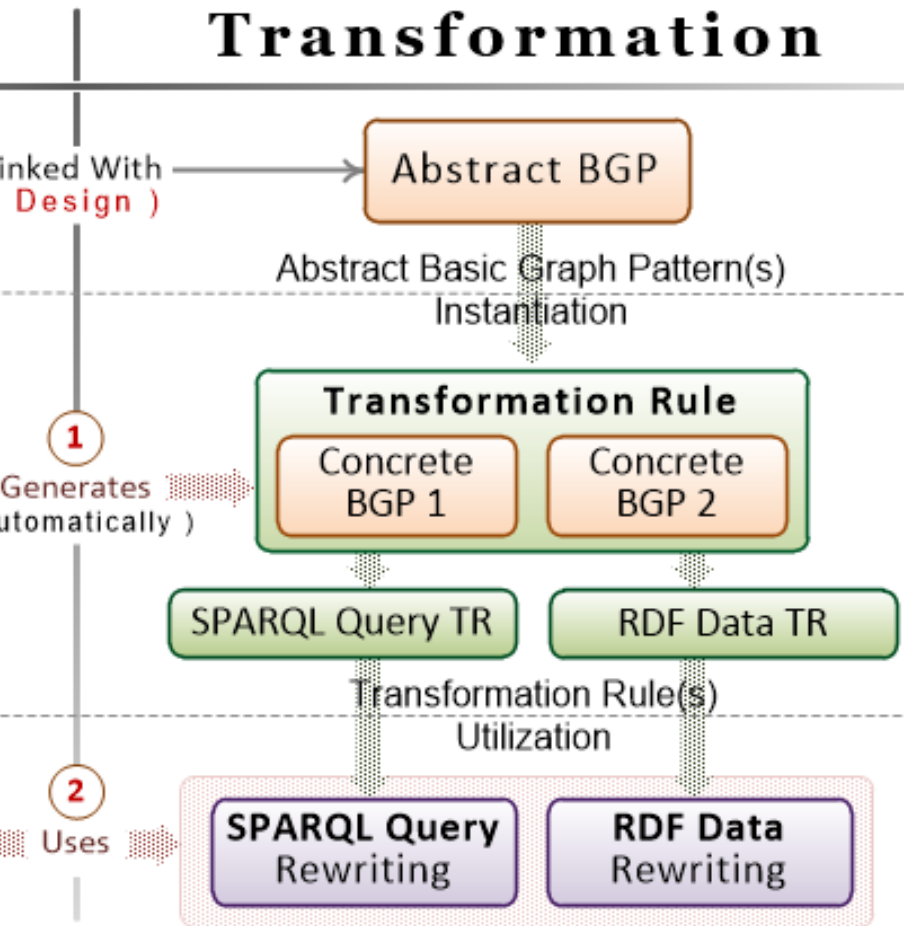
- User
  - Forms the SPARQL query (i.e., SELECT or CONSTRUCT)
  - Terminology: **User Model & Vocabularies**
- **Mediator**
  - Rewrites both SPARQL query and RESULT (i.e., RDF data)
  - Based only on the Mapping Rules specified
- DB SPARQL Endpoint
  - Executes the SPARQL query provided
  - Terminology: **Data Model & DB vocabularies**

# Query & Result Rewriting – Overview





# Abstract BGP & Transformation Rules



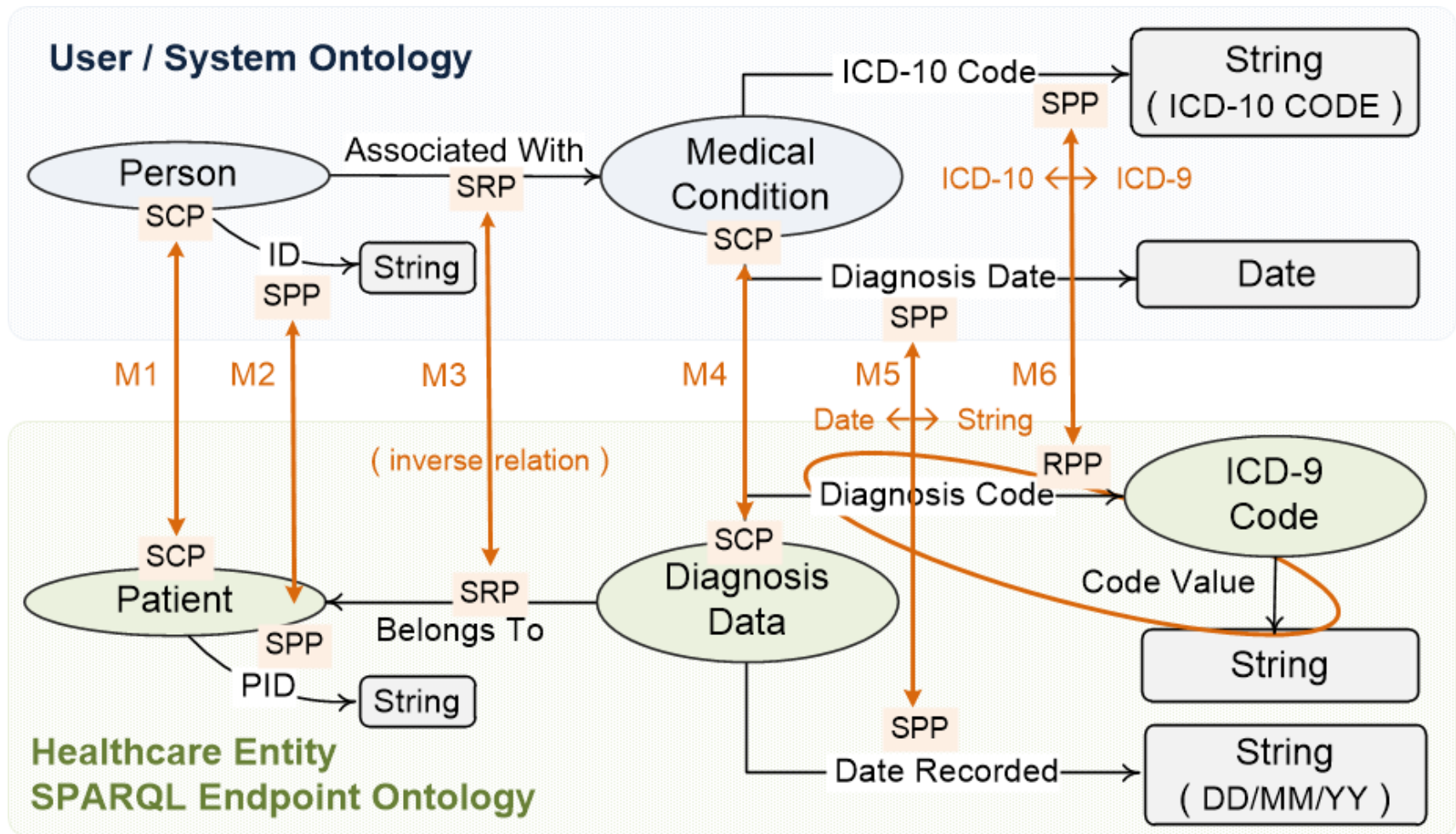
## Abstract BGP

- Provides a triple-based representation of an Ontology Pattern
- Example: for a Simple Property  
[sth1] {property-uri} [sth2]

## Transformation Rule

- Specifies the changes that should take place in a graph-based representation of the corresponding data
- Example:  
[s1] PropA [s2]  $\leftrightarrow$  [s1] PropB [s2]  
[s2] : Data Transformation ?

# Example (1/3) – Mapping Rules



## Example (2/3) – Transformation Rules

ID	MR	TR: Left Side	TR: Right Side
TR1	M1	[sth] rdf:type usr:Person	[sth] rdf:type db:Patient
TR2	M2	[sth1] usr:ID [sth2]	[sth1] db:PID [sth2]
TR3	M3	[sth1] usr:associatedWith [sth2]	[sth2] db:belongsTo [sth1]
TR4	M4	[sth] rdf:type usr:MedicalCondition	[sth] rdf:type db:DiagnosisData
TR5	M5	[sth1] usr:icd10code [sth2]	[sth1] db:diagnosisCode [tmp] . [tmp] db:codeValue [sth2]
		Data Transformation: [sth2]:ICD-10 to/from ICD-9	
TR6	M6	[sth1] usr:diagnosisDate [sth2]	[sth1] user:dateRecorded [sth2]
		Data Transformation: [sth2]:Date to/from String	

# Example (3/3) – Query & Result Rewriting

Query

## USER – SPARQL Query

CONSTRUCT {

?person **rdf:type** user:Person .  
 ?person user:ID ?id .  
 ?person user:associatedWith ?medCond .  
 ?medCond **rdf:type** user:MedicalCondition .  
 ?medCond user:diagnosisDate ?diagDate .

} WHERE {

?person **rdf:type** user:Person .  
 ?person user:ID ?id .  
 ?person user:associatedWith ?medCond .  
 ?medCond **rdf:type** user:MedicalCondition .  
 ?medCond user:diagnosisDate ?diagDate .  
 ?medCond user:icd10code ?medCondValue .

**FILTER** ( ?medCondValue = "I21"

}

## SPARQL Endpoint – SPARQL Query

CONSTRUCT {

?person **rdf:type** db:Patient .  
 ?person db:PID ?id .  
 ?medCond db:belongsTo ?person .  
 ?medCond **rdf:type** db:DiagnosisData .  
 ?medCond db:dateRecorded ?diagDate .

} WHERE {

?person **rdf:type** db:Patient .  
 ?person db:PID ?id .  
 ?medCond db:belongsTo ?person .  
 ?medCond **rdf:type** :DiagnosisData .  
 ?medCond db:dateRecorded ?diagDate .  
 ?medCond db:diagnosisCode ?tmp .  
 ?tmp db:codeValue ?medCondValue .

**FILTER** ( ?medCondValue = "410" ) }

Results

## USER – RDF Data

db:Patient1 **rdf:type** user:Person .  
 db:Patient1 user:ID "P14312"^^xsd:string .  
 db:Patient1 user:acossiatedWith db:PDdiagData1 .  
 db:PDdiagData1 **rdf:type** user:MedicalCondition .  
 db:PDdiagData1 user:dateDiagnosed "2014-02-11"^^xsd:date .

## SPARQL Endpoint – RDF Data

db:Patient1 **rdf:type** db:Patient .  
 db:Patient1 db:PID "P14312"^^xsd:string .  
 db:PDdiagData1 db:belongTo :Patient1 .  
 db:PDdiagData1 **rdf:type** db:DiagnosisData .  
 db:PDdiagData1 db:dateRecorded "11/02/14"^^xsd:string .

# Discussion

- Ontology Alignment and Query / Result Rewriting
  - Global Ontology may be **partially linked** with Local Ontology
  - Information Loss during both Query and Results Rewriting
    - Explicitly mention to the end users any modification (either change or omission) made during both SPARQL query and Result rewriting processes.
- Semantics
  - On condition that there is **a)** 1:1 among vocabularies and **b)** all the Global Ontology terms have been mapped to the corresponding ones, the **semantics** of the rewritten SPARQL query (and hence data retrieved) do not change !
- Mapping Issues: Semantically Related Terms
  - e.g. age (including date recorded) and date/year of birth

# Conclusion

- A novel approach for Query Answering purposes
  - Enable users to access a Relation DB or RDF graph using their own Models and Vocabularies
- Users specify correspondence through a semi-automatic process based on the OAT (step 1)
  - Enable users to instantiate one or more Ontology Patterns for specifying the entities of a Mapping Rule
- Mediator rewrites SPARQL queries formed and data retrieved to the appropriate ones (step 2)
  - Automatically produces the corresponding Transformation Rules for both SPARQL query and RDF data rewriting purposes



# Thank You !

