DartGrid

Vijetha Shivarudraiah
Sai Phalgun Tatavarthy

CSc 8711

Georgia State University

Semantic Web

- Focused on machines
- *"a web talking to machines"*

The Grid

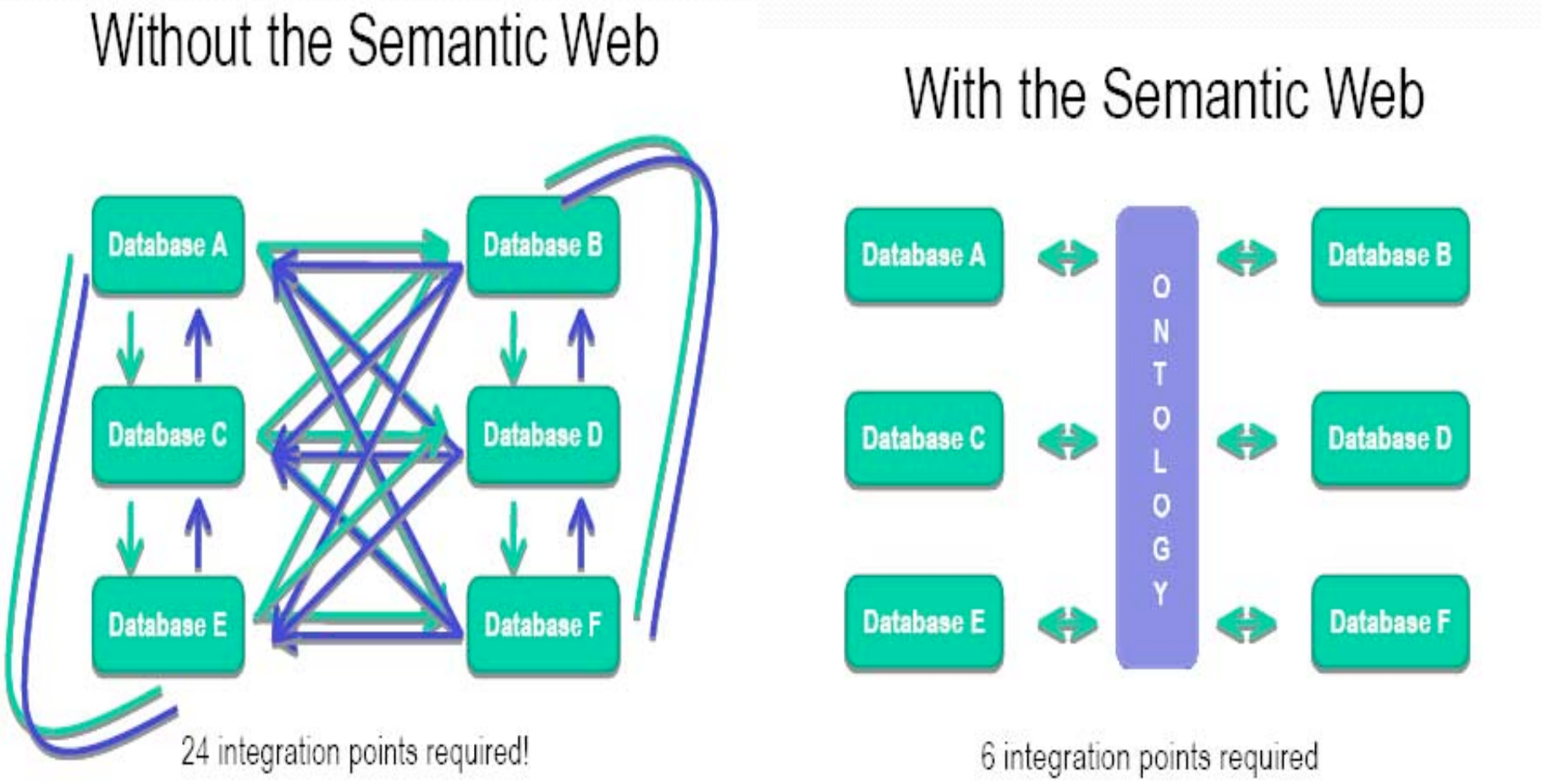
- Super virtual computer
- Many networked loosely coupled computers
- Work together to perform humongous tasks

Semantic Grid

- Data heterogeneity – semantic web technologies
- Challenge – to design **framework** to **collaborate** data from various sources

- Data in big organizations – stored in **relational databases** - heterogeneous
- support **integration** of **heterogeneous** relational databases using semantic web technologies

How Ontology helps?



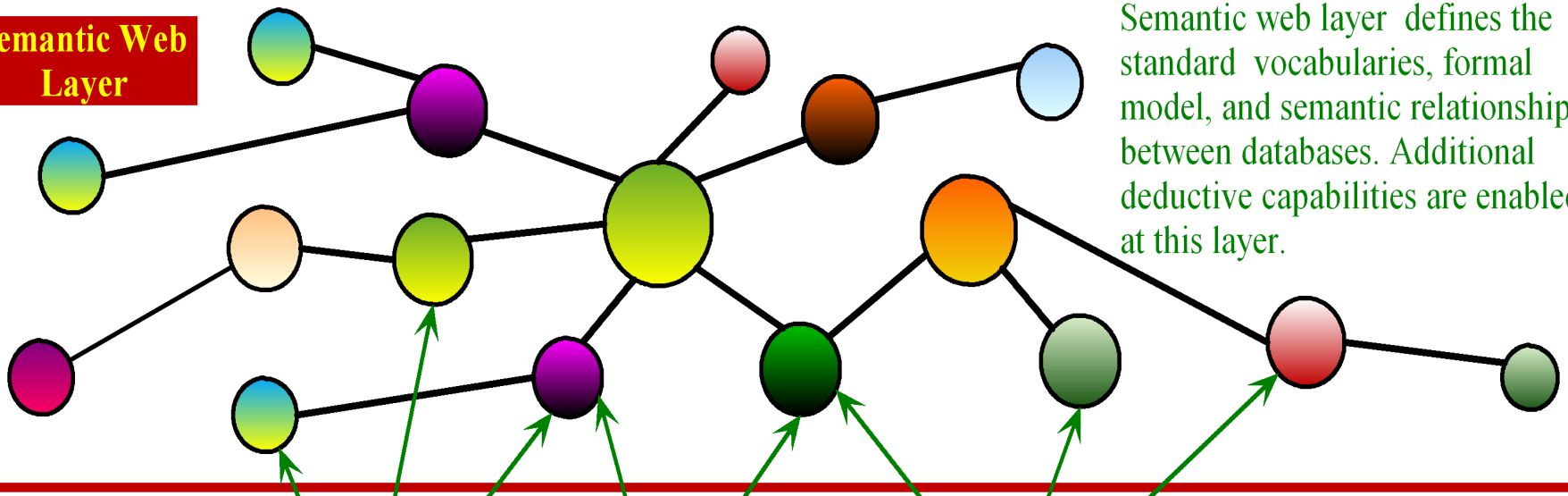
Towards a Semantic Web of RD

User Layer

Semantic Browse, Semantic Query, Semantic Search, Semantic Navigation, Semantic Programming

Persons or programs only interact with the semantic web layer, and query is defined over the shaired ontologies.

Semantic Web Layer



Semantic web layer defines the standard vocabularies, formal model, and semantic relationships between databases. Additional deductive capabilities are enabled at this layer.

Relational Databases Layer

	A	B	C
	1	...	2
	...	...	...
	...		

	D	E	F
	...		
	...	...	...

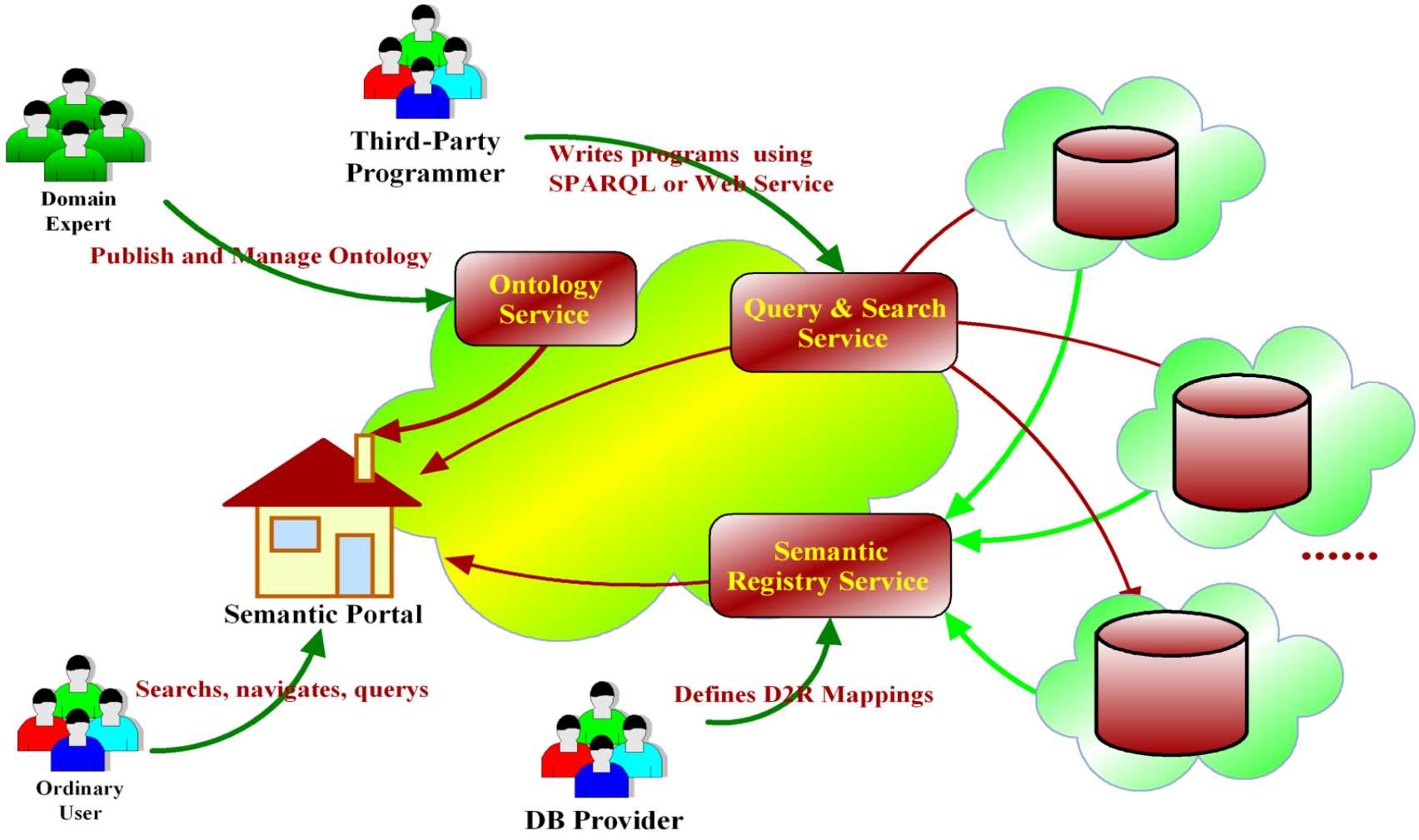
	G	H	I
	...		
		...	...

Relational databases are mapped to the semantic web layer, and new databases should be added at anytime by anyone, and should be independent of the semantic web layer.

DartGrid

- implementation of Semantic Grid
- application development **framework**
- together with a set of practical semantic **tools**
- to facilitate **integration** of heterogeneous relational databases
- leverages upon technologies from both the Semantic web and the Grid

System Architecture



Key Components in DartGrid

- **Ontology Service**
  - expose the shared ontologies
- **Semantic Registration Service**
  - maintains the semantic mapping information
- **Semantic Query Service**
  - to process SPARQL semantic queries
- **Search Service**
  - supports full-text search in all databases

Semantic Tools

- DartMapping
  - Visualized mapping tool
  - heterogeneous relational schemas $\xrightarrow[\text{Mapping}]{\text{Semantic}}$ RDF/OWL ontologies.
- DartQuery
  - Ontology based query interface
  - SPARQL semantic queries $\xrightarrow{\text{rewrite}}$ SQL queries
- DartSearch
  - ontology-based search engine
  - to make full-text search **over all databases**

Semantic Mapping

- two legacy relational databases
  - W3C and ZJU (*Zhejiang University*)
  - about their employees and projects
- integrate them by the FOAF ontology

Target Scheme: foaf Ontology

- Person
- Project
- Organization

Source Relational Schemes

W3C source: w3c:emp(?en,?em,?pn,?ph,?fon)

ZJU source: zju: emp(?en,?em).
zju:emp_pro(?en,?pn)
zju:pro_org(?pn,?fon)
zju:org(?fon,?foh)

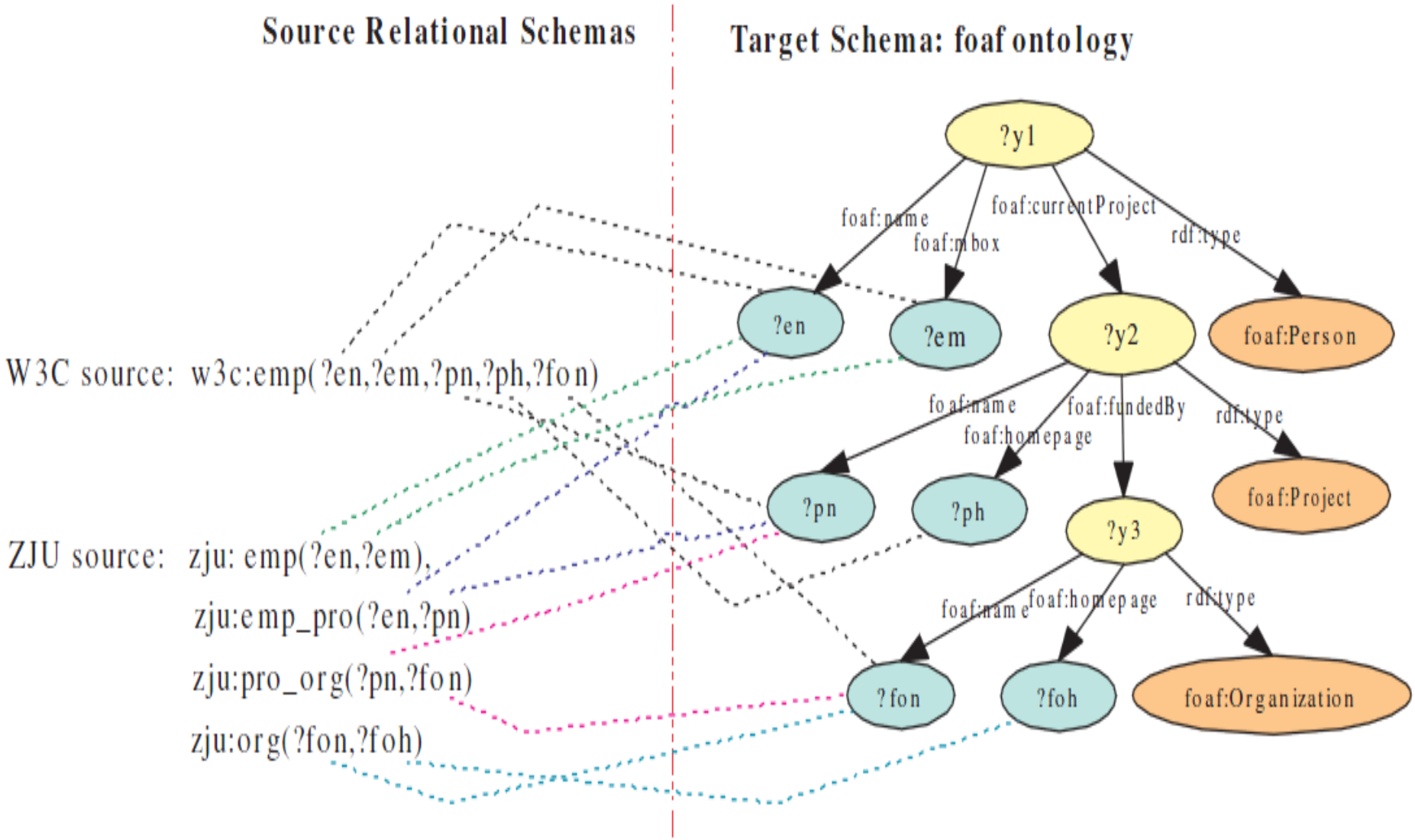
- emp (empName, empMail, projectName, projectHomePage, fundingOrganization)

- emp (empName, empMail)
- emp_org (empName , projectName)
- pro_org (projectName, fundingOrganization)
- org (fundingOrganization, fundingOrgHomePage)

RDF Views

- Define each relational table in the source as a view over the RDF ontologies.
- Such views are called RDF Views – 2 parts
- left part -- view head -- a relational predicate.
- right part -- view body -- a set of RDF triples

```
W3C Source:
V1:  w3c:emp(?en,?em,?pn,?ph,?fon)  :-
          (?y1, rdf:type, foaf:Person),
          (?y1, foaf:name, ?en),
          (?y1, foaf:mobx, ?em),
          (?y1, foaf:currentProject, ?y2),
          (?y2, rdf:type,foaf:Project),
          (?y2, foaf:name,?pn),
          (?y2,foaf:homepage,?ph),
          (?y2,foaf:fundedBy, ?y3),
          (?y3, rdf:type,foaf:Organization),
          (?y3, foaf:name, ?fon).
```


Source Relational Schemas
Target Schema: foaf ontology
W3C source: w3c:emp(?en,?em,?pn,?ph,?fon)
ZJU source: zju: emp(?en,?em),
zju:emp_pro(?en,?pn)
zju:pro_org(?pn,?fon)
zju:org(?fon,?foh)
?y1
foaf:name
foaf:mbox
foaf:currentProject
rdf:type
?en
?em
?y2
foaf:Person
foaf:name
foaf:homepage
foaf:fundedBy
rdf:type
?pn
?ph
?y3
foaf:Project
foaf:name
foaf:homepage
rdf:type
?fon
?foh
foaf:Organization

Mapping Example

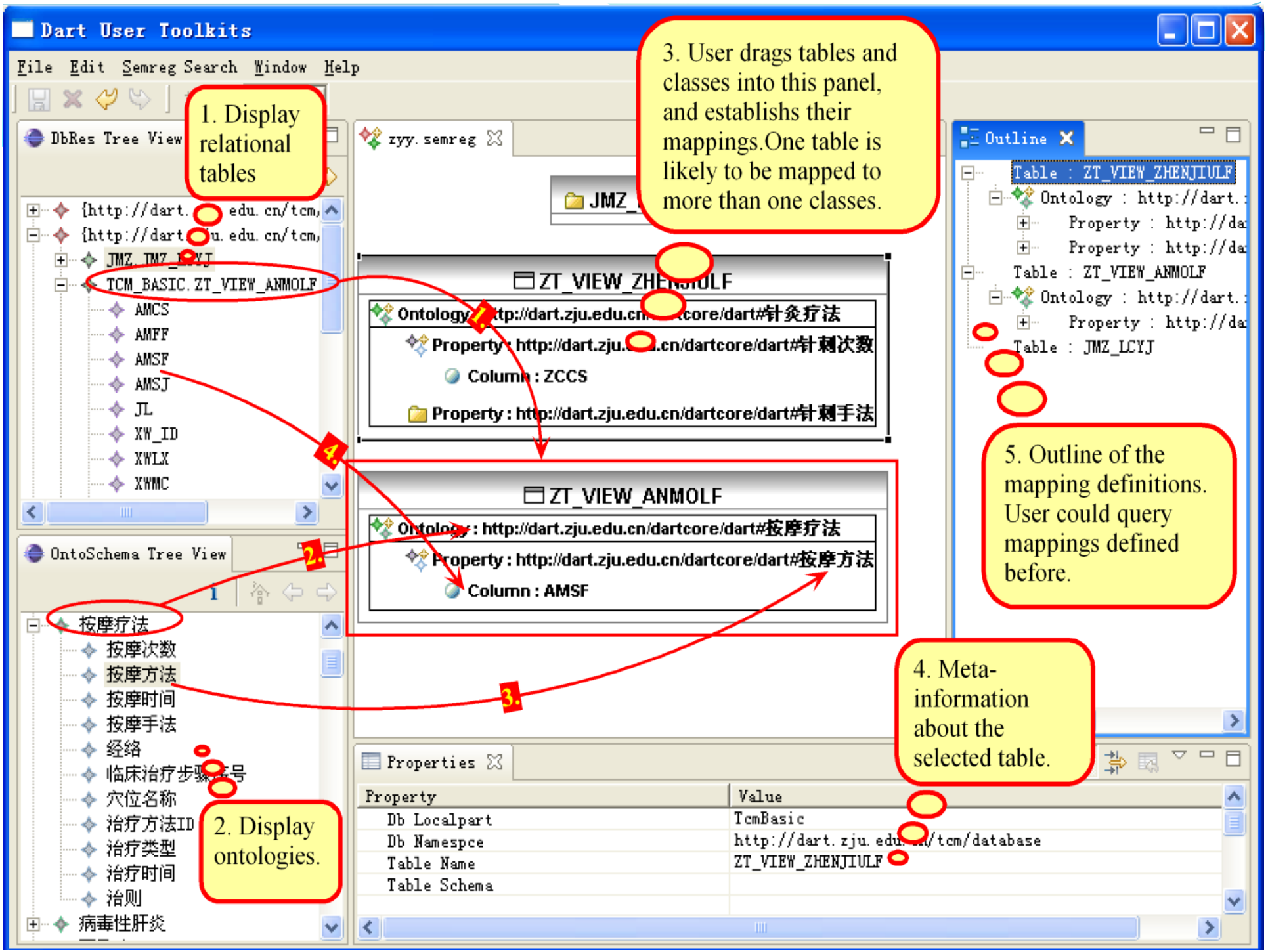
```
Relational Tuple:
---------------------
 w3c:emp("DanBrickley","danbri@w3.org",
         "SWAD","http://swad.org","EU");

Yielded RDF triples by Applying V1:
-----------------------------------
    _:bn1 rdf:type  foaf:Person;
        foaf:name "Dan Brickley";
        foaf:mbox "danbri@w3.org";
        foaf:currentProject _:bn2.
    _:bn2 rdf:type  foaf:Project;
        foaf:name "SWAD";
        foaf:homepage "http://swad.org";
        foaf:fundedBy _:bn3.
    _:bn3 rdf:type foaf:Organization;
        foaf:name "EU".
```

DartMapping

- *DBRes panel displays the relational schemas*
- *OntoSchem* panel displays the shared ontology
- *Mapping Panel visually displays the* mappings from relational schemas to ontologies

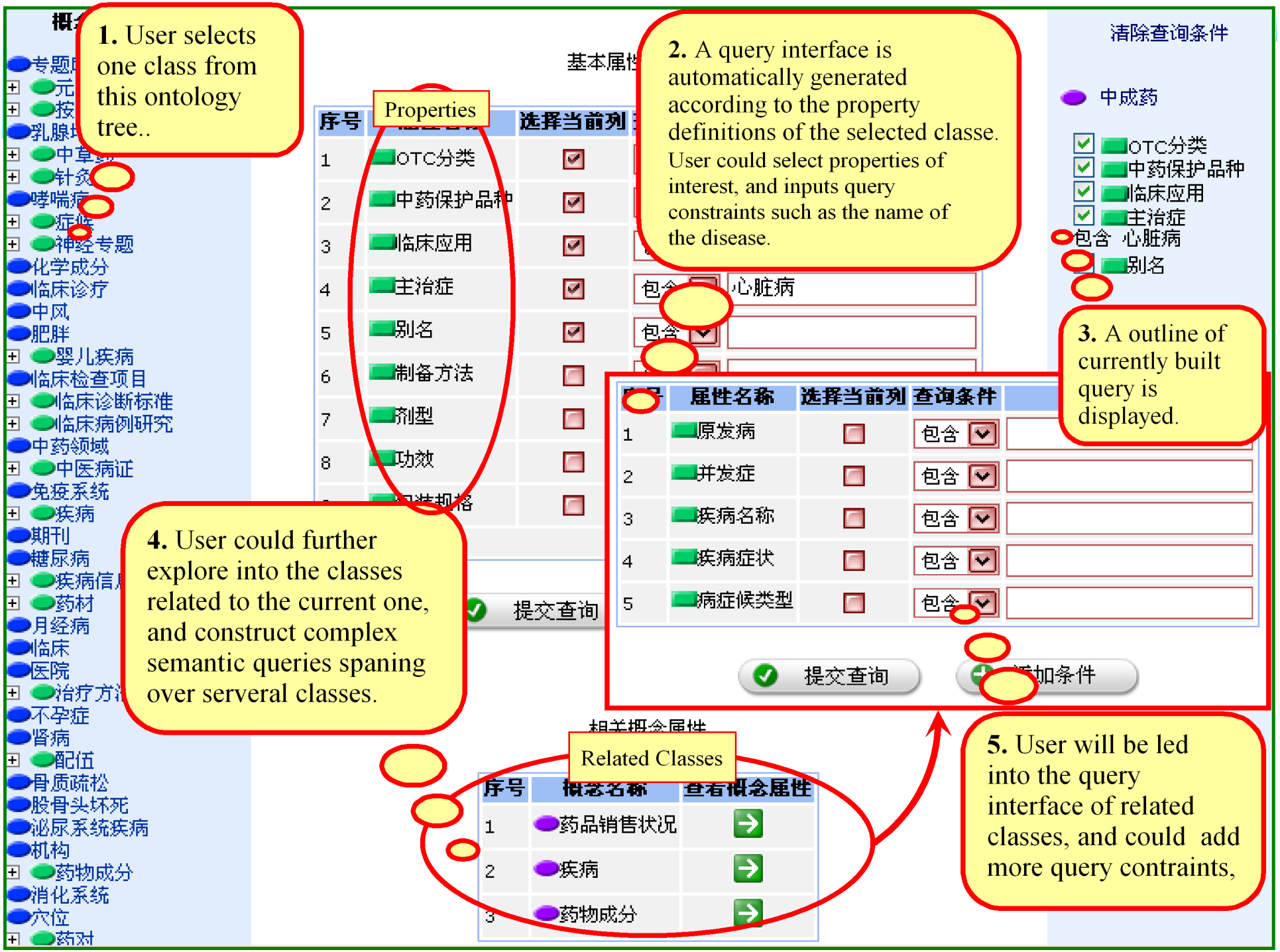
- Generates registration entry – submit to Semantic Registration Service
- *Outline panel to browse and query* mapping information


Dart User Toolkits
File Edit Semreg Search Window Help
1. Display relational tables
DbRes Tree View
TCM_BASIC.ZT_VIEW_ANMOLF
AMCS
AMFF
AMSF
AMSJ
JL
XW_ID
XWLX
XWMC
zyy.semreg
3. User drags tables and classes into this panel, and establishs their mappings.One table is likely to be mapped to more than one classes.
ZT_VIEW_ZHENJIULF
Ontology : http://dart.zju.edu.cn/dartcore/dart#针灸疗法
Property : http://dart.zju.edu.cn/dartcore/dart#针刺次数
Column : ZCCS
Property : http://dart.zju.edu.cn/dartcore/dart#针刺手法
ZT_VIEW_ANMOLF
Ontology : http://dart.zju.edu.cn/dartcore/dart#按摩疗法
Property : http://dart.zju.edu.cn/dartcore/dart#按摩方法
Column : AMSF
1.
2.
3.
4.
OntoSchema Tree View
按摩疗法
按摩次数
按摩方法
按摩时间
按摩手法
经络
临床治疗步骤序号
穴位名称
治疗方法ID
治疗类型
治疗时间
治则
病毒性肝炎
2. Display ontologies.
Outline
Table : ZT_VIEW_ZHENJIULF
Ontology : http://dart.
Property : http://da
Property : http://da
Table : ZT_VIEW_ANMOLF
Ontology : http://dart.
Property : http://da
Table : JMZ_LCYJ
5. Outline of the mapping definitions. User could query mappings defined before.
4. Meta-information about the selected table.
Properties
Property
Value
Db Localpart
TcmBasic
Db Namespce
http://dart.zju.edu.cn/tcm/database
Table Name
ZT_VIEW_ZHENJIULF
Table Schema

DartQuery

- Ontology based **semantic query user interface**
- Browse ontology tree – select classes
- Query form of properties of the classes generated
- User can select the properties and input constraints

- Semantic query generated – submitted to Semantic Query Service
- Query rewritten into set of SQL queries – using mapping views – Semantic Registration Service


1. User selects one class from this ontology tree..
Properties
2. A query interface is automatically generated according to the property definitions of the selected classe. User could select properties of interest, and inputs query constraints such as the name of the disease.
3. A outline of currently built query is displayed.
4. User could further explore into the classes related to the current one, and construct complex semantic queries spaning over serveral classes.
Related Classes
5. User will be led into the query interface of related classes, and could add more query contraints,
清除查询条件
中成药
OTC分类
中药保护品种
临床应用
主治症
包含 心脏病
别名
基本属性
序号
选择当前列
1 OTC分类
2 中药保护品种
3 临床应用
4 主治症
5 别名
6 制备方法
7 剂型
8 功效
包含
心脏病
提交查询
属性名称
查询条件
1 原发病
2 并发症
3 疾病名称
4 疾病症状
5 病症候类型
包含
提交查询
相关概念属性
序号
概念名称
查看概念属性
1 药品销售状况
2 疾病
3 药物成分
专题
元
按
乳腺
中草
针灸
哮喘病
症候
神经专题
化学成分
临床诊疗
中风
肥胖
婴儿疾病
临床检查项目
临床诊断标准
临床病例研究
中药领域
中医病证
免疫系统
疾病
期刊
糖尿病
疾病信
药材
月经病
临床
医院
治疗方
不孕症
肾病
配伍
骨质疏松
股骨头坏死
泌尿系统疾病
机构
药物成分
消化系统
穴位
药对

Results Interface

- User can navigate through all the related entries by following the semantic links
- The relations between search results and those discovered by the semantic links are derived from the semantic layer

Results Interface

中医药数据库网格查询结果

共查到11028条记录　共1103页　当前第1页 每页显示10条

以分栏方式显示结果

下一页 最后一页

查询条件基本信息

清除查询条件

Semantic Links

- 临床研究对象
- 对照组
- 疾病
- 治疗方法
- 临床诊疗
- 临床检查项目

1. User selects one data entry which will be highlighted.

2. By following these links, user could get all those data objects semantically related to the current one.

3. Note: the relations between the current object and those "discovered" by following the semantic link are derived through the semantic web layer.

4. User could keep navigating through an unending set of databases as long as they are semantically connected .

序号	临		病例采集时间	传染途径	年龄组			女性观察例数	临床研究名称
1	武汉大学中南医院中西医结合科 430071	治疗组 $对照组	1995/01-2000/01		成年人（1 $中年人（ $老年人（65-79）	专题			中西医结合治疗癌性胸腔积液的临床体会
2	山东 沂南县人民医院 276300		1999/09-2000/06		新生儿（<1个月）$婴儿（1-23个月）$儿童，学龄前（2-5）$儿童（6-12）	呼吸专题			
3	唐山市人民医院儿科 063001	中西医结合组 $西药组	1988/4-1991/4		婴儿（1-23个月）$儿童，学龄前（2-5）$儿童（6-12）	呼吸专题			
4	石家庄铁路医院传染科 050000	治疗组 $对照组			青年人（13-18）$成年人（19-44）$中年人（45-64）$老年人（65-79）	呼吸专题			
5	湖南省中医药研究院附属医院		1990/01-1996/12		成年人（19-44）$中年人（45-64）$老年人（65-79）	呼吸专题			
	湖南省中医药研				成年人（19-44）	呼吸			

DartSearch

- Google-like search interface
- accepts one or more keywords
- makes a complete full-text search in all databases

- Navigate the search results by following the semantic links – as in Query Interface
- Ranking – based on relevance of keywords
- Links lead to the semantic query interface – get more accurate results


1. User input a keyword, and trigger a full-text search over all databases.
4. The search system could generate a suggestive list of concepts ranked according to their relevance to the keywords.
5. users could further explore into the query interface of those concepts, and specify a semantic query to get more accurate and appropriate information
Concept Ranking
Sematnic Links
2. Being similar to query, by following these links, user could get all those data objects semantically related to the current one.
3. User could keep navigating through an unending set of databases as long as they are semantically connected .
提交查询

About DartGrid

- developed by Zhejiang University of China
- toolkit was first introduced in 2004
- used to build VO for Traditional Chinese Medicine (TCM)
- 70 legacy TCM databases by a formal TCM ontology with over 70 classes and 800 properties

Pros

- It greatly facilitate developers to interconnect distributed located legacy databases using richer semantics,
- To provide ontology-based query, search and navigation services as one huge distributed database,
- To add additional deductive capabilities on the top to increase the usability and reusability of data

Unsolved Issues on Mapping

1) Redundancy among different database schemas,
2) Inconsistence between two database schemas,
3) Alternative ways to map n-ary (n>2) *relation into RDF/OWL model*.

Thank You!