基于知识图谱的反问推演

刘荣

北京中科汇联信息技术有限公司

liur@huilan.com
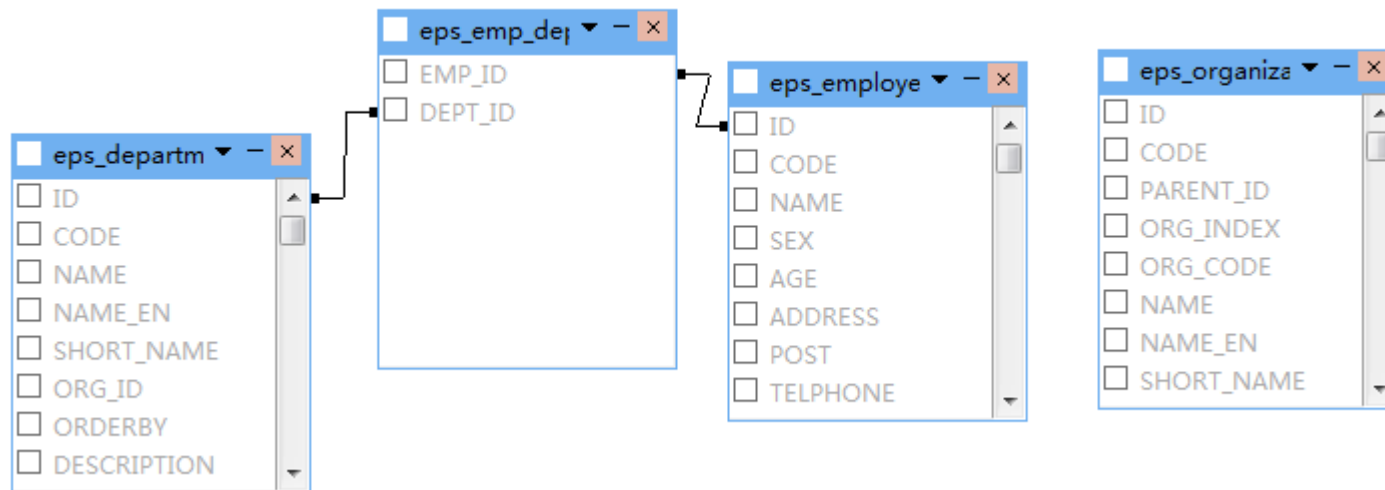
中科汇联
Huilan Infotech

一切为了用户体验！

目录

知识图谱构建

用户数据库DB   用户无数据库

D2R tool

Mapping 文件   知识图谱编辑工具

领域知识/人工修改

中间映射图(RDF)   D2R dump   RDF数据/虚拟RDF

数据库
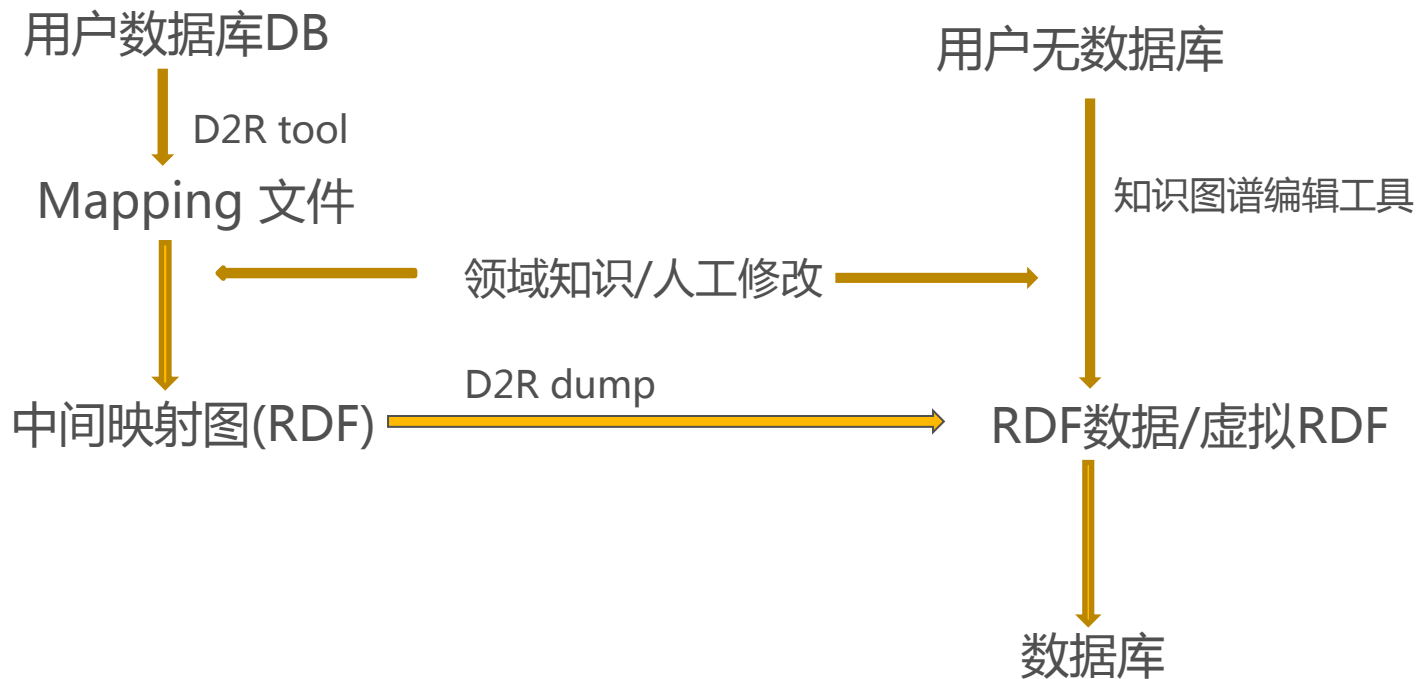
知识图谱构建—DB转RDF

➤ 原始数据库

1. 有四张表，包含一张关系表
2. Department与employee建立关系，通过关系表，organization没有与Employee建立关系

知识图谱构建—DB转RDF

➢ RDF数据

● D2R

1. 生成映射文件
2. 将虚拟RDF转成物理RDF
3. 可以进行可视化浏览（网页）
4. 与Jena Api 结合，进行查询

Tools
D2R Server
generate-mapping
d2r-query
dump-rdf
D2RQ + Jena API

employee

Property	Value
vocab:eps_emp_dept	\<http://localhost:2020/resource/eps_department/1038032\>
vocab:eps_employee_APP_ID	1
vocab:eps_employee_AVAILABLE	0
vocab:eps_employee_CODE	001
vocab:eps_employee_CREATE_DATE	2014-02-24 15:34:27
vocab:eps_employee_CREATE_USER_ID	1
vocab:eps_employee_DEPT_ID	1038032
vocab:eps_employee_EMAIL	zhangsan@00.com
vocab:eps_employee_FLAG	1
vocab:eps_employee_ID	1038035
vocab:eps_employee_MARITAL_STATUS	1
vocab:eps_employee_NAME	张三
vocab:eps_employee_ORG_ID	1038030
vocab:eps_employee_SEX	1
vocab:eps_employee_TYPE	official
rdfs:label	eps_employee #1038035
rdf:type	vocab:eps_employee

employee

organization

Property	Value
vocab:eps_organization_APP_ID	1
vocab:eps_organization_CODE	test
vocab:eps_organization_CREATE_DATE	2014-05-30 10:34:12
vocab:eps_organization_CREATE_USER_ID	1
vocab:eps_organization_DESCRIPTION	测试机构
vocab:eps_organization_FLAG	1
vocab:eps_organization_ID	2120
vocab:eps_organization_LEAF	1
vocab:eps_organization_NAME	测试机构
vocab:eps_organization_NAME_EN	test
vocab:eps_organization_ORDERBY	5 (xsd:decimal)
vocab:eps_organization_ORG_CODE	test
vocab:eps_organization_ORG_INDEX	-1,2120
vocab:eps_organization_ORG_LEVEL	1 (xsd:decimal)
vocab:eps_organization_PARENT_ID	-1
vocab:eps_organization_SHORT_NAME	测试机构
rdfs:label	eps_organization #2120
rdf:type	vocab:eps_organization

organization

department

Property	Value
vocab:eps_department_APP_ID	1
vocab:eps_department_CODE	SXXGKC
vocab:eps_department_CREATE_DATE	2014-02-24 15:16:36
vocab:eps_department_CREATE_USER_ID	1
vocab:eps_department_DEPT_INDEX	-1,1038032
vocab:eps_department_DEPT_LEVEL	1 (xsd:decimal)
vocab:eps_department_FLAG	1
vocab:eps_department_ID	1038032
vocab:eps_department_LEAF	1
vocab:eps_department_NAME	省信息公开处
vocab:eps_department_NAME_EN	由32个以内的英文字母组成
vocab:eps_department_ORDERBY	1 (xsd:decimal)
vocab:eps_department_ORG_ID	1038030
vocab:eps_department_PARENT_ID	-1
is vocab:eps_emp_dept of	\<http://localhost:2020/resource/eps_employee/1038035\>
is vocab:eps_emp_dept of	\<http://localhost:2020/resource/eps_employee/1038036\>
is vocab:eps_emp_dept of	\<http://localhost:2020/resource/eps_employee/1038037\>
is vocab:eps_emp_dept of	\<http://localhost:2020/resource/eps_employee/1038050\>
rdfs:label	eps_department #1038032
rdf:type	vocab:eps_department

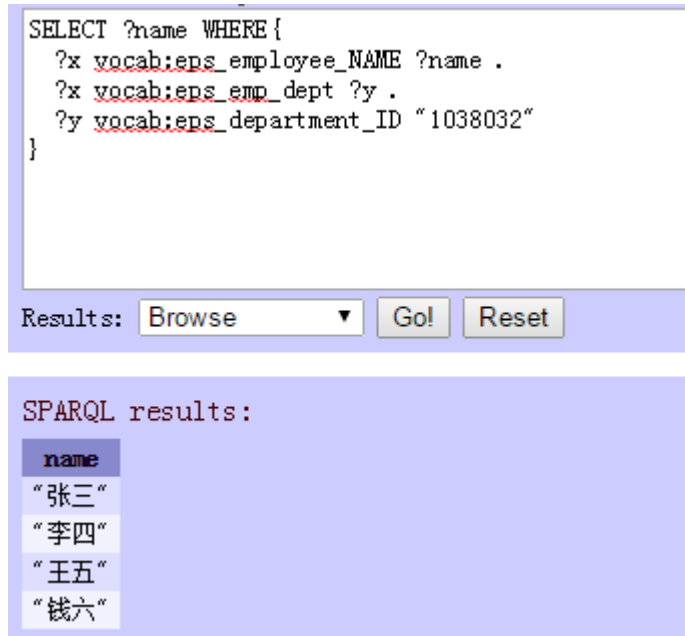**department**

中科汇联 Huilan Infotech 一切为了用户体验！

知识图谱构建—DB转RDF

➢ RDF数据

● Sparql查询

```
SELECT ?name WHERE{
  ?x vocab:eps_employee_NAME ?name .
  ?x vocab:eps_emp_dept ?y .
  ?y vocab:eps_department_ID "1038032"
}
```



```
SELECT ?name WHERE{
  ?x vocab:eps_employee_NAME ?name .
  ?x vocab:eps_emp_dept ?y .
  ?y vocab:eps_department_ID "1038032"
}
```

Results: Browse ▼ Go! Reset

SPARQL results:

name
"张三"
"李四"
"王五"
"钱六"

中科汇联 Huilan Infotech 一切为了用户体验!

知识图谱构建—DB转RDF

➤ RDF数据修改

● 添加关系

```
#n:m relationship of employee and organization
map:eps_emp_org a d2rq:PropertyBridge;
        d2rq:belongsToClassMap map:eps_employee;
        d2rq:property vocab:eps_emp_org;
        d2rq:refersToClassMap map:eps_organization;
        d2rq:join "eps_employee.ORG_ID => eps_organization.ID";
```

● 添加实体，并添加相关属性
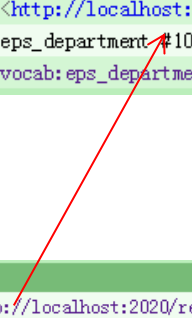
```
# Table eps_job
map:eps_job a d2rq:ClassMap;
        d2rq:dataStorage map:database;
        d2rq:uriPattern "eps_job";
        d2rq:class vocab:eps_job;
        d2rq:classDefinitionLabel "eps_job";
map:eps_job_name a d2rq:PropertyBridge;
    d2rq:belongsToClassMap map:eps_job;
    d2rq:property vocab:name;
    d2rq:constantValue "job";
    .
```

添加job类，并添加name属性

中科汇联 Huilan Infotech | 一切为了用户体验!

Table 1 (left)

Property	Value
is vocab:eps_emp_org of	\<http://localhost:2020/resource/eps_employee/1038035>
is vocab:eps_emp_org of	\<http://localhost:2020/resource/eps_employee/1038036>
is vocab:eps_emp_org of	\<http://localhost:2020/resource/eps_employee/1038037>
is vocab:eps_emp_org of	\<http://localhost:2020/resource/eps_employee/1038038>
is vocab:eps_emp_org of	\<http://localhost:2020/resource/eps_employee/1038039>
is vocab:eps_emp_org of	\<http://localhost:2020/resource/eps_employee/1038040>
is vocab:eps_emp_org of	\<http://localhost:2020/resource/eps_employee/1038050>
is vocab:eps_employee_ORG_ID of	\<http://localhost:2020/resource/eps_employee/1038035>
is vocab:eps_employee_ORG_ID of	\<http://localhost:2020/resource/eps_employee/1038036>
is vocab:eps_employee_ORG_ID of	\<http://localhost:2020/resource/eps_employee/1038037>
is vocab:eps_employee_ORG_ID of	\<http://localhost:2020/resource/eps_employee/1038038>
is vocab:eps_employee_ORG_ID of	\<http://localhost:2020/resource/eps_employee/1038039>
is vocab:eps_employee_ORG_ID of	\<http://localhost:2020/resource/eps_employee/1038040>
is vocab:eps_employee_ORG_ID of	\<http://localhost:2020/resource/eps_employee/1038050>
vocab:eps_organization_APP_ID	1
vocab:eps_organization_CODE	SZWZZ
vocab:eps_organization_CREATE_DATE	2014-02-24 15:04:19
vocab:eps_organization_CREATE_USER_ID	1
vocab:eps_organization_FLAG	1
vocab:eps_organization_ID	1038030
vocab:eps_organization_LEAF	1
vocab:eps_organization_NAME	省政务中心
vocab:eps_organization_ORDERBY	3 (xsd:decimal)
vocab:eps_organization_ORG_CODE	SZWZZ
vocab:eps_organization_ORG_INDEX	-1,1038030
vocab:eps_organization_ORG_LEVEL	1 (xsd:decimal)
vocab:eps_organization_PARENT_ID	-1
rdfs:label	eps_organization #1038030
rdf:type	vocab:eps_organization

Table 2 (top right)

Property	Value
vocab:eps_department_APP_ID	1
vocab:eps_department_CODE	SXXGKC
vocab:eps_department_CREATE_DATE	2014-02-24 15:16:36
vocab:eps_department_CREATE_USER_ID	1
vocab:eps_department_DEPT_INDEX	-1,1038032
vocab:eps_department_DEPT_LEVEL	1 (xsd:decimal)
vocab:eps_department_FLAG	1
vocab:eps_department_ID	1038032
vocab:eps_department_LEAF	1
vocab:eps_department_NAME	省信息公开处
vocab:eps_department_NAME_EN	由32个以内的英文字母组成
vocab:eps_department_ORDERBY	1 (xsd:decimal)
vocab:eps_department_ORG_ID	1038030
vocab:eps_department_PARENT_ID	-1
is vocab:eps_emp_dept of	\<http://localhost:2020/resource/eps_employee/1
is vocab:eps_emp_dept of	\<http://localhost:2020/resource/eps_employee/1
is vocab:eps_emp_dept of	\<http://localhost:2020/resource/eps_employee/1
is vocab:eps_emp_dept of	\<http://localhost:2020/resource/eps_employee/1
rdfs:label	eps_department #1038032
rdf:type	vocab:eps_department

Table 3 (bottom right)

Property	Value
vocab:eps_emp_dept	\<http://localhost:2020/resource/eps_department/1038032>
vocab:eps_emp_org	\<http://localhost:2020/resource/eps_organization/103803
vocab:eps_employee_APP_ID	1
vocab:eps_employee_AVAILABLE	0
vocab:eps_employee_CODE	001
vocab:eps_employee_CREATE_DATE	2014-02-24 15:34:27
vocab:eps_employee_CREATE_USER_ID	1
vocab:eps_employee_DEPT_ID	1038032
vocab:eps_employee_EMAIL	zhangsan@00.com
vocab:eps_employee_FLAG	1
vocab:eps_employee_ID	1038035
vocab:eps_employee_MARITAL_STATUS	1
vocab:eps_employee_NAME	张三
vocab:eps_employee_ORG_ID	\<http://localhost:2020/resource/eps_organization/103803
vocab:eps_employee_SEX	1
vocab:eps_employee_TYPE	official
rdfs:label	eps_employee #1038035
rdf:type	vocab:eps_employee

知识图谱构建—可视化编辑软件

➤ RDF数据查询

- Protégé:开源
- Isaviz:开源
- Altova SchemaAgent (x64)/ Altova SemanticWorks® 2010：收费
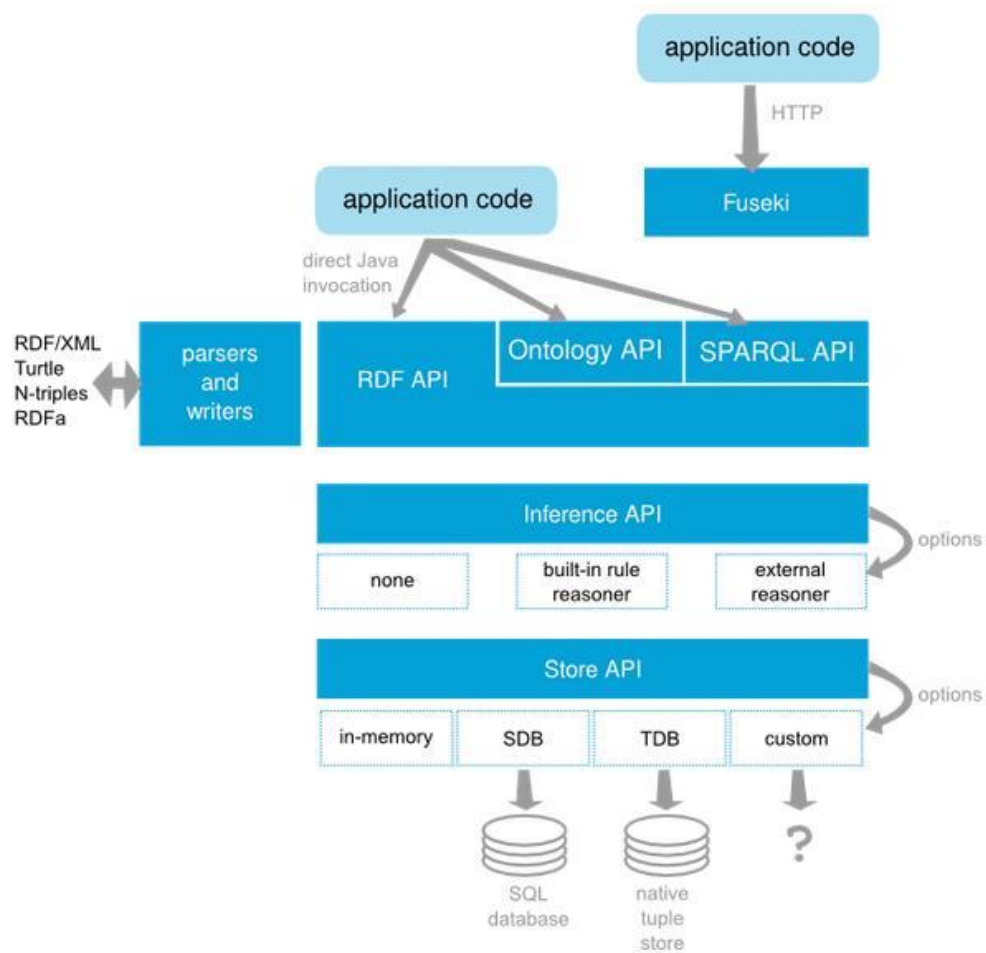
Site:http://d2rq.org/jena

知识图谱存储—RDF存储

➢ RDF数据转成物理RDF

dump-rdf -f N-TRIPLE -b http://localhost:2020/ -m mapping-iswc.ttl -o iswc.nt
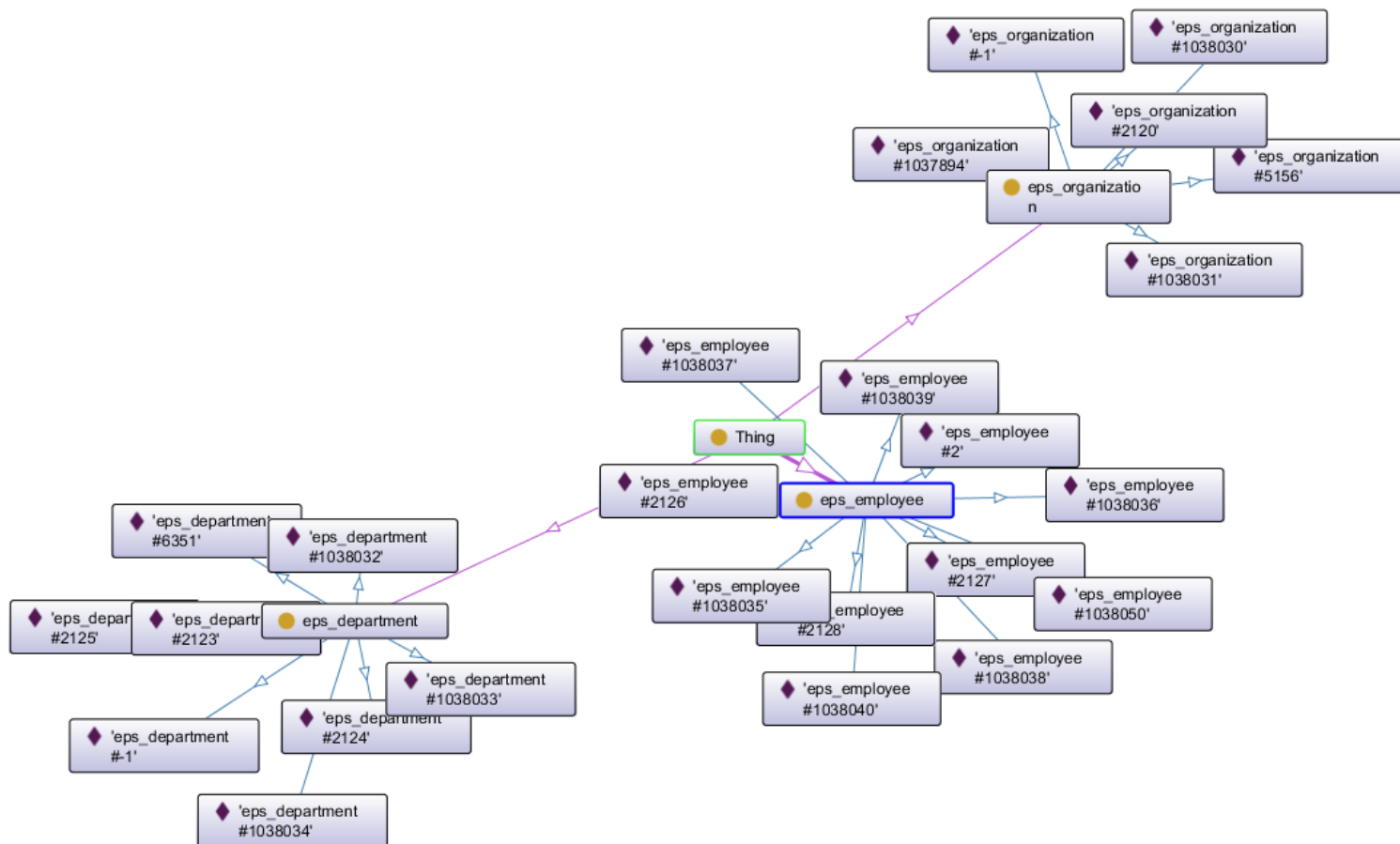
知识图谱存储—RDF存储

➤ RDF数据存储

知识图谱存储—RDF存储

➤ RDF数据存储

- Jena支持单机数据库

  TDB—Jena自己

SDB Databases Supported

Product	Version
Oracle 10g	Including OracleXE
Microsoft SQL Server 2005	Including MS SQL Express
DB2 9	Including DB2 9 Express
PostgreSQL	v8.2
MySQL	v5.0.22
Apache Derby	v10.2.2.0
H2	1.0.71
HSQLDB	1.8.0

- Jena支持分布式数据库

  Jena-hbase: A Distributed, Scalable and Efficient RDF Triple Store

中科汇联 Huilan Infotech 一切为了用户体验！

知识图谱存储—RDF存储

➤ RDF数据存储

- ● Neo4j—图数据库

  1. Neo4J支持RDF，SPARQL等扩展，**sparql-plugin**
  2. 数十亿的节点数

知识图谱查询—D2R虚拟查询

➤ RDF数据查询

- Jena

  Using D2RQ with Jena's Model API

  Using D2RQ with Jena's Graph API

  Executing SPARQL queries against a ModelD2RQ

- Http&&D2R

Site:http://d2rq.org/jena

知识图谱查询—D2R虚拟查询

➤ RDF数据查询

- Jena

Using D2RQ with Jena's Model API

Using D2RQ with Jena's Graph API

Executing SPARQL queries against a ModelD2RQ

Site:http://d2rq.org/jena

知识图谱构建—可视化编辑软件