

Интеграция данных реляционных баз данных на базе технологий семантической паутины

Студент группы ДА-61

Гомольский Е. В.

Цель работы

- исследование архитектур систем интеграции данных на базе технологий семантической паутины;
- прототипирование системы интеграции данных;

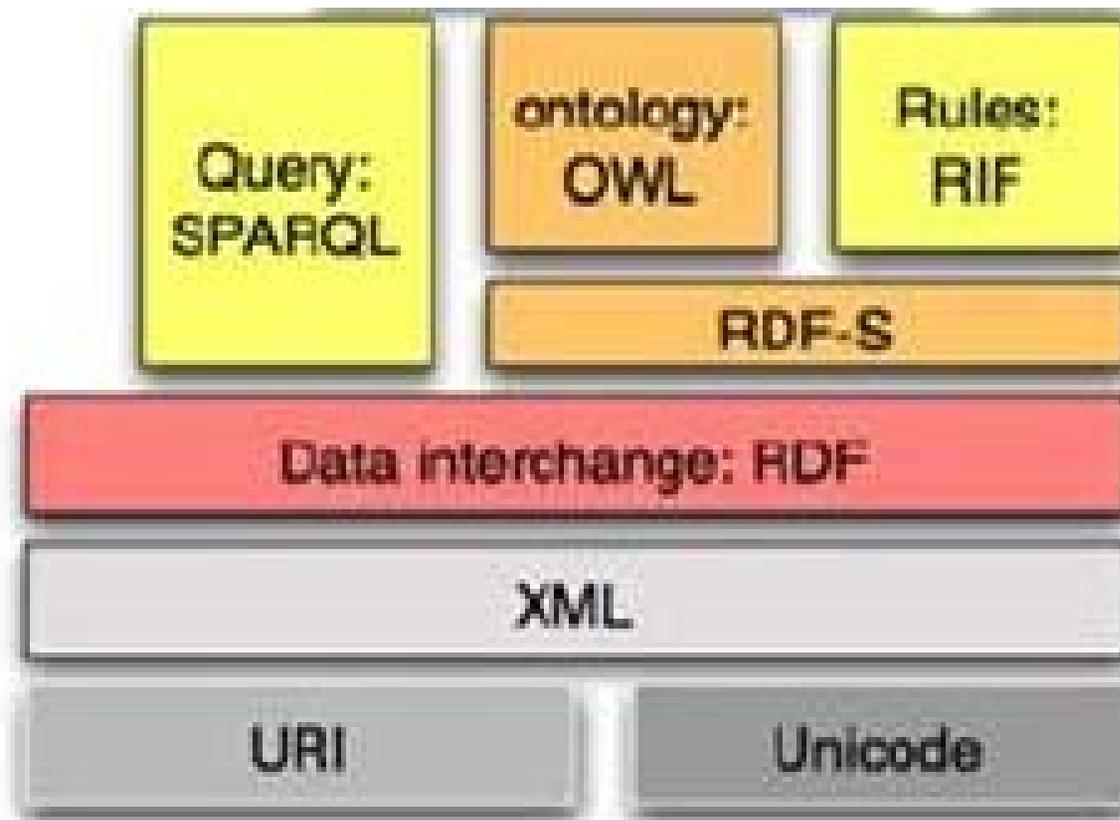
Интеграция данных

Задача интеграции данных заключается в соединении данных из различных источников и предоставлении пользователю единого (унифицированного) представления этих данных, в том числе возможности извлечения интересующей пользователя информации по запросу.

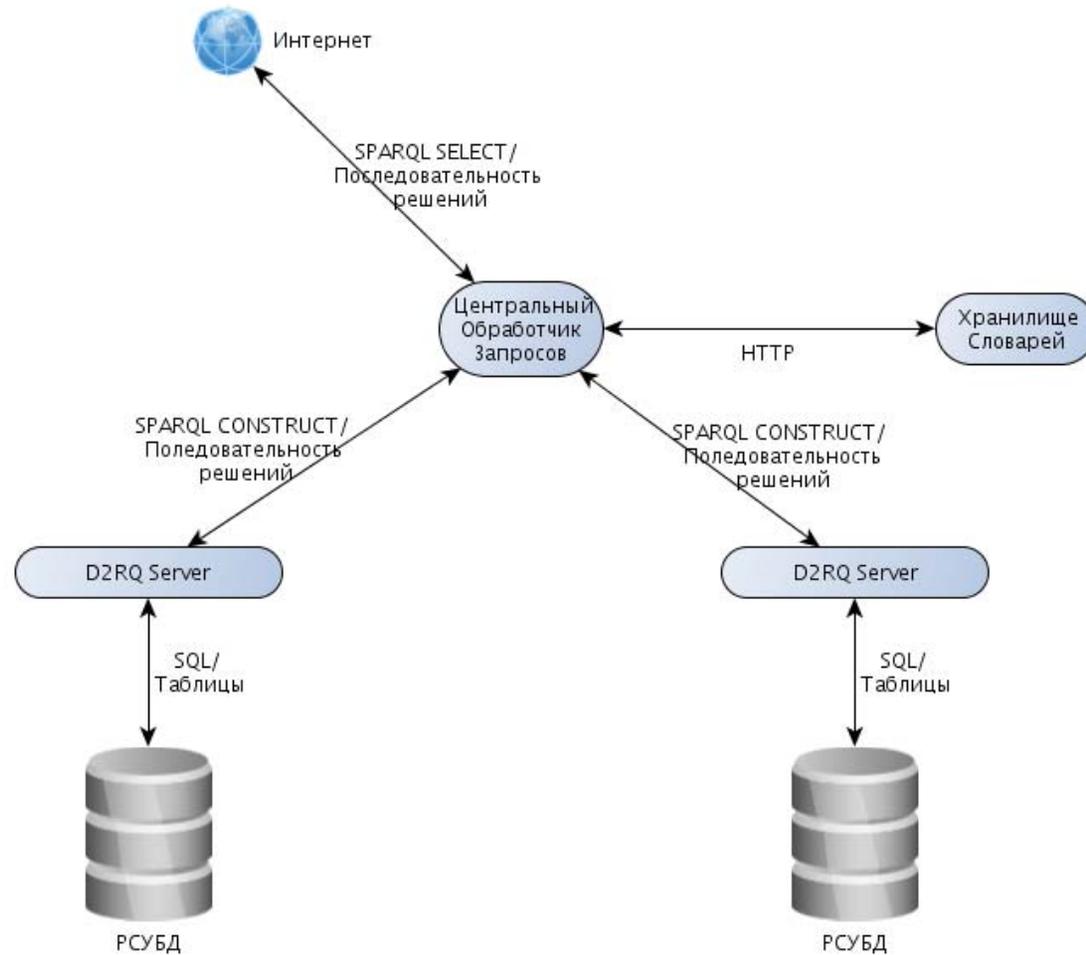
Типы несоответствия схем данных

- Конфликты неоднородности ;
- Конфликты именованя;
- Семантические конфликты;
- Структурные конфликты;

Семантическая паутина



Архитектура системы



Интерфейс системы

Snorql: Exploring <http://localhost:8080/webdis/sparql>

SPARQL:

PREFIX xsd: <<http://www.w3.org/2001/XMLSchema#>>

PREFIX gav: <<http://dl.dropbox.com/u/30752769/disconfigs/triconf/gav.owl#>>

```
SELECT *  
WHERE  
{  
  ?country gav:gav_Country_Capital ?capital.  
  ?country gav:gav_Country_Population ?country_population.  
  ?capital gav:gav_City_Name "Kyiv"^^xsd:string.  
}
```

Results:

SPARQL results:

country	capital	country_population
http://213.111.121.16:2004/resource/Country/UKR	http://213.111.121.16:2004/resource/City/3426	50456000

Powered by [D2R Server](#)

Тестовые базы данных (1)

CitiesAndCountries
- CityID
- CityName
- CityDistrict
- CityPopulation
- CityIsCapital
- CountryCode
- CountryName
- CountryRegion
- CountryPopulation
- CountryLifeExpectancyInDays
- CountryContinent

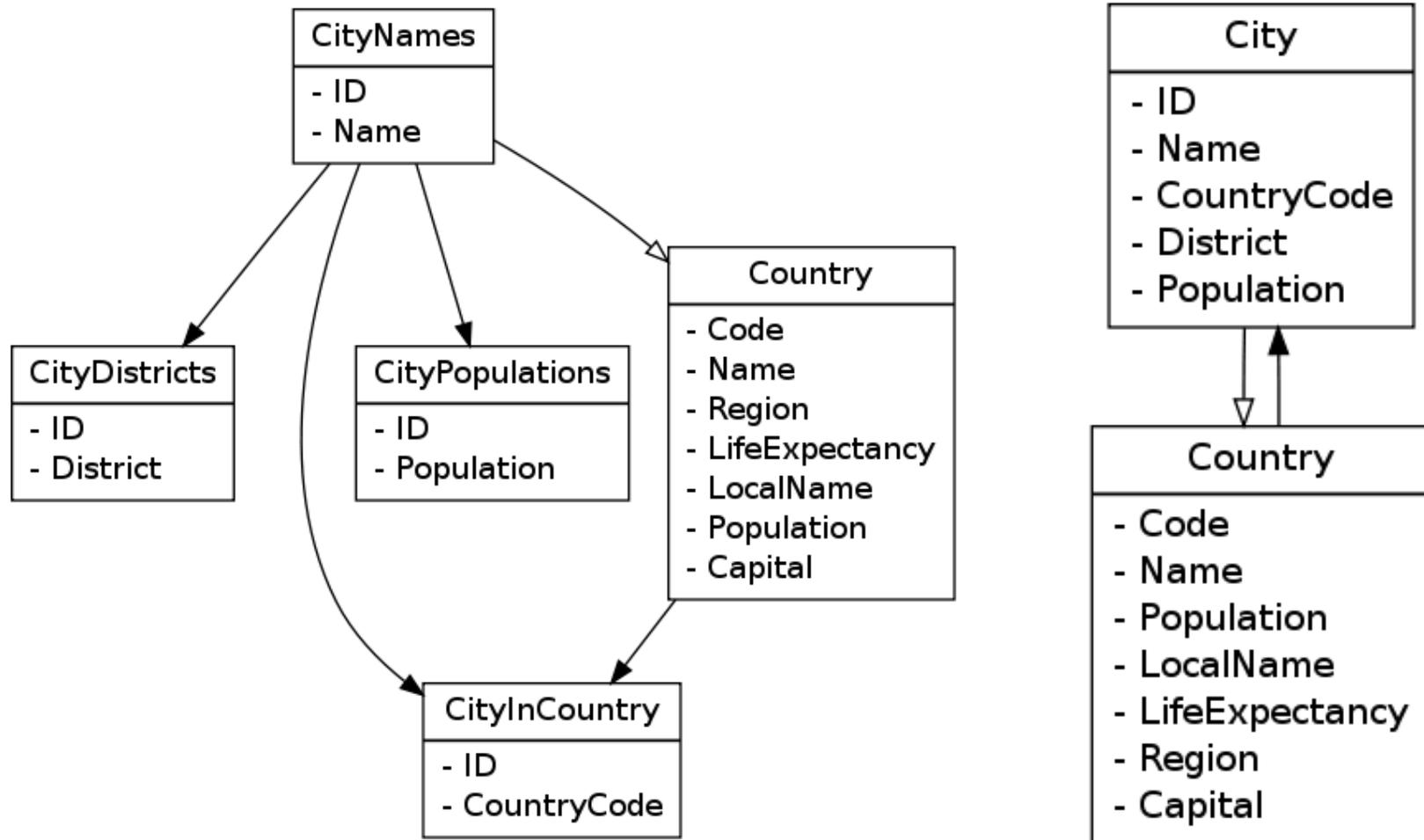
Cities
- ID
- Name
- CountryCode
- District
- Population

Countries
- Code
- Name
- Region
- Population
- Continent
- LifeExpectancy

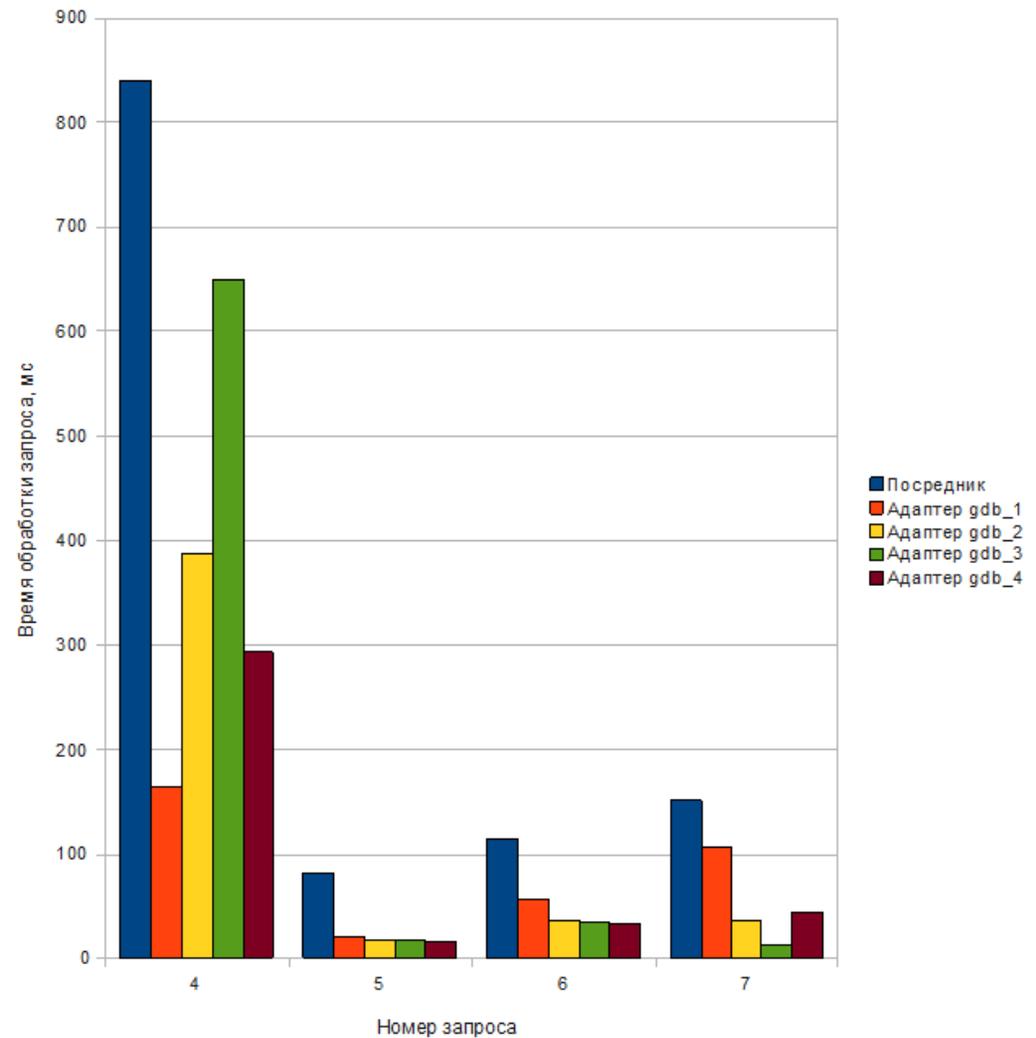
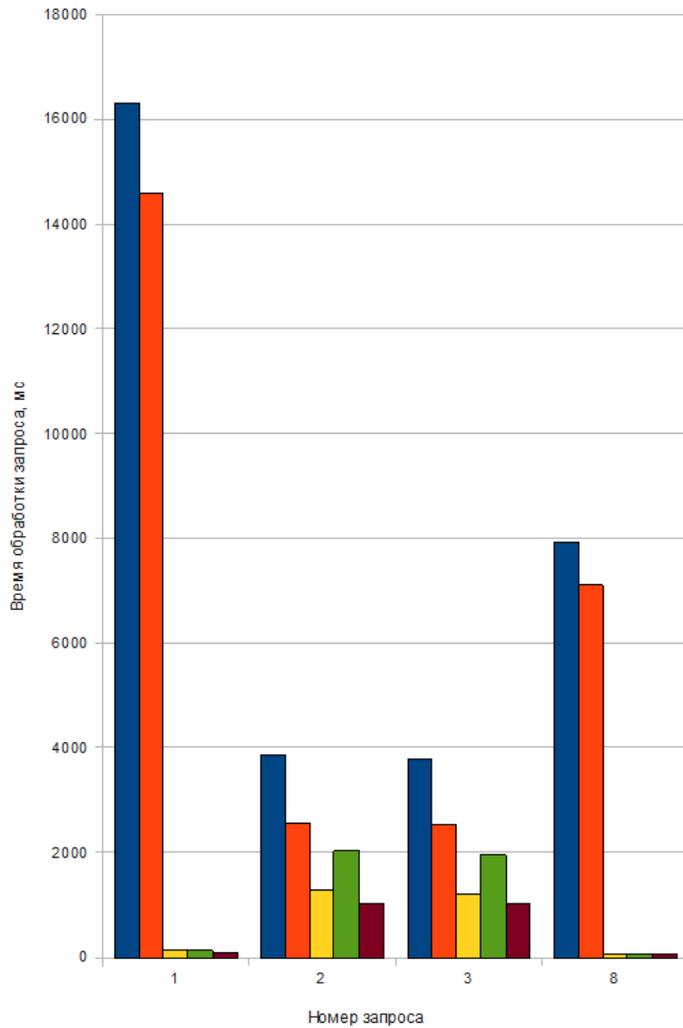
CapitalsOfCountries
- ID
- CountryCode



Тестовые базы данных (2)



Производительность системы



Выводы

- Технологии семантической паутины предоставляют широкие возможности по интеграции данных с различных источников, переводя все источники данных к единой RDF модели, за счет посредников и позволяя рассматривать данных, представленные в них, как единый RDF граф;
- Ключевым понятием RDF модели является понятие ресурса, которое решает проблему глобальной идентификации объектов, возникающей при переходе с уровня одной базы данных к глобальному уровню;

Спасибо за внимание