

НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Кафедра Систем информатики

МЕТОДЫ И СРЕДСТВА УСТАНОВЛЕНИЯ СООТВЕТСТВИЯ МЕЖДУ ОНТОЛОГИЯМИ

Выполнила:

Студентка группы 0208

Селиванова Надежда Павловна

Научный руководитель:

Апанович Зинаида Владимировна

Новосибирск
2013

SEMANTIC WEB

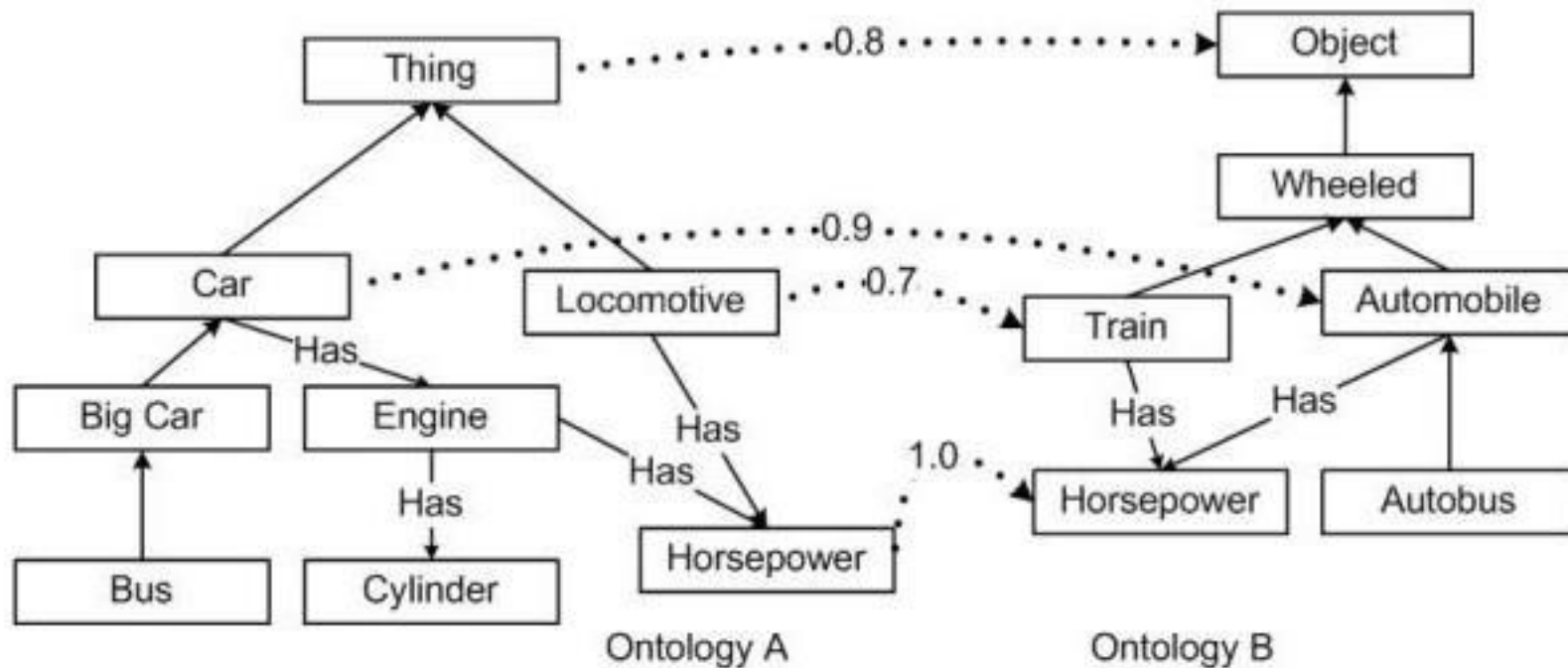
Semantic Web – направление развития Интернет, с целью представления информации в виде пригодном для машинной обработки.

Онтология – это структурированное описание некоторой предметной области.



ОТОБРАЖЕНИЕ ОНТОЛОГИЙ

Отображение онтологий (ontology mapping) - деятельность по установлению соответствия между несколькими онтологиями или нахождение семантических связей подобных элементов из разных онтологий.



ТЕСТОВЫЕ ДАННЫЕ

АКТ Reference Ontology – представляет знания о людях, проектах, публикациях, географических данных .

```
<owl:ObjectProperty rdf:ID="owned-by">
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#Information-Bearing-Object"/>
        <owl:Class rdf:about="#Technology"/>
        <owl:Class rdf:about="#Method"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
  <rdfs:range rdf:resource="#Legal-Agent"/>
  <rdfs:isDefinedBy rdf:resource="#&base;"/>
</owl:ObjectProperty>

<owl:Class rdf:ID="Abstract-Information">
  <rdfs:comment>Information in general, independent of an object
  <rdfs:subClassOf rdf:resource="#&support;Intangible-Thing"/>
  <rdfs:isDefinedBy rdf:resource="#&base;"/>
</owl:Class>
```

Онтология неспецифических сущностей (ОНС) – открытый архив Сибирского отделения Российской Академии Наук

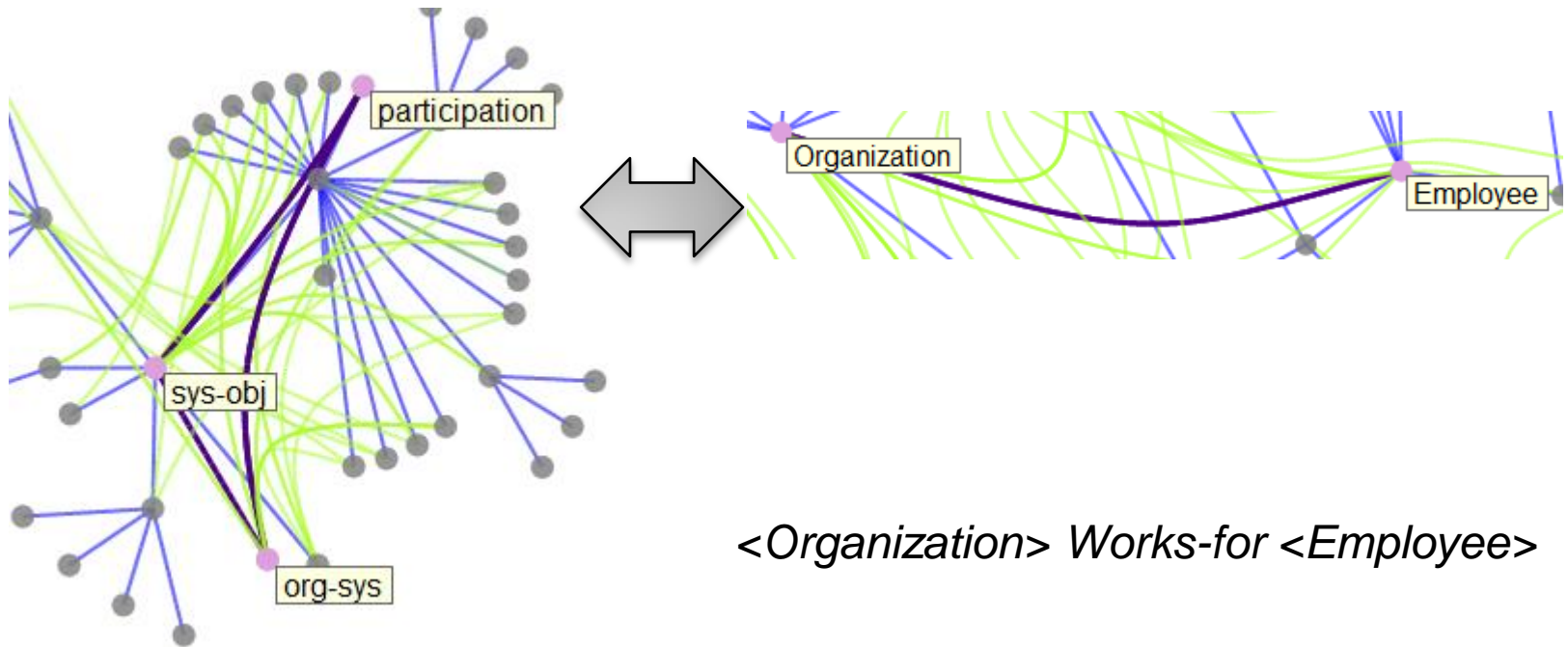
```
<owl:Class rdf:ID="sys-obj">
  <rdfs:subClassOf rdf:resource="#Thing" />
</owl:Class>
<owl:DatatypeProperty rdf:ID="name">
  <rdfs:label xml:lang="ru">имя</rdfs:label>
  <rdfs:label xml:lang="en">name</rdfs:label>
  <rdfs:domain rdf:resource="#sys-obj" />
  <rdfs:range rdf:resource="#text" />
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:ID="description">
  <rdfs:label xml:lang="ru">описание</rdfs:label>
  <rdfs:label xml:lang="en">description</rdfs:label>
  <rdfs:domain rdf:resource="#sys-obj" />
  <rdfs:range rdf:resource="#text" />
</owl:DatatypeProperty>
<owl:Class rdf:ID="person">
  <rdfs:label xml:lang="ru">Персона</rdfs:label>
  <rdfs:label xml:lang="en">Person</rdfs:label>
  <rdfs:subClassOf rdf:resource="#sys-obj" />
</owl:Class>
```



ОСОБЕННОСТЬ ОНТОЛОГИИ ОНС

СТРУКТУРА ИЗ ОНС

СТРУКТУРА ИЗ АКТ



<Organization> Works-for <Employee>

<org-sys> in-org <participation> participant <sys-obj>



ПРОГРАММНЫЕ СРЕДСТВА, ПРЕДНАЗНАЧЕННЫЕ ДЛЯ ОТОБРАЖЕНИЯ ОНТОЛОГИЙ

Преимущества Agreement Maker:

- Точность найденных соответствий;
- Большой выбор алгоритмов;
- Возможность накладывать и объединять результаты, полученные различными алгоритмами;

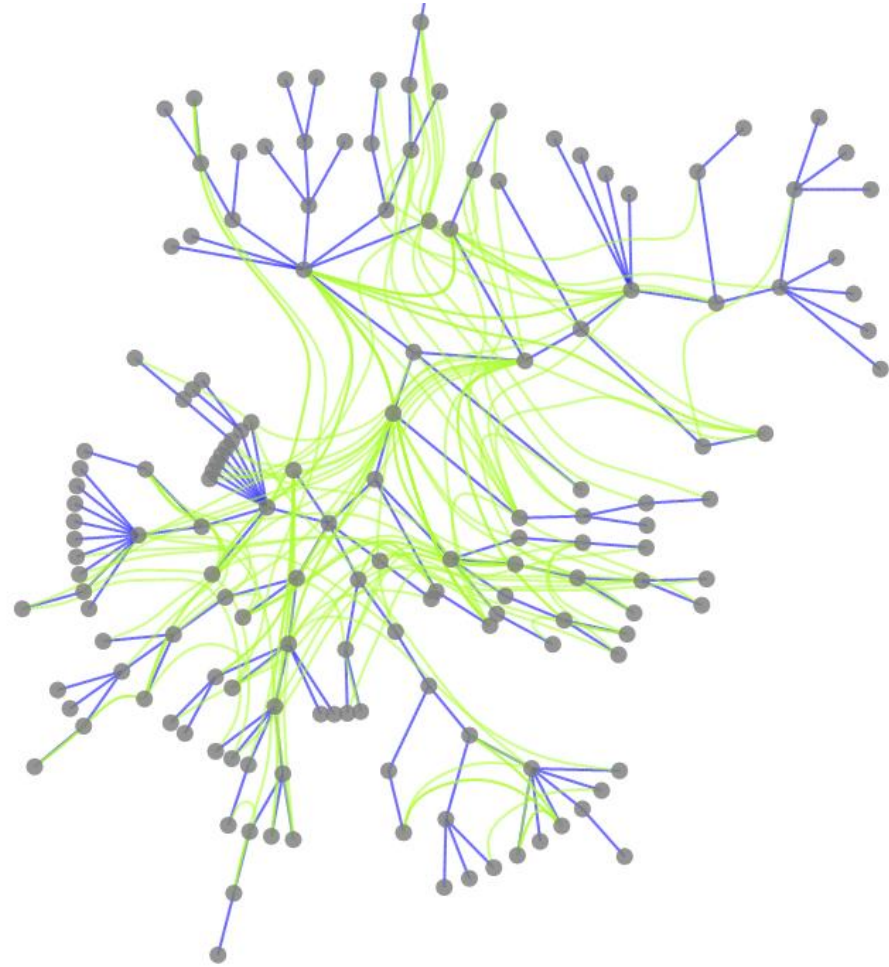
The screenshot displays the AgreementMaker application window. The main area shows two ontologies side-by-side with numerous green lines representing matches between their classes and properties. The matches are annotated with confidence percentages such as 100%, 99%, and 98%. The left ontology includes terms like 'is-recipient', 'learn', 'aspirant', 'learn-special', 'pupil', 'student', 'location', 'naming', 'org-relatives', 'participation', 'post-address', 'reflection', 'reg-phone-code', 'scanned-page', 'sys-obj', 'archive', 'collection', and 'cassette'. The right ontology includes terms like 'Award', 'Degree', 'Academic-Degree', 'Financial-Award', 'Employment-Contract-Type', 'Gender', 'Generic-Area-Of-Interest', 'Business-Area', 'Research-Area', 'Method', 'Organization-Size', 'Quantity', 'Physical-Quantity', 'Amount-Of-Money', 'Duration', 'Software-Status', 'Time-Position', and 'Time-Interval'. At the bottom, the 'Matchers Control Panel' is visible, featuring buttons for Match, New, Copy, Delete, Clear All, Reference Evaluation, Quality Evaluation, Export, Import, and Tuning. Below these buttons is a table with columns for Color, Index, Name, Show/Hide, Threshold, S-Relations, T-Relations, Input Matchers, Modified, Align Class, Align Prop, Performance(ms), and Fou... The table contains three rows of data:

Color	Index	Name	Show/Hide	Threshold	S-Relations	T-Relations	Input Matchers	Modified	Align Class	Align Prop	Performance(ms)	Fou...
Green	0	User Manual Matching	<input checked="" type="checkbox"/>	1%	ANY	ANY	N/A	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	9	0
Yellow	1	(Test) All One Similarities	<input checked="" type="checkbox"/>	100%	1	1	N/A	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	214	135
Red	2	(Test) Random Similarities	<input checked="" type="checkbox"/>	70%	1	1	(Test) All One Similarities	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	24	135



МЕТОДЫ ОТОБРАЖЕНИЯ ОНТОЛОГИЙ

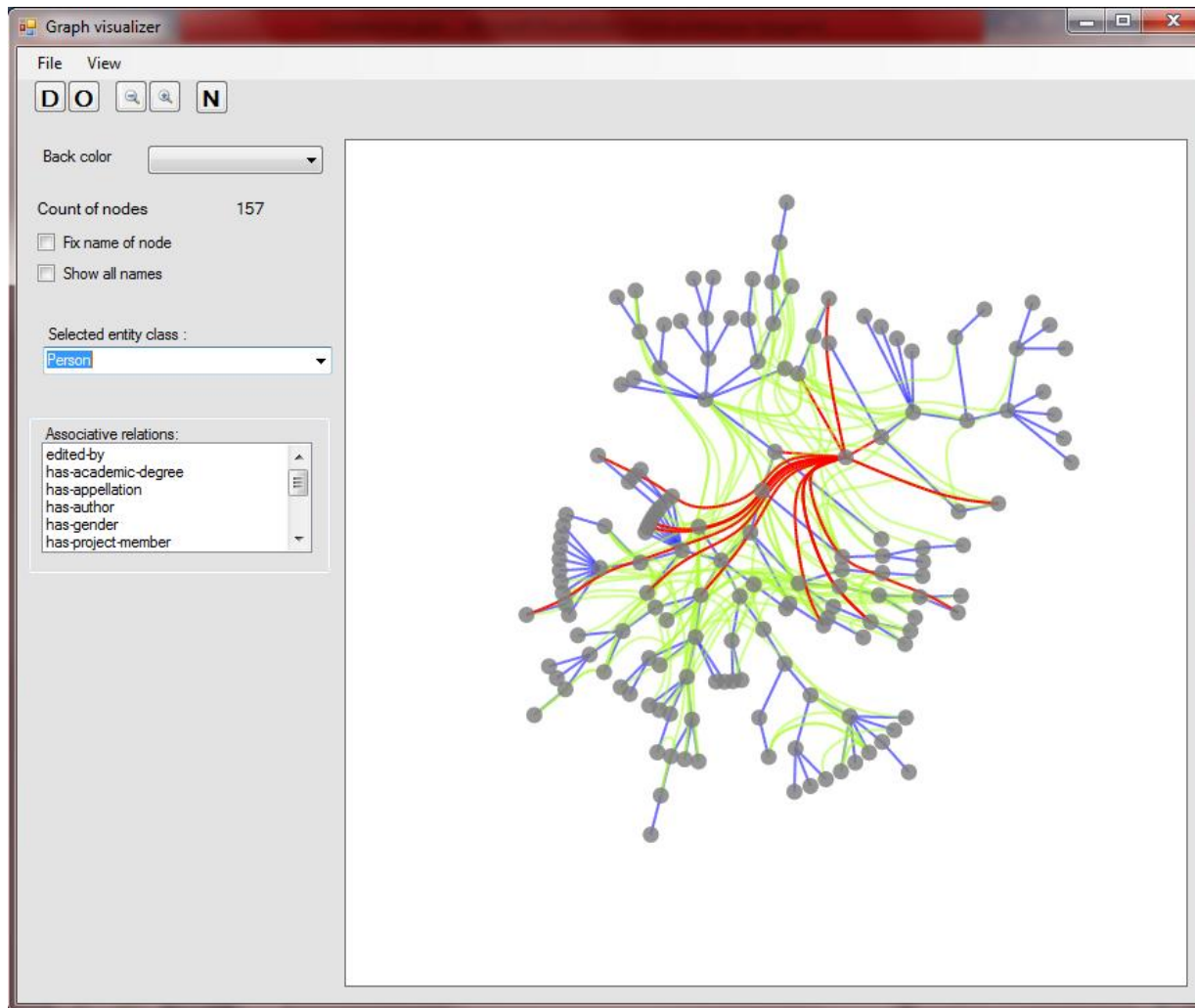
- **Лингвистический /
Лексический /
Текстовый анализ**
- Структурный анализ
- Экстенсиональный
(статистический)
анализ
- Логический анализ



ПРИМЕР ГРАФА ОНТОЛОГИИ НА 157 ВЕРШИН



ВИЗУАЛИЗАЦИИ ОНТОЛОГИИ И ИНФОРМАЦИОННОГО НАПОЛНЕНИЯ СЕМАНТИЧЕСКИХ СИСТЕМ



ЭТАПЫ РАЗРАБОТКИ МОДУЛЯ ДЛЯ ОТОБРАЖЕНИЯ ОНТОЛОГИЙ

1. Адаптация программы "Визуализации онтологии и информационного наполнения семантических систем" для внедрения модуля.
2. Разработка алгоритма лексического анализа.
3. Создание инструмента для установления соответствия между онтологиями.
4. Получение данных по построенным соотношениям. Решением данной задачи является разработка SPARQL запроса с автоматической подстановкой классов и отношений, и сохранением полученных данных.



АДАПТАЦИЯ ПРИЛОЖЕНИЯ ВИЗУАЛИЗАЦИИ ОНТОЛОГИИ И ИНФОРМАЦИОННОГО НАПОЛНЕНИЯ СЕМАНТИЧЕСКИХ СИСТЕМ

The screenshot displays a 'Graph visualizer' application window. The interface includes a menu bar with 'File' and 'View', a toolbar with icons for 'D', 'O', search, and 'N', and a 'Back color' dropdown. The main area is split into two panels, each showing a network graph. The left panel shows a graph for the 'authority' class with 44 nodes. The right panel shows a graph for the 'Employee' class with 157 nodes. Both graphs feature nodes connected by edges, with some nodes highlighted in red and green. The right panel also includes a 'Count of nodes' field (157), a 'Fix name of node' checkbox (checked), a 'Selected entity class' dropdown (Employee), and a list of 'Associative relations' (has-contract-type, is_parent, works-for, works-in-unit). A 'Similar class' section shows 'Lexical' and 'person (syn: person)'. The bottom left corner lists various ontology classes such as 'University-Faculty', 'Meeting-Taking-Place', 'Government-Organization', etc.

АЛГОРИТМ ЛЕКСИЧЕСКОГО АНАЛИЗА

КРИТЕРИЙ:

Если метки двух сущностей подобны, то эти сущности подобны.

ВХОДНЫЕ ДАННЫЕ:

- Исходный класс первой онтологии;
- Список классов второй онтологии;

ВЫХОДНЫЕ ДАННЫЕ:

- Список подобных классов;

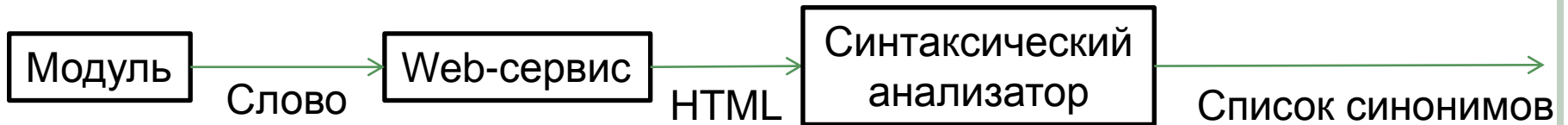
ПРИМЕРЫ НАЗВАНИЙ КЛАССОВ ОНТОЛОГИИ:

- Abstract-Information (*Подслова: abstract, information*);
- Information-Bearing-Object (*Подслова: information, bearing, object*);
- Recorded-Video (*Подслова: recorded, video*);
- Publication-Reference (*Подслова: publication, reference*).



ТЕЗАУРУС СИНОНИМОВ

СХЕМА ПОЛУЧЕНИЯ СИНОНИМОВ



Thesaurasize

A really fast thesaurus... with lots of synonyms.



Thesaurasize Random

city (29 synonym found)

Sort city by: [Alphabetically](#) | [Length](#) | [Random](#) | [Definition](#)

[banlieue](#)
[barrio](#)
[boom town](#)
[bourg](#)
[burg](#)
[burgh](#)
[burghal](#)
[citified](#)
[civic](#)
[conurbation](#)
[exurb](#)
[exurbia](#)
[faubourg](#)
[ghost town](#)
[greater city](#)
[interurban](#)
[market town](#)
[megalopolis](#)
[municipal](#)
[oppidan](#)
[Stadt](#)
[spread city](#)
[suburb](#)
[urban](#)
[urban complex](#)
[urban sprawl](#)
[urbs](#)
[ville](#)
[ward](#)

History

This word has been viewed 858 times.

+1 0

Like 0 Send

Tweet 0

Related searches for city:

- [Flickr](#)
- [Folvo Pronunciation](#)
- [Google](#)
- [Panlexicon](#)
- [Thesaurus](#)
- [Twitter](#)
- [Visuwords](#)
- [Wordnik](#)

Most Recent

- [thesaurasize](#)
- [civil disobedience](#)
- [interloped](#)
- [UNDER THE RADAR](#)
- [misconstrue](#)
- [lucky shot](#)
- [stand across](#)
- [...](#)

City

banlieue
barrio
boom town
bourg
burg
burgh
burghal
citified
civic
conurbation
exurb



КРИТЕРИИ ПОДОБИЯ СЛОВ

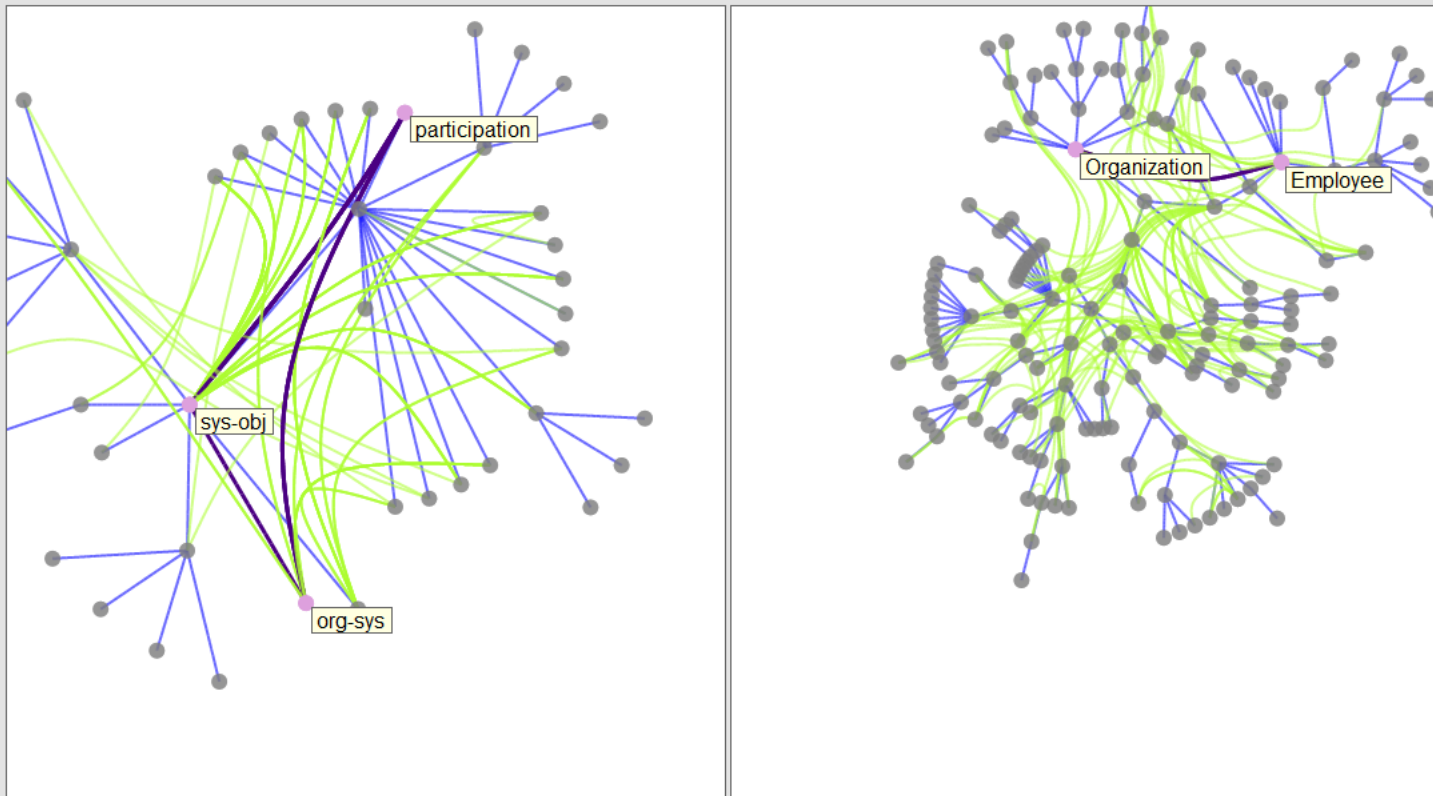
1. Полное совпадение (*city - city*)
2. Синонимичные термины (*city - town*)
3. Слово является подсловом (*org - organization*)
4. Не точное совпадение, рассчитывается по расстоянию Левенштейна (*organization - organizations*)

РАССТОЯНИЕ ЛЕВЕНШТЕЙНА -

это минимальное количество операций вставки одного символа, удаления одного символа и замены одного символа на другой, необходимых для превращения одной строки в другую.



ИНСТРУМЕНТ УСТАНОВЛЕНИЯ СООТВЕТСТВИЯ



Instrument matching

Fix mathing

Save Map

Send request

Create Map

Delete Map

Import in RDF

sys-obj
participant
participation
Sub-class
in-org
org-sys

Organization
works for
Employee

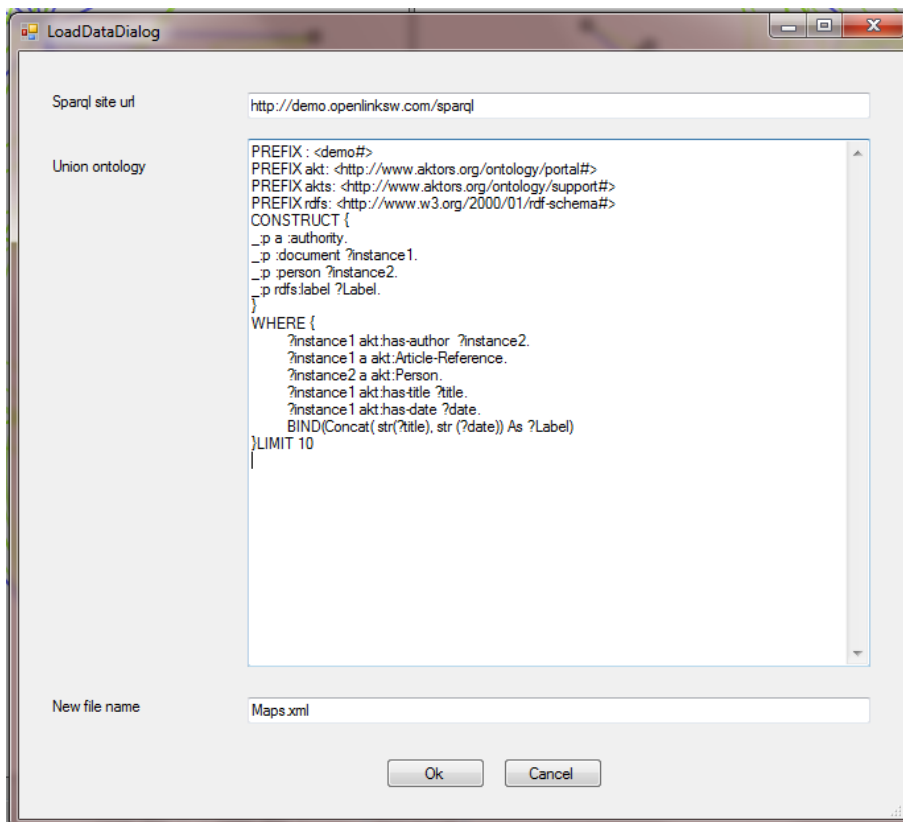
<person> <Person>



ПОЛУЧЕНИЕ ДАННЫХ ПО ПОСТРОЕННЫМ СООТНОШЕНИЯМ

SPARQL — язык запросов к данным, представленным по модели RDF, а также протокол для передачи этих запросов и ответов на них. Был создан как язык запросов для Semantic Web.

ПРИМЕР ЗАПРОСА СООВТОРСТВА, ПОСТРОЕННОГО АВТОМАТИЧЕСКИ



ШАБЛОН SPARQL ЗАПРОСА

PREFIX : <demo#>

PREFIX akt: <http://www.aktors.org/ontology/portal#>

PREFIX akts: <http://www.aktors.org/ontology/support#>

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

CONSTRUCT {

_:p a :Class4.

_:p rdfs:label ?Label.

_:p :relation2 ?instance1.

_:p :relation3 ?instance2.

}

WHERE {

?instance1 akt:relation1 ?instance2.

?instance1 a akt:Class1.

?instance2 a akt:Class2.

?instance1 akt:label ?instance1_label.

?instance2 akts:label ?instance2_label.

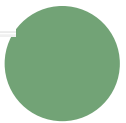
BIND(Concat(str(?instance1_label), str (?instance2_label)) As ?Label)

}



РЕЗУЛЬТАТ РАБОТЫ ЗАПРОСА

```
<?xml version="1.0" encoding="utf-8" ?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:ns2="demo#" >
  <rdf:Description rdf:nodeID="b4047498">
    <rdf:type rdf:resource="demo#authority" />
    <rdfs:label>Computer Security - A Survey.http://www.aktors.org/ontology/date#1972</rdfs:label>
    <ns2:document rdf:resource="http://dblp.rkbexplorer.com/id/conf/afips/Browme76" />
    <ns2:person rdf:resource="http://dblp.rkbexplorer.com/id/people-acc54203a8dd06eaceb4e76d2809279e-f2549dc515e0350bc01566555b38d4c8" />
  </rdf:Description>
  <rdf:Description rdf:nodeID="b4047499">
    <rdf:type rdf:resource="demo#authority" />
    <rdfs:label>Theoretical Improvements in Algorithmic Efficiency for Network Flow Problems.http://www.aktors.org/ontology/date#1972</rdfs:label>
    <ns2:document rdf:resource="http://dblp.rkbexplorer.com/id/conf/aussois/EdmondsK01" />
    <ns2:person rdf:resource="http://dblp.rkbexplorer.com/id/people-70045bbb52077b5673c04fbb565a254-8bc26de2c3fdf5b0d8a16a31f66787f5" />
  </rdf:Description>
  <rdf:Description rdf:nodeID="b4047500">
    <rdf:type rdf:resource="demo#authority" />
    <rdfs:label>Theoretical Improvements in Algorithmic Efficiency for Network Flow Problems.http://www.aktors.org/ontology/date#1972</rdfs:label>
    <ns2:document rdf:resource="http://dblp.rkbexplorer.com/id/conf/aussois/EdmondsK01" />
    <ns2:person rdf:resource="http://dblp.rkbexplorer.com/id/people-bd1f1b832dbadfc0beea876e9e6b6e18-f8f47b01222e6d5ab1a6a98c8901b6c0" />
  </rdf:Description>
  <rdf:Description rdf:nodeID="b4047501">
    <rdf:type rdf:resource="demo#authority" />
    <rdfs:label>Equivalence of Programs with Structured Variables.http://www.aktors.org/ontology/date#1972</rdfs:label>
    <ns2:document rdf:resource="http://dblp.rkbexplorer.com/id/conf/focs/AhoU70" />
    <ns2:person rdf:resource="http://dblp.rkbexplorer.com/id/people-c8c0bd2001d087784a5e09b622223d1a-4fb4a90250453df6a4c29af9e92f7f8e" />
  </rdf:Description>
  <rdf:Description rdf:nodeID="b4047502">
    <rdf:type rdf:resource="demo#authority" />
    <rdfs:label>Equivalence of Programs with Structured Variables.http://www.aktors.org/ontology/date#1972</rdfs:label>
    <ns2:document rdf:resource="http://dblp.rkbexplorer.com/id/conf/focs/AhoU70" />
    <ns2:person rdf:resource="http://dblp.rkbexplorer.com/id/people-fafc0f08750dc34d161ae9159d06b938-1786714398f5994c6e4e8221321c08a0" />
  </rdf:Description>
</rdf:RDF>
```



ПРОДЕЛАННАЯ РАБОТА

- Изучены подразделы предметной области Semantic Web;
- Исследованы программные средства, позволяющие различные операции с онтологиями;
- Изучены существующие методы для отображения онтологий;
- Разработан и реализован лексический алгоритм установления соответствия между классами;
- Разработаны новые функциональные возможности для программы "Визуализации онтологии и информационного наполнения семантических систем", позволяющие устанавливать соответствия в интерактивном режиме;
- Добавлена возможность автоматического формирования SPARQL запроса для получения данных по построенным соотношениям;
- Результаты были протестированы на онтологиях ОНС и АКТ.



СПАСИБО ЗА ВНИМАНИЕ!

