

**SOEN 691B Semantic Computing**

**Winter 2013/2014**

**Assignment #2**

**Report**

**Presented by: Anish Chavan**

**Studentid :6421180**

Changes in assignment 1:

Few changes werè made to the assignment .

Gazetteer was modified .Few changes were made to the rules in organization jape rule in order to increase the precision.

dataset of documents was changed

source of new data documents http://www.cs.cmu.edu/afs/cs/project/theo-20/www/data/

correlation was added for detecting the sameA relation bet ween similar meaning words

Co-reference used ;pronomial,nominal and annies co- reference.

**Schema Diagram:owl**

**Schema Description:**

Tool used :Protégé.

Schema is designed in owl which includes owl classes and properties.

Schema consists of main classes for entities has root class called Thing .Thing is the root class in owl.

The entities classes which come under root are UniversityPosition,OrganizationalUnit,

University Position.

University Person is the subclass of FOAF:Person .An already defined schema which provides with already defined classes and properties.

University on other hand also is the subclass of the FOAF:organization .

The schema also has two Object properties .

***belongs\_to:***This Object property has domain and range .Dmoin and range are the RDFS feature added to the property

*belongs\_to* consist of Universuty Person as domain and University as range.

***has\_position:***Imilaar to belongs\_to has postion also uses RDFS domain and RDFS range

has\_posiion has University person as domain and range is university position.

In project the *newo.owl* .file is the schema which is loded in the onamodel using default ontmodel

**RDFExport:**

The linking is done by reading the annotation from the annotation set and creating the

OntModel class and reading th e required ont model class from the shcema.Schema has been as mentioned ab ove has already been read in the OntModel. object.

OntModel onto = ModelFactory.*createOntologyModel*( OntModelSpec.*OWL\_MEM*, **null** );

onto(newo.owl)

The individual of each clas is created and the class is URI is acquired from create Individual("",classname)

The individual is augmented with the property of type in which the class is mentioned .one more property is also added to the an Individual with FOAF.Organization as a super class and this individuals subclass of that class.

**Example of the University**

////

OntClass equipe = onto.getOntClass(prefix + "University" );

Individual m=onto.createIndividual(prefix +mName ,equipe);

m.addProperty(RDF.*type*,equipe).addProperty(RDFS.*subClassOf*,FOAF.*Organization* ).addProperty(OWL.*sameAs*,DBPEDIA +mName);

Prefix is common url in an university is class defined in schema. The individual is added with the class as property type. plus mentioned also the subclass of the FOAF.Organization.

Exported data:

<rdf:Description rdf:about="http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#Emory\_University">

<rdf:type rdf:resource="http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#University"/>

***<rdfs:subClassOf rdf:resource="http://xmlns.com/foaf/0.1/Organization"/>*** -------->subclass of orgnaization

<owl:sameAs>http://dbpedia.org/resource/Emory\_University</owl:sameAs> ------------>**Dbpedi links**

</rdf:Description>

Similarly its done for University person,UniverityPosition ,Organizational unit .

Following are the examples of each respectievely.

**Example of University Person**

*University person is subclass of foaf person*

<rdf:Description rdf:about="http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#J.+Ullman">

*<rdfs:subClassOf rdf:resource="http://xmlns.com/foaf/0.1/Person"/> --->subclass of person*

<rdf:type rdf:resource="http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityPerson"/>

</rdf:Description>

**Example of UniversityPosition:**

***Instance of the class university position:***

<rdf:Description rdf:about="http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#Editor">

<rdf:type rdf:resource="http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityPosition"/>

</rdf:Description>

**belongs\_to:**

As discussed earlier belongs to is the object property which has domain and range

The domain and range are university person and university respectievely.The instance of the belongs to object property is created and the additional property is added to the object property.

following is the implementation of belongs\_to

ObjectProperty belongs\_to=onto.getObjectProperty(prefix+"belongs\_\_to");

Random or = **new** Random();

Individual m2=onto.createIndividual(prefix+or.nextInt(99999) +name,belongs\_to);

m2.addProperty(RDFS.*domain*,UniversityPerson.getURI()+domain);

m2.addProperty(RDFS.*range*,University.getURI()+range);

m2.addProperty(RDF.*predicate*, "belongsto");

Emaple output of belongs to from output file

<*rdf:Description rdf:about="http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#22118University\_of\_Siegen">*

*<rdf:type rdf:resource="http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#belongs\_\_to"/>*

*<rdfs:domain>http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityPersonLeben</rdfs:domain>*

*<rdfs:range>http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityUniversity\_of\_Siegen</rdfs:range>*

*<rdf:predicate>belongsto</rdf:predicate>*

</rdf:Description>

***Each instances is created unique by using the random number Plus the rdf type is uri of belong t\_o***

**has\_postition:**

As discussed earlier belongs to is the object property which has domain and range

The domain and range are university person and universityposition respectievely.The instance of the belongs to object property is created and the additional property is added to the object property.

following is the implementation of has\_position

ObjectProperty belongs\_to=onto.getObjectProperty(prefix+" has\_position ");

Random or = **new** Random();

Individual m2=onto.createIndividual(prefix+or.nextInt(99999) +name, has\_position);

m2.addProperty(RDFS.*domain*,UniversityPerson.getURI()+domain);

m2.addProperty(RDFS.*range*,University.getURI()+range);

m2.addProperty(RDF.*predicate*, " has\_position ");

Emaple output of belongs to from output file

<*rdf:Description rdf:about="http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#22118University\_of\_Siegen">*

*<rdf:type rdf:resource="http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#* has\_position *"/>*

*<rdfs:domain>http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityPersonLeben</rdfs:domain>*

*<rdfs:range>http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityUniversity\_of\_Siegen</rdfs:range>*

*<rdf:predicate>belongsto</rdf:predicate>*

</rdf:Description>

***Each instances is created unique by using the random number Plus the rdf type is uri of has\_position***

**LiinkingStrategy:**

RDF export is done by reading the the annotation in a for loop from the annotation set .The main

idea is to retrieve the annotation in the string format and replace the spaces in the string with underscore which is generally standard for url.

{String content="";

Integer id= annot.getId();

content = doc.getContent().getConten(annot.getStartNode().getOffset(),

annot.getEndNode().getOffset()).toString();

String mName= content.replace(" ", "\_");

String InstanceURI = URLEncoder.encode(content,"UTF-8");

OntClass equipe = onto.getOntClass(prefix + "University" );

Individual m=onto.createIndividual(prefix +InstanceURI ,equipe);

Thus the each annotationn from selected annotation set is collected and is stored in content as a string format..

This String format is converted in to a link by adding the prefix to it .

**Resource URIgeneration**:

linking is done by globally defined prefix variable

prefix: "http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#";

prefix +InstanceURI and resource is creatted using this link .Thus unique for each instance.

**Dbpedia Linking:**

Dbpedia linking is done by using owl SameAS feature. link si created in same manner but instea of prefix variable new global variable is declared which is used to add the link like this .

String DBPEDIA="http://dbpedia.org/resource/";

Individual m=onto.createIndividual(prefix +mName ,equipe);

String foaforg=FOAF.*Organization*.getURI();

m.addProperty(RDF.*type*,equipe).addProperty(RDFS.*subClassOf*,FOAF.*Organization* ).addProperty(**OWL.*sameAs*,DBPEDIA +mName);**

**Correlation linking:**

Corelation is done by linking all the related they are obtained from the matches and also add the co refrencing objects to the list.

The co frencing list is created for each annotaiton and annotation category is created .

*FeatureMap map = annot.getFeatures();*

***if****((map.get("matches"))!=****null****)*

*{*

*annotCategory=((ArrayList<Long>) map.get("matches"));*

***for****(Object k :annotCategory )*

*{*

***if****(!list.contains((Integer)k))*

*{*

*list.add((Integer)k);*

***Each corefrencing anotation is retrieved from the id list and is stored inn rdf with owl sameAas feature.***

**for** (Object m:annotCategory)

{

Annotation antecedent = blacklist1.get((Integer) m);

String content1 = doc.getContent().getContent

(antecedent.getStartNode().getOffset()

***try*** *{*

*String uri2=URLEncoder.encode(content2,"UTF-8");*

*}* ***catch*** *(UnsupportedEncodingException e) {*

*e.printStackTrace();*

*}*

*onto.createResource(prefix+content1).addProperty(OWL.sameAs, prefix +content2);*

***Example:output***

***<owl:sameAs>http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#David</owl:sameAs>***

***<owl:sameAs>http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#Professor David Warner</owl:sameAs>***

**TripleStore implementation:**

Triple store implementation consists of Individual éinstacnes of the class defimned in schema.

Each Instance is added to its type what class it is member of.I f the class is subclass of any other class special property is added to define it .

for example:

***<rdf:Description rdf:about="http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#Swarthmore\_College">***

***<rdf:type rdf:resource="http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#University"/>***

***<rdfs:subClassOf rdf:resource="http://xmlns.com/foaf/0.1/Organization"/>***

***<owl:sameAs>http://dbpedia.org/resource/Swarthmore\_College</owl:sameAs>***

***<j.0:name>Swarthmore\_College</j.0:name>***

***</rdf:Description>***

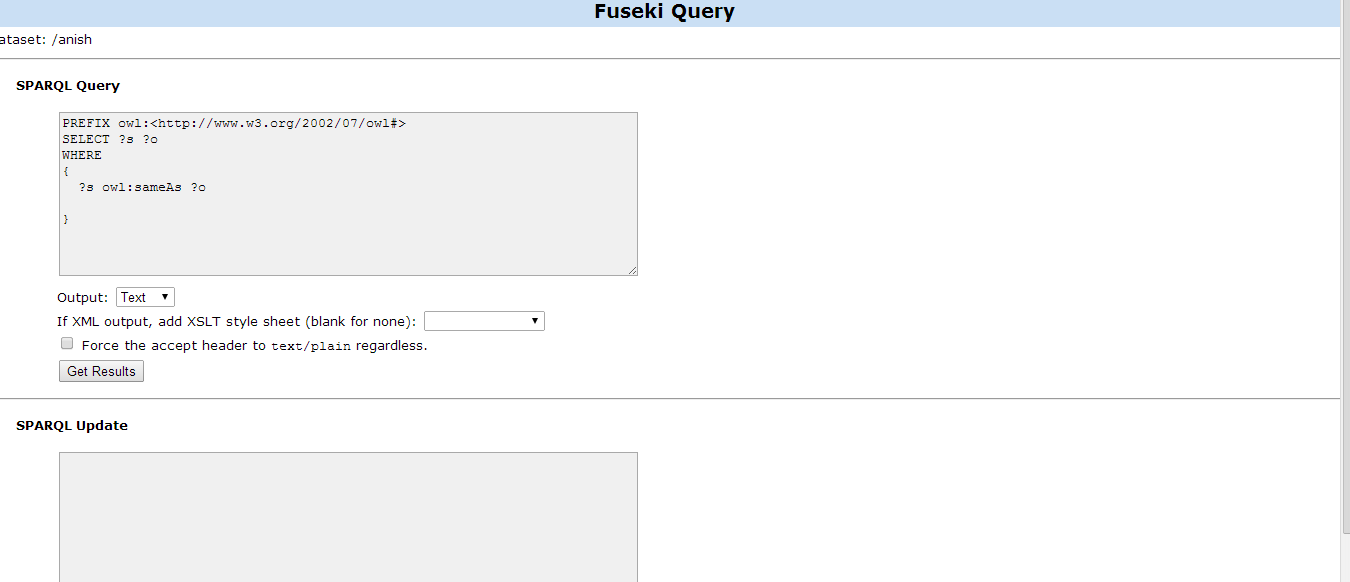
same is done form object property but adding the domain and range of respectieve class

**TOTAL number of triplets** per document :approx 300

in ttoal triplest 1765( per document triplet output on screen )

**Sparqlendpoint :**

The spqrql enpoint was was hosted on local server ,open source fuseki server .



**SPARQL queries:**

**Q1 display the names of all university**

**Follwoing inds the sparql query**

PREFIX owl:<http://www.w3.org/2002/07/owl#>

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

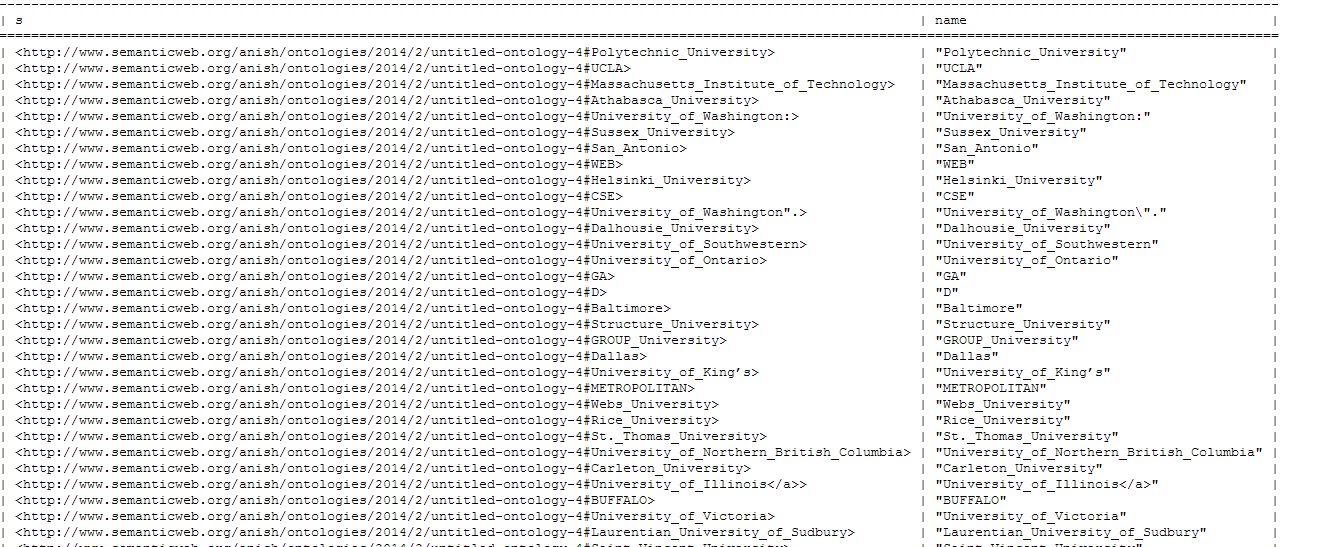
PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>

PREFIX j.0:<http://xmlns.com/foaf/0.1/>

SELECT ?s ?name

WHERE{ ?s rdf:type <http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#University>.

?s j.0:name ?name}



**Q2 Find the names of all university employees, together with their position (if available) and**

**organizational unit (if available)**

PREFIX owl:<http://www.w3.org/2002/07/owl#>

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

PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>

SELECT

?name ?uniperson ?position

WHERE{?name rdf:type <http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#belongs\_\_to>.?name rdfs:domain ?uniperson.

?x rdf:type <http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#has\_position>.

?x rdfs:domain ?uniperson.

?x rdfs:range ?positio}

**output**

-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

| name | uniperson | position |

====================================================================================================================================================================================================================================================================================================================================================

| <http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#50645University\_of\_Siegen> | "http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityPersonLeben" | "http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityPositionLecturer" |

| <http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#50645University\_of\_Siegen> | "http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityPersonLeben" | "http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityPositionLecturer" |

| <http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#42457D> | "http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityPersonGAR" | "http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityPositionAssociate Professor" |

| <http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#32151Humboldt\_University> | "http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityPersonGoel" | "http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityPositionLecturer" |

| <http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#32151Humboldt\_University> | "http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityPersonGoel" | "http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityPositionLecturer" |

| <http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#26907University\_of\_Chester> | "http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityPersonPrincipal" | "http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityPositionVice-Chancellor" |

| <http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#26907University\_of\_Chester> | "http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityPersonPrincipal" | "http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityPositionPrincipal" |

| <http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#26907University\_of\_Chester> | "http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityPersonPrincipal" | "http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityPositionVice-Chancellor" |

| <http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#26907University\_of\_Chester> | "http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityPersonPrincipal" | "http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#UniversityPositionVice-Chancellor" |

| <http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#26907University\_of\_Chester> |

**Q3 For a given university name, provide all information from your knowledge base, as well as**

**DBPedia**

PREFIX owl:<http://www.w3.org/2002/07/owl#>

PREFIX owl:<http://www.w3.org/2002/07/owl#>

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

PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>

PREFIX j.0:<http://xmlns.com/foaf/0.1/>

PREFIX owl:<http://www.w3.org/2002/07/owl#>

PREFIX owl:<http://www.w3.org/2002/07/owl#>

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

PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>

SELECT ?obj ?o ?DBPEDIAdata

WHERE {

{

<http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#St.\_Thomas\_University> ?predi ?obj.

}

UNION{

<http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#St.\_Thomas\_University>

owl:sameAs ?o

SERVICE <http://dbpedia.org/sparql>

{

<http://dbpedia.org/resource/St.\_Thomas\_University> ?q ?DBPEDIAdata

}}}

**Output**

--------------------

| obj | o | DBPEDIAdata |

====================================================================================================================================================================================================================

| <http://www.semanticweb.org/anish/ontologies/2014/2/untitled-ontology-4#University> | | |

| <http://xmlns.com/foaf/0.1/Organization> | | |

| "http://dbpedia.org/resource/St.\_Thomas\_University" | | |

| "St.\_Thomas\_University" | | |

| | "http://dbpedia.org/resource/St.\_Thomas\_University" | owl:Thing |

| | "http://dbpedia.org/resource/St.\_Thomas\_University" | <http://dbpedia.org/ontology/Agent> |

| | "http://dbpedia.org/resource/St.\_Thomas\_University" | <http://dbpedia.org/ontology/Organisation> |

| | "http://dbpedia.org/resource/St.\_Thomas\_University" | <http://dbpedia.org/ontology/University> |

| | "http://dbpedia.org/resource/St.\_Thomas\_University" | <http://dbpedia.org/ontology/EducationalInstitution> |

| | "http://dbpedia.org/resource/St.\_Thomas\_University" | <http://schema.org/Organization> |

| | "http://dbpedia.org/resource/St.\_Thomas\_University" | <http://schema.org/EducationalOrganization> |

| | "http://dbpedia.org/resource/St.\_Thomas\_University" | <http://schema.org/CollegeOrUniversity> |

| | "http://dbpedia.org/resource/St.\_Thomas\_University" | <http://www.wikidata.org/entity/Q7591999> |

| | "http://dbpedia.org/resource/St.\_Thomas\_University" | 929770 |

| | "http://dbpedia.org/resource/St.\_Thomas\_University" | 513634996 |

| | "http://dbpedia.org/resource/St.\_Thomas\_University" | "St. Thomas University"@en |

| | "http://dbpedia.org/resource/St.\_Thomas\_University" | <http://dbpedia.org/resource/St.\_Thomas\_University\_(New\_Brunswick)> |

| | "http://dbpedia.org/resource/St.\_Thomas\_University" | <http://dbpedia.org/resource/St.\_Thomas\_University,\_Japan> |

| | "http://dbpedia.org/resource/St.\_Thomas\_University" | <http://dbpedia.org/resource/St.\_Thomas\_University\_(Florida)> |

| | "http://dbpedia.org/resource/St.\_Thomas\_University" | <http://en.wikipedia.org/wiki/St.\_Thomas\_University> |

| | "http://dbpedia.org/resource/St.\_Thomas\_University" | <http://en.wikipedia.org/wiki/St.\_Thomas\_University?oldid=513634996> |

| | "http://dbpedia.org/resource/St.\_Thomas\_University" | "9.47157e-06"^^<http://www.w3.org/2001/XMLSchema#float> |

| | "http://dbpedia.org/resource/St.\_Thomas\_University" | 9 |

| | "http://dbpedia.org/resource/St.\_Thomas\_University" | 5 |

--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

**Link Evaluation:**

Out of 180 ,175 of univer sity has a link.

The coverage of your external links (%) 97.7777

***The 25 university are listed below randomly from different text .***

|  |  |  |
| --- | --- | --- |
| http://dbpedia.org/resource/University\_of\_Hamburg</TextWithNodes> | | |
| http://dbpedia.org/resource/University\_of\_Wisconsin-Madison | | |
| "http://dbpedia.org/resource/University\_of\_Vienna," | |  |
| http://dbpedia.org/resource/University\_of\_Toronto | |  |
| http://dbpedia.org/resource/University\_of\_Cambridge | |  |
| http://dbpedia.org/resource/University\_of\_Wisconsin-Milwaukee | | |
| http://dbpedia.org/resource/Wayne\_State\_University |  |  |
| http://dbpedia.org/resource/University\_of\_California | |  |
| http://dbpedia.org/resource/University\_of\_Bolton | |  |
| http://dbpedia.org/resource/Johns\_Hopkins\_University | |  |
| http://dbpedia.org/resource/Cambridge\_University | |  |
| http://dbpedia.org/resource/Istanbul\_University |  |  |
| http://dbpedia.org/resource/Schiller\_University |  |  |
| http://dbpedia.org/resource/University\_of\_Hamburg; | |  |
| http://dbpedia.org/resource/University\_of\_Birmingham | |  |
| http://dbpedia.org/resource/Swansea\_University |  |  |
| http://dbpedia.org/resource/Brookes\_University |  |  |
| http://dbpedia.org/resource/University\_of\_Pittsburgh | |  |
| http://dbpedia.org/resource/Ruskin\_University |  |  |
| http://dbpedia.org/resource/University\_of\_Plymouth | |  |
| http://dbpedia.org/resource/University\_of\_North\_Carolina | | |
| "http://dbpedia.org/resource/University\_of\_Stirling," | |  |
| http://dbpedia.org/resource/University\_of\_the\_Arts | |  |
| http://dbpedia.org/resource/University\_of\_Worcester | |  |
| http://dbpedia.org/resource/Durham\_University |  |  |

Precision is number of links accurate to the context of the text .

here precision is 96% as they all mean to be universities which are actually being addressed to except university of arts which on dbpedia pints to school of arts. Thus resulting in the ambiguity of the links.

To solve this ambiguity extra measure can be taken .like check more of the context of the text along with annotation .plus the location if mentioned in the text can be used to filter the ambiguity