



The
University
Of
Sheffield.



FREyA: an Interactive Way of Querying Linked Data using Natural Language

Danica Damljanović, Milan Agatonović, Hamish Cunningham

contact: danica@dcs.shef.ac.uk

NATURAL LANGUAGE INTERFACES AND LINKED OPEN DATA

- ⊙ What are NLI's?
- ⊙ Challenges:
 - ⊙ NL understanding/grammar
 - ⊙ ambiguity/expressiveness
 - ⊙ knowledge structure/portability
 - small ontologies, large ontologies, multiple ontologies, Linked Open Data?

SHIFT IN CHALLENGES

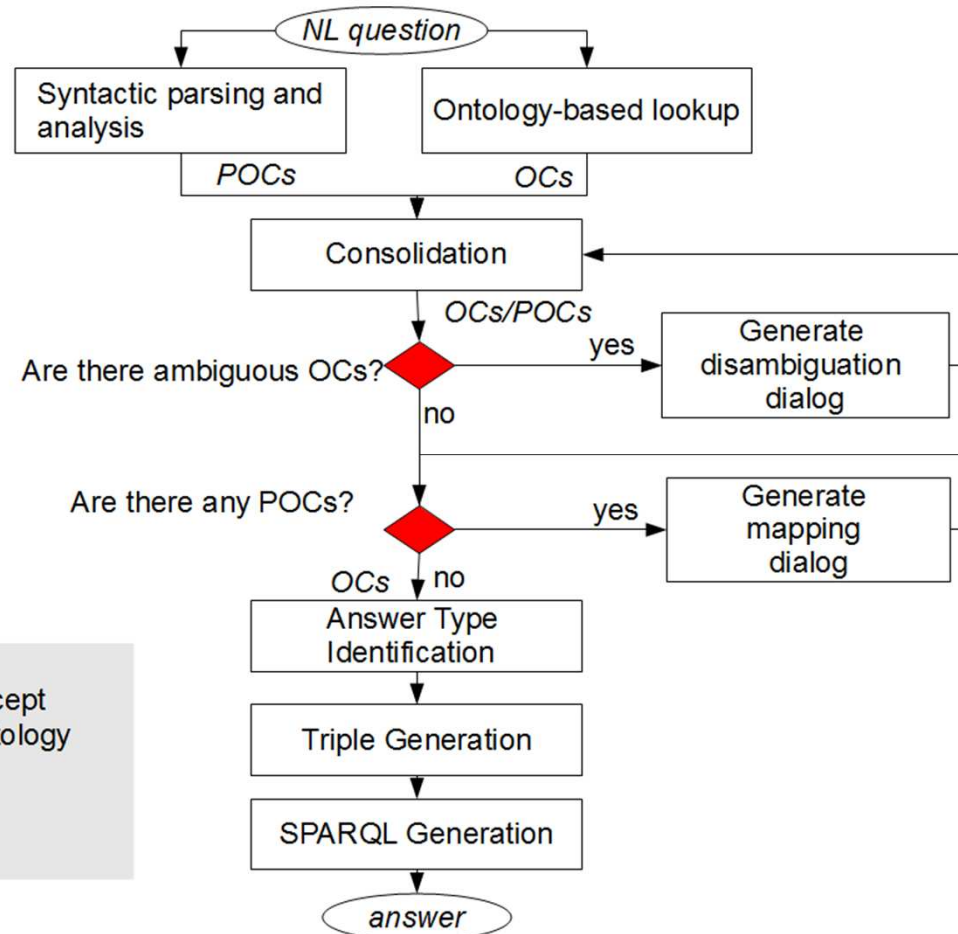
- ⊙ Portability 5 years ago vs. today?
 - ⊙ Heterogeneity, incompleteness, redundancy

Closed-domain	Open-domain
Find the answer	Disambiguate the answer
Language ambiguity	Data ambiguity
Increase recall (Wordnet, etc.)	Increase precision



- ⊙ Can one system support both?

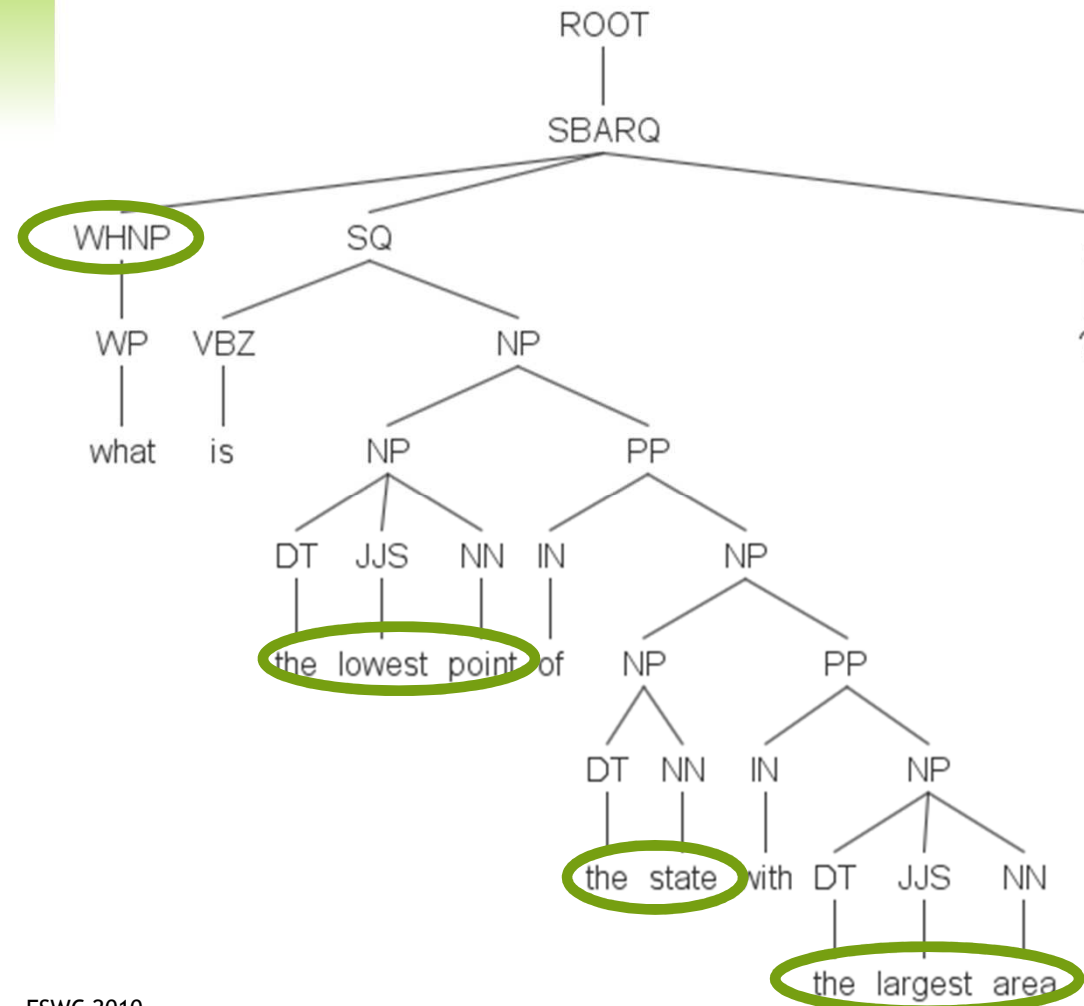
FREYA WORKFLOW



Legend:

OC: Ontology Concept
POC: Potential Ontology Concept

FINDING POCs



FINDING OCs

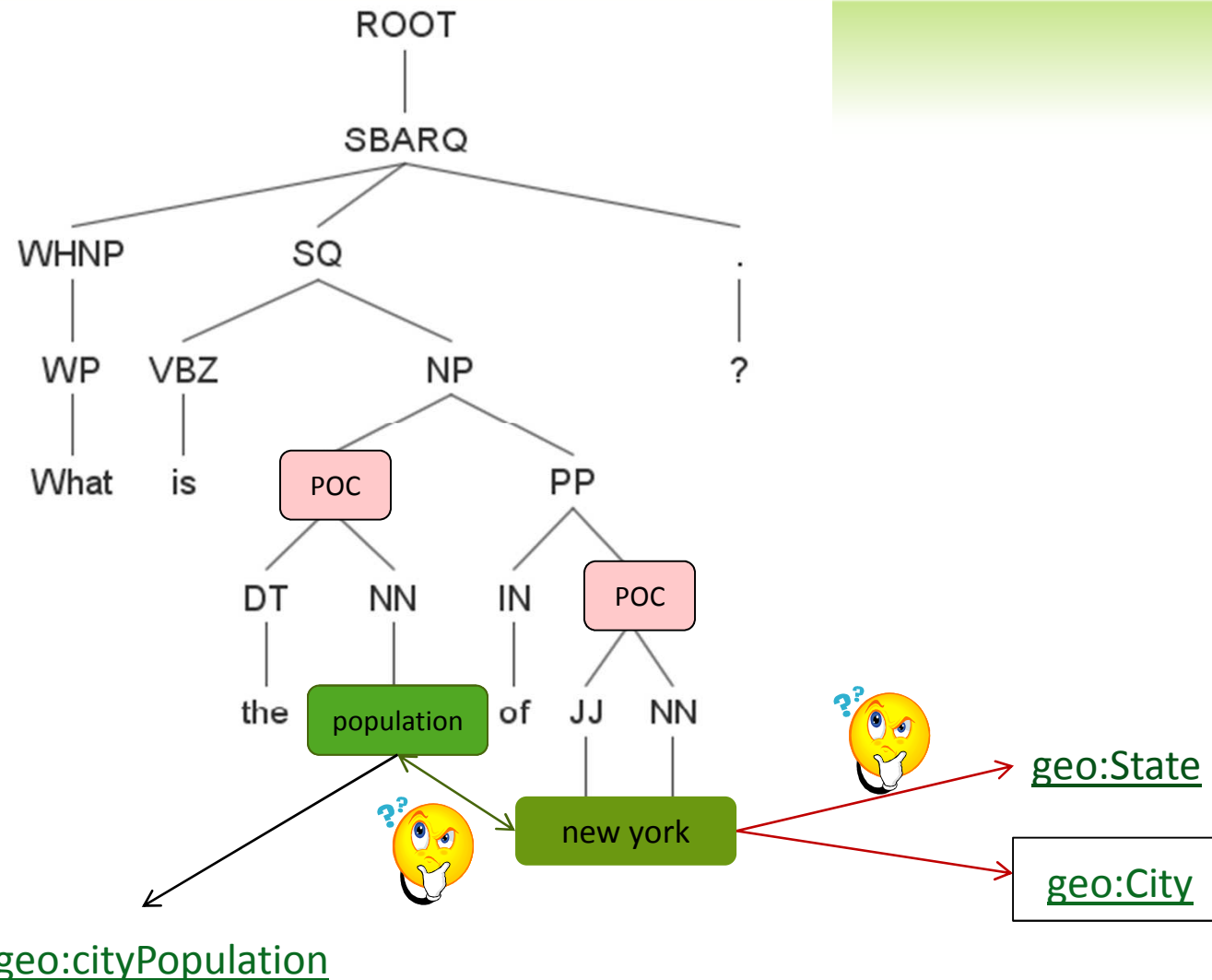
what is the lowest point of the **state** with the largest area?

The screenshot shows a web application interface for finding Open Classes (OCs). The interface includes a search bar with the text "Lookup", a table of search criteria, and a button to open a search tool.

URI	Value	Action
URI	http://www.mooney.net/geo#State	X
type	class	X
		X

▶ Open Search & Annotate tool

MAPPING POC TO OCs



NEW YORK IS A CITY

Query: What is the population of new york?

Submit

I struggle with new york. Is 'new york' related to:

city

state

Query: What is the population of new york?

Submit

I struggle with population. Is 'population' related to:

city population

state

is city of

none



NEW YORK IS A STATE

Query:

I struggle with new york. Is 'new york' related to:

city
state



Query:

I struggle with population. Is 'population' related to:

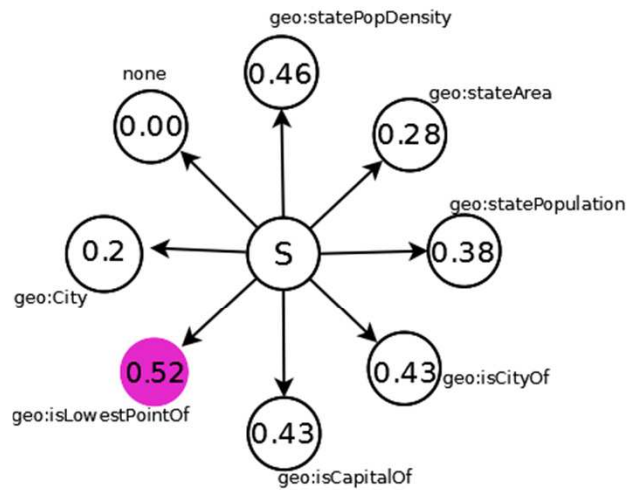
state population
state area
state pop density
none



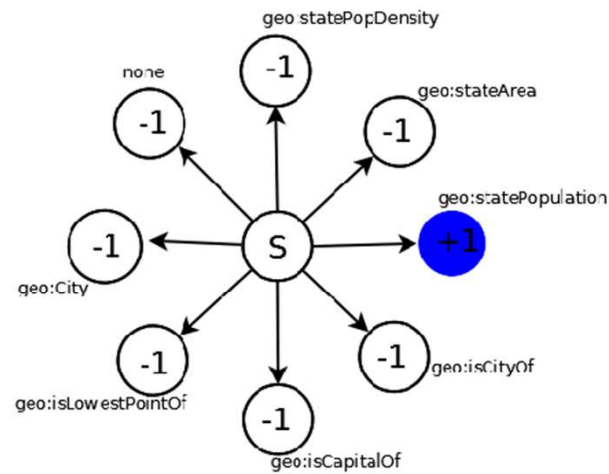
state (1) → new york (state population) → 49100

IF		THEN	
POC	OC (context)	candidate OC	function
new york		geo:State	-
new york		geo:City	-
population	geo:State	geo:statePopulation	-
population	geo:City	geo:cityPopulation	-

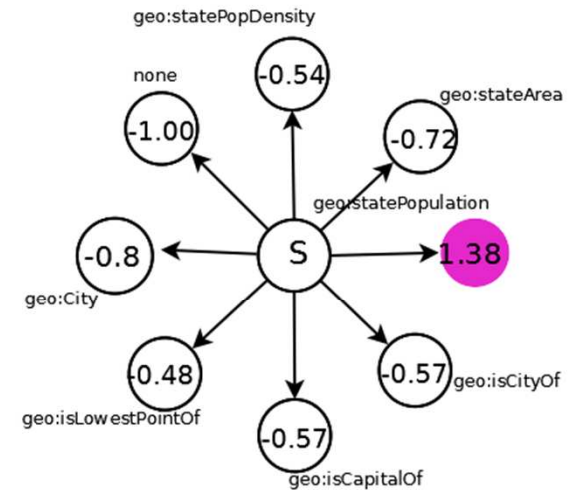
LEARNING



a) INITIAL RANKING



b) REINFORCEMENT BASED ON THE USER
SELECTING geo:statePopulation



c) RANKING AFTER THE USER
SELECTS geo:statePopulation

QUERYING LINKED DATA WITH FREYA: THE USUAL CYCLE

```
for i=1 to n {  
  Initialise the system using dataset  $A_i$   
  (forceDialog)  
  Train the system by asking questions  
  Save learningModel $_i$   
}  
Intialise the system (automatic mode) by  
loading learningModel $_i$ ,  $i=1,n$   
connect to the repository containing all  $A_i$   
datasets, where  $i=1, n$ 
```

- ◎ Initialisation
- ◎ System performance
 - ◎ Precision, recall, f-measure using MusicBrainz and DBPedia datasets
 - ◎ Mean Reciprocal Rank to assess the effect of learning mechanism
- ◎ Analysis of failures

INITIALISATION AND THE DATASETS SIZE

	MusicBrainz	DBpedia
explicit statements	14 926 841	328 318 709
statements	19 202 664	372 110 845
entities	5 490 237	96 515 478
SPARQL queries	30	361623
initialisation time	1380s (0.38h)	182779s (50.77h)

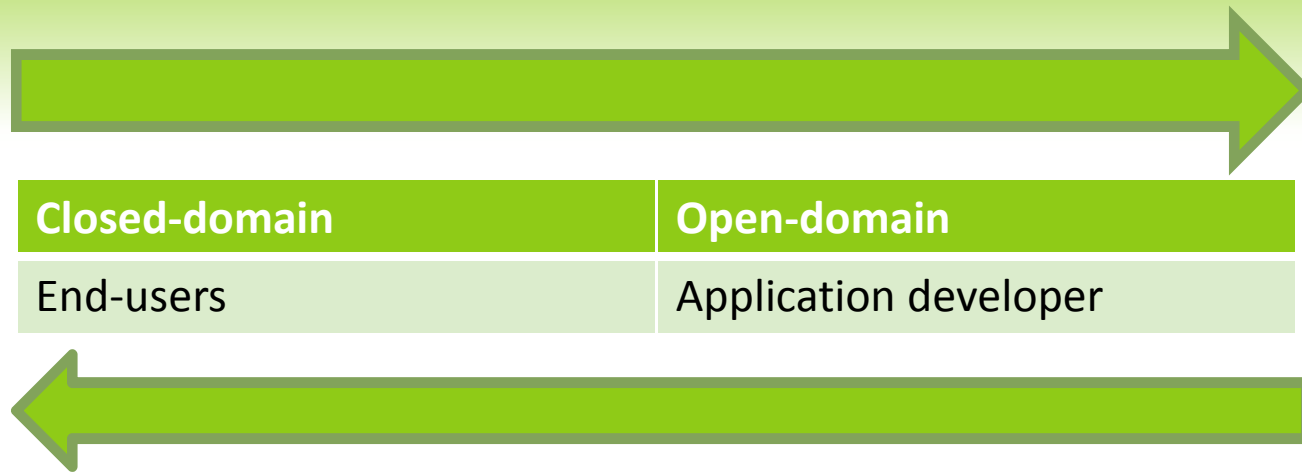
RESULTS: F-MEASURE STATISTICS

	MusicBrainz		DBPedia	
	Training	Testing	Training	Testing
Precision	0.75/0.77	0.66/0.8	0.74/0.85	0.49/0.63
Recall	0.66/0.68	0.54/0.66	0.58/0.66	0.42/0.54
F-measure	0.70/0.74	0.59/0.71	0.67/0.72	0.45/0.58
not supported questions	6	9	11	7
reformulated questions	1	6	4	6
avg #dialogs per question	3.4	3.65	2.7	2.85
partially correct questions	1	1	3	12

LEARNING

	MusicBrainz	DBPedia
	untrained/trained	untrained/trained
MRR	0.63/0.68	0.52/0.54

CONCLUSION



- ⊙ Output:
 - ⊙ Correct answer OR
 - ⊙ Identifying the flaws in the data?
- ⊙ Ranking/disambiguation algorithms to improve MRR

THANK YOU FOR YOUR ATTENTION! QUESTIONS?



Thanks to Ivan Peikov from Ontotext who helped with the configuration of OWLIM which was necessary for performing the experiments reported in this paper.

Contact: danica@dcs.shef.ac.uk

DEMO

United States geography:

<http://gate.ac.uk/sale/dd/movies/mooney/html/freya.html>

MusicBrainz:

<http://gate.ac.uk/sale/dd/movies/mb/html/musicbrainz.html>

DBpedia:

<http://gate.ac.uk/sale/dd/movies/dbpedia/html/small/freya.html>