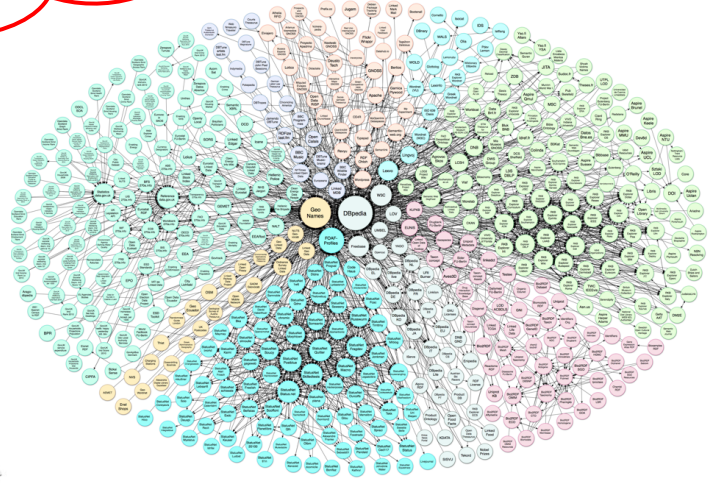
HAWK - Hybrid Question Answering

Ricardo Usbeck, Axel-Cyrille Ngonga Ngomo, Lorenz Bühmann & Christina Unger
University Leipzig & University Bielefeld

Motivation

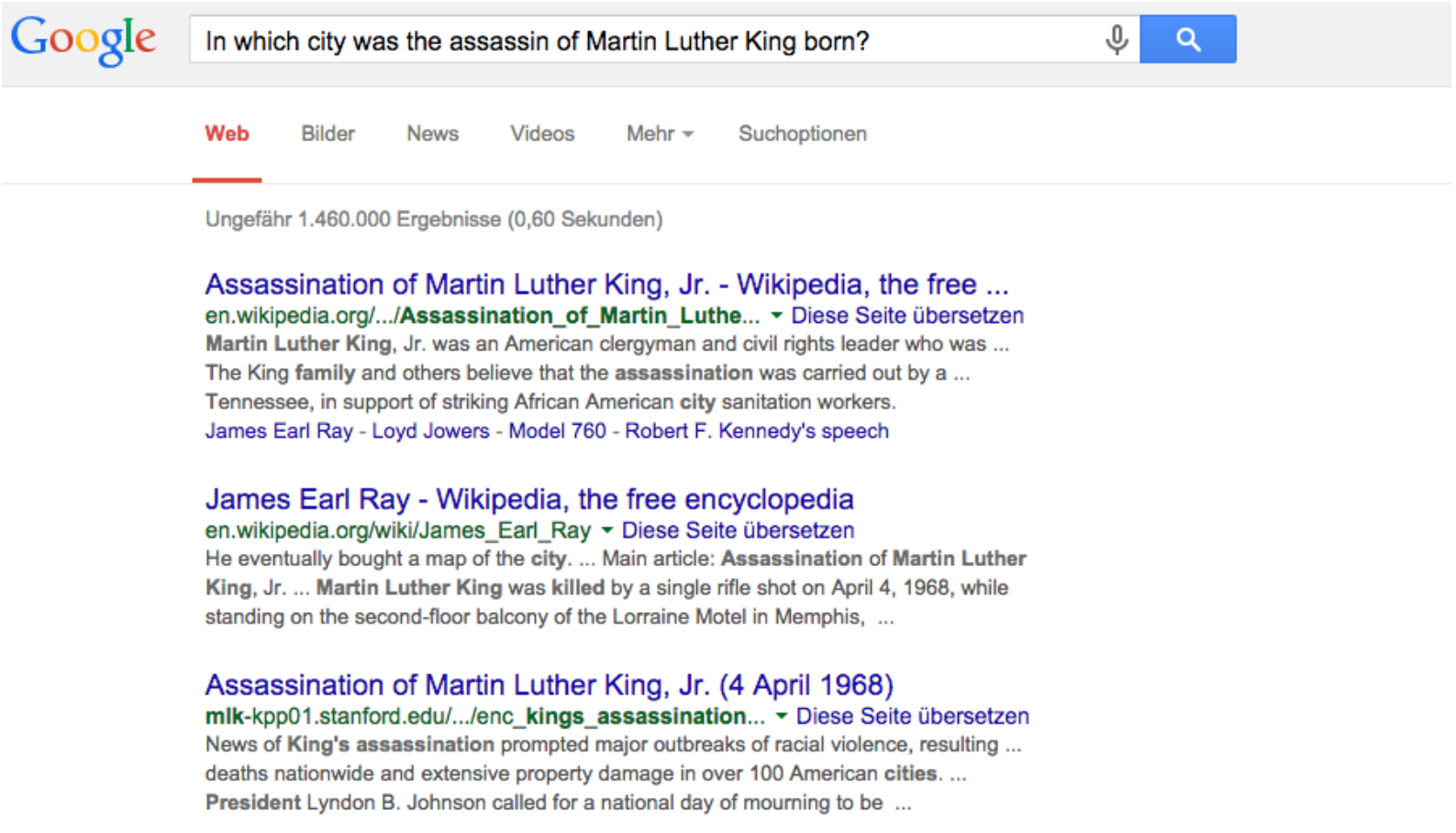
Document Web

In which city was the assassin of Martin Luther King born?



Linked Data Web

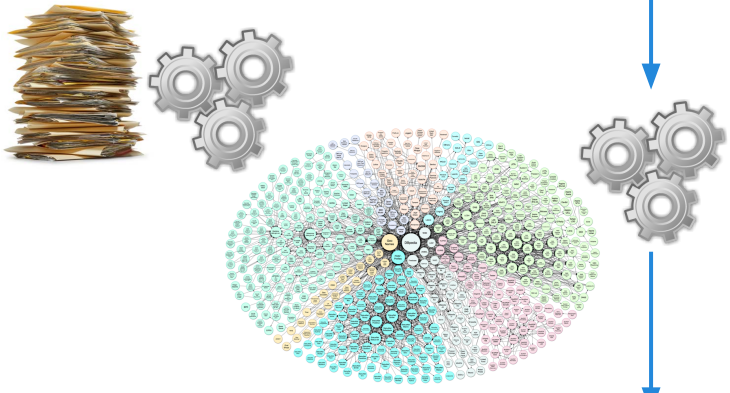
Motivation


Google
In which city was the assassin of Martin Luther King born?
Web
Bilder
News
Videos
Mehr
Suchoptionen
Ungefähr 1.460.000 Ergebnisse (0,60 Sekunden)
Assassination of Martin Luther King, Jr. - Wikipedia, the free ...
en.wikipedia.org/.../Assassination_of_Martin_Luthe... Diese Seite übersetzen
Martin Luther King, Jr. was an American clergyman and civil rights leader who was ...
The King family and others believe that the assassination was carried out by a ...
Tennessee, in support of striking African American city sanitation workers.
James Earl Ray - Loyd Jowers - Model 760 - Robert F. Kennedy's speech
James Earl Ray - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/James_Earl_Ray Diese Seite übersetzen
He eventually bought a map of the city. ... Main article: Assassination of Martin Luther King, Jr. ... Martin Luther King was killed by a single rifle shot on April 4, 1968, while standing on the second-floor balcony of the Lorraine Motel in Memphis, ...
Assassination of Martin Luther King, Jr. (4 April 1968)
mlk-kpp01.stanford.edu/.../enc_kings_assassination... Diese Seite übersetzen
News of King's assassination prompted major outbreaks of racial violence, resulting ...
deaths nationwide and extensive property damage in over 100 American cities. ...
President Lyndon B. Johnson called for a national day of mourning to be ...

Hawk - Hybrid Question Answering

Query: In which city was the assassin of Martin Luther King born?

User Query

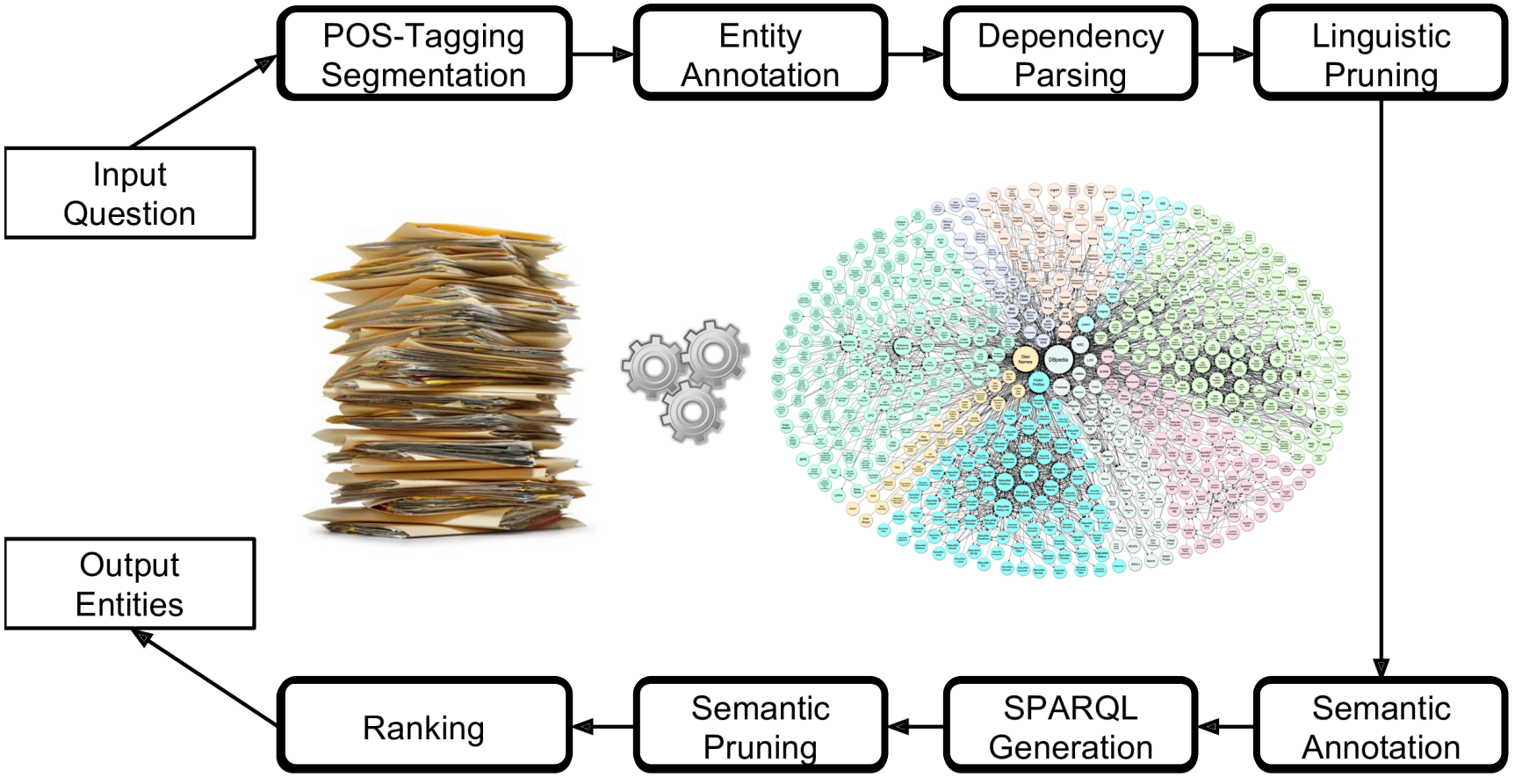


Hybrid Question Answering Framework

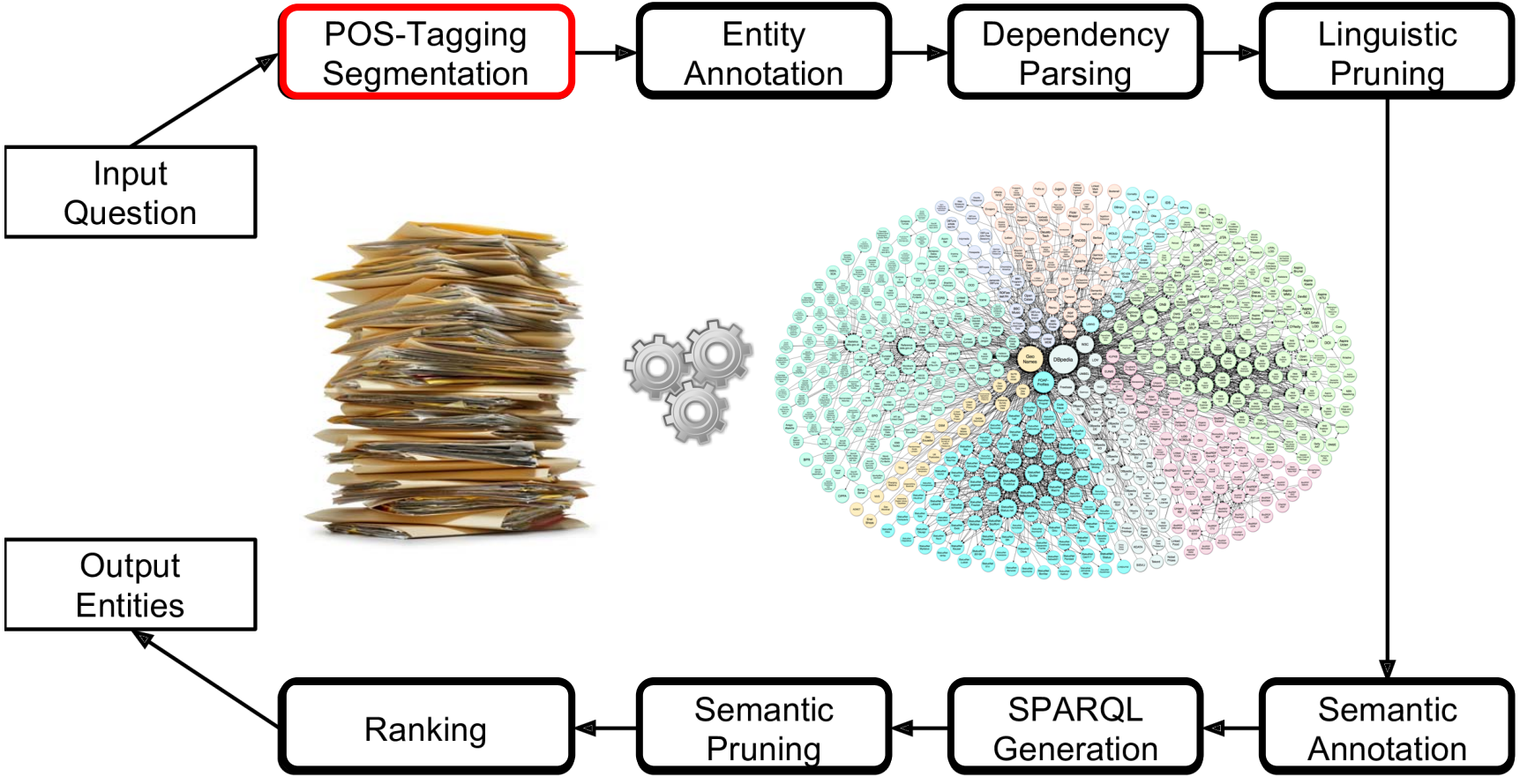
<http://dbpedia.org/resource/Alton,_Illinois>

Hybrid Answer

Hawk - Hybrid Question Answering



Hawk - Hybrid Question Answering



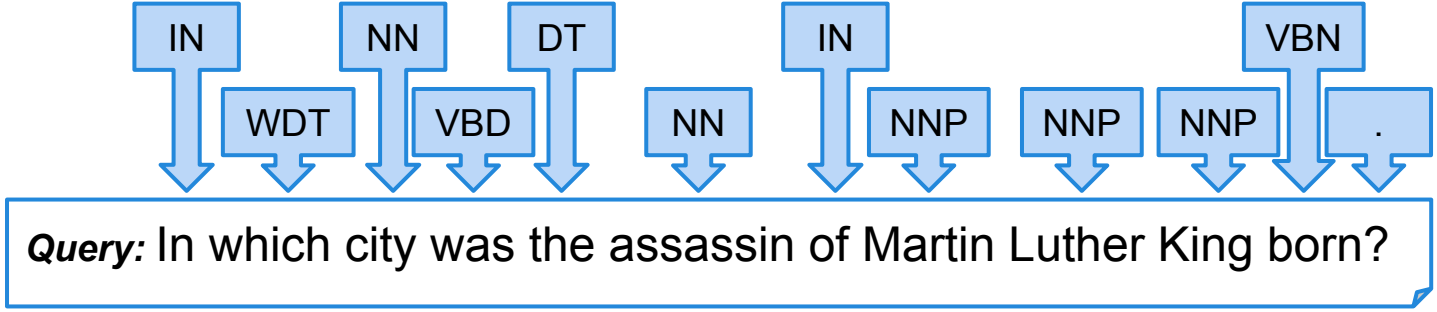
Hawk -
POS Tagging / Segmentation

Query: In which city was the assassin of Martin Luther King born?

- Powered by the ClearNLP API
- Segmentation based on WhiteSpace Tokenizer

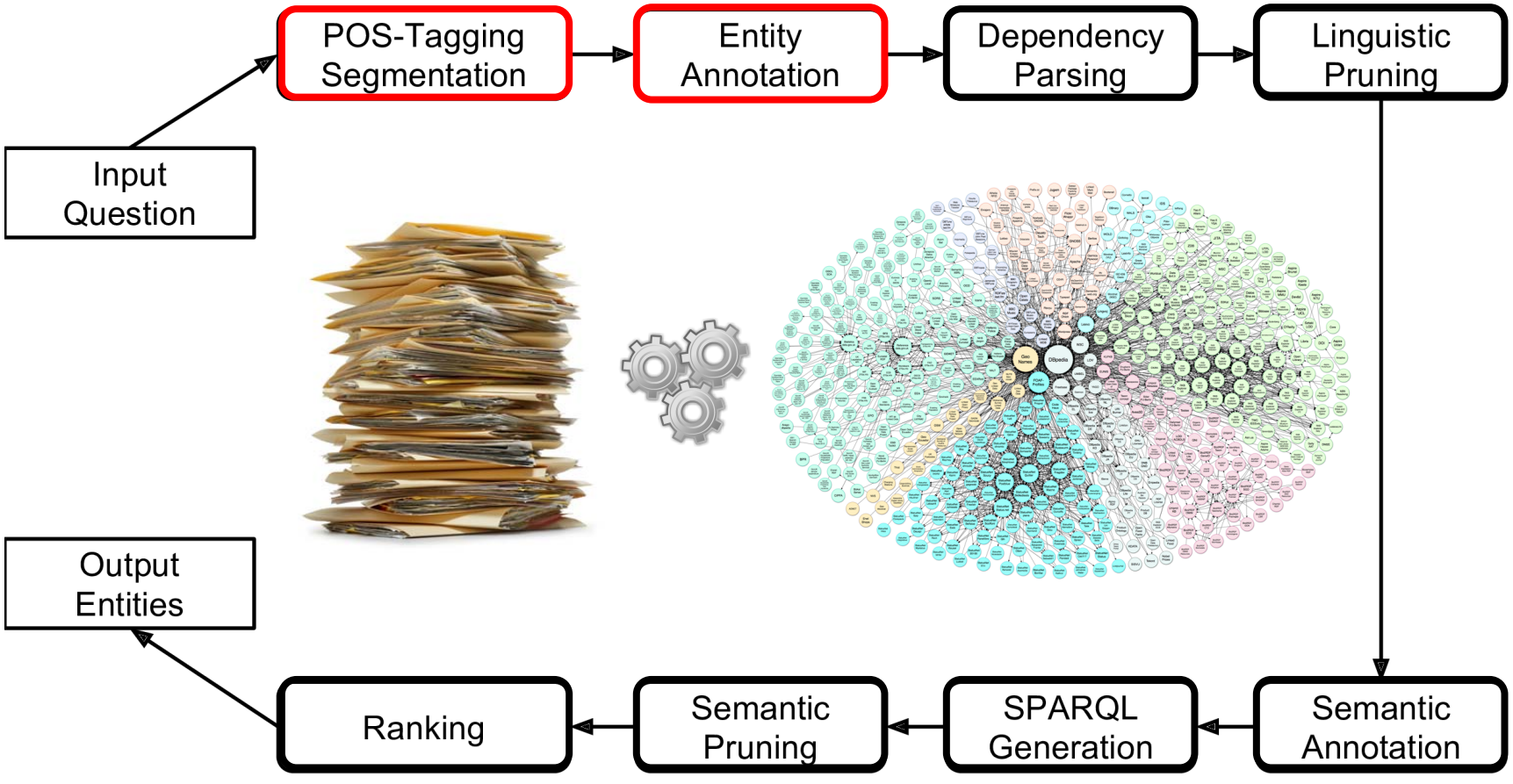
Hawk -

POS Tagging / Segmentation

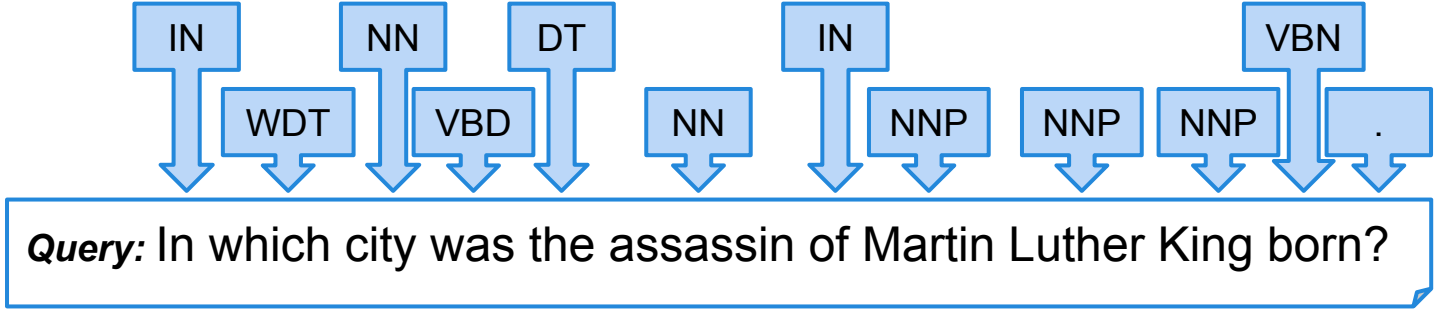


- Powered by the ClearNLP API
- Segmentation based on WhiteSpace Tokenizer

Hawk - Hybrid Question Answering

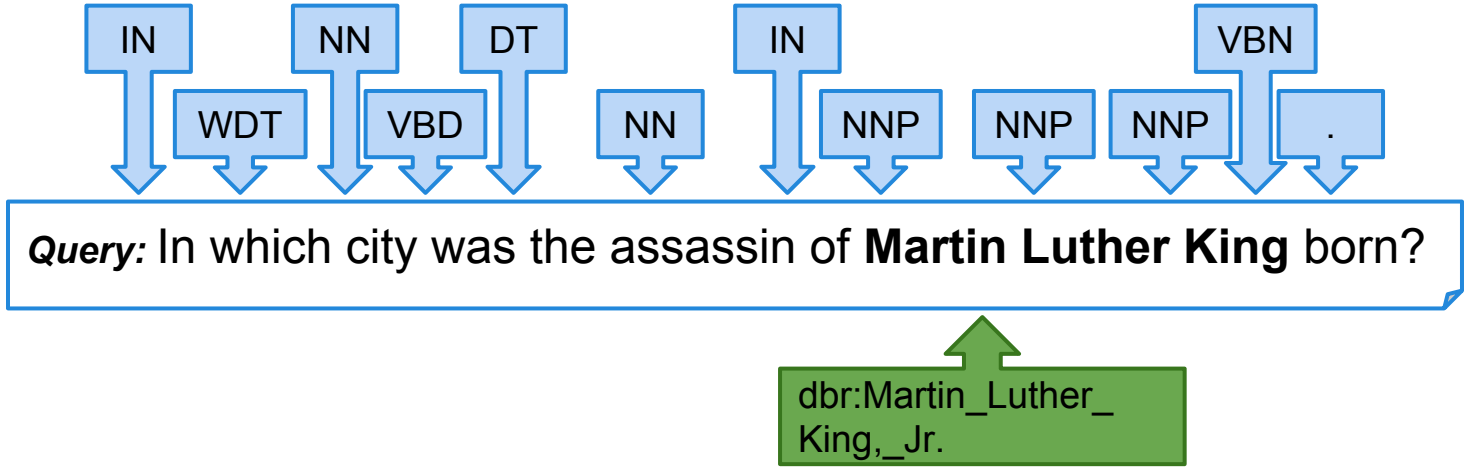


Hawk -
Entity Annotation



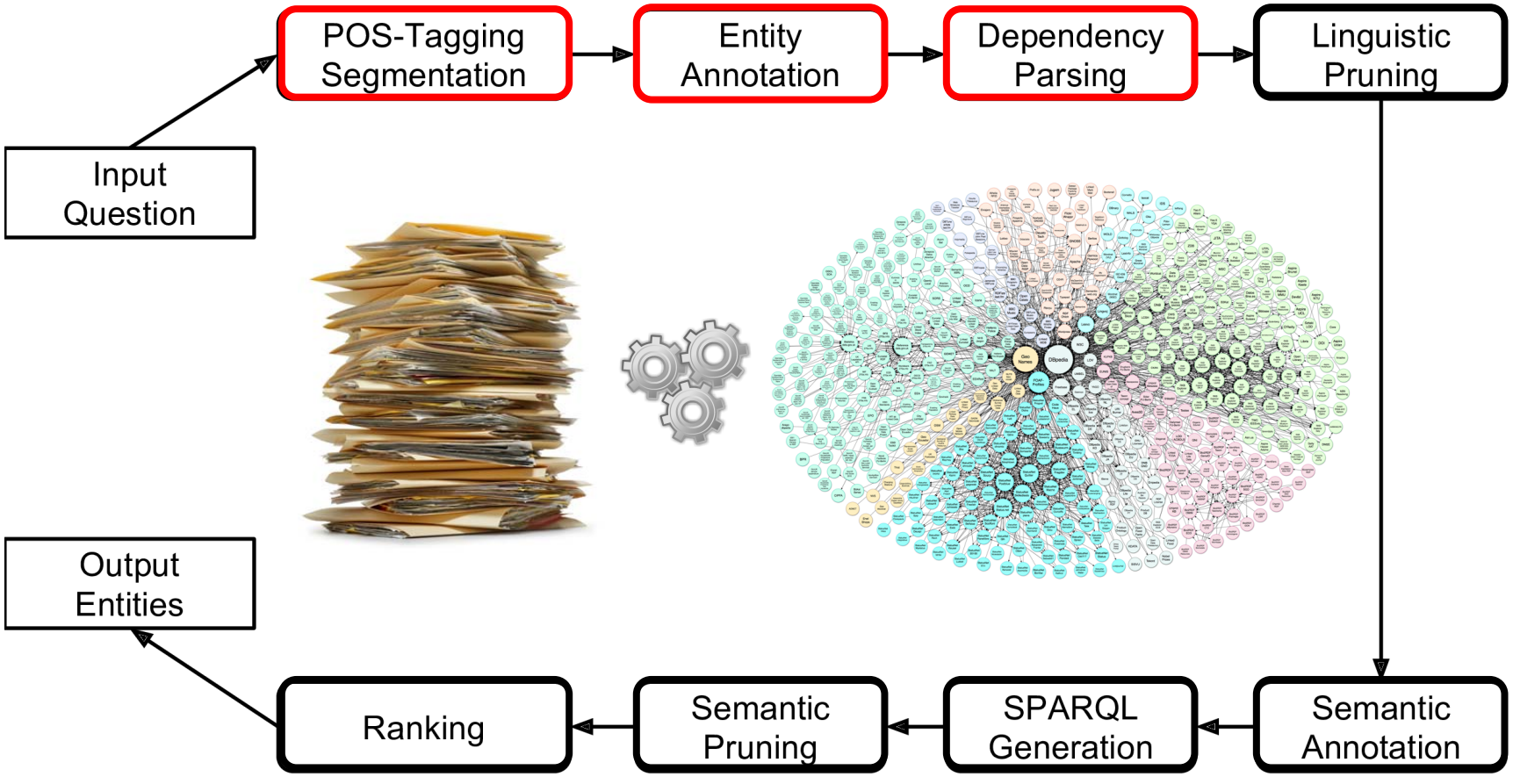
- Analysed the influence of 4 different annotation systems

Hawk -
Entity Annotation

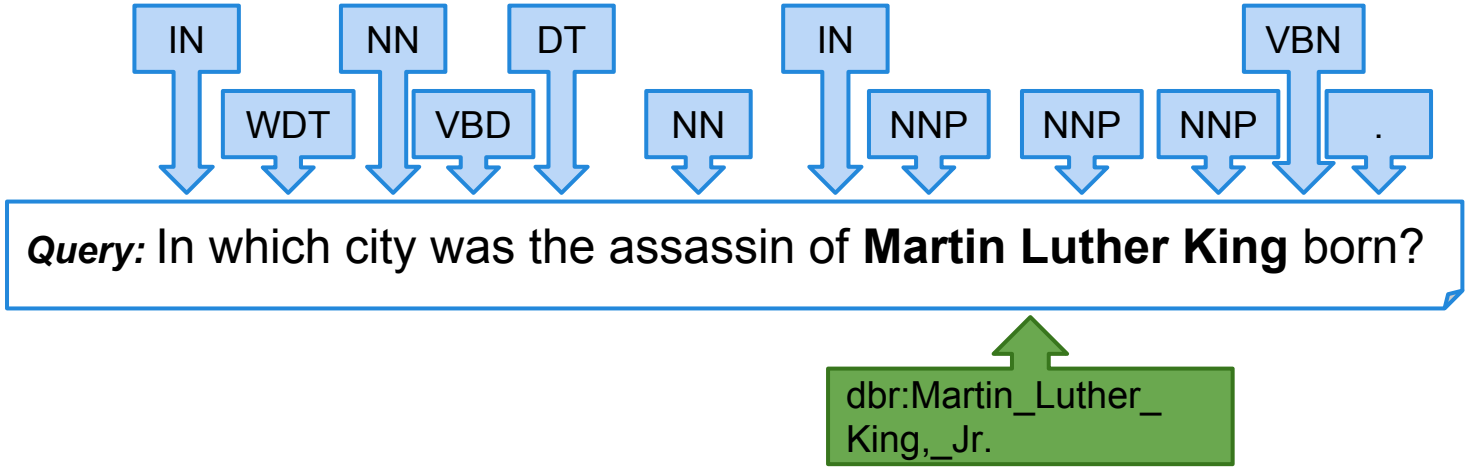


- Analysed the influence of 4 different annotation systems
- Tested against the union of annotation systems as well as an “optimal” annotator

Hawk - Hybrid Question Answering

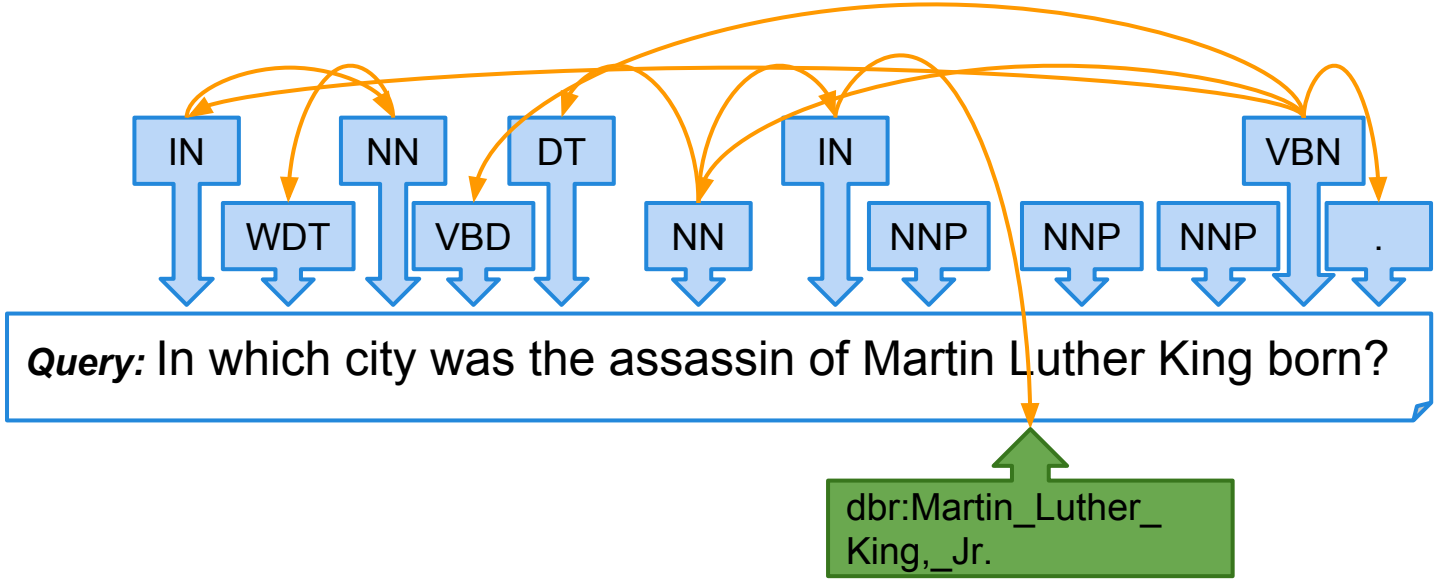


Hawk -
Dependency Parsing



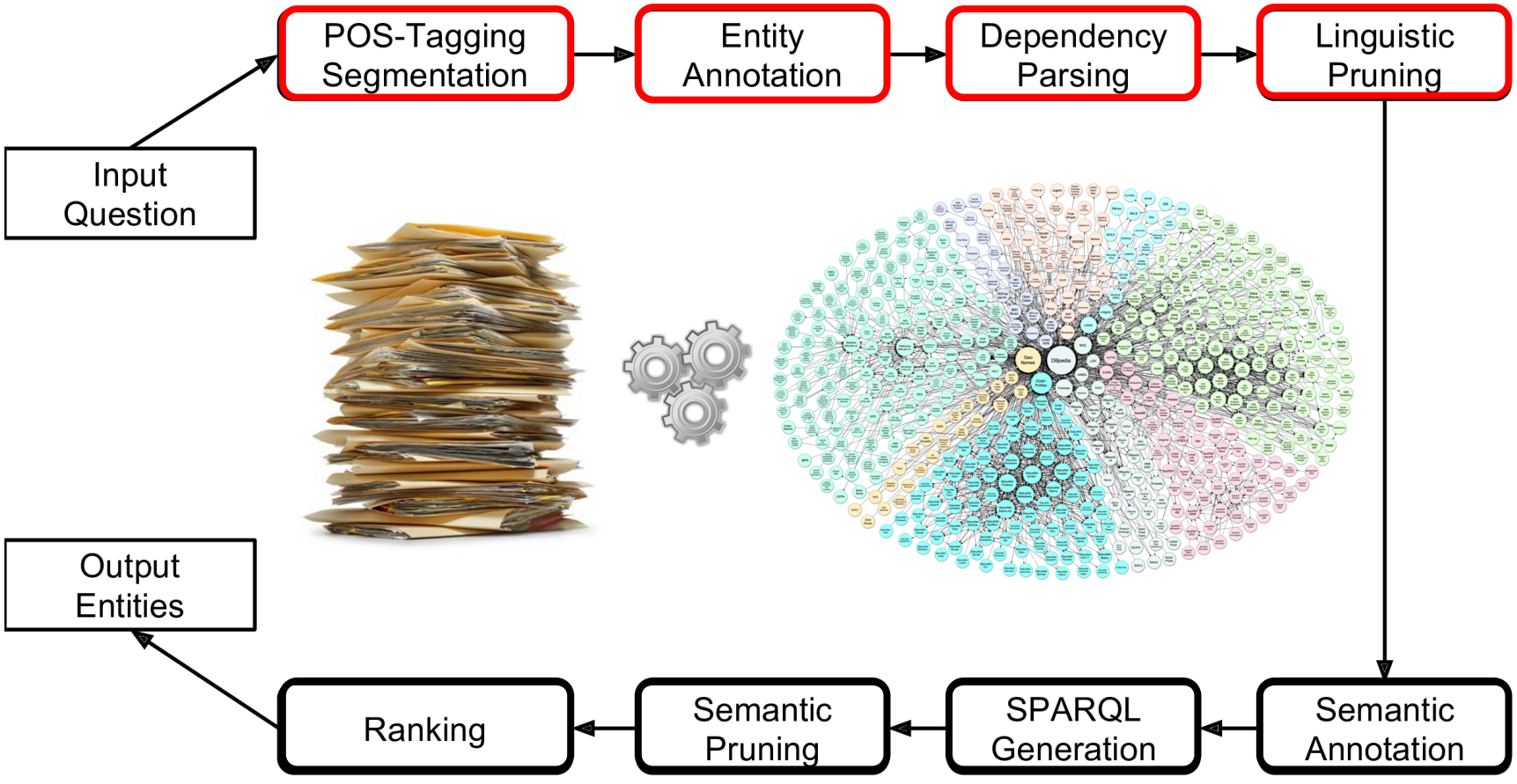
- Powered by the ClearNLP API
- Predicate-Argument Tree

Hawk -
Dependency Parsing

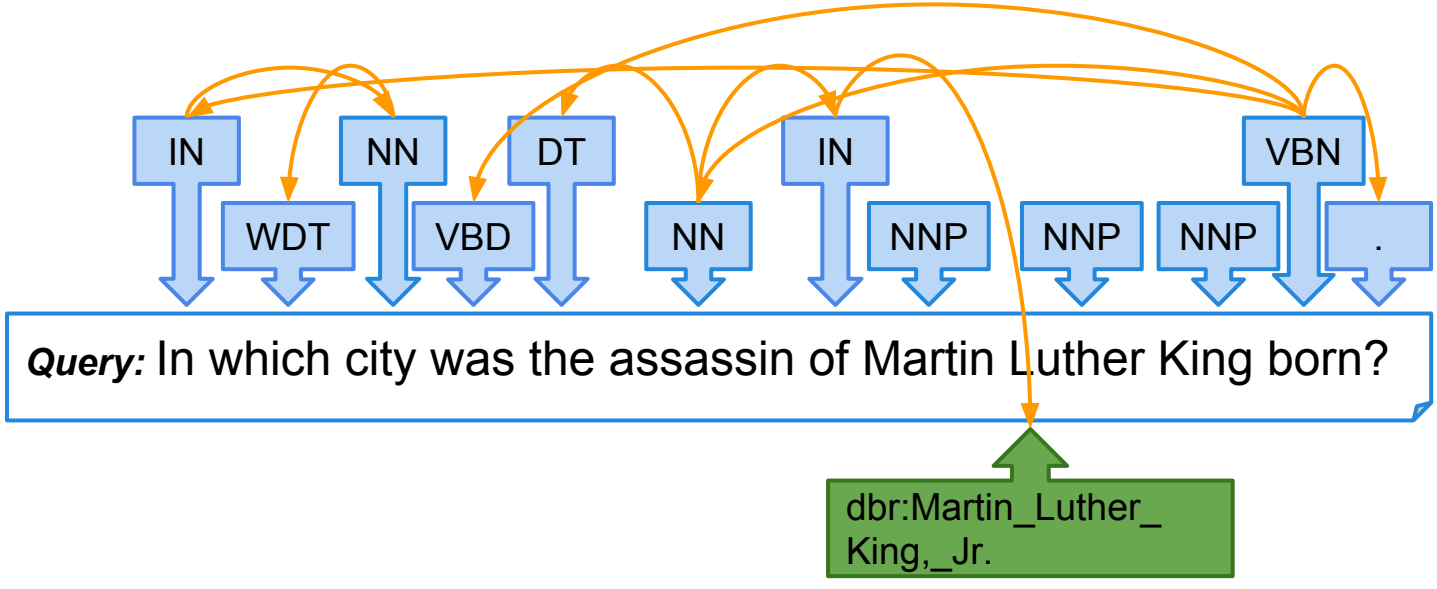


- Powered by the ClearNLP API
- Predicate-Argument Tree

Hawk - Hybrid Question Answering

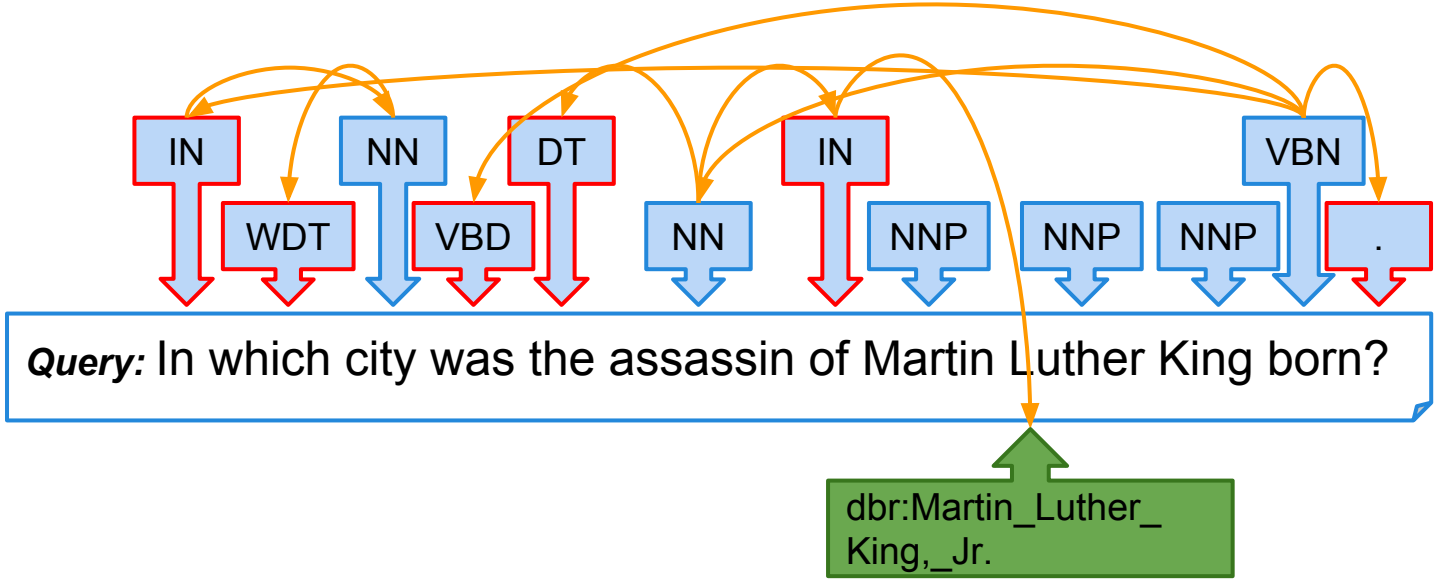


Hawk -
Linguistic Pruning



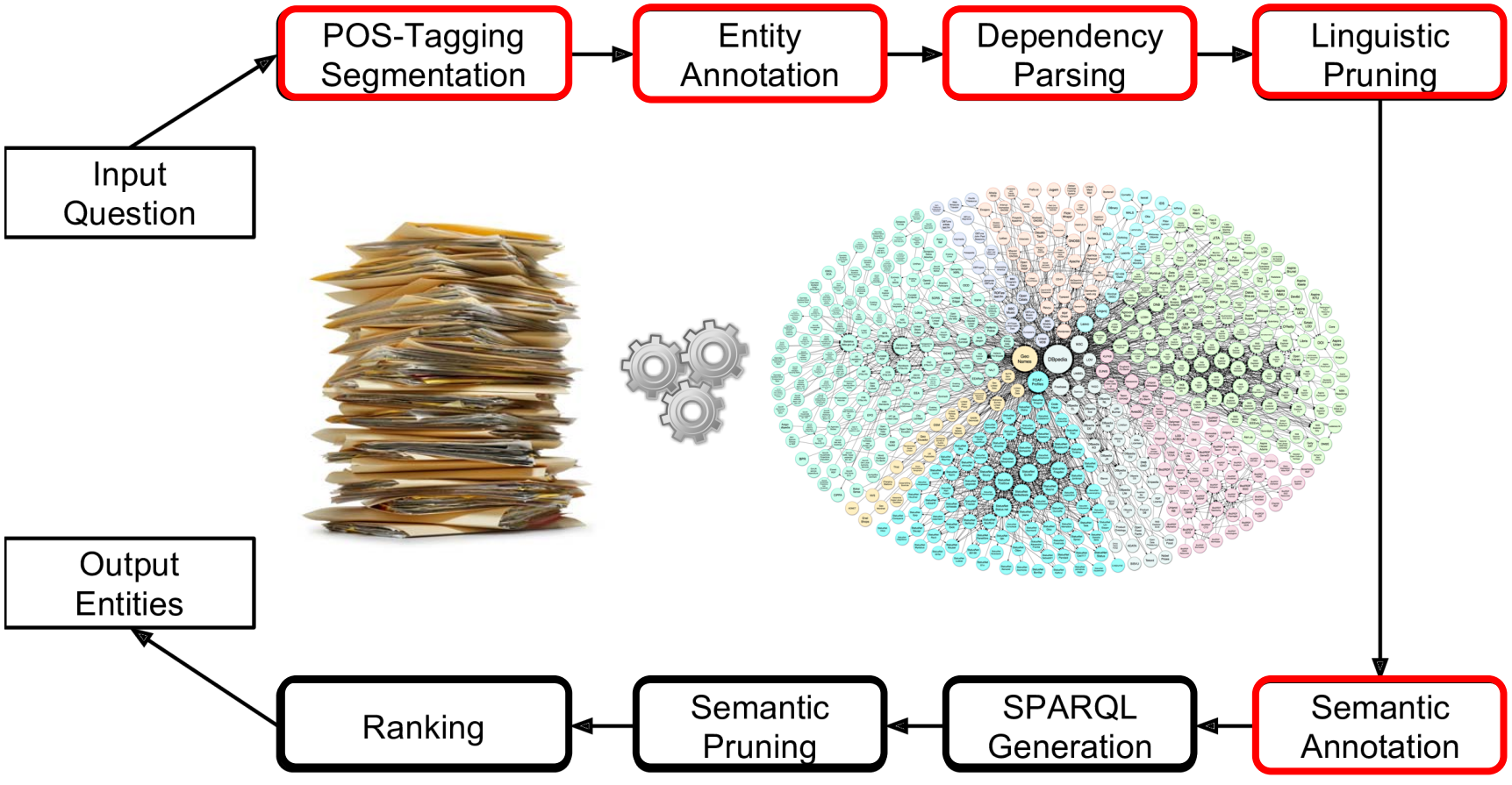
- Heuristic-based pruning of non-semantic words

Hawk -
Linguistic Pruning



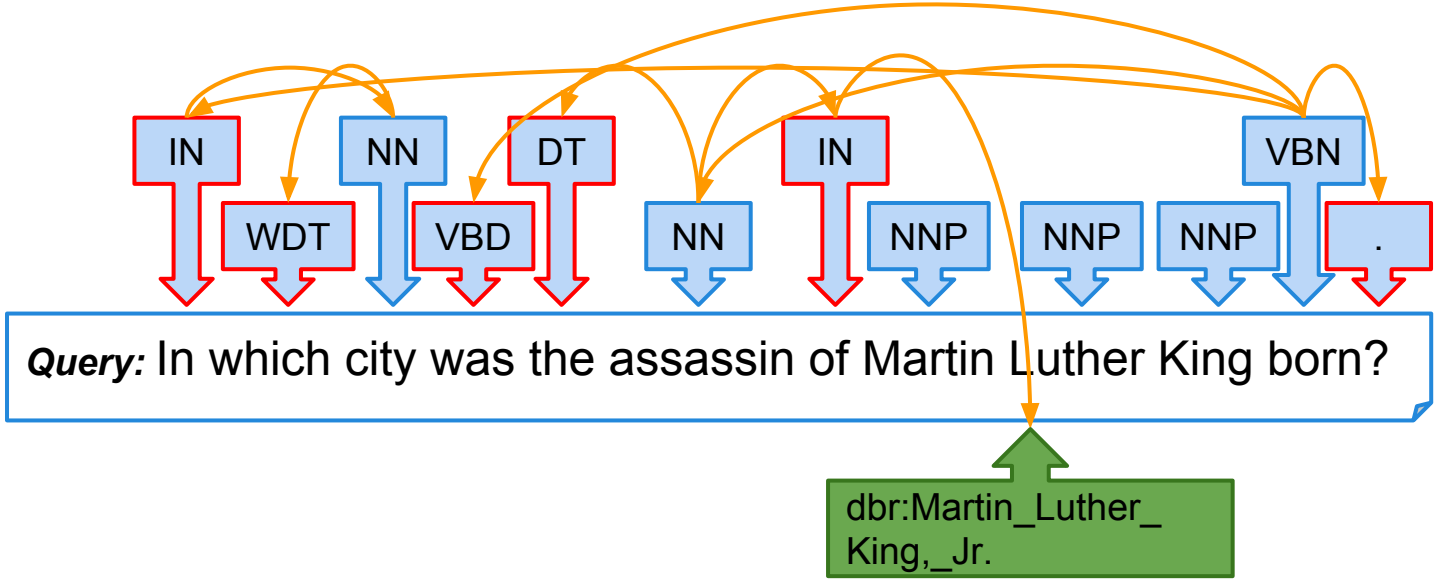
- Heuristic-based pruning of non-semantic words

Hawk - Hybrid Question Answering



Hawk -

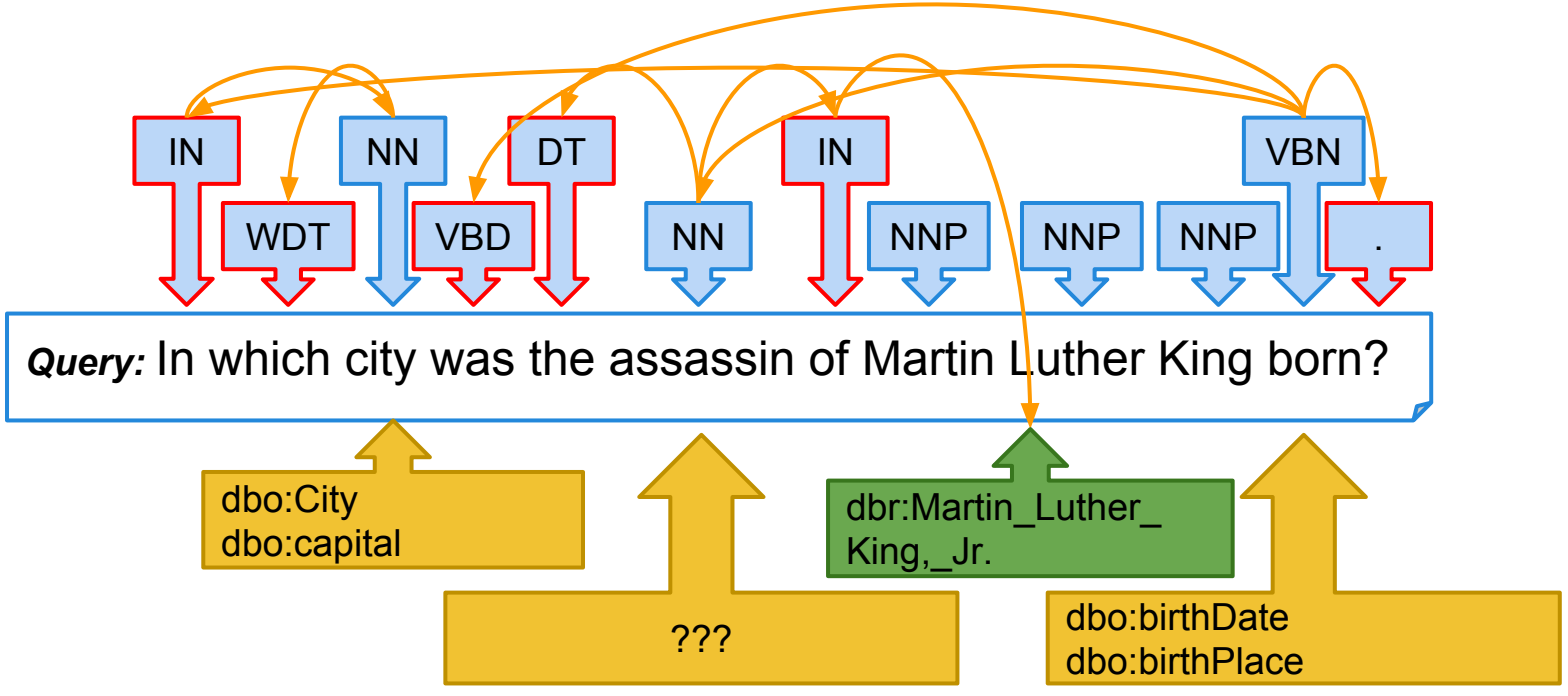
Semantic Annotation



- Dictionary-based surface form search
- Ontology-based search for classes and properties

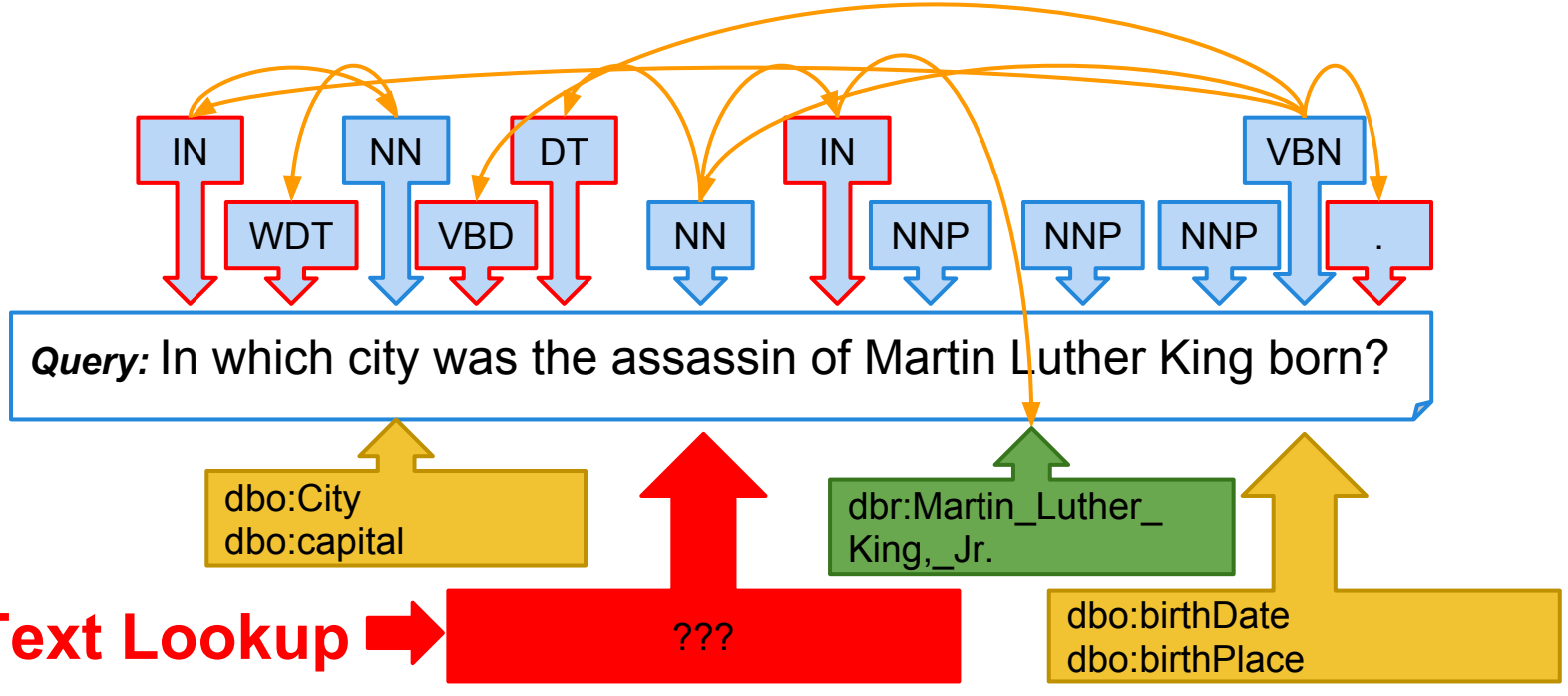
Hawk -

Semantic Annotation



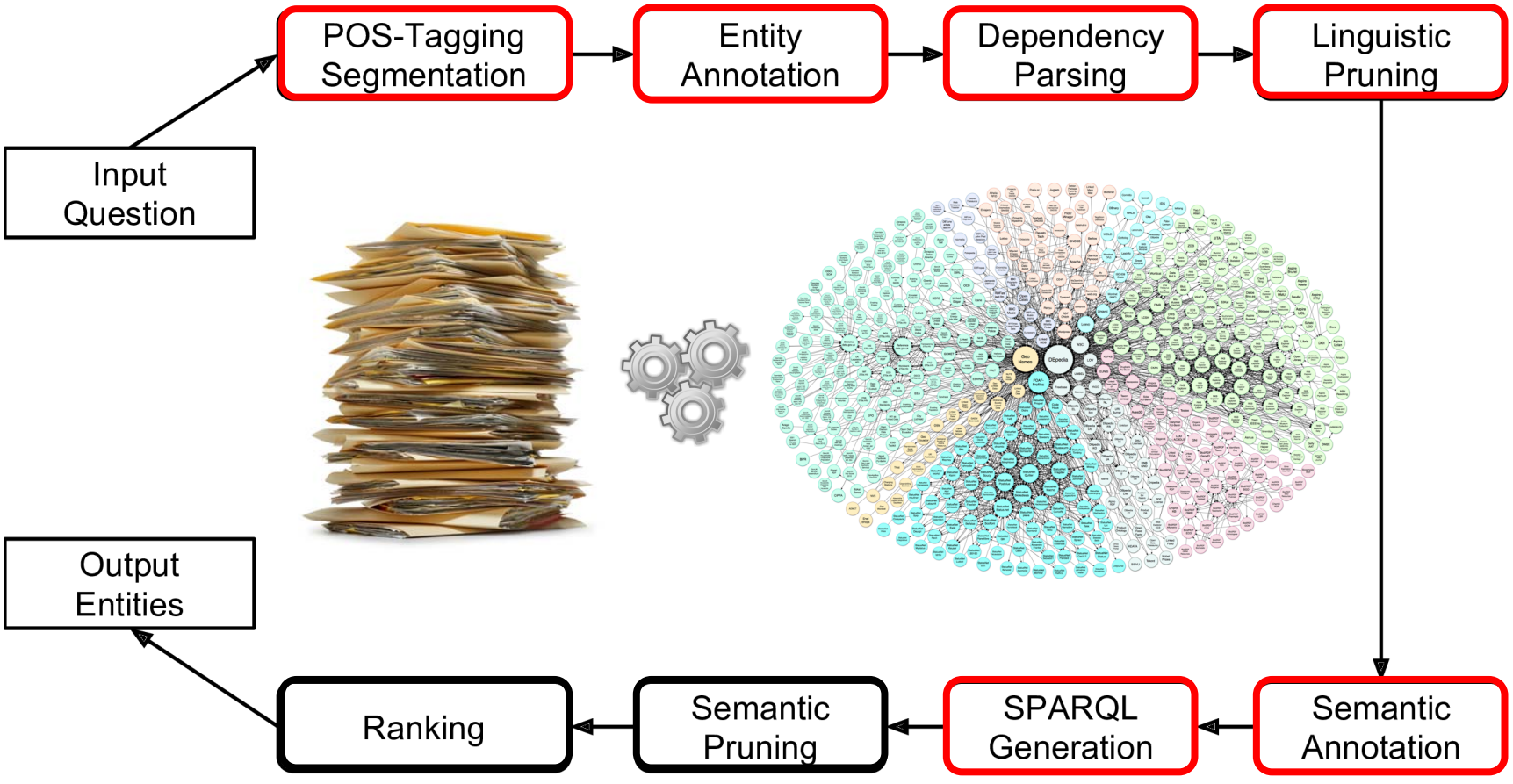
- Dictionary-based surface form search
- Ontology-based search for classes and properties

Hawk -
Semantic Annotation

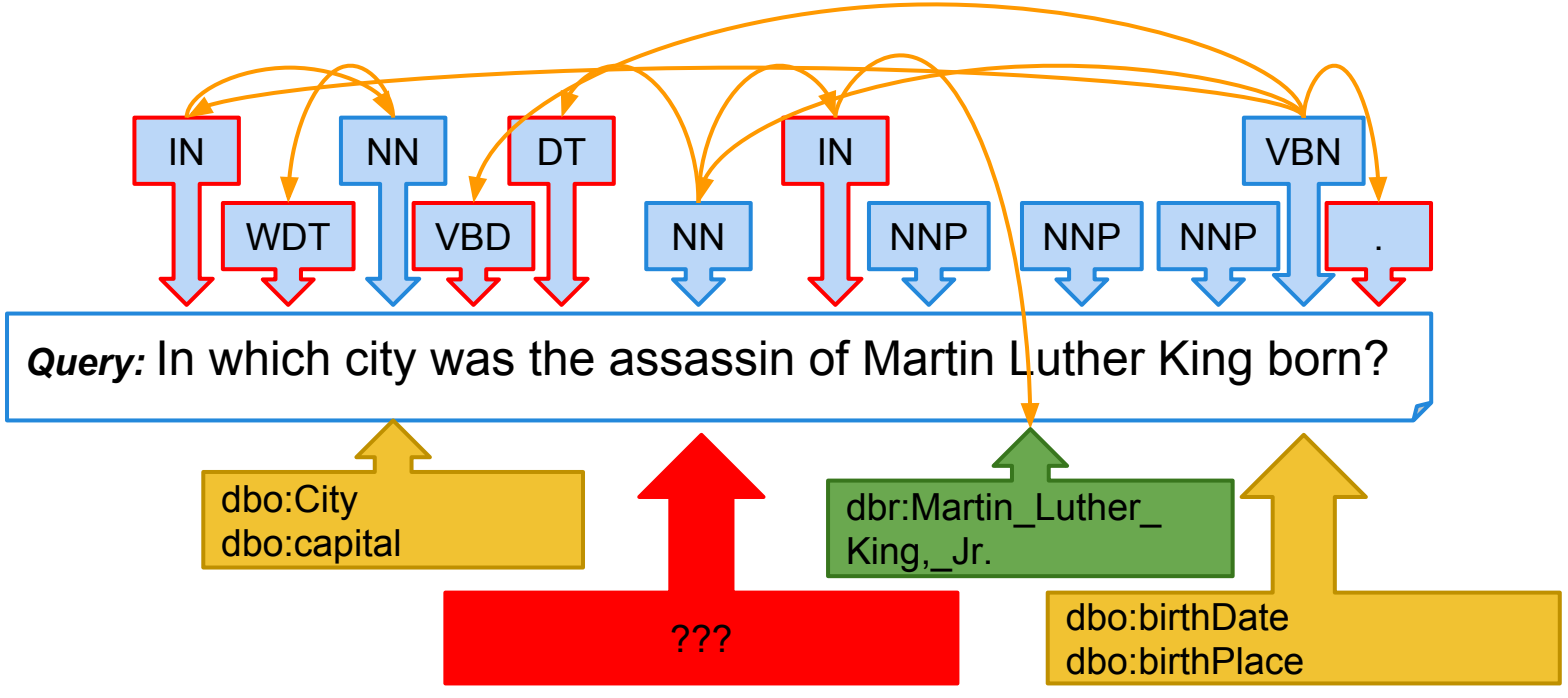


- Dictionary-based surface form search
- Ontology-based search for classes and properties

Hawk - Hybrid Question Answering

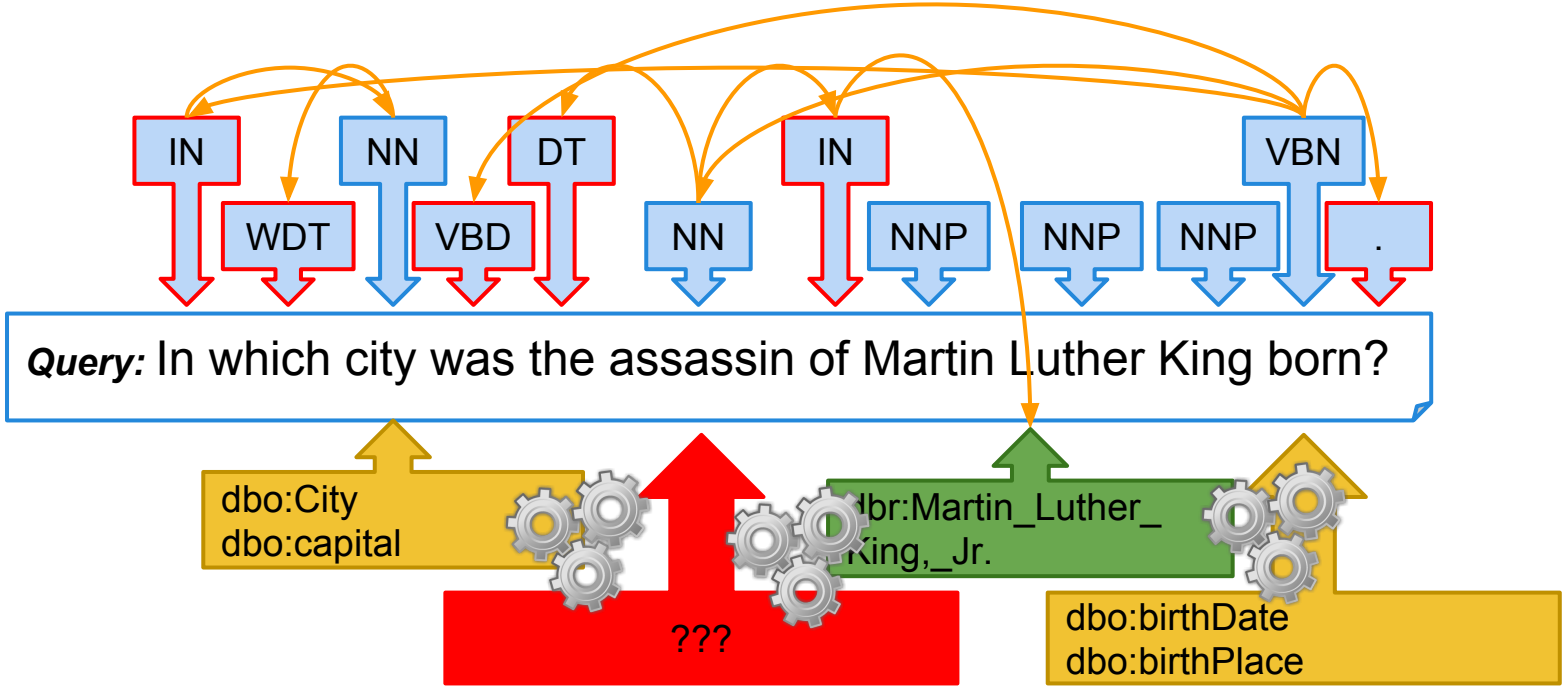


Hawk -
SPARQL Generation



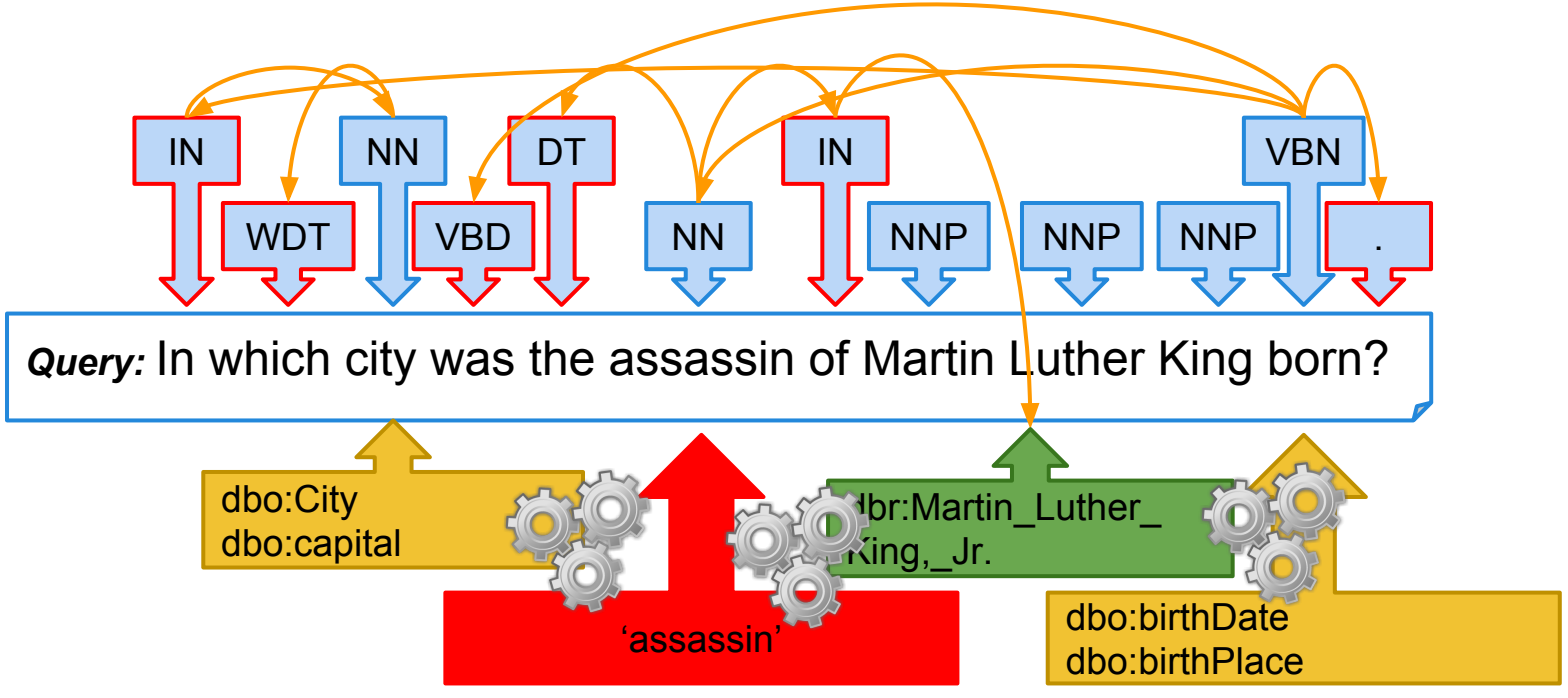
- Build for each node a triple pattern (in subject, predicate and object position) or
- Use Cross-Product of *N* free variables in SPARQL query
- Full-Text Filter based on CBD of entity

Hawk -
SPARQL Generation



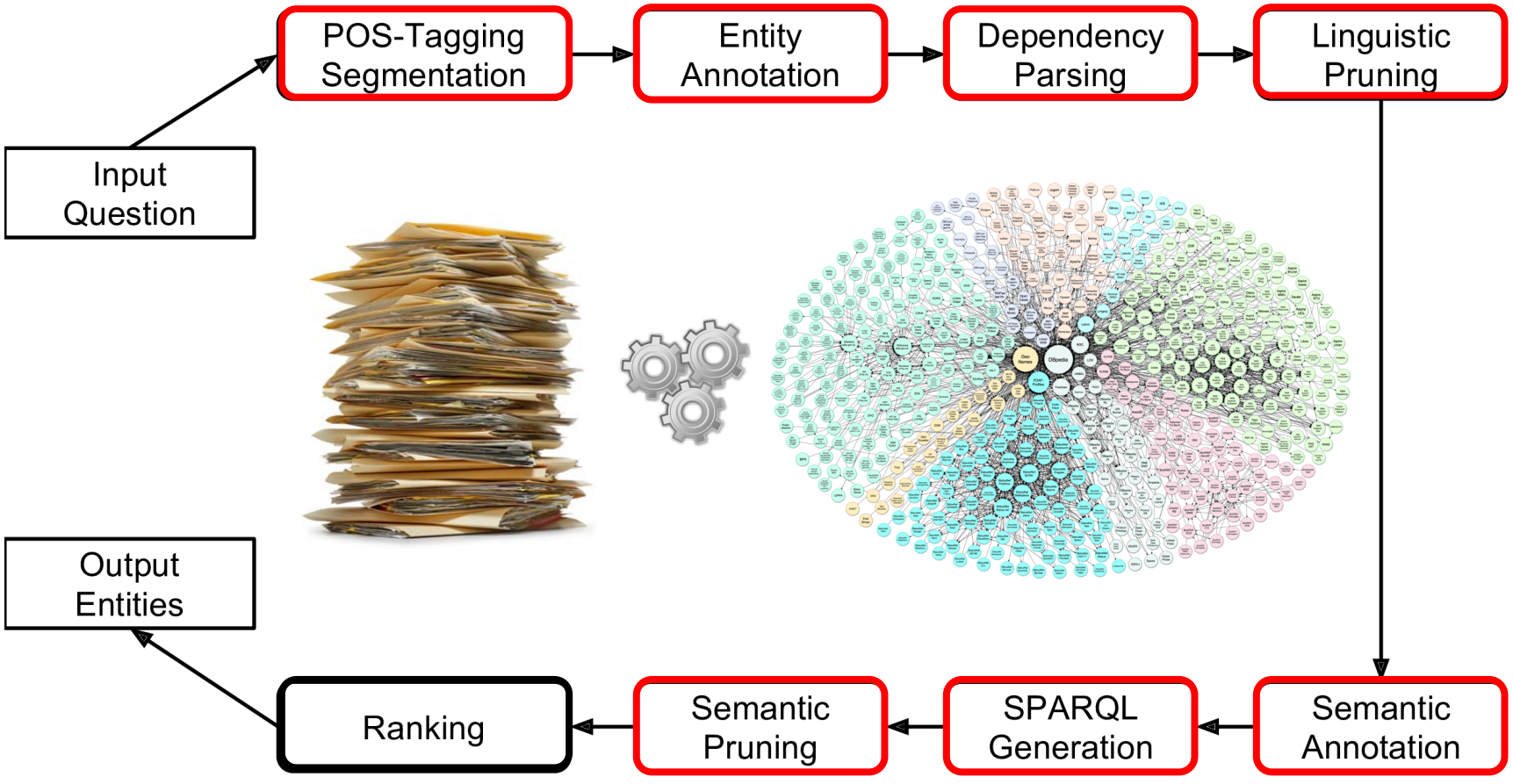
- Build for each node a triple pattern (in subject, predicate and object position) or
- Use Cross-Product of *N* free variables in SPARQL query
- Full-Text Filter based on CBD of entity

Hawk -
SPARQL Generation

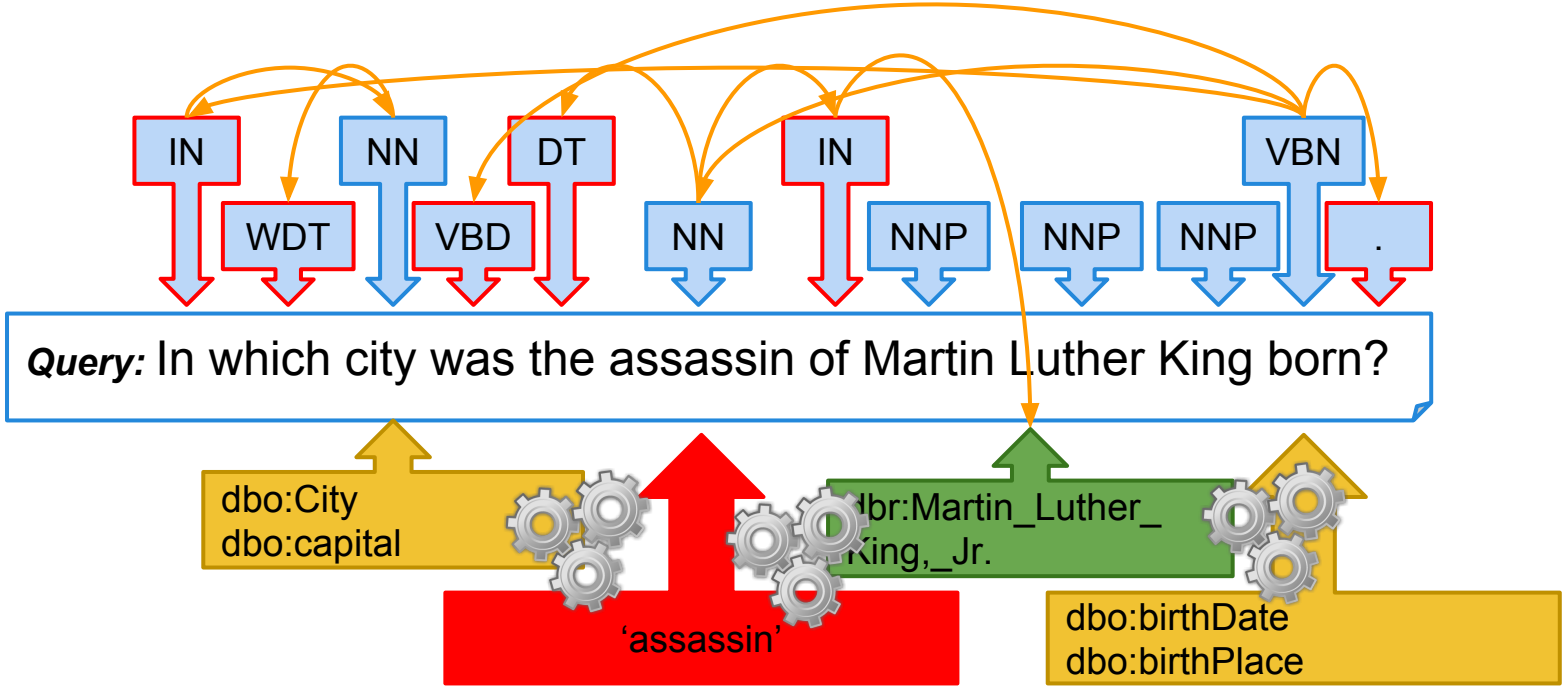


- Build for each node a triple pattern (in subject, predicate and object position) or
- Use Cross-Product of *N* free variables in SPARQL query
- Full-Text Filter based on CBD of entity

Hawk - Hybrid Question Answering

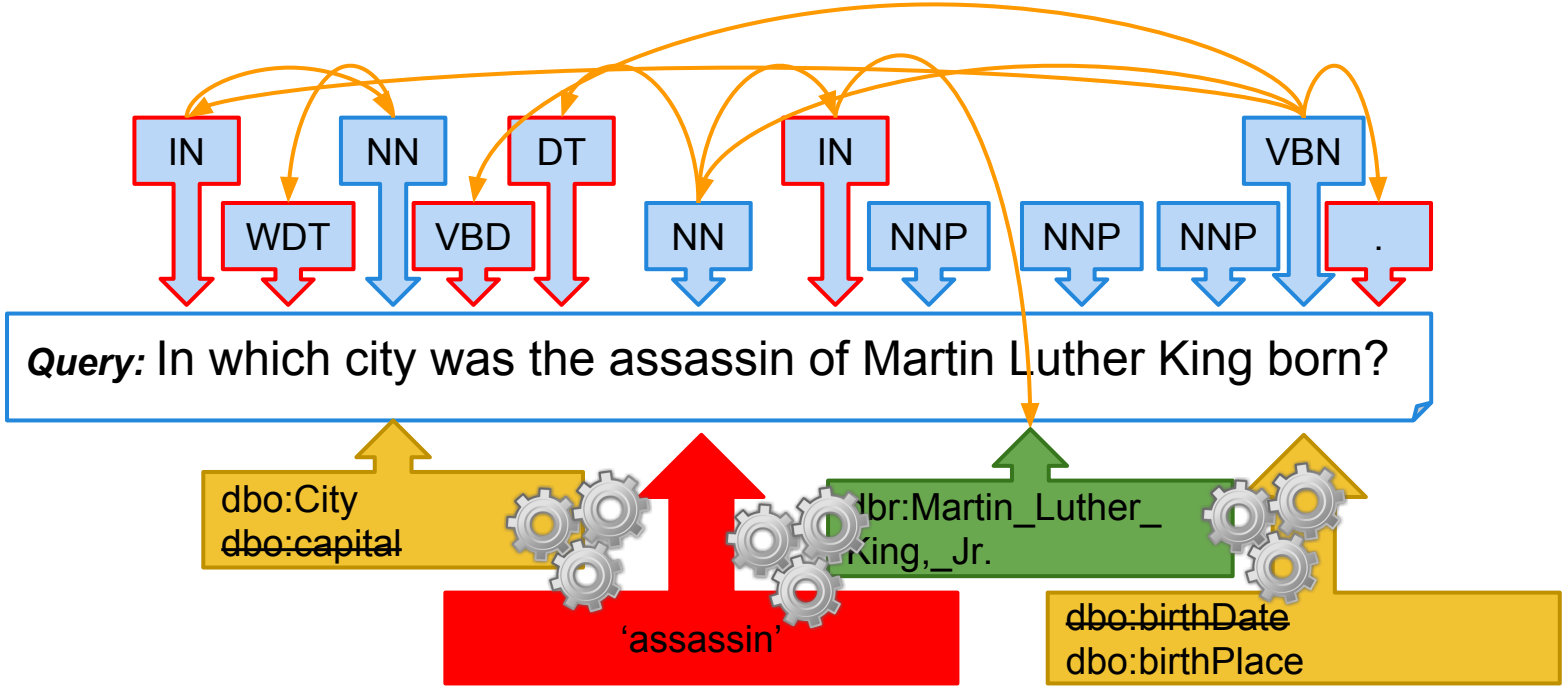


Hawk -
Semantic Pruning



- Discard invalid SPARQL queries via heuristics
  - Assume Disjointness of Classes

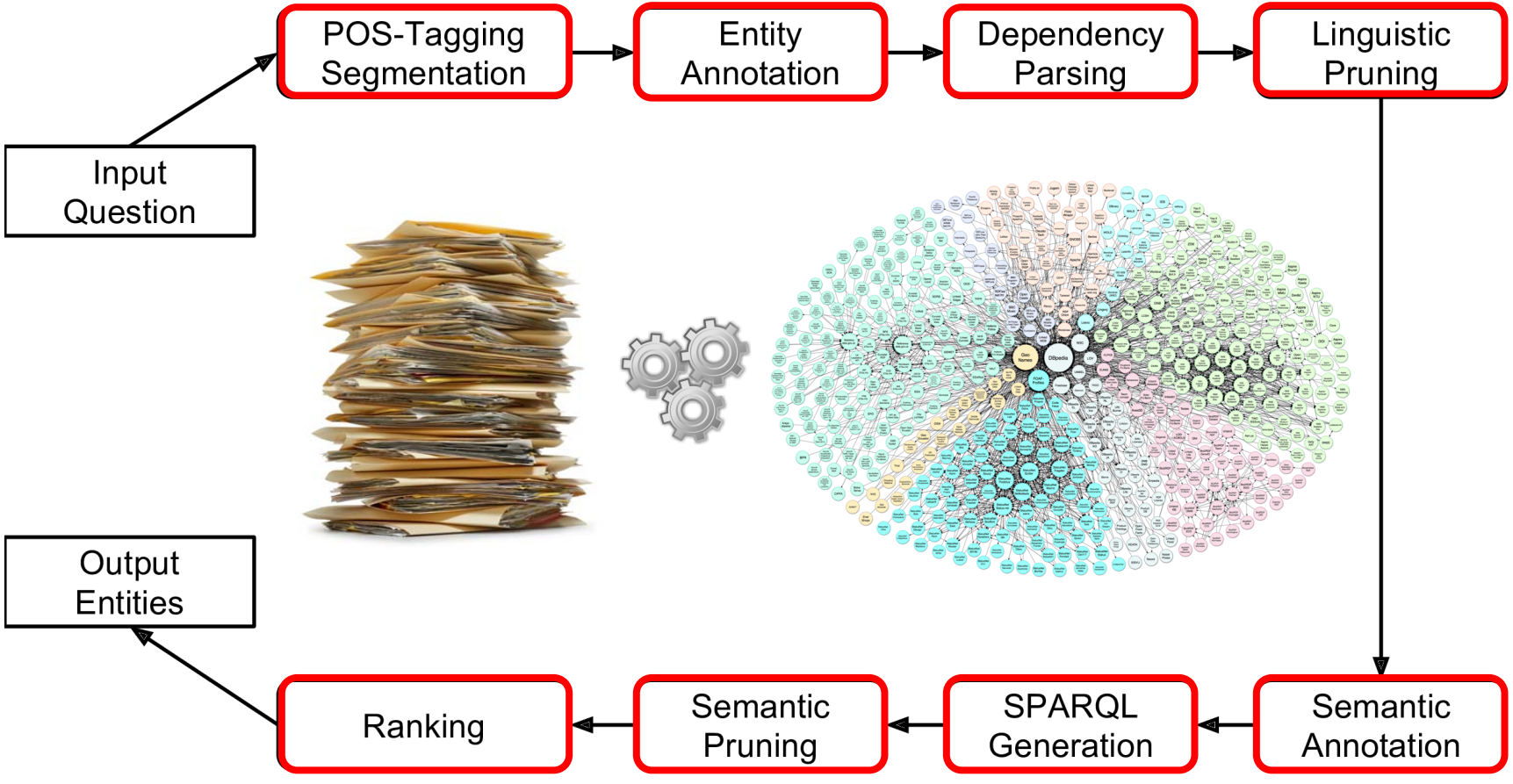
Hawk -
Semantic Pruning



- Discard invalid SPARQL queries via heuristics
  - Assume Disjointness of Classes

Hawk -

Hybrid Question Answering as a first step to a hybrid query engine



Hawk -
Semantic Ranking

Query: In which city was the assassin of Martin Luther King born?

```
SELECT DISTINCT ?proj WHERE {
?const text:query     (dbo:abstract 'assassin') .
?proj  dbo:birthPlace ?const .
?proj  ?pbridge       dbr:Martin_Luther_King,_Jr. .
}
```

```
SELECT DISTINCT ?proj WHERE {
?const text:query     (dbo:abstract 'assassin') .
?proj  dbo:birthPlace ?const .
?proj  ?pbridge       ?entity .
}
```

- Feature-based Ranking
  - Number of bound variables etc.

Hawk -
Semantic Ranking

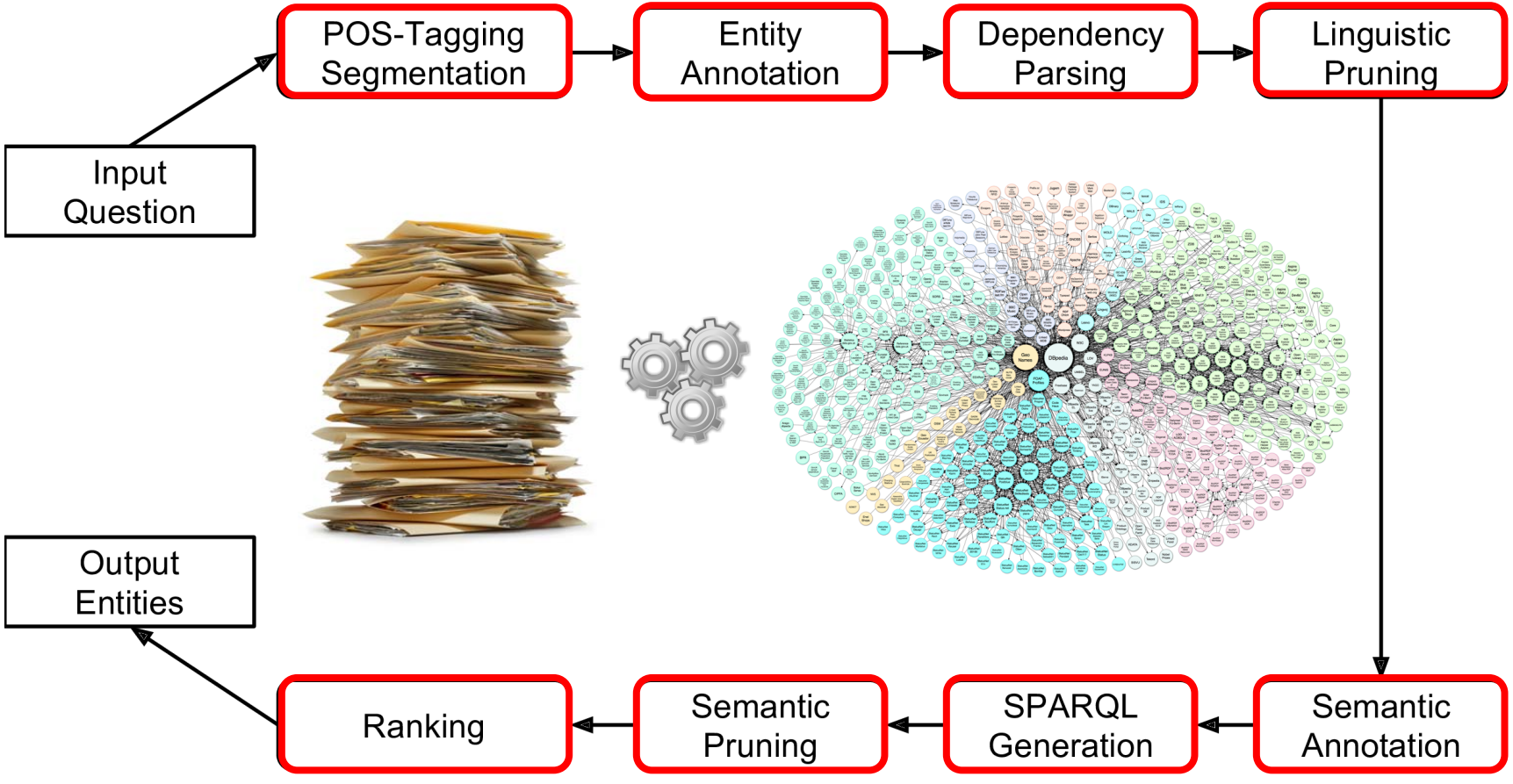
Query: In which city was the assassin of Martin Luther King born?

```
SELECT DISTINCT ?proj WHERE {
?const  text:query      (dbo:abstract ‘assassin’) .
?const  dbo:birthPlace ?proj .
?const  ?pbridge        dbr:Martin_Luther_King,_Jr.
}
```

~~SELECT DISTINCT ?proj WHERE {~~
~~?const text:query (dbo:abstract ‘assassin’) .~~
~~?proj dbo:birthPlace ?const .~~
~~?proj ?pbridge ?entity .~~
~~}~~

- Feature-based Ranking
  - Number of bound variables etc.

HAWK - Evaluation



HAWK - Evaluation

Setup

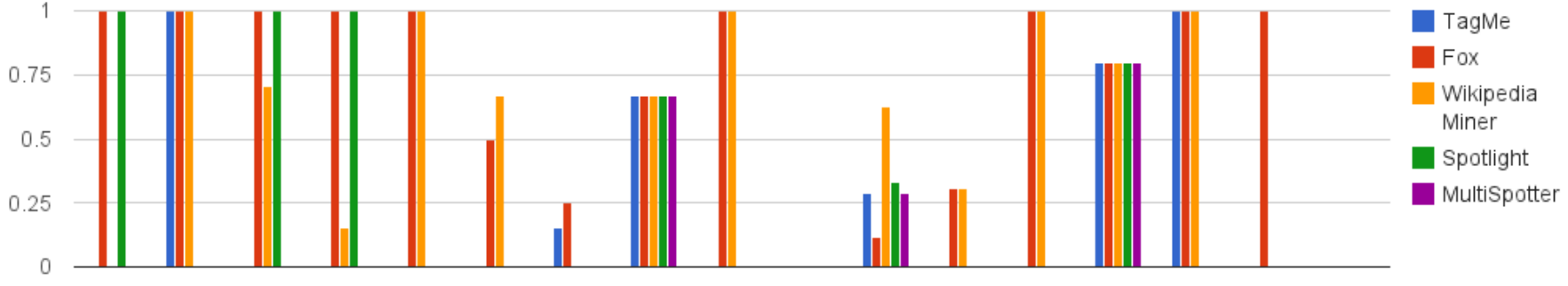
- QALD 4
  - 25 hybrid train + 10 hybrid test questions
  - Discard non-entity, aggregation and yago-Type demanding questions => 17 questions from train dataset

HAWK - Evaluation

Entity Annotator Influence

=> Assume an optimal ranking

- FOX = F-measure of 0.68
- Optimal annotator = F-measure of 0.58
- Union annotator = F-measure of 0.10



HAWK - Evaluation

ID	Question	F-measure	Precision	Recall
1	Give me the currencies of all G8 countries.	0.0	0.0	0.0
2	In which city was the assassin of Martin Luther King born?	1.0	1.0	1.0
3	Which anti-apartheid activist graduated from the University of South Africa?	1.0	1.0	1.0
5	Which recipients of the Victoria Cross died in the Battle of Arnhem?	0.8	0.67	1.0
6	Where did the first man in space die?	1.0	1.0	1.0
8	Which members of the Wu-Tang Clan took their stage name from a movie?	0.31	0.18	1.0
9	Which writers had influenced the philosopher that refused a Nobel Prize?	0.71	0.56	1.0
11	Who composed the music for the film that depicts the early life of Jane Austin?	0.0	0.0	0.0
14	Which horses did The Long Fellow ride?	1.0	1.0	1.0
15	Of the people that died of radiation in Los Alamos, whose death was an accident?	0.67	1.0	0.5
16	Which buildings owned by the crown overlook the North Sea?	0.25	0.14	1.0
17	Which buildings in art deco style did Shreve, Lamb and Harmon design?	0.5	0.33	1.0
18	Which birds are protected under the National Parks and Wildlife Act?	1.0	1.0	1.0
19	Which country did the first known photographer of snowflakes come from?	1.0	1.0	1.0
20	List all the battles fought by the lover of Cleopatra.	1.0	1.0	1.0
22	Which actress starring in the TV series Friends owns the production company Coquette Productions?	1.0	1.0	1.0
23	Dakar is the capital of which country member of the African Union?	1.0	1.0	1.0

- Evaluation based on QALD 4
- Assuming an optimal ranking

HAWK - Evaluation

ID	Question	F-measure	Precision	Recall
1	Give me the currencies of all G8 countries.	0.0	0.0	0.0
2	In which city was the assassin of Martin Luther King born?	1.0	1.0	1.0
3	Which anti-apartheid activist graduated from the University of South Africa?	1.0	1.0	1.0
5	Which recipients of the Victoria Cross died in the Battle of Arnhem?	0.8	0.67	1.0
6	Where did the first man in space die?	1.0	1.0	1.0
8	Which members of the Wu-Tang Clan took their stage name from a movie?	0.31	0.18	1.0
9	Which writers had influenced the philosopher that refused a Nobel Prize?	0.71	0.56	1.0
11	Who composed the music for the film that depicts the early life of Jane Austin?	0.0	0.0	0.0
14	Which horses did The Long Fellow ride?	1.0	1.0	1.0
15	Of the people that died of radiation in Los Alamos, whose death was an accident?	0.67	1.0	0.5
16	Which buildings owned by the crown overlook the North Sea?	0.25	0.14	1.0
17	Which buildings in art deco style did Shreve, Lamb and Harmon design?	0.5	0.33	1.0
18	Which birds are protected under the National Parks and Wildlife Act?	1.0	1.0	1.0
19	Which country did the first known photographer of snowflakes come from?	1.0	1.0	1.0
20	List all the battles fought by the lover of Cleopatra.	1.0	1.0	1.0
22	Which actress starring in the TV series Friends owns the production company Coquette Productions?	1.0	1.0	1.0
23	Dakar is the capital of which country member of the African Union?	1.0	1.0	1.0

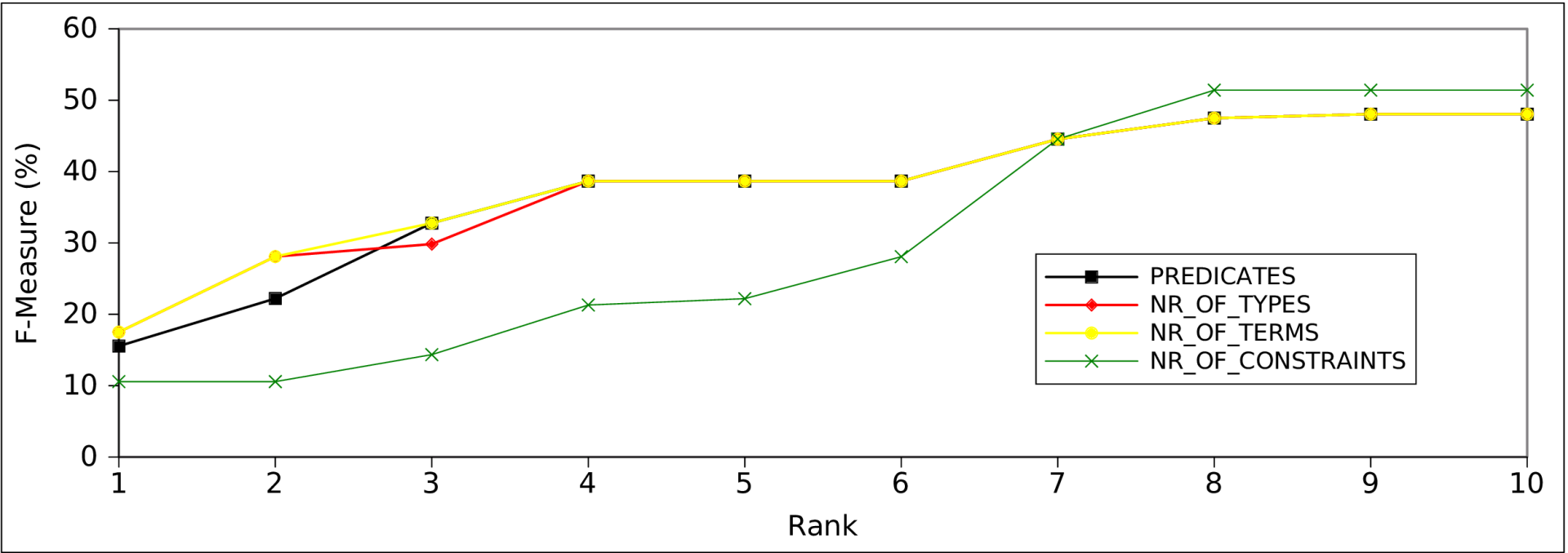
- Evaluation based on QALD 4
- Assuming an optimal ranking

- Precision = 0.70
- Recall = 0.85
- F-measure = 0.72

HAWK - Evaluation

Evaluation of Feature-based Ranking

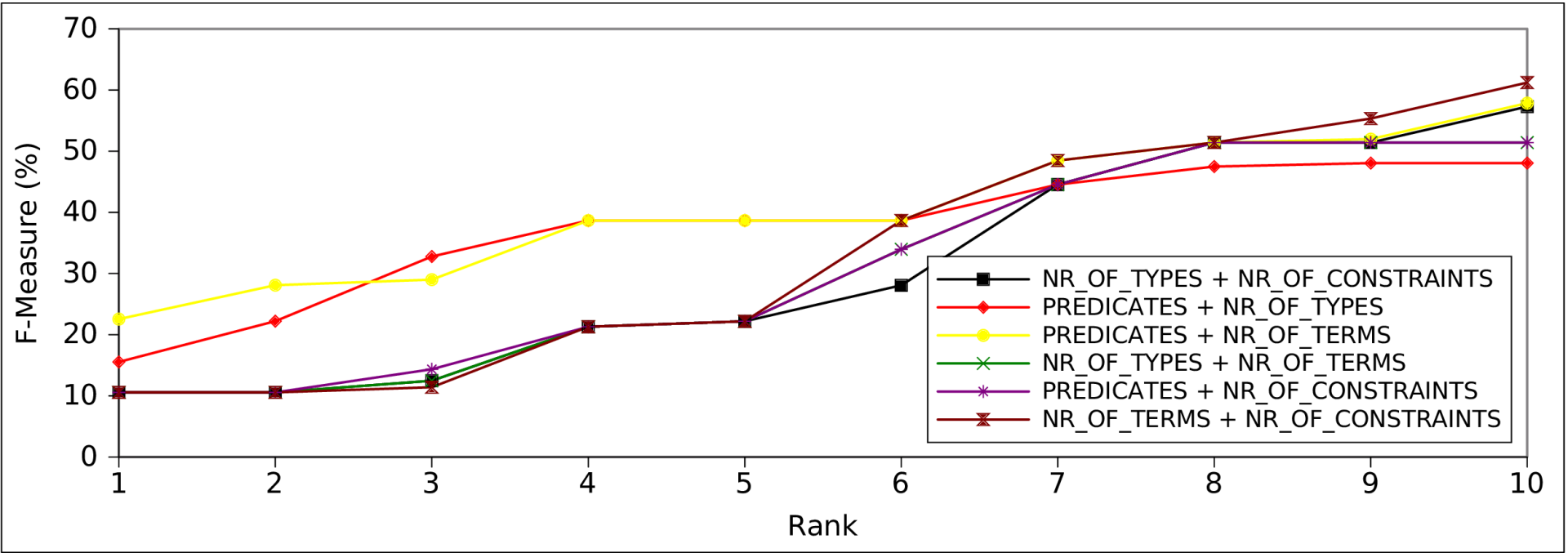
- Rank = Number of generated SPARQL queries



HAWK - Evaluation

Evaluation of Feature-based Ranking

- Rank = Number of generated SPARQL queries



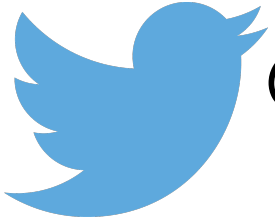
HAWK - Conclusion

Novel, modular, hybrid question answering system

- Challenges:
  - Training + benchmark datasets small
  - Answering other types of queries
  - Improve ranking mechanism
  - Extend the QA ecosystem

HAWK - Conclusion

Thank you! Questions?



@akswgroup

http://aksw.org/Projects/HAWK
http://github.org/aksw/hawk

Ricardo Usbeck
Axel Ngonga
Augustusplatz 10
Leipzig, Germany
http://aksw.org

HAWK - Evaluation

Detailed results per question

ID	Recall	Precision	F1-measure
301	1.0	1.0	1.0
303	1.0	0.5	0.67
305	0.5	0.5	0.5
306	1.0	1.0	1.0
308	0.5	0.5	0.5
309	1.0	1.0	1.0
314	0.86	1.0	0.92
315	0.5	1.0	0.67
317	1.0	0.33	0.5
318	0.67	0.67	0.67
319	1.0	1.0	1.0
320	0.23	0.42	0.29
322	1.0	1.0	1.0
323	1.0	1.0	1.0
334	1.0	1.0	1.0

- Evaluation based on QALD 5
- Assuming an optimal ranking

- Train dataset:
  - Precision = 0.30
  - Recall = 0.30
  - F-measure = 0.30
- Test dataset:
  - Only able to answer 1 Question correctly