Query Interpretations from Entity-Linked Segmentations

WSDM 2022

Vaibhav Kasturia[1]

Marcel Gohsen[2]

Matthias Hagen[3]

[1,3] MARTIN-LUTHER-UNIVERSITÄT HALLE-WITTENBERG

[2] Bauhaus-Universität Weimar

[2,3] Webis

What is the intent?

©KASTURIA/GOHSEN/HAGEN 2022

What is the intent?

Is it . . .

```
source | of the | Nile
```

©KASTURIA/GOHSEN/HAGEN 2022

What is the intent?

Is it . . . or . . . ?

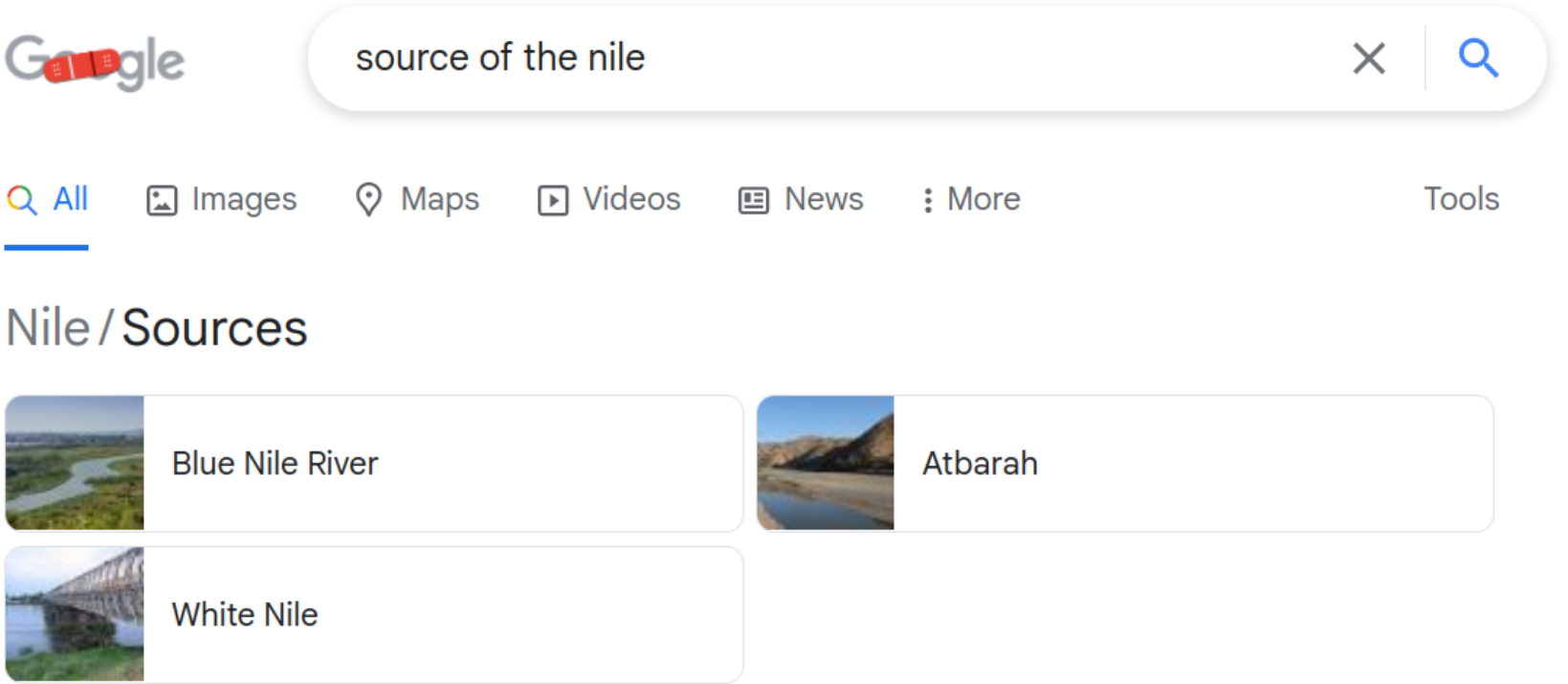
`source | of the | Nile`

`Source_of_the_Nile`

©KASTURIA/GOHSEN/HAGEN 2022

What is the intent?

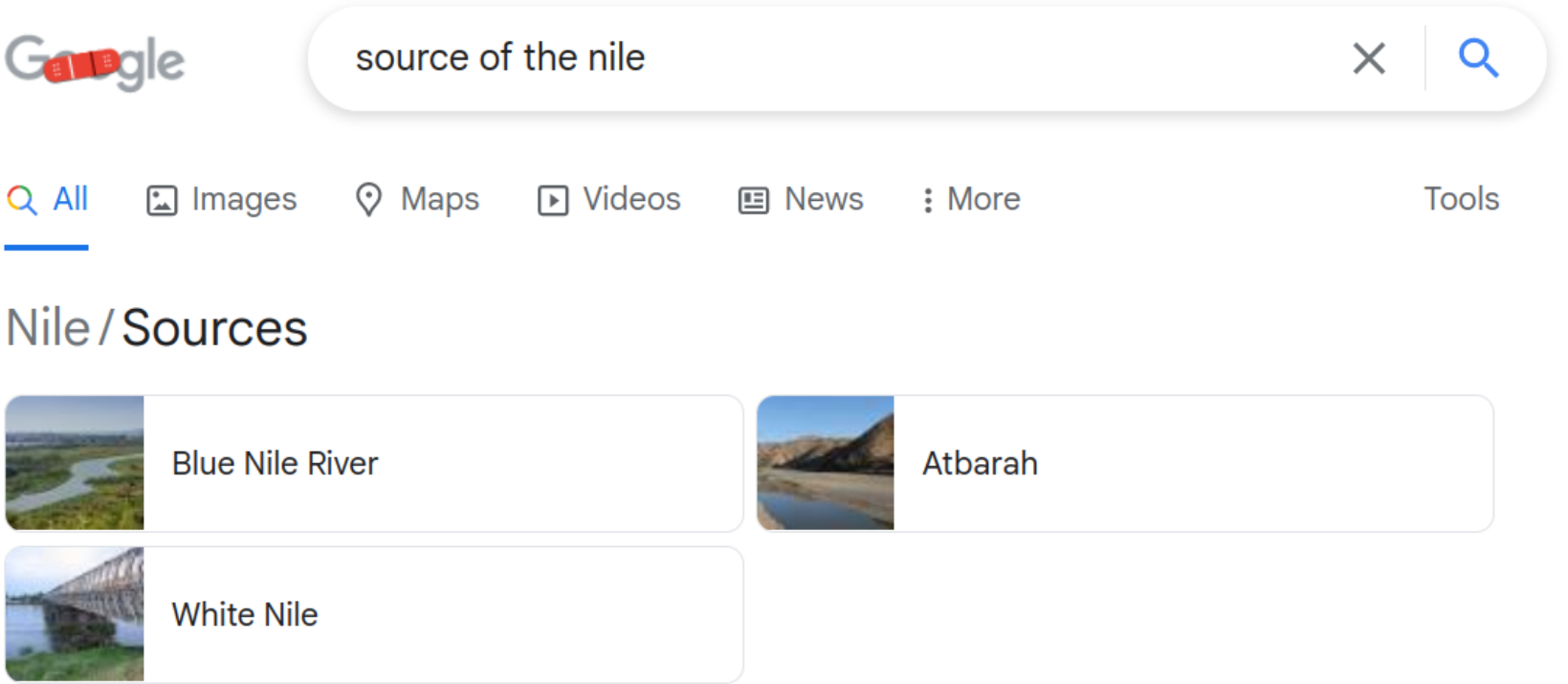
What does Google think?



©KASTURIA/GOHSEN/HAGEN 2022

What is the intent?

What does Google think?



⇒ Where is the board game?

©KASTURIA/GOHSEN/HAGEN 2022

Computational Problems

Entity Linking

`source of the nile`

Explicit entities:

©KASTURIA/GOHSEN/HAGEN 2022

Computational Problems

Entity Linking

source of the nile

Explicit entities:

Nile

©KASTURIA/GOHSEN/HAGEN 2022

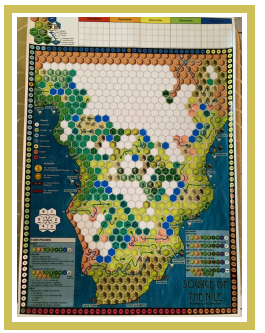
Computational Problems

Entity Linking

`source of the nile`

Explicit entities:

`Nile`



`Source_of_the_Nile`

©KASTURIA/GOHSEN/HAGEN 2022

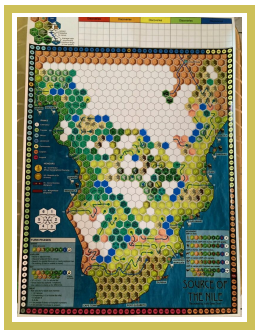
Computational Problems

Entity Linking

source of the nile

Explicit entities:

Nile



Source_of_the_Nile

...

Implicit entities:

©KASTURIA/GOHSEN/HAGEN 2022

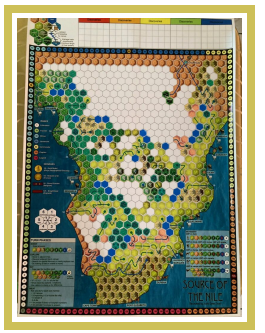
Computational Problems

Entity Linking

source of the nile

Explicit entities:

Nile



Source_of_the_Nile

...

Implicit entities:

White_Nile

©KASTURIA/GOHSEN/HAGEN 2022

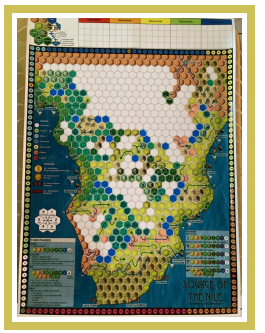
Computational Problems

Entity Linking

source of the nile

Explicit entities:

Nile



Source_of_the_Nile

...

Implicit entities:

White_Nile

Blue_Nile

©KASTURIA/GOHSEN/HAGEN 2022

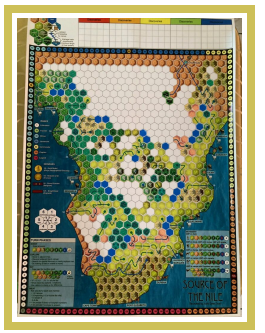
Computational Problems

Entity Linking

source of the nile

Explicit entities:

Nile



Source_of_the_Nile

...

Implicit entities:

White_Nile

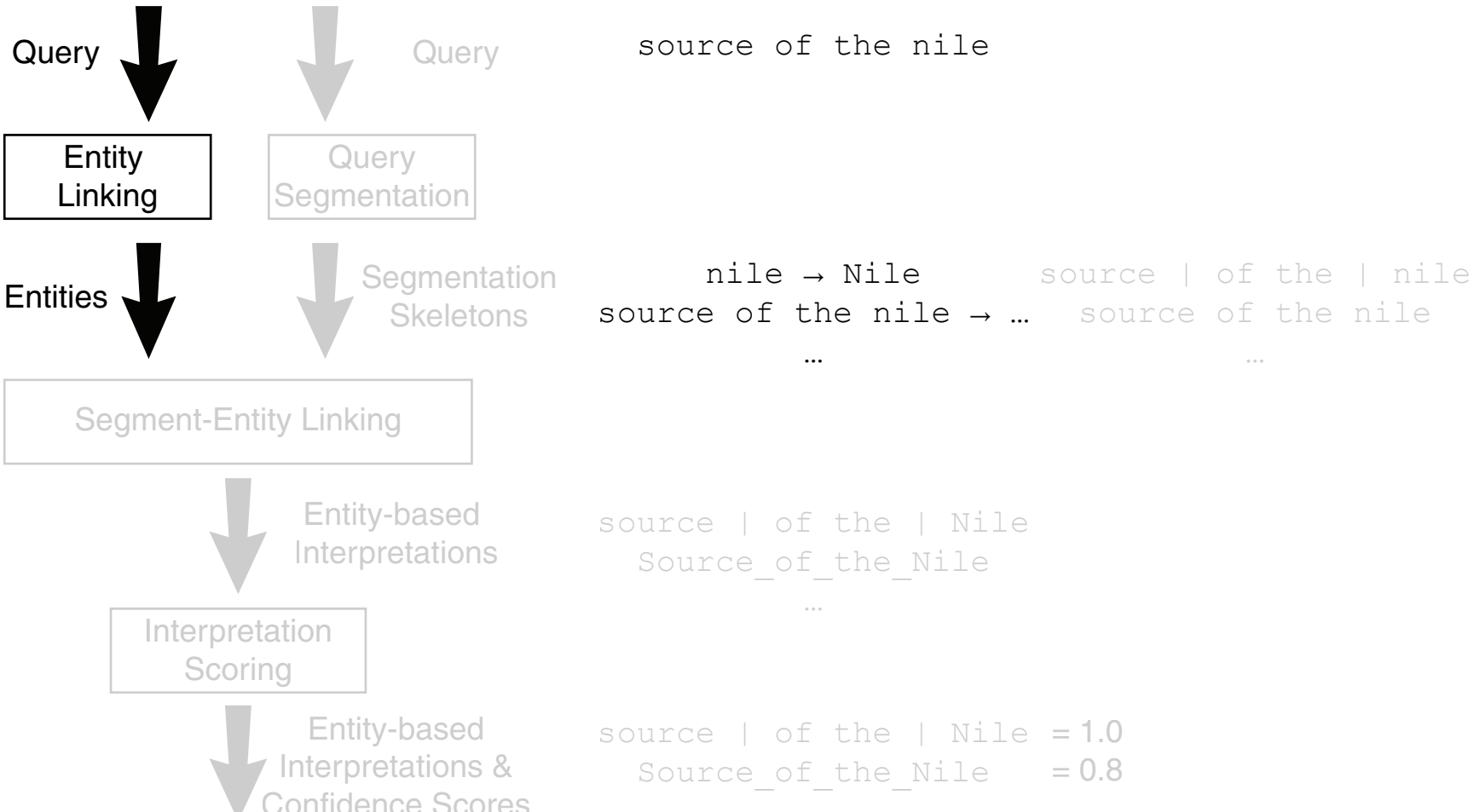
Blue_Nile

...

©KASTURIA/GOHSEN/HAGEN 2022

Computational Problems

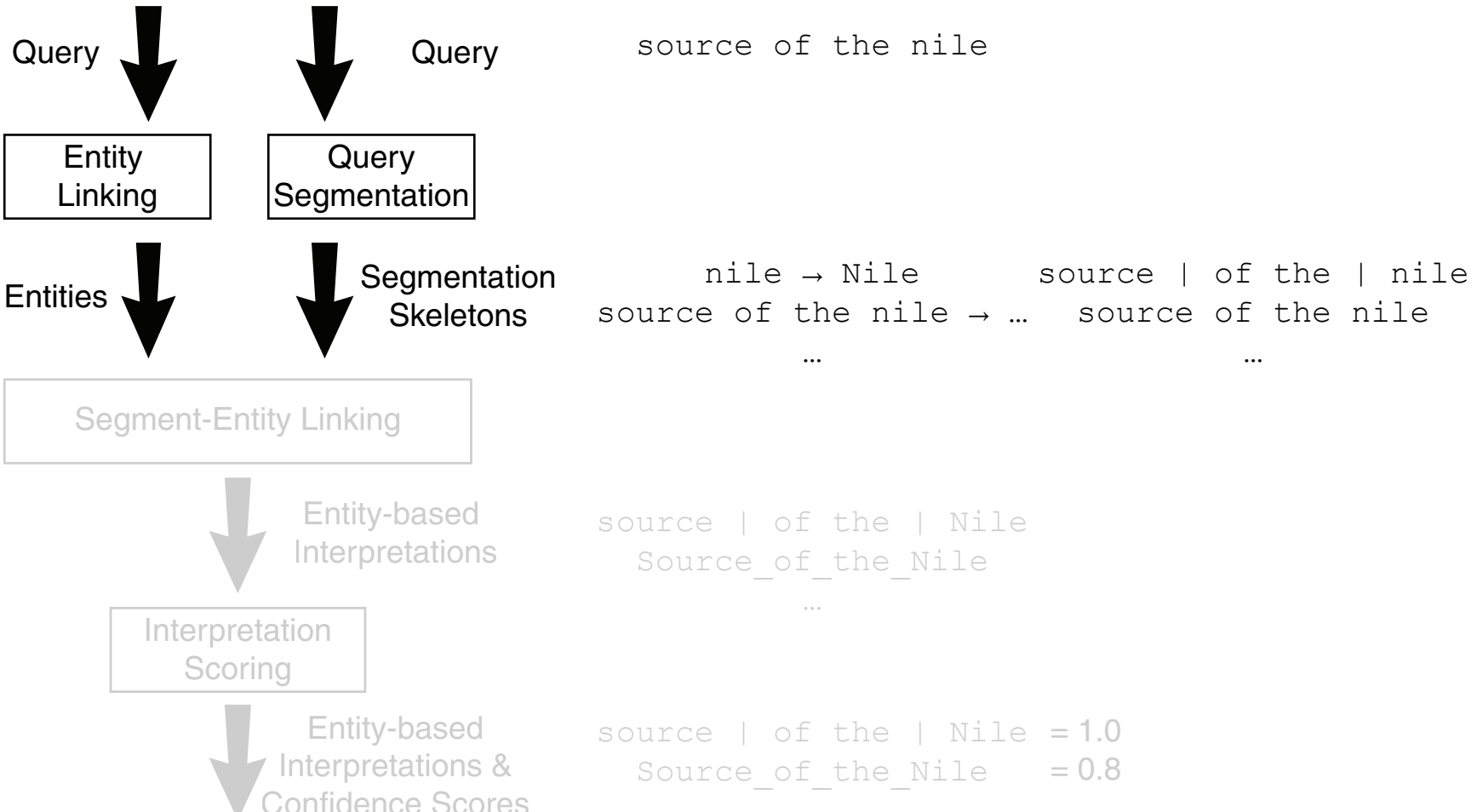
Query Interpretation



©KASTURIA/GOHSEN/HAGEN 2022

Computational Problems

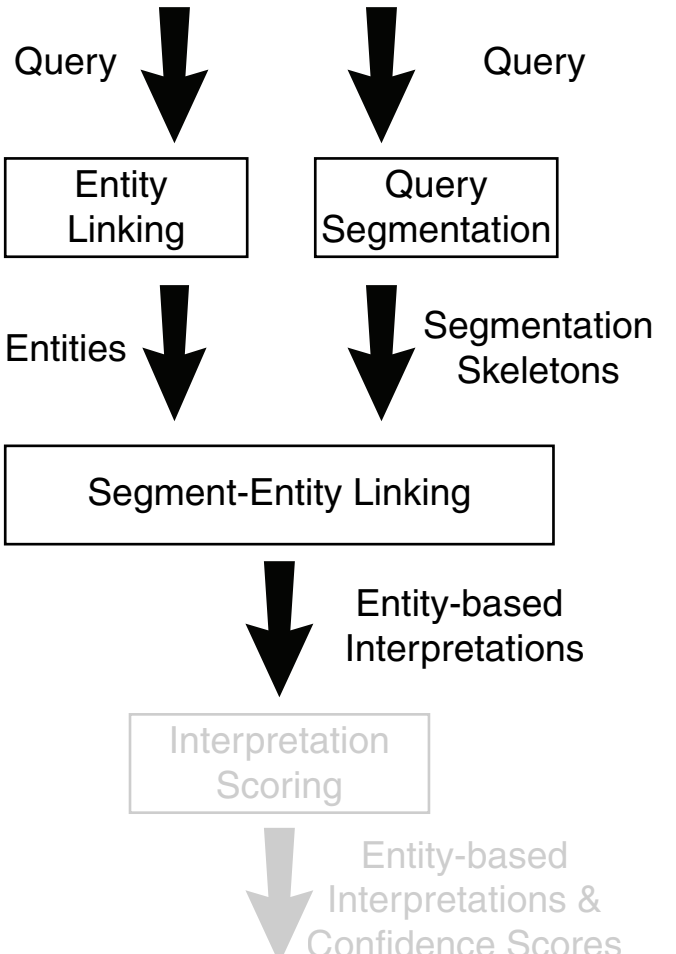
Query Interpretation



©KASTURIA/GOHSEN/HAGEN 2022

Computational Problems

Query Interpretation

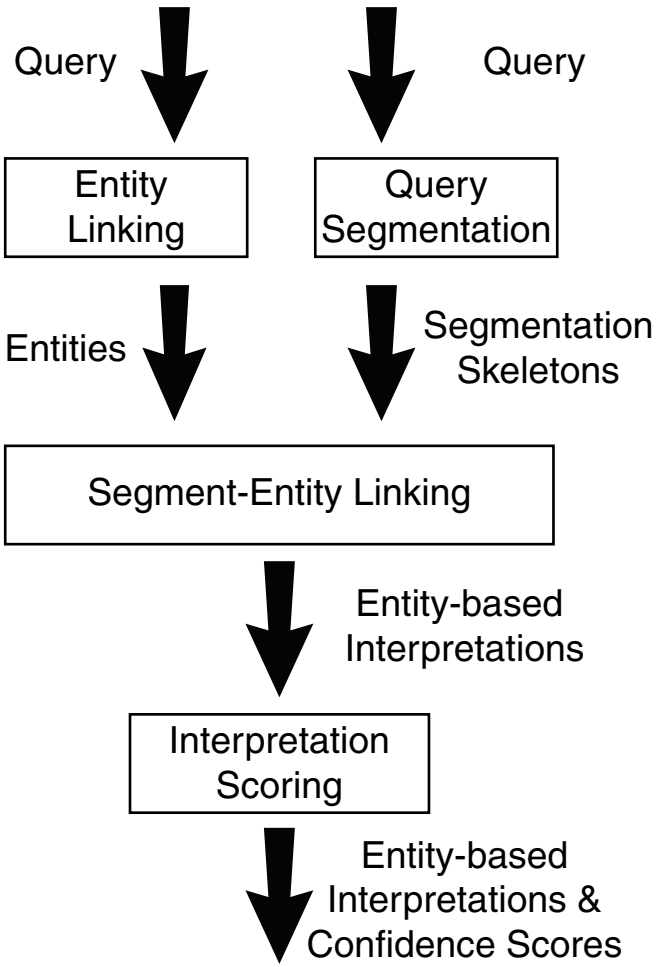



```
source of the nile

nile → Nile                    source | of the | nile
source of the nile → …           source of the nile
…                                …

source | of the | Nile
  Source_of_the_Nile
…

source | of the | Nile  = 1.0
  Source_of_the_Nile    = 0.8
…
```


©KASTURIA/GOHSEN/HAGEN 2022

Computational Problems

Query Interpretation




```
source of the nile

nile → Nile                    source | of the | nile
source of the nile → …           source of the nile
…                                …

source | of the | Nile
  Source_of_the_Nile
…

source | of the | Nile  = 1.0
  Source_of_the_Nile    = 0.8
…
```


©KASTURIA/GOHSEN/HAGEN 2022

Evaluation

Experiment: Query Interpretation

Comparison to pairs of *greedy interpretation finding* [Hasibi et al., 2014] and various entity linking approaches.

Data:

- Test split of our new query interpretation dataset (544 queries)
- Interpretations that are at least “moderately likely”

©KASTURIA/GOHSEN/HAGEN 2022

Evaluation

Query Interpretation Results

	Complete Matches			Time
	Recall	Precision	F_1	(ms)
Our approach	**0.295**	**0.336**	**0.283**	47
Dexter	0.230	0.312	0.246	282
Nordlys EL	0.189	0.278	0.207	1,533
Radboud EL	0.144	0.199	0.155	200
Smaph	0.176	0.243	0.190	116,425
Dandelion	0.166	0.226	0.177	74
TagMe	0.165	0.216	0.175	99
Babelfy	0.112	0.160	0.124	49
TextRazor	0.098	0.131	0.105	367
FEL	0.133	0.173	0.141	**22**
Ambiverse	0.007	0.011	0.009	53

©KASTURIA/GOHSEN/HAGEN 2022

Conclusion

Contributions:

- Entity-based query interpretation
- Combination annotated dataset: 2,800 queries
- Recall-oriented explicit entity recognizer
- Segmentation-based query interpretation method
- Comparison to existing entity linking and interpretation methods

©KASTURIA/GOHSEN/HAGEN 2022

Conclusion

Contributions:

- Entity-based query interpretation
- Combination annotated dataset: 2,800 queries
- Recall-oriented explicit entity recognizer
- Segmentation-based query interpretation method
- Comparison to existing entity linking and interpretation methods

Future work:

- Implicit entity recognition
- Query interpretation for conversational systems

©KASTURIA/GOHSEN/HAGEN 2022

Conclusion

Contributions:

- Entity-based query interpretation
- Combination annotated dataset: 2,800 queries
- Recall-oriented explicit entity recognizer
- Segmentation-based query interpretation method
- Comparison to existing entity linking and interpretation methods

Future work:

- Implicit entity recognition
- Query interpretation for conversational systems

Thank you! :-)

©KASTURIA/GOHSEN/HAGEN 2022