

Integrating Query Interpretation Components into the Information Retrieval Experiment Platform

WOWS 2024



**Marcel
Gohsen**

**Bauhaus-
Universität
Weimar**



**Benno
Stein**



Webis

Query Interpretation



source of the Nile



Query Interpretation

🔍 source of the Nile ✕



Query Interpretation

🔍 source of the Nile ✕



OR



Query Interpretation



source of the Nile



Nile / Sources



Blue Nile River



Atbarah



White Nile

Query Interpretation



source of the Nile



Nile / Sources



Blue Nile River



Atbarah



White Nile



Wikipedia

<https://en.wikipedia.org/wiki/Nile>

Nile

In particular, the Nile is the **primary water source of Egypt, Sudan and South Sudan**. ... Additionally, the Nile is an important economic river, supporting ...
[White Nile](#) · [Blue Nile](#) · [Nile Delta](#) · [River source](#)



Query Interpretation

🔍 source of the Nile



Nile / Sources



Blue Nile River



Atbarah



White Nile



Wikipedia

<https://en.wikipedia.org/wiki/Nile> ⋮

Nile

In particular, the Nile is the **primary water source of Egypt, Sudan and South Sudan**. ... Additionally, the Nile is an important economic river, supporting ...
[White Nile](#) · [Blue Nile](#) · [Nile Delta](#) · [River source](#)



Britannica

[https://www.britannica.com/.../Rivers & Canals](https://www.britannica.com/.../Rivers_and_Canals) ⋮

Nile River | Delta, Map, Basin, Length, Facts, Definition ...

Its reputed source is a **spring, considered holy by the Ethiopian Orthodox Church**, from which a small stream, the Abay, flows down to Lake Tana (T'ana), a fairly ...
[Physiography](#) · [White Nile River](#) · [Nile River basin](#) · [Climate and hydrology](#)



Query Interpretation

🔍 source of the Nile ✕



source of the Nile

OR



source of the Nile

Query Interpretation

🔍 source of the Nile ✕



OR



⟨source | of the | Nile⟩

⟨source of the Nile⟩

Query Interpretation

🔍 source of the Nile ✕



OR

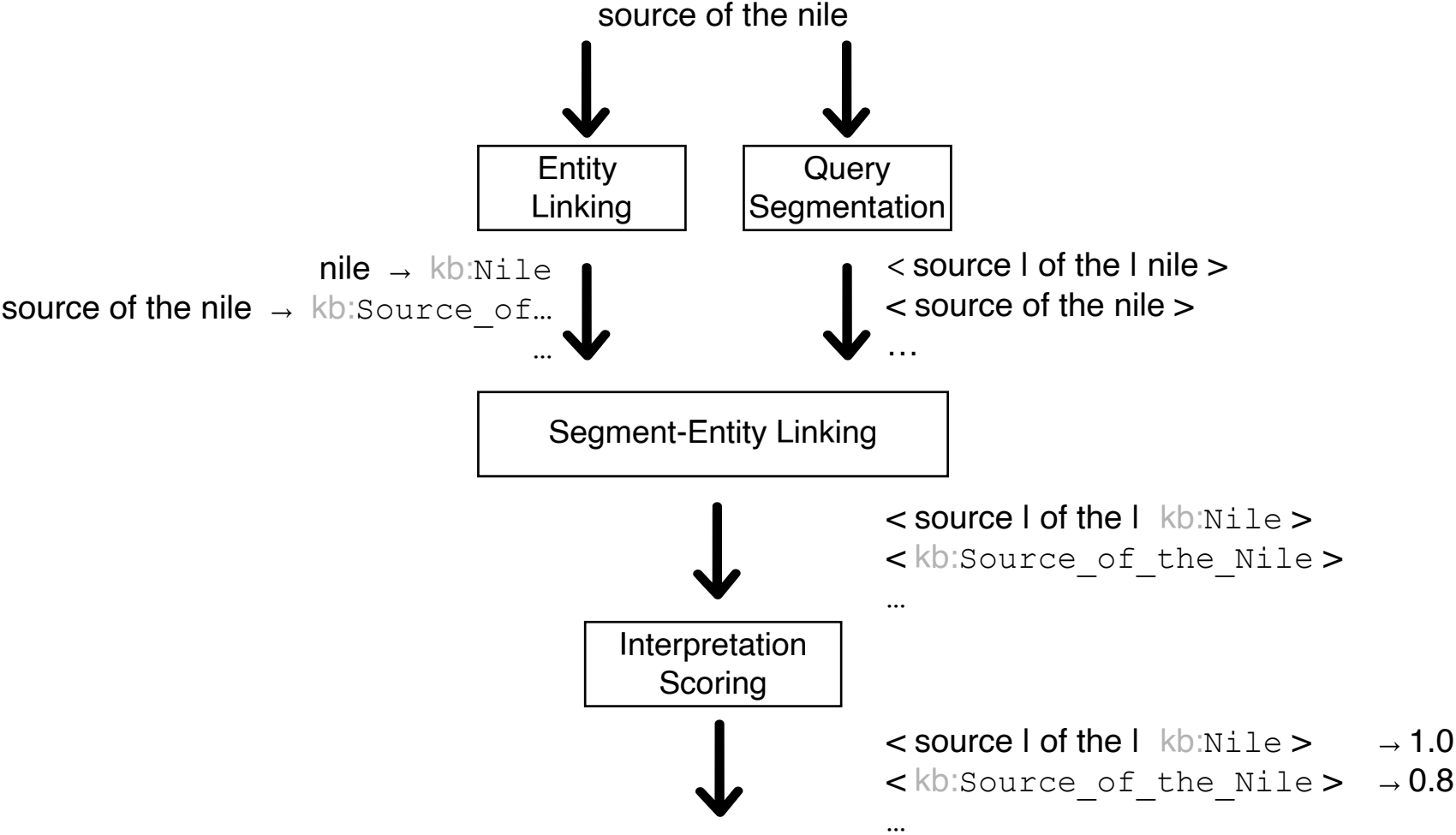


`<source | of the | kb:Nile>`

`<kb:Source_of_the_Nile>`

Query Interpretation

Kasturia et al. 2022



Entity Linking

Kasturia et al. 2022

Nile

[Article](#) [Talk](#)

From Wikipedia, the free encyclopedia

(Redirected from [Nile river](#))

For other uses, see [Nile \(disambiguation\)](#).

Entity index:

- `nile` → [`kb:Nile,...`]
- ...

Entity Linking

Kasturia et al. 2022

Nile

[Article](#) [Talk](#)

From Wikipedia, the free encyclopedia

(Redirected from [Nile river](#))

For other uses, see [Nile \(disambiguation\)](#).

Entity index:

- ❑ `nile` → [`kb:Nile`,...]
- ❑ `nile river` → [`kb:Nile`, ...]
- ❑ ...

Entity Linking

Kasturia et al. 2022

Nile

[Article](#) [Talk](#)

From Wikipedia, the free encyclopedia

(Redirected from [Nile river](#))

For other uses, see [Nile \(disambiguation\)](#).

Entity index:

- ❑ `nile` → [`kb:Nile`, `kb:Nile_(band)`, ...]
- ❑ `nile river` → [`kb:Nile`, ...]
- ❑ ...

Entity Linking

Kasturia et al. 2022

Nile

[Article](#) [Talk](#)

From Wikipedia, the free encyclopedia

(Redirected from [Nile river](#))

For other uses, see [Nile \(disambiguation\)](#).

Entity index:

- ❑ `nile` → [`kb:Nile`, `kb:Nile_(band)`, ...]
- ❑ `nile river` → [`kb:Nile`, ...]
- ❑ ...

Entity linking process

1. Segment query into all possible n-grams
2. Request entity index for each segment
3. Compute commonness scores for each segment-entity pair
4. Remove all entities with commonness of zero

Query Segmentation

Hagen et al. 2012, Hagen et al. 2013

Rank all possible segmentations by the sum of pre-computed segment scores.

The segment score of a segment s :

- s **is not** a Wikipedia title
 - Occurrence frequency according to Google n-grams
- s **is** a Wikipedia title
 - $1 +$ Occurrence frequency of most frequent 2-gram in s

Query Segmentation

Hagen et al. 2012, Hagen et al. 2013

Rank all possible segmentations by the sum of pre-computed segment scores.

The segment score of a segment s :

- s **is not** a Wikipedia title
 - Occurrence frequency according to Google n-grams
- s **is** a Wikipedia title
 - $1 +$ Occurrence frequency of most frequent 2-gram in s

Cut off segmentations with a ratio of less than 0.66:

Rank	Segmentation	Score	Ratio
1	⟨source of the nile⟩	496.6 million	-
2	⟨source of the nile⟩	333.4 million	0.67
3	⟨source of the nile⟩	35.6 million	0.11
⋮			

Query Interpretation

Combine segmentations with linked entities to build interpretation candidates.

Rank interpretation candidates by a weighted sum of

- ❑ Entity commonness
- ❑ Relatedness between entities
- ❑ Context score between an entity and unlinked segments

Relatedness and context scores are computed with Wikipedia-based joint word-entity embeddings.¹

¹Wikipedia2Vec [wikipedia2vec.github.io/wikipedia2vec/](https://github.com/tylertomlinson/wikipedia2vec/)

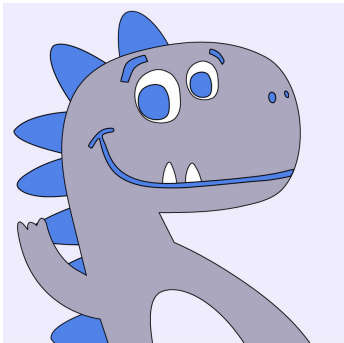
Query Interpretation

Combine segmentations with linked entities to build interpretation candidates.

Rank interpretation candidates by a weighted sum of

- ❑ Entity commonness
- ❑ Relatedness between entities
- ❑ Context score between an entity and unlinked segments

Relatedness and context scores are computed with Wikipedia-based joint word-entity embeddings.¹



The query entity linking and interpretation approaches are available as dockerized software and have been integrated into TIREx.

¹Wikipedia2Vec [wikipedia2vec.github.io/wikipedia2vec/](https://github.com/wikipedia2vec/wikipedia2vec/)

Query Interpretation Analytics

As components of TIREx,...

- 89,289 entities in 2,544 queries from 31 datasets and
- 2,304 interpretations of 1,225 queries from 18 datasets

... have been identified, which can be used and accessed through TIREx.

Query Interpretation Analytics

As components of TIREx,...

- 89,289 entities in 2,544 queries from 31 datasets and
- 2,304 interpretations of 1,225 queries from 18 datasets

... have been identified, which can be used and accessed through TIREx.

As part of our preliminary analysis, we have found that ...

- most queries contain entities (98%)
- the number of entities in a query correlates with its length ($\rho = 0.63$)
- most queries are ambiguous (1.8 interpretations per query)
- the number of interpretations does not correlate with the number of relevant documents ($\rho \approx 0$)

Query Interpretation Examples

Example queries from TREC Web Track 2009 and 2012 datasets.

Rank	Interpretation	Score
<i>For query “obama family tree”</i>		
1	⟨kb:Barack_Obama family tree⟩	0.77
2	⟨kb:Barack_Obama kb:Family_Tree⟩	0.50
3	⟨obama kb:Family_Tree⟩	0.37
<i>For query “pork tenderloin”</i>		
1	⟨kb:Pork_tenderloin⟩	0.92
<i>For query “last supper painting”</i>		
1	⟨kb:The_Last_Supper_(Leonardo) painting⟩	1.50
2	⟨last supper kb:Painting⟩	1.16
3	⟨kb:The_Last_Supper_(Leonardo) kb:Painting⟩	1.12

Summary

We contributed. . .

- ❑ query entity linking and interpretation software as TIREx components
- ❑ 89,289 entity candidates and 2,304 interpretations for 31 and 18 datasets
- ❑ preliminary analytics of entities and interpretations

Summary

We contributed...

- ❑ query entity linking and interpretation software as TIREx components
- ❑ 89,289 entity candidates and 2,304 interpretations for 31 and 18 datasets
- ❑ preliminary analytics of entities and interpretations

Thank you! Find more information in the paper and in the following repositories.



[www.github.com/webis-de/
query-entity-linking](https://www.github.com/webis-de/query-entity-linking)



[www.github.com/webis-de/
query-interpretation](https://www.github.com/webis-de/query-interpretation)

Evaluation

Query Entity Linking

	MicR	MicR*	MacR	MacR*	MicP	MacP
<i>Entity Linking Tools</i>						
Our approach	0.838	0.859	0.668	0.670	0.035	0.126
Nordlys ER	0.735	0.776	0.543	0.548	0.002	0.009
TagMe	0.333	0.410	0.385	0.401	0.328	0.399
Babelfy	0.320	0.398	0.383	0.398	0.293	0.289
Smaph	0.314	0.390	0.399	0.413	0.431	0.463
Dandelion	0.302	0.373	0.414	0.428	0.431	0.500
Nordlys EL	0.293	0.359	0.579	0.593	0.780	0.731
Dexter	0.267	0.332	0.359	0.372	0.481	0.462
FEL	0.250	0.309	0.313	0.324	0.273	0.333
TextRazor	0.216	0.265	0.372	0.380	0.511	0.445
Radboud EL	0.213	0.263	0.498	0.507	0.789	0.627
Falcon	0.204	0.251	0.226	0.234	0.397	0.368
Ambiverse	0.011	0.013	0.259	0.259	0.750	0.263
<i>Entity Recognition Tools</i>						
AWS Comprehend	0.229	–	0.476	–	0.604	0.616
MITIE	0.114	–	0.358	–	0.797	0.463
Flair NER	0.129	–	0.374	–	0.787	0.487
LingPipe NER	0.109	–	0.321	–	0.497	0.410
DeepPavlov	0.048	–	0.269	–	0.478	0.305
Stanford NER	0.011	–	0.257	–	0.563	0.261
OpenNLP	0.000	–	0.246	–	0.000	0.246
No-Entity Baseline	0.000	0.000	0.246	0.246	0.000	0.246

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”

Evaluation

Query Interpretation Results

	Complete Matches			Time (ms)
	Recall	Precision	F ₁	
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
Babelify	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