

YAGO: a Multilingual Knowledge Base from Wikipedia, Wordnet, and Geonames

Thomas Rebele¹, Fabian Suchanek¹, Johannes Hoffart²,
Joanna Biega², Erdal Kuzey², Gerhard Weikum²



¹ Télécom ParisTech ² Max Planck Institute for Informatics

2016-09-09

What is a knowledge base?

YAGO:
Multilingual
Knowledge Base

Thomas Rebele

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

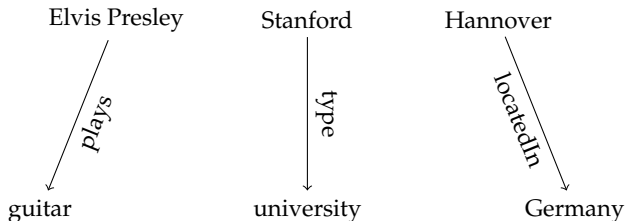
How was YAGO
constructed?

How precise is
YAGO?

How to use
YAGO?

Who uses YAGO?

Future work



What is a knowledge base?

YAGO:
Multilingual
Knowledge Base

Thomas Rebele

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

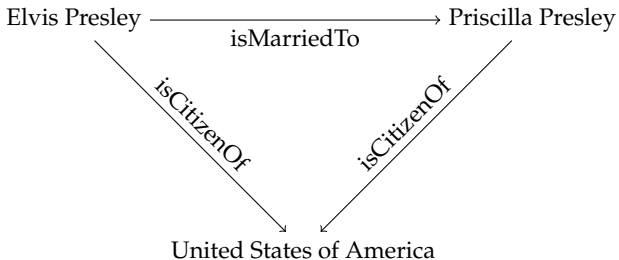
How was YAGO
constructed?

How precise is
YAGO?

How to use
YAGO?

Who uses YAGO?

Future work



Where do knowledge bases help?

YAGO:
Multilingual
Knowledge Base

Thomas Rebele

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

How was YAGO
constructed?

How precise is
YAGO?

How to use
YAGO?

Who uses YAGO?

Future work

- ▶ question answering
- ▶ semantic search
- ▶ intelligent personal assistants
- ▶ investigative journalism
- ▶ text analysis
- ▶ machine translation

Hence, everybody does it

- ▶ Google: Knowledge Graph
- ▶ Amazon: Evi
- ▶ Microsoft: Satori

How did YAGO evolve?

YAGO:
Multilingual
Knowledge Base

Thomas Rebele

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

How was YAGO
constructed?

How precise is
YAGO?

How to use
YAGO?

Who uses YAGO?

Future work

- 2006 idea: combine entities from Wikipedia with WordNet taxonomy
- 2007 YAGO: first version
- 2008 extract facts from infoboxes
- 2010 YAGO2: temporal and geographical meta facts
- 2013 YAGO2s: modular architecture
- 2015 YAGO3: entities and facts from 10 languages



What does YAGO know?

YAGO:
Multilingual
Knowledge Base

Thomas Rebele



<Barack Obama> $\xrightarrow{\text{<wasBornOnDate>}}$ "1961-08-04"^^xsd:date

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

How was YAGO
constructed?

How precise is
YAGO?

How to use
YAGO?

Who uses YAGO?

Future work

Image source: <https://wikipedia.de>

What does YAGO know?

YAGO:
Multilingual
Knowledge Base

Thomas Rebele

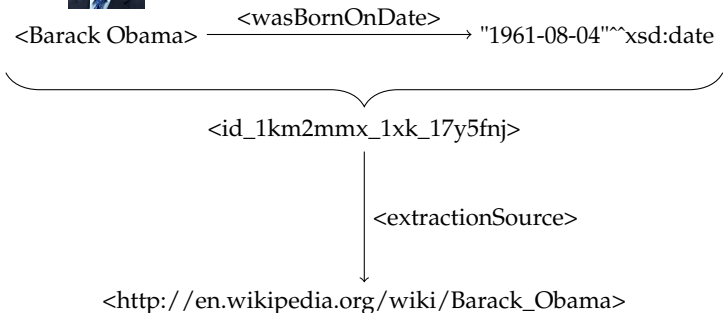


Image source: <https://wikipedia.de>

What does YAGO know?

YAGO:
Multilingual
Knowledge Base

Thomas Rebele

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

How was YAGO
constructed?

How precise is
YAGO?

How to use
YAGO?

Who uses YAGO?

Future work

theme name	number of facts
taxonomy-related facts	95 m
simplified taxonomy	17 m
main facts	55 m
GeoNames facts	39 m
meta-facts	203 m
multilingual class labels	787 k
maps to other knowledge bases	4 m
raw information Wikipedia	296 m
redirect links and anchor texts	471 m

Table: Themes of YAGO

How was YAGO constructed? - Sources

YAGO:
Multilingual
Knowledge Base

Thomas Rebele

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

How was YAGO
constructed?

Sources

Wikipedia

WordNet

GeoNames

Extraction process

How precise is
YAGO?

How to use
YAGO?

Who uses YAGO?

Future work

- Wikipedia



WIKIPEDIA
The Free Encyclopedia

- WordNet

WordNet
A lexical database for English

- GeoNames



GeoNames

How was YAGO constructed? - Wikipedia

infobox sourcecode:

```
...  
| years2 = 2003–2009 | clubs2 = [[Manchester ↵  
    United F.C.]]
```

```
...
```

YAGO:
Multilingual
Knowledge Base

Thomas Rebele

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

How was YAGO
constructed?

Sources

Wikipedia

WordNet

GeoNames

Extraction process

How precise is
YAGO?

How to use
YAGO?

Who uses YAGO?

Future work

How was YAGO constructed? - Wikipedia

YAGO:
Multilingual
Knowledge Base

Thomas Rebele

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

How was YAGO
constructed?

Sources

Wikipedia

WordNet

GeoNames

Extraction process

How precise is
YAGO?

How to use
YAGO?

Who uses YAGO?

Future work

infobox sourcecode:

```
...  
| years2 = 2003–2009 | clubs2 = [[Manchester ↵  
    United F.C.]]  
...
```

extracted facts:

```
<id_12bk998_1ul_134t38k> ↵  
    <Cristiano_Ronaldo>    <playsFor> ↵  
    <Manchester_United_F.C.>  
  
<id_111nk7s_1co_nvjcl> ↵  
    <id_12bk998_1ul_134t38k> ↵  
    <occursSince>    "2003-##-##"^^xsd:date  
  
<id_111nk7s_1cp_14wq6p3> ↵  
    <id_12bk998_1ul_134t38k> ↵  
    <occursUntil>    "2009-##-##"^^xsd:date
```

How was YAGO constructed? - WordNet

YAGO:
Multilingual
Knowledge Base

Thomas Rebele

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

How was YAGO
constructed?

Sources

Wikipedia

WordNet

GeoNames

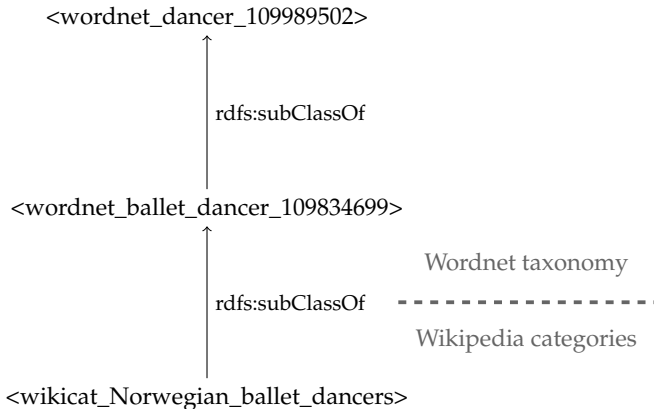
Extraction process

How precise is
YAGO?

How to use
YAGO?

Who uses YAGO?

Future work



How was YAGO constructed? - GeoNames

YAGO:
Multilingual
Knowledge Base

Thomas Rebele

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

How was YAGO
constructed?

Sources

Wikipedia

WordNet

GeoNames

Extraction process

How precise is
YAGO?

How to use
YAGO?

Who uses YAGO?

Future work

- ▶ 7m geographical entities
- ▶ provides links to Wikipedia

How was YAGO constructed? - Extraction process

YAGO:
Multilingual
Knowledge Base

Thomas Rebele

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

How was YAGO
constructed?

Sources

Wikipedia

WordNet

GeoNames

Extraction process

How precise is
YAGO?

How to use
YAGO?

Who uses YAGO?

Future work

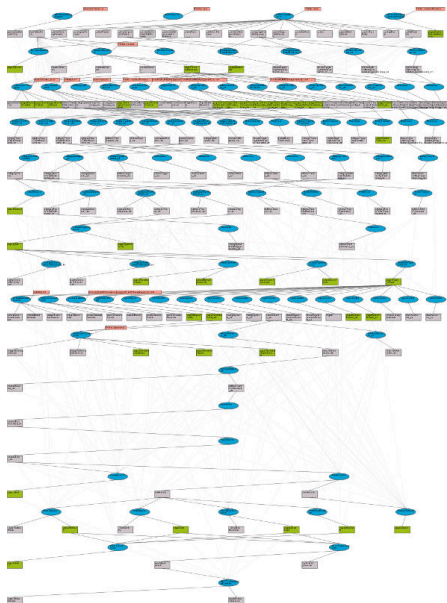
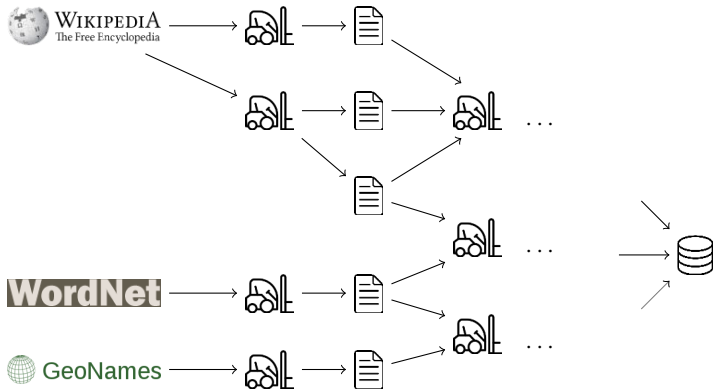


Figure: YAGO architecture

How was YAGO constructed? - Extraction process



YAGO:
Multilingual
Knowledge Base
Thomas Rebele

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

How was YAGO
constructed?

Sources

Wikipedia

WordNet

GeoNames

Extraction process

How precise is
YAGO?

How to use
YAGO?

Who uses YAGO?

Future work

How was YAGO constructed? - Extraction process

YAGO:
Multilingual
Knowledge Base
Thomas Rebele

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

How was YAGO
constructed?

Sources

Wikipedia

WordNet

GeoNames

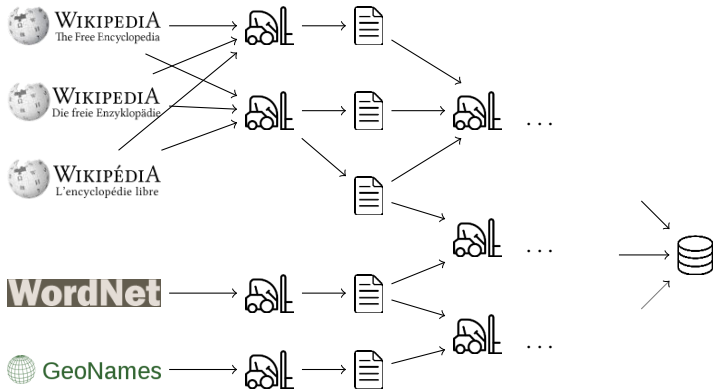
Extraction process

How precise is
YAGO?

How to use
YAGO?

Who uses YAGO?

Future work



intermediate extractors

- ▶ clean facts
- ▶ deduplicate entities and facts
- ▶ check consistency

How precise is YAGO?

YAGO3 Evaluation - Current Standings

Overall State of the Evaluation

98.07% of 4412 evaluations were judged to be correct. This gives a weighted average Wilson center of 95.03% (4.19 % width)

Evaluation Results for Relations

Evaluation Target	Evaluations	Correct	Ratio (%)	Wilson Center (%)	Wilson Width (%)	Progress
<happenedIn>	87	87	100	97.89	2.11	
<byTransport>	120	119	99.17	97.64	2.21	
<hasExpenses>	135	133	98.52	97.18	2.42	
<hasExport>	60	60	100	96.99	3.01	
<hasISBN>	59	59	100	96.94	3.06	
<exports>	58	58	100	96.89	3.11	
<isMarriedTo>	57	57	100	96.84	3.16	

Figure: Screenshot of evaluation result

YAGO:
Multilingual
Knowledge Base

Thomas Rebele

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

How was YAGO
constructed?

How precise is
YAGO?

How to use
YAGO?

Who uses YAGO?

Future work

How to use YAGO? - Formats

YAGO:
Multilingual
Knowledge Base
Thomas Rebele

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

How was YAGO
constructed?

How precise is
YAGO?

How to use
YAGO?

Formats

Tools

Who uses YAGO?

Future work

download:

- ▶ homepage of Max-Planck Institute for Informatics
(w3id.org/yago)
- ▶ Datahub (datahub.io/dataset/yago)

formats:

- ▶ **TSV**

```
<id_n4kpe2_oyl_126n5hd> <Elvis_Presley> ↔  
    <wasBornIn>         <Tupelo , _Mississippi>
```

- ▶ **TTL**

```
#@ <id_n4kpe2_oyl_126n5hd>  
<Elvis_Presley> <wasBornIn> ↔  
    <Tupelo , _Mississippi> .
```


How to use YAGO? - Tools

YAGO:
Multilingual
Knowledge Base

Thomas Rebele

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

How was YAGO
constructed?

How precise is
YAGO?

How to use
YAGO?

Formats

Tools

Who uses YAGO?

Future work

YAGO 2 spotlx

Query

Id	Subject	Property	Object	Time	
?id0:	<input type="text" value="?x"/>	<input type="text" value="rdf:type"/>	<input type="text" value="scientist"/>	<input type="text" value=""/>	
?id1:	<input type="text" value="?x"/>	<input type="text" value="rdf:type"/>	<input type="text" value="politician"/>	<input type="text" value=""/>	
?id2:	<input type="text" value="?x"/>	<input type="text" value="<wasBornIn>"/>	<input type="text" value="?p"/>	<input type="text" value="before"/>	<input type="text" value="1900"/>
?id3:	<input type="text" value=""/>	<input type="text" value=""/>	<input type="text" value=""/>	<input type="text" value=""/>	
?id4:	<input type="text" value=""/>	<input type="text" value=""/>	<input type="text" value=""/>	<input type="text" value=""/>	

Results

	Id	Subject	Property	Object
1	null	<Albert Gallatin>	rdf:type	<wordnet.scientist.1105606>
	null	<Albert Gallatin>	rdf:type	<wordnet.politician.1104512>

Figure: SPOTLX Browser [▷ w3id.org/yago/spotlx](http://w3id.org/yago/spotlx)

How to use YAGO? - Tools

YAGO:
Multilingual
Knowledge Base

Thomas Rebele

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

How was YAGO
constructed?

How precise is
YAGO?

How to use
YAGO?

Formats

Tools

Who uses YAGO?

Future work

FEEDBACK

SPARQL

Javascript

HTML

Matlab

Python

R

Ruby

PHP

Wiki

Endpoint :

<https://linkeddata1.calcul.u-psud.fr/sparql>

Copy

Query :

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX yago: <http://yago-knowledge.org/resource/>
```

Copy

```
SELECT ?person, ?date
WHERE
{
  ?actor rdfs:label "actor"@eng .
  ?person rdf:type ?actor ;
    yago:isCitizenOf yago:France ;
    yago:wasBornOnDate ?date .
  FILTER (?date < "1931-01-10"^^xsd:dateTime)
  FILTER NOT EXISTS { ?person yago:diedOnDate ?x }
  FILTER NOT EXISTS { ?person yago:diedIn ?y }
}
```

Figure: SPARQL endpoint ▷ w3id.org/yago/sparql

Who uses YAGO? - DBpedia



YAGO:
Multilingual
Knowledge Base

Thomas Rebele

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

How was YAGO
constructed?

How precise is
YAGO?

How to use
YAGO?

Who uses YAGO?

DBpedia

IBM Watson

AIDA

Semantic Culturomics

Future work

- ▶ research and community project
- ▶ knowledge base built from Wikipedia
- ▶ imported YAGO's taxonomy

Who uses YAGO? - IBM Watson

YAGO:
Multilingual
Knowledge Base

Thomas Rebele



What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

How was YAGO
constructed?

How precise is
YAGO?

How to use
YAGO?

Who uses YAGO?

DBpedia

IBM Watson

AIDA

Semantic Culturomics

Future work

- ▶ participated in *Jeopardy* (TV quizz show)
- ▶ first place (competitors: human champions)
- ▶ used YAGO's type hierarchy

Who uses YAGO? - AIDA



YAGO:
Multilingual
Knowledge Base

Thomas Rebele

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

How was YAGO
constructed?

How precise is
YAGO?

How to use
YAGO?

Who uses YAGO?

DBpedia

IBM Watson

AIDA

Semantic Culturomics

Future work



Figure: Named Entity Disambiguation with AIDA

▷ github.com/yago-naga/aida

Who uses YAGO? - Semantic Culturomics

in cooperation with Le Monde

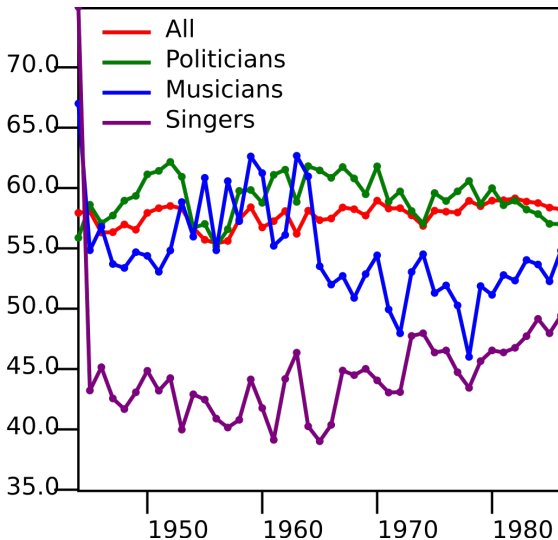


Figure: Mentions in Le Monde by occupation [▷ w3id.org/yago/mpi/le-monde](http://w3id.org/yago/mpi/le-monde)

YAGO:
Multilingual
Knowledge Base

Thomas Rebele

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

How was YAGO
constructed?

How precise is
YAGO?

How to use
YAGO?

Who uses YAGO?

DBpedia

IBM Watson

AIDA

Semantic Culturomics

Future work

Who uses YAGO? - Semantic Culturomics

in cooperation with Le Monde

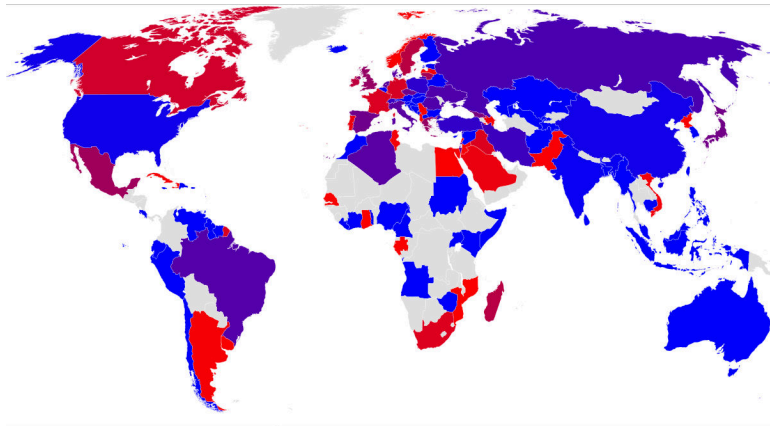


Figure: Importance of the capital [▷ w3id.org/yago/mpi/le-monde](http://w3id.org/yago/mpi/le-monde)

YAGO:
Multilingual
Knowledge Base

Thomas Rebele

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

How was YAGO
constructed?

How precise is
YAGO?

How to use
YAGO?

Who uses YAGO?

DBpedia

IBM Watson

AIDA

Semantic Culturomics

Future work

Future work

YAGO:
Multilingual
Knowledge Base

Thomas Rebele

What is a
knowledge base?

Where do
knowledge bases
help?

How did YAGO
evolve?

What does YAGO
know?

How was YAGO
constructed?

How precise is
YAGO?

How to use
YAGO?

Who uses YAGO?

Future work

- ▶ improve release cycle
- ▶ textual extension
- ▶ commonsense knowledge
- ▶ intensional knowledge
- ▶ completeness