

# Multilingual Disambiguation of Named Entities Using Linked Data

79



### AGDISTIS - AGNOSTIC DISAMBIGUATION OF NAMED ENTITIES USING LINKED OPEN DATA

Ricardo Usbeck<sup>1,2</sup>, Axel-Cyrille Ngonga Ngomo<sup>1</sup>, Michael Röder<sup>1,2</sup>, Andreas Both<sup>2</sup>  
<sup>1</sup> University of Leipzig, Germany, <sup>2</sup> R & D, Unister GmbH, Germany | {usbeck|ngonga}@informatik.uni-leipzig.de

#### Overview

AGDISTIS –

- constitutes an accurate and scalable framework for disambiguating named entities that is agnostic to the underlying knowledge base (KB).
- it outperforms the state of the art by up to 29% *F*-measure.
- has quadratic time complexity.
- is evaluated on eight well-known and diverse open-source datasets.
- combines the HTS algorithm [1] with label expansion and string similarity measures.

#### Demo

**AGDISTIS**  
Multilingual Disambiguation of Named Entities Using Linked Data

Home About Data Usage Comment Line Usage

German Examples English Examples English Untranslated Chinese Examples

Mark the entities with square brackets.

AGDISTIS is a software tool for disambiguating named entities (NEs) in text. It is based on the Free Base of Knowledge (FB) and the DBpedia Spotlight. It is one of the most accurate tools in the world and the most efficient one in terms of execution time. It is available under the GNU General Public License (GPL). For more information, please visit the AGDISTIS website.

Get Dataset Download

Figure 1: Demo available at <http://agdistis.aksw.org/demo>

#### Example

Input document:  
Example 1 Barack Obama arrived this afternoon in Washington, D.C. President Obama's wife accompanied him.

#### AGDISTIS vs. State of the Art

Dataset	Approach	F1-measure	Precision	Recall
AIDA/CO-NLL-TestB	TagMe 2	0.565	0.58	0.551
	DBpedia Spotlight	0.341	0.308	0.384
	AGDISTIS	<b>0.596</b>	<b>0.642</b>	<b>0.556</b>
AQUAINT	TagMe 2	0.457	0.412	0.514
	DBpedia Spotlight	0.26	0.178	0.48
	AGDISTIS	<b>0.547</b>	<b>0.777</b>	0.422
ITB	TagMe 2	0.408	0.416	0.4
	DBpedia Spotlight	<b>0.46</b>	0.434	<b>0.489</b>
	AGDISTIS	0.31	<b>0.646</b>	0.204
MSNBC	TagMe 2	0.466	0.431	0.508
	DBpedia Spotlight	0.331	0.317	0.347
	AGDISTIS	<b>0.761</b>	<b>0.796</b>	<b>0.729</b>

Table 1: Performance of AGDISTIS, DBpedia Spotlight and TagMe 2 on four different datasets using micro F1-measure (F1).

#### Architecture

#### Acknowledgements

Parts of this work were supported by the ESF and the Free State of Saxony and the FP7 project GeoKnow (GA No. 318159).

AKSW UNISTER ESF

<http://agdistis.aksw.org/demo>