Techreport for GERBIL 1.2.2 - V1

Michael Röder, Ricardo Usbeck, Axel-Cyrille Ngonga Ngomo

February 21, 2016

Current Development of GERBIL

Recently, we released the latest version 1.2.2 of GERBIL [16][1].

Experiment Types and Improved Diagnostics

Since the initial release of GERBIL, we added four tasks to the list of available experiment types. First, we supported the OKE Challenge 2015 [9] by adding the first and second task, i.e, named entity recognition, typing and linking as well as entity type annotation (cf. CETUS [13]), to GERBIL's experiments types. Here, we also integrated a hierarchical f-measure for the evaluation of the new typing task. The challenge requirements led to the introduction of the new concept of sub-tasks. Thus, users are now able to analyze whether the linking or the recognition step of an annotator caused the most problems. Second, we directly derived two tasks, namely Entity Recognition and Entity Typing from the two tasks above.

In addition, GERBIL now contains improved diagnostic capabilities such as the possibility for runtime measurements. Moreover, we added the calculation of correlations of dataset features and annotator performance as well as different measures for better analysis of the annotator performance, such as the distinction between entities known to a knowledge base or emerging entities [4].

We removed the separation between Sa2KB and A2KB as well as between Sc2KB and C2KB. The usage of confidence scores is not a part of the A2KB and C2KB experiments. If an annotator adds confidence scores to its annotations, GERBIL searches a threshold that optimizes the micro F1-score.

Datasets

The implementation of the OKE challenge tasks also added six new datasets designed for the challenge, see Table 1. The datasets were manually created and approved by at least two domain experts and contain NIF-based annotations for RDF entities and classes, cf. the OKE challenge documentation for further details [9]. Overall, GERBIL now contains 19 individual datasets available to 7 experiment types, see Table 2.

[1] `https://github.com/AKSW/gerbil/releases/tag/v1.2.2`

Dataset	Avg. Entities	Avg. Document Length	#Documents	#Entities
OKE 2015 Task 1				
Example set	4.000	22.333	3	12
Evaluation dataset	6.574	30.337	101	664
Gold standard sample	3.589	20.484	95	341
OKE 2015 Task 2				
Example set	3.000	27.500	2	6
Evaluation dataset	3.081	36.919	99	305
Gold standard sample	3.030	19.222	99	300

Table 1: Features of novel datasets and their documents. As the names suggest, the different datasets can be used either for Task 1 or Task 2 of the OKE Challenge.

	A2KB, C2KB, D2KB, Entity Recognition	**Entity Typing**	**OKE Task 1**	**OKE Task2**
AIDA/CoNLL-Complete	✓			
AIDA/CoNLL-Test A	✓			
AIDA/CoNLL-Test B	✓			
AIDA/CoNLL-Training	✓			
AQUAINT	✓			
DBpediaSpotlight	✓			
IITB	✓			
KORE50	✓			
MSNBC	✓			
Microposts 2014-Test	✓			
Microposts 2014-Train	✓			
N3-RSS-500	✓			
N3-Reuters-128	✓			
OKE 2015 Task 1 evaluation dataset	✓	✓	✓	
OKE 2015 Task 1 example set	✓	✓	✓	
OKE 2015 Task 1 gold standard sample	✓	✓	✓	
OKE 2015 Task 2 evaluation dataset				✓
OKE 2015 Task 2 example set				✓
OKE 2015 Task 2 gold standard sample				✓

Table 2: Novel datasets and their availability to the experiments types.

Annotators

With the current version 1.2.2 we added six new annotators compared to the initial GERBIL version 1.0, see Table 3. While CETUS, CETUS_FOX [13] and FRED [1] where added to take part in the OKE challenge 2015, FREME e-Entity[2], `entityclassifier.eu` [2] and AIDA [5] where added to enhance the spectrum of A2KB annotators.

[2] `https://github.com/freme-project/e-Entity`

		BAT-Framework	GERBIL 1.0	GERBIL 1.2.1
[7]	Wikipedia Miner	✓	✓	✓
[11]	Illionois Wikifier	✓		
[6]	Spotlight	✓	✓	✓
[3]	TagMe 2	✓	✓	✓
[14]	KEA		✓	✓
[10]	WAT		✓	✓
[15]	AGDISTIS	(✓)	✓	✓
[8]	Babelfy		✓	✓
[12]	NERD-ML		✓	✓
	NIF-based Annotator		✓	✓
[5]	AIDA	✓		✓
[2]	`entityclassifier.eu`		✓	✓
	FREME e-Entity			✓
[13]	CETUS/CETUS_FOX			✓
[1]	FRED			✓

Table 3: Overview of implemented annotator systems. Brackets indicate the existence of the implementation of the adapter but also the inability to use it in the live system.

Contribution to the Community

One of GERBIL's main goals was to provide the community with an online benchmarking tool that provides archivable and comparable experiment URIs. Thus, the impact of the framework can be measure by analyzing the interactions on the platform itself. Since its first public release on the 17th October 2014 until the 15th February 2015, 1.824 experiments were started on the platform containing more than 12.466 tasks for annotator-dataset pairs. One interesting aspect is the usage of the different provided systems, especially the heavy exploitation of the possibility to test NIF-based webservices, see Table 4.

References

[1] S. Consoli and D. Recupero. Using FRED for Named Entity Resolution, Linking and Typing for Knowledge Base Population. In F. Gandon, E. Cabrio, M. Stankovic, and A. Zimmermann, editors, *Semantic Web Evaluation Challenges*, volume 548 of *Communications in Computer and Information Science*, pages 40–50. Springer International Publishing, 2015.

[2] M. Dojchinovski and T. Kliegr. Entityclassifier.eu: Real-time classification of entities in text with wikipedia. In *Proceedings of the European Conference on Ma-*

Annotator	Number of Tasks
NIF-based Annotators	2519
Babelfy	958
DBpedia Spotlight	922
TagMe 2	811
WAT	787
Kea	763
Wikipedia Miner	714
NERD-ML	639
Dexter	587
AGDISTIS	443
Entityclassifier.eu NER	410
FOX	352
Cetus	1

Table 4: Number of tasks executed per annotator. By caching results we did not need to execute 12466 tasks but only 9906.

chine Learning and Principles and Practice of Knowledge Discovery in Databases, ECMLPKDD'13, pages 1–1. Springer-Verlag, 2013.

[3] P. Ferragina and U. Scaiella. Fast and accurate annotation of short texts with wikipedia pages. *IEEE software*, 29(1), 2012.

[4] J. Hoffart, Y. Altun, and G. Weikum. Discovering emerging entities with ambiguous names. In *Proceedings of the 23rd International Conference on World Wide Web*, WWW '14, pages 385–396, New York, NY, USA, 2014. ACM.

[5] J. Hoffart, M. A. Yosef, I. Bordino, H. Fürstenau, M. Pinkal, M. Spaniol, B. Taneva, S. Thater, and G. Weikum. Robust Disambiguation of Named Entities in Text. In *Conference on Empirical Methods in Natural Language Processing, EMNLP 2011, Edinburgh, Scotland*, pages 782–792, 2011.

[6] P. N. Mendes, M. Jakob, A. Garcia-Silva, and C. Bizer. DBpedia Spotlight: Shedding Light on the Web of Documents. In *I-Semantics*, 2011.

[7] D. Milne and I. H. Witten. Learning to link with wikipedia. In *17th ACM CIKM*, pages 509–518, 2008.

[8] A. Moro, A. Raganato, and R. Navigli. Entity linking meets word sense disambiguation: A unified approach. *TACL*, 2, 2014.

[9] A.-G. Nuzzolese, A. Gentile, V. Presutti, A. Gangemi, D. Garigliotti, and R. Navigli. Open knowledge extraction challenge. In F. Gandon, E. Cabrio, M. Stankovic, and A. Zimmermann, editors, *Semantic Web Evaluation Challenges*, volume 548

of *Communications in Computer and Information Science*, pages 3–15. Springer International Publishing, 2015.

[10] F. Piccinno and P. Ferragina. From TagME to WAT: a new entity annotator. In *Proceedings of the first international workshop on Entity recognition & disambiguation*, 2014.

[11] L. Ratinov, D. Roth, D. Downey, and M. Anderson. Local and global algorithms for disambiguation to wikipedia. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 1375–1384, Portland, Oregon, USA, June 2011. Association for Computational Linguistics.

[12] G. Rizzo, M. van Erp, and R. Troncy. Benchmarking the extraction and disambiguation of named entities on the semantic web. In *9th LREC*, 2014.

[13] M. Röder, R. Usbeck, R. Speck, and A.-C. Ngonga Ngomo. CETUS – A Baseline Approach to Type Extraction. In *1st Open Knowledge Extraction Challenge at 12th European Semantic Web Conference (ESWC 2015)*, 2015.

[14] N. Steinmetz and H. Sack. Semantic multimedia information retrieval based on contextual descriptions. In *The Semantic Web: Semantics and Big Data*. 2013.

[15] R. Usbeck, A.-C. Ngonga Ngomo, S. Auer, D. Gerber, and A. Both. AGDISTIS – Graph-Based Disambiguation of Named Entities using Linked Data. In *International Semantic Web Conference (ISWC)*. 2014.

[16] R. Usbeck, M. Röder, A.-C. Ngonga Ngomo, C. Baron, A. Both, M. Brümmer, D. Ceccarelli, M. Cornolti, D. Cherix, B. Eickmann, P. Ferragina, C. Lemke, A. Moro, R. Navigli, F. Piccinno, G. Rizzo, H. Sack, R. Speck, R. Troncy, J. Waitelonis, and L. Wesemann. GERBIL – General Entity Annotation Benchmark Framework. In *24th WWW conference*, 2015.