

diaNED: Time-Aware Named Entity Disambiguation for Diachronic Corpora

Prabal Agarwal¹, Jannik Strötgen^{1,2}, Luciano del Corro³, Johannes Hoffart³, Gerhard Weikum¹

¹Max Planck Institute for Informatics, Saarland Informatics Campus, Saarbrücken, Germany

²Bosch Center for Artificial Intelligence, Renningen, Germany

³Ambiverse GmbH, Saarbrücken, Germany

{pagarwal, weikum}@mpi-inf.mpg.de

jannik.stroetgen@de.bosch.com

{luciano, johannes}@ambiverse.com

Abstract

Named Entity Disambiguation (NED) systems perform well on news articles and other texts covering a specific time interval. However, NED quality drops when inputs span long time periods like in archives or historic corpora. This paper presents the first time-aware method for NED that resolves ambiguities even when mention contexts give only few cues. The method is based on computing temporal signatures for entities and comparing these to the temporal contexts of input mentions. Our experiments show superior quality on a newly created diachronic corpus.¹

1 Introduction

Problem. *Schumacher convinced to win on Sunday.* When this news headline is fed into modern tools for Named Entity Disambiguation (NED), virtually all of them would map the mention *Schumacher* onto the former Formula One champion *Michael Schumacher*, as the best-fitting entity from a Wikipedia-centric knowledge base (KB). However, knowing that *Sunday* refers to August 14, 1949, i.e., ignoring the surface form but exploiting normalized information, it becomes clear that the text actually refers to the German politician *Kurt Schumacher*. State-of-the-art NED methods (see surveys by Hachey et al. (2013), Ling et al. (2015), Shen et al. (2015)) tend to miss this because they are designed and trained for temporally focused input corpora such as current news, and do not cope well with longitudinal archives and other diachronic corpora that span decades. Standard NED benchmarks from CoNLL and TAC do not reflect this difficulty either.

¹The diaNED corpus and the temporal signatures of entities are publicly available: <https://www.mpi-inf.mpg.de/yago-naga/dianed/>.

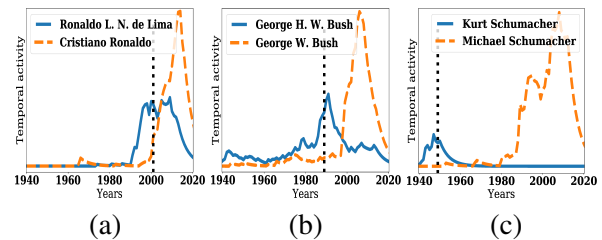


Figure 1: Temporal signatures of candidate entities for the following three sample sentences (vertical lines represent temporal contexts):

- Ronaldo comeback cut to 14 minutes.* (2001)
- Bush to stress domestic issues in speech.* (1989)
- Schumacher convinced to win on Sunday.* (1949)

What is needed here is a better way of capturing *temporal context*, for both the mention *Schumacher* and each of the candidate entities. Figure 1 illustrates “time profiles” for sample entities with highly ambiguous names. Normalized temporal information from the input context, such as *Sunday* (1949-08-14), can provide additional cues for proper disambiguation. The problem addressed in this paper is how to model and capture temporal contexts and how to enhance NED with this novel asset.

Contribution. Our approach to this problem is to compute *temporal signatures* for entities in the KB, and to use these as expressive features when comparing candidate entities against the context of an input mention. Temporal signatures are embeddings that reflect the importance of different years for entities. They are automatically constructed by extracting and normalizing temporal expressions in entity descriptions such as Wikipedia articles. Analogously, temporal signals are captured in the contexts of textual mentions and represented by embeddings.

The time-aware NED method that we devise with these features can robustly cope with inputs

from diachronic corpora. We propose a new evaluation benchmark, based on the New York Times Archive, spanning more than 20 years, and the history collection *historynet.com*, spanning several centuries. Our experiments demonstrate that time-aware NED substantially outperforms some of the best standard NED tools.

2 Temporal Signatures and Contexts

Better context representation improves disambiguation quality (see, e.g., Shen et al. (2015)). As the underlying entity descriptions (e.g., Wikipedia articles) are not only textually but also temporally related to their mentions, we enrich the context representation with a temporal dimension, which no prior work handles explicitly.

We model the temporal dimension by embedding vectors. The embeddings represent the *temporal signatures* of entities in a KB and the *temporal contexts* of entity mentions in text in a joint vector space. Then, the similarity between them quantifies their temporal relatedness.

Temporal vector space. We use 2,050 dimensions (years 1 AD to 2050) to define the vector space. Coarser and finer granularities than years could be used, but finer ones (e.g., days) are rarely needed for NED and coarser granularities (e.g., centuries) are too vague.²

Temporal signatures of entities. We use the temporal tagger HeidelTime (Strötgen and Gertz, 2010; Strötgen and Gertz, 2015) to extract and normalize date expressions from an entity’s Wikipedia page³ and aggregate them by years. This results in a count-based temporal vector $\mathbf{t}_e^{cb} = (t_{0001}, \dots, t_i, \dots, t_{2050})$ where t_i is the total number of temporal expressions extracted from e ’s Wikipedia page referring to year i . Temporal expressions of finer granularities are mapped to respective years and expressions of coarser granularities than year are currently ignored.

As the count-based vectors may overfit to the entity descriptions and to avoid discontinuity in the temporal signatures, we apply exponential smoothing and get smoothed temporal vectors

²In an analysis of temporal expressions extracted with HeidelTime from the Wikipedia corpus (August 2016 dump), we find that there are on average 18.500 expressions per year value (with year values ranging from 0001 AD to 2050 AD) in contrast to only 9.64 expressions per day value (with day values ranging from 0001-01-01 to 2050-12-31). Therefore, using year level identifiers to define our temporal vector space results in short and non-sparse temporal vectors.

³August 2016 Wikipedia dump

$\mathbf{t}_e^s = (t_{0001}^s, \dots, t_i^s, \dots, t_{2050}^s)$ such that $t_i^s = \alpha \cdot t_i^{cb} + (1 - \alpha) \cdot t_{i-1}^s$, for $i > 0001$ where α is the smoothing factor with $0 \leq \alpha \leq 1$. For further smoothing, this procedure can be recursively applied n times. In experiments, we set $\alpha = 0.2$ and $n = 2$ based on cross-validation.

Temporal contexts of entity mentions. We exploit temporal expressions in the surrounding text of entity mentions and the texts’ publication dates. In news-style articles, entities are likely to be related to the document creation time (dct), while dates in the content are important for other types of documents (Strötgen and Gertz, 2016).

Temporal vectors for mentions \mathbf{t}_m are thus a combination of a one-hot temporal vector $\mathbf{t}_m^{dct} = (0, \dots, t_i, \dots, 0)$ where $t_i=1$ if i is the dct’s year, and $\mathbf{t}_m^{content}$ containing dates extracted by a temporal tagger in the immediate context of the mention (e.g., in the same sentence or paragraph), aggregated by year. \mathbf{t}_m^{dct} and $\mathbf{t}_m^{content}$ are linearly combined as $\mathbf{t}_m = \lambda \cdot \mathbf{t}_m^{dct} + (1 - \lambda) \cdot \mathbf{t}_m^{content}$ where λ (with $0 \leq \lambda \leq 1$) weights the components.

Relatedness. We calculate the temporal relatedness between a mention and all candidate entities as the cosine similarity between \mathbf{t}_m and \mathbf{t}_e .

3 Time-aware NED

To test the importance of time-awareness for NED, we use two settings. We enhance a basic NED system and a state-of-the-art system by enriching both with temporal signatures and contexts.

diaNED-1, as basic NED system, uses a mention-entity prior reflecting entity prominence and a keyphrase-based language model for the similarity of mention and entity contexts (as suggested by Hoffart et al. (2011)). These components are cast into edge weights for a graph over which the final disambiguation is computed. Hyper-parameters for the relative influence of the two components are tuned using an SVM.

We added the temporal dimension to the feature set and retrained the model accordingly to get new feature weights.

diaNED-2 based on Yamada et al. (2016): This is a learning-to-rank-based model. Besides mention-entity priors and string-similarity features, it uses word and entity embeddings trained in a joint vector space to model context and coherence. The intuition is that a good candidate entity vector must be close to the word and entity vectors appearing in the same context.

Yamada et al. (2016) measures entity context by averaging the word vectors of the proper noun neighbors and calculating the cosine similarity with each candidate entity. Similarly, the coherence between entities is measured by computing the cosine similarity between candidates and the average of the other entities in the neighborhood.

diaNED-2 enhances this model as follows. We compute the cosine similarity between the mention’s and the candidate entities’ temporal vectors, and normalize the time relatedness scores across candidate entities. Finally, all similarity features are used to train a binary classifier with gradient-boosted decision trees. The top-ranked candidate entity in each pool of candidates is assigned to the mention being evaluated.

4 A Diachronic NED Data Set

Datasets for NED evaluation contain articles published within a short period. Consequently, all mentions share a temporal context making it difficult to evaluate temporal variability. CoNLL-AIDA (Hoffart et al., 2011) are newswire articles from 1996, TAC 2010 (Ji et al., 2010) news and forum articles from 2004–2007, and Microposts-2014 (Cano et al., 2014) tweets from 2011.

To account for this limitation, we create a new diachronic benchmark containing documents with heterogeneous temporal context. As in Microposts-2014, we limit documents to single sentences and headlines from *HistoryNet.com* (*HN*) and *The New York Times corpus* (*NYT*). For the annotation process, we followed the entity annotation guidelines, which have been used for annotating CoNLL-AIDA (Hoffart et al., 2011).

HN is an online resource of world history with information on popular historical topics. Its section *Today in History* contains short texts on what happened on a specific day with a total of 7,061 facts/events (excluding *born today*). Using Stanford NER (Finkel et al., 2005), we extracted 13,773 entity mentions and randomly selected 350 of them. We annotated all entity mentions in respective sentences with their Wikipedia ids. After removing NER errors and out-of-KB entities, the dataset contains 865 gold entity mentions in 334 sentences. Examples are: “*Conrad II* claims the throne in *France*” from 1032 or “*The Old Pretender*, son of *James III* dies” from 1766.

NYT contains more than 1.5 million documents published between 1987 and 2007. After apply-

ing the same procedure, the dataset contains 368 manually annotated mentions in 290 news headlines. Examples are “*Arafat’s* Faction is Said to Avoid Guerrilla Actions” from 1989 or “*U.N.* Aide to Meet *Milosevic*, Angering Some” from 1999.

As *HN* texts come without further context, entity mentions are rather explicit. Entity mentions in *NYT*’s headlines are more ambiguous as more information is available in the articles and the entities are mostly, at the time of publication, prominent and obvious to the reader.

Finally, we created a third subset from the 7,061 documents of *HistoryNet.com* with 13,773 entity mentions. It contains the sentences with all the entity mentions which are linked to different entities by diaNED-2 depending on whether it uses its time-awareness or not, i.e., whether diaNED-2 is trained with or without the temporal feature. This set (*HN-timediff*) contains 567 manually annotated entities from 547 documents. It is the most challenging subset as all entity mentions are difficult to disambiguate.

5 Evaluation

To evaluate the importance of temporal information in NED, we focus in our analysis on the newly created *diaNED* corpus. As standard NED datasets CoNLL-AIDA and TAC 2010 contain only articles published within a short period of time, they are not suited for evaluating time-aware NED (cf. Section 4), and experiments on these datasets showed no significant differences between using diaNED-1 and diaNED-2 with or without their time-awareness features.

Note that the temporal contexts in the *HN* sentences and the *NYT* headlines of the *diaNED* corpus are part of the metadata. Thus, to ensure a fair comparison among all systems, we added the temporal contexts in the form of year information to all documents to allow the non-time-aware systems to exploit the temporal context in case the respective year number occurs as part of the entities’ textual context.⁴

5.1 Intra-system Comparison

As described above, we (re-)implemented two NED systems as diaNED-1 and diaNED-2. To al-

⁴Disambiguation quality of non-time-aware systems was generally lower without this additional information. The *diaNED* corpus contains all sentences with and without year information so that evaluation results can be reproduced for both settings.

Feature set	HN subset		NYT subset	
	w/o time	w time	w/o time	w time
Prior	72.26	80.48*	38.14	54.24*
Context	63.63	66.10*	48.31	62.71*

* significant over w/o time (Welch’s t-test at level of 0.01)

Table 1: Micro-accuracy of diaNED-1 with and without time-awareness feature.

Feature set	HN subset		NYT subset	
	w/o time	w time	w/o time	w time
Base	89.44	90.23*	85.81	87.36*
String	89.40	90.00*	86.28	87.07*
Context	91.10	91.81*	87.07	88.34*
Coherence	91.16	91.98*	86.83	88.69*

* significant over w/o time (Welch’s t-test at level of 0.01)

Table 2: Micro-accuracy of diaNED-2 with and without time-awareness feature.

low the systems to adapt to the diachronic corpus without considering temporal information explicitly, we retrained the systems so that appropriate weights are learnt for each standard feature. Due to the rather small size of the diaNED corpus, we use bootstrapping (i.e., train and evaluate on 50 randomly shuffled versions of the corpus) with and without using the time-awareness feature.

diaNED-1. Table 1 shows micro-accuracy for our basic NED system on the *HN* and *NYT* sets of diaNED. Significant gains are achieved when combining the prior and context features with the time-awareness feature. This demonstrates that NED systems with standard features can be improved by making them time-aware.

diaNED-2. Table 2 shows micro-accuracy for our re-implementation of Yamada et al. (2016)’s initial features with and without the time-awareness feature. As can be seen in the table, adding the temporal feature improves the results significantly in each setting on both sets, which demonstrates that even state-of-the-art systems can be improved by making them time-aware.

5.2 Inter-system Comparison

In Table 3, we compare the time-aware NED approach diaNED-2 to various NED tools available via GERBIL (v. 1.2.5) (Usbeck et al., 2015) and to the recent work by Gupta et al. (2017). As all systems are used with standard settings, we also trained diaNED-2 on standard NED training data (CoNLL-AIDA) with the temporal context of entity mentions being the respective article’s year

system	HN	NYT
xLisa-NGRAM (Zhang and Rettinger, 2014)	87.07	66.30
WAT (Ferragina and Scaiella, 2012)	82.26	70.95
PBOH (Ganea et al., 2016)	90.26	71.75
FREME NER (Dojchinovski and Kliegr, 2013)	48.50	45.27
FRED (Consoli and Recupero, 2015)	23.18	15.44
FOX (Speck and Ngomo, 2014)	77.85	54.25
Dexter (Ceccarelli et al., 2013)	69.66	49.12
DBpedia Spotlight (Mendes et al., 2011)	56.92	61.91
AIDA (Hoffart et al., 2011)	82.35	70.14
AGDISTIS (Usbeck et al., 2014)	70.77	50.14
Gupta et al. (2017)	62.82	43.33
reimpl. of (Yamada et al., 2016)	90.87	72.55
diaNED-2 w time	91.68	76.09

Table 3: F1-scores of various systems on the *HN* and *NYT* subsets of the diaNED benchmark.

	overall	person	location	organization
time-agnostic	27.51	9.63	40.07	33.77
time-aware	42.50	39.91	45.22	40.26

Table 4: Micro-accuracy of diaNED-2 on *HN-timediff* with and without time-awareness feature.

of publication. However, due to the differences in what kind of entities the systems consider and what kind of candidate entity lookup dictionaries they use, the systems are not directly comparable and the performance differences should be interpreted with a grain of salt. Nevertheless, time-awareness further increases the distance between (Yamada et al., 2016) and the second best system significantly, which demonstrates its usefulness for NED.

5.3 Type-based Analysis

To gain further insights about the importance of time-awareness, we analyzed the results of diaNED-2 with and without temporal feature on the *HN-timediff* set of our benchmark (Table 4). On these particularly challenging documents, the time-awareness feature helps to improve NED quality for all entity types. While location and organization entities moderately benefit, there is a huge performance increase for person entities. The explanation that person entities benefit most could be that person entities have comparably short life spans and are thus most time-sensitive.

6 Related Work

Starting with the early work of Bunescu and Paşca (2006), Cucerzan (2007), Mihalcea and Csomai (2007), and Milne and Witten (2008),

NED methods and tools have been greatly advanced and become mature. Many systems use a combination of (i) local features like string similarities, lexico-syntactic characteristics and context between mentions and candidate entities and (ii) global features like the coherence among a set of selected entities. The inference over this feature space is typically performed by probabilistic graphical models, learning-to-rank techniques or algorithms related to such models (see, e.g., Ratinov et al. (2011), Hoffart et al. (2011), Ferragina and Scaiella (2012), Cheng and Roth (2013), Guo and Barbosa (2014), Durrett and Klein (2014), Chisholm and Hachey (2015), Pershina et al. (2015), Lazic et al. (2015), Nguyen et al. (2016), Globerson et al. (2016), Eshel et al. (2017), and Ganea and Hofmann (2017)). The GERBIL framework (Usbeck et al., 2015) provides a unified way of evaluating a wide variety of NED tools and services.

A recent line of work uses representational learning to characterize contexts through embeddings (e.g., He et al. (2013), Sun et al. (2015), Francis-Landau et al. (2016), Yamada et al. (2016), Gupta et al. (2017), Yamada et al. (2017)). These approaches naturally lend themselves towards inference by neural networks such as LSTMs. In our experiments, the Neural Text-Entity Encoder by Yamada et al. (2016) serves as state-of-the-art baseline.

While temporal information was used as a global feature to compute coherence between entity lifespans (Hoffart et al., 2013), no prior work on named entity disambiguation made explicit use of temporal information as a local feature. However, the value of time has been shown in a variety of other information extraction tasks, such as relation extraction (UzZaman et al., 2013; Mirza and Tonelli, 2016), event extraction (Kuzey et al., 2016; Spitz and Gertz, 2016), and slot filling (Ji et al., 2011; Surdeanu et al., 2011; Surdeanu, 2013), as well as in the context of information retrieval (Berberich et al., 2010; Agarwal and Strötgen, 2017) and fact checking (Popat et al., 2017). In this paper, inspired by the importance of temporal information for many NLP tasks, we analyzed its value for NED.

7 Conclusions and Ongoing Work

We proposed the first NED method with explicit consideration of temporal background. As demon-

strated in our experiments, this time-awareness improves NED quality over diachronic texts that span long time periods. The diaNED dataset and the temporal signatures of entities are publicly available.⁵

Currently, we integrate a strategy for handling out-of-KB entities to determine how temporal affinity may help in the *nil* detection problem. Furthermore, we plan large-scale experiments with distant supervision data which will also allow to evaluate the effectiveness of considering temporal expressions in the context of the entity mentions as further temporal context information. Finally, using a multilingual temporal tagger (Strötgen and Gertz, 2015), the value of time for NED could be studied for further languages.

Acknowledgments

The authors thank the anonymous reviewers for their valuable comments and suggestions and Ikuya Yamada for providing the word-entity embeddings and his help in the implementation of one of the baseline systems.

References

- Prabal Agarwal and Jannik Strötgen. 2017. [Tiwiki: Searching Wikipedia with Temporal Constraints](#). In *Proceedings of the 7th Temporal Web Analytics Workshop*, TempWeb'17, pages 1595–1600.
- Klaus Berberich, Srikanta Bedathur, Omar Alonso, and Gerhard Weikum. 2010. [A Language Modeling Approach for Temporal Information Needs](#). In *Proceedings of the 32nd European Conference on Advances in Information Retrieval*, ECIR'10, pages 13–25.
- Razvan Bunescu and Marius Paşca. 2006. [Using Encyclopedic Knowledge for Named Entity Disambiguation](#). In *Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics*, EACL'06, pages 9–16.
- Amparo E. Cano, Giuseppe Rizzo, Andrea Varga, Matthew Rowe, Milan Stankovic, and Aba-Sah Dadzie. 2014. [Making Sense of Microposts \(#Microposts2014\): Named Entity Extraction & Linking Challenge](#). In *Proceedings of the 4th Workshop on Making Sense out of Microposts*, Microposts'14, pages 54–60.
- Diego Ceccarelli, Claudio Lucchese, Salvatore Orlando, Raffaele Perego, and Salvatore Trani. 2013. [Dexter: An Open Source Framework for Entity](#)

⁵<https://www.mpi-inf.mpg.de/yago-naga/dianed/>

- Linking.** In *Proceedings of the 6th International Workshop on Exploiting Semantic Annotations in Information Retrieval*, ESAIR'13, pages 17–20.
- Xiao Cheng and Dan Roth. 2013. **Relational Inference for Wikification.** In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, EMNLP'13, pages 1787–1796.
- Andrew Chisholm and Ben Hachey. 2015. **Entity Disambiguation with Web Links.** *Transactions of the Association for Computational Linguistics*, 3:145–156.
- Sergio Consoli and Diego Reforgiato Recupero. 2015. **Using FRED for Named Entity Resolution, Linking and Typing for Knowledge Base Population.** In *Proceedings of the Semantic Web Evaluation Challenge*, SemWebEval'15, pages 40–50.
- Silviu Cucerzan. 2007. **Large-Scale Named Entity Disambiguation Based on Wikipedia Data.** In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, EMNLP-CoNLL'07, pages 708–716.
- Milan Dojchinovski and Tomáš Kliegr. 2013. **Entityclassifier.eu: Real-Time Classification of Entities in Text with Wikipedia.** In *Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, ECML-PKDD'13, pages 654–658.
- Greg Durrett and Dan Klein. 2014. **A Joint Model for Entity Analysis: Coreference, Typing, and Linking.** *Transactions of the Association for Computational Linguistics*, 2:477–490.
- Yotam Eshel, Noam Cohen, Kira Radinsky, Shaul Markovitch, Ikuya Yamada, and Omer Levy. 2017. **Named Entity Disambiguation for Noisy Text.** In *Proceedings of the 21st Conference on Computational Natural Language Learning*, CoNLL'17, pages 58–68.
- Paolo Ferragina and Ugo Scaiella. 2012. **Fast and Accurate Annotation of Short Texts with Wikipedia Pages.** *IEEE Software*, 29(1):70–75.
- Jenny Rose Finkel, Trond Grenager, and Christopher Manning. 2005. **Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling.** In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, ACL'05, pages 363–370.
- Matthew Francis-Landau, Greg Durrett, and Dan Klein. 2016. **Capturing Semantic Similarity for Entity Linking with Convolutional Neural Networks.** In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, NAACL-HLT'16, pages 1256–1261.
- Octavian-Eugen Ganea, Marina Ganea, Aurelien Lucchi, Carsten Eickhoff, and Thomas Hofmann. 2016. **Probabilistic Bag-Of-Hyperlinks Model for Entity Linking.** In *Proceedings of the 25th International Conference on World Wide Web*, WWW'16, pages 927–938.
- Octavian-Eugen Ganea and Thomas Hofmann. 2017. **Deep Joint Entity Disambiguation with Local Neural Attention.** In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, EMNLP'17, pages 2609–2619.
- Amir Globerson, Nevena Lazic, Soumen Chakrabarti, Amarnag Subramanya, Michael Ringgaard, and Fernando Pereira. 2016. **Collective Entity Resolution with Multi-Focal Attention.** In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, ACL'16, pages 621–631.
- Zhaochen Guo and Denilson Barbosa. 2014. **Robust Entity Linking via Random Walks.** In *Proceedings of the 23rd ACM International Conference on Information and Knowledge Management*, CIKM'14, pages 499–508.
- Nitish Gupta, Sameer Singh, and Dan Roth. 2017. **Entity Linking via Joint Encoding of Types, Descriptions, and Context.** In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, EMNLP'17, pages 2681–2690.
- Ben Hachey, Will Radford, Joel Nothman, Matthew Honnibal, and James R. Curran. 2013. **Evaluating Entity Linking with Wikipedia.** *Artificial Intelligence*, 194:130–150.
- Zhengyan He, Shujie Liu, Mu Li, Ming Zhou, Longkai Zhang, and Houfeng Wang. 2013. **Learning Entity Representation for Entity Disambiguation.** In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, ACL'13, pages 30–34.
- Johannes Hoffart, Fabian M. Suchanek, Klaus Berberich, and Gerhard Weikum. 2013. **YAGO2: A Spatially and Temporally Enhanced Knowledge Base from Wikipedia.** *Artificial Intelligence*, 194:28–61.
- Johannes Hoffart, Mohamed Amir Yosef, Ilaria Bordino, Hagen Fürstenu, Manfred Pinkal, Marc Spaniol, Bilyana Taneva, Stefan Thater, and Gerhard Weikum. 2011. **Robust Disambiguation of Named Entities in Text.** In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, EMNLP'11, pages 782–792.
- Heng Ji, Ralph Grishman, and Hoa Dang. 2011. **Overview of the TAC2011 Knowledge Base Population Track.** In *Proceedings of the 4th Text Analysis Conference*, TAC'11.
- Heng Ji, Ralph Grishman, Hoa Trang Dang, Kira Grifitt, and Joe Ellis. 2010. **Overview of the TAC 2010 Knowledge Base Population Track.** In *Proceedings of the 3rd Text Analysis Conference*, TAC'10.

- Erdal Kuzey, Vinay Setty, Jannik Strötgen, and Gerhard Weikum. 2016. *As Time Goes By: Comprehensive Tagging of Textual Phrases with Temporal Scopes*. In *Proceedings of the 25th International Conference on World Wide Web, WWW'16*, pages 915–925.
- Nevena Lazic, Amarnag Subramanya, Michael Ringgaard, and Fernando Pereira. 2015. *Plato: A Selective Context Model for Entity Resolution*. *Transactions of the Association for Computational Linguistics*, 3:503–515.
- Xiao Ling, Sameer Singh, and Daniel Weld. 2015. *Design Challenges for Entity Linking*. *Transactions of the Association for Computational Linguistics*, 3:315–328.
- Pablo N. Mendes, Max Jakob, Andrés García-Silva, and Christian Bizer. 2011. *DBpedia Spotlight: Shedding Light on the Web of Documents*. In *Proceedings of the 7th International Conference on Semantic Systems, I-Semantics'11*, pages 1–8.
- Rada Mihalcea and Andras Csomai. 2007. *Wikify!: Linking Documents to Encyclopedic Knowledge*. In *Proceedings of the 16th ACM Conference on Information and Knowledge Management, CIKM'07*, pages 233–242.
- David Milne and Ian H. Witten. 2008. *Learning to Link with Wikipedia*. In *Proceedings of the 17th ACM Conference on Information and Knowledge Management, CIKM'08*, pages 509–518.
- Paramita Mirza and Sara Tonelli. 2016. *CATENA: Causal and Temporal Relation Extraction from Natural Language Texts*. In *Proceedings the 26th International Conference on Computational Linguistics, COLING'16*, pages 64–75.
- Dat Nguyen, Martin Theobald, and Gerhard Weikum. 2016. *J-NERD: Joint Named Entity Recognition and Disambiguation with Rich Linguistic Features*. *Transactions of the Association for Computational Linguistics*, 4:215–229.
- Maria Pershina, Yifan He, and Ralph Grishman. 2015. *Personalized Page Rank for Named Entity Disambiguation*. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT'15*, pages 238–243.
- Kashyap Popat, Subhabrata Mukherjee, Jannik Strötgen, and Gerhard Weikum. 2017. *Where the truth lies: Explaining the credibility of emerging claims on the web and social media*. In *Proceedings of the 26th International Conference on World Wide Web Companion, WWW'17*, pages 1003–1012.
- Lev Ratinov, Dan Roth, Doug Downey, and Mike Anderson. 2011. *Local and Global Algorithms for Disambiguation to Wikipedia*. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, ACL-HLT'11*, pages 1375–1384.
- Wei Shen, Jianyong Wang, and Jiawei Han. 2015. *Entity Linking with a Knowledge Base: Issues, Techniques, and Solutions*. *IEEE Transactions on Knowledge and Data Engineering*, 27(2):443–460.
- René Speck and Axel-Cyrille Ngonga Ngomo. 2014. *Named Entity Recognition Using FOX*. In *Proceedings of the International Semantic Web Conference, ISWC-PD'14*, pages 85–88.
- Andreas Spitz and Michael Gertz. 2016. *Terms over LOAD: Leveraging Named Entities for Cross-Document Extraction and Summarization of Events*. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR'16*, pages 503–512.
- Jannik Strötgen and Michael Gertz. 2010. *HeidelTime: High Quality Rule-based Extraction and Normalization of Temporal Expressions*. In *Proceedings of the 5th International Workshop on Semantic Evaluation, SemEval'10*, pages 321–324.
- Jannik Strötgen and Michael Gertz. 2015. *A Baseline Temporal Tagger for all Languages*. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, EMNLP'15*, pages 541–547.
- Jannik Strötgen and Michael Gertz. 2016. *Domain-sensitive Temporal Tagging*. Synthesis Lectures on Human Language Technologies, Morgan & Claypool Publishers.
- Yaming Sun, Lei Lin, Duyu Tang, Nan Yang, Zhenzhou Ji, and Xiaolong Wang. 2015. *Modeling Mention, Context and Entity with Neural Networks for Entity Disambiguation*. In *Proceedings of the 24th International Conference on Artificial Intelligence, IJCAI'15*, pages 1333–1339.
- Mihai Surdeanu. 2013. *Overview of the TAC2013 Knowledge Base Population Evaluation: English Slot Filling and Temporal Slot Filling*. In *Proceedings of the 6th Text Analysis Conference, TAC'13*.
- Mihai Surdeanu, Sonal Gupta, John Bauer, David McClosky, Angel X. Chang, Valentin I. Spitskovsky, and Christopher D. Manning. 2011. *Stanford's Distantly-Supervised Slot-Filling System*. In *Proceedings of the 4th Text Analysis Conference, TAC'11*.
- Ricardo Usbeck, Axel-Cyrille Ngonga Ngomo, Michael Röder, Daniel Gerber, Sandro Athaide Coelho, Sören Auer, and Andreas Both. 2014. *AGDISTIS - Graph-Based Disambiguation of Named Entities Using Linked Data*. In *Proceedings of the 13th International Semantic Web Conference, ISWC'14*, pages 457–471.
- Ricardo Usbeck, Michael Röder, Axel-Cyrille Ngonga Ngomo, Ciro Baron, Andreas Both, Martin Brümmer, Diego Ceccarelli, Marco Cornolti, Didier Cherix, Bernd Eickmann, Paolo Ferragina, Christiane Lemke, Andrea Moro, Roberto Navigli,

- Francesco Piccinno, Giuseppe Rizzo, Harald Sack, René Speck, Raphaël Troncy, Jörg Waitelonis, and Lars Wesemann. 2015. [GERBIL: General Entity Annotator Benchmarking Framework](#). In *Proceedings of the 24th International Conference on World Wide Web, WWW'15*, pages 1133–1143.
- Naushad UzZaman, Hector Llorens, Leon Derczynski, James Allen, Marc Verhagen, and James Pustejovsky. 2013. [SemEval-2013 Task 1: TempEval-3: Evaluating Time Expressions, Events, and Temporal Relations](#). In *Proceedings of the 7th International Workshop on Semantic Evaluation, SemEval'13*, pages 1–9.
- Ikuya Yamada, Hiroyuki Shindo, Hideaki Takeda, and Yoshiyasu Takefuji. 2016. [Joint Learning of the Embedding of Words and Entities for Named Entity Disambiguation](#). In *Proceedings of the 20th SIGNLL Conference on Computational Natural Language Learning, CoNLL'15*, pages 250–259.
- Ikuya Yamada, Hiroyuki Shindo, Hideaki Takeda, and Yoshiyasu Takefuji. 2017. [Learning Distributed Representations of Texts and Entities from Knowledge Base](#). *Transactions of the Association for Computational Linguistics*, 5:397–411.
- Lei Zhang and Achim Rettinger. 2014. [X-LiSA: Cross-lingual Semantic Annotation](#). *Proc. VLDB Endow.*, 7(13):1693–1696.