

Bachelorarbeit: Multiwort-Bedeutungsauflösung bei Anforderungen

Thomas Bartel, betreut von Tobias Hey

IPD Tichy, Fakultät für Informatik

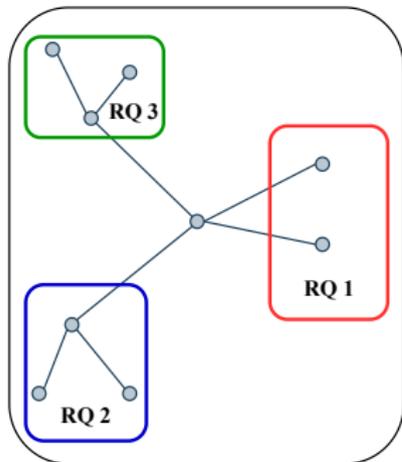


- Rückverfolgbarkeitsinformationen helfen bei Wartung und Pflege von Software
- Grundvoraussetzung: Verständnis der Anforderungen in natürlicher Sprache

Beispiel für einen Multiwort-Ausdruck (MWA)

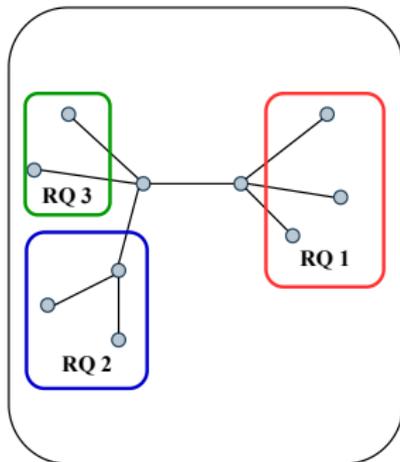
The network is unavailable because of a **denial of service attack**

Anforderungsmodell



Anforderungsspezifikationen

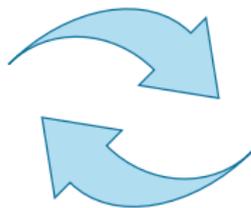
Quelltextmodell



Quelltext

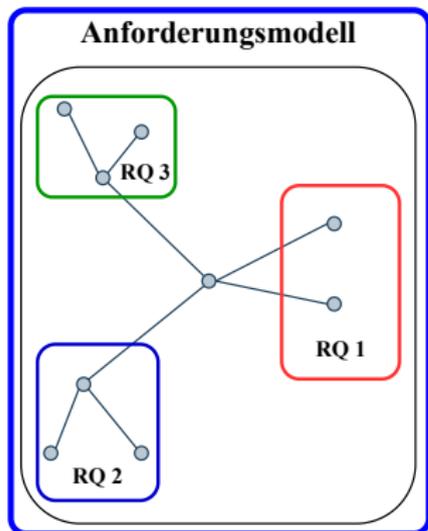


Dokumentation

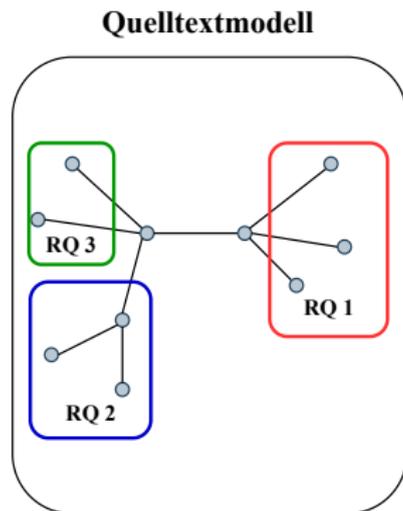
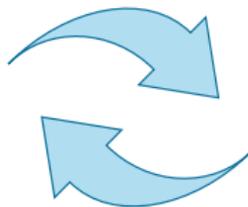


Programmanalyse

INDIRECT



Anforderungsspezifikationen



Programmanalyse

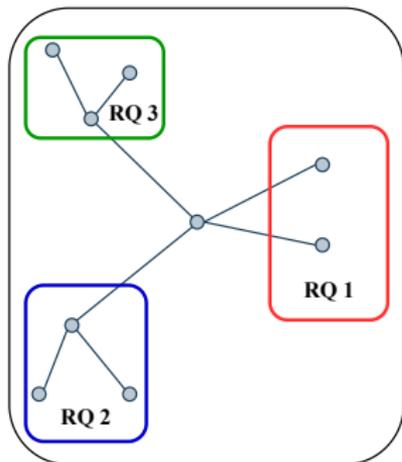


Quelltext

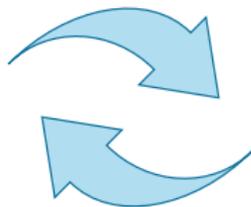


Dokumentation

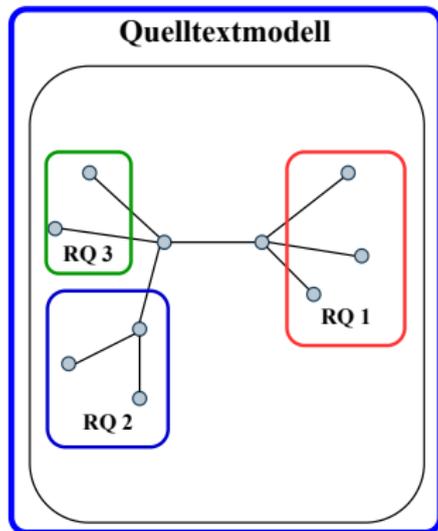
Anforderungsmodell



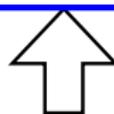
Anforderungsspezifikationen



Quelltextmodell



Programmanalyse

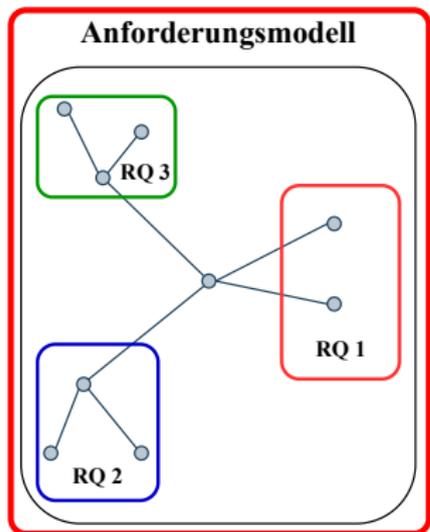


Quelltext

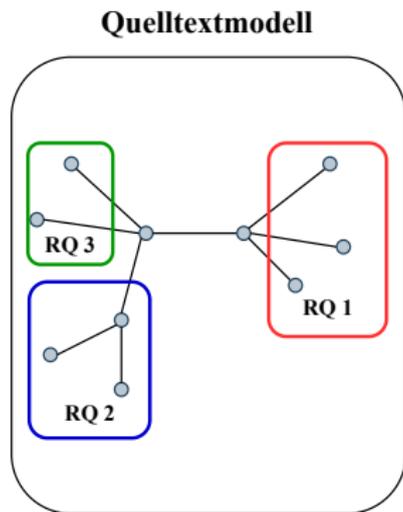
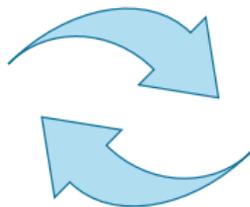


Dokumentation

INDIRECT



Anforderungsspezifikationen



Programmanalyse



Quelltext



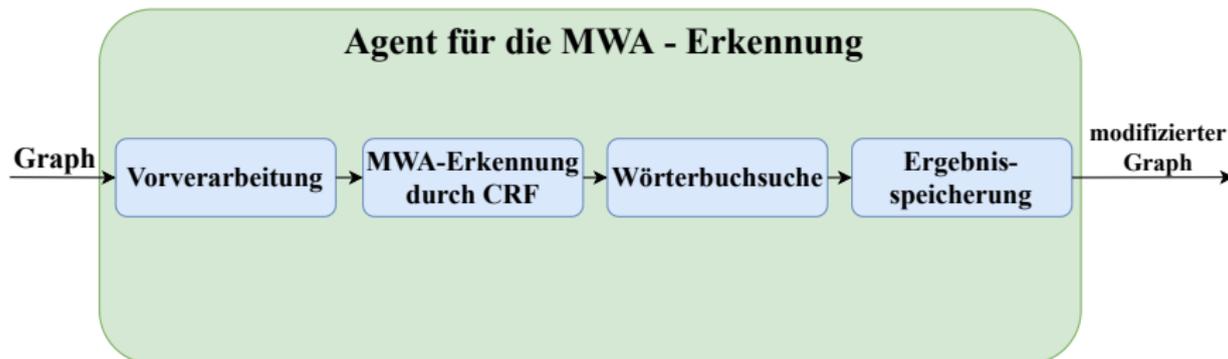
Dokumentation

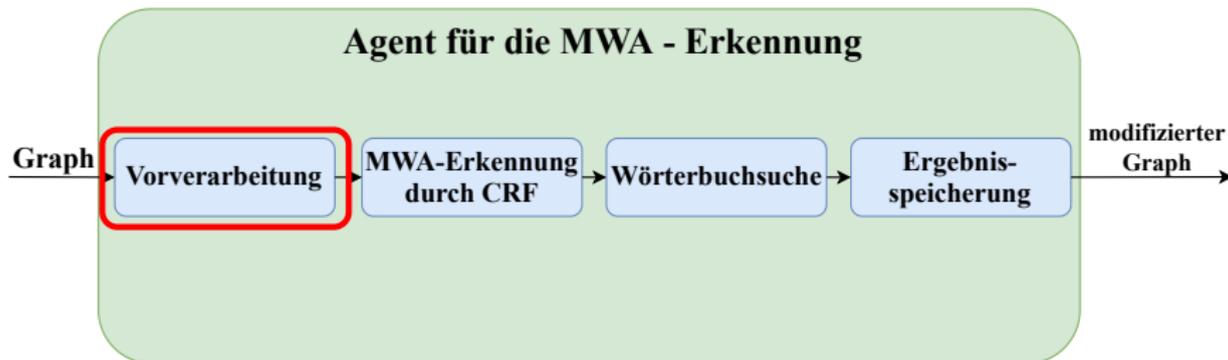
Erkennung von Multiwort-Ausdrücken

- Hosseini et al. [HSL16]
 - Erkennung von MWAs anhand eines doppelt verketteten Zufallsfeldes
- Cordeiro et al. [CRV16]
 - Erkennung von MWAs anhand von Wortartmarkierungsmustern

Wortbedeutungsauflösung

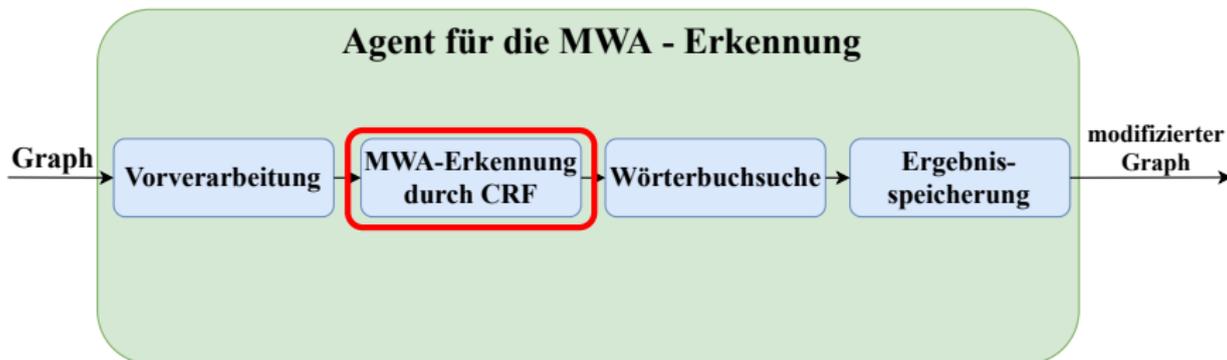
- Agirre und Soroa [AS09; ALS18]
 - Graphbasierte Wortbedeutungsauflösung anhand eines personalisierten PageRanks (UKB)
- Moro et al. [MRN14]
 - Graphbasierte Wortbedeutungsauflösung mit Wikipedia und WordNet (Babelfy)





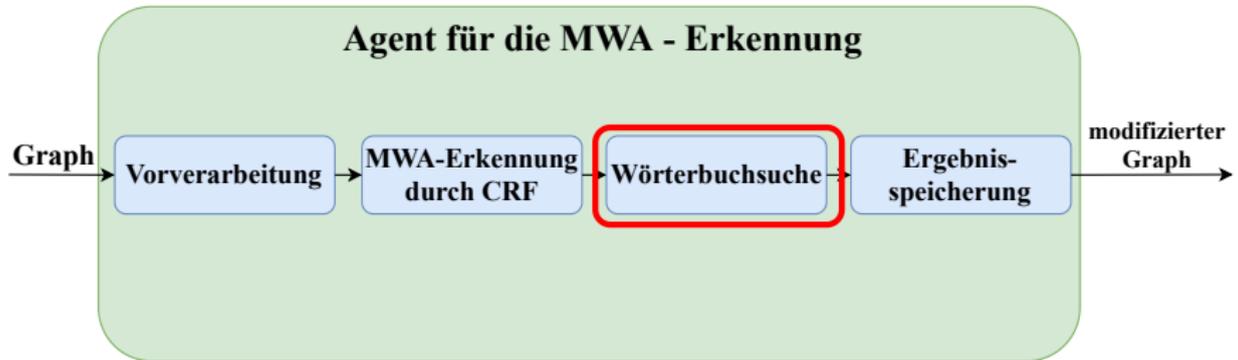
Vorverarbeitung

- Auslesen der Anforderungen (Wörter, Lemmata, Wortartmarkierungen)
- Eingabedatei erzeugen



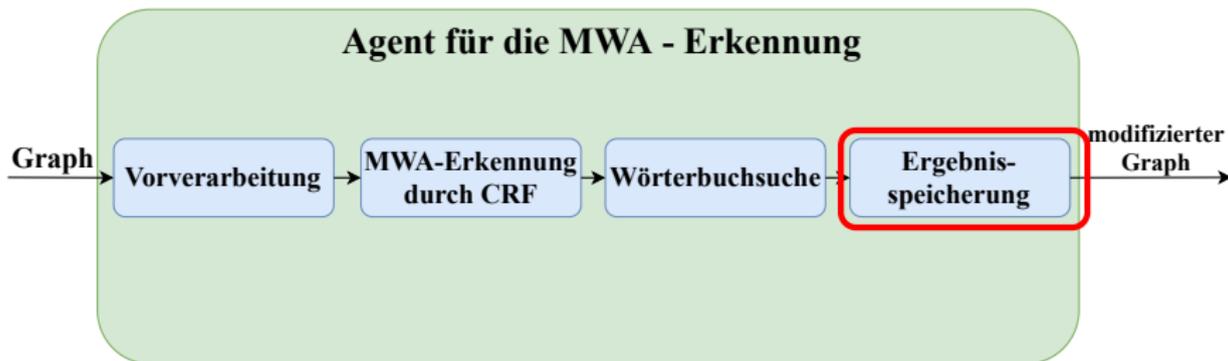
MWA-Erkennung durch CRF

- Markieren von MWAs durch trainiertes CRF

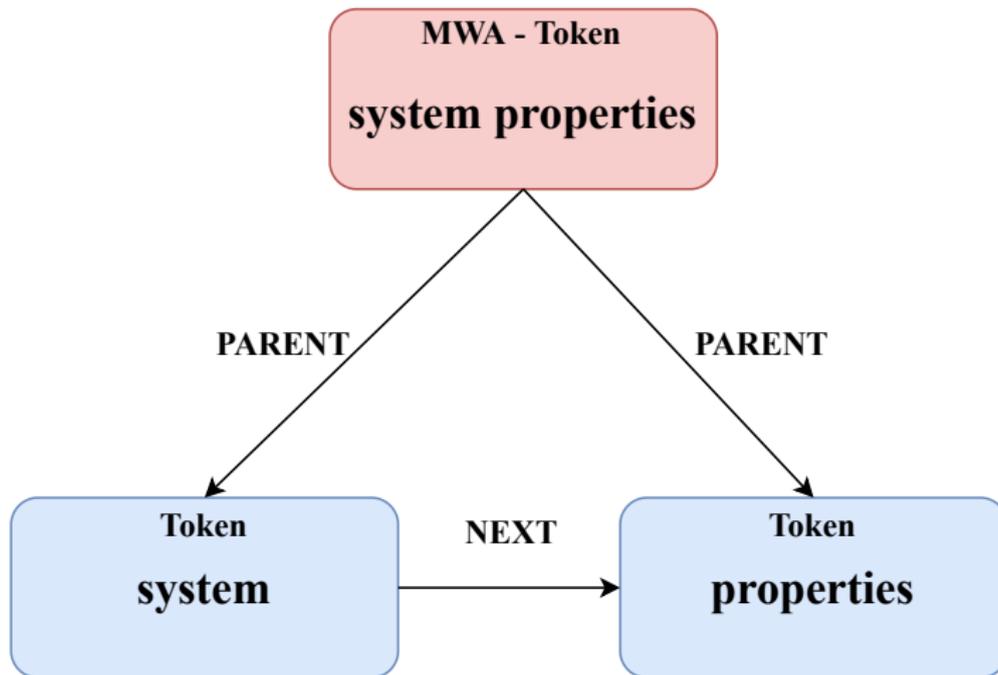


Wörterbuchsuche

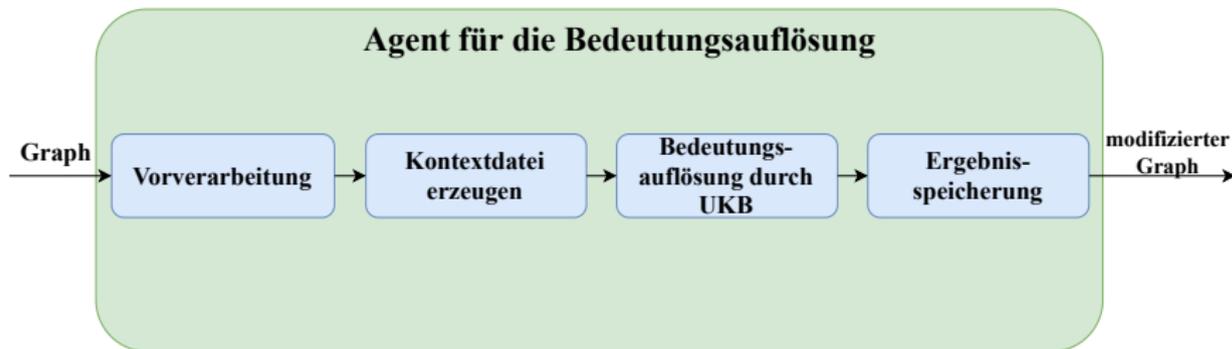
- Suche von Adjektiv + Nomen Kombinationen im Wörterbuch



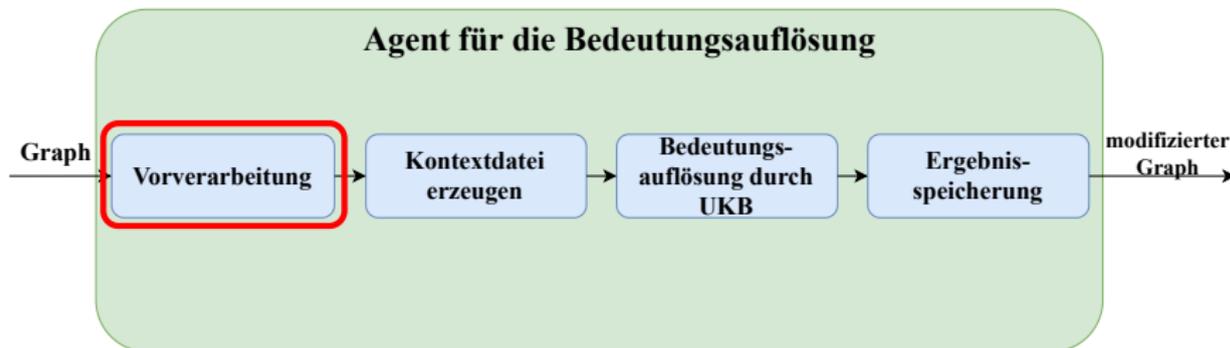
Graph nach der Ergebnisspeicherung



Entwurf: Bedeutungsauflösung mit Berücksichtigung von MWAs



Entwurf: Bedeutungsauflösung mit Berücksichtigung von MWAs



Vorverarbeitung

- Annotation der Token mit Suchwörtern

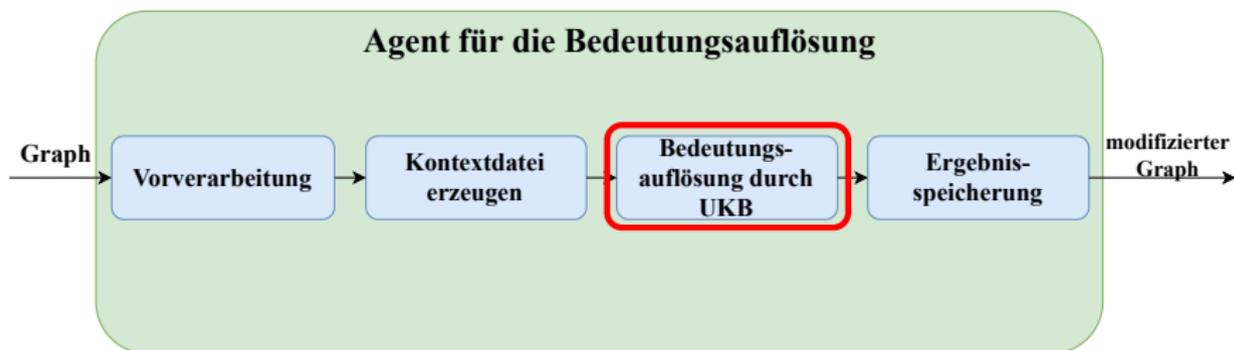
Entwurf: Bedeutungsauflösung mit Berücksichtigung von MWAs



Kontextdatei erzeugen

- Erzeugen einer Kontextdatei mit relevanten Wörtern einer Anforderung

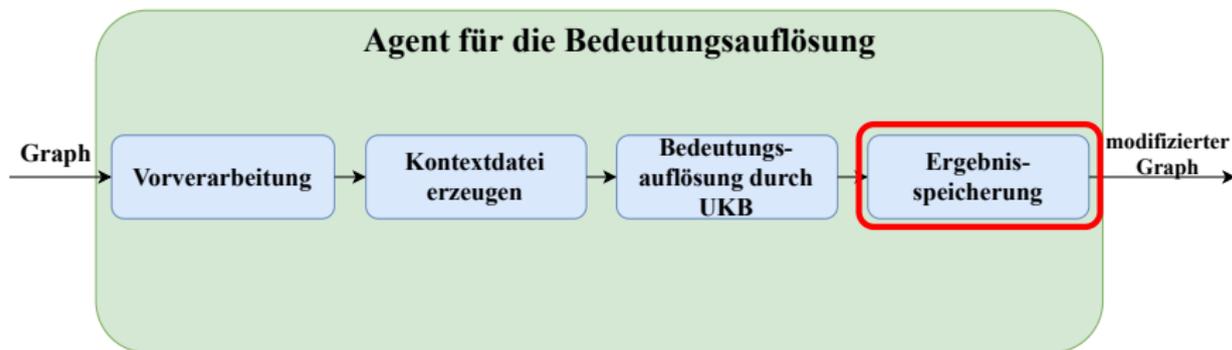
Entwurf: Bedeutungsauflösung mit Berücksichtigung von MWAs



Bedeutungsauflösung durch UKB

- Graph einer Wissensbasis verwenden, um Wörtern einer Kontextdatei eine Bedeutung zuzuordnen

Entwurf: Bedeutungsauflösung mit Berücksichtigung von MWAs



- Insgesamt 978 Anforderungen aus 4 Datensätzen [HHD09; HDS06; CHCC03; CH+07]
- MWAs annotiert
- Bedeutungen von Nomen und Eigennamen annotiert

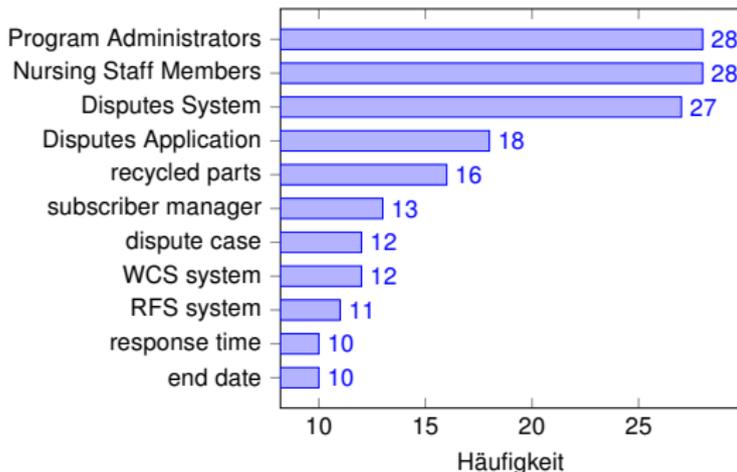


Abbildung: Häufigste MWAs im Evaluationsdatensatz

Tabelle: Ergebnisse der 10-fachen Kreuzvalidierung ohne Wörterbuchsuche

Testdatensatz	Präzision	Ausbeute	F ₁ -Wert
1	0.865	0.714	0.783
2	0.841	0.791	0.815
3	0.839	0.703	0.765
4	0.801	0.602	0.688
5	0.941	0.874	0.906
6	0.845	0.839	0.842
7	0.926	0.830	0.875
8	0.885	0.777	0.827
9	0.815	0.586	0.682
10	0.887	0.855	0.871
Gesamt	0.865	0.757	0.805

Tabelle: Ergebnisse der 10-fachen Kreuzvalidierung mit Suche im Wörterbuch

Testdatensatz	Präzision	Ausbeute	F ₁ -Wert
1	0.860	0.730	0.790
2	0.843	0.799	0.820
3	0.833	0.721	0.773
4	0.793	0.617	0.694
5	0.941	0.879	0.909
6	0.840	0.846	0.843
7	0.920	0.852	0.885
8	0.870	0.811	0.839
9	0.794	0.602	0.684
10	0.871	0.855	0.863
Gesamt	0.857	0.771	0.810

Tabelle: Evaluation der DBpedia Graphen (mit Teilbedeutungen/mit Fehlern der Wissensbasis)

Graph	PageRank	Häufigkeit	Präzision		Ausbeute		F ₁ -Wert	
Hyperonyme + Kategorien	ppr	ja	0.298	0.299	0.275	0.276	0.286	0.287
		nein	0.383	0.392	0.354	0.362	0.368	0.376
	ppr_w2w	ja	0.516	0.527	0.477	0.486	0.496	0.506
		nein	0.478	0.485	0.442	0.448	0.459	0.466
Hyperonyme	ppr	ja	0.452	0.454	0.407	0.409	0.428	0.430
		nein	0.483	0.488	0.435	0.439	0.458	0.462
	ppr_w2w	ja	0.411	0.421	0.371	0.379	0.390	0.399
		nein	0.359	0.383	0.323	0.345	0.340	0.363
Wikipedia MFS	-	-	0.360	-	0.329	-	0.344	-
Babelify (S)	-	-	0.517	-	0.292	-	0.373	-
Babelify	-	-	0.476	-	0.292	-	0.362	-

Tabelle: Evaluation des WordNet 3.1 Graphen (mit Teilbedeutungen/mit Fehlern der Wissensbasis)

Kontext	PageRank	Häufigkeit	Präzision		Ausbeute		F ₁ -Wert	
Adj., Adv., V, N, MWA	ppr	ja	0.559	0.558	0.510	0.508	0.533	0.532
		nein	0.373	0.378	0.340	0.345	0.356	0.361
	ppr_w2w	ja	0.509	0.510	0.464	0.465	0.485	0.486
		nein	0.411	0.402	0.375	0.366	0.392	0.383
N, MWA	ppr	ja	0.556	0.554	0.507	0.505	0.530	0.528
		nein	0.371	0.379	0.338	0.345	0.354	0.361
	ppr_w2w	ja	0.509	0.522	0.464	0.476	0.485	0.498
		nein	0.420	0.418	0.383	0.382	0.401	0.399
Adj., Adv., V, N, MWA (N)	ppr	ja	0.575	0.573	0.521	0.519	0.547	0.545
		nein	0.379	0.385	0.344	0.349	0.361	0.366
	ppr_w2w	ja	0.525	0.523	0.475	0.474	0.499	0.497
		nein	0.429	0.422	0.389	0.382	0.408	0.401
N, MWA (N)	ppr	ja	0.571	0.568	0.517	0.514	0.543	0.540
		nein	0.375	0.384	0.340	0.348	0.357	0.365
	ppr_w2w	ja	0.523	0.537	0.474	0.487	0.497	0.511
		nein	0.436	0.437	0.395	0.396	0.414	0.415
WordNet 3.1 MFS	-	-	0.504	-	0.460	-	0.481	-

Tabelle: Vergleich der kombinierten Ergebnisse (DBpedia/mit Teilebedeutungen/mit Fehlern der Wissensbasis)

Graph	PageRank	Häufigkeit	Präzision	Ausbeute	F ₁ -Wert
Hypernyme + Kategorien	ppr	ja	0.278	0.258	0.268
		nein	0.359	0.334	0.346
	ppr_w2w	ja	0.488	0.453	0.470
		nein	0.450	0.418	0.433
Hypernyme	ppr	ja	0.425	0.385	0.404
		nein	0.456	0.413	0.433
	ppr_w2w	ja	0.382	0.346	0.363
		nein	0.334	0.303	0.318
Wikipedia MFS	-	-	0.324	0.298	0.310

Tabelle: Vergleich der kombinierten Ergebnisse (WordNet/mit Teilbedeutungen/mit Fehlern der Wissensbasis)

Kontextwortarten	PageRank	Häufigkeit	Präzision	Ausbeute	F ₁ -Wert
Adj., Adv., V, N, MWA	ppr	ja	0.536	0.489	0.511
		nein	0.358	0.327	0.342
	ppr_w2w	ja	0.488	0.446	0.466
		nein	0.400	0.365	0.382
N, MWA	ppr	ja	0.533	0.486	0.508
		nein	0.355	0.324	0.339
	ppr_w2w	ja	0.488	0.445	0.466
		nein	0.404	0.368	0.385
Adj., Adv., V, N, MWA (N)	ppr	ja	0.549	0.500	0.523
		nein	0.364	0.331	0.347
	ppr_w2w	ja	0.501	0.457	0.478
		nein	0.414	0.377	0.395
N, MWA (N)	ppr	ja	0.545	0.496	0.519
		nein	0.359	0.327	0.342
	ppr_w2w	ja	0.499	0.454	0.475
		nein	0.417	0.379	0.397
WordNet 3.1 MFS	-	-	0.463	0.422	0.442

- Implementierung eines Systems zur Wortbedeutungsauflösung mit Berücksichtigung von MWAs
 - Erkennung von MWAs durch ein einfach verkettetes Zufallsfeld
 - Bedeutungsauflösung durch UKB mit den Wissensbasen DBpedia und WordNet 3.1
- Ergebnisse
 - F_1 -Wert von bis zu 0.81 bei der Erkennung von MWAs
 - F_1 -Wert von maximal 0.496 bei der Bedeutungsauflösung mit DBpedia
 - F_1 -Wert von maximal 0.547 bei der Bedeutungsauflösung mit WordNet 3.1
- Ausblick
 - Trainingsdatensatz des Zufallsfeldes erweitern (weniger Homogenität)
 - DBpedia Datensätze auf weitere mögliche Graphen untersuchen
 - Gewichte bei der Bedeutungsauflösung verändern

- Agirre, Eneko, Oier Lopez de Lacalle und Aitor Soroa (Juli 2018). “The risk of sub-optimal use of Open Source NLP Software: UKB is inadvertently state-of-the-art in knowledge-based WSD”. In: *Proceedings of Workshop for NLP Open Source Software (NLP-OSS)*. Melbourne, Australia: Association for Computational Linguistics, S. 29–33. DOI: 10.18653/v1/W18-2505. URL: <https://www.aclweb.org/anthology/W18-2505> (besucht am 12. 12. 2019).
- Agirre, Eneko und Aitor Soroa (März 2009). “Personalizing PageRank for Word Sense Disambiguation”. In: *Proceedings of the 12th Conference of the European Chapter of the ACL (EACL 2009)*. Athens, Greece: Association for Computational Linguistics, S. 33–41. URL: <https://www.aclweb.org/anthology/E09-1005> (besucht am 16. 12. 2019).

- Cleland-Huang, J., C.K. Chang und M. Christensen (Sep. 2003).
“Event-based traceability for managing evolutionary change”. In: *IEEE Transactions on Software Engineering* 29.9, S. 796–810. ISSN: 2326-3881. DOI: 10.1109/TSE.2003.1232285.
- Cleland-Huang, Jane u. a. (März 2007). *nfr*. type: dataset. DOI: 10.5281/zenodo.268542. URL: <https://zenodo.org/record/268542> (besucht am 08. 04. 2020).

- Cordeiro, Silvio, Carlos Ramisch und Aline Villavicencio (Juni 2016). “UFRGS&LIF at SemEval-2016 Task 10: Rule-Based MWE Identification and Predominant-Supersense Tagging”. In: *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*. San Diego, California: Association for Computational Linguistics, S. 910–917. DOI: 10.18653/v1/S16-1140. URL: <https://www.aclweb.org/anthology/S16-1140> (besucht am 10. 12. 2019).
- Hayes, Jane Huffman, Alex Dekhtyar und Senthil Karthikeyan Sundaram (2006). “Advancing Candidate Link Generation for Requirements Tracing: The Study of Methods”. In: *IEEE Trans. Software Eng.* 32.1, S. 4–19. DOI: 10.1109/TSE.2006.3. URL: <http://dx.doi.org/10.1109/TSE.2006.3>.

- Holbrook, E. Ashlee, Jane Huffman Hayes und Alex Dekhtyar (Aug. 2009). “Toward Automating Requirements Satisfaction Assessment”. In: *2009 17th IEEE International Requirements Engineering Conference*. ISSN: 2332-6441, S. 149–158. DOI: 10.1109/RE.2009.10.
- Hosseini, Mohammad Javad, Noah A. Smith und Su-In Lee (2016). “UW-CSE at SemEval-2016 Task 10: Detecting Multiword Expressions and Supersenses using Double-Chain Conditional Random Fields”. en. In: *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*. San Diego, California: Association for Computational Linguistics, S. 931–936. DOI: 10.18653/v1/S16-1143. URL: <http://aclweb.org/anthology/S16-1143> (besucht am 10. 12. 2019).

Moro, Andrea, Alessandro Raganato und Roberto Navigli (2014). “Entity Linking meets Word Sense Disambiguation: a Unified Approach”. In: *Transactions of the Association for Computational Linguistics 2*, S. 231–244. URL: <https://www.aclweb.org/anthology/Q14-1019>.