



University of Stuttgart
Germany



Automatische Bedeutungswandelerkennung in der Computerlinguistik

Ein Überblick

12. März 2025

Dominik Schlechtweg

Institut für maschinelle Sprachverarbeitung, Universität Stuttgart

Überblick

Einleitung

Wortverwendungsgraphen

Erkennung unbekannter Bedeutungbeschreibungen

Anwendungen

Abschluss

Referenzen

- ▶ Bedeutungswandelerkennung
 - ▶ **Ziel:** Automatisierung der Analyse von Wortbedeutungsänderungen über die Zeit
 - (1) ***Mäuse** und Ratten sind selbstverständlich mit den europäischen Schiffen auch hierher gekommen.*
 - (2) *Deshalb eignet sich die **Maus** auch für verschiedene Betriebssysteme neben Windows und macOS.*

Zwei Messparadigmen

1. Wortverwendungsgraphen

(Schlechtweg, 2023)

- ▶ vergleicht **Korpus** mit **Korpus**
- ▶ Erweiterung der **Word Sense Induction**

(Schütze, 1998)

2. Erkennung unbekannter Bedeutungsbeschreibungen

(Erk, 2006; Fedorova et al., 2024)

- ▶ vergleicht **Korpus** mit **Wörterbuch**
- ▶ Erweiterung der **Word Sense Disambiguation**

(Navigli, 2009)

Wortverwendungsgraphen

Menschliche Bedeutungswandelmessung


A	1824	and taking a knife from her pocket, she opened a vein in her little arm ,	😊
B	1842	And those who remained at home had been heavily taxed to pay for the arms , ammunition;	✖
C	1860	and though he saw her within reach of his arm , yet the light of her eyes seemed as far off	😊
		...	
D	1953	overlooking an arm of the sea which, at low tide, was a black and stinking mud-flat	🗿
E	1975	twelve miles of coastline lies in the southwest on the Gulf of Aqaba, an arm of the Red Sea.	🗿
F	1985	when the disembodied arm of the Statue of Liberty jets spectacularly out of the	😊

Tabelle 1: Diachrones Beispielkorpus.

Wortverwendungspaare

- (A) [...] and taking a knife from her pocket, she opened a vein in her little **arm**, and dipping a feather in the blood, wrote something on a piece of white cloth, which was spread before her. 😊
- (D) It stood behind a high brick wall, its back windows overlooking an **arm** of the sea which, at low tide, was a black and stinking mud-flat [...]

Skala der semantischen Nähe



4: Identisch
3: Enger Bezug
2: Entfernter Bezug
1: Kein Bezug

Tabelle 2: DUREl-Skala des semantischen Bezugs.

Graphrepräsentation

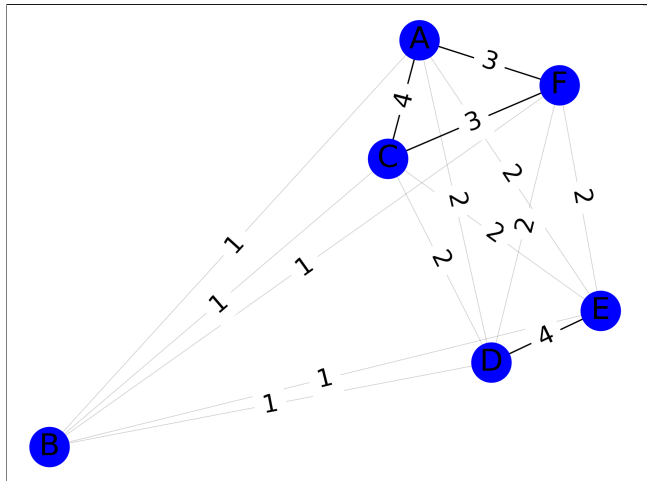


Abbildung 1: Wortverwendungsgraph des englischen Wortes *arm*.

Clustering

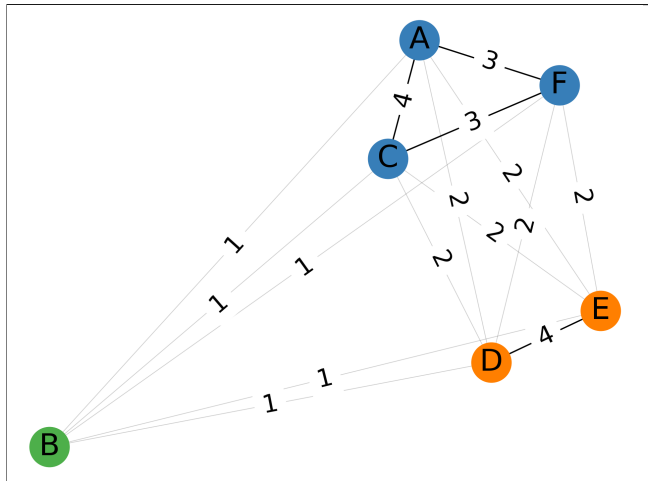
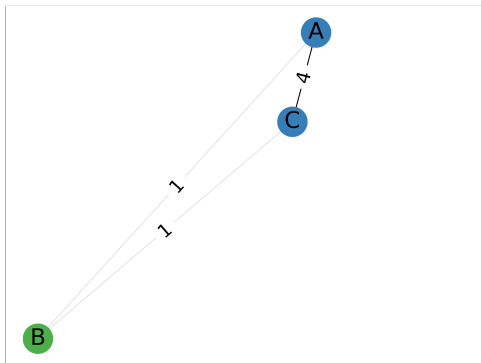
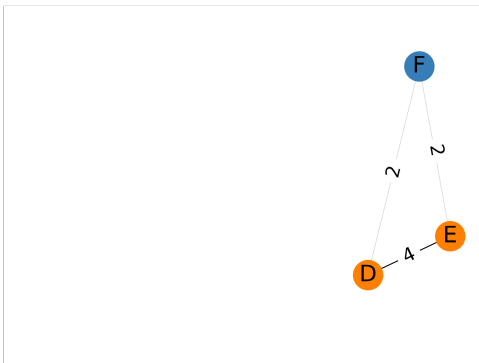


Abbildung 2: Wortverwendungsgraph des englischen Wortes *arm*. $D = (3, 2, 1)$.

Bedeutungswandel



$t_1, D_1 = (2, 0, 1)$



$t_2, D_2 = (1, 2, 0)$

Wandelwerte

- ▶ **binärer Wandel** (Verlust und Hinzugewinn von Bedeutungen)
- ▶ **gradierter Wandel** (Änderungen in Bedeutungsvorkommenswahrscheinlichkeiten)

Evaluationstasks

- Task 1 Binäre Klassifikation:** Für eine Menge von Zielworten, sage den binären Wandelwert vorher.
- Task 2 Ranking:** Reihe eine Menge von Zielworten nach ihrem gradierten Wandelwert.

(Schlechtweg et al., 2020)

Beispiel: Schwedisch *ledning*¹

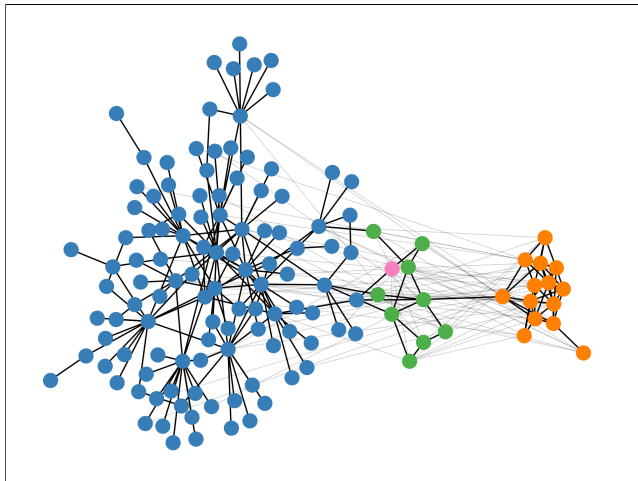


Abbildung 4: WUG des schwedischen Wortes *ledning*.

¹Datensätze verfügbar unter <https://www.ims.uni-stuttgart.de/data/wugs>

Beispiel: Schwedisch *ledning*

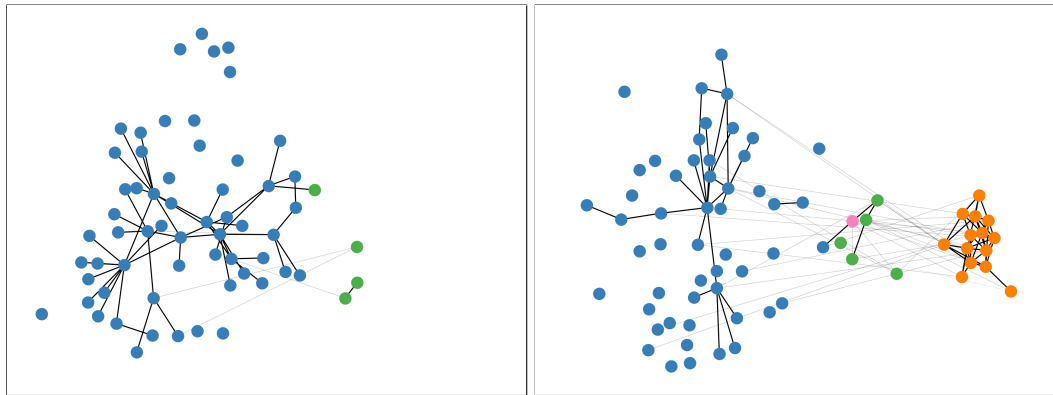


Abbildung 5: WUGs des schwedischen Wortes *ledning*: Subgraphen für die erste Zeitperiode G_1 (links) und die zweite Zeitperiode G_2 (rechts). $D_1 = (58, 0, 4, 0)$, $D_2 = (52, 14, 5, 1)$, $B(w) = 1$ und $G(w) = 0.34$.

Beispiel: Deutsch *Eintagsfliege*

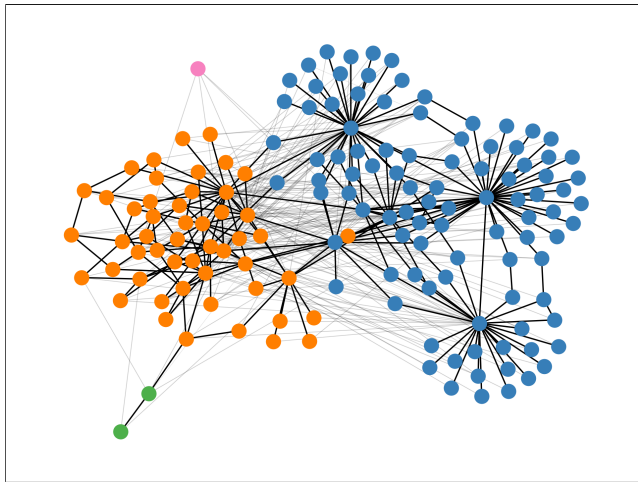


Abbildung 6: WUG des deutschen Wortes *Eintagsfliege*.

Beispiel: Deutsch *Eintagsfliege*

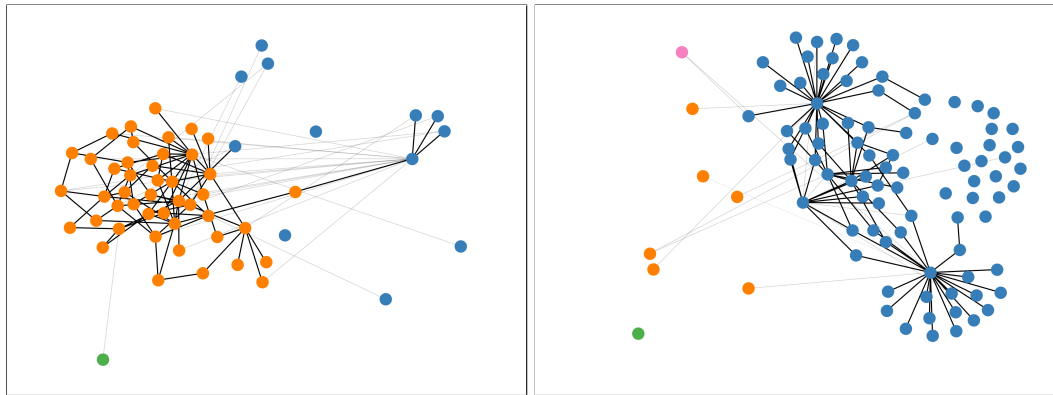


Abbildung 7: WUG of German *Eintagsfliege*: Subgraphen für die erste Zeitperiode G_1 (links) und die zweite Zeitperiode G_2 (rechts). $D_1 = (12, 45, 0, 1)$, $D_2 = (85, 6, 1, 1)$, $B(w) = 0$ Und $G(w) = 0.66$.

Zusammenfassung der Annotationsschritte

1. Annotation der semantischen Nähe
2. Clustering
3. Wandelmessung

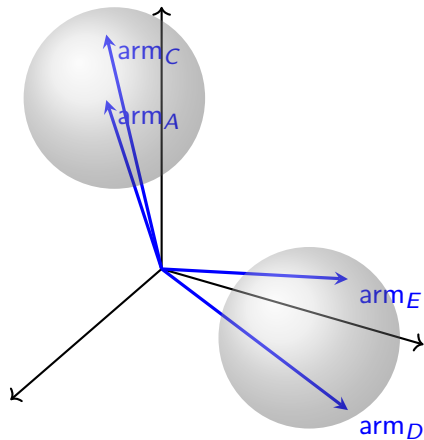
Zusammenfassung der Annotationsschritte mit Tasks

1. Annotation der semantischen Nähe ↔ **Word-in-Context Task**
2. Clustering ↔ **Word Sense Induction**
3. Wandelmessung ↔ **Bedeutungswandelerkennung** (inklusive vorheriger Tasks)

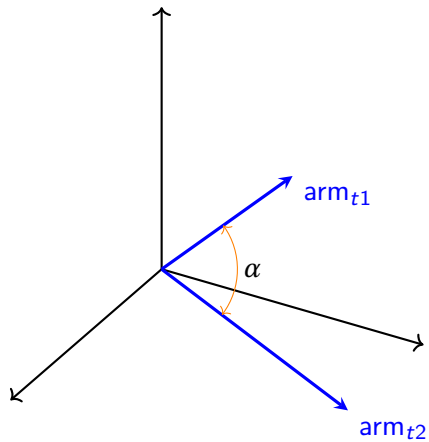
Automatische Bedeutungswandelmessung

- ▶ Ein typisches **verwendungsbasiertes** (token) Modell besteht aus
 1. Modell der semantischen Nähe (z. B. Ähnlichkeit zwischen kontextsensitiven Embeddings)
 2. Clustering-Algorithmus (optional)
 3. Wandelmaß
 - ▶ modellieren den menschlichen Messungsprozess
 - ▶ ein Vektor pro Verwendung (BERT, ELMo)
 - ▶ kleine Sprachmodelle (LMs)
- ▶ Ein typisches **wortbasiertes** (type) Modell besteht aus
 1. einer semantischen Repräsentation pro Wort (Wortvektor)
 2. Algorithmus zur zeitlichen Abbildung
 3. Wandelmaß
 - ▶ modellieren **nicht** den menschlichen Messungsprozess
 - ▶ ein gemittelter Vektor pro Wort (Word2Vec, GloVe)
 - ▶ noch kleinere Sprachmodelle

Einfaches verwendungsbasiertes Modell



Einfaches wortbasiertes Modell



SOTA Modellkomponenten

► SOTA Modellkomponenten pro Modellierungsebene:

1. **semantische Nähe:** DeepMistake, XL-Lexeme, GlossReader

(Arefyev et al., 2021; Arefyev & Rachinskiy, 2021; Cassotti et al., 2023)

2. **Clustering:** Agglomerative, Spectral, Correlation, Stochastic Blockmodel

(cf. Schlechtweg et al., 2024)

3. **Wandelmaß:** Clusterverlust/-hinzugewinn, Schwellenwert, Jensen-Shannon-Distanz, Mittlere paarweise Distanz

(Kutuzov & Giulianelli, 2020; Lin, 1991)

Modelle der semantischen Nähe

- ▶ aka **Word-in-Context**-Modelle (Pilehvar & Camacho-Collados, 2019)
- ▶ schätzen den Grad der semantischen Nähe/Wahrscheinlichkeit der Bedeutungsgleichheit
- ▶ SOTA: DeepMistake, XL-Lexeme, GlossReader (Arefyev et al., 2021; Arefyev & Rachinskiy, 2021; Cassotti et al., 2023)
- ▶ optimiert auf binären, mehrsprachigen Datensätzen (Martelli et al., 2021; Pilehvar & Camacho-Collados, 2019; Raganato et al., 2020)
- ▶ führten zu großen Erkennungsgenauigkeitssprüngen (Kutuzov & Pivovarova, 2021)
- ▶ können auf ordinale Bewertungen abgebildet werden (Schlechtweg et al., 2025)
- ▶ Trainingsdatenbeispiel:
 - (A) [...] and taking a knife from her pocket, she opened a vein in her little **arm**, and dipping a feather in the blood, wrote something on a piece of white cloth, which was spread before her. 😊
 - (D) It stood behind a high brick wall, its back windows overlooking an **arm** of the sea which, at low tide, was a black and stinking mud-flat [...]. 🚫
Label: 0 (binär), 2 (ordinal)

Modelle der semantischen Nähe

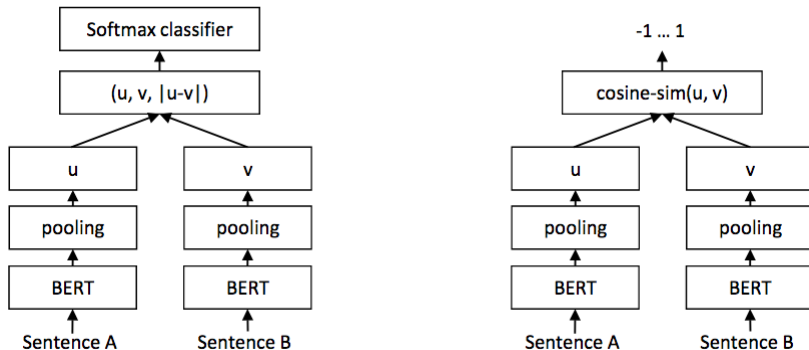
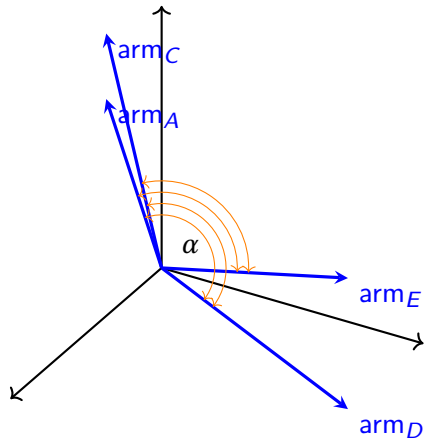


Abbildung 8: S-BERT (Reimers & Gurevych, 2019) Trainingsarchitektur, benutzt für XL-Lexeme.

SOTA-Modell für gradierten Wandel: APD



Sprache	binär	Modell	gradiert	Modell
Chinesisch			.73	XL-Lex.+APD
Englisch	.70 (.67/.75)	BERT+HDBSCAN	.89	XL-Lex.+APD
Deutsch	.70 (.60/.82)	SGNS+thres.	.84	XL-Lex.+APD
Norwegisch			.76	XL-Lex.+PRT
Russisch			.86	XL-Lex.+APD
Schwedisch	.64 (.47/1.0)	XLM-R+K-means	.81	XL-Lex.+APD
Spanisch	.72 (.62/.86)	GlossR.+APD+thres.	.74	GlossR.+APD

Tabelle 3: SOTA-Genauigkeit der Bedeutungswandelerkennung (Cassotti et al., 2023; Periti & Tahmasebi, 2024; Rachinskiy & Arefyev, 2022; Schlechtweg et al., 2020; Zamora-Reina et al., 2022). Werte geben F1 (P/R) for binären Wandel und Spearman für gradierten Wandel.

Zusammenfassung

- ▶ **Vorteile:**

- ▶ keine Abhängigkeit von existierenden Wortbedeutungsdefinitionen
- ▶ eher explizite Annotationskriterien

- ▶ **Nachteile:**

- ▶ fragliches Konzept der semantischen Nähe
- ▶ quadratische Erhöhung des Annotationsbedarfs
- ▶ abhängig von Clusteringalgorithmus

- ▶ **Offene Fragen:**

- ▶ Clustering & binärer Wandel
- ▶ Anwendung
- ▶ mehrere Zeitperioden
- ▶ Erkennung nicht-wohlgeformter Verwendungen
- ▶ Fehleranalyse

(Choppa et al., 2025)

Erkennung unbekannter Bedeutungsbeschreibungen

Menschliche Messung unbekannter Bedeutungsbeschreibungen

- A and taking a knife from her pocket, she opened a vein in her little **arm**,
- B And those who remained at home had been heavily taxed to pay for the **arms**, ammunition;
- C and though he saw her within reach of his **arm**, yet the light of her eyes seemed as far off
- D overlooking an **arm** of the sea which, at low tide, was a black and stinking mud-flat
- E twelve miles of coastline lies in the southwest on the Gulf of Aqaba, an **arm** of the Red Sea.
- F when the disembodied **arm** of the Statue of Liberty jets spectacularly out of the

Lemma	Definition
arm	a human limb weapon system

Abbildung 9: Beispielkorpus (links) und Wörterbucheintrag (rechts).

Verwendungs-Bedeutungs-Vergleich

(D) It stood behind a high brick wall, its back windows overlooking an **arm** of the sea which, at low tide, was a black and stinking mud-flat [...]

Bedeutung 1 a human limb

Bedeutung 2 weapon system

Bedeutung 3 [NEUE BEDEUTUNG]

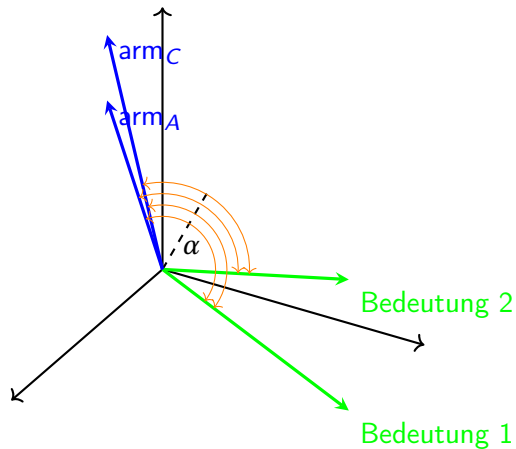
Evaluationstasks

- ▶ kürzlicher **shared task** führt zwei Evaluationstasks ein (Fedorova et al., 2024)
- ▶ **Daten:** Finnisch, Russisch, Deutsch
- ▶ **Task:** Gegeben eine Menge historischer Verwendungs-Bedeutungs-Paare als Referenz. . .
 1. **Erkennung unbekannter Wortbedeutungen**
 - ▶ disambiguiere Verwendungen, die zu bekannten Wortbedeutungen gehören (evaluiert mit F1)
 - ▶ clustere eine Menge von Verwendungen in Wortbedeutungen (evaluiert mit ARI)
 2. **Wortbedeutungsgenerierung**
 - ▶ generiere eine Wortbedeutungsdefinition für unbekannte Bedeutungen (evaluiert durch BERTScore/BLEU)

Automatische Messung unbekannter Wortbedeutungen

- ▶ Ein typisches Modell besteht aus
 1. Modell der semantischen Nähe (diesmal zwischen Verwendungen und Bedeutungen)
 2. binärer Unbekanntheitsklassifikator
 3. Clustering-Algorithmus (optional)

Einfaches Modell: Schwellenwert auf Kosinus-Ähnlichkeit



SOTA Modellkomponenten

- ▶ SOTA-Modellkomponenten pro Modellierungsebene: (Kokosinskii, Kuklin & Arefyev, 2024)
 1. **semantische Ähnlichkeit**: GlossReader (Arefyev & Rachinskiy, 2021)
 2. **Klassifikator**: logistische Regression auf Vektordistanz-Features
 3. **Clustering**: Agglomerative

Resultate

Sprache	F1	Modell	ARI	Modell
Deutsch	.76	GlossR.	.54	GlossR.
Finnisch	.76	GlossR.	.65	GlossR.+LR+Agglom.
Russisch	.75	GlossR.	.26	GlossR.+Agglom.

Tabelle 4: SOTA-Genauigkeit auf Erkennung unbekannter Wortbedeutungen (Fedorova et al., 2024; Kokosinskii et al., 2024).

Automatische Wortbedeutungsgenerierung

- ▶ typischerweise ein **Large Language Model** (mit Decoder)

(Open AI, 2022; Touvron et al., 2023)

- ▶ z. B. GPT, Llama, DeepSeek
- ▶ **vortrainiert** auf großen Text Datenmengen mit Language-Modeling-Task
- ▶ optional werden sie **gefinetuned** mit spezifischen Task-Instruktionen
- ▶ oft ist **Prompt-Engineering** notwendig

Prompt-Beispiel

[Instruction]:

Imagine that you are a lexicographer, given a headword {target_word} in {lang}, write the dictionary definition of its usage in the following quotations:

1. First quotation
2. Second quotation

[Requirements]:

The definition you create should be brief. A maximum of ten words is allowed. The definition ends at the first period.

[Response]:

Definition (string): {definition}

Abbildung 10: Ma et al. (2024)'s Prompt für GPT-3.5-turbo.

Sprache	BLEU+BERTScore	Model
Deutsch	.33	unb.
Finnisch	.41	GPT-3.5
Russisch	.73	WSD

Tabelle 5: SOTA-Genauigkeit auf Bedeutungsdefinitionsgenerierungstask (Fedorova et al., 2024; Kokosinskii et al., 2024).

Beispiele

- ▶ **Verwendung:** *Якобы тут, возле Черниховки, находится радиоактивный могильник.*
'Angeblich gibt es hier in der Nähe von Tschernikhovka ein radioaktives **Endlager**.'
- ▶ **Definition:** *спец. место захоронения радиоактивных отходов; специальное сооружение для такого захоронения.*
'Spezial-Deponie für radioaktive Abfälle; Eine besondere Struktur für eine solche Lagerung.'
- ▶ **generierte Definition mit GPT-3.5-turbo:** *Могильник - место захоронения радиоактивных отходов или погибших.*
'Могильник - der Ort, an dem radioaktive Abfälle oder tote Personen entsorgt werden.'
- ▶ **Evaluation:** BLEU: 21.2, BERTScore: 0.79

Beispiel

Prompt: ... The representative of the United States Government speaks about seeds for rice with, increased yield which the United States will send to Vietnam," Mr. Thuy said, "but even at this moment American troops and American **planes** are attacking our harvest, are now destroying the economic infrastructure of South Vietnam." What is the definition of **plane**?

Definition mit mT0: An aircraft.

Definition mit GPT-3.5-turbo: An aircraft that is used for military purposes, typically for combat or warfare.

Definition mit GPT-o1-preview: An airplane; a powered flying vehicle with fixed wings and a weight greater than that of the air it displaces, used for transportation or military purposes.

Definition im OED: An aeroplane.

► **Hauptproblem:** Granularität der Definitionen

Prompt: It follows, that if the point in space be upon the horizontal **plane**, its distance with regard to this last named **plane** will be zero or nothing, and the vertical A a will be zero also. What is the definition of **plane**?

Definition mit mT0: A surface without outline; a flat or almost flat surface.

Definition im OED: A flat geometrical surface which has the property that every straight line joining any two points of the surface lies wholly in the surface.

Zusammenfassung

- ▶ **Vorteile:**

- ▶ Wörterbücher sind reichhaltige Ressource
- ▶ praktische Relevanz
- ▶ Wandelvorsagen werden interpretierbar

- ▶ **disadvantages:**

- ▶ benötigt Wortbedeutungsdefinitionen
- ▶ inkonsistente Kriterien für deren Erstellung

- ▶ **offene Fragen:**

- ▶ klare Evaluation
- ▶ interpretierbare Evaluation für Generierung

(Kutuzov, Fedorova, Schlechtweg & Arefyev, 2024)

- ▶ Fehleranalyse
- ▶ Anwendung
- ▶ Einbindung bedeutungsannotierter Verwendungen

(Lautenschlager et al., 2024)

Anwendungen

Anwendungen

► Wortverwendungsgraphen

- Erkennung starker Veränderungen in deutschen historischen Korpora mit effizienten wortbasierten Methoden (Kurtyigit et al., 2021)
- Erkennung unbekannter Bedeutungen in schwedischem Wörterbuch durch Vergleich der Anzahl automatisch erkannter Korpusbedeutungen mit Wörterbuchbedeutungen (Sander et al., 2024; Sköldbberg, Virk, Sander, Hengchen & Schlechtweg, 2024)

► Erkennung unbekannter Wortbedeutungen

- Erkennung unbekannter Wortbedeutungen in englischem und schwedischem Wörterbuch mit Modellen der semantischen Nähe (Lautenschlager et al., 2024)
- finegetunte LLMs für Bedeutungsgenerierung im Englischen, Deutschen, Russischen, Norwegischen (Kutuzov et al., 2024)
- Anwendung von SOTA-Modellen zur Priorisierung von Wörterbuchupdatevorschlägen im ODE

Abschluss

- ▶ zwei Messparadigmen in der automatischen Bedeutungswandelerkennung
 - ▶ **Wortverwendungsgraphen** (inspiriert durch WSI)
 - ▶ hohe Genauigkeit für **gradierten Wandel**
 - ▶ niedrige Genauigkeit für **binären Wandel**, keine solide Evaluation
 - ▶ **Erkennung unbekannter Bedeutungsbeschreibungen** (inspiriert durch WSD)
 - ▶ unklare Genauigkeit, besonders in Anwendungsszenarien
- ▶ viele **offene Fragen**:
 - ▶ Wie können wir Bedeutungcluster genauer modellieren?
 - ▶ Wie soll die Erkennung unbekannter Bedeutungsbeschreibungen evaluiert werden?
 - ▶ Wie können Modelle effektiv und effizient in Anwendungsszenarien eingesetzt werden?

Referenzen I

- Arefyev, N., Fedoseev, M., Protasov, V., Homskiy, D., Davletov, A. & Panchenko, A. (2021). DeepMistake: Which senses are hard to distinguish for a word-in-context model. In (Bd. 2021-June, S. 16–30). Zugriff auf <https://www.dialog-21.ru/media/5491/arefyevnplusetal133.pdf>
- Arefyev, N. & Rachinskiy, M. (2021). Zero-shot cross-lingual transfer of a gloss language model for semantic change detection. In (Bd. 2021-June, S. 578–586). doi: 10.28995/2075-7182-2021-20-578-586
- Cassotti, P., Siciliani, L., de Gemmis, M., Semeraro, G. & Basile, P. (2023, Juli). XI-lexeme: Wic pretrained model for cross-lingual lexical semantic change. In *Proceedings of the 61th annual meeting of the association for computational linguistics*. Online: Association for Computational Linguistics.
- Choppa, T., Roth, M. & Schlechtweg, D. (2025, jan). Predicting median, disagreement and noise label in ordinal word-in-context data. In M. Roth & D. Schlechtweg (Hrsg.), *Proceedings of context and meaning: Navigating disagreements in nlp annotation* (S. 65–77). Abu Dhabi, UAE: International Committee on Computational Linguistics. Zugriff auf <https://aclanthology.org/2025.comedi-1.6/>
- Erk, K. (2006, Juni). Unknown word sense detection as outlier detection. In R. C. Moore, J. Bilmes, J. Chu-Carroll & M. Sanderson (Hrsg.), *Proceedings of the human language technology conference of the NAACL, main conference* (S. 128–135). New York City, USA: Association for Computational Linguistics. Zugriff auf <https://aclanthology.org/N06-1017>
- Fedorova, M., Mickus, T., Partanen, N., Siewert, J., Spaziani, E. & Kutuzov, A. (2024, aug). AXOLOTL'24 shared task on multilingual explainable semantic change modeling. In N. Tahmasebi et al. (Hrsg.), *Proceedings of the 5th workshop on computational approaches to historical language change* (S. 72–91). Bangkok, Thailand: Association for Computational Linguistics. Zugriff auf <https://aclanthology.org/2024.lchange-1.8> doi: 10.18653/v1/2024.lchange-1.8
- Kokosinskii, D., Kuklin, M. & Arefyev, N. (2024, aug). Deep-change at AXOLOTL-24: Orchestrating WSD and WSI models for semantic change modeling. In N. Tahmasebi et al. (Hrsg.), *Proceedings of the 5th workshop on computational approaches to historical language change* (S. 168–179). Bangkok, Thailand: Association for Computational Linguistics. Zugriff auf <https://aclanthology.org/2024.lchange-1.16> doi: 10.18653/v1/2024.lchange-1.16
- Kurtyigit, S., Park, M., Schlechtweg, D., Kuhn, J. & Schulte im Walde, S. (2021, aug). Lexical Semantic Change Discovery. In *Proceedings of the 59th annual meeting of the association for computational linguistics and the 11th international joint conference on natural language processing (volume 1: Long papers)*. Online: Association for Computational Linguistics. Zugriff auf <https://aclanthology.org/2021.acl-long.543/>

Referenzen II

- Kutuzov, A., Fedorova, M., Schlechtweg, D. & Arefyev, N. (2024, may). Enriching word usage graphs with cluster definitions. In N. Calzolari, M.-Y. Kan, V. Hoste, A. Lenci, S. Sakti & N. Xue (Hrsg.), *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)* (S. 6189–6198). Torino, Italia: ELRA and ICCL. Zugriff auf <https://aclanthology.org/2024.lrec-main.546>
- Kutuzov, A. & Giulianelli, M. (2020). UiO-UvA at SemEval-2020 Task 1: Contextualised Embeddings for Lexical Semantic Change Detection. In *Proceedings of the 14th international workshop on semantic evaluation*. Barcelona, Spain: Association for Computational Linguistics.
- Kutuzov, A. & Pivovarov, L. (2021). Rushifteval: a shared task on semantic shift detection for russian. *Komp'yuternaya Lingvistika i Intellektual'nye Tekhnologii: Dialog conference*.
- Lautenschlager, J., Hengchen, S., Sköldberg, E. & Schlechtweg, D. (2024). *Detection of non-recorded word senses in English and Swedish*. Zugriff auf <https://arxiv.org/abs/2403.02285>
- Lin, J. (1991, 01). Divergence measures based on the shannon entropy. *IEEE Transactions on Information Theory*, 37, 145–151.
- Ma, X., Schlechtweg, D. & Zhao, W. (2024, aug). Presence or absence: Are unknown word usages in dictionaries? In N. Tahmasebi et al. (Hrsg.), *Proceedings of the 5th workshop on computational approaches to historical language change* (S. 42–54). Bangkok, Thailand: Association for Computational Linguistics. Zugriff auf <https://aclanthology.org/2024.lchange-1.5>
- Martelli, F., Kalach, N., Tola, G. & Navigli, R. (2021, aug). SemEval-2021 task 2: Multilingual and cross-lingual word-in-context disambiguation (MCL-WiC). In A. Palmer, N. Schneider, N. Schluter, G. Emerson, A. Herbelot & X. Zhu (Hrsg.), *Proceedings of the 15th international workshop on semantic evaluation (semeval-2021)* (S. 24–36). Online: Association for Computational Linguistics. Zugriff auf <https://aclanthology.org/2021.semeval-1.3> doi: 10.18653/v1/2021.semeval-1.3
- Navigli, R. (2009). Word sense disambiguation: a survey. *ACM Computing Surveys*, 41 (2), 1–69.
- Open AI. (2022). *Introducing chatgpt*. Zugriff auf <https://openai.com/index/chatgpt/> (Accessed: 2024-10-01)
- Periti, F. & Tahmasebi, N. (2024). A systematic comparison of contextualized word embeddings for lexical semantic change. In *Proceedings of the 2024 conference of the north american chapter of the association for computational linguistics: Human language technologies (volume 1: Long papers)* (S. 4262–4282). Association for Computational Linguistics. Zugriff auf <http://dx.doi.org/10.18653/v1/2024.naacl-long.240> doi: 10.18653/v1/2024.naacl-long.240
- Pilehvar, M. T. & Camacho-Collados, J. (2019, Juni). WiC: the word-in-context dataset for evaluating context-sensitive meaning representations. In *Proceedings of the 2019 conference of the north American chapter of the association for computational linguistics: Human language technologies, volume 1 (long and short papers)* (S. 1267–1273). Minneapolis, Minnesota: Association for Computational Linguistics. doi: 10.18653/v1/N19-1128

Referenzen III

- Rachinskiy, M. & Arefyev, N. (2022). GlossReader at LSCDiscovery: Train to select a proper gloss in english – discover lexical semantic change in spanish. In *Proceedings of the 3rd international workshop on computational approaches to historical language change*. Dublin, Ireland: Association for Computational Linguistics.
- Raganato, A., Pasini, T., Camacho-Collados, J. & Pilehvar, M. T. (2020, nov). XL-WiC: A multilingual benchmark for evaluating semantic contextualization. In B. Webber, T. Cohn, Y. He & Y. Liu (Hrsg.), *Proceedings of the 2020 conference on empirical methods in natural language processing (emnlp)* (S. 7193–7206). Online: Association for Computational Linguistics. Zugriff auf <https://aclanthology.org/2020.emnlp-main.584> doi: 10.18653/v1/2020.emnlp-main.584
- Reimers, N. & Gurevych, I. (2019, nov). Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In K. Inui, J. Jiang, V. Ng & X. Wan (Hrsg.), *Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (emnlp-ijcnlp)* (S. 3982–3992). Hong Kong, China: Association for Computational Linguistics. Zugriff auf <https://aclanthology.org/D19-1410> doi: 10.18653/v1/D19-1410
- Sander, P., Hengchen, S., Zhao, W., Ma, X., Sköldbberg, E., Virk, S. M. & Schlechtweg, D. (2024). The DUREl Annotation Tool: Using fine-tuned LLMs to discover non-recorded senses in multiple languages. In *Proceedings of the Workshop on Large Language Models and Lexicography at 21st EURALEX International Congress Lexicography and Semantics*. Zugriff auf https://www.cjvt.si/wp-content/uploads/2024/10/LLM-Lex_2024_Book-of-Abstracts.pdf
- Schlechtweg, D. (2023). *Human and computational measurement of lexical semantic change* (Dissertation, University of Stuttgart, Stuttgart, Germany). Zugriff auf <http://dx.doi.org/10.18419/opus-12833>
- Schlechtweg, D., Choppa, T., Zhao, W. & Roth, M. (2025, jan). CoMeDi shared task: Median judgment classification & mean disagreement ranking with ordinal word-in-context judgments. In M. Roth & D. Schlechtweg (Hrsg.), *Proceedings of context and meaning: Navigating disagreements in nlp annotation* (S. 33–47). Abu Dhabi, UAE: International Committee on Computational Linguistics. Zugriff auf <https://aclanthology.org/2025.comedi-1.4/>
- Schlechtweg, D., McGillivray, B., Hengchen, S., Dubossarsky, H. & Tahmasebi, N. (2020). SemEval-2020 Task 1: Unsupervised Lexical Semantic Change Detection. In *Proceedings of the 14th International Workshop on Semantic Evaluation*. Barcelona, Spain: Association for Computational Linguistics. Zugriff auf <https://www.aclweb.org/anthology/2020.semeval-1.1/>
- Schlechtweg, D., Zamora-Reina, F. D., Bravo-Marquez, F. & Arefyev, N. (2024). Sense through time: Diachronic word sense annotations for word sense induction and lexical semantic change detection. *Language Resources and Evaluation*. Zugriff auf <https://doi.org/10.1007/s10579-024-09771-7>
- Schütze, H. (1998, März). Automatic word sense discrimination. *Computational Linguistics*, 24 (1), 97–123.

Referenzen IV

- Sköldberg, E., Virk, S. M., Sander, P., Hengchen, S. & Schlechtweg, D. (2024). Revealing semantic variation in Swedish using computational models of semantic proximity: Results from lexicographical experiments. In *Proceedings of the 21st EURALEX International Congress Lexicography and Semantics*. Zugriff auf <https://euralex.org/publications/revealing-semantic-variation-in-swedish-using-computational-models-of-semantic-proximity-results-from-lexicographical-experiments/>
- Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., ... Lample, G. (2023). *Llama: Open and efficient foundation language models*.
- Zamora-Reina, F. D., Bravo-Marquez, F. & Schlechtweg, D. (2022). LSCDiscovery: A shared task on semantic change discovery and detection in Spanish. In *Proceedings of the 3rd international workshop on computational approaches to historical language change*. Dublin, Ireland: Association for Computational Linguistics. Zugriff auf <https://aclanthology.org/2022.lchange-1.16/>