

Bestimmung der semantischen Funktion von Sätzen in Anforderungsbeschreibungen

Dana Tomova

Betreut von Tobias Hey

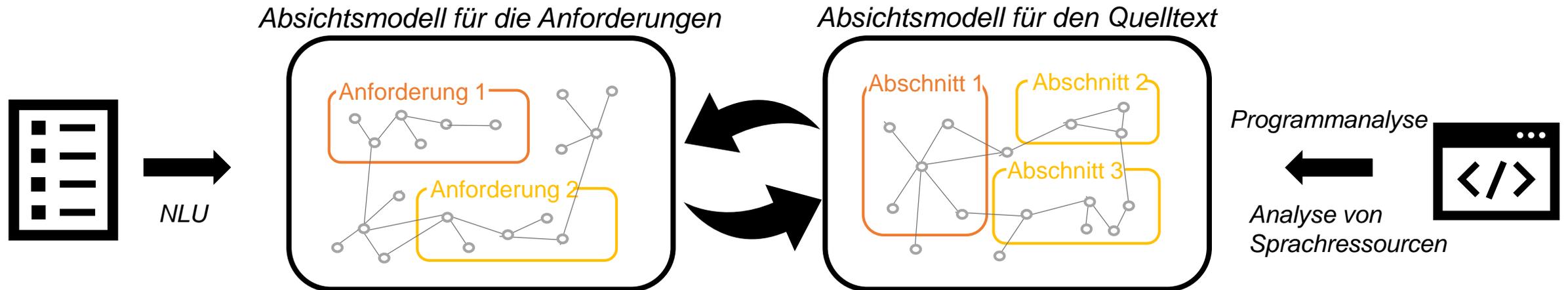
IPD Tichy, Fakultät für Informatik



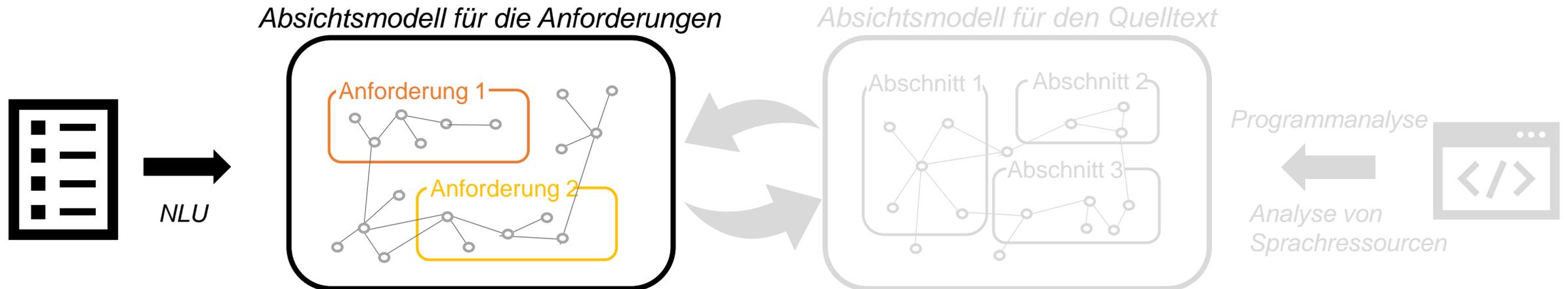
Motivation

- Rückverfolgbarkeit von Anforderungen
- Verständnis der Semantik der Anforderungen
- Zustandsbeschreibung ⇒ Passwords must be at least six characters long.
- Aktionsbeschreibung ⇒ Customers shall be able to write movie reviews.

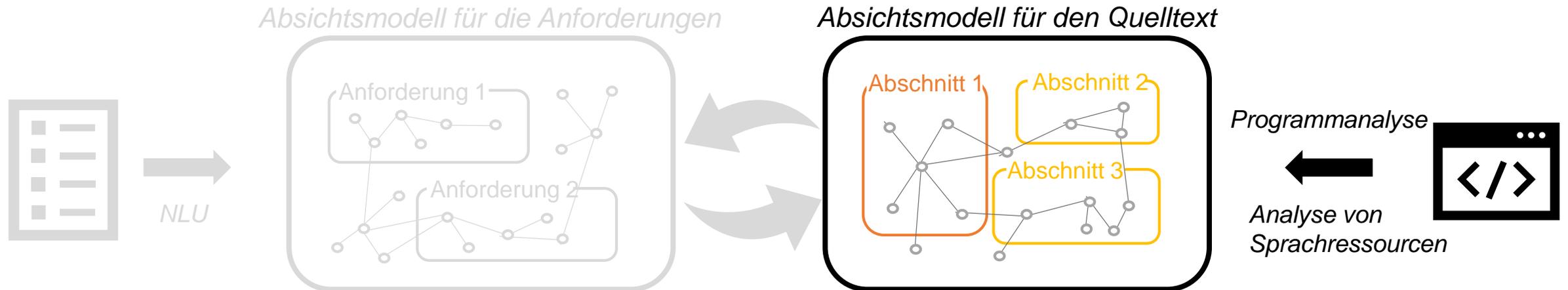
Gesamtprojekt INDIRECT



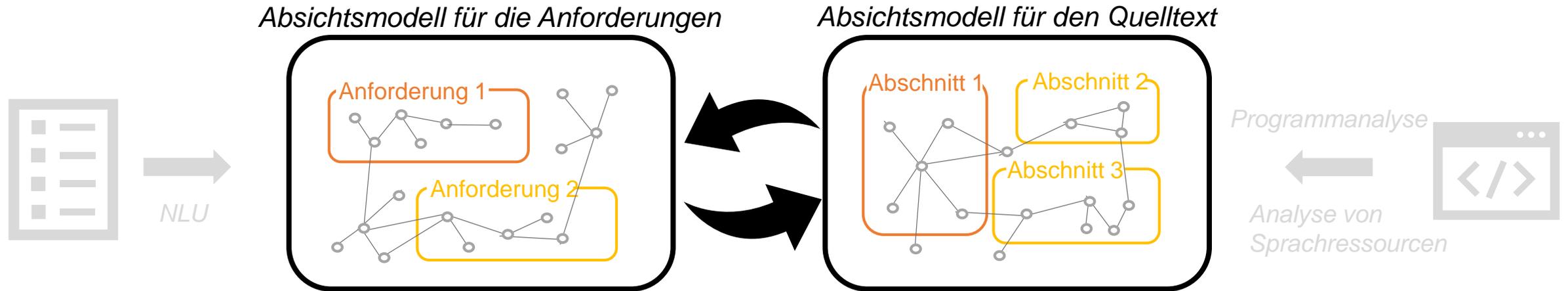
Gesamtprojekt INDIRECT



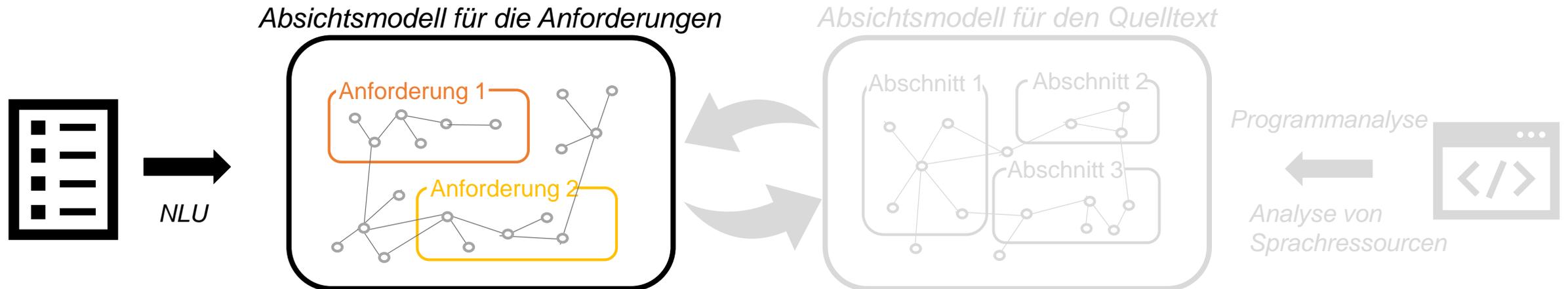
Gesamtprojekt INDIRECT



Gesamtprojekt INDIRECT



Gesamtprojekt INDIRECT



- Ziel dieser Arbeit: Bestimmung von Satzkategorien
 - Aktionsbeschreibung, Ereignis, Zustandsbeschreibung, Aggregation

Verwandte Arbeiten

Klassifikation von Anforderungen

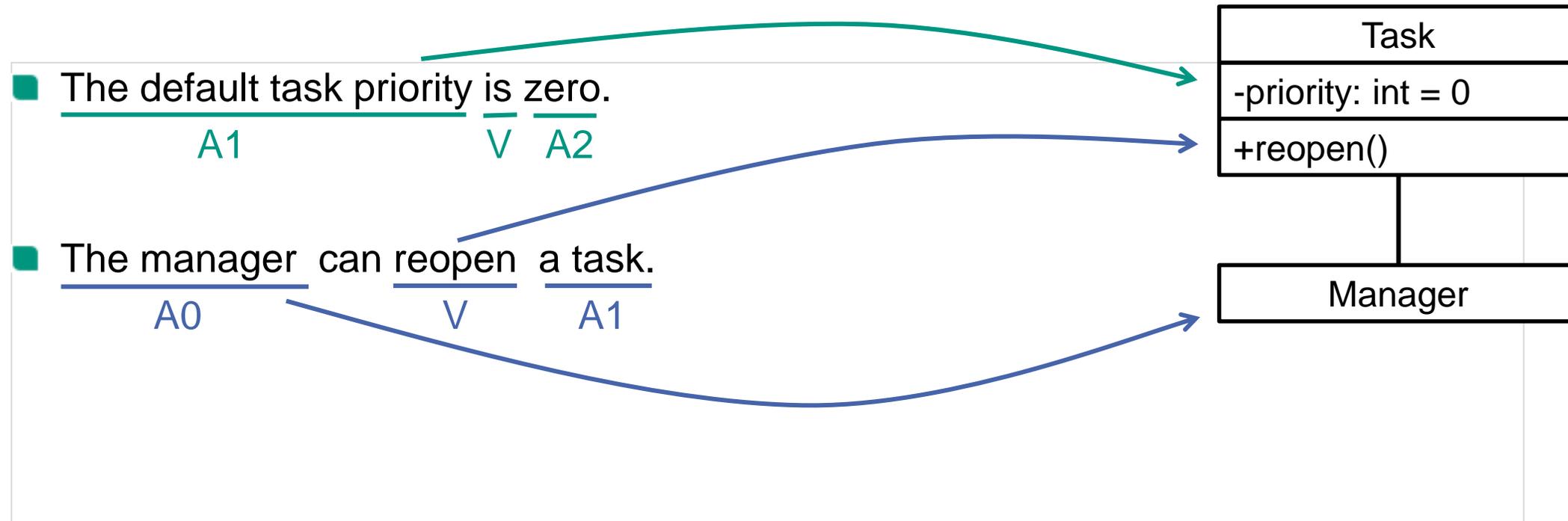
- Funktionale / nichtfunktionale Anforderungen [KM17] [VR11]
- Identifizierung von Domain-Aktionen [RM18]

Satzklassifikation

- Fragentypklassifizierung, Stimmungsvorhersage, Filmkritiken
- Word embeddings, CNN [KM16] [Kim14]

Festlegung der Kategorien

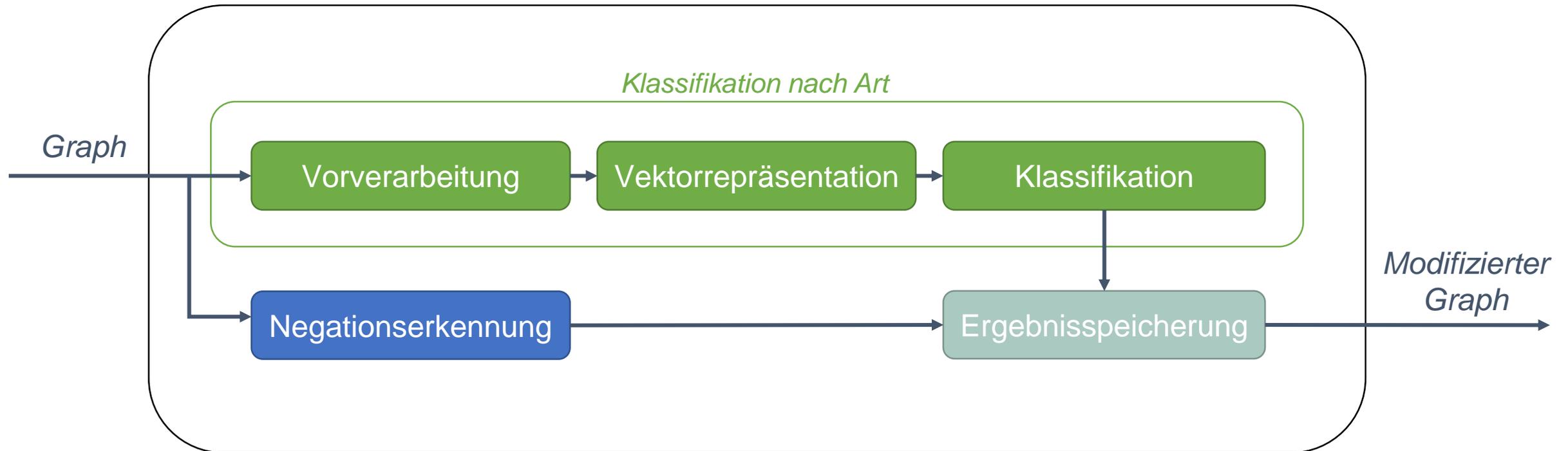
- Analyse von 200 Beispielsätzen
 - Zustände und Prozesse



Kategorien

- Vier Hauptkategorien
 - Aktionsbeschreibung (*action*): The user can edit their contact information.
 - Ereignis (*occurrence*): New students shall become members of the library.
 - Zustandsbeschreibung (*state*): Students shall be members of the library.
 - Aggregation (*aggregation*): Every student shall have a library account.
- Unterscheidung zwischen positiven und negativen Formen
 - Negative Aktionsbeschreibung: The user cannot change their date of birth.
- Acht Klassen: *positive_action*, *negative_action*, *positive_occurrence*, *negative_occurrence*, *positive_state*, *negative_state*, *positive_aggregation*, *negative_aggregation*

Bestimmung von Satzkategorien

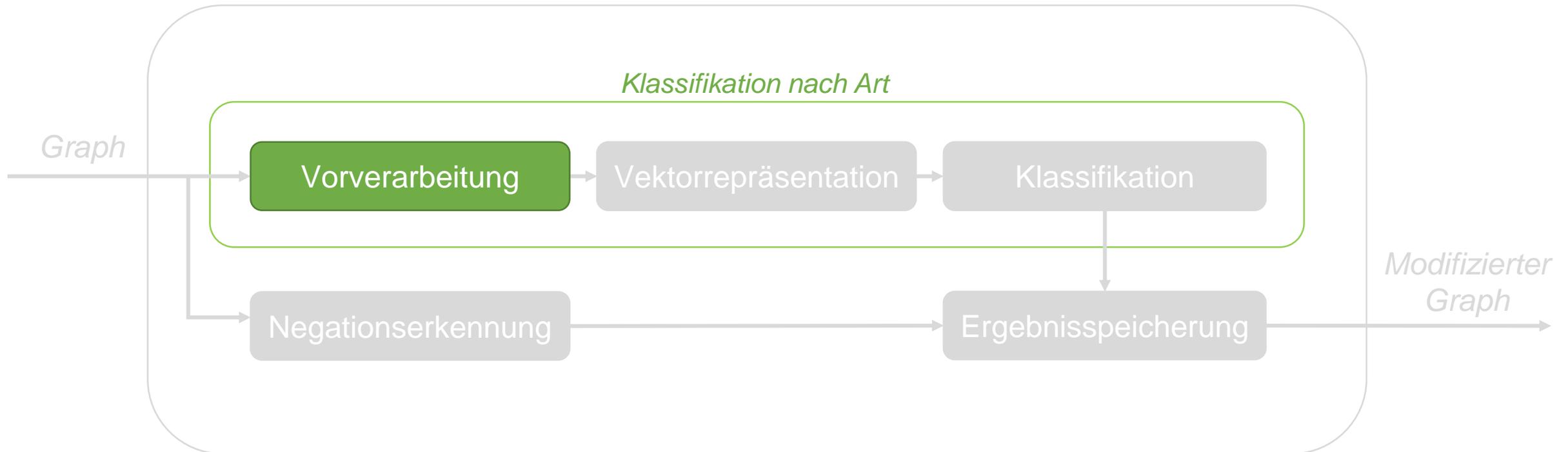


Bestimmung von Satzkategorien



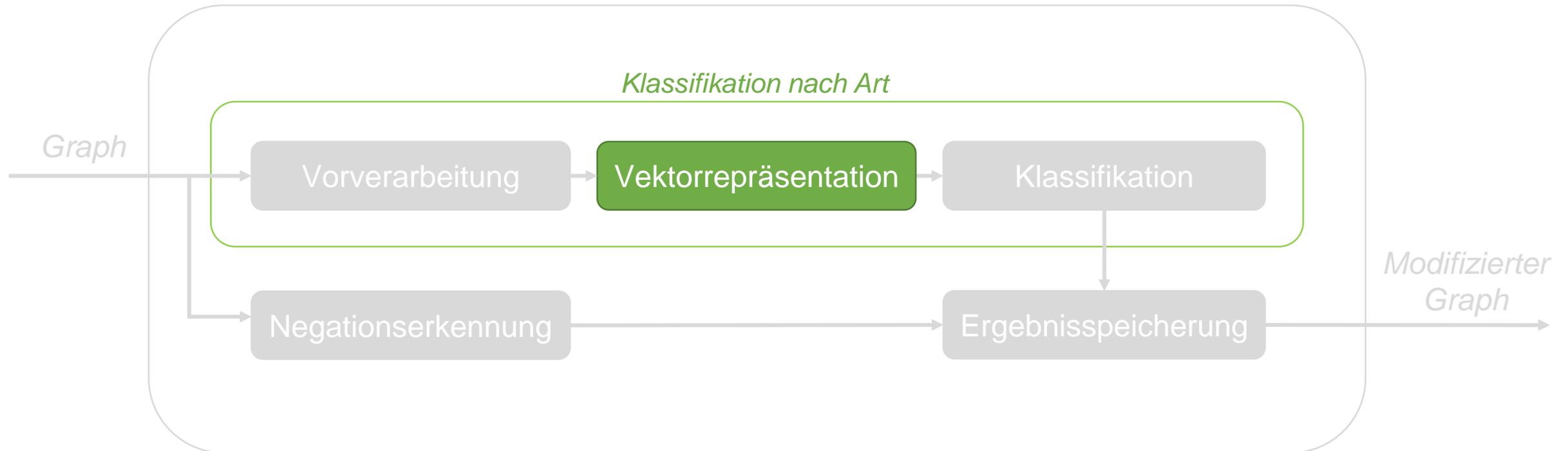
- Verwendung von Abhängigkeitsbäumen

Bestimmung von Satzkategorien

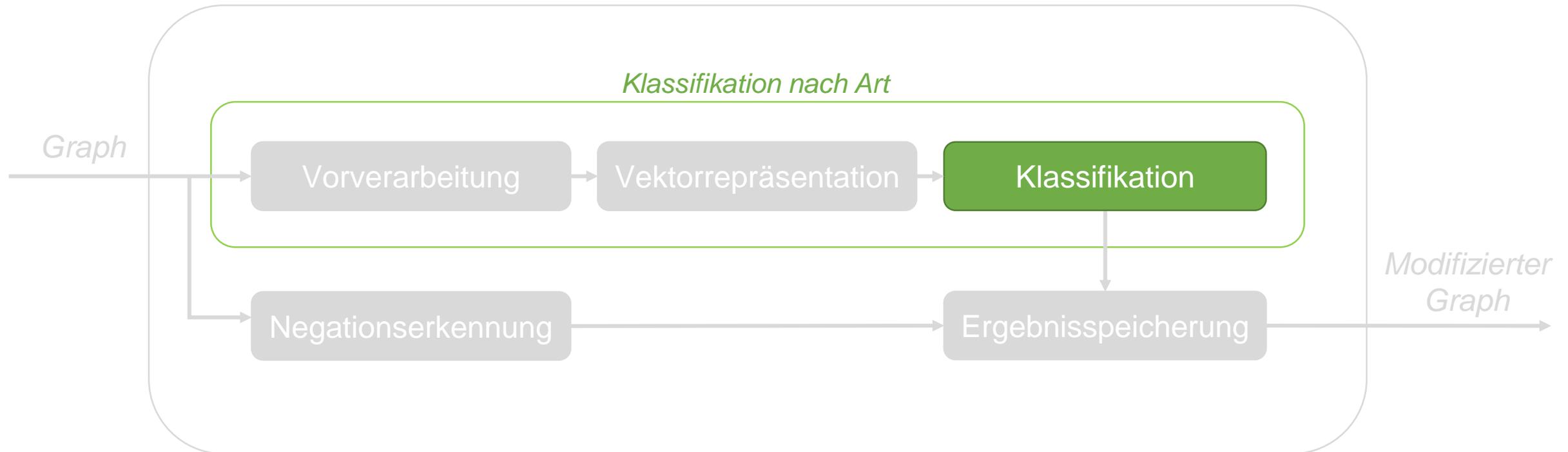


- Entfernung von Satzzeichen + Umwandlung in Kleinbuchstaben

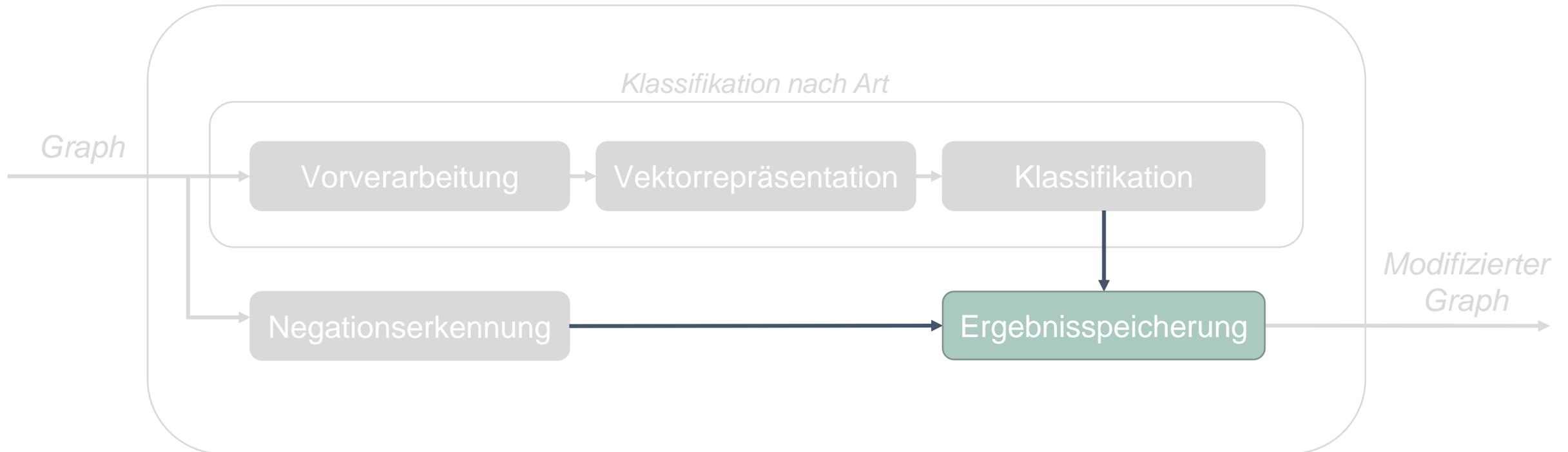
Bestimmung von Satzkategorien



Bestimmung von Satzkategorien



Bestimmung von Satzkategorien

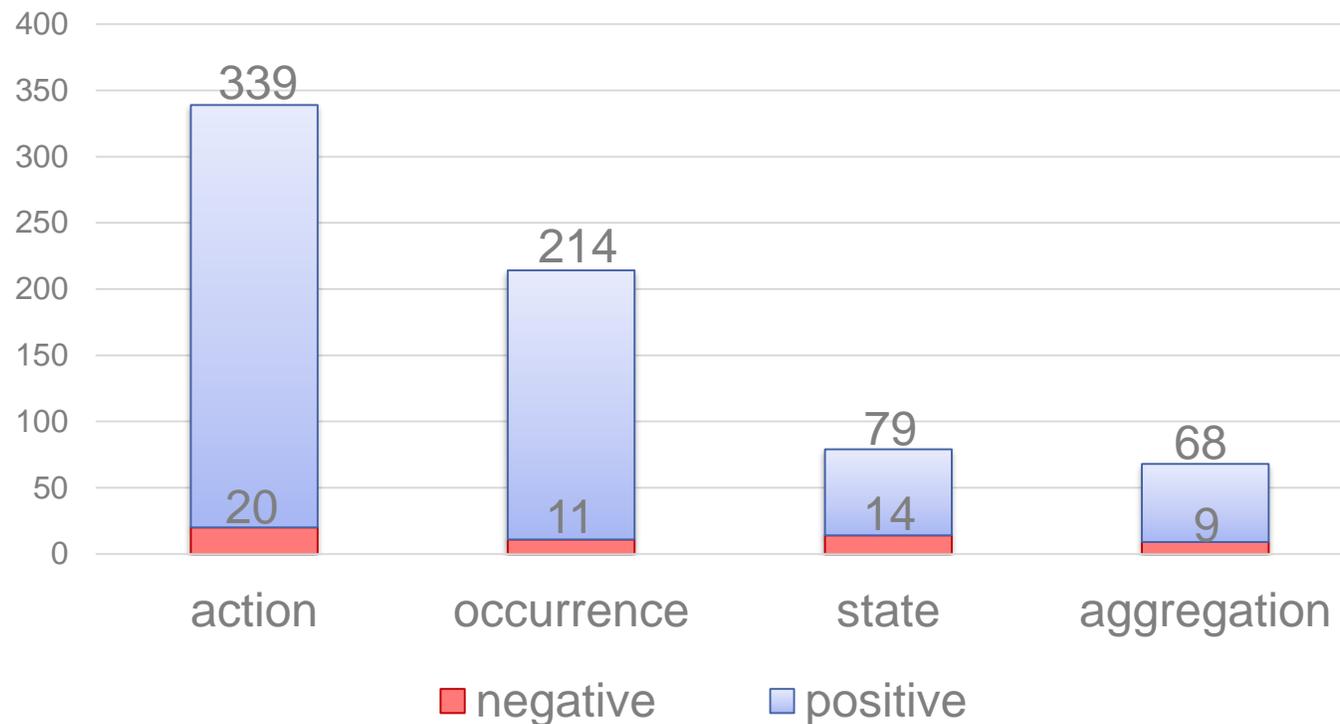


- Neue Knoten in den Graphen einfügen

Evaluation: Datensatz

- 700 Sätze aus 5 Projekten
- Zufällige Auswahl + manuelle Überprüfung

Aufteilung nach Kategorien



Evaluation: Ergebnisse auf dem gesamten Datensatz

- Klassifikation nach semantischen Kategorien: Genauigkeit

Merkmale	Logistische Regression	Stützvektormaschine	Random Forest
Bag-Of-Words	0,87	0,87	0,84
Bag-Of-N-Grams: Zeichen*	0,89	0,90	0,89
Bag-Of-N-Grams: Wörter	0,87	0,88	0,78
TF-IDF	0,86	0,85	0,84
Worteinbettungen: GloVe	0,83	0,83	0,72
Worteinbettungen: GoogleNews-Word2Vec	0,85	0,84	0,73
Doc2Vec	0,84	0,87	0,86

*Zeichenketten der Länge von 2 bis 5 Zeichen

- Negationserkennung: Genauigkeit 1,00

Evaluation

- Übertragbarkeit der Verfahren auf neue Projekte untersuchen
- Aufteilung des Datensatzes nach Projekten
 - Trainieren: 4, Test: 1

Projekt	Aktion	Ereignis	Zustand	Aggregation	Gesamt
NFR	139	43	14	15	202
SR Risk Prediction	56	16	6	17	98
iTrust	88	83	37	22	230
Warc	24	23	15	8	70
IceBreaker	32	55	7	6	100

Evaluation: Ergebnisse auf iTrust

■ Genauigkeit

Merkmale	Logistische Regression	Stützvektormaschine	Random Forest
Bag-Of-Words	0,71	0,72	0,62
Bag-Of-N-Grams: Zeichen	0,72	0,72	0,79
Bag-Of-N-Grams: Wörter	0,68	0,69	0,51
TF-IDF	0,68	0,66	0,63
Worteinbettungen: GloVe	0,67	0,69	0,56
Worteinbettungen: GoogleNews-Word2Vec	0,74	0,71	0,56
Doc2Vec	0,61	0,60	0,62

Evaluation: Ergebnisse auf iTrust

- Random Forest: Genauigkeit 0,79
- Klassifizierungsbericht

Klasse	Präzision	Ausbeute	F ₁ -Maß	Anzahl
action	0,80	0,94	0,86	88
occurrence	0,76	0,82	0,79	83
state	0,71	0,41	0,52	37
aggregation	1,00	0,73	0,84	22

Evaluation

■ Tests auf einzelnen Projekten

Testprojekt	Beste Konfiguration	Genauigkeit
iTrust	Bag-Of-N-Grams: Zeichen + RF	0,79
NFR	Bag-Of-N-Grams: Zeichen + SVM	0,93
SR Risk Prediction	TF-IDF + SVM	0,88
Warc	Bag-Of-N-Grams: Zeichen + LR	0,81
IceBreaker	Bag-Of-N-Grams: Zeichen + LR	0,96

Fazit

- Ziel: Klassifikation von Sätzen in Anforderungsbeschreibungen nach ihrer Semantik
- Ansatz
 - Erkennung von Satzkategorien: Bag-Of-N-Grams auf Zeichen, Random Forest
 - Negationserkennung mittels Abhängigkeitsbäume
- Ergebnisse
 - Genauigkeit 0,89 auf dem gesamten Datensatz
 - Genauigkeit 0,79 auf iTrust
- Ausblick
 - Erweiterung des Datensatzes
 - Anreicherung mit semantischen Rollen
 - Verwendung von Worteinbettungen und einem neuronalen Netzwerk
 - Aufteilung von zusammengesetzten Sätzen

If the answer to the security question is correct,

Zustandsbeschreibung

the user can change their password.

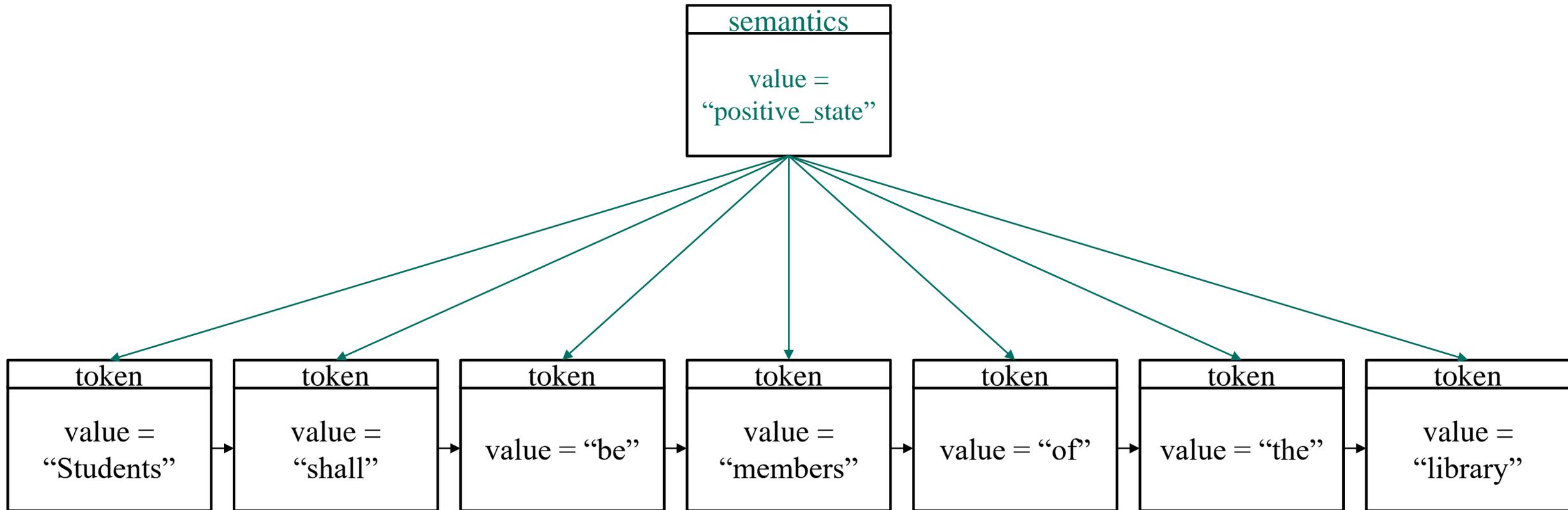
Aktionsbeschreibung

Literatur

- Hey19] Hey, Tobias: Intent-driven Requirements-to-Code Traceability. In: *ICSE'19 DS: Doctoral Symposium of the 2019 IEEE/ACM 41th International Conference on Software Engineering*, 2019
- [VR11] Vlas, R. ; Robinson, W. N.: A Rule-Based Natural Language Technique for Requirements Discovery and Classification in Open-Source Software Development Projects. In: *2011 44th Hawaii International Conference on System Sciences*, 2011
- [KM17] Kurtanovic, Zijad; Maalej, Walid: Automatically Classifying Functional and Non-functional Requirements Using Supervised Machine Learning. In: *2017 IEEE 25th International Requirements Engineering Conference (RE)*. Lisbon, Portugal: IEEE, September 2017.
- [Kim14] Kim, Yoon: Convolutional Neural Networks for Sentence Classification, 2014
- [KM16] Komninos, Alexandros; Manandhar, Suresh: Dependency Based Embeddings for Sentence Classification Tasks. In: *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. San Diego, California: Association for Computational Linguistics, Juni 2016

Literatur

- [CHMLP07] Cleland-Huang, Jane ; Mazrouee, Sepideh ; Liguio, Huang ; Port, Dan: *nfr*.
<http://dx.doi.org/10.5281/zenodo.268542>. Version: März 2007
- [sNZ18] shaukat, Zain ; Naseem, Rashid ; Zubair, Muhammad: Software Requirement Risk Prediction Dataset. <http://dx.doi.org/10.5281/zenodo.1209601> . Version: März 2018
- <http://www.coest.org/>, zuletzt besucht am 11.11.2019



Evaluation: Ergebnisse auf iTrust

- Stützvektormaschine: Genauigkeit 0,72
- Klassifizierungsbericht

Klasse	Präzision	Ausbeute	F_1 -Maß	Anzahl
action	0,82	0,77	0,80	88
occurrence	0,71	0,75	0,73	83
aggregation	0,89	0,77	0,83	22
state	0,44	0,49	0,46	37

Stützvektormaschine

aggregation	17	2	1	2
action	2	68	7	11
state	0	7	18	12
occurrence	0	6	15	62
	aggregation	action	state	occurrence

Random Forest

aggregation	16	5	0	1
action	0	83	1	4
state	0	6	15	16
occurrence	0	10	5	68
	aggregation	action	state	occurrence

Evaluation: Ergebnisse auf iTrust

- Random Forest: Genauigkeit 0,79
- Negationserkennung: Genauigkeit 1,00
- Zusammenfassung der Ergebnisse

Klasse	Präzision	Ausbeute	F ₁ -Maß	Anzahl
positive_action	0,82	0,96	0,89	80
negative_action	0,60	0,75	0,67	8
positive_occurrence	0,82	0,82	0,82	77
negative_occurrence	0,42	0,83	0,56	6
positive_aggregation	1,00	0,80	0,89	15
negative_aggregation	1,00	0,57	0,73	7
positive_state	0,67	0,38	0,49	26
negative_state	0,83	0,45	0,59	11

Evaluation: Ergebnisse auf NFR

■ Genauigkeit

Technik	Logistische Regression	Stützvektormaschine	Random Forest
Bag-Of-Words	0,89	0,90	0,87
Bag-Of-N-Grams: Zeichen	0,92	0,93	0,91
Bag-Of-N-Grams: Wörter	0,88	0,90	0,86
TfIdf	0,89	0,89	0,89
Worteinbettungen: GloVe	0,80	0,79	0,77
Worteinbettungen: GoogleNews-Word2Vec	0,83	0,79	0,77
Doc2Vec	0,79	0,80	0,80

Evaluation: Ergebnisse auf Risk prediction

■ Genauigkeit

Technik	Logistische Regression	Stützvektormaschine	Random Forest
Bag-Of-Words	0,87	0,86	0,78
Bag-Of-N-Grams: Zeichen	0,85	0,84	0,87
Bag-Of-N-Grams: Wörter	0,86	0,86	0,74
TfIdf	0,85	0,88	0,84
Worteinbettungen: GloVe	0,85	0,84	0,68
Worteinbettungen: GoogleNews-Word2Vec	0,85	0,85	0,69
Doc2Vec	0,73	0,78	0,77

Evaluation: Ergebnisse auf Warc

■ Genauigkeit

Technik	Logistische Regression	Stützvektormaschine	Random Forest
Bag-Of-Words	0,74	0,74	0,69
Bag-Of-N-Grams: Zeichen	0,81	0,77	0,77
Bag-Of-N-Grams: Wörter	0,73	0,76	0,69
TfIdf	0,74	0,69	0,73
Worteinbettungen: GloVe	0,66	0,63	0,56
Worteinbettungen: GoogleNews-Word2Vec	0,79	0,70	0,60
Doc2Vec	0,57	0,59	0,60

Evaluation: Ergebnisse auf IceBreaker

■ Genauigkeit

Technik	Logistische Regression	Stützvektormaschine	Random Forest
Bag-Of-Words	0,86	0,82	0,85
Bag-Of-N-Grams: Zeichen	0,96	0,95	0,91
Bag-Of-N-Grams: Wörter	0,90	0,90	0,86
TfIdf	0,83	0,83	0,84
Worteinbettungen: GloVe	0,80	0,81	0,76
Worteinbettungen: GoogleNews-Word2Vec	0,90	0,85	0,74
Doc2Vec	0,82	0,83	0,78