

FP3 - Extraktion von Argumentstrukturen

Forschungsfrage:

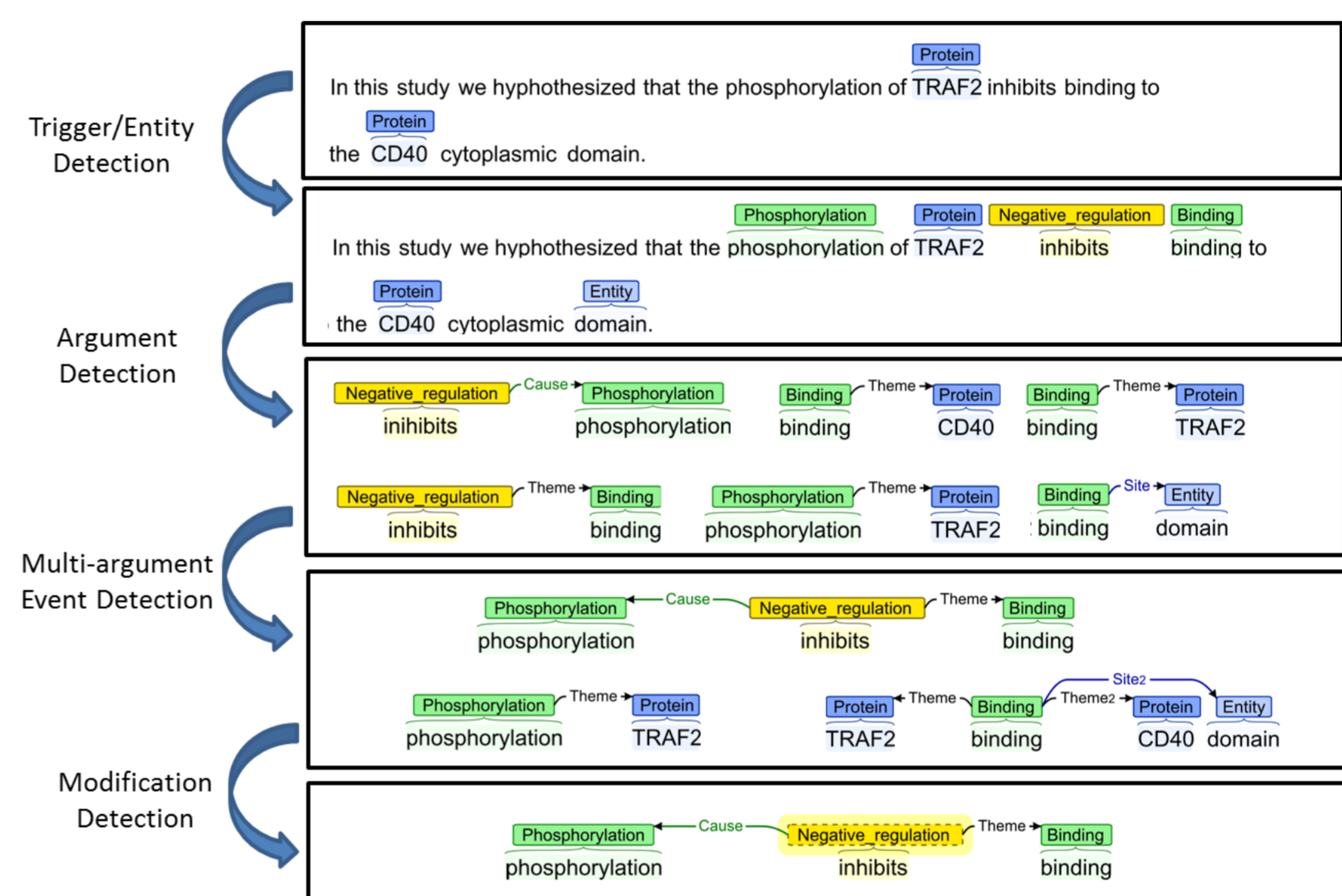
- Vorkommen von Begriffen kein hinreichendes Kriterium für Relevanz eines Dokuments
→ Extraktion von Argumentstrukturen
- Herausforderungen
 - Implizite Argumentation [8]
 - Dokumente in mehreren Sprachen
 - Spezifische Terminologie
 - Individueller Stil (verkürzt, Rechtschreibfehler)

- Konstante Befunde unter Dabrafenib und Trametinib, daher Fortführung der Therapie.
(Befund) (Medikament) (Medikament) (Rel) (Therapie)
- Größenzunahme unter Dabrafenib und Trametinib. Therapieabbruch.
(Befund) (Medikament) (Medikament) (Therapie)
- Konstante Befunde unter Dabrafenib/Trametinib. Therapieabbruch wegen Nebenwirkungen.
(Befund) (Medikament) (Medikament) (Therapie) (Rel)

Beispielannotationen in einem Arztbrief

Stand der Technik:

- Extraktion expliziter Argumentationsstrukturen
- Bisher hauptsächlich in argumentativen Aufsätzen [7], sozialen Medien [9], und medizinische Diagnostik [6]
- Implizite Argumentation noch schwierig
- Hilfsmittel: (partielle) semantische Äquivalenz von Aussagen [1,4]



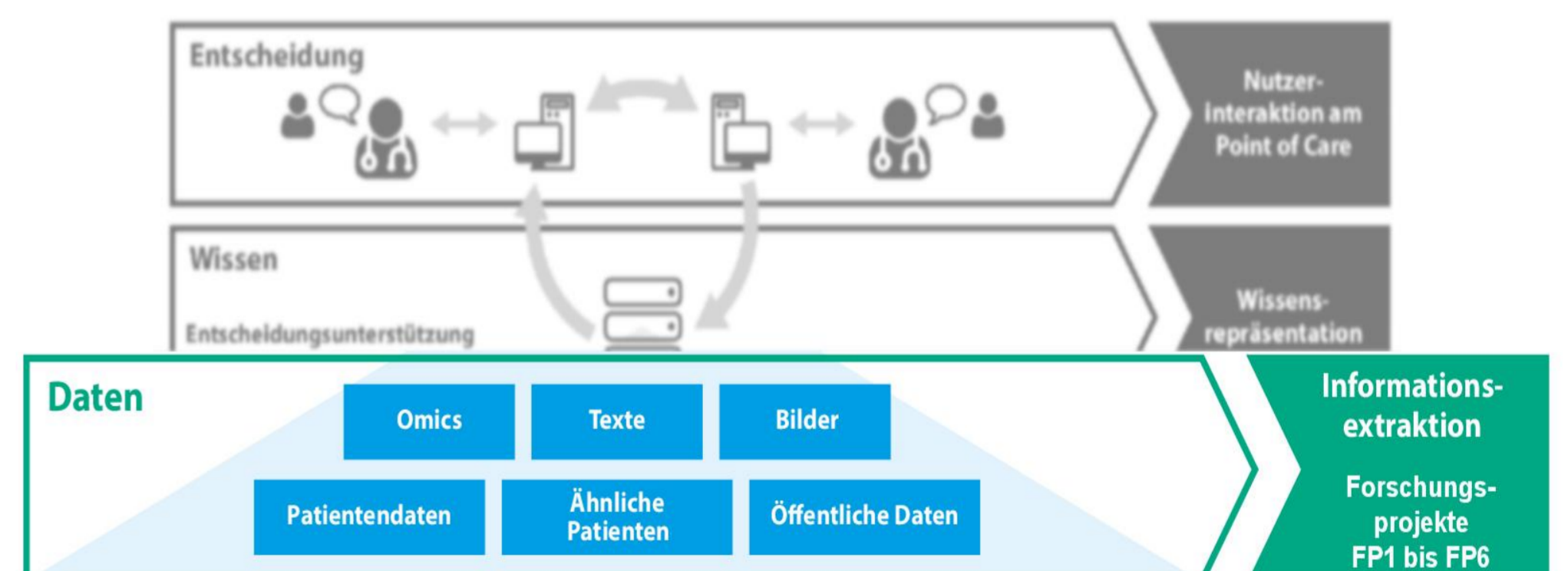
Extraktion von Argumenten in biomedizinischen Artikeln [5]

Lösungsansatz:

- Zusammenarbeit mit FP2
→ Zugriff auf mit Terminologie annotierte Texte
- Erstellung eines Referenzkorpus
 - Abdeckung aller relevanten Dokumentklassen
 - Annotation von Argumentstrukturen
 - Möglichst frei veröffentlichen (→ Anonymisierung)
- Entwicklung und Evaluation der Extraktionsalgorithmen auf Grundlage des Korpus

Integration:

- Integration in DKPro/UIMA Textanalyse Architektur [2,3]
- Re-trainierbar auf neue Daten/Anwendungsbereiche
- Anbindung über SHIP APP Schnittstelle (JVM)



Prof. Dr.-Ing.
Torsten Zesch

Universität Duisburg-Essen



Prof. Dr. med.
Dirk Schadendorf

Universitätsklinikum Essen

Literatur

- Benikova D, Zesch T., Same same, but different: Compositionality of paraphrase granularity levels, In Proc. of RANLP, 2017.
- Daumke P., Heitmann, K. U., Heckmann, S., Martiinez-Costa, C., & Schulz, S. Clinical Text Mining on FHIR. *Studies in Health Technology and Informatics*, 264, 83–87, 2019
- Horsmann T., Zesch T., DeepTC – An Extension of DKPro Text Classification for Fostering Reproducibility of Deep Learning Experiments, In Proceedings of LREC 2018.
- Levy O., Zesch T., Dagan I., Gurevych I., Recognizing Partial Textual Entailment, In Proceedings of ACL, 2013.
- Miwa M., Thompson, P., McNaught, J., Kell D.B., Ananiadou, S. Extracting semantically enriched events from biomedical literature. *BMC Bioinformatics*, 13:108, 2012.
- Schulz, C., Meyer, C. M., Kiesewetter, J., Sailer, M., Bauer, E., Fischer, M. R., Fischer, F., Gurevych, I., Analysis of Automatic Annotation Suggestions for Hard Discourse-Level Tasks in Expert Domains. In Proceedings of ACL (pp. 2761–2772), 2019
- Stab, C., Gurevych, I. Parsing argumentation structures in persuasive essays. *Computational Linguistics* 43.3, 619–659, 2017.
- Wojatzki M., Zesch T., Stance-based Argument Mining – Modeling Implicit Argumentation Using Stance, In Proc. of KONVENS, 2016.
- Wojatzki M., Zesch T. Itl.uni-due at SemEval-2016 Task 6: Stance Detection in Social Media Using Stacked Classifiers. In Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016); San Diego, California.