

Sachbericht

MLwin — Maschinelles Lernen in Wissensgraphen

Volker Tresp, Hinrich Schütze, Peter Fasching, Daniel Cremers, Afshin Sadeghi,
Patricia Oppelt, Matthias Schubert, Mehdi Ali, Hanna Hübner, Vladimir Golkov,
Yinchong Yang

Das diesem Bericht zugrunde liegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01IS18050 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei der Autorin/beim Autor.



Projektlaufzeit: 01. Oktober 2018 – 31. Januar 2022

Kooperationspartner

Siemens AG (FKZ: 01IS18050A)

Projektleitung: Volker Tresp (Verbundkoordinator)
Adresse: Otto-Hahn-Ring 6, 81739 München
Tel.: +49 1735293072
E-Mail: Volker.Tresp@siemens.com

LMU (FKZ: 01IS18050C)

Projektleitung: Prof. Dr. Hinrich Schütze
Adresse: Oettingenstr. 67, 80538 München
Tel.: +49 89 2180 9720
E-Mail: hs2016@cislmu.org

TUM (FKZ: 01IS18050E)

Projektleitung: Prof. Dr. Daniel Cremers
Adresse: Boltzmannstr. 3, 85748 Garching bei München
Tel.: +49 89 289-17755

E-Mail: cremers@tum.de

Uni Bonn (FKZ: 01IS18050D)

Projektleitung: Prof. Dr. Jens Lehmann
Adresse: Friedrich-Hirzebruch-Allee 5, 53115 Bonn
Tel.: +49 228 73-4315
E-Mail: jens.lehmann@cs.uni-bonn.de

UKER (FKZ: 01IS18050B)

Projektleitung: Prof. Dr. Peter A. Fasching
Adresse: Universitätsstraße 21-23, 91054 Erlangen
Tel.: 09131 8533553
E-Mail: peter.fasching@uk-erlangen.de

FhG IAIS (FKZ: 01IS18050F)

Projektleitung: Afshin Sadeghi
Adresse: Schloss Birlinghoven, 1, 53757 Sankt Augustin
Tel.: +49 2241 14-2812
E-Mail: afshin.sadeghi@iais.fraunhofer.de

IFG (FKZ: 01IS18050G)

Projektleitung: Prof. Dr. Patricia G. Oppelt
Adresse: Hindenburgstr. 50, 91054 Erlangen
Tel.: 09131 91880613
E-Mail: management@ifg-erlangen.de

Inhaltsverzeichnis

1	Verwendung der Zuwendungsmittel sowie Darstellung der inhaltlichen Arbeiten und erzielten Ergebnisse	4
2	Zahlenmäßiger Nachweis	20
3	Notwendigkeit und Angemessenheit der geleisteten Arbeiten	20
4	Nutzen und Verwertbarkeit der Ergebnisse sowie zukünftiger Verwertungsplan	22
5	Zwischenzeitlicher technischer Fortschritt auf dem Gebiet des Vorhabens	23
6	Nachweis erfolgter oder geplanter Veröffentlichungen der Ergebnisse	24
7	Anhang: Quantitative Forschungsergebnisse	37

Abkürzungsverzeichnis

KI	Künstliche Intelligenz (Artificial Intelligence)
ML	Maschinelles Lernen (Machine Learning)
DL	Deep Learning
NLP	Natural Language Processing
CV	Computer Vision
KG	Knowledge Graph
AP	Arbeitspaket

1 Verwendung der Zuwendungsmittel sowie Darstellung der inhaltlichen Arbeiten und erzielten Ergebnisse

AP 2: Maschinelles Lernen und Data Mining in KGs

Wir haben in einem iterativen Ansatz die Software PyKEEN¹ (Python KnowlEdge EmbeddiNGs) [6] entwickelt. PyKEEN ermöglicht das effiziente Trainieren und Evaluieren von Knowledge Graph Embedding Modellen. Die erste Version umfasste 10 etablierte Modelle, einen Trainingsansatz, Random Search als Hyperparameteroptimierungs-Algorithmus, und zwei Evaluationsmetriken [6]. Die modulare Architektur von PyKEEN vereinfachte die Erweiterung um weitere Komponenten. Die erste Version wurde bereits erfolgreich in verschiedenen Anwendungen eingesetzt, darunter auch bei Bayer Crop Science, um Repräsentationen für Knoten und Kanten für Wissensgraphen zu erlernen, die Feldexperimente beschreiben [6]. In einer weiteren Anwendung wurden Äquivalenzen und Hierarchien zwischen biologischen Pathways berechnet, die in unterschiedlichen biomedizinischen Datenbanken (KEGG, WikiPathways, Reactome) enthalten sind [5].

Durch die Erfahrungswerte, die wir durch die Nutzung von PyKEEN erhalten haben und ersten Forschungsergebnissen, die darauf hindeuten, dass die Performanz von Knowledge Graph Embedding Modellen nicht ausschließlich auf die Modellarchitektur beruht [46], entworfen wir PyKEEN in einem zweiten Schritt neu mit dem Fokus auf vollständiger Flexibilität bei der Zusammenstellung von Knowledge Graph Embedding Modellen [4]. Die neu konzipierte Version von PyKEEN umfasst derzeit 44 Modelle, 15 Kostenfunktionen, zwei Trainingsansätze, fünf Regularisierungsansätze, 44 Evaluationsmetriken, und 41 integrierte Datensätze. In PyKEEN, wird ein Knowledge Graph Embedding Modell als eine Komposition von folgenden vier Komponenten definiert: Interaktionsmodell (Modellarchitektur), Trainingsansatz, Kostenfunktion, und die Modellierung (bzw. nicht Modellierung) von expliziten inversen Tripel. Diese grundlegende Abstraktion, die sich von der ersten Version von PyKEEN unterscheidet, ermöglicht es den Effekt jeder einzelner Komponente oder das Zusammenspiel mehrere Komponenten auf die Performanz des Knowledge Graph Embedding Modells zu untersuchen.

Wir verwendeten PyKEEN, um eine Reproduzierbarkeitsstudie und die umfangreichste Link-Prediction Benchmarkingstudie basierend auf Knowledge Graph Embedding Modellen durchzuführen. Wir untersuchten im ersten Schritt, ob und welche Link-Prediction Experimente reproduziert werden können. Teil unserer Studien waren die Datensätze FB15K, FB15K-237, WN18, und WN18RR, und 15 Knowledge Graph Embedding Modelle. Wir fanden heraus, dass für FB15K vier von 13, für FB15K-237 zwei von drei, für WN18 5 von 13, und für WN18RR drei von fünf bis zu einem Toleranzwert reproduziert werden konnten [3].

Im zweiten Schritt führten wir die umfangreichste Link-Prediction-Benchmarking-Studie basierend auf Knowledge Graph Embedding Modellen durch. Wir untersuchten die Fragestellung, ob die Link-Prediction Performanz eines Knowledge Graph Embedding Modells ausschließlich von der Modellarchitektur abhängt wie in der Vergangenheit oftmals implizit angenommen wurde, oder von der Kombination aus Interaktionsmodell, Trainingsansatz, Kostenfunktion, und die Modellierung (bzw. nicht Modellierung von inversen Relationen). Unsere Studie umfasste mehr als 73.683 Experimente, die 24.804 GPU-Stunden in Anspruch nahmen. Wir führten unsere Experimente auf den Datensätzen Kinships, WN18RR, FB15K-237, und YAGO3-10 aus. Unsere Ergebnisse verdeutlichen, dass oftmals nicht ausschließlich das Interaktionsmodell allein die Leistung eines Knowledge Graph Embedding Modells allein bestimmt, sondern das Zusammenspiel dessen vier Komponenten. Interaktionsmodelle, die als Baseline betrachtet werden können bei passender Konfiguration mit State-of-The-Art Interaktionsmodellen konkurrieren bzw. übertreffen. Tabelle 11 fasst die Hauptergebnisse unserer Benchmarkingstudie zusammen. Es existiert keine Konfiguration, die für jeden Datensatz am besten abschnitt. Stattdessen existieren je nach Datensatz mehrere Konfigurationen, die ähnlich gut abschneiden. Des Weiteren existiert keine Kostenfunktion, die immer das beste Ergebnis erhält. Unsere Studie zeigte zusätzlich auf, dass die Modellierung von expliziten inversen Relationen insbesondere in Kombination des LCWA-Trainingsansatzes performant ist.

¹ <https://github.com/pykeen/pykeen>

Abschließend konnten wir für mehrere Interaktionsmodelle Konfigurationen ermitteln, die bessere Ergebnisse lieferten als ursprünglich publizierten. Dazu gehörte auch eine neue Konfiguration, die (zusammen mit dem Modell Graph Attenuated Attention Networks [102]) neue State-of-The-Art-Ergebnisse auf dem Datensatz WN18RR lieferte.

Ferner wurde ein hybrider wissensgetriebener und datengetriebener Ansatz entwickelt, um expressive Repräsentation von Patienten zu lernen. Dazu wurde ein biomedizinischer Wissensgraph mit Patientendaten angereichert, um darauf expressive Patientenrepräsentation basierend auf KGEMs zu lernen. Die gelernten Repräsentationen wurden zur finalen Klassifikation zwischen gesunden und kognitiv beeinträchtigten Patienten verwendet. Es konnte gezeigt werden, dass die gelernten Patientenrepräsentationen die Klassifikation maßgeblich verbessern[10].

Des Weiteren wurden sechs neue hyper-relationale Benchmark-Datensätze erstellt, die die Evaluation von induktiven Link-Prediction Ansätzen innerhalb von hyper-relationalen Wissensgraphen ermöglichen [2]. Hyper-relationale Wissensgraphen enthalten zusätzlich zu den Tripeln Meta-Informationen auf Tripelebene. Die Datensätze sind in unserem Artikel „Improving Inductive Link Prediction Using Hyper-Relational Facts [2]“ für welchen wir auf der International Semantic Web Conference 2021 den „Best Paper Award“ gewonnen haben beschrieben und unter <https://zenodo.org/record/6581274#.YtFMysZCRhE> verfügbar.

Außerdem wurde das KGE Modell LiteralE mitentwickelt, das numerische und textuelle Literale der Knoten berücksichtigt wodurch Kanten in Wissensgraphen welche Literale enthalten präziser vorhergesagt werden können. Der Artikel wurde in der „International Semantic Web Conference“ veröffentlicht, und das Modell wurde in PyKEEN integriert.

Ferner wurde PyTorch Lightning in unsere Software PyKEEN integriert wodurch das Multi-GPU-Training von Knowledge Graph Embedding Modellen ermöglicht wird. Dadurch können effizient Embeddings für große Wissensgraphen berechnet werden.

Fraunhofer IAIS hat im Jahr 2019 das MDE [77] Knowledge Graph Embedding Modell entwickelt. Es wurde in einem wissenschaftlichen Paper beschrieben, das auf der European Conference on Artificial Intelligence (ECAI) angenommen wurde. Das Modell ermöglicht das implizite Lernen von vier Beziehungsmustern. Diese Beziehungsmuster, die logischen Regeln ähneln, ermöglichen eine bessere Vorhersage versteckter Beziehungen in Knowledge Graph Embeddings (KGEs).

Tabelle 12 zeigt die Verbesserung der Verbindungsvorhersageergebnisse durch MDE auf dem Benchmark von WN18RR und FB15k-237 und Tabelle 13 zeigt die Ergebnisse von MDE auf den Datensätzen der Länder, die ein Benchmark zur Bewertung der Inferenz von Zusammensetzungsmustern sind. Dieses Ergebnis zeigt die bessere Wirksamkeit von MDE für Beziehungen mit diesem Beziehungsmuster.

Das Fraunhofer IAIS hat das GFA-NN [76] entwickelt, ein Modell zur Einbettung von Wissensgraphen, das auf dem Lernen von Graphenmerkmalen basiert und ein neuronales Netz ist. Die Arbeit wurde auf der European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2021) vorgestellt.

Das GFA-NN-Modell für die Einbettung von Wissensgraphen berücksichtigt nicht nur das Lernen von Beziehungen, sondern auch das Lernen von Grapheneigenschaften, wie die relative Position von Knoten und deren Grad. Das Modell ist effizient bei der Vorhersage von Verbindungen in der konventionellen transduktiven Einstellung und auch dann, wenn die abgefragten Entitäten nicht im Trainingsdatensatz enthalten sind (induktive Einstellung). Es übertrifft auch den Stand der Technik bei der Vorhersage von Links in großen Datensätzen.

Tabelle 14 und Tabelle 15 fassen die Leistung von GFA-NN und anderen KGE-Modellen in der transduktiven Link-Vorhersageaufgabe zusammen. Wir stellen fest, dass GFA-NN andere moderne KGEs bei WN18RR übertrifft und konkurrenzfähige Ergebnisse bei FB15k-237 liefert. Tabelle 15 zeigt die *MRR*-Evaluierungsergebnisse für den vergleichsweise großen biologischen Datensatz namens *ogbl-biokg*. In diesem Benchmark ist die Anzahl der Entitäten und Trainingsbeispiele viel größer als in den Datensätzen WN18rr und FB15k-237.

Die Fähigkeit, aufmerksame Merkmalseinbettungen zu lernen, ist bei dieser transduktiven Link- Vorhersageaufgabe entscheidend. Während die besten KGEs nur ein MRR von 0,8105 im Validierungsdatensatz und 0,8095 im Testdatensatz erreichen können, erreicht GFA-NN 0,901 in beiden Datensätzen und verbessert damit den Stand der Technik um 9 Prozent.

Dieser große Abstand zwischen den Ergebnissen unterstützt die Annahme, dass eigenschaftsbezogene Einbettungen in größeren realen Netzwerken die bisherigen Methoden übertreffen.

Tabelle 16 fasst die Hit@10-Ranking-Leistung von GFA-NN im Vergleich zu Methoden zusammen, die für die induktive Linkvorhersage spezifiziert wurden. Obwohl wir GFA-NN nicht explizit für diese Aufgabe entwickelt haben, stellen wir fest, dass GFA-NN in dieser Umgebung sehr konkurrenzfähig ist und die besten induktiven Lernmodelle in den meisten Fällen übertrifft. Dieses Ergebnis stützt unsere Hypothese, dass Wissensgrapheneinbettungen, die auf Positions- und Zentralitätseigenschaften achten, für Vorhersageaufgaben in anspruchsvollen Umgebungen, d.h. für induktive Linkvorhersagen, von Vorteil sind.

Eine der größten Herausforderungen im Arbeiten mit Wissensgraphen ist die Extraktion und Verarbeitung von Wissensgraphen aus unstrukturierten Datenquellen. Geschieht dies im Rahmen von automatischen Verfahren kann es dabei vorkommen, dass mehrere semantisch unterschiedliche Beziehungen ein und demselben Beziehungstyp zugeordnet werden. Zum Beispiel, geschieht dies bei der Extraktion von Relationen aus Verben mit unterschiedlichen Wortbedeutungen. Im Rahmen des Arbeitspakets wurden zunächst erste Methoden zur Dissoziation von Relationen im Rahmen der Masterarbeit „Relation Disambiguation in Knowledge Graphs“ untersucht. Anschließend wurde die Methode V-Coder [26] entwickelt, um Beziehungen zu identifizieren, die Entitäten aus verschiedenen Domänen miteinander verbinden. Diese Beziehungen können als mehrdeutig betrachtet werden und sind Kandidaten für eine Zerlegung. Ähnlich wie bei der Adaptiven Lerntheorie (ART) lernt unser Modell neue Muster aus dem Wissensgraphen, indem es die Einheiten in einem sogenannten Competitive Layer bekannten Mustern zuweist und neue Muster erzeugt sofern Instanzen zu keinem der bereits bekannten Muster passen. Die Evaluierung auf realen Datensätzen von Freebase, Yago und NELL zeigt, dass der V-Coder nicht nur in der Lage ist, Links aus beschädigten Eingabedaten wiederherzustellen, sondern auch, dass die semantische Zerlegung von Beziehungen häufig zu einer Vereinfachung des Link Prediction Problems führt, da die vorherzusagenden Kategorien klarer voneinander getrennt werden können [17],

Ein häufig auftretendes Problem bei Graphischen Neuronalen Netzen (GNN) ist als “Over-Smoothing” bekannt. Dabei gleichen sich die Darstellungen der Knoten des Eingabegraphen mit steigender Anzahl an Message-Passing Iterationen immer weiter an bis sie quasi ununterscheidbar werden. Auch die neuesten Modelle, die attention-basierte *Graph Transformer Layers* (GTLs) verwenden, sind immer noch auf der schichtweisen Berechnung von GNNs beschränkt, der solche Effekte nicht verhindern können. Im Rahmen des Arbeitspakets wurde daher das Verfahren *Graph Shell Attention* (SEA) entwickelt, das mit Hilfe eines Routing-Mechanismus diesen Effekt entgegenwirken soll. Konkret werden verschiedene Repräsentationen für Knoten gelernt, die sich bezüglich der Größe der Nachbarschaft des betrachteten Knotens unterscheiden. Intuitiv gewinnen die Modelle durch die Verwendung mehrerer Experten an Ausdruckskraft, so dass die Repräsentation eines Knotens ausschließlich auf Knoten basiert, die sich im rezeptiven Feld eines Experten befinden. Wir evaluieren unsere Architektur auf verschiedenen Referenzdatensätzen und zeigen konkurrenzfähige Ergebnisse, während wir die Anzahl der Parameter im Vergleich zu State-of-the-Art-Modellen allerdings drastisch reduziert hat [18]. Das Verfahren wird auf der European Conference for Machine Learning (ECML/PKDD) 2022 vorgestellt [25].

In einem temporalen Wissensgraphen werden Fakten durch Quadrupel aus Subjekt, Prädikat, Objekt und Zeitangabe beschrieben. Das Link-Prediction Problem in temporalen Wissensgraphen ist die Vorhersage aller zutreffenden Relationstypen, die in einem gegebenen Zeitintervall zwischen zwei Entitäten gelten. Zur Generierung von interpretierbaren Lösungen für dieses Problem wurde das Verfahren AppTek [24] entwickelt, das mittels Reinforcement Learning semantische Topologien lernt, mit denen die unbekannt Beziehung zweier Entitäten in einem Anfragezeitraum vorhergesagt werden kann. Dabei lernt AppTek zunächst welche Arten von Relationen für die Vorhersage der Beziehung zwischen Subjekt und Objekt eines Fakts relevant sind. Durch die Verwendung von Reinforcement Learning ist das Verfahren dabei

besser in der Lage relevante Informationen für bestimmte Typen von Relationen zu selektieren und irrelevante auszublenden. Die so entstandenen semantischen Topologien können dann analysiert werden, um die Vorhersage des Algorithmus besser zu verstehen. Unsere Experimente zeigen die gute Vorhersageperformanz von AppTek 19 und dass die durch das Verfahren gefundenen Topologien Aufschluß über die Gründe der Vorhersage geben.

Siemens AG hat darüber hinaus einen neuartigen Empfehlungsdienst NECTR (Englisch: Neural Encoders Combined with Tensor Decompositions for Recommendations) entwickelt [39], welches aus einem Autoencoder und einer bilinearen Faktorisierung eines Adjazenzensors eines Wissensgraphen besteht. Die Tensor-Faktorisierung kodiert technische Eigenschaften der Artikel, die durch einen Wissensgraphen modelliert werden. Der Autoencoder erfasst (nicht-lineare) Interaktionen der einzelnen Module, die zusammen das System bilden. Beide Komponenten werden durch gemeinsam gelernte Gewichte gekoppelt. Das NECTR wurde auf einem Datensatz für industrielle Automatisierungslösungen getestet. Dabei konnte gezeigt werden, dass NECTR bessere Ergebnisse als gängige Empfehlungsdienste wie z.B. nicht-negative Matrix-Faktorisierungen oder reine Tensor-Faktorisierungen erzielt.

AP 3: Textanalyse und -generierung mit KGs:

Im Rahmen des Projekts wurden Encoder-Decoder-Modelle zur Transformation von Wissensgraphen in Text und umgekehrt entwickelt. Wissensgraphen können sich von einer Domäne zur anderen stark unterscheiden. Daher werden überwachte Ansätze sowohl für die Graph-zu-Text-Generierung als auch für die Text-zu-Graph-Wissensextraktion (semantisches Parsing) immer unter einem Mangel an domänenspezifischen parallelen Graph-Text-Daten leiden; gleichzeitig ist die Anpassung eines Modells, das für eine andere Domäne trainiert wurde, oft unmöglich, da es nur wenige oder gar keine Überschneidungen bei Entitäten und Beziehungen gibt. Diese Situation erfordert einen Ansatz, der keine großen Mengen an annotierten Daten benötigt und daher nicht auf Techniken zur Domain Adaption angewiesen ist, um in verschiedenen Domänen gut zu funktionieren.

Zu diesem Zweck wurde an der LMU der erste Ansatz zur unüberwachten Texterzeugung aus KGs vorgestellt [87, 88]. Des Weiteren wurde auch gezeigt wie dieser Ansatz auch für unüberwachtes semantisches Parsing verwendet werden kann. Der vorgeschlagene Ansatz wurde mit WebNLG v2.1 und einem neuen Benchmark, der Szenengraphen von Visual Genome nutzt, evaluiert. Dabei wurde gezeigt, dass das vorgeschlagene System beide Text-Graph-Konvertierungsmethoden ohne manuelle Anpassung des Datensatzes übertrifft. Tabelle 20 zeigt die Qualität der aus Wissensgraphen generierten Texte. In zusätzlichen Experimenten wurden die Auswirkungen verschiedener unüberwachter Lernziele untersucht.

Des Weiteren wurde ein neues Modell der Graph-zu-Textgenerierung entwickelt, das auf dem Text-zu-Text-Modell Transformer aufbaut. Unsere Arbeiten ergänzen dieses Modell um eine Graphkodierungskomponente. Dieser Ansatz wurde zunächst als Preprint [82] und anschließend auf einem peer-reviewten Workshop [83] veröffentlicht. Dabei konnte experimentell festgestellt werden, dass das vorgeschlagene neue Modell effizienter arbeitet als vergleichbare Modelle, die auf Graph Neural Networks basieren (vgl. Tabelle 21).

Zusätzlich wurde ein Übersichtsartikel publiziert, der verschiedene Möglichkeiten vergleicht, die Eingabe eines Transformer-Modells zu strukturieren [17]. Im Rahmen des Projekts wurde ebenfalls die Anwendung vortrainierter Sprachmodelle auf das Problem der Graph-zu-Textgenerierung erforscht [73]. Durch die Betreuung der Masterarbeit "Exploring Data Augmentation Techniques for Unsupervised Graph-Text Conversion" wurden in dieses Thema auch Studierende eingebunden.

Ein weiteres Thema des Arbeitspakets ist die Erforschung von Methoden und Embedding-Modellen für Textklassifikation und grammatische Analyse in Low-Resource-Szenarien. Der praktische Erfolg eines Großteils der maschinellen Sprachverarbeitung hängt von der Verfügbarkeit von Trainingsdaten ab. In realen Szenarien sind Trainingsdaten jedoch oft knapp, nicht zuletzt, weil in vielen Anwendungsdomänen Daten nur begrenzt verfügbar und sehr spezifisch sind. In [89] werden verschiedene Methoden verglichen, um mit diesem Problem umzugehen. Dabei werden Vorgehensweisen für die Entwicklung von NLP-Anwendungen, in denen nur eine kleine Menge an gelabelten Trainingsdaten für eine bestimmte

Domäne verfügbar ist, vorgeschlagen. Obwohl das Transfer-Lernen mit vortrainierten Sprachmodellen Vergleichsmethoden bei allen Aufgaben übertrifft, schneiden die Alternativen nicht viel schlechter ab und erfordern gleichzeitig einen viel geringeren Rechenaufwand, was die finanziellen und ökologischen Kosten erheblich reduziert. Im Rahmen des Projekts wurde daher der Trade-Off zwischen Leistung und Rechenaufwand mehrerer Alternativen untersucht. Einige dieser Modelle können dabei bis zu 175-tausendmal schneller trainiert werden und benötigen dabei keine GPU. Ein Vergleich wird in Tabelle 22 gezeigt.

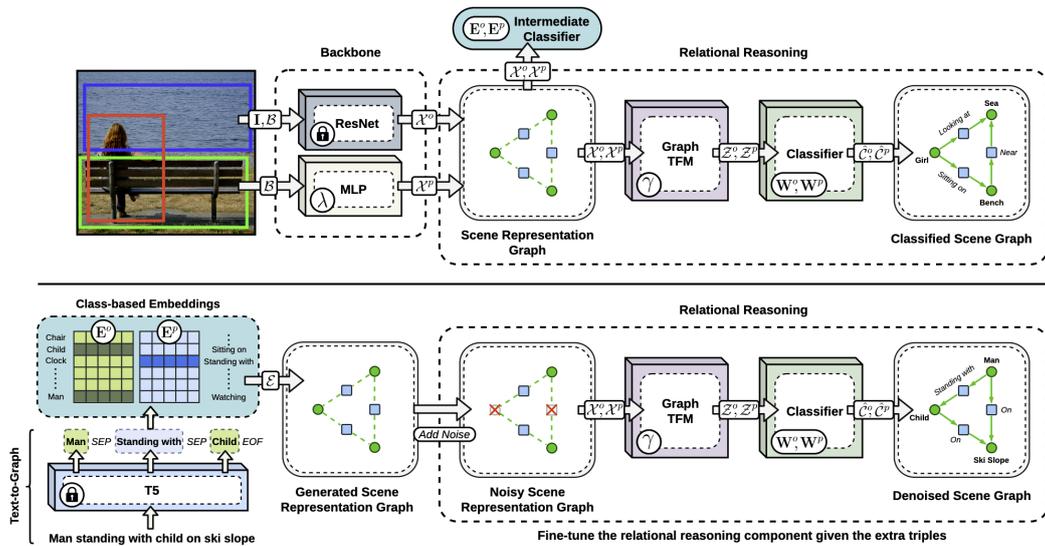
Im Rahmen des Arbeitspakets wurde ein Modell zur Wissensextraktion aus Texten der Form Subjekt-Prädikat-Objekt entwickelt [86, 85, 84]. Die Bearbeitung dieser Form der Texte ist besonders herausfordernd, da die Semantik der Verben bzw. der Wissensgraph-Relationen eine zentrale Rolle spielt und bisher effektive Modelle dafür fehlten. Lexikalische Inferenz im Kontext (LliK) ist die Aufgabe, textuelle Folgerungen zwischen zwei sehr ähnlichen Sätzen zu erkennen, d.h. Sätzen, die sich nur in einem Ausdruck unterscheiden. Sie kann daher als eine Variante der natürlichsprachlichen Inferenzaufgabe betrachtet werden, die sich auf lexikalische Semantik konzentriert. Im Rahmen des Projekts wurden die ersten Ansätze in diesem Gebiet formuliert und evaluiert, die auf vortrainierten Sprachmodellen basieren. Es wurden die folgenden drei Ansätze vorgeschlagen: (i) ein few-shot NLI-Klassifikator, (ii) ein Ansatz zur Beziehungsinduktion, der auf handgefertigten Mustern basiert, die die Semantik der lexikalischen Inferenz ausdrücken, und (iii) eine Variante von (ii) mit Mustern, die automatisch aus einem Korpus extrahiert wurden. Alle vorgeschlagenen Ansätze übertreffen den bisherigen Stand der Technik (Abbildung 7) und zeigen das Potenzial von vortrainierten Sprachmodellen für LliK.

Im Bereich Text und Bilddaten wurde der Einsatz von Szenegraphen für die Generierung von Bildunterschriften erforscht. Dabei hat sich die Diskretisierung in Objekte und konkrete visuelle Relationen als der Hauptvorteil einer Graph basierten Zwischenrepräsentation auf dem Weg von Bild zu Text herauskristallisiert. Dadurch konnten vor allem Fortschritte bei der Vermeidung von Halluzinationen erreicht werden. Diese Arbeit wurde als Preprint auf Arxiv veröffentlicht [64]. Des Weiteren wurden Techniken zur automatischen Erstellung von Wissensgraphen aus Texten untersucht. Insbesondere wurden neue Ansätze zur Ermittlung von Inferenzbeziehungen zwischen Wissensgraph-Relationen entwickelt.

Etablierte Verfahren betrachten die Vorhersage von Beziehungen (Link Prediction) in Wissensgraphen als rein transduktive Aufgabe, die keine Schlussfolgerungen über ungesehene Entitäten zulässt. Daher gab es innerhalb der Forschungsgemeinschaft ein signifikantes Interesse an der Entwicklung von halb- und vollinduktiven Methoden, die Schlussfolgerungen über unbekannte und neue Entitäten ermöglichen. Existierende Ansätze berücksichtigen jedoch nur triple-basierte Wissensgraphen, während ihre reichhaltigeren Gegenstücke, hyperrelationale Wissensgraphen (z.B. Wikidata), nicht ausreichend untersucht wurden. Im Rahmen des Projekts wurde eine Kategorisierung verschiedener induktiver Verfahren aufgestellt und die Vorteile des Einsatzes von hyperrelationalen Wissensgraphen bei einer Vielzahl von halb- und vollinduktiven Linkvorhersageaufgaben, die durch die jüngsten Fortschritte bei Graph Neuronalen Netzen unterstützt werden, untersucht. Unsere Experimente (vgl. Table 23) auf einer neuen Menge an Benchmarkdatensätzen zeigen, dass Qualifizierer über typisierten Kanten zu Leistungsverbesserungen von 6 % absolutem Gewinn für die Hits@10-Metrik im Vergleich zu reinen Triple-basierten Verfahren führen. Für diese Arbeit [2] erhielten die Projektmitarbeiter den "Best Paper Award" auf der International Semantic Web Conference 2021.

Szenegraphen stellen relationales Wissen dar, das mit symbolischen Daten aus Texten oder Wissensgraphen modelliert werden kann. Das Trainieren von Klassifikationsmodellen von Szenegraphen erfordert eine große Menge an kommentierten Bilddaten. Während die Annotation von Bildern einen hohen Arbeitsaufwand erfordert, ist das Sammeln von textuellen Beschreibungen natürlicher Szenen weniger aufwändig. In [91] wird untersucht, ob textuelle Szenenbeschreibungen die annotierten Bilddaten ersetzen können. Zu diesem Zweck wurde eine neue Architektur zur Klassifikation von Szenegraphen entwickelt, die nicht nur anhand von kommentierten Bildern, sondern auch anhand von symbolischen Daten trainiert wird. In Abbildung 1 wird ein Überblick über die Architektur dargestellt. In dieser Architektur werden die symbolischen Entitäten zunächst auf ihre entsprechenden bildbasierten Repräsentationen abgebildet und dann in die relationale Reasoning-Pipeline eingespeist. Auch wenn eine strukturierte

Abbildung 1: Netzwerkarchitektur zu [91]



Form des Wissens, wie sie in Wissensgraphen vorkommt, nicht immer verfügbar ist, können wir diese aus unstrukturierten Texten mithilfe eines Transformer-basierten Sprachmodells generieren. Wir zeigen, dass wir durch Fine-Tuning der Klassifizierungspipeline mit dem aus Texten extrahierten Wissen, ca. 8x genauere Ergebnisse bei der Klassifizierung von Szenengraphen, ca. 3x genauere Ergebnisse bei der Objektklassifizierung und ca. 1,5x bessere Ergebnisse bei der Prädikatsklassifizierung erreichen können, obwohl das Verfahren verglichen mit Vergleichsverfahren aus dem Supervised Learning mit nur 1% der annotierten Bilder auskommt. Exemplarisch werden Vergleichsexperimente zur Vorhersage von Fakten in Tabelle 24 dargestellt.

Final wurde im Rahmen des Arbeitspakets ein neuer Algorithmus entwickelt, der visuelle Bildinformation lesen kann, um das extrahierte Wissen strukturell zu repräsentieren und abzuspeichern, um weitere Aufgaben zu unterstützen.

Das extrahierte Wissen sollte strukturell repräsentiert und gespeichert werden, um weitere Aufgaben zu unterstützen. Aus diesem Grund betrachten wir beide Aufgaben als eine und haben eine neue Methode zur Beantwortung von Fragen mit visuellem Bezug (visual question answering / VQA) entwickelt. Dabei sind die Fragen in natürlicher Sprache verfasst und beziehen sich auf den Inhalt eines Bildes. Im letzten Jahr haben wir die ersten Resultate eines Prototyps im Rahmen eines Kurzpapers beim ICML Workshop GRL+ publiziert [37]. VQA ist eine ambitionierte Aufgabe an der Schnittstelle von Computerlinguistik und Computer Vision. Die Darstellung der Bilder durch Szenen Graphen erlaubt es uns außerdem Methoden des maschinellen Lernens auf Wissensgraphen anzuwenden. Wir kombinieren moderne Methoden der Computerlinguistik und Computer Vision wie z.B. Transformer oder DetectoRS. Unser Szenen Graph Reasoning Modul basiert auf der Kantenvorhersagemethode MINERVA. Wir zeigen mithilfe von manuell erstellten Szenengraphen auf dem anspruchsvollen Datensatz GQA, dass unsere Methode in etwa menschliches Leistungsniveau erreicht, welches aktuell eine obere Schranke für das Leistungsvermögen maschineller Lernmethoden darstellt. Auf unseren eigenen, automatisch generierten Szenengraphen übertrifft unsere Methode andere Pfadmethoden die auf demselben Szenengraphen operieren.

AP 4: Bild und Videoanalyse mit KGs:

Zur Wiedererkennung von geometrischen Formen, beispielsweise zur Navigation in der Umgebung, werden sogenannte Korrespondenzen zwischen den geometrischen Formen berechnet. Für die Repräsentation der Geometrie kommen dabei Dreiecksgitter zum Einsatz, also ungerichtete Graphen mit im 3D-Raum eingebetteten Knoten. Bei der Berechnung von Korrespondenzen wird zwischen intrinsischen und extrinsischen Methoden unterschieden. Intrinsische Methoden verwenden die innere Struktur des

Graphen unabhängig von der 3D-Einbettung, wohingegen extrinsische Methoden die 3D-Einbettung verwenden. Unsere Arbeit “Smooth Shells: Multi-Scale Shape Registration with Functional Maps” [18] ist eine der ersten, die intrinsische und extrinsische Methoden kombiniert. Hierzu arbeiten wir im Produkt-raum der intrinsischen und extrinsischen Repräsentationen. Ferner definieren wir Repräsentationen geometrischer Formen, die die ursprüngliche geometrische Form auf mehreren Auflösungsstufen von grober zu feiner Auflösung approximieren. Zunächst werden für grobe Approximationen Korrespondenzen berechnet und feinere Details iterativ hinzugefügt. Durch Symmetrien bedingte Uneindeutigkeiten werden durch eine Markowketten-Monte-Carlo-Initialisierung aufgelöst. Das Ergebnis sind glattere und realistischere Korrespondenzen im Vergleich zu bisherigen Methoden. Unsere Arbeit wurde auf der CVPR-Tagung 2020 für einen Vortrag ausgewählt.

Auf dieser Methode aufbauend haben wir die Methode “Deep Shells” [20] entwickelt. Dabei haben wir den langsamen Initialisierungsschritt durch eine schnellere, gelernte Initialisierung ersetzt. Zu diesem Zweck haben wir systematisch eine differenzierbare hierarchische Matching-Pipeline aus entropieregularisiertem optimalen Transport hergeleitet. Dadurch konnten wir lokale Merkmalsextraktion basierend auf glatten abgeschnittenen spektralen Faltungskernen verwenden. Letztere basieren auf den Eigenfunktionen des Laplace-Beltrami-Operators auf der durch einen Graph (Dreiecksgitter) approximierten Objektoberfläche. Solche Faltungskerne sind relativ robust gegenüber verschiedenen Diskretisierungen der Oberfläche zu Graphen (Dreiecksgittern). Unsere unüberwachte Lernmethode liefert bessere Ergebnisse auf mehreren Datensätzen als bisherige Methoden, sogar überwachte Methoden. Ferner generalisiert unsere Methode gut auf Testdatensätze, die sich stark vom Trainingsdatensatz unterscheiden. Unsere Methode wurde als Paper auf der NeurIPS-Tagung 2020 vorgestellt.

In Zusammenarbeit mit Facebook AI Research haben wir als Nachfolgemethode die Methode “Neuro-Morph” [19] entwickelt, die besonders gut strukturelle Korrespondenzen erkennen kann. Unsere Methode kombiniert dabei Graph-Convolutions (basierend auf EdgeConv) mit aus Knoten des gesamten Graphen global gepoolten Merkmalen. Während der Trainingsphase verwenden wir Strafterme in der Kostenfunktion, die unnötige Deformationen der durch Graphen approximierten Objektoberflächen möglichst zu vermeiden versuchen und geodätische Distanzen auf diesen Oberflächen möglichst zu erhalten versuchen. Als Ergebnis bekommt man in nur einer feed-forward-Anwendung des neuronalen Netzes nicht nur Korrespondenzen zwischen Formen, sondern auch eine Interpolation zwischen Formen. Die Ergebnisse dieser unüberwacht trainierten Methode auf mehreren Benchmarks sind gleichauf und teilweise besser als die anderer Methoden, sogar im Vergleich zu den besten überwacht trainierten Methoden. Diese Arbeit haben wir auf der CVPR-Tagung 2021 vorgestellt. Ein Vergleich der Methoden ist in Abbildung 8 zu finden.

Bei der Erkennung von Objekten zur Navigation in der geometrischen Umgebung können die zu erkennenden Objekte auf verschiedenste Arten teilweise verdeckt sein. In unserer Arbeit “Implicit Shape Completion via Adversarial Shape Priors” [81] rekonstruiert ein neuronales Netz die fehlende Geometrie in Form einer vorzeichenbehafteten Distanzfunktion.

In unserer Arbeit “Neural Online Graph Exploration” [13] formulieren wir die Exploration von Graphen (zur Exploration der geometrischen Umgebung in Graphenrepräsentation) als bestärkendes Lernen. Wir verwenden Graph Convolutional Networks und Direct Future Prediction. Der Agent behält besuchte Knoten im Gedächtnis und wählt die nächsten Schritte basierend auf diesem Gedächtnis und der gelernten Strategie. Soweit wir wissen ist das der erste Ansatz für Online-Graphenexploration, der auf maschinellem Lernen basiert. Unser Agent lernt dabei Strategien, die besser sind als die von populären Algorithmen. In den Experimenten wenden wir die Methode unter anderem auf Straßennetze mit verschiedener Struktur an, nämlich die von München, Oxford und San Francisco. Eine vereinfachte Darstellung ist in Abbildung 2 zu finden.

Gemeinsam mit der Arbeitsgruppe von Prof. Hinrich Schütze (LMU) haben wir gezeigt [64], dass die automatische Erstellung von Szenengraphen (Graphen, die das Wissen über die Beziehungen zwischen den Objekten in einem Bild) helfen kann, bessere Textbeschreibungen von Bildern zu erzeugen. Wir verwenden tiefe neuronale Netze mit einem Aufmerksamkeitsmechanismus für Graphen. Die Abbildung von Bildern zu Szenengraphen und von Szenengraphen zu Text wird Ende-zu-Ende trainiert. Die explizite

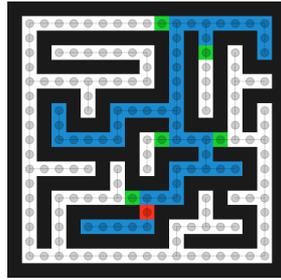


Abbildung 2: Ein Agent (rot), der Graphen exploriert, behält bereits besuchte Knoten (blau) sowie die Grenzknoten (grün), die als nächstes besucht werden können, im Gedächtnis. Die Grenzknoten stellen die Grenze zwischen unbekannter (weiß) und bereits besuchter Umgebung dar. Das Ziel ist, alle Knoten so schnell wie möglich zu entdecken und zu besuchen. Unsere Methode (“Neural Online Graph Exploration”, arXiv 2020) lernt dabei Strategien, die besser sind als die von populären Algorithmen.

Zwischenrepräsentation von Objektbeziehungen als Szenegraphen hilft dem neuronalen Netz, die Szene besser zu verstehen. Die Ergebnisse sind besser als die von Methoden, die Textbeschreibungen direkt aus Bildern generieren.

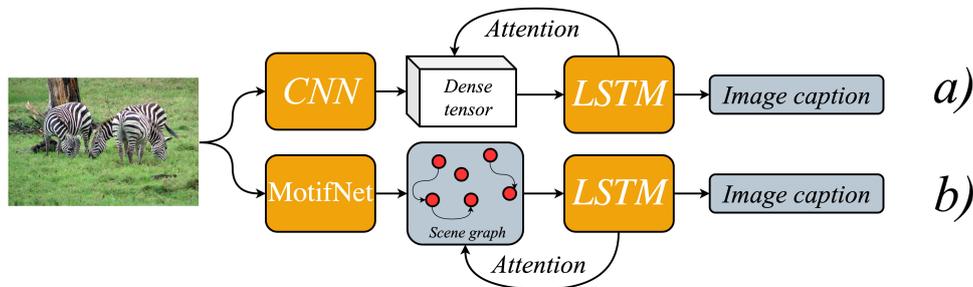


Abbildung 3: Vergleich der Datenprozessierung bei bisherigen Methoden zur Textannotation von Bildern (oben) und unserer auf Szenegraphen basierender Methode (unten) [64]. Unsere Methode erstellt einen Szenegraphen als Zwischenschritt, um die Beziehungen zwischen Objekten besser zu modellieren.

In Zusammenarbeit mit Sony Europe in Stuttgart haben wir eine Text-to-Speech-Methode publiziert [21]. Ein wichtiges Alleinstellungsmerkmal unserer Methode ist, dass sie Kontrolle über Faktoren der Sprache wie Tonhöhe, Rhythmus, Sprechlautstärke und Klangfarbe erlaubt (aber nicht erfordert). Ein zentraler Bestandteil unseres neuronalen Netzwerks ist dabei die sogenannte differenzierbare digitale Signalverarbeitung, ein neuer mit neuronalen Netzen kompatibler Ansatz, der bisher nur für Musik aber nicht für Sprachaudio verwendet wurde. Unsere Methode produziert realistische Sprache.

In Zusammenarbeit mit der Vanderbilt University haben wir das Paper “3D Deep Learning for Biological Function Prediction from Physical Fields” auf der 3DV-Tagung 2020 vorgestellt [29]. In dieser Arbeit sagen wir die biologische Funktion von Molekülen verschiedener Größe (kleinen potentiellen Pharmazeutika und großen Proteinen) voraus. Wir zeigen, dass dank Verwendung dichter physikalischer Felder (Elektronendichte und elektrostatisches Potential), die sich aus dem Graphen der Atombindungen ergeben, diesen umgeben und die intermolekularen Wechselwirkungen prägen, bessere Vorhersagen der Funktion der Moleküle mittels neuronaler Netze möglich sind. In einem weiteren Paper [28] mit der Vanderbilt University haben wir Molekülfunktion ohne diese Felder sondern stattdessen mit klassischen Deskriptoren der Atombindungsgraphen durchgeführt, aber dafür mit neu entwickelten ML-Kostenfunktionen, die die Besonderheiten der Moleküldatensätze besser berücksichtigen. Die Ergebnisse waren sind als die üblicher ML-Kostenfunktionen.

Ca. ab dem Jahr 2018 gab es in der Forschungscommunity einen Durchbruch in der Prozessierung der geometrischen Umgebung: es wurden neue (sogenannte rotationsäquivalente) Schichten neuronaler Netze hergeleitet, die Merkmale in geometrischen Daten mathematisch-garantiert gleich gut erkennen

unabhängig davon, in welcher Orientierung (also wie gedreht) die Merkmale im Computer gespeichert sind. Die Ergebnisse solcher Methoden sind besser, denn das neuronale Netz hat dann nicht die zusätzliche Schwierigkeit, die vielen möglichen Orientierungen jedes Merkmals einzeln zu lernen. Die besten dieser Schichten sind dabei auch trainierbar und im Sinne bisheriger neuronaler Netze möglichst allgemein (z.B. trainierbare affine Schichten, trainierbarer Aufmerksamkeits-Mechanismus, usw.). Da einige Gemeinsamkeiten und Unterschiede zwischen diesen Methoden und ihre Wichtigkeit nicht offensichtlich waren, haben wir diese Gemeinsamkeiten und Unterschiede sowie die Vor- und Nachteile der einzelnen Methoden systematisiert und samt Überblick zu den verwendeten mathematischen Konzepten in einem Survey-Paper [59] veröffentlicht.

Darauf aufbauend konnten wir diese rotationsäquivalenten neuronalen Netze auf weitere Arten von Daten verallgemeinern. Bisherige neuronale Netze bieten beispielsweise Äquivarianz unter zweidimensionalen Rotationen für zweidimensionale Daten, oder Äquivarianz unter dreidimensionalen Rotationen für dreidimensionale Daten. Dies haben wir verallgemeinert auf Äquivarianz unter dreidimensionalen Rotationen für sechsdimensionale Diffusions-Magnetresonanztomographie-Daten (Diffusions-MRT-Daten). Diffusions-MRT-Daten enthalten zu jedem Voxel im dreidimensionalen Ortsraum Signale im dreidimensionalen diffusionskodierenden Raum (Raum der Diffusionsrichtungen und -wichtungen). Diese Signale charakterisieren die Mikrostruktur z.B. von biologischem Gewebe, denn Moleküle diffundieren verschieden in verschiedene Richtungen und tasten dabei die Mikrostruktur ab, wodurch sich je nach Mikrostruktur verschiedene Diffusionsstatistiken ergeben. Die Äquivarianz unseres neuen neuronalen Netzes ist angebracht, denn Merkmale wie z.B. Nervenfasern können in verschiedenen Orientierungen vorliegen, und verschiedene Orientierungen äußern sich in Diffusions-MRT-Daten als Orientierungen im Orts- und Diffusionsraum. Unsere Methode liefert bessere Ergebnisse als bisherige Methoden. Wir haben sie auf der ISMRM-Tagung 2021 vorgestellt [66]. Eine stark erweiterte Version dieser Arbeit haben wir außerdem auf arXiv publiziert [65] und wir bereiten sie derzeit für eine Journal-Einreichung vor.

Gemeinsam mit UZ Brüssel und der Universität von Antwerpen haben wir unsere Rekonstruktionsmethode namens "Holistic Reconstruction" für Diffusions-MRT-Daten mit einer verbesserten Datenakquisition basierend auf Fast Spin Echo kombiniert. Dadurch ist schnellere Datenakquisition mit besserer Bildqualität und weniger Deformationsartefakten möglich. Erste Versuche wurden bei einem gemeinsamen Journal-Paper [68] unternommen. Ein ausführlicher Vergleich verschiedener Sampling-Strategien wurde auf der ISMRM-Tagung 2021 vorgestellt [67].

Unsere Analyse und Verbesserung von Informationsflussgraphen von ML-Methoden hat zu vielversprechenden ersten Ergebnissen geführt. Beispielsweise im Bereich Domain Adaptation haben unsere Methoden mit neuen Informationsflussgraphen bessere Ergebnisse erzielt als bisherige Methoden. Publikationen sind in Planung.

Ferner wurde ein Ansatz zur Extraktion eines Szenendiagramms entwickelt, das Tiefeninformationen mittels eines unabhängigen CNNs vorhersagt und die so gewonnenen 3D-Informationen zur Extraktion von Relationen verwendet. Die Ergebnisse dieses Ansatzes wurden in einem Preprint [26] der Öffentlichkeit zugänglich gemacht und später auf der International Conference on Pattern Recognition veröffentlicht [90]. Eine weitere Methode zur Extraktion von Szenegraphen aus Bilddaten wurde im Rahmen des Projekts entwickelt und auf dem ICML Workshop "Bridge Between Perception and Reasoning: Graph Neural Networks & Beyond" veröffentlicht [5]. Außerdem wurde das Verfahren Graphhopper [50] entwickelt. Es ermöglicht visuelle Fragebeantwortung auf dem neuesten Stand der Technik. Es basiert auf einem neuem Reinforcement-Learning-Verfahren, das interpretierbare Szenegraphen exploriert.

Wir haben Ansätze zur Wissensextraktion auf Bildern auf die Anwendung auf Videos erweitert. Hierbei zeigte sich allerdings, dass eine einfache Erweiterung nicht ausreichend ist, um der Dynamik der Objektbeziehungen gerecht zu werden. Außerdem wurde ein Videosegmentierungsverfahren entwickelt, das mit Box- und Bewegungsannotationen arbeitet [33].

Ferner werden Methoden vorgestellt, wie sich diese Szenegraphen für visuelle Anfragebearbeitung einsetzen lassen. Es wurde eine Preprint zu diesem Thema veröffentlicht [36]. Die "Relation Transformer"-Architektur wurde entwickelt, die sowohl den lokalen also auch den globalen Kontext bei der Erstellung von Szenegraphen berücksichtigt. Ein Preprint, der dieses Ergebnis beschreibt wurde auf Arxiv.org

veröffentlicht [52].

Wir haben eine Methode namens Relationformer zur Generierung von Graphen unter simultaner Interaktion von Entitäten, Prädikaten und ihrer wechselseitigen Hierarchie entwickelt [94].

An der Schnittstelle zwischen der Extraktion von strukturierten Informationen aus Bildern und Videos (Arbeitspaket 4) und der Generierung von Texten (wie sie im Arbeitspaket 3 entwickelt werden) wurden die Ergebnisse beider Arbeitspakete kombiniert, um textuelle Beschreibungen aus Bildern zu entwickeln. OODformer [51] wurde entwickelt und ist ein objektkontextbasiertes Verfahren zur Erkennung von Ausreißern.

AP 5: Perspektiven zu AI und Kognitionsforschung

Perzeption und Memory sind zwei grundlegende Funktionalitäten eines AI Systems. In diesem AP untersuchen wir Bezüge zur AI aber auch zur Kognitionsforschung. Dies entspricht dem internationalen Trend, AI und Kognition in einem engen Zusammenhang zu betrachten [34]. Besonders wichtig ist zu analysieren, inwieweit die Ergebnisse aus AP3 und AP4 die Grundlage für Entscheidungsunterstützungssysteme bilden können. Wir können hier auf eigene Vorarbeiten aufbauen [99] wollen aber auch internationale Entwicklungen integrieren. Die Arbeiten bauen speziell auf AP3 und AP4 auf. Die Ergebnisse bilden eine Grundlage zur Entscheidungsmodellierung in AP6.

In vergangenen Veröffentlichungen ist gezeigt worden, dass auf der Basis einer Tensor Approximation eines KG ein semantisches Gedächtnis realisiert werden kann, analog zum semantischen Gedächtnis des menschlichen Gehirns. Das semantische Gedächtnis betrifft Fakten, die wir wissen. Hier wollen wir dieses Konzept weiterentwickeln. Zentral ist hier die genaue Repräsentation der Gedächtnisspuren (Memes), über latente Zustände der Entitäten und Relationen. Ein Fokus sind Ansätze, die auf der Idee randomisierter Repräsentationen beruhen. Eine wichtige Fragestellung betrifft die Unterscheidung einer reinen Gedächtnisfunktion zu einer Gedächtnisfunktion, die es erlaubt, über gespeicherte Informationen hinweg zu generalisieren. Ein weiterer besonderer Schwerpunkt betrifft die Repräsentation von Entitäten, von Klassen, sowie Klassenhierarchien. Ein episodisches Gedächtnis über Fakten, an die wir uns erinnern, kann über vierdimensionale Tensoren modelliert werden. Ein aktives Forschungsgebiet, mit dem wir uns beschäftigen wollen, ist der Bezug zwischen semantischem und episodischem Gedächtnis. Eine Theorie ist, dass das semantische Gedächtnis durch eine zeitliche Integration des episodischen Gedächtnis realisiert werden kann. Hier stellt sich ebenfalls die Frage, inwieweit ein reines Abspeichern genügt, oder ob über episodische Erinnerungen ebenfalls eine Generalisierung stattfinden muss.

Wir haben verschiedene Modelle zur Modellierung und Schlussfolgerung episodischer Wissensgraphen vorgeschlagen und untersucht. Episodische Wissensgraphen, auch als zeitliche oder zeitabhängige Wissensgraphen bezeichnet, sind umfangreiche Ereignisdatenbanken, die dynamische und multirelationale Daten beschreiben können. Beispielsweise sind die zwei ereignisbasierte Wissensgraphen: das Global Database of Events, Language, and Tone (GDELT) und das Integrated Crisis Early Warning System (ICEWS), die Aufmerksamkeit erregt haben. Wie der Name sagt, die GDELT Datenbank besteht aus sich entwickelnden Wissen über die Interaktionen zwischen Ländern und Organisationen auf der ganzen Welt. Die ICEWS Datenbank enthält Informationen von nationalen und internationalen Krisen. Inferenz mit zeitlichen Wissensgraphen ermöglicht es, vorherzusagen, welche Ereignisse in der Zukunft eintreten werden und insbesondere, wenn diese Ereignisse eintreten werden. [32] ist die erste Forschungsarbeit, die den Hawkes Prozess verwendet (siehe Abbildung 4), um die zugrunde liegende Dynamik von Entitäts- und Ereignisinteraktionen in episodischen Wissensgraphen zu erfassen. Das GHNN Modell besteht aus zwei Modulen. Ein Graph Neural Network (GNN) Modul aggregiert Nachbarschaftsinformationen von Ereignissen, die zum gleichen Zeitstempel eingetreten sind.

Außerdem wird ein zeitkontinuierliches LSTM Modul eingesetzt, um die Intensitätsfunktion im Hawkes Prozess zur Modellierung des Auftretens eines Ereignistyps abzuschätzen. Wir testen das GHNN Modell an zwei episodischen Wissensgraphen, GDELT und ICEWS. Wobei zeigen die experimentellen Ergebnisse, dass das vorgeschlagene GHNN Modell die anderen Methoden in Entitäts- und Zeitvorhersagen übertrifft.

Ein episodisches Gedächtnis kann als Trainingsmaterial dienen, um Fähigkeiten zu lernen und zu

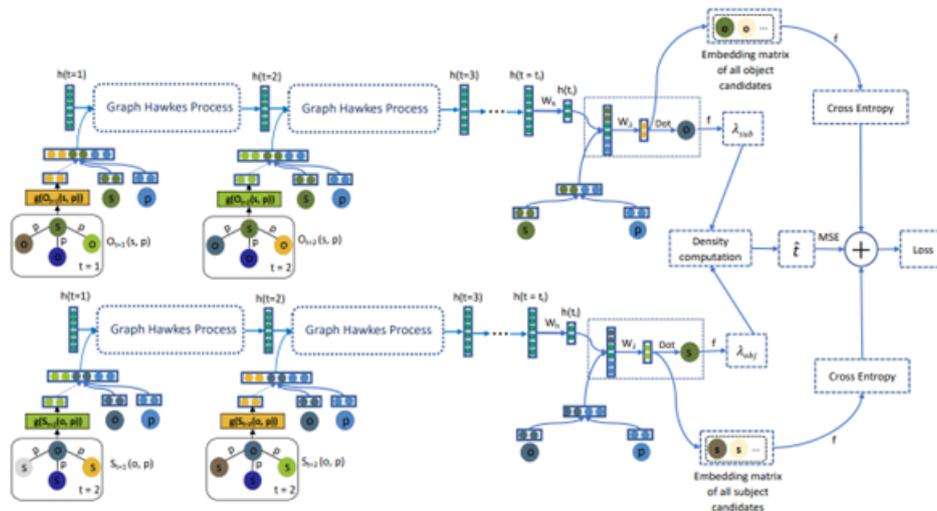


Abbildung 4: Architektur des entwickelten Graph-Hawkes-Neural-Networks

verbessern. Einige Theorien zum episodischen Gedächtnis sehen diese Funktionalität als zentral an, im Gegensatz zum Transfer in ein semantisches Gedächtnis.

Moderne Wissensgraphen bestehen oft aus großer Anzahl an Fakten und diese wächst rapid während konventionelle ML Modelle nicht immer skalieren kann. Unser Beitrag ist der erste Quantum-Algorithmus fürs Modellieren von großen Wissensgraphen-Daten in Form von Tensoren. In [61] wird ein Quantum-Algorithmen entwickelt, um das Training und Inferenz in Wissensgraphen zu beschleunigen, unter Anwendung intrinsisch paralleler Berechnungsfähigkeit von Quantum-Computation.

In [62] Wir benutzen intrinsisch parallele Berechnungsfähigkeit von Quantum-Computation, um das Training und Inferenz in Wissensgraphen zu beschleunigen.

Alternative Gedächtnismodelle sind in jüngerer Zeit entwickelt worden, wie die Memory Models von DeepMind. Hier wollen wir die unterschiedliche Gedächtnismodelle vergleichen. Die wichtigste biologische Rolle von einem Gedächtnissystemen ist dessen Unterstützung in der Verbesserung von Entscheidungen. Hier untersuchen wir, wie gedächtnisgesteuerte Modelle zu besseren Entscheidungen führen können. Insbesondere werden wir die Verbindung zur Entscheidungsunterstützung im klinischen Anwendungsszenario untersuchen. Wir werden weiter untersuchen, inwieweit ein episodisches Gedächtnisnetzwerk das Reinforcement Lernen verbessern kann. Genauer gesagt, wie kann ein episodisches Gedächtnisnetzwerk die Schätzung von Value Funktion und die Strategie des Agenten verbessern? Auf diese Weise könnten wir die Wechselwirkungen zwischen den Belohnungszentren und dem Hippocampus im Gehirn besser verstehen.

Wissensgraphen haben zahlreiche Anwendungen in verschiedenen Teilgebieten der künstlichen Intelligenz wie z.B. das Auflösen von Entitäten in der Computerlinguistik, visuelle Relation Erkennung beim Bildverstehen, oder beim kollaborativen Filtern. Dabei können viele Aufgabenstellungen als Vorhersage von neuen Kanten (also Fakten) formuliert werden. Den dominanten Ansatz bilden Methoden, die eine Einbettung von Entitäten und Relationen in Vektorräumen vornehmen und auf diesen Einbettungen Klassifikatoren trainieren. Während sich die Methoden sowohl in der Art der Einbettung und als auch der Klassifikatoren unterscheiden, haben sie meistens gemein, dass sie am Ende einen Klassifikationsscore für einen gegebenen Input-Fakt berechnen. Dabei ist für den Benutzer meistens nicht ersichtlich, welche Fakten und Strukturen zu dem Score beigetragen haben. Daher ergibt sich ein sogenanntes Black Box Problem, insbesondere in sensiblen Anwendungsfeldern (z.B. dem medizinischen Bereich).

Wir haben die neue Methode R2D2 (Reveal Relations using Debate Dynamics) [38] zur Klassifikation von Fakten in Wissensgraphen entwickelt, um das Black Box Problem anzugehen. Die zugrunde liegende Idee ist, dass die Klassifikation von unbekanntem Fakten als Debatte zwischen zwei konkurrierenden Agenten

modelliert wird. Dabei hat ein Agent die Aufgabe, Argumente (Pfade auf dem Wissensgraphen) zu extrahieren, die als Evidenz für den Wahrheitsgehalt des Faktes sprechen sollen. Der andere Agent soll Argumente für die gegenteilige Position finden. Am Ende der Debatte (nachdem alle Argumente extrahiert wurden) werden die Argumente von einem weiteren Akteur, dem sogenannten Richter, verarbeitet. Dieser trifft auf der Basis der Argumente eine Klassifikationsentscheidung. Dieses Vorgehen hat den Vorteil, dass der Benutzer sich die Argumente anschauen und die Entscheidung des Richters nachvollziehen kann. Unsere Experimente auf Benchmark Datensätzen zeigen, dass R2D2 vergleichbare oder bessere Ergebnisse als gängige Black Box Methoden erzielt, dabei aber zusätzliche Erklärungen liefert. Die Erklärbarkeit wurde anhand einer Benutzerumfrage mit positivem Ergebnis untersucht.

Primitive Lebewesen haben ein einfaches Neuronales Netzwerk, welches Sensorik über Zwischenneuronen mit Aktuatorneuronen verknüpft. Ein Episodisches Gedächtnis erlaubt den Bezug zu besonderen vergangenen Situationen von Relevanz. Ein Semantisches Gedächtnis liefert weitere Generalisierungsmöglichkeiten und ermöglicht das Erlernen in unerfahrenen Situationen. Hier untersuchen wir, wie sich das einfache Neuronale System der Amphibien zu kompliziertes Neuronales System der Primaten entwickelte, um sich an immer komplizierte Umgebungen anzupassen. Darüber hinaus studieren wir, wie semantisches Gedächtnis, das vom Neocortex unterstützt ist, aus vom Hippocampus unterstützten episodischen Gedächtnis sich entwickeln konnte.

Der allmähliche Übergang vom episodischen Gedächtnis zum semantischen Gedächtnis ermöglicht es dem Agenten, den Mechanismus hinter seiner Umgebung wirklich zu nachzuvollziehen. Es erlaubt dem Agenten, sich sehr schnell an neue Situationen anzupassen.

Die aktuelle automatisierte ICD-9 Klassifizierungsaufgabe wird von RNN- und CNN- basierende Modelle dominiert. In dieser Arbeit konzentrieren wir uns auf die Transformer-basierende Modelle und investigieren wir die Perspektive von Explainability von den Vorhersagen.

Die Technik, strukturierte Wissen automatisch aus Textdaten ableiten zu können, ist die Grundlage eines Gedächtnissystems. Für diesen Zweck haben wir in [23] innovative Lösungen entwickelt, und diese anhand einer Klassifizierungsaufgabe des ICD-9 Code evaluiert. Wir haben an ICD-9 Code Klassifizierung mit der Anwendung von Transformer-basierende Modelle gearbeitet.

Für den ICD-9 Klassifizierungsaufgaben möchten wir die Performanz von neue Transformer-basierende Modelle mit anderen state-of-the-art RNN- und CNN- basierende Modelle vergleichen. Darüber hinaus haben wir eine Methode vorgeschlagen, um die Explainability von den Transformer-basierende Modelle zu verbessern. Wir haben verschiedene Transformer-basierende Modelle auf die ICD-9 Klassifizierungsaufgaben angewendet, z.B., BERT, H-BERT, Longformer. Darüber hinaus haben wir einen description-based label attention classifier vorgeschlagen, um die Explainability für einzelne Vorhersage zu verbessern.

AP 6: Klinische Lernende Entscheidungsunterstützung

Das Institut für Frauengesundheit (IFG) GmbH ist eine wissenschaftlich ausgerichtete Clinical Research Organisation, welche klinische Studien und molekulare Registerstudien durchführt. Des Weiteren werden Datenmanagement und Verwaltung von Bioproben für verschiedenste Studien nationaler Studiengruppen übernommen. Ein besonderer Fokus hierbei sind das Monitoring und die Kuration von Daten. Daten-Kuration ist ein aufwendiger Prozess in klinischen Studien. Im Rahmen des MLwin Projektes wurden von der IFG GmbH Methoden erarbeitet, um diese Kuration von Daten effizienter zu gestalten. Neben der Eignung einzelner Algorithmen für Datenanalyse und maschinelles Lernen spielt insbesondere die Qualität der verwendeten Informationen im Analyse- bzw. Lernprozess eine entscheidende Rolle. Daher wurde zunächst ein flexibles System zur Evaluierung der Dokumentationsgüte etabliert. Dieses ermittelt anhand frei definierbarer Abfragen unterschiedlicher Granularität (pro Datensatz, pro Formulareintrag) einzelne Marker zu Vollständigkeit und Korrektheit der Dokumentation. Erkannte Defekte werden im Dokumentationssystem zur Bearbeitung durch medizinische Dokumentare und das Monitoring markiert sowie deren Behebung nachvollzogen und in einem tabellarischen Bericht aufbereitet dargestellt. Durch eine Kategorisierung der einzelnen Defekte (z.B. Vollständigkeit und/oder Korrektheit von Stammdaten, Tumordokumentation, Medikationsdokumentation) wird eine abgestufte Einschätzung der Dokumentati-

onsqualität ermöglicht, um gezielt einzelne Datensätze für eine spezifische Auswertung auswählen zu können. Die Grundlage hierfür bildet ein konzeptuelles, logisches (Daten-)modell, welches Patienten, deren Erkrankung, Diagnosen und Therapien unabhängig von einer konkreten Datenquelle zu erfassen erlaubt. Neben der technischen Realisierung im Rahmen von Datenbanktabellen, wird dieses Modell zudem als Kommunikations- und Schulungsgrundlage verwendet, um alle beteiligten Stakeholder über Entitäten, deren Beziehungen zueinander sowie Variablen und deren Wertebereiche in Kenntnis zu setzen. Eine vereinfachte Darstellung dieses Modells mit ausgewählten Biomarkern und Tumorentitäten ist in Abbildung 5 dargestellt. Ergänzt wird das Modell durch ein Data Dictionary, dessen Erstellung im direkten, technischen Entwicklungsprozess von Datenintegration und -extraktion für Analysezwecke integriert ist.

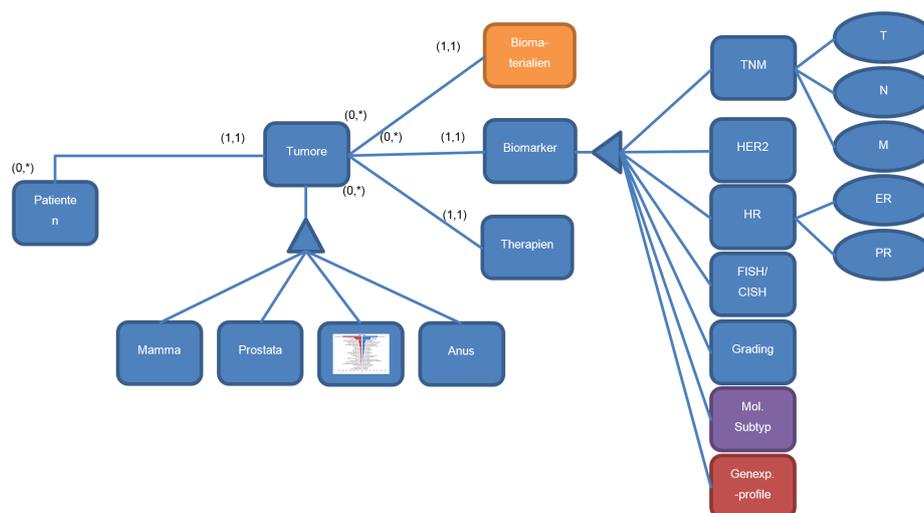


Abbildung 5: Vereinfachte Darstellung des konzeptuellen, logischen Datenmodells zur Etablierung eines gemeinsamen Verständnisses über Datensätze, Variablen und Beziehungen

Um Daten für Auswertungen und weitere Verarbeitungen verfügbar zu machen ist die Zusammenführung verschiedener Datensätze (z.B. klinische und genomische Daten) und die initiale Strukturierung im Rahmen der Daten-Kuration essentiell. Hierbei stellen genetische Daten, die über Freitextfelder dokumentiert werden und eine große Vielfalt an Variablen einnehmen können, eine große Schwierigkeit dar. Im Rahmen des MLwin Projektes wurde gezielt an der Aufarbeitung dieser Datentypen sowie der Harmonisierung verschiedener Datenquellen, gearbeitet. Die Daten der PRAEGNANT-Studie wurden in einem ORACLE-basierten Datenferneingabesystem (eCRF; secutrial®, IAS GmbH, Berlin) dokumentiert. Hierbei wurden unter anderem genetische Daten aus den klinischen Routineuntersuchungen erfasst. Aufgrund der hohen Anzahl unterschiedlichster Mutationsvarianten erfolgte die Dokumentation manuell in Freitextfeldern. Allerdings ist diese Art der Dokumentation mit einer hohen Fehlerrate behaftet. Für die Aufarbeitung und Zusammenführung notwendiger Daten wurde deshalb von der IFG GmbH ein halbautomatischer Extract, Transform, Load- (ETL-) Prozess etabliert, der die wissenschaftliche Verwendung und Auswertung der genetischen Daten ermöglicht (Abbildung 6). Das UKER hat hierbei die biometrische Einarbeitung der genomischen Daten übernommen. Hierdurch war es nun möglich exakte genetische Veränderungen zu extrahieren und diese mit externen Quellen (z.B. ClinVar) zu verknüpfen und so die Mutationen zu kategorisieren und für wissenschaftliche Fragestellungen heranzuziehen. Dieser Prozess ermöglichte die Extraktion von Mutationsvarianten mit einer Präzision von 99,8% und einem Recall von 95%. [75]

Darüber hinaus wurde im Rahmen des Projektes Studiendaten auch den anderen Projektpartnern zur Verfügung gestellt. Hierfür wurden die Ansprüche der Partner erarbeitet, sowie die ethischen und

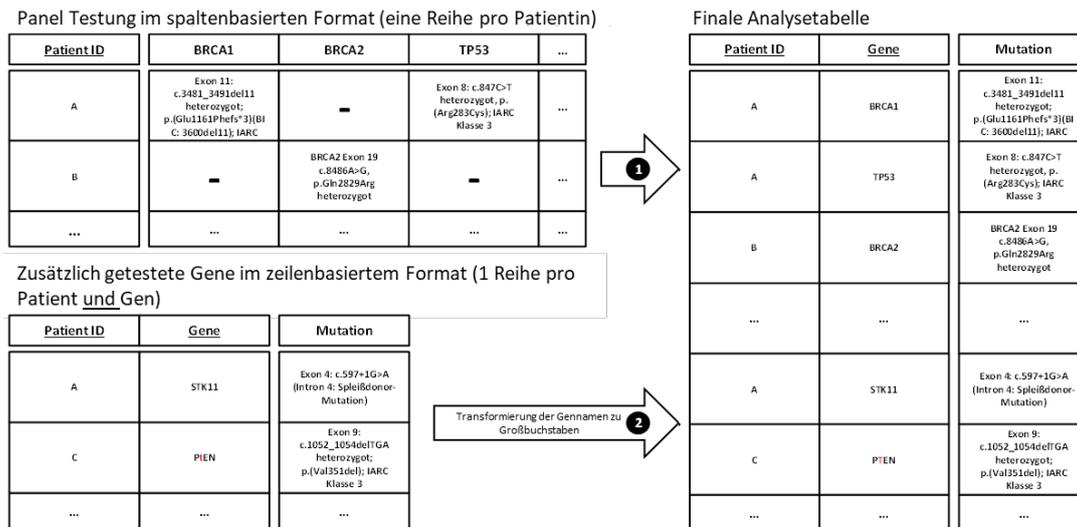


Abbildung 6: Schema des ETL-Prozesses zum Aufbau einer speziell für den Matching-Prozess angepassten Repräsentation der eCRF-Daten: (1) Transformation von spaltenbasierten Mutationsinformationen der wichtigsten Gene in ein zeilenbasiertes Format und (2) Hinzufügen von Nicht-Standard-Genmutationen angeben in zeilenbasiertem Format mit Umwandlung von Gennamen in Großbuchstaben.

datenschutzrechtlichen Belange evaluiert. Zur Erarbeitung einer Analysedatei auf Basis umfangreicher SecuTrial-basierter Datentabellen wurden SQL-Skripte entwickelt, Algorithmen implementiert und die Daten transformiert, um sie in aufgearbeiteter Form den Projektpartnern zur Verfügung zu stellen. Zur Aufarbeitung und Zusammenführung notwendiger Daten wurde ein standardisierter ETL-Prozess etabliert, der die wissenschaftliche Verwendung aller notwendigen klinischen Daten aus der PRAEGNANT Registerstudie ermöglicht. Dieser ETL-Prozess kann auch für zukünftige wissenschaftliche Projekte angewandt und entsprechend der Fragestellungen individuell angepasst werden. Die Prozesse der aktiven, laufenden Datenverwaltung wurden so angepasst, dass das Abrufen der Daten für Forschungszwecke unter Einbezug maximalen Qualitätserhalts und Nutzungswertes und gleichzeitiger Einsparung finanzieller und zeitlicher Kapazitäten durchgeführt werden kann. Hierbei kann jederzeit eine individuelle Datenbereitstellung angepasst auf die wissenschaftliche Fragestellung und unabhängig vom Exportzeitpunkt durchgeführt werden. Es wurden insgesamt drei Datenübergaben im Rahmen von AP6 des Projektes durchgeführt.

Für diese Datenexporte wurde vom Projektpartner UKER die klinisch relevanten und für die Fragestellung notwendigen Datenquellen identifiziert und zugehörige Data Dictionaries erarbeitet. Die klinischen Datensätze wurde im Detail mit dem Kooperationspartner Siemens besprochen sowie die Projektziele für die Analyse mittels maschinellem Lernen erarbeitet. Hierbei ging es vor allem um die klinisch-wissenschaftliche Expertise wie der Datensatz zu behandeln ist und welche Fragestellungen primär wichtig sind für die weiteren Analysen um einen fundierten Erkenntnisgewinn durch das maschinelle Lernen aus den Daten zu generieren.

Das Siemens Team hat an den zur Verfügung gestellten Datensätzen quantitative Analyse sowie Modellierungen durchgeführt.

Zunächst wurde ein Pipeline zur Daten-Vorverarbeitung, Visualisierung und Benchmarking entwickelt. In den Daten sind insgesamt 12 Tabellen mit Informationen über 1801 Patienten mit metastasiertem Brustkrebs. Die Daten wurden gesammelt zwischen 1976 und 2019 und enthalten 83 numerische, kategorische und textuelle Attribute. 46 Attribute davon sind von statischer Natur (wie z.B. demografische Information), während die restlichen haben Zeitstempel und stellen klinische Ereignisse dar. Diese Daten ermöglichen unsere Modellierung der Therapie-Entscheidung und Endpoints. Zu den Hauptaufgaben gehören u.a. kategorischen Attributen zu kodieren, fehlende Daten zu imputieren, klinische Ereignisse zeitlich zu sortieren, sowie Text-Attribute kodieren mit state-of-the-art NLP Technologien wie BERT,

BioBert zu kodieren.

Man möchte evaluieren, wie gut unterschiedliche prädiktive Methoden aus maschinellem Lernen die klinische Entscheidungsprozesse modellieren können. Dabei spielen mehrere Faktoren wie Genauigkeit der Prädiktionen, Erklärbarkeit, Robustheit und Unsicherheiten der Modelle eine Rolle. Ein ähnlicher Datensatz wie unter Punkt 2.1.2.3 wurde benutzt, um die Modelle zu bauen, trainieren und evaluieren. Im Laufe des Projekts wurde gewisse Aktualisierungen daran durchgeführt. Die Datenstruktur bleibt unverändert, aber neue Datenpunkte, also die Patienten wurden inkludiert sowie alte wurden entfernt. Daher sind die Ergebnisse nicht direkt vergleichbar.

In Tabelle 9 sind die Test-Ergebnisse zusammengefasst. Die ersten 3 Spalten beziehen sich auf Experimente, die nur mit den statischen Features durchgeführt wurden. Die letzten 3 Spalten werden Ergebnisse berichtet, welche auch die zeitlichen Features berücksichtigen. Es ist keine einzige Methode zu identifizieren, welche alle anderen an Prädiktionsqualitäten übertreffen kann. Jedoch zeigt der Gaussian Process systematisch bessere Ergebnisse.

Die vielversprechenden Performanzen der GP Modelle führten zu weiteren Forschungsarbeiten, wo wir genauer analysiert haben, i) inwiefern lassen sich GP Modelle auf höhere Datenvolumne und Dimensionalitäten skalieren und ii) inwiefern kann man die Unsicherheitsquantifizierung des GPs verifizieren. Da die UKER Datensätze von relativer einfacher Natur und kleinerem Umfang sind, haben wir unsere Experimente an öffentlich zugänglichen großen und unstrukturierten Daten durchgeführt.

Spezifisch interessieren wir uns für Vorhersagemethoden, die in der Lage sind, erklärbarer Unsicherheit der Vorhersage zu generieren. In sicherheitskritischen Situationen wie therapeutische Entscheidung kann eine hohe Unsicherheit zur menschlichen Intervention führen und mögliche Fehlentscheidung verhindern. Wir leiten unsere Unsicherheit der Vorhersagen im Grunde genommen aus Ähnlichkeiten zwischen den Patienten ab. Mit anderen Worten, unsere Methode entdeckt seltene Patientenfälle und übersetzt diese „Seltenheit“ zu Unsicherheit auf eine skalierbare und mathematisch solide Weise. Die vorgeschlagene Methode gehörte zum sogenannten Deep Kernel Learning. D.h., wir nutzen moderne Deep Neuralnetzwerk als ein Feature Extractor, um die wichtigen Repräsentationen für alle Probe zu lernen. Mit den gelernten Repräsentationen wenden wir die moderne sparse GPs an, z.B., Sparse Variational Gaussian Process und Parametric Predictive Gaussian Process.

Spezifisch haben wir neue Vorhersagemethoden entwickelt, die in der Lage sind, erklärbarer Unsicherheit der Vorhersage zu generieren. In sicherheitskritischen Situationen wie therapeutische Entscheidung kann eine hohe Unsicherheit zur menschlichen Intervention führen und mögliche Fehlentscheidung verhindern. Wir leiten unsere Unsicherheit der Vorhersagen im Grunde genommen aus Ähnlichkeiten zwischen den Patienten ab. Mit anderen Worten, unsere Methode entdeckt seltene Patientenfälle und übersetzt diese „Seltenheit“ zu Unsicherheit auf eine skalierbare und mathematisch solide Weise. Die vorgeschlagene Methoden [105, 104] gehören zum sogenannten Deep Kernel Learning. D.h., wir nutzen moderne Deep Neuralnetzwerk als ein Feature Extractor, um die wichtigen Repräsentationen für alle Probe zu lernen. Entwickelt und getestet werden unter anderem Convolutional und Recurrent Neural Networks. Mit den gelernten Repräsentationen wenden wir die moderne sparse GPs an, z.B., Sparse Variational Gaussian Process und Parametric Predictive Gaussian Process.

Die Unsicherheitsanalyse stellt eine wichtige Komponente eines vertrauenswürdigen klinischen Entscheidungsunterstützungssystems dar. Denn die Unsicherheit, welche mathematisch verifizierbar ist, könnte als Konfidenz der Vorhersage interpretiert werden. Der menschliche Arzt ist dadurch in der Lage, sich mehr auf Patienten zu konzentrieren, für welche unsichere Vorhersagen produziert werden.

Wir haben unsere Methoden anhand zwei klinischer X-Ray Bild-Datensätze und eines EHR-Datensatzes. Die Aufgaben bei den Bild-Datensätzen bestehen darin, das Alter von den Patienten bzw. die Koordinaten der Läsion vorherzusagen. Anhand der EHR-Daten wollen wir die Länge des Aufenthalts der Patienten in Intensivstation zu modellieren. Wir evaluieren die Präzision unserer Vorhersagen MAD und RMSE, und die Plausibilität der Unsicherheit mit einer neuen Methode. In Abbildungen 10 und 11 findet man jeweils die Test-Ergebnisse der Modellierung sowie Verifikationen der quantifizierten Unsicherheiten, jeweils anhand der Bone-Age und Lesion Localization Datensätze.

Für die EHR Daten haben wir die Überlebensdauer von Patienten modelliert und unsere Ergebnisse mit klassischen Methoden wie Cox Regression und AFT Modellen verglichen. Die Ergebnisse findet man in Abbildung

Diese Ergebnisse zeigen zum einen, dass unsere neuen Methoden, welche die Stärke von Unsicherheitsquantifizierung des GPs und die Stärke der Repräsentationslernens der Neuronalen Netze kombinieren, state-of-the-art Ergebnisse in der Vorhersagequalität erzielen. Zum anderen haben wir gezeigt, dass die Unsicherheitsquantifizierung tatsächlich mit den Vorhersagenqualitäten korreliert sind.

Darüber hinaus haben wir auch generative Neuronale Netze entwickelt, um fehlende Daten in EHR Daten zu imputieren. Das Essentielle an dem Algorithmus besteht darin, Generative Adversarial Neural Network so zu modifizieren, damit diese auch kategorische Features welche diskret kodiert sind, auch generieren können.

Um maschinelles Lernen auch in der Patientenversorgung aktiv einzusetzen und nutzbar zu machen, hat das UKER zusammen mit den Projektpartnern LMU und IFG GmbH eine klinische Fragestellung erarbeitet, die einen Patienten-nahen kurzfristigen Benefit bringen kann. Im Rahmen dieses klinischen Szenarios sollen Audioaufnahmen und Transkripte der Aufnahmen von Arzt-Patienten Gesprächen genutzt werden, um mittels maschinellem Lernen A) die Tumorboard- und Therapieentscheidung auf Basis dieser Kommunikation zu verbessern, B) das Gespräch durch Indikation von Informationslücken zu optimieren und C) eine automatisch gesteuerte Dokumentation der Gesprächsergebnisse in klinischen Dokumentationssystemen oder Studiendatenbanken zu ermöglichen. Für die Durchführung dieser Untersuchungen wurde ein Studienprotokoll (COMMITMENT Studie: *Communication Improvement with Machine Learning methods – A study to improve patient-physician communication in standard healthcare situations*) entwickelt und das Votum von der zuständigen Ethikkommission eingeholt. Im Rahmen der COMMITMENT Studie wurde die Kommunikationsführung so optimiert, dass eine standardisierte Auswertung mittels maschinellem Lernen ermöglicht wird. Insgesamt konnten 80 Arzt-Patienten Gespräche aufgezeichnet und entsprechende Fragebögen zur Usability, Feasibility und Akzeptanz vom Einsatz Technologien des maschinellen Lernens in der Arzt-Patienten Kommunikation ausgewertet werden (Abbildung 13). Eine Publikation zur Feasibility und Akzeptanz ist aktuell in Arbeit [97].

AP 7: Community & Benchmarking

Die FHG entwickelte ein Modell zur Einbettung verteilter Wissensgraphen. Dieses Modell basiert auf der MDE, die für verteilte Systeme angepasst wurde und auf Apache Ignite für die gemeinsame Nutzung von Speicher während der Parallelverarbeitung beruht.

FHG hat das entwickelte Modell mit PyTorch-Biggraph [56] verglichen und deshalb die parallelisierte Trainingsmethode HOGWILD [71] integriert, die sich dort als effizient erwiesen hat. FHG hat das verteilte Wissensgraphen-Einbettungsmodell in das SANSA-Stack-Framework (²) integriert. Der Quellcode ist in <https://github.com/SANSA-Stack/SANSA-Stack/tree/dev-ML-Ignite-KGE/sansa-ml> frei verfügbar.

Auf der Grundlage der FAIR-Prinzipien wird die HOBBIT-Bewertungsplattform erweitert und online zur Verfügung gestellt. Die Plattform ermöglicht ein umfassendes Benchmarking von KGE-Modellen auf der Grundlage der Metriken "Hits@10" und "Mean Reciprocal Rank".

Ein Tutorial wurde erstellt, um die Nützlichkeit der Plattform und ihre Anwendung auf Methoden zur Einbettung von Wissensgraphen zu demonstrieren ³. Diese auf der HOBBIT-Plattform basierenden Einbettungsbenchmarks wurden auf der Konferenz SEMANTiCS 2021 vorgestellt [80].

Darüber hinaus wird ein umfangreicher Satz von Datensätzen entwickelt, um Einbettungsmodelle auf der Grundlage individueller Beziehungsmuster zu bewerten. Diese Datensätze umfassen 4 verschiedene Beziehungsmuster und beinhalten sowohl induktive als auch transduktive Einstellungen. Dieser Datensatz wurde zum Benchmarking modernster Linkprädiktoren verwendet. Die daraus resultierende Arbeit wird auf der NeurIPS 2021 Konferenz veröffentlicht [79].

In dieser Studie haben wir eine Teilmenge von Daten aus dem Standard FB15K und WN18 mit Symmetrie-

²<https://sansa-stack.net/>

³<https://github.com/mlwin-de/BenchEmbedd>

, Invers-, Antisymmetrie- und Inferenzmustern extrahiert. Dann haben wir aus jeder Teilmenge einen vollständig induktiven und einen transduktiven Datensatz sowie zwei weitere angepasste Datensätze extrahiert. Im ersten Satz ist der Prozentsatz der induktiven und transduktiven Tripel festgelegt, und im zweiten Satz hat jedes Tripel eine induktive Entität und die andere Entität ist transduktiv. Die Beschreibung dieser Einstellungen findet sich im Folgenden.

Induktiv: Wir haben vier Datensätze mit der induktiven Einstellung erstellt, bei der die Entitäten der Test- und Trainingsdatensätze vollständig disjunkt sind. Wir haben diese vollständig disjunkten Datensätze von FB15K und WN18 genommen und sie dann in relationale Datensätze unterteilt, so dass wir einen Satz von acht Datensätzen erhielten. Tabelle 28 enthält die Statistiken unserer Datensätze. **Transduktiv:** Das transduktive Setting hat gemeinsame Entitäten in Trainings- und Testdatensätzen. Daher sind die Entitäten dem Modell bereits bekannt, was die Vorhersage wesentlich erleichtert. Tabelle 25 enthält die Statistiken der acht Datensätze des transduktiven Typs. **Kopf-Schwanz-Verhältnis induktiv:** Wir haben eine Reihe von Datensätzen erstellt, indem wir entweder den Kopf oder das Ende jedes Triplets im Trainingsdatensatz vor dem Testdatensatz verborgen haben. Auf diese Weise erhalten wir einen halbinduktiven Datensatz, bei dem jedes Triplet ungesehen bleibt. Tabelle 26 zeigt die Statistiken dieser Datensätze. **Prozentuale Bildung:** In unserer Studie haben wir die Hälfte der Testtripel mit induktiven Einstellungen und die andere Hälfte mit der transduktiven Einstellung erzeugt. Wir wenden diese prozentuale Basisdatengenerierung auf jede Kategorie an. Tabelle 27 beschreibt die Statistiken der 50%-Datensätze.

In den Tabellen 29 werden die AUC-PR-Ergebnisse mehrerer moderner Link-Vorhersagemodelle für unsere extrahierten Tabellen aufgeführt. Erweiterte Ergebnisblätter sind unter https://github.com/mlwin-de/relational_pattern_benchmarking verfügbar.

Fraunhofer IAIS hat zusammen mit den Projektpartnern Uni Bonn und MLU zwei MLwin KGRL-Workshops⁴ (New Trends in Representation Learning with Knowledge Graphs) organisiert und eine Zusage erhalten. Die Workshops fanden im Rahmen der ECML PKDD Konferenz (European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases) statt.

Der erste Workshop fand im Jahr 2019 statt und der zweite ebenfalls im Rahmen der ECML PKDD, jedoch online im Jahr 2021.

2 Zahlenmäßiger Nachweis

Das Förderprojekt wurde weitgehend entsprechend der ursprünglichen Planung bearbeitet. Im ersten Laufzeitjahr mussten noch Kapazitäten gewonnen und eingestellt, sowie Daten gesammelt und geteilt werden. Daher wurde anfangs mit geringer Kapazität an den Arbeitspaketen gearbeitet. Aufgrund der pandemischen Lage seit 2019 wurde das Projekt kostenneutral um 4 Monate verlängert.

Die Tabellen 1, 2, 3, 6, 4, 6, 7, 8, 10 weisen die geleisteten jährlichen Ausgaben der jeweiligen Partner zusammen.

3 Notwendigkeit und Angemessenheit der geleisteten Arbeiten

Keiner der Partner hätte die Arbeiten für sich alleine durchführen können, da alle Kompetenzen im Konsortium für einen Gesamterfolg notwendig waren. Wenn man das Anwendungsszenario betrachtet, so konnte nur mit allen Partnern im Projekt die Integration von Big Data Analysen und Methoden des maschinellen Lernens im klinischen Kontext angegangen werden. UKE hätte den Aufwand nicht durch Hausstellen im Universitätsklinikum leisten können. Durch eine Förderung konnte dieser Bereich weiterentwickelt und gefördert werden. Hierfür standen zum Beginn des Projekts keine alternativen Mittel zur

⁴<https://sites.google.com/view/kgrlfr-workshop/home>

Zeitraum (Geschäftsjahr)	Jährliche Kosten in €	Förderanteil(40%)
2018	44.736,43	17.894,57
2019	476.008,58	190.403,43
2020	435.241,62	174.096,65
2021	254.640,36	101.856,14
2022	46.338,24	18.535,30

Posten	geplante Ausgaben in €	entstandene Kosten
Personalkosten	1.354.795,00	1.256.992,23
Materialkosten		
Fremdleistung		
Reisekosten	18.000,00	
sonstige Kosten		
Summe	1.372.795,00	1.256.922,23

Tabelle 1: Die geleisteten jährlichen Ausgaben und mit dem Förderanteil 40% (oben), sowie die geplanten und tatsächlichen entstandenen Ausgaben (unten) des Projektpartners Siemens AG

Posten	geplante Ausgaben in €	entstandene Kosten
Wissenschaftliche Mitarbeiter	577.740,00	605.560,06
Studentische Hilfskräfte	0,00	0,00
Reisekosten	28.530,00	3.108,17
Anschaffungen von Hardware	0,00	0,00
Summe	606.270,00	608.668,23

Tabelle 2: Die geplanten und tatsächlichen entstandenen Ausgaben des Projektpartners LMU

Verfügung. Für Siemens war es wichtig, dass die verschiedenen Kompetenzen der unterschiedlichen Partner im Projekt gebündelt wurden, da Siemens zwar punktuell sehr weit ist, aber viele der Themen nicht abdecken kann. Die hohe technische Komplexität, der hohe Innovationsgrad und der interdisziplinäre Charakter des Verbundprojekts stellten ein wissenschaftliches und technisches Risiko auch für das FhG IAIS dar. Das betraf insbesondere die Integration der MLwin-Tasks in eine Benchmarking-Plattform und die Implementierung der Methoden im Open Source SANSA-Framework. Für diese zeit- und kostspielige Aufgaben standen keine alternativen Mittel zur Verfügung, so dass das Vorhaben ausschließlich durch die Zuwendung im Rahmen des Förderprogramms "IKT 2020 – Forschung für Innovationen" realisiert würden. Für die universitären Partner UBO, TUM und LMU stellte das Konsortium eine einzigartige Möglichkeit dar, die Forschungsthemen an den Anwendungen und in Zusammenarbeit mit internationalen Großfirmen wie Siemens auf eine neue Ebene zu heben. Die FhG eröffnete für die Unis wichtige Verwertungsschienen. Das UKE konnte einzigartigen Einblick in die Probleme der klinischen Praxis anbieten und einzigartige Daten zur Verfügung stellen. Bei den Universitäten wurde angestrebt, dass alle für das Projekt eingestellten Mitarbeiter im Rahmen des Projekts eine Promotion erwerben. Die war ein wichtiger Faktor für die wirtschaftliche Erschließung von Forschungskapazitäten an deutschen Hochschulen. Des Weiteren ermöglichte es derartige Promotionsstellen, hochbegabten Absolventen/innen Erfahrungen mit der Nutzung von aktuellen Forschungsergebnissen zu sammeln. Hierdurch wurde zum einem die Verfügbarkeit von hochqualifizierten Arbeitskräften am Standort Deutschland gefördert, als auch das Klima für Firmenneugründungen verbessert.

Zeitraum (Geschäftsjahr)	Jährliche Kosten in €	Förderanteil(50%)
2019	83.482,17	41.741,09
2020	57.741,75	28.870,88
2021	55.448,66	27.724,33
2022	0,00	0,00

Tabelle 3: Die geleisteten jährlichen Ausgaben und mit dem Förderanteil 50% für die Partner TUM

Posten	geplante Ausgaben in €	entstandene Kosten
Wissenschaftliche Mitarbeiter	198.693,00	196.672,58
Studentische Hilfskräfte	0,00	0,00
Reisekosten	8.000,00	0,00
Anschaffungen von Hardware	0,00	0,00
Summe	206.693,00	196.672,58

Tabelle 4: Die geplanten und tatsächlichen entstandenen Ausgaben des Projektpartners TUM

4 Nutzen und Verwertbarkeit der Ergebnisse sowie zukünftiger Verwertungsplan

Durch den von der IFG GmbH verbesserten Datenkurationsprozess können in Zukunft finanzielle und zeitliche Kapazitäten eingespart und so eine wirtschaftlich konkurrenzfähige Dienstleistung angeboten werden. Dies ermöglicht den Aufbau neuer Geschäftsverhältnisse mit anderen CROs oder Vertretern der Pharmabranche. Insbesondere die technische Fähigkeit zur Verarbeitung von genetischen Daten aus Freitextfelder klinischer Studiendokumentationssysteme stellt einen nationalen und internationalen wirtschaftlichen Vorteil dar. Wir gehen davon aus, dass die Ausweitung des Datenkurationsprozesses auf Daten aus digitalen, mobilen Quellen (z.B. App-Daten, Sensordaten etc.) in naher Zukunft umgesetzt werden kann. Der zukünftige Verwertungsplan beinhaltet die Weiterentwicklung des im Rahmen von MLWin umgesetzten technischen Prozesses, sodass ein Abgleich von Studiendaten mit der elektronischen Patientenakte durch digitales Monitoring ermöglicht wird. Die Beteiligung an nationalen und internationalen Anschlussprojekten wurde bereits erfolgreich initiiert.

Das UKER wird die wissenschaftlichen Ergebnisse aus der COMMITMENT Studie im Rahmen von Folgeprojekten nutzen. In diesem Zuge sind KI-gestützte realtime Auswertungen von Arzt-Patienten-Gesprächen über Audio- und Videoaufnahmen geplant. Hierbei sollen Verfahren der Gesichtserkennung inkludiert werden, um Gefühle und Reaktionen während des Gespräches zu detektieren. Hierdurch soll der Algorithmus weiter optimiert und die Gesprächsführung durch KI-gestützte Hinweise unterstützt werden.

Siemens AG legt viel Wert darauf, umfangreiche Domain-wissen im Zusammenhang mit den Produkten mit den entwickelten Technologien auf strukturierterer Weise abzuspeichern, darzustellen, und neue Insights daraus zu extrahieren. Insbesondere von Interesse ist die neue Möglichkeit, unterschiedlichen Datenquellen wie Text- und Bild-Daten zu fusionieren. Beispielsweise wird evaluiert, Produkt-Informationen als KG darzustellen und zu analysieren, so dass die Kunden individualisierte Produkt-Empfehlungen erhalten. In Kollaboration mit UKER entwickelt Siemens auch fortgeschrittene Methoden zum Vorhersagen therapeutischer Entscheidungen. Im Fokus steht u.a., dass die KI diese Vorhersagen optimiert bzgl. des Outcomes der Patienten, statt einfacher Likelihood jeder Entscheidung. Eine andere Innovation besteht darin, dass die KI aus den fusionierten klinischen Daten erlernt, jeder individuellen Vorhersage eine mathematisch begründete Unsicherheitsgröße zuzuordnen. Diese könnte von den menschlichen Ärzten als Konfidenz oder Vertrauenswürdigkeit der KI interpretiert werden.

Die Modellierung symbolischen Wissens über Wissensgraphen findet in verschiedenen Richtungen des maschinellen Lernens Anwendungen. An der LMU sollen die erzielten Ergebnisse daher bei der Einwerbung neuer Drittmittelprojekte in weiteren Bereichen als Nachweis der Einschlägigkeit der Projektpartner

Zeitraum (Geschäftsjahr)	Jährliche Kosten in €	Förderanteil(50%)
2018	933,24	14.036,84
2019	83.659,16	63.350,00
2020	66.161,17	71.489,00
2021	71.638,91	74.343,13
2022	7.211,00	12.282,03

Tabelle 5: Die geleisteten jährlichen Ausgaben und mit dem Förderanteil 100% für die Partner UBO

Posten	geplante Ausgaben in €	entstandene Kosten
Wissenschaftliche Mitarbeiter	197.421,00	194.756,05
Studentische Hilfskräfte	16.730,00	16.218,49
Reisekosten	9.350,00	5.235,49
Anschaffungen von Hardware	12.000,00	13.393,45
Summe	235.501,00	229.603,48

Tabelle 6: Die geplanten und tatsächlichen entstandenen Ausgaben des Projektpartners UBO

Verwendung finden.

Das vom Fraunhofer IAIS im Rahmen des Projekts entwickelte MDE-Einbettungsmodell hat Anwendungen in der digitalen Medizin gefunden, indem es zeigt, wie das Modell dazu beiträgt, Ärzte mit den neuesten Updates in ihrem Bereich zu verbinden [78]. Darüber hinaus wird das im Projekt [76] entwickelte Modell des neuronalen Netzes GFA-NN bereits für die Datenintegration und die Verknüpfung von Vorschlägen zwischen Unternehmensverbrauchern und Organisationsabteilungen im Projekt Open Integration Hub Plus (OIH Plus) verwendet. Dieses Projektkonsortium wird vom BMWK gefördert und hat bisher das OIH-Framework geschaffen. Das Fraunhofer IAIS ist als Mitglied dieses Konsortiums an der Entwicklung dieser industriellen Anwendung beteiligt.

5 Zwischenzeitlicher technischer Fortschritt auf dem Gebiet des Vorhabens

ML und Data Mining in KGs: Die Generierung, Validierung und Nutzung von Graphdaten und KGs, im Besonderen, erfreute sich während der Projektlaufzeit reger Aufmerksamkeit verschiedener weiterer Spitzengruppen im Bereich maschinelles Lernen. Ein kürzlich erschienener Überblick über die allgemeine Behandlung von Graphen im Bereich Deep Learning ist in [106] zu finden. Es wurden neue Verfahren in den Bereichen Repräsentationslernen für allgemeine KGs [95, 47, 54] und zeitliche KGs [55, 98, 110, 45] vorgeschlagen. Insbesondere wurden mehrere Methoden zur Verwendung von Methoden des Reinforcement Learnings vorgeschlagen [14, 92, 57], um unbekannte Beziehungen zwischen Entitäten in KGs vorherzusagen.

Im Rahmen des Projekts entwickelte das Fraunhofer IAIS die Benchmarking-Plattform *BenchEmbedd*⁵ für ein zuverlässiges Benchmarking von Knowledge Graph Embedding-Modellen. Wir haben diese Plattform später in unserer anderen Studie verwendet, die in NeurIPS veröffentlicht wurde.⁶ In der Folgestudie haben wir einen umfangreichen leckfreien Beziehungsmuster-Benchmarking-Datensatz generiert und modernste Knowledge Graph Embedding-Modelle mit Hilfe des extrahierten Datensatzes und *BenchEmbedd* einem Benchmarking unterzogen. Wir planen, die Studie alle zwei Jahre zu wiederholen und neuere Modelle in den Benchmark zu integrieren. Unser veröffentlichter Bericht enthält derzeit einen umfangreichen Satz von Knowledge Graph Embedding-Modellen, die unter https://github.com/mlwin-de/relational_pattern_benchmarking zugänglich sind.

⁵<https://github.com/mlwin-de/BenchEmbedd>

⁶<https://openreview.net/pdf?id=Gln7zxMffae>

Zeitraum (Geschäftsjahr)	Jährliche Kosten in €	Förderanteil(100%)
2018	25.355,49	25.355,49
2019	72.277,84	72.277,84
2020	81.355,22	81.355,22
2021	65.822,69	65.822,69
2022	3.349,38	3.349,38

Tabelle 7: Die geleisteten jährlichen Ausgaben und mit dem Förderanteil 100% für die Partner UKER

Posten	geplante Ausgaben in €	entstandene Kosten
Wissenschaftliche Mitarbeiter	237.807,00	238.253,33
Studentische Hilfskräfte	11.375,00	9.907,30
Reisekosten	-	-
Anschaffungen von Hardware	-	-
Summe	249.182,00	248.160,63

Tabelle 8: Die geplanten und tatsächlichen entstandenen Ausgaben des Projektpartners UKER

Textanalyse mit KGs: Im Bereich der Textanalyse und Sprachverarbeitung wurden auch außerhalb des Projekts KGs zur Strukturierung von in Text enthaltenen Informationen von unterschiedlichsten Forschenden untersucht. Zum Beispiel wurden neue Verfahren zur Ableitung von KGs aus Textdaten vorgeschlagen [22, 60]. Insbesondere sind hier Arbeiten [11, 69, 93] über die inhärente Struktur von Wissen in gängigen Sprachmodellen von hoher Relevanz. Des Weiteren wurden Modelle zur Generierung von textuellen Beschreibungen aus KGs entwickelt [49, 63, 72, 74].

Bild- und Videoanalyse mit KGs:

Fortgeschrittene Modellierung von Patientendaten Um die longitudinale Eigenschaft der EHR Daten zu fassen werden in letzter Zeit oft Rekurrente Neuronale Netze verwendet [70]. [96] und [12] haben gezeigt, dass man von den neuesten Entwicklungen in Deep Learning für die Analyse medizinischer Bilddaten profitieren kann. Wir konnten auch von den Fortschritten in Unsicherheitsquantifizierung wie [44] und [103] profitieren. **Modellierung von Wissensgraphen** Die multi-hop Reasoning Methode MINERVA wurde entwickelt von [14]. [35] hat eine der aktuellen Methoden zur Translations-basierten Empfehlung vorgeschlagen, welche Long-Term Gedächtnis der Benutzenden ausnutzt. Es wurde gezeigt, dass man Quantum-Algorithmen anwenden kann, um Tensor-Faktorisierung durch High-Order SVD zu realisieren [30], was die Grundlage für viele Wissensgraph-Techniken bilden könnte. **Transfer in die Klinik** Methoden des Deep Learnings finden nur langsam Einsatz in der klinischen Praxis, weil die Algorithmen oftmals als schwierig zu interpretieren – im Sinne des Black-Box-Phänomens – eingeordnet werden. Kürzlich wurde eine Graph Layer-wise Relevance Propagation (GLRP) Methode entwickelt, die Tumorboard-Entscheidungen auf Basis von Graph Convolutional Neural Networks (GCN) erklärbar macht [40]. Dies könnte in Zukunft dazu beitragen, dass die Einstellung klinischen Fachpersonals zu Deep-Learning-Algorithmen positiv wird.

6 Nachweis erfolgter oder geplanter Veröffentlichungen der Ergebnisse

Es gab im Projekt folgende Veröffentlichungen.

Zeitraum (Geschäftsjahr)	Jährliche Kosten in €	Förderanteil(60%)
2019	117.524,64	70.514,784
2020	71.973,80	43.184,28
2021	7.287,04	4.372,224
2022	-	-

Tabelle 9: Die geleisteten jährlichen Ausgaben (inkl. Pauschalzuschlag) und mit dem Förderanteil 60% für die Partner IFG

Posten	geplante Ausgaben in €	entstandene Kosten
Wissenschaftliche Mitarbeiter	189.389,00	196.604,68
Studentische Hilfskräfte	-	-
Reisekosten	2.750,00	180,80
Anschaffungen von Hardware	-	-
Summe	192.139,00	196.785,48
Bundesanteil (60% Förderquote)	115.283,00	115.283,00

Tabelle 10: Die geplanten und tatsächlichen entstandenen Ausgaben des Projektpartners IFG

Veröffentlichungen

- Christian M. M. Frey, Yunpu Ma und Matthias Schubert. „APPTeK: Agent-Based Predicate Prediction in Temporal Knowledge Graphs“. In: *CoRR* abs/2110.14284 (2021). arXiv: 2110.14284
- Christian M. M. Frey und Matthias Schubert. *V-Coder: Adaptive AutoEncoder for Semantic Disclosure in Knowledge Graphs*. 2022. DOI: 10.48550/ARXIV.2208.01735
- Christian M. M. Frey, Yunpu Ma und Matthias Schubert. „SEA: Graph Shell Attention in Graph Neural Networks“. In: *Machine Learning and Knowledge Discovery in Database ECML/PKDD*. 2022
- Mehdi Ali, Hajira Jabeen, Charles Tapley Hoyt und Jens Lehmann. „The KEEN Universe - An Ecosystem for Knowledge Graph Embeddings with a Focus on Reproducibility and Transferability“. In: *ISWC (2)*. Bd. 11779. Lecture Notes in Computer Science. Springer, 2019, S. 3–18
- Mehdi Ali, Max Berrendorf, Charles Tapley Hoyt, Laurent Vermue, Sahand Sharifzadeh, Volker Tresp und Jens Lehmann. „PyKEEN 1.0: A Python Library for Training and Evaluating Knowledge Graph Embeddings“. In: *J. Mach. Learn. Res.* 22 (2021), 82:1–82:6
- Mehdi Ali, Charles Tapley Hoyt, Daniel Domingo-Fernández, Jens Lehmann und Hajira Jabeen. „BioKEEN: a library for learning and evaluating biological knowledge graph embeddings“. In: *Bioinform.* 35.18 (2019), S. 3538–3540
- Mehdi Ali, Max Berrendorf, Charles Tapley Hoyt, Laurent Vermue, Mikhail Galkin, Sahand Sharifzadeh, Asja Fischer, Volker Tresp und Jens Lehmann. „Bringing light into the dark: A large-scale evaluation of knowledge graph embedding models under a unified framework“. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2021)
- Mehdi Ali, Max Berrendorf, Mikhail Galkin, Veronika Thost, Tengfei Ma, Volker Tresp und Jens Lehmann. „Improving Inductive Link Prediction Using Hyper-relational Facts“. In: *ISWC*. Bd. 12922. Lecture Notes in Computer Science. Springer, 2021, S. 74–92
- Vinay Srinivas Bharadhwaj, Mehdi Ali, Colin Birkenbihl, Sarah Mubeen, Jens Lehmann, Martin Hofmann-Apitius, Charles Tapley Hoyt und Daniel Domingo-Fernández. „CLEP: a hybrid data- and knowledge-driven framework for generating patient representations“. In: *Bioinform.* 37.19 (2021), S. 3311–3318

- Martin Schmitt und Hinrich Schütze. „SherLLiC: A Typed Event-Focused Lexical Inference Benchmark for Evaluating Natural Language Inference“. In: *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*. Association for Computational Linguistics, 2019, S. 902–914. DOI: 10.18653/v1/p19-1086
- Martin Schmitt und Hinrich Schütze. „Unsupervised Text Generation from Structured Data“. In: *CoRR abs/1904.09447* (2019). arXiv: 1904.09447
- Leonardo F. R. Ribeiro, Martin Schmitt, Hinrich Schütze und Iryna Gurevych. „Investigating Pre-trained Language Models for Graph-to-Text Generation“. In: *CoRR abs/2007.08426* (2020). arXiv: 2007.08426
- Martin Schmitt, Leonardo F. R. Ribeiro, Philipp Dufter, Iryna Gurevych und Hinrich Schütze. „Modeling Graph Structure via Relative Position for Better Text Generation from Knowledge Graphs“. In: *CoRR abs/2006.09242* (2020). arXiv: 2006.09242
- Martin Schmitt, Sahand Sharifzadeh, Volker Tresp und Hinrich Schütze. „An Unsupervised Joint System for Text Generation from Knowledge Graphs and Semantic Parsing“. In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP 2020, Online, November 16-20, 2020*. Association for Computational Linguistics, 2020, S. 7117–7130. DOI: 10.18653/v1/2020.emnlp-main.577
- Philipp Dufter, Martin Schmitt und Hinrich Schütze. „Position information in transformers: An overview“. In: *Computational Linguistics* (2021), S. 1–31
- Maximilian Mozes, Martin Schmitt, Vladimir Golkov, Hinrich Schütze und Daniel Cremers. „Scene Graph Generation for Better Image Captioning?“ In: *CoRR abs/2109.11398* (2021). arXiv: 2109.11398
- Martin Schmitt, Leonardo F. R. Ribeiro, Philipp Dufter, Iryna Gurevych und Hinrich Schütze. „Modeling Graph Structure via Relative Position for Text Generation from Knowledge Graphs“. In: *Proceedings of the Fifteenth Workshop on Graph-Based Methods for Natural Language Processing (TextGraphs-15)*. Mexico City, Mexico: Association for Computational Linguistics, Juni 2021, S. 10–21. DOI: 10.18653/v1/2021.textgraphs-1.2
- Martin Schmitt und Hinrich Schütze. „Language Models for Lexical Inference in Context“. In: *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*. Online: Association for Computational Linguistics, Apr. 2021, S. 1267–1280. DOI: 10.18653/v1/2021.eacl-main.108
- Martin Schmitt und Hinrich Schütze. „Continuous Entailment Patterns for Lexical Inference in Context“. In: *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 7-11 November, 2021*. Association for Computational Linguistics, 2021, S. 6952–6959. DOI: 10.18653/v1/2021.emnlp-main.556
- Marina Sedinkina, Martin Schmitt und Hinrich Schütze. „Domain Adaptation for Sparse-Data Settings: What Do We Gain by Not Using Bert?“ In: *CoRR abs/2203.16926* (2022)
- Suprosanna Shit u. a. „Relationformer: A Unified Framework for Image-to-Graph Generation“. In: *CoRR abs/2203.10202* (2022). DOI: 10.48550/arXiv.2203.10202. arXiv: 2203.10202
- M. Eisenberger, Z. Lähner und D. Cremers. „Smooth Shells: Multi-Scale Shape Registration with Functional Maps“. In: *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*. 2020

- M. Eisenberger, A. Toker, L. Leal-Taixé und D. Cremers. „Deep Shells: Unsupervised Shape Correspondence with Optimal Transport“. In: *34th Conference on Neural Information Processing Systems (NeurIPS)*. 2020
- M. Eisenberger, D. Novotny, G. Kerchenbaum, P. Labatut, N. Neverova, D. Cremers und A. Vedaldi. „NeuroMorph: Unsupervised Shape Interpolation and Correspondence in One Go“. In: *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*. 2021
- Abhishek Saroha, Marvin Eisenberger, Tarun Yenamandra und Daniel Cremers. „Implicit Shape Completion via Adversarial Shape Priors“. In: *arXiv preprint arXiv:2204.10060* (2022)
- I Chiotellis und D Cremers. „Neural Online Graph Exploration“. In: *arXiv preprint arXiv:2012.03345* (2020). arXiv: 2012.03345 [cs.LG]
- G Fabbro, V Golkov, T Kemp und D Cremers. „Speech Synthesis and Control Using Differentiable DSP“. In: *arXiv preprint arXiv:2010.15084* (2020). arXiv: 2010.15084 [eess.AS]
- V. Golkov, M. J. Skwark, A. Mirchev, G. Dikov, A. R. Geanes, J. Mendenhall, J. Meiler und D. Cremers. „3D Deep Learning for Biological Function Prediction from Physical Fields“. In: *International Conference on 3D Vision (3DV)*. 2020. arXiv: 1704.04039
- V. Golkov, A. Becker, D. T. Plop, D. Čuturilo, N. Davoudi, J. Mendenhall, R. Moretti, J. Meiler und D. Cremers. „Deep Learning for Virtual Screening: Five Reasons to Use ROC Cost Functions“. In: *arXiv preprint arXiv:2007.07029* (2020). arXiv: 2007.07029
- L. Della Libera, V. Golkov, Y. Zhu, A. Mielke und D. Cremers. „Deep Learning for 2D and 3D Rotatable Data: An Overview of Methods“. In: *arXiv preprint arXiv:1910.14594* (2019). arXiv: 1910.14594
- P. Müller, V. Golkov, V. Tomassini und D. Cremers. „Rotation-Equivariant Deep Learning for Diffusion MRI (short version)“. In: *International Society for Magnetic Resonance in Medicine (ISMRM) Annual Meeting*. 2021
- P. Müller, V. Golkov, V. Tomassini und D. Cremers. „Rotation-Equivariant Deep Learning for Diffusion MRI“. In: *arXiv preprint* (2021). arXiv: 2102.06942 [cs.CV]
- M. Naeyaert, J. Aelterman, J. Van Audekerke, V. Golkov, D. Cremers, A. Pižurica, J. Sijbers und M. Verhoye. „Accelerating in vivo fast spin echo high angular resolution diffusion imaging with an isotropic resolution in mice through compressed sensing“. In: *Magnetic Resonance in Medicine* 85.3 (2020), S. 1397–1413. DOI: 10.1002/mrm.28520. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.28520>
- M Naeyaert, V Golkov, D Cremers, J Sijbers und M Verhoye. „Faster and better HARDI using FSE and holistic reconstruction“. In: *International Society for Magnetic Resonance in Medicine (ISMRM) Annual Meeting*. 2021
- Marcel Hildebrandt, Hang Li, Rajat Koner, Volker Tresp und Stephan Günnemann. „Scene Graph Reasoning for Visual Question Answering“. In: *CoRR abs/2007.01072* (2020). arXiv: 2007.01072
- Rajat Koner, Poulami Sinhamahapatra, Volker Tresp und Stephan Günnemann. „Scenes and Surroundings: Scene Graph Generation using Relation Transformer“. In: *ICML workshop on Bridge Between Perception and Reasoning: Graph Neural Networks and Beyond*. 2020
- Rajat Koner, Poulami Sinhamahapatra und Volker Tresp. „Relation Transformer Network“. In: *CoRR abs/2004.06193* (2020). arXiv: 2004.06193
- Rajat Koner, Poulami Sinhamahapatra, Karsten Roscher, Stephan Günnemann und Volker Tresp. „OODformer: Out-Of-Distribution Detection Transformer“. In: *32nd British Machine Vision Conference 2021, BMVC 2021, Online, November 22-25, 2021*. BMVA Press, 2021, S. 209

- Rajat Koner, Hang Li, Marcel Hildebrandt, Deepan Das, Volker Tresp und Stephan Günnemann. „Graphhopper: Multi-hop Scene Graph Reasoning for Visual Question Answering“. In: *The Semantic Web – ISWC 2021*. Cham: Springer International Publishing, 2021, S. 111–127. ISBN: 978-3-030-88361-4
- S. Sharifzadeh, S. Baharlou, M. Berrendorf, R. Koner und V. Tresp. „Improving Visual Relation Detection using Depth Maps“. In: *2020 25th International Conference on Pattern Recognition (ICPR)*. Los Alamitos, CA, USA: IEEE Computer Society, Jan. 2021, S. 3597–3604. DOI: 10.1109/ICPR48806.2021.9412945
- Tanveer Hannan, Rajat Koner, Jonathan Kobold und Matthias Schubert. „Box Supervised Video Segmentation Proposal Network“. In: *CoRR abs/2202.07025 (2022)*. arXiv: 2202.07025
- Sahand Sharifzadeh, Sina Moayed Baharlou, Martin Schmitt, Hinrich Schütze und Volker Tresp. „Improving Scene Graph Classification by Exploiting Knowledge from Texts“. In: *Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence, IAAI 2022, The Twelfth Symposium on Educational Advances in Artificial Intelligence, EAAI 2022 Virtual Event, February 22 - March 1, 2022*. AAAI Press, 2022, S. 2189–2197
- Zhen Han, Yunpu Ma, Yuyi Wang, Stephan Günnemann und Volker Tresp. „Graph hawkes neural network for forecasting on temporal knowledge graphs“. In: *Automated Knowledge Base Construction (AKBC), 2020*
- Yunpu Ma und Volker Tresp. „Quantum Machine Learning on Knowledge Graphs“. In: Bd. 2. 03. *ACM Transactions on Quantum Computing*, 2021
- Yunpu Ma, Volker Tresp, Liming Zhao und Yuyi Wang. „Variational quantum circuit model for knowledge graph embedding“. In: Bd. 2. 7-8. *Wiley Online Library*, 2019, S. 1800078
- Marcel Hildebrandt, Jorge Andres Quintero Serna, Yunpu Ma, Martin Ringsquandl, Mitchell Joblin und Volker Tresp. „Reasoning on knowledge graphs with debate dynamics“. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Bd. 34. 04. 2020, S. 4123–4131
- Malte Feucht, Zhiliang Wu, Sophia Althammer und Volker Tresp. „Description-based Label Attention Classifier for Explainable ICD-9 Classification“. In: *arXiv preprint arXiv:2109.12026 (2021)*
- Zhiliang Wu, Yinchong Yang, Jindong Gu und Volker Tresp. „Quantifying predictive uncertainty in medical image analysis with deep kernel learning“. In: *2021 IEEE 9th International Conference on Healthcare Informatics (ICHI)*. IEEE. 2021, S. 63–72
- Zhiliang Wu, Yinchong Yang, Peter A Fashing und Volker Tresp. „Uncertainty-Aware Time-to-Event Prediction using Deep Kernel Accelerated Failure Time Models“. In: *Machine Learning for Healthcare Conference*. PMLR. 2021, S. 54–79

Geplante Veröffentlichungen

- Matthias Ruebner u. a. *Routine genotyping within PRAEGNANT breast cancer registry - Automated mutation identification from routine genotyping reports, physician letters and pathological reports*. 2022
- TBN. *Feasibility and acceptance of the use of audio recordings from routine physician-patient dialogues for implementation into machine learning algorithms*. 2022

Softwarepakete

- <https://github.com/the/most/awesome/package>

Literatur

- [1] Emmanuel Abbe. „Community Detection and Stochastic Block Models“. In: *Found. Trends Commun. Inf. Theory* (Juni 2018). ISSN: 1567-2190. DOI: 10.1561/01000000067.
- [2] Mehdi Ali, Max Berrendorf, Mikhail Galkin, Veronika Thost, Tengfei Ma, Volker Tresp und Jens Lehmann. „Improving Inductive Link Prediction Using Hyper-relational Facts“. In: *ISWC*. Bd. 12922. Lecture Notes in Computer Science. Springer, 2021, S. 74–92.
- [3] Mehdi Ali, Max Berrendorf, Charles Tapley Hoyt, Laurent Vermue, Mikhail Galkin, Sahand Sharifzadeh, Asja Fischer, Volker Tresp und Jens Lehmann. „Bringing light into the dark: A large-scale evaluation of knowledge graph embedding models under a unified framework“. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2021).
- [4] Mehdi Ali, Max Berrendorf, Charles Tapley Hoyt, Laurent Vermue, Sahand Sharifzadeh, Volker Tresp und Jens Lehmann. „PyKEEN 1.0: A Python Library for Training and Evaluating Knowledge Graph Embeddings“. In: *J. Mach. Learn. Res.* 22 (2021), 82:1–82:6.
- [5] Mehdi Ali, Charles Tapley Hoyt, Daniel Domingo-Fernández, Jens Lehmann und Hajira Jabeen. „BioKEEN: a library for learning and evaluating biological knowledge graph embeddings“. In: *Bioinform.* 35.18 (2019), S. 3538–3540.
- [6] Mehdi Ali, Hajira Jabeen, Charles Tapley Hoyt und Jens Lehmann. „The KEEN Universe - An Ecosystem for Knowledge Graph Embeddings with a Focus on Reproducibility and Transferability“. In: *ISWC (2)*. Bd. 11779. Lecture Notes in Computer Science. Springer, 2019, S. 3–18.
- [7] Ivana Balazevic, Carl Allen und Timothy M. Hospedales. „Multi-relational Poincaré Graph Embeddings“. In: *NeurIPS*. 2019, S. 4465–4475.
- [8] Ivana Balazevic, Carl Allen und Timothy M. Hospedales. „TuckER: Tensor Factorization for Knowledge Graph Completion“. In: *EMNLP/IJCNLP (1)*. Association for Computational Linguistics, 2019, S. 5184–5193.
- [9] Dominique Beaini, Saro Passaro, Vincent Létourneau, William L. Hamilton, Gabriele Corso und Pietro Liò. „Directional Graph Networks“. In: *CoRR* (2020). arXiv: 2010.02863.
- [10] Vinay Srinivas Bharadhwaj, Mehdi Ali, Colin Birkenbihl, Sarah Mubeen, Jens Lehmann, Martin Hofmann-Apitius, Charles Tapley Hoyt und Daniel Domingo-Fernández. „CLEP: a hybrid data- and knowledge-driven framework for generating patient representations“. In: *Bioinform.* 37.19 (2021), S. 3311–3318.
- [11] Zied Bouraoui, José Camacho-Collados und Steven Schockaert. „Inducing Relational Knowledge from BERT“. In: *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*. AAAI Press, 2020, S. 7456–7463.
- [12] Ting Chen, Simon Kornblith, Mohammad Norouzi und Geoffrey Hinton. „A simple framework for contrastive learning of visual representations“. In: *International conference on machine learning*. PMLR. 2020, S. 1597–1607.
- [13] I Chiotellis und D Cremers. „Neural Online Graph Exploration“. In: *arXiv preprint arXiv:2012.03345* (2020). arXiv: 2012.03345 [cs.LG].
- [14] Rajarshi Das, Shehzaad Dhuliawala, Manzil Zaheer, Luke Vilnis, Ishan Durugkar, Akshay Krishnamurthy, Alex Smola und Andrew McCallum. „Go for a walk and arrive at the answer: Reasoning over paths in knowledge bases using reinforcement learning“. In: *arXiv preprint arXiv:1711.05851* (2017).
- [15] Tim Dettmers, Pasquale Minervini, Pontus Stenetorp und Sebastian Riedel. „Convolutional 2D Knowledge Graph Embeddings“. In: *AAAI*. AAAI Press, 2018, S. 1811–1818.

- [16] Tim Dettmers, Pasquale Minervini, Pontus Stenetorp und Sebastian Riedel. „Convolutional 2d knowledge graph embeddings“. In: *Thirty-Second AAAI Conference on Artificial Intelligence*. 2018.
- [17] Philipp Dufter, Martin Schmitt und Hinrich Schütze. „Position information in transformers: An overview“. In: *Computational Linguistics (2021)*, S. 1–31.
- [18] M. Eisenberger, Z. Löhner und D. Cremers. „Smooth Shells: Multi-Scale Shape Registration with Functional Maps“. In: *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*. 2020.
- [19] M. Eisenberger, D. Novotny, G. Kerchenbaum, P. Labatut, N. Neverova, D. Cremers und A. Vedaldi. „NeuroMorph: Unsupervised Shape Interpolation and Correspondence in One Go“. In: *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*. 2021.
- [20] M. Eisenberger, A. Toker, L. Leal-Taixé und D. Cremers. „Deep Shells: Unsupervised Shape Correspondence with Optimal Transport“. In: *34th Conference on Neural Information Processing Systems (NeurIPS)*. 2020.
- [21] G Fabbro, V Golkov, T Kemp und D Cremers. „Speech Synthesis and Control Using Differentiable DSP“. In: *arXiv preprint arXiv:2010.15084 (2020)*. arXiv: 2010.15084 [eess.AS].
- [22] Angela Fan, Claire Gardent, Chloé Braud und Antoine Bordes. „Using Local Knowledge Graph Construction to Scale Seq2Seq Models to Multi-Document Inputs“. In: *EMNLP*. 2019.
- [23] Malte Feucht, Zhiliang Wu, Sophia Althammer und Volker Tresp. „Description-based Label Attention Classifier for Explainable ICD-9 Classification“. In: *arXiv preprint arXiv:2109.12026 (2021)*.
- [24] Christian M. M. Frey, Yunpu Ma und Matthias Schubert. „APPTeK: Agent-Based Predicate Prediction in Temporal Knowledge Graphs“. In: *CoRR abs/2110.14284 (2021)*. arXiv: 2110.14284.
- [25] Christian M. M. Frey, Yunpu Ma und Matthias Schubert. „SEA: Graph Shell Attention in Graph Neural Networks“. In: *Machine Learning and Knowledge Discovery in Database ECML/PKDD*. 2022.
- [26] Christian M. M. Frey und Matthias Schubert. *V-Coder: Adaptive AutoEncoder for Semantic Disclosure in Knowledge Graphs*. 2022. DOI: 10.48550/ARXIV.2208.01735.
- [27] Rishab Goel, Mehran Seyed Kazemi, Marcus Brubaker und Pascal Poupart. „Diachronic Embedding for Temporal Knowledge Graph Completion“. In: *National Conference on Artificial Intelligence (AAAI) (2020)*.
- [28] V. Golkov, A. Becker, D. T. Plop, D. Čuturilo, N. Davoudi, J. Mendenhall, R. Moretti, J. Meiler und D. Cremers. „Deep Learning for Virtual Screening: Five Reasons to Use ROC Cost Functions“. In: *arXiv preprint arXiv:2007.07029 (2020)*. arXiv: 2007.07029.
- [29] V. Golkov, M. J. Skwark, A. Mirchev, G. Dikov, A. R. Geanes, J. Mendenhall, J. Meiler und D. Cremers. „3D Deep Learning for Biological Function Prediction from Physical Fields“. In: *International Conference on 3D Vision (3DV)*. 2020. arXiv: 1704.04039.
- [30] Lejia Gu, Xiaoqiang Wang und Guofeng Zhang. „Quantum higher order singular value decomposition“. In: *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*. IEEE. 2019, S. 1166–1171.
- [31] Will Hamilton, Zhitao Ying und Jure Leskovec. „Inductive Representation Learning on Large Graphs“. In: *Advances in Neural Information Processing Systems*. Hrsg. von I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan und R. Garnett. 2017.
- [32] Zhen Han, Yunpu Ma, Yuyi Wang, Stephan Günnemann und Volker Tresp. „Graph hawkes neural network for forecasting on temporal knowledge graphs“. In: *Automated Knowledge Base Construction (AKBC)*, 2020.
- [33] Tanveer Hannan, Rajat Koner, Jonathan Kobold und Matthias Schubert. „Box Supervised Video Segmentation Proposal Network“. In: *CoRR abs/2202.07025 (2022)*. arXiv: 2202.07025.

- [34] Demis Hassabis, Dharshan Kumaran, Christopher Summerfield und Matthew Botvinick. „Neuroscience-inspired artificial intelligence“. In: *Neuron* 95.2 (2017), S. 245–258.
- [35] Ruining He, Wang-Cheng Kang und Julian McAuley. „Translation-based recommendation“. In: *Proceedings of the eleventh ACM conference on recommender systems*. 2017, S. 161–169.
- [36] Marcel Hildebrandt, Hang Li, Rajat Koner, Volker Tresp und Stephan Günnemann. „Scene Graph Reasoning for Visual Question Answering“. In: *CoRR* abs/2007.01072 (2020). arXiv: 2007.01072.
- [37] Marcel Hildebrandt, Hang Li, Rajat Koner, Volker Tresp und Stephan Günnemann. „Scene graph reasoning for visual question answering“. In: *arXiv preprint arXiv:2007.01072* (2020).
- [38] Marcel Hildebrandt, Jorge Andres Quintero Serna, Yunpu Ma, Martin Ringsquandl, Mitchell Joblin und Volker Tresp. „Reasoning on knowledge graphs with debate dynamics“. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Bd. 34. 04. 2020, S. 4123–4131.
- [39] Marcel Hildebrandt, Swathi Shyam Sunder, Serghei Mogoreanu, Mitchell Joblin, Akhil Mehta, Ingo Thon und Volker Tresp. „A recommender system for complex real-world applications with nonlinear dependencies and knowledge graph context“. In: *European Semantic Web Conference*. Springer. 2019, S. 179–193.
- [40] Chereda Hryhorii, Bleckmann Annalen, Menck Kerstin, Perera-Bel Julia, Stegmaier Philip, Auer Florian, Kramer Frank, Leha Andreas und Beissbarth Tim. „Explaining decisions of graph convolutional neural networks: patient-specific molecular subnetworks responsible for metastasis prediction in breast cancer“. In: *GENOME MEDICINE* 1756-994X (2021).
- [41] Weihua Hu, Matthias Fey, Marinka Zitnik, Yuxiao Dong, Hongyu Ren, Bowen Liu, Michele Catasta und Jure Leskovec. „Open Graph Benchmark: Datasets for Machine Learning on Graphs“. In: *NeurIPS*. Hrsg. von Hugo Larochelle, Marc’Aurelio Ranzato, Raia Hadsell, Maria-Florina Balcan und Hsuan-Tien Lin. 2020.
- [42] Weihua Hu, Matthias Fey, Marinka Zitnik, Yuxiao Dong, Hongyu Ren, Bowen Liu, Michele Catasta und Jure Leskovec. „Open Graph Benchmark: Datasets for Machine Learning on Graphs“. In: *arXiv preprint arXiv:2005.00687* (2020).
- [43] John J Irwin, Teague Sterling, Michael M Mysinger, Erin S Bolstad und Ryan G Coleman. „ZINC: A Free Tool to Discover Chemistry for Biology“. eng. In: *Journal of chemical information and modeling* (2012). ISSN: 1549-9596.
- [44] Martin Jankowiak, Geoff Pleiss und Jacob Gardner. „Parametric gaussian process regressors“. In: *International Conference on Machine Learning*. PMLR. 2020, S. 4702–4712.
- [45] Jaehun Jung, Jinhong Jung und U Kang. „Learning to Walk across Time for Interpretable Temporal Knowledge Graph Completion“. In: *KDD ’21: The 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Virtual Event, Singapore, August 14-18, 2021*. Hrsg. von Feida Zhu, Beng Chin Ooi und Chunyan Miao. ACM, 2021, S. 786–795. DOI: 10.1145/3447548.3467292.
- [46] Rudolf Kadlec, Ondrej Bajgar und Jan Kleindienst. „Knowledge Base Completion: Baselines Strike Back“. In: *Rep4NLP@ACL*. Association for Computational Linguistics, 2017, S. 69–74.
- [47] Seyed Mehran Kazemi und David Poole. „Simple Embedding for Link Prediction in Knowledge Graphs“. In: *Proceedings of the 32nd International Conference on Neural Information Processing Systems*. NIPS’18. Montréal, Canada: Curran Associates Inc., 2018, S. 4289–4300.
- [48] Thomas N. Kipf und Max Welling. „Semi-Supervised Classification with Graph Convolutional Networks“. In: *Proceedings of the 5th International Conference on Learning Representations (ICLR)*. ICLR ’17. Palais des Congrès Neptune, Toulon, France, 2017.
- [49] Rik Koncel-Kedziorski, Dhanush Bekal, Yi Luan, Mirella Lapata und Hannaneh Hajishirzi. „Text Generation from Knowledge Graphs with Graph Transformers“. In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*. Hrsg. von Jill Burstein, Christy Doran und Thamar Solorio. Association for Computational Linguistics, 2019, S. 2284–2293. DOI: 10.18653/v1/n19-1238.

- [50] Rajat Koner, Hang Li, Marcel Hildebrandt, Deepan Das, Volker Tresp und Stephan Günnemann. „Graphhopper: Multi-hop Scene Graph Reasoning for Visual Question Answering“. In: *The Semantic Web – ISWC 2021*. Cham: Springer International Publishing, 2021, S. 111–127. ISBN: 978-3-030-88361-4.
- [51] Rajat Koner, Poulami Sinhamahapatra, Karsten Roscher, Stephan Günnemann und Volker Tresp. „OODformer: Out-Of-Distribution Detection Transformer“. In: *32nd British Machine Vision Conference 2021, BMVC 2021, Online, November 22-25, 2021*. BMVA Press, 2021, S. 209.
- [52] Rajat Koner, Poulami Sinhamahapatra und Volker Tresp. „Relation Transformer Network“. In: *CoRR abs/2004.06193 (2020)*. arXiv: 2004.06193.
- [53] Rajat Koner, Poulami Sinhamahapatra, Volker Tresp und Stephan Günnemann. „Scenes and Surroundings: Scene Graph Generation using Relation Transformer“. In: *ICML workshop on Bridge Between Perception and Reasoning: Graph Neural Networks and Beyond*. 2020.
- [54] Devin Kreuzer, Dominique Beaini, William L. Hamilton, Vincent Létourneau und Prudencio Tossou. *Rethinking Graph Transformers with Spectral Attention*. 2021. arXiv: 2106.03893 [cs.LG].
- [55] Timothée Lacroix, Guillaume Obozinski und Nicolas Usunier. „Tensor Decompositions for Temporal Knowledge Base Completion“. In: *International Conference on Learning Representations*. 2020.
- [56] Adam Lerer, Ledell Wu, Jiajun Shen, Timothee Lacroix, Luca Wehrstedt, Abhijit Bose und Alex Peysakhovich. „Pytorch-biggraph: A large scale graph embedding system“. In: *Proceedings of Machine Learning and Systems 1 (2019)*, S. 120–131.
- [57] Shuangyin Li, Heng Wang, Rong Pan und Mingzhi Mao. „MemoryPath: A deep reinforcement learning framework for incorporating memory component into knowledge graph reasoning“. In: *Neurocomputing* 419 (2021), S. 273–286. DOI: 10.1016/j.neucom.2020.08.032.
- [58] Zixuan Li, Xiaolong Jin, Wei Li, Saiping Guan, Jiafeng Guo, Huawei Shen, Yuanzhuo Wang und Xueqi Cheng. „Temporal Knowledge Graph Reasoning Based on Evolutional Representation Learning“. In: *CoRR abs/2104.10353 (2021)*. arXiv: 2104.10353.
- [59] L. Della Libera, V. Golkov, Y. Zhu, A. Mielke und D. Cremers. „Deep Learning for 2D and 3D Rotatable Data: An Overview of Methods“. In: *arXiv preprint arXiv:1910.14594 (2019)*. arXiv: 1910.14594.
- [60] Yi Luan, Luheng He, Mari Ostendorf und Hannaneh Hajishirzi. „Multi-Task Identification of Entities, Relations, and Coreference for Scientific Knowledge Graph Construction“. In: *Proc. Conf. Empirical Methods Natural Language Process. (EMNLP)*. 2018.
- [61] Yunpu Ma und Volker Tresp. „Quantum Machine Learning on Knowledge Graphs“. In: Bd. 2. 03. *ACM Transactions on Quantum Computing*, 2021.
- [62] Yunpu Ma, Volker Tresp, Liming Zhao und Yuyi Wang. „Variational quantum circuit model for knowledge graph embedding“. In: Bd. 2. 7-8. *Wiley Online Library*, 2019, S. 1800078.
- [63] Diego Marcheggiani und Laura Perez-Beltrachini. „Deep Graph Convolutional Encoders for Structured Data to Text Generation“. In: *Proceedings of the 11th International Conference on Natural Language Generation*. Tilburg University, The Netherlands: Association for Computational Linguistics, Nov. 2018, S. 1–9. DOI: 10.18653/v1/W18-6501.
- [64] Maximilian Mozes, Martin Schmitt, Vladimir Golkov, Hinrich Schütze und Daniel Cremers. „Scene Graph Generation for Better Image Captioning?“ In: *CoRR abs/2109.11398 (2021)*. arXiv: 2109.11398.
- [65] P. Müller, V. Golkov, V. Tomassini und D. Cremers. „Rotation-Equivariant Deep Learning for Diffusion MRI“. In: *arXiv preprint (2021)*. arXiv: 2102.06942 [cs.CV].
- [66] P. Müller, V. Golkov, V. Tomassini und D. Cremers. „Rotation-Equivariant Deep Learning for Diffusion MRI (short version)“. In: *International Society for Magnetic Resonance in Medicine (ISMRM) Annual Meeting*. 2021.

- [67] M Naeyaert, V Golkov, D Cremers, J Sijbers und M Verhoye. „Faster and better HARDI using FSE and holistic reconstruction“. In: *International Society for Magnetic Resonance in Medicine (ISMRM) Annual Meeting*. 2021.
- [68] M. Naeyaert, J. Aelterman, J. Van Audekerke, V. Golkov, D. Cremers, A. Pižurica, J. Sijbers und M. Verhoye. „Accelerating in vivo fast spin echo high angular resolution diffusion imaging with an isotropic resolution in mice through compressed sensing“. In: *Magnetic Resonance in Medicine* 85.3 (2020), S. 1397–1413. DOI: 10.1002/mrm.28520. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.28520>.
- [69] Fabio Petroni, Tim Rocktäschel, Sebastian Riedel, Patrick S. H. Lewis, Anton Bakhtin, Yuxiang Wu und Alexander H. Miller. „Language Models as Knowledge Bases?“ In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP 2019, Hong Kong, China, November 3-7, 2019*. Hrsg. von Kentaro Inui, Jing Jiang, Vincent Ng und Xiaojun Wan. Association for Computational Linguistics, 2019, S. 2463–2473. DOI: 10.18653/v1/D19-1250.
- [70] Sanjay Purushotham, Chuizheng Meng, Zhengping Che und Yan Liu. „Benchmarking deep learning models on large healthcare datasets“. In: *Journal of biomedical informatics* 83 (2018), S. 112–134.
- [71] Benjamin Recht, Christopher Re, Stephen Wright und Feng Niu. „Hogwild!: A lock-free approach to parallelizing stochastic gradient descent“. In: *Advances in neural information processing systems* 24 (2011).
- [72] Leonardo F. R. Ribeiro, Claire Gardent und Iryna Gurevych. „Enhancing AMR-to-Text Generation with Dual Graph Representations“. In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. Hong Kong, China: Association for Computational Linguistics, Nov. 2019, S. 3183–3194. DOI: 10.18653/v1/D19-1314.
- [73] Leonardo F. R. Ribeiro, Martin Schmitt, Hinrich Schütze und Iryna Gurevych. „Investigating Pretrained Language Models for Graph-to-Text Generation“. In: *CoRR* abs/2007.08426 (2020). arXiv: 2007.08426.
- [74] Leonardo F. R. Ribeiro, Yue Zhang, Claire Gardent und Iryna Gurevych. „Modeling Global and Local Node Contexts for Text Generation from Knowledge Graphs“. In: *CoRR* abs/2001.11003 (2020). arXiv: 2001.11003.
- [75] Matthias Ruebner u. a. *Routine genotyping within PRAEGNANT breast cancer registry - Automated mutation identification from routine genotyping reports, physician letters and pathological reports*. 2022.
- [76] Afshin Sadeghi, Diego Collarana, Damien Graux und Jens Lehmann. „Embedding Knowledge Graphs Attentive to Positional and Centrality Qualities“. In: *European Conference on Machine Learning and Data Mining, ECML PKDD 2021*. 2021.
- [77] Afshin Sadeghi, Damien Graux, Hamed Shariat Yazdi und Jens Lehmann. „MDE: Multiple Distance Embeddings for Link Prediction in Knowledge Graphs“. In: *ECAI*. 2020.
- [78] Afshin Sadeghi und Jens Lehmann. „Linking Physicians to Medical Research Results via Knowledge Graph Embeddings and Twitter“. In: *Machine Learning and Knowledge Discovery in Databases*. Hrsg. von Peggy Cellier und Kurt Driessens. Cham: Springer International Publishing, 2020, S. 622–630. ISBN: 978-3-030-43823-4.
- [79] Afshin Sadeghi, Hirra Abdul Malik, Diego Collarana und Jens Lehmann. „Relational Pattern Benchmarking on the Knowledge Graph Link Prediction Task“. In: *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*. 2021.
- [80] Afshin Sadeghi, Xhulia Shahini, Martin Schmitz und Jens Lehmann. „BenchEmbedd: A FAIR Benchmarking tool for knowledge graph Embeddings“. In: *SEMANTiCS 2021 Demo Session*. <http://ceur-ws.org/>, 2021.

- [81] Abhishek Saroha, Marvin Eisenberger, Tarun Yenamandra und Daniel Cremers. „Implicit Shape Completion via Adversarial Shape Priors“. In: *arXiv preprint arXiv:2204.10060* (2022).
- [82] Martin Schmitt, Leonardo F. R. Ribeiro, Philipp Dufter, Iryna Gurevych und Hinrich Schütze. „Modeling Graph Structure via Relative Position for Better Text Generation from Knowledge Graphs“. In: *CoRR abs/2006.09242* (2020). arXiv: 2006.09242.
- [83] Martin Schmitt, Leonardo F. R. Ribeiro, Philipp Dufter, Iryna Gurevych und Hinrich Schütze. „Modeling Graph Structure via Relative Position for Text Generation from Knowledge Graphs“. In: *Proceedings of the Fifteenth Workshop on Graph-Based Methods for Natural Language Processing (TextGraphs-15)*. Mexico City, Mexico: Association for Computational Linguistics, Juni 2021, S. 10–21. DOI: 10.18653/v1/2021.textgraphs-1.2.
- [84] Martin Schmitt und Hinrich Schütze. „Continuous Entailment Patterns for Lexical Inference in Context“. In: *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 7-11 November, 2021*. Association for Computational Linguistics, 2021, S. 6952–6959. DOI: 10.18653/v1/2021.emnlp-main.556.
- [85] Martin Schmitt und Hinrich Schütze. „Language Models for Lexical Inference in Context“. In: *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*. Online: Association for Computational Linguistics, Apr. 2021, S. 1267–1280. DOI: 10.18653/v1/2021.eacl-main.108.
- [86] Martin Schmitt und Hinrich Schütze. „SherLLiC: A Typed Event-Focused Lexical Inference Benchmark for Evaluating Natural Language Inference“. In: *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*. Association for Computational Linguistics, 2019, S. 902–914. DOI: 10.18653/v1/p19-1086.
- [87] Martin Schmitt und Hinrich Schütze. „Unsupervised Text Generation from Structured Data“. In: *CoRR abs/1904.09447* (2019). arXiv: 1904.09447.
- [88] Martin Schmitt, Sahand Sharifzadeh, Volker Tresp und Hinrich Schütze. „An Unsupervised Joint System for Text Generation from Knowledge Graphs and Semantic Parsing“. In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP 2020, Online, November 16-20, 2020*. Association for Computational Linguistics, 2020, S. 7117–7130. DOI: 10.18653/v1/2020.emnlp-main.577.
- [89] Marina Sedinkina, Martin Schmitt und Hinrich Schütze. „Domain Adaptation for Sparse-Data Settings: What Do We Gain by Not Using Bert?“ In: *CoRR abs/2203.16926* (2022).
- [90] S. Sharifzadeh, S. Baharlou, M. Berrendorf, R. Koner und V. Tresp. „Improving Visual Relation Detection using Depth Maps“. In: *2020 25th International Conference on Pattern Recognition (ICPR)*. Los Alamitos, CA, USA: IEEE Computer Society, Jan. 2021, S. 3597–3604. DOI: 10.1109/ICPR48806.2021.9412945.
- [91] Sahand Sharifzadeh, Sina Moayed Baharlou, Martin Schmitt, Hinrich Schütze und Volker Tresp. „Improving Scene Graph Classification by Exploiting Knowledge from Texts“. In: *Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence, IAAI 2022, The Twelveth Symposium on Educational Advances in Artificial Intelligence, EAAI 2022 Virtual Event, February 22 - March 1, 2022*. AAAI Press, 2022, S. 2189–2197.
- [92] Yelong Shen, Jianshu Chen, Po-Sen Huang, Yuqing Guo und Jianfeng Gao. „M-Walk: Learning to Walk over Graphs using Monte Carlo Tree Search“. In: *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, December 3-8, 2018, Montréal, Canada*. Hrsg. von Samy Bengio, Hanna M. Wallach, Hugo Larochelle, Kristen Grauman, Nicolò Cesa-Bianchi und Roman Garnett. 2018, S. 6787–6798.

- [93] Taylor Shin, Yasaman Razeghi, Robert L. Logan IV, Eric Wallace und Sameer Singh. „AutoPrompt: Eliciting Knowledge from Language Models with Automatically Generated Prompts“. In: *CoRR* abs/2010.15980 (2020). arXiv: 2010.15980.
- [94] Suprosanna Shit u. a. „Relationformer: A Unified Framework for Image-to-Graph Generation“. In: *CoRR* abs/2203.10202 (2022). DOI: 10.48550/arXiv.2203.10202. arXiv: 2203.10202.
- [95] Zhiqing Sun, Zhi-Hong Deng, Jian-Yun Nie und Jian Tang. „RotatE: Knowledge Graph Embedding by Relational Rotation in Complex Space“. In: *International Conference on Learning Representations*. 2019.
- [96] Muhammed Talo. „Pneumonia detection from radiography images using convolutional neural networks“. In: *2019 27th Signal Processing and Communications Applications Conference (SIU)*. IEEE. 2019, S. 1–4.
- [97] TBN. *Feasibility and acceptance of the use of audio recordings from routine physician-patient dialogues for implementation into machine learning algorithms*. 2022.
- [98] Komal K. Teru, Etienne G. Denis und William L. Hamilton. „Inductive Relation Prediction by Subgraph Reasoning“. In: *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event*. Bd. 119. Proceedings of Machine Learning Research. PMLR, 2020, S. 9448–9457.
- [99] Volker Tresp und Yunpu Ma. „The tensor memory hypothesis“. In: *arXiv preprint arXiv:1708.02918* (2017).
- [100] Théo Trouillon, Johannes Welbl, Sebastian Riedel, Éric Gaussier und Guillaume Bouchard. „Complex embeddings for simple link prediction“. In: *International Conference on Machine Learning*. 2016, S. 2071–2080.
- [101] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò und Yoshua Bengio. „Graph Attention Networks“. In: *6th International Conference on Learning Representations* (2017).
- [102] Rui Wang, Bicheng Li, Shengwei Hu, Wenqian Du und Min Zhang. „Knowledge Graph Embedding via Graph Attenuated Attention Networks“. In: *IEEE Access* 8 (2020), S. 5212–5224.
- [103] Andrew Gordon Wilson, Zhiting Hu, Ruslan Salakhutdinov und Eric P Xing. „Deep kernel learning“. In: *Artificial intelligence and statistics*. PMLR. 2016, S. 370–378.
- [104] Zhiliang Wu, Yinchong Yang, Peter A Fashing und Volker Tresp. „Uncertainty-Aware Time-to-Event Prediction using Deep Kernel Accelerated Failure Time Models“. In: *Machine Learning for Healthcare Conference*. PMLR. 2021, S. 54–79.
- [105] Zhiliang Wu, Yinchong Yang, Jindong Gu und Volker Tresp. „Quantifying predictive uncertainty in medical image analysis with deep kernel learning“. In: *2021 IEEE 9th International Conference on Healthcare Informatics (ICHI)*. IEEE. 2021, S. 63–72.
- [106] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang und Philip S. Yu. „A Comprehensive Survey on Graph Neural Networks“. In: *IEEE Transactions on Neural Networks and Learning Systems* 32.1 (2021), S. 4–24. DOI: 10.1109/TNNLS.2020.2978386.
- [107] Keyulu Xu, Weihua Hu, Jure Leskovec und Stefanie Jegelka. „How Powerful are Graph Neural Networks?“ In: *CoRR* (2018). arXiv: 1810.00826.
- [108] Bishan Yang, Wen-tau Yih, Xiaodong He, Jianfeng Gao und Li Deng. „Embedding Entities and Relations for Learning and Inference in Knowledge Bases“. In: *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*. 2015.
- [109] Chengxuan Ying, Tianle Cai, Shengjie Luo, Shuxin Zheng, Guolin Ke, Di He, Yanming Shen und Tie-Yan Liu. „Do Transformers Really Perform Bad for Graph Representation?“ In: *arXiv preprint arXiv:2106.05234* (2021).

- [110] Cunchao Zhu, Muhao Chen, Changjun Fan, Guangquan Cheng und Yan Zhang. „Learning from History: Modeling Temporal Knowledge Graphs with Sequential Copy-Generation Networks“. In: *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021*. AAAI Press, 2021, S. 4732–4740.

7 Anhang: Quantitative Forschungsergebnisse

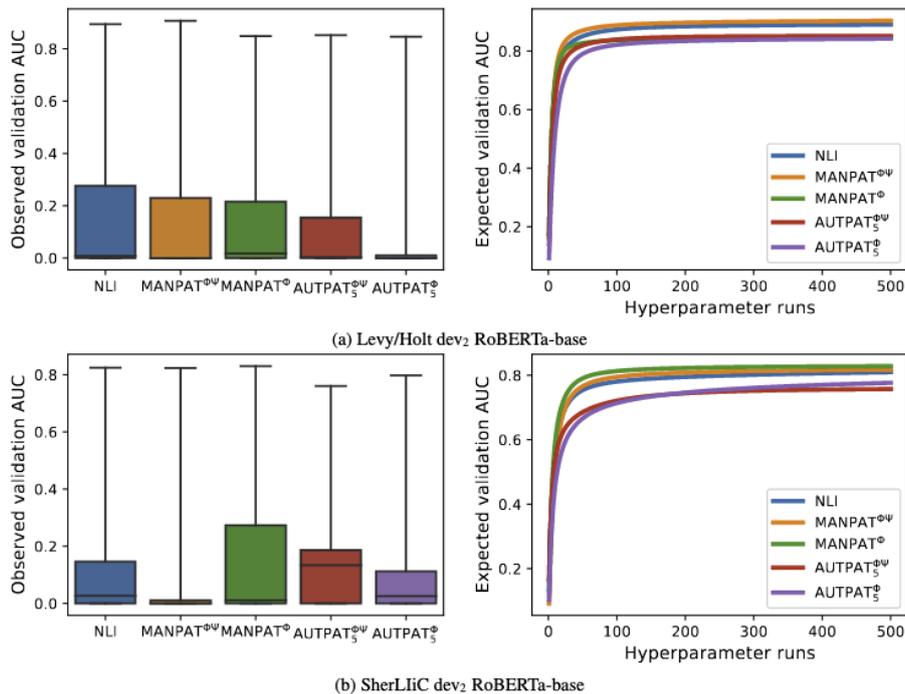


Abbildung 7: Performanz für die lexical inference in context Aufgabe der von uns entwickelten Methoden auf verschiedenen Datensätzen und über verschiedene Hyperparameterläufe (links). Erwartete Validierungsleistung pro Anzahl der getesteten Hyperparameterkonfigurationen (rechts). Die Leistung wird als Fläche unter der Precision-Recall-Kurve für Precision-Werte $\geq 0,5$ gemessen. Die Boxen repräsentieren 75 % der jeweiligen Datenpunkte.

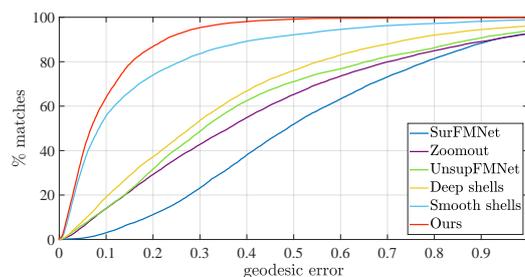


Abbildung 8: Vergleich von Methoden zur Berechnung von Korrespondenzen zwischen als Graphen (Dreiecksgitter) repräsentierten Objekten des SHREC20-Datensatzes, auf den nur unüberwachte Methoden anwendbar sind. Unsere Methoden NeuroMorph (“Ours”) [19], Deep Shells [20] und Smooth Shells [18] liefern bessere Ergebnisse (kleinere Fehler) als die besten bisherigen Methoden.

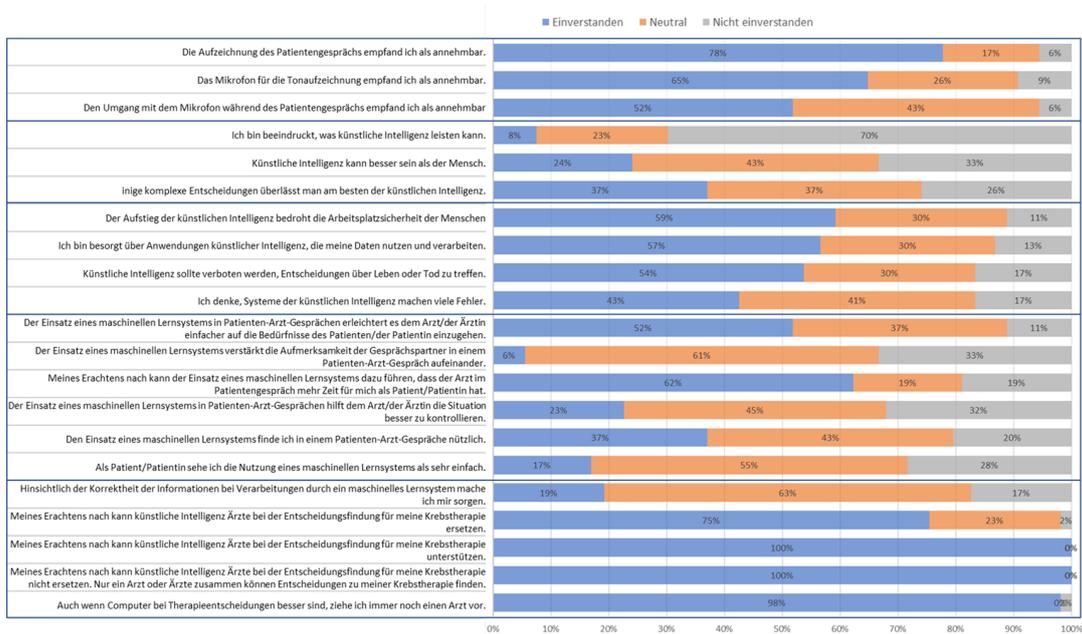


Abbildung 13: Präliminäre Ergebnisse der Patientinnen-Befragung im Rahmen der COMMITMENT Studie

Tabelle 11: Zusammenfassung der Benchmarkingergebnisse (Tabelle entnommen aus [3]).

<i>Interaktionsmodelle</i>	
RotatE	Unter den zehn besten Interaktionsmodellen bei allen Datensätze.
MuRE	Unter den zehn besten Interaktionsmodelle bei den Datensätzen WN18RR, FB15K-237, and YAGO3-10.
ConvE	Unter den zehn besten Interaktionsmodelle bei dem Datensätzen Kinships und FB15K-237 (wurde nicht auf YAGO3-10 evaluiert).
ComplEx	Unter den zehn besten Interaktionsmodelle bei den Datensätzen Kinships und YAGO3-10.
TuckER	Unter den zehn besten Interaktionsmodelle bei den Datensätzen Kinships, und FB15K-237 (wurde nicht auf YAGO3-10 evaluiert).
DistMult	Unter den zehn besten Interaktionsmodelle bei den Datensatz FB15K-237.
QuatE	Unter den zehn besten Interaktionsmodelle bei dem Datensatz YAGO3-10.
TransE	Unter den zehn besten Interaktionsmodelle bei dem Datensatz WN18RR.
Strucured Embedding	Unter den zehn besten Interaktionsmodelle bei dem Datensatz Kinships.
<i>Kostenfunktionen</i>	
BCEL	Unter den zehn besten Kostenfunktionen bei allen Datensätze.
NSSAL	Unter den zehn besten Kostenfunktionen bei allen Datensätze.
SPL	Unter den zehn besten Kostenfunktionen bei allen Datensätzen.
CEL	Unter den zehn besten Kostenfunktionen bei den Datensätzen Kinships und FB15K-237 (wurde nicht auf YAGO3-10 evaluiert).
MRL	Unter den zehn besten Kostenfunktionen bei dem Datensatz Kinships.
<i>Trainingsansätze</i>	
SLCWA	Unter den zehn besten Trainingsansätzen bei allen Datensätze.
LCWA	Unter den zehn besten Trainingsansätzen bei den Datensätze Kinships, WN18RR und FB15K-237 (wurde nicht auf YAGO3-10 evaluiert).
<i>Explizites Modellieren von Inversen Relationen</i>	
	Ist in der Regel in Kombination mit dem LCWA-Trainingsansatz von Vorteil.
<i>Konfigurationen</i>	
Performanz	Eine geeignete Kombination aus Interaktionsmodell, Trainingsansatz, Kostenfunktion und Wahl der expliziten Modellierung inverser Relationen ist entscheidend für die Leistung, z. B. kann TransE mit mehreren State-of-The-Art-Interaktionsmodellen auf WN18RR konkurrieren, wenn eine geeignete Konfiguration gewählt wird. Es gibt keine einzelne Konfiguration, die für alle Datensätze am besten funktioniert.
Varianz	Einige Interaktionsmodelle weisen eine hohe Varianz über verschiedene Konfigurationen hinweg auf, z. B. RotatE bei YAGO3-10.
<i>Allgemeines</i>	
SOTA	Für WN18RR erzielen wir mit einer RotatE-Konfiguration (zusammen mit Graph Attenuated Attention Networks [102]) neue State-of-The-Art Hits@10-Ergebnisse (60,09% Hits@10). Zusätzlich hinaus haben wir eine TransE-Konfiguration ermittelt, die die meisten der veröffentlichten SOTA-Ergebnisse übertrifft (56,98% Hits@10). Auf der Grundlage unserer Ergebnisse möchten wir herausheben, dass die Hyperparameter-Werte für die vielversprechendsten Konfigurationen für die weiteren Benchmarking-Datensätze weiter untersucht werden sollten.
Verbesserung	Für ConvE (56,33% im Vergleich zu 52,00% [15]), MuRE (57,90% im Vergleich zu 55,50% [7]) und TuckER (56,09% im Vergleich zu 52,6% [8]) übertreffen wir die ursprünglich publizierten Ergebnisse durch die Auswahl geeigneter Konfigurationen und Hyperparameter auf WN18RR.

Model	WN18RR			FB15k-237		
	MR	MRR	Hit@10	MR	MRR	Hit@10
DistMult [108]	5110	0.43	0.49	254	0.241	0.419
ComplEx [100]	5261	0.44	0.51	339	0.247	0.428
ConvE [16]	5277	0.46	0.48	246	0.316	0.491
RotatE [95]	3340	0.476	0.571	177	0.338	0.533
MDE	2629	0.457	0.536	189	0.288	0.484
MDE _{NN}	3165	0.432	0.531	-	-	-
MDE _{adv}	3219	0.458	0.560	203	0.344	0.531

Tabelle 12: Verbindungsvorhersageergebnisse von MDE auf WN18RR und FB15k-237 im Vergleich zum Stand der Technik. Die besten sind fett gedruckt.

Model	Countries(AUC-PR)		
	S1	S2	S3
DistMult [108]	1.00 ± 0.00	0.72 ± 0.12	0.52 ± 0.07
ComplEx [100]	0.97 ± 0.02	0.57 ± 0.10	0.43 ± 0.07
ConvE [16]	1.00 ± 0.00	0.99 ± 0.01	0.86 ± 0.05
RotatE [95]	1.00 ± 0.00	1.00 ± 0.00	0.95 ± 0.00
MDE	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00

Tabelle 13: Ergebnisse der MDE für die Countries Datensätze.

Tabelle 14: Vergleich der Ergebnisse von GFA-NN mit den aktuellen WN18RR und FB15k-237. Die besten Ergebnisse sind fett gedruckt.

Model	WN18RR			FB15k-237		
	MR	MRR	Hit@10	MR	MRR	Hit@10
ComplEx-N3	-	0.48	0.57	-	0.37	0.56
QuatE ²	-	0.482	0.572	-	0.366	0.556
TuckER	-	0.470	0.526	-	0.358	0.544
COMPGCN	3533	0.479	0.546	197	0.355	0.535
RotatE	3340	0.476	0.571	177	0.338	0.533
MDE	3219	0.458	0.536	203	0.344	0.531
GFA-NN	3390	0.486	0.575	186	0.338	0.522

Tabelle 15: MRR-Ergebnisse von GFA-NN und anderen neueren Modellen für den großen biologischen Wissensgraphen ogbl-biokg. (Ergebnisse früherer Modelle stammen von [41].)

Method	Validation	Test
TRANSE	0.7456	0.7452
DISTMULT	0.8055	0.8043
COMPLEX	0.8105	0.8095
ROTATE	0.7997	0.7989
GFA-NN	0.9011	0.9011

Tabelle 16: Hit@10-Ergebnisse von GFA-NN für induktive Datensätze.

Model	WN18RR- v_3 -ind	WN18RR- v_4 -ind	NELL-995- v_1 -ind	NELL-995- v_4 -ind
NeuralLP	0.4618	0.6713	0.4078	0.8058
DRUM	0.4618	0.6713	0.5950	0.8058
RuleN	0.5339	0.7159	0.5950	0.6135
GraiL	0.5843	0.7341	0.5950	0.7319
GFA-NN	0.5893	0.7355	0.9500	0.7722

Tabelle 17: Veränderung der Link-Prediction Ergebnisse nach Zerlegung mehrdeutiger Relationen mit dem V-Coder auf dem *FB15K-237*, dem *NELL-955* und dem *YAGO3-10* Datensätzen.

	FB15k-237 & FB15K-237-10splits						NELL-995 & NELL-995-10splits						YAGO3-10 & YAGO3-10-1split					
	MRR		Hits@10		Hits@1		MRR		Hits@10		Hits@1		MRR		Hits@10		Hits@1	
DistMult	.250	.250	.392	.394	.178	.178	.221	.247	.321	.352	.160	.179	.316	.317	.522	.523	.212	.213
TransE	.287	.288	.476	.477	.190	.191	.301	.302	.429	.431	.222	.222	.264	.265	.469	.469	.158	.160
RotatE	.251	.251	.426	.427	.164	.163	.334	.338	.414	.423	.288	.287	.143	.142	.252	.254	.083	.084
Complex	.260	.264	.435	.442	.174	.174	.267	.285	.356	.383	.216	.227	.409	.407	.605	.609	.310	.306
Analogy	.301	.302	.467	.468	.219	.219	.142	.150	.266	.277	.091	.092	.345	.345	.550	.553	.241	.241
Simple	.283	.285	.445	.447	.202	.204	.223	.223	.318	.323	.173	.173	.481	.498	.656	.663	.389	.411

Tabelle 18: Vergleich von SEA mit dem Stand der Technik

ZINC			OGBG-MOLHIV			PATTERN		
Model	#params.	MAE	Model	#params.	%ROC-AUC	Model	#params.	% ACC
GCN [48]	505K	0.367	GCN-GRAPHNORM [48]	526K	76.06	GCN [48]	500K	71.892
GIN [107]	509K	0.526	GIN-VN [107]	3.3M	77.80	GIN [107]	100K	85.590
GAT [101]	531K	0.384	DGN [9]	114K	79.05	GAT [101]	526K	78.271
SAN [54]	508K	0.139	Graphormer-FLAG [109]	47.0M	80.51	GraphSage [31]	101K	50.516
Graphormer-SLIM [109]	489K	0.122	Vanilla GTL	386K	78.06	SAN [54]	454K	86.581
Vanilla GTL	83K	0.227	SEA-GTL	347K	79.53	Vanilla GTL	82K	84.691
SEA-GTL	347K	0.212	SEA-AGGREGATED	133K	80.18	SEA-GTL	132K	85.006
SEA-AGGREGATED	112K	0.215	SEA-2-HOP	511K	80.01	SEA-AGGREGATED	69K	57.557
SEA-2-HOP	430K	0.159	SEA-2-HOP-AUG	594K	79.08	SEA-2-HOP	48K	86.768
SEA-2-HOP-AUG	709K	0.189				SEA-2-HOP-AUG	152K	86.673

(a) ZINC [43]

(b) ogbg-molhiv [42]

(c) PATTERN [1]

Tabelle 19: Performanz von AppTek auf den *ICEWS14*, *ICEWS05-15*, and *YAGO15K* Datensätzen.

↓ Model Dataset →	ICEWS14			ICEWS05-15			YAGO15K		
	MRR	Hits@10	Hits@1	MRR	Hits@10	Hits@1	MRR	Hits@10	Hits@1
TNTComplex [55]	0.2691	0.5266	0.1701	0.2398	0.4775	0.1327	<u>0.5852</u>	0.9699	<u>0.3168</u>
DE-Simple [27]	0.5074	<u>0.7573</u>	0.3882	0.4289	0.6551	0.3162	0.2749	0.4228	0.1852
DE-TransE [27]	0.2068	0.3119	0.1364	0.1928	0.3486	0.1074	0.3906	0.4827	0.3120
DE-DistMult [27]	0.4613	0.7370	0.3261	0.3197	0.5243	0.2129	0.2459	0.4298	0.1457
RE-GCN [58]	0.4756	0.7524	<u>0.3535</u>	<u>0.4750</u>	<u>0.7483</u>	0.3524	-	-	-
APPTeK	<u>0.4889</u>	0.7749	0.3421	0.4844	0.8246	<u>0.3222</u>	0.6041	<u>0.9547</u>	0.3537

Tabelle 20: Evaluation unserer Methoden zur Textgenerierung aus Wissensgraphen [88].

graph → text	Visual Genome						WebNLG					
	BLEU		METEOR		CHRF++		BLEU		METEOR		CHRF++	
	val	test	val	test	val	test	val	test	val	test	val	test
R _{graph} →text	5.9	5.9	28.2	28.1	43.4	43.3	18.3	18.3	33.5	33.6	55.0	55.2
Ours w/ sampled noise	19.8	19.5	31.4	31.2	50.9	50.7	39.1	37.7	35.4	35.5	61.9	62.1
Ours w/ composed noise	23.2	23.2	33.0	32.9	53.7	53.6	30.8	30.5	30.2	30.0	53.1	52.8
Ours supervised	26.5	26.4	32.3	32.2	53.7	53.6	35.1	34.4	39.6	39.5	64.1	64.0

Tabelle 21: Ergebnisse zur Graph Text Generierung nach [83].

(a) Agenda Datensatz					(b) WebNLG Datensatz				
	BLEU	METEOR	CHRF++	#P		BLEU	METEOR	CHRF++	#P
Ours	17.80 \pm 0.31	22.07 \pm 0.23	45.43 \pm 0.39	36.3	Ours	61.15 \pm 0.22	43.38 \pm 0.17	75.43 \pm 0.19	5.3
GT	14.30 \pm 1.01	18.80 \pm 0.28	–	–	UPF-FORGe	40.88	40.00	–	–
GT+RBS	15.1 \pm 0.97	19.5 \pm 0.29	–	–	Melbourne	54.52	41.00	70.72	–
CGE-LW	18.01 \pm 0.14	22.34 \pm 0.07	46.69 \pm 0.17	69.8	Adapt	60.59	44.00	76.01	–
					Graph Conv.	55.90	39.00	–	4.9
					GTR-LSTM	58.60	40.60	–	–
					E2E GRU	57.20	41.00	–	–
					CGE-LW-LG	63.69 \pm 0.10	44.47 \pm 0.12	76.66 \pm 0.10	10.4

Tabelle 22: Trainingszeit und Geschwindigkeit von vortrainierten Sprachmodellen gegenüber Bag-of-Words Repräsentationen auf dem BioNLP und dem Twitter Datensätzen für Part-Of-Speech tagging. t_c steht für die Trainingszeit auf einer CPU und t_g steht für die Trainingszeit mit GPU in Sekunden. Für SENNA (W) sind beide Zeiten identisch, da hier keine GPU benötigt wird.

(a) Twitter					(b) BioNLG				
model	t_c	t_g	acc	acc/ t_g	model	t_c	t_g	acc	acc/ t_g
ALBERT	56985	7598	91.7	0.0120	ALBERT	55345	15813	99.8	0.0063
RoBERTa	42012	5524	94.9	0.0172	RoBERTa	37638	10912	99.8	0.0091
DistilBERT	22842	3013	94.1	0.0312	DistilBERT	20844	7149	99.9	0.0139
BERT	41472	4597	94.8	0.0206	BERT	37260	8477	99.9	0.0117
Flair	2610	348	92.5	0.32	Flair	1883	538	99.0	0.18
W (SENNA)	7.5	7.5	88.8	11.84	BioBERT	23520	6720	98.3	0.0146
W+Subsp	83.9	83.9	88.5	1.05	W (SENNA)	5.7	5.7	94.3	16.54
					W+Subsp	41.96	41.96	95.2	2.27

Model	#QP	WD20K (66) V1					WD20K (66) V2				
		AMR(%)	MRR(%)	H@1(%)	H@5(%)	H@10(%)	AMR(%)	MRR(%)	H@1(%)	H@5(%)	H@10(%)
BLP	0	34.96	2.10	0.45	2.29	4.44	45.29	1.56	0.27	1.88	3.35
CompGCN	0	35.99	5.80	2.38	8.93	12.79	47.24	<u>2.56</u>	<u>1.17</u>	3.07	4.46
QBLP	0	35.30	3.69	1.30	4.85	7.14	42.48	0.94	0.08	0.79	1.82
StarE	2	37.72	6.84	3.24	9.71	<u>13.44</u>	52.78	2.62	0.74	<u>3.55</u>	<u>5.78</u>
StarE	4	38.91	<u>6.40</u>	<u>2.83</u>	<u>8.94</u>	13.39	51.93	5.06	2.09	7.34	9.82
StarE	6	38.20	6.87	3.46	8.98	13.57	47.01	4.42	2.04	5.73	8.97
QBLP	2	<u>30.37</u>	3.70	1.26	4.90	8.14	53.67	1.39	0.41	1.66	2.59
QBLP	4	30.84	3.20	0.90	4.00	7.14	37.10	2.08	0.38	2.20	4.92
QBLP	6	26.34	4.34	1.66	5.53	9.25	<u>39.12</u>	1.95	0.41	2.15	4.10

Tabelle 23: Ergebnisse der Link Prediction nach [2] auf dem FI WD20K (66) (V1 & V2) Datensatz. #QP beschreibt die Anzahl der qualifizierenden Paare, die für jedes Statement verwendet wurden. Das beste Ergebnis ist fett gedruckt und das zweitbeste Ergebnis unterstrichen.

Tabelle 24: Erwartungswert und Standardabweichung von Precision, Recall und F1-Wert der vorhergesagten Fakten aus Text über vier zufälligen Aufteilungen in Test und Trainingsdaten. Die Ergebnisse der Vergleichsverfahren wurden mit Hilfe der von den Autoren offiziell zur Verfügung gestellten Implementierung generiert.

Method	Precision		Recall		F1	
	1%	10%	1%	10%	1%	10%
R _{text→graph}	1.92 \pm 0.00	1.86 \pm 0.01	1.87 \pm 0.00	1.81 \pm 0.01	1.89 \pm 0.00	1.84 \pm 0.01
SSGP	14.86 \pm 0.01	14.52 \pm 0.02	18.47 \pm 0.01	18.05 \pm 0.02	16.47 \pm 0.01	16.09 \pm 0.02
CopyNet	29.20 \pm 0.13	30.77 \pm 0.49	27.19 \pm 0.28	29.79 \pm 0.29	28.16 \pm 0.21	30.27 \pm 0.34
T5	33.37 \pm 0.11	33.81 \pm 0.08	31.06 \pm 0.18	32.45 \pm 0.33	32.17 \pm 0.13	33.12 \pm 0.16

Tabelle 25: Transductive Setting Datasets

Transductive		Train	Test	Valid	Total Entities	Total Relations	% of Original Dataset
FB15K	Symmetry	5781	1399	1416	2823	52	1.452
	Anti-Symmetry	20711	128	109	2471	143	3.537
	Inverse	31332	750	750	10988	696	5.544
	Inference	70226	104	111	10500	377	11.895
WN18	Symmetry	1449	362	363	2030	5	1.436
	AntiSymmetry	6366	190	168	3393	15	4.441
	Inverse	4364	750	750	5765	17	3.873
	Inference	2027	12	10	3009	18	1.353

Tabelle 26: Head-Tail Inductive Setting Datasets

Head-Tail Inductive		Train	Test	Valid	Total Entities	Total Relations	% of Original Dataset
FB15K	Symmetry	5621	989	990	3632	52	1.283
	Anti-Symmetry	15404	10795	10795	9241	470	6.247
	Inverse	24176	4701	4701	12065	794	5.670
	Inference	13845	5898	5636	9671	665	4.286
WN18	Symmetry	1630	185	186	2447	5	1.322
	AntiSymmetry	30000	5603	5603	23786	18	27.217
	Inverse	5421	621	621	7843	17	4.401
	Inference	462	84	85	815	15	0.417

Tabelle 27: 50% Inductive Setting Datasets

Percentage Based Inductive		Train	Test	Valid	Total Entities	Total Relations	% of Original Dataset
FB15K	Symmetry	4677	445	444	3219	51	0.940
	Anti-Symmetry	14904	11603	11608	9911	472	6.436
	Inverse	6124	779	779	4958	600	1.297
	Inference	5840	5249	5263	7031	446	2.761
WN18	Symmetry	2009	253	253	3285	5	1.661
	AntiSymmetry	22208	5330	5329	21666	18	21.709
	Inverse	7613	678	678	10785	17	5.924
	Inference	685	58	64	1235	15	0.533

Tabelle 28: Inductive Setting Datasets

Inductive		Train	Test	Valid	Total Entities	Total Relations	% of Original Dataset
FB15K	Symmetry	4254	542	542	3447	51	0.901
	Anti-Symmetry	12930	3494	3884	8304	433	3.429
	Inverse	4753	2568	2568	7745	641	1.670
	Inference	3489	2824	2745	6083	611	1.530
WN18	Symmetry	2322	272	272	4344	5	1.893
	AntiSymmetry	16650	4698	4697	20552	18	17.203
	Inverse	8728	903	904	13842	17	6.958
	Inference	844	99	99	1639	15	0.688

Tabelle 29: Mean AUC-PR performance of the LP methods.

Type of dataset	WN18				FB15K			
	Metric (AUC-PR)							
	Inductive	Transductive	Head/Tail Ratio	Percentage Based(50%)	Inductive	Transductive	Head/Tail Ratio	Percentage Based(50%)
DistMult	0.5015	0.7203	0.5456	0.6634	0.4995	0.6255	0.5468	0.5755
TransE	0.4688	0.6239	0.5169	0.5643	0.4489	0.6685	0.5969	0.5266
RotatE	0.4542	0.7497	0.5840	0.6231	0.4282	0.6697	0.5829	0.5336
MDE	0.4865	0.9624	0.6426	0.5952	0.4474	0.8709	0.5937	0.5310
GraL	0.5000	0.9960	0.5013	0.6433	0.5004	0.9975	0.6105	0.5926
CompGCN	0.5531	0.9269	0.5713	0.5596	0.6368	0.9976	0.6473	0.6768
QuatE	0.4426	0.6518	0.6518	0.7382	0.5867	0.9618	0.6563	0.8208

Kurzbericht

MLwin — Maschinelles Lernen in Wissensgraphen

1. Dezember 2022

Projektlaufzeit: 01. Oktober 2018 – 31. Januar 2022
Allgemeines Förderkennzeichen: 01IS18050

1 Aufgabenstellung

Intelligente Systeme wie das menschliche Gehirn verarbeiten Informationen sowohl auf einer sensorischen subsymbolischen Ebene, als auch auf einer abstrakteren diskreten symbolischen Ebene. Beide Ebenen wechselwirken miteinander: aus subsymbolischen Quellen wird abstraktes Wissen extrahiert, und abstraktes a priori Wissen ist notwendig, um Sensordaten zu analysieren und zu verstehen. Diese Dichotomie findet man in der Kognition, in der Analyse von unstrukturierten Informationen, wie auch in industriellen Anwendungen. Maschinelles Lernen (ML), im besonderen Deep Learning, hat seine Stärken primär in der Analyse unstrukturierter Daten; typischerweise werden unstrukturierte Daten in sehr einfache diskrete Aussagen transformiert. Während im ML die sensorische Ebene zunehmend reichhaltiger wird, bleibt die strukturierte Seite eher simplistisch. Im Gegensatz zu vorangegangenen Anstrengungen zum Lernen in strukturierten Daten, sind für Wissensgraphen (oder englisch Knowledge Graphs (KGs)) leistungsfähige Lernalgorithmen entwickelt worden. Eine besonders interessante Perspektive ergibt sich aus der Möglichkeit, dass sich auf Basis von Hintergrundwissen und unstrukturierten Informationen in Kombination mit modernen Techniken des Deep Learnings prädiktive Systeme entwickeln lassen, die eine lernende Entscheidungsunterstützung realisieren. Das MLwin-Konsortium bündelt die Kompetenzen und die Vorarbeiten der beteiligten Partner aus den unterschiedlichen wissenschaftlichen und technischen Fachrichtungen, um das Thema Maschinelles Lernen mit Wissensgraphen substantiell weiterzuentwickeln und auf ein neues Fundament zu stellen. Ziel ist es, die Forschung in Deutschland als international führend zu diesen Themen zu etablieren und die Ergebnisse der Industrie und anderen Wissenschaften zu vermitteln. Ein zentrales Thema betrifft die Weiterentwicklungen und Konsolidierung der KG-Lernalgorithmen. Wir erweitern existierende Ansätze, um die Generalisierung auf neue Entitäten und Domänen zu verbessern. Weiterhin werden Data-Mining-Algorithmen entwickelt, die zur Analyse der zugrundeliegenden Strukturen in KGs dienen. Eine KG-unterstützte Informationsextraktion aus Texten und das halbautomatische Befüllen von KGs ist Thema eines weiteren APs. Neben der Analyse von Texten entwickeln wir KG-unterstützte Informationsextraktion aus Bildern und Videos. Schließlich beleuchten wir den Bezug zur Artificial Intelligence (AI) und zur Kognitionsforschung.

2 Stand von Wissenschaft und Technik

ML und Data Mining in KGs: ML in KGs ist aus verschiedenen Traditionen heraus entstanden, wie dem Inductive Logic Programming (ILP) und dem Statistical Relational Learning (SRL). Als besonders erfolgreich haben sich Ansätze erwiesen, die auf einer Faktorisierung des aus dem KG abgeleiteten Adjacency Tensors beruhen. Dieser Ansatz ist im Siemens-Team entstanden und von diesem und anderen Forschungsgruppen weiterentwickelt worden.

Textanalyse mit KGs: Deep Learning hat die automatische Analyse und Generierung von natürlicher Sprache revolutioniert. Dies gilt auch für Forschungsthemen, die für Wissensgraphen relevant sind. Für

das Befüllen von Wissensgraphen konnte gezeigt werden, dass Deep-Learning-Architekturen (insbesondere Recurrent Neural Networks (RNNs) und Convolutional Neural Networks (CNNs)) nicht-neuronalen Befüllungsverfahren überlegen sind.

Bild- und Videoanalyse mit KGs: Deep Learning konnte als erstes seine Leistungsfähigkeit in der automatisierten Analyse von Bildern und Videos demonstrieren. Das extrahierte Wissen ist in der Regel jedoch eher einfach und beschränkt sich beispielsweise auf Klassenzugehörigkeiten von Objekten im Bild. Ein aktives Forschungsthema ist die automatische Textannotation von Bildern. Die Extraktion von höherwertiger strukturierter Information aus Bildern und Videos steht noch am Anfang, wird jedoch zunehmend von der Forschung aufgegriffen.

Perspektiven zu AI und Kognitionsforschung: Führende internationale Gruppen stellen zunehmend den Bezug her zwischen den Entwicklungen der modernen Artificial Intelligence (AI) und der Kognitionsforschung. Am Sichtbarsten ist diese Entwicklung bei DeepMind. Auch die Arbeiten in unserem Vorhaben können auf kognitive Fähigkeiten des menschlichen Gehirns abgebildet werden. So kann man einen Bezug herstellen zwischen einem semantischen KG und dem semantischen Gedächtnis des menschlichen Gehirns, und zwischen einem episodischen KG und dem episodischen Gedächtnis.

Klinische Entscheidungsunterstützung mit KGs: Die Digitalisierung erfasst zunehmend das Gesundheitswesen. Ein wesentlicher Aspekt ist, dass klinische Daten analysiert werden, um Entscheidungsprozesse in den Kliniken zu verbessern. Einen solchen Ansatz haben die Teams von Siemens und des UKER im Smart Data Projekt Klinische Datenintelligenz (KDI) und in anderen Projekten verfolgt. Anhand des Beispiels „Entscheidungsunterstützung im Tumor Board“ wurde in Zusammenarbeit mit dem UKER und dem IFG demonstriert, dass RNNs trainiert werden konnten, Entscheidungen und z. B. die Prognose von Krebspatienten vorherzusagen, zur Entscheidungsunterstützung einzusetzen und hierbei Ergebnisse zu erzielen, die größere Akzeptanz finden als Entscheidungen der Ärzte im Komitee.

Ablauf des Vorhabens Das Projektvorhaben wurde von den Partnern innerhalb einer Laufzeit von drei Jahren und vier Monaten durchgeführt. Die Projektarbeiten begannen mit ersten Vorarbeiten im Oktober 2018 und liefen intensiv seit dem Kick-Off Workshop des Konsortiums im Januar 2019. Aufgrund der Covid-Pandemie wurde das Projekt kostenneutral um 4 Monate verlängert. Das Projekt endete erfolgreich am 31.01.2022. Alle Meilensteine und Deliverables wurde erreicht.

3 Wesentliche Ergebnisse und Zusammenarbeit im Projekt

Ein wesentlicher Beitrag in **AP 2** stellt die Entwicklung der Software PyKEEN dar, welches effizientes Trainieren und Evaluieren von Knowledge Graph Embedding Modellen ermöglicht. Es wurde auch verwendet, um die Reproduzierbarkeit zu analysieren, und die Link-Prediction Benchmarkingstudie basierend auf Knowledge Graph Embedding Modellen durchzuführen. Es wurde in **AP 3** zunächst ein Encoder-Decoder-Modelle zur Transformation von Wissensgraphen in Text und umgekehrt entwickelt. An der LMU wurde der erste Ansatz zur unüberwachten Texterzeugung aus KGs vorgestellt. Des Weiteren wurde ein neues Modell der Graph-zu-Textgenerierung entwickelt. Zu den wesentlichen Ergebnissen in **AP 4** gehören neue Methoden zur Schätzung von Korrespondenzen zwischen als Dreiecksgitter (ungerichtete Graphen mit im 3D-Raum eingebetteten Knoten) repräsentierten 3D-Formen. Gemeinsam mit dem Projektpartner Hinrich Schütze (LMU) wurde gezeigt, dass die automatische Erstellung von Szenengraphen helfen kann, bessere Textbeschreibungen von Bildern zu erzeugen. **AP 5** untersucht die Bezüge zur AI sowie zur Kognitionsforschung, insbesondere inwieweit die Ergebnisse aus AP3 und AP4 die Grundlage für Entscheidungsunterstützungssysteme bilden können. Wir haben neue Methoden zur Modellierung und Schlussfolgerung episodischer Wissensgraphen vorgeschlagen und untersucht. Als Anwendungsszenarien gelten u.a. Klassifikationsaufgaben für Fakten in Wissensgraphen und Textdaten. Die Ergebnisse bilden eine Grundlage zur Entscheidungsmodellierung in AP6. Im **AP 6** wurde Kuration von klinischen Daten effizienter implementiert. Diese Studiendaten wurden den Projektpartnern zur Verfügung gestellt. Hierfür wurden die Ansprüche der Partner erarbeitet, sowie die ethischen und datenschutzrechtlichen Belange evaluiert. Das Siemens Team hat an den zur Verfügung gestellten Datensätzen quantitative Analyse sowie Modellierungen durchgeführt. Das UKER Team hat zusammen mit den Projektpartnern LMU und IFG GmbH eine klinische Fragestellung erarbeitet, die einen Patienten-nahen kurzfristigen Benefit bringen kann. In **AP 7** wurde ein Modell zur Einbettung verteilter Wissensgraphen entwickelt. Das Modell wurde mit PyTorch-Biggraph verglichen und die parallelisierte Trainingsmethode HOGWILD integriert.