



MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Knowledge Engineering:
Modellierung multimodaler terminologischer Ressourcen
am Beispiel des ophthalmologischen Bereichs“

verfasst von / submitted by

Andrea Ferri

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Master of Arts (MA)

Wien, 2021 / Vienna 2021

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

UA 070 348 331

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Masterstudium Translation Italienisch Deutsch

Betreut von / Supervisor:

Ass.-Prof. Mag. Dr. Dagmar Gromann, BSc

Danksagung / Ringraziamenti

An dieser Stelle möchte ich mich bei allen bedanken, die mich bei der Fertigstellung meiner Masterarbeit unterstützt haben. Mein ganz besonderer Dank gilt meiner Betreuerin Ass.-Prof. Mag. Dr. Dagmar Gromann, BSc für ihre ständige Unterstützung, ihre wertvollen Ratschläge und das Vertrauen, das sie mir bei der Durchführung dieser Studie geschenkt hat.

Des Weiteren möchte ich mich ganz herzlich bei meinen Studienkolleginnen und Studienkollegen und liebe Freundinnen und Freunde Francesca, Gaia, Júlia, Lucas, Kamila, Monica, Stefano, ohne die meine Wiener-Erfahrung nicht die gleiche gewesen wäre.

Vielen Dank Kamila und EP Lucas, dass ihr meine Masterarbeit korrekturgelesen habt.

Un ringraziamento particolare va anche alla mia famiglia e a tutti i miei amici in Italia, che rappresentano la mia forza e le mie radici, senza le quali non potrei spiccare il volo.

Danke! Grazie!

Abbildungsverzeichnis

Abbildung 1. Semiotisches Dreieck.

Abbildung 2. Beispiel für ein Bestandssystem.

Abbildung 3. Wissensrepräsentationsmethoden.

Abbildung 4. Eine Ontologie.

Abbildung 5. Semantische Reichhaltigkeit.

Abbildung 6. Eine andere Ontologie.

Abbildung 7. Ontologie mit is-a-Beziehung.

Abbildung 8. Stilisierter Mensch.

Abbildung 9. Ein Kreis.

Abbildung 10. Ein Kreis im Konsekutivdolmetschen.

Abbildung 11. Begriff *Water* in EcoLexicon.

Abbildung 12. Beispiel für Relationssystem.

Abbildung 13. Seite Ophthalmology auf MedPix®.

Abbildung 14. Szenario 13, Silicone globe injection, *Topic*.

Abbildung 15. Szenario 13, Silicone globe injection, *Case*.

Abbildung 16. Drei Ebenen in der Anatomie.

Abbildung 17. CT w/contrast (IV).

Abbildung 18. CT – noncontrast.

Abbildung 19. MR – T1W – noncontrast.

Abbildung 20. MR - T2 weighted.

Abbildung 21. MR - FLAIR.

Abbildung 22. MR - T2 FLAIR w/Contrast.

Abbildung 23. MR - T1W w/Gd (fat suppressed).

Abbildung 24. Gross photograph.

Abbildung 25. HE - Med Power (~50-200x).

Abbildung 26. HE - High Power (>200x).

Abbildung 27. MR – STIR.

Abbildung 28. VariMed-Suchleiste.

Abbildung 29. Begriff auf Spanisch und Englisch.

Abbildung 30. Ergebnisse für einen Begriff in VariMed.

Abbildung 31. Begriffe mit spanischer Äquivalenz in VariMed.

Abbildung 32. Ergebnisse für cyst in VariMed.

Abbildung 33. Begriffssystem mit von VariMed vorgeschlagenen Beziehungen.

Abbildung 34. Resultierendes Begriffssystem.

Abbildung 35. Begriff *grave* im TBX-Format

Inhaltsverzeichnis

Einleitung	8
1. Grundlagen	11
1.1 Grundlagen der Terminologie.....	11
1.1.1 Formen der Terminologearbeit.....	12
1.1.2 Der Beruf der Terminologinnen und Terminologen.....	13
1.1.3 Grundelemente der Terminologie.....	14
1.1.3.1 Gegenstand.....	15
1.1.3.2 Begriffe.....	16
1.1.3.3 Benennung.....	16
1.1.4 Arten von Begriffsbeziehungen.....	17
1.1.4.1 Intension und Extension.....	18
1.1.5 Begriffssysteme.....	19
1.1.6 Beziehungen zwischen Benennungen und Begriffen.....	20
1.1.6.1 Synonymie.....	20
1.1.6.2 Antonymie.....	21
1.1.6.3 Ambiguität.....	22
1.1.6.4 Äquivalenz.....	23
1.2 Wissensmanagement, Wissensmodellierung und Wissensrepräsentation.....	24
1.2.1 Semantische Netze.....	27
1.2.1.1 Taxonomie, Thesaurus und Ontologie.....	28
1.2.1.2 Wissensvernetzung durch Ontologien.....	32
1.2.1.3 Terminologische Begriffssysteme.....	34
1.2.1.4 Darstellungsformate von semantischen Netzen.....	35
1.3 Multimodalität.....	36

1.4 Medizinische Fachsprache.....	39
1.4.1 Medizinische Datenbanken.....	42
1.5 ISO Normungen.....	43
2. Aktueller Stand der Forschung.....	44
3. Methode.....	51
3.1 Erster Ansatz zu MedPix®.....	51
3.1.1 Die Ophthalmologie.....	51
3.2 Analyse der 15 Szenarien.....	53
3.3 Analyse der Schlüsselwörter.....	61
3.4 Erste Begriffsbeziehungen mit Hilfe von VariMed.....	65
3.5 Arten von Beziehungen für das Endprojekt.....	69
3.5.1 Ausgewähltes Darstellungsformat.....	71
4. Ergebnisse.....	72
4.1 Erstellung des Begriffssystems.....	72
4.2 Aufbau von Text-Bild-Beziehungen.....	77
5. Diskussion.....	80
6. Conclusio.....	82
7. Bibliographie.....	83
7.1 Literaturquellen.....	83
7.2 Linksammlung.....	87
7.3 Videoquellen.....	88

Einleitung

Die vorliegende Masterarbeit analysiert einen multimodalen Datensatz, der sowohl terminologische Informationen als auch dazugehörige Bilder enthält. Der Fokus und die Zielsetzung dieser Masterarbeit liegt auf der Modellierung multimodaler Informationen in terminologischen Ressourcen.

Die Welt erfährt gegenwärtig beispiellose Zeiten ständiger und schneller Entwicklung, die von zwei Haupteigenschaften geprägt sind: die rasante Zunahme von Informationen und die wachsende Zusammenarbeit an neuen Technologien (Drewer et al. 2017a). Unmengen von Daten werden tagein tagaus verarbeitet. Um am besten von dem erwähnten Szenario profitieren zu können, kann die Künstliche Intelligenz implementiert werden, sodass Menschen und Maschinen sich einander helfen können. Diese Unmengen von Daten sind Informationen, die das Wissen bilden. Das Wissen kann in der derzeitigen Epoche ebenso *digitalisiert* werden. Diesbezüglich kommt die Wissensmodellierung (Knowledge Engineering) ins Spiel, die das Wissen in Form von maschinenlesbaren Ontologien repräsentiert (Drewer et al. 2017a). Auf diese Art und Weise werden Informationen von Maschinen interpretiert und miteinander in Verbindung gesetzt, nicht mehr lediglich von Menschen verstanden (Dobratz 2008).

Der Bereich, auf den sich die vorliegende Masterarbeit bezieht, ist die Medizin bzw. die Ophthalmologie, gerade wegen ihres multimodalen Charakters und Hochspezialisierungsgrads. Folglich wird die angewendete Methode erläutert.

Zunächst einmal wurde eine Durchsicht der medizinischen Datenbanken durchgeführt, um den geeigneten Datensatz für die vorliegende Masterarbeit zu finden. MedPix®, der ausgewählte multimodale medizinische Datensatz, hat alle Bedürfnisse erfüllt, weil er reich an textuellen Informationen und Bildern ist. Infolgedessen wurde eine Analyse des Datensatzes durchgeführt, indem sowohl Bilder als auch Text-Informationen extrahiert worden sind, um modelliert zu werden. Infolge der Extraktion wurden die Begriffe in VariMed gesucht, einem anderen Projekt bezüglich multimodaler medizinischer Terminologie. Mit Hilfe von VariMed wurden die ersten Beziehungen erstellt, die übrigen Beziehungen wurden mit Hilfe von medizinischen Wörterbüchern und textuellen Informationen, die im Datensatz vorhanden sind, aufgebaut. Letztendlich wurde die modellierte Terminologie in Form eines terminologischen Begriffssystems dargestellt.

Die vorgeschlagene Masterarbeit zielt darauf ab, mögliche Repräsentationsmodelle bzw. -formate für eine multimodale Terminologie zu analysieren. Bevor diese empirische Untersuchung vorgestellt wird, werden zuerst die theoretischen Grundlagen im theoretischen

Teil der vorliegenden Masterarbeit vorgestellt. Zuerst wird das Grundwissen durch die ersten Kapitel über die Grundlagen der Terminologie, einige Formen der Terminologearbeit sowie kurze Erwähnungen über den Beruf der Terminologinnen und Terminologen erworben. Erst danach wird das semiotische Dreieck präsentiert, indem die Grundelemente der Terminologie erläutert werden. Infolgedessen werden die unterschiedlichen Arten von Begriffsbeziehungen, Beziehungen zwischen Benennungen und Begriffen sowie mögliche Begriffssysteme vorgestellt. Zuletzt wird das Thema der Multimodalität und der Wissensrepräsentation mit Hilfe von semantischen Netzen erläutert. Letztendlich wird in das Thema der medizinischen Fachsprache bzw. der medizinischen Datenbanken eingegangen. Im Kapitel des aktuellen Stands der Forschung werden einige Publikationen präsentiert, die sich entweder mit medizinischer Terminologie oder mit multimodalen Repräsentationssystemen befassen.

Angesichts des eben erwähnten Ziels dieser Masterarbeit wird die Forschungsfrage wie folgt formuliert: Welche Haupteigenschaften und zu betrachtende Kernaspekte sind bei der Modellierung einer multimodalen (Text-Bild-Beziehungen) Terminologie zu beachten, welche aus terminologischen Informationen aus Text und dazugehörigen Bildern zusammengestellt wird?

Die jetzigen Hypothesen sehen vor, dass es möglich sein wird, terminologische Modelle eines technischen multimodalen Datensatzes zu finden und, dass es gemeinsame Variablen geben wird, die auf gemeinsamen Eigenschaften basieren, so wie das Wortfeld (bedeutungsähnliche Wörter), den Bezugsgrad zwischen Bild und Text und die Beziehungen zwischen Nebeninformationen bzw. zusätzlichen Informationen der Patientinnen und Patienten, die den Kontext prägen, in welchem sich die multimodale Terminologie befindet.

Wie bereits erwähnt ist demgemäß das Ziel dieser Arbeit die Erstellung einer multimodalen Modellierung für terminologische Informationen, die am Beispiel der Augenheilkunde angewandt wird und als Lernstütze für Medizinstudierende, als terminologische Datenbank für Krankenhäuser und als unterstützendes Material für Laiinnen und Laien und ebenso als Stütze für Übersetzerinnen und Übersetzer im Medizinbereich fungiert. Eine solche Wissensmodellierung soll die Datenverarbeitung und den Wissenserwerb dank einem höheren Informationsgrad erleichtern. Es handelt sich um einen neuen bzw. adaptierten Ansatz der Wissensmodellierung und der Multimodalität (Zusammenhang von textuellen Ressourcen und Bildern). Dies führt zu einem bedeutsamen Beitrag zur Forschungslandschaft in der multimodalen Wissensmodellierung und einer Erweiterung der derzeitigen Terminologiemodelle. Darüber hinaus könnte diese Arbeit zu einem

Darstellungsformat beitragen, das zu paralleler Text- und Bildverarbeitung in Systemen der Künstlichen Intelligenz verwendet werden könnte.

1. Grundlagen

Dieses Kapitel soll einen Überblick über die Hauptthemen dieser Masterarbeit bieten, welche Terminologiewissenschaft, in dem das semiotische Dreieck eingeführt wird und darüber hinaus Begriffe, Begriffsbeziehungen und Begriffsrepräsentationssysteme erläutert werden, die verschiedenen semantischen Netze, Wissensmanagement, Wissensmodellierung, Wissensrepräsentation sowie multimodale Ressourcen und medizinische Fachsprache. Die Reihenfolge der erwähnten Bereiche erfolgt nach einem Kriterium, das als „Trichter-Kriterium“ bezeichnet werden könnte, das heißt vom allgemeineren Bereich (Terminologiewissenschaft) bis zum letzten und spezifischsten Feld (medizinische Fachsprache bzw. Datenbanken). Auf diese Art und Weise wird das Grundwissen vorgestellt, um die nachfolgende Arbeit und die damit verbundenen Techniken und Methoden besser zu beschreiben.

1.1 Grundlagen der Terminologie

Die Terminologiewissenschaft, auch Terminologielehre genannt, beschäftigt sich mit dem Fachwortschatz bzw. mit Begriffen und deren Bezeichnungen in den Fachsprachen. Es handelt sich hierbei um die rein theoretische Beschäftigung mit dem Thema der Terminologie (KÜDES 2018).

Während eine kontinuierliche und unaufhaltsame Spezialisierung in einer Vielzahl von Bereichen stattfindet, muss der Mensch auch mit den Nachteilen einer zunehmend vernetzten Welt umgehen (Drewer et al. 2017a). Unterschiedliche Kulturen und Sprachen treten immer häufiger in Kontakt. Dies erfordert eine klar definierte terminologische Klärung, um eventuelle Missverständnisse zu vermindern und, um die Genauigkeit in komplexen und verflochtenen Netzen zu gewährleisten. Genau um die Festlegung dieser klaren, deutlichen und einheitlichen Termini geht es bei der Terminologearbeit (KÜDES 2018). Im Gegensatz zur Terminologiewissenschaft, die sich mit der Theorie beschäftigt, befasst sich die Terminologearbeit mit der praktischen Anwendung der Terminologie, das heißt die praktische Erarbeitung, Bearbeitung, Verwaltung, Darstellung und Verbreitung (Pulitano 2006). In das Thema der Terminologearbeit wird allerdings in den nächsten Kapiteln tiefer eingegangen.

1.1.1 Formen der Terminologiearbeit

Die Terminologie ist die Gesamtheit der Fachtermini bzw. des Fachwortschatzes, die bzw. der in einem Fachgebiet verwendet werden bzw. wird. Häufig ist es wichtig, die korrekte Verwendung dieser Fachsprache zu beachten, um die Entstehung von Missverständnissen zu vermeiden (KÜDES 2018).

Es existieren zwei Hauptkategorien von Terminologiearbeit: zweisprachig bzw. mehrsprachig und einsprachig. Zweisprachige bzw. mehrsprachige Terminologiearbeit ist die bekannteste Form, weil sie dazu dient, die Arbeit der Übersetzerinnen und Übersetzer und der Dolmetscherinnen und Dolmetscher zu erleichtern. Allerdings ist einsprachige Terminologiearbeit für bestimmte Situationen ausreichend. Die Unternehmensrealität ist ein ausgezeichnetes Beispiel dafür, weil eine terminologische Verwaltung mit Hilfe einer einsprachigen Terminologiearbeit für solche Realität notwendig sein kann. Bezüglich der internen Kommunikation gibt es Abteilungen, sowie Entwicklung, Marketing, Produktdesign, technische Dokumentation, Übersetzung usw., die von einer einheitlichen internen Kommunikation profitieren können (KÜDES 2018). Natürlich hat eine konsistente Kommunikation eine positive Auswirkung auf den Außenauftritt. Die Glaubwürdigkeit des Unternehmens gegenüber den Kundinnen und Kunden wird durch eine eindeutige Kommunikation gestärkt. Dies führt zur Kundenbindung und zu einem besseren Image des Unternehmens. Konsistente Terminologiearbeit kann sogar die Übersetzungskosten reduzieren, weil die Terminologie des Unternehmens stets einheitlich wiederverwendet wird (Drewer et al. 2011).

Es existieren auch weitere Formen von Terminologiearbeit, wie zum Beispiel punktuell oder systematisch, allgemein oder speziell und deskriptiv oder präskriptiv (Drewer et al. 2011). Die punktuelle Terminologiearbeit beschäftigt sich mit der Lösung einzelner terminologiebezogener Probleme und zielt darauf ab, Benennungen, Neologismen oder Fachtermini zu erarbeiten, die in existierenden Terminologiebeständen fehlen. Die systematische Terminologiearbeit fordert zwei Arten von Untersuchungen: die textbezogene und die sachgebietsbezogene Untersuchung. Die textbezogene Untersuchung ist eine vorbereitende terminologische Untersuchung bei der Übersetzung von Fachtexten. Die sachgebietsbezogene Untersuchung ist eine terminologische Untersuchung aller relevanten Begriffe eines Sachgebiets und deren Beziehungen und Darstellung (Drewer et al. 2011). Bei sachgebietsbezogenen Untersuchungen werden nicht nur die einzelnen Begriffe bearbeitet, sondern auch ihre Nachbarbegriffe (KÜDES 2018). Auf Grund des hohen

Spezialisierungsgrads sachgebietsbezogener Untersuchungen werden in der Regel Expertinnen und Experten eingesetzt, um die Korrektheit der Begriffe und deren Beziehungen zu überprüfen (KÜDES 2018).

Die allgemeine Terminologearbeit ist die sprach- oder fachübergreifende Terminologearbeit sowie das Ergebnis eines Abstraktionsprozesses aus speziellen Terminologien. Spezielle Terminologearbeit beschäftigt sich mit der Fachterminologie in einem bestimmten Fachgebiet (Arntz et al. 2014).

Deskriptive Terminologearbeit beschäftigt sich mit den Termini zu einer bestimmten Zeit und in einem spezifischen Fachgebiet, ohne sie zu bewerten. Präskriptive Terminologearbeit wird in der Regel erst nach der deskriptiven Terminologearbeit durchgeführt und beschäftigt sich mit der Vorbereitung einer Vorzugsbenennung für jeden Begriff, um den gesamten Wortschatz zu vereinheitlichen (Drewer et al. 2011).

In Hinblick auf die multimodale Terminologearbeit, die den Kern der vorliegenden Masterarbeit bildet, werden nicht nur textuelle Ressourcen, sondern auch zusätzliche Ressourcen anderer Art, das heißt bildliche Ressourcen, eingesetzt. Dieses Thema wird in den nächsten Unterkapiteln vertieft.

1.1.2 Der Beruf der Terminologinnen und Terminologen

Aufgrund der oben erwähnten Umstände und, infolgedessen, der Notwendigkeit für Terminologearbeit hat sich in den letzten Jahrzehnten ein relevantes Interesse für die Terminologiewissenschaft entwickelt, das zur Entstehung des Berufs der Terminologinnen und Terminologen geführt hat (Pulitano 2006).

Im Jahr 1958 wurden die ersten Terminologinnen und Terminologen in der Europäischen Kommission eingesetzt. In Kanada wurde diese professionelle Rolle seit dem Anfang der 1970er Jahre eingesetzt, um die offizielle Terminologie der Bundesverwaltung zu sammeln und zu verbreiten (Pulitano 2006). Allerdings hat dieser Beruf erst in den letzten Jahrzehnten des 20. Jahrhunderts an Relevanz gewonnen (Pulitano 2006).

Die Terminologinnen und Terminologen zielen darauf ab, die Fachkommunikation zu verbessern. Der Beruf besteht aus der Verarbeitung, Beschreibung, Verwaltung und Verbreitung von Fachterminologie, die mit vielfältigen Informationen und Fachkenntnissen umgeht (Pulitano 2006). Es handelt sich um einen Beruf, der Gemeinsamkeiten sowohl mit der Sprachwissenschaft (Dolmetschen, Lexikographie, Übersetzen etc.), als auch mit der

Informatik, einigen wissenschaftlichen Disziplinen, der Kognitionswissenschaft und der Normalisierung teilt (Pulitano 2006).

Terminologinnen und Terminologen beschäftigen sich mit der Verarbeitung terminologischer Ressourcen nach einer gut definierten Methodik. Diesbezüglich beschäftigen sich Terminologinnen und Terminologen mit der Durchsicht der Unterlagen, der Erarbeitung der Konzepte und Begriffe, und der Verfassung der Terminologie, die die Terminologearbeit zusammenstellt (Pulitano 2006). Häufig müssen Terminologinnen und Terminologen Tätigkeiten terminologischer Planung ausüben, in dem sie neue Termini, Namen für Produkte, Funktionsbezeichnungen erfinden, um die Terminologie zu normalisieren (Pulitano 2006). Darüber hinaus tragen Terminologinnen und Terminologen zur Verbreitung terminologischer Sammlungen bei, indem sie sich um terminologische Datenbanken, die Evaluierung und Entwicklung terminologischer Software und die Verwaltung terminologischer Datensätze kümmern (Pulitano 2006). Außerdem gehen Terminologinnen und Terminologen Beratungs- und Weiterbildungsaktivitäten nach. Es handelt sich um Beratungsaktivitäten sowohl für direkte Benutzerinnen und Benutzer terminologischer Datenbanken als auch für die Erstellung von Terminologiedienstleistungen, sowie Kurse für die Lehre solcher Kompetenzen, sodass die Schlüsselqualifikationen im Bereich der Terminologiewissenschaft erworben werden können (Pulitano 2006).

1.1.3 Grundelemente der Terminologie

In der Terminologiewissenschaft ist es wichtig, zwischen *Begriff* und *Benennung* zu unterscheiden, da diese beiden Begriffe im allgemeinen Sprachgebrauch mit *Wort* ausgedrückt werden. Gerade um dieses mögliche Missverständnis zu überwinden, wurde das semiotische Dreieck (siehe Abb. 1) konzipiert (Arntz et al. 2014). Das semiotische Dreieck zielt darauf ab hervorzuheben, dass die Benennung sich nicht direkt auf den Gegenstand beziehen kann, sondern lediglich mittels eines Vermittlers: den Begriff (Arntz et al. 2014).

Vor der Erstellung einer Terminologearbeit ist es wichtig, die Bedeutung der folgenden grundlegenden Elemente einer Terminologie zu klären: Gegenstand, Begriff und Benennung. Der Gegenstand ist sowohl das physische als auch das immaterielle Objekt in der realen Welt mit den dazugehörigen typischen Eigenschaften (Drewer et al. 2017b). Der Begriff ist die Vorstellung eines Gegenstandes in unserem Kopf. Diese Vorstellung wird durch Abstraktion infolge der Enkulturation bzw. Lebenserfahrungen gewonnen. Die „Summe“ der Eigenschaften aller *erfahrenen* Gegenstände bildet die Vorstellung zu einem Gegenstand bzw. einem Begriff.

Unter Benennung hingegen versteht man die sprachliche Repräsentation des Begriffes (Schwarz-Friesel et al. 2014). Folglich wird dieses Konzept mit Hilfe eines konkreten Beispiels bzw. Objektes erläutert.

Die gemeinsamen Eigenschaften vieler Tische der realen Welt bilden Merkmale, die einen Begriff definieren. Ein Tisch hat zum Beispiel in der Regel vier Beine und eine flache Ebene. Sobald diese Eigenschaften eines Gegenstandes durch Abstraktion unseren Begriff geprägt haben, ist es nicht relevant, ob ein Tisch ein oder zehn Beine hat, oder ob es um einen kreisförmigen, zeitgenössischen Tisch aus Kristall oder ein Museumsstück aus Holz aus dem 18. Jahrhundert geht. Die Hauptsache ist, dass der Tisch über alle Merkmale verfügt, um als Tisch bezeichnet werden zu können. Während die Benennung die Art und Weise ist, wie ein bestimmter Begriff sprachlich dargestellt wird. Anders ausgedrückt versteht man unter Benennung die Buchstabenfolge, die auf einen bestimmten Begriff zurückführt (Schwarz-Friesel et al. 2014). Abb. 1 zeigt das semiotische Dreieck, welches die Beziehungen zwischen Gegenstand, Begriff und Benennung hervorhebt.

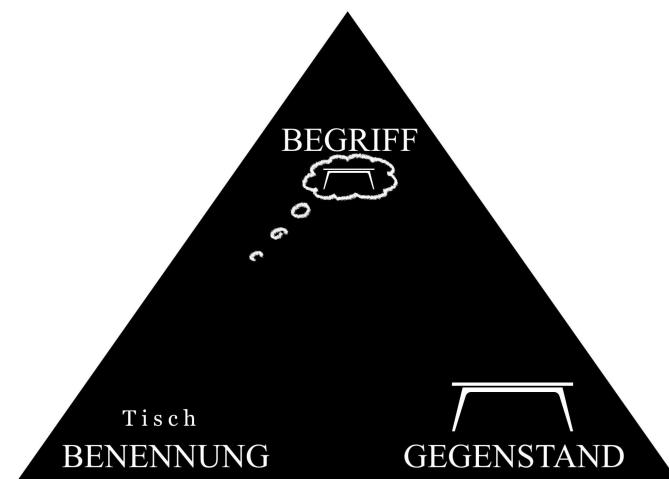


Abbildung 1. Semiotisches Dreieck

1.1.3.1 Gegenstand

Wie im vorigen Unterkapitel erklärt wurde, ist ein Gegenstand ein Ausschnitt der Realität in der wir leben. Der Gegenstand ist aber nicht unbedingt etwas Materielles bzw. ein konkretes Objekt, sondern kann ein Gegenstand ein immaterielles Stück der Realität repräsentieren (Drewer et al. 2017b). Drewer et al. (2017b) führen ein Beispiel für die diversen Arten von Gegenstand an. Eine *Tastatur* ist ein Beispiel für einen konkreten Gegenstand, die *Silbentrennung* ist ein Beispiel für einen abstrakten Gegenstand und *Festplatte formatieren* ist ein Beispiel für einen Sachverhalt oder einen Vorgang (Drewer et al. 2017b). Die Tastatur, die

Silbentrennung oder das Formatieren der Festplatte existieren aber nicht als solche in unserer Welt. Sie sind nämlich das Ergebnis der Übereinstimmung unterschiedlicher Eigenschaften (auch Merkmale genannt), die zu Begriffen zusammengefasst werden (Drewer et al. 2017b).

1.1.3.2 Begriffe

Es existieren verschiedene Arten von Begriffen: Allgemein- oder Individualbegriffe und konkrete oder abstrakte Begriffe (Tatievskaja 2003). Allgemeinbegriffe sind abstrakte Begriffe, die sich auf eine Vielfalt von Gegenständen beziehen, die Gemeinsamkeiten teilen. Allgemeinbegriffe können beispielsweise Mensch oder Tier sein, oder, um nicht vom medizinischen Bereich abzuweichen, könnte der Begriff *Krebsregister* ein Allgemeinbegriff sein (Pedersen et al. 2002). Unter Individualbegriff werden Begriffe verstanden, die sich auf einzelne Gegenstände beziehen. Ein medizinisches Beispiel dafür könnte der Begriff *Epidemiologisches Krebsregister Sachsen* sein (Pedersen et al. 2002).

1.1.3.3 Benennung

Wie bereits erwähnt versteht man unter Benennung die Bezeichnung eines Gegenstandes und die Art und Weise, wie ein Begriff in Buchstabenfolge ausgedrückt wird, sowohl durch ein Wort als auch durch mehrere Wörter, die den Begriff repräsentieren (Schwarz-Friesel et al. 2014). Die Bezeichnung ist der Oberbegriff der Benennung und umschließt ebenso unter anderem Symbolen, Formeln, Emoticons, Emojis. Wobei die Bezeichnung alles das ist, was nicht ausschließlich aus Buchstaben zusammengesetzt ist, bezeichnet man aber einen Begriff, versteht man unter Benennung lediglich die Ausdrucksformen in Buchstabenfolgen eines Begriffes. Aus diesem Grund können einem Begriff mehrere Benennungen zugeordnet werden und man spricht in diesem Fall von Synonymie (falls in derselben Sprache) (Arntz et al. 2014).

Dies könnte zum Beispiel eine chemische Formel sein, die sowohl aus Buchstaben als auch aus Zeilen und Zahlen besteht, oder Symbolen wie das Lebensgefahrsymbol mit Schädel und gekreuzten Knochen (Drewer et al. 2017b). Es existieren zwei Hauptkategorien von Benennung: Einwort- und Mehrwortbenennung. Einwortbenennungen sind Benennungen, die aus einem Wort bestehen. Nomen-Komposita (zusammengesetzte Wörter) in der deutschen Sprache werden ebenso als Einwortbenennungen klassifiziert (Drewer et al. 2017b). Beispiele dafür sind sowohl das Substantiv *Garten* als auch die Einwortbenennung bzw. das Kompositum *Blumentopferde* (Drewer et al. 2017b). Mehrwortbenennungen sind alle Benennungen, die mehrere Wörter bzw. Substantive, Adjektive und Verben benötigen, um einen Gegenstand zu

bezeichnen (Drewer et al. 2017b). Ein Beispiel dafür könnte die Mehrwortbenennung *Drucker mit Einzelblatteinzug* sein (Drewer et al. 2017b).

Wie bereits erwähnt ist die Benennung die sprachliche Darstellung eines Begriffes (Drewer et al. 2017b). Bei Fachsprachen ist es notwendig, dass eine eindeutige Beziehung zwischen Begriff und Benennung besteht (Drewer et al. 2017b). In diesem Fall spricht man von Eineindeutigkeit. Dieses Prinzip sieht vor, dass es lediglich eine Benennung pro Begriff und einen Begriff pro Benennung gibt (Drewer et al. 2017b). In manchen Fällen ist es schwer, das eben erwähnte Eineindeutigkeitsprinzip zu verfolgen, weil die Synonymie und Ambiguität Hindernisse darstellen können (Drewer et al. 2017b).

1.1.4 Arten von Begriffsbeziehungen

Begriffe können durch Beziehungen in Verbindung gesetzt werden, sodass Begriffe, die gemeinsame Merkmale haben und in einer Gruppe klassifizierbar sind, mit einer anderen thematischen Begriffsgruppe verknüpft werden können (Drewer et al. 2017b).

Die hierarchischen Beziehungen (IS-A, PART-OF) können in zwei Untergruppen aufgeteilt werden: Bestands- und Abstraktionsbeziehungen. Bestandsbeziehungen (auch Teil-Ganzes-Beziehung oder partitive Beziehung genannt) sind von einem übergeordneten und einem untergeordneten Begriff charakterisiert. Der übergeordnete Begriff repräsentiert das Ganze (z.B. einen Hund) und der untergeordnete Begriff repräsentiert ein Teil vom übergeordneten Begriff (z.B. seinen Schwanz) (Drewer et al. 2017b). Als Beispiel für dieses System siehe Abb. 2.

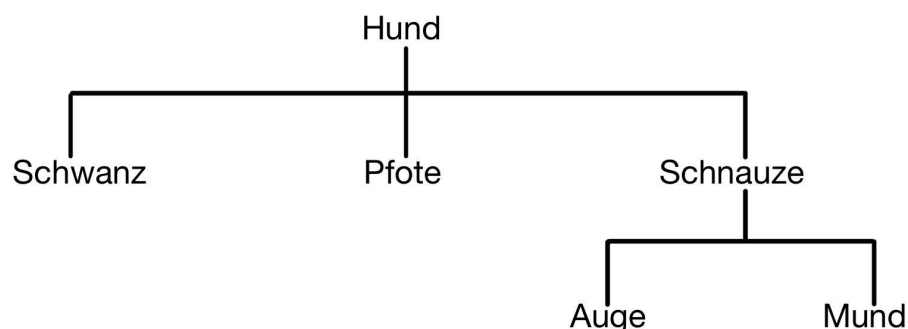


Abbildung 2. Beispiel für ein Bestandssystem (Drewer et al. 2017b: 10)

Abstraktionsbeziehungen (auch generische, logische Beziehung oder *genericRelation* genannt) sehen vor, dass der untergeordnete Begriff über alle Eigenschaften des übergeordneten Begriffs verfügt, aber zusätzlich eine weitere Eigenschaft besitzt (z.B. Handy und Smartphone). Ein Smartphone ist ein Handy, verfügt aber über zusätzliche Merkmale, die es *komplexer* im Vergleich zu einem üblichen Handy machen (Drewer et al. 2017b).

Systeme, die lediglich von Bestandsbeziehungen charakterisiert sind, werden Bestandssysteme (oder *partitiveRelation*) genannt (Drewer et al. 2017b). Bei Systemen, die lediglich von Abstraktionsbeziehungen charakterisiert sind, spricht man von Abstraktionssystemen (Drewer et al. 2017b). Allerdings ist es wichtig zu klären, dass es sehr selten der Fall ist, dass ein System lediglich aus Abstraktionsbeziehungen besteht. Die Realität ist eben ein sehr komplexes System verflochtener Informationen und kann kaum lediglich von einem oder dem anderen System dargestellt werden. Aus diesem Grund werden die oben erwähnten Systeme oft miteinander verknüpft. Man spricht in diesem Fall von gemischten Begriffssystemen (Drewer et al. 2017b). Bei sehr komplexen Darstellungen eines Systems, wo viele Begriffe thematisch miteinander verbunden sind, spricht man von einem Begriffsfeld (Drewer et al. 2017b).

Neben den hierarchischen Beziehungen existieren ebenso die sogenannten nicht-hierarchischen Beziehungen (MADE-OF, HAS-FUNCTION), die für die Terminologearbeit weniger sinnvoll als die hierarchischen Beziehungen sind (Drewer et al. 2017b). Bei solchen Beziehungen können zum Beispiel Ursache-Wirkung-Beziehung oder chronologische Abfolgen repräsentiert werden (Drewer et al. 2017b). Darüber hinaus gibt es in jedem Bereich je nach Bedarf unterschiedliche Arten von Beziehungen. Im Medizinbereich sind OCCURS-AT, LOCATED-AT, CAUSED-BY (*spatialRelation*, *casualRelation* (Nuopponen 2014)) die Beziehungen, die am meisten verwendet werden. Sie sind nämlich notwendig, um das Gebiet zu definieren, wo sich ein Körperteil befindet oder wo eine Erkrankung auftritt bzw. auftreten kann. Die CAUSED-BY Relation fügt den Auslöser der Krankheit und die ausgelöste Erkrankung zusammen (Nuopponen 2014). Allerdings werden alle Arten von Beziehungen ausführlicher im zweiten Teil der vorliegenden Masterarbeit erläutert.

1.1.4.1 Intension und Extension

Unter Extension versteht die Semantik alle Gegenstände, die unter den Schirm eines bestimmten Begriffs fallen können. Als Intension werden die Merkmale der Gegenstände verstanden (Menne 1991). Um ein Beispiel zu nennen kann Platon erwähnt werden. Platon

prägte den Begriff *Mensch*, indem er bestimmte, dass der Mensch ein „zweifüßiges ungefedertes Lebewesen“ sei (Menne 1991: 91). Werden Merkmale einem Gegenstand zugeschrieben, dann spricht man von Intension des Begriffs (Menne 1991). Die Extension des Begriffs ist einfach der Mensch, der infolge der Intension über die Eigenschaften verfügt, um als Mensch angesehen zu werden, ohne zusätzliche erklärende Eigenschaften zu spezifizieren. Dank der Intension kann einem Begriff ein konkreter Gegenstand klar zugeordnet werden, wenn dieser alle Merkmale dieses Begriffs erfüllt (Menne 1991).

1.1.5 Begriffssysteme

Mit Hilfe von graphischen Begriffssystemen können Begriffe mit relevanten gemeinsamen Eigenschaften in Verbindung gesetzt und dargestellt werden (Drewer et al. 2017b). Trotz der nicht immer eindeutigen Begriffsbeziehungen müssen Begriffsrepräsentationssysteme immer eindeutig ablesbar sein. In vielen Fällen sind diese Systeme nämlich nicht nur von A nach B ablesbar, sondern es können mehrere Richtungen zu berücksichtigen sein, um den Graphen abzulesen und richtig interpretieren zu können, zum Beispiel von A nach B, aber auch von B nach A und dann B nach C, aber nicht A nach C. Aus diesem Grund lassen sich die Begriffsbeziehungen in diversen Formen mit Hilfe unterschiedlicher Repräsentationsformate darstellen (KÜDES 2018). Ein Beispiel solcher Systeme kann die Baumstruktur sein, die eine freie Zuordnung der Begriffe ermöglicht. Ein weiteres Beispiel ist das hierarchische System, das diese Flexibilität nicht erlaubt (KÜDES 2018).

Begriffsbeziehungen können hierarchisch oder nicht-hierarchisch sein. Falls jeder Begriff nur einen Oberbegriff hat, spricht man von Monohierarchie. Wenn ein Begriff mehrere Oberbegriffe hat, spricht man von Polyhierarchie (Drewer et al. 2017b).

Es kann der Fall sein, dass ein Sachgebiet sowohl abstrakte als auch konkrete Begriffe beinhaltet, und diese Begriffe könnten in ungleichen Beziehungen zueinander verknüpft sein. In diesem Fall ist es wichtig, ein hierarchisches System zu beseitigen und die Begriffsbeziehungen mit Hilfe eines anderen Systems darzustellen, das die Begriffsbeziehungen unterschiedlicher Art besser aufwerten und hervorheben kann (KÜDES 2018).

1.1.6 Beziehungen zwischen Benennungen und Begriffen

Nachdem nun die Begriffsbeziehungen erläutert wurden, können jetzt die Beziehungen zwischen Begriff und Benennung besprochen werden. Sehr wichtig hierbei ist, dass das Prinzip der Eineindeutigkeit gilt (Drewer et al. 2017b). Das heißt, vor dem Beginn der Erstellung einer Terminologiarbeit, muss die Terminologin bzw. der Terminologe eine einzige Benennung jedem einzelnen Begriff zuschreiben, soweit das möglich ist (Drewer et al. 2017b). Wie bereits erwähnt, ist dies nicht immer möglich, aber notwendig, um den Grad der Ambiguität zu verringern (Drewer et al. 2017b) und wird oft in Form von Vorzugsbenennungen umgesetzt.

1.1.6.1 Synonymie

Synonyme bedeutet in der Linguistik Bedeutungsgleichheit zwischen Wörtern. In der Semantik wird Synonymie auch mit dem Begriff Äquivalenzrelation bezeichnet. Aus terminologischer Sicht sieht die Synonymie laut Drewer et al. (2017b) vor, dass mehrere Benennungen einem Begriff zugeordnet werden. In natürlichen Sprachen existieren Synonyme als Ergebnis eines natürlichen Prozesses bzw. der Entwicklung der Sprachen und dienen dazu, die Sprache zu bereichern (Drewer et al. 2017b). In fachlichen Kontexten ist dies nicht der Fall. Synonyme existieren in Fachsprachen als Ergebnis eines verflochtenen Wissens, in dem sich die Terminologie gerade entwickelt und noch nicht normalisiert worden ist (Drewer et al. 2017b). Verschiedene Benennungen werden auf austauschbare Weise verwendet bis sich eine einzige Benennung etabliert hat und die anderen *verkörpert* hat (Drewer et al. 2017b). Ein weiterer zu betrachtender Aspekt bei der Synonymie ist die Abkürzung. Es existieren nämlich gleiche Abkürzungen, die in der Langform eine komplett andere Bedeutung tragen (Drewer et al. 2017b). Beispiel dafür ist die Abkürzung PKW, die sowohl für den Personenkraftwagen als auch für die Publizistik- und Kommunikationswissenschaft steht.

Die Synonymie wird nach Cruse (2004) in drei Sorten klassifiziert: absolute Synonymie, propositionale Synonymie und Plesionymie. Unter absoluter Synonymie wird eine völlige Austauschbarkeit von Wörtern in allen Kontexten verstanden. Ein Beispiel dafür sind die Verben „beginnen“ und „anfangen“, die in allen kommunikativen Situationen austauschbar sind, ohne dass sich der Sinn ändert. Dies ist allerdings ein seltener Fall, weil zwei Wörter, obwohl Synonyme, in den meisten Fällen eine leicht abweichende Bedeutung tragen (Roth 2006). Die propositionale Synonymie liegt vor, wenn zwei Ausdrücke ausgetauscht werden, ohne dass sich der Sinn des Satzes verändert. Allerdings ist das Ergebnis eines solchen Austausches ein nicht hundertprozentig idiomatischer Satz. Dies bedeutet, dass die zwei Sätze

nun die gleiche Denotation haben (bei der keine Bedeutungsverschiebung stattfindet), aber nicht mehr über dieselbe Konnotation verfügen (die Bedeutung wird aufgrund zusätzlicher Aspekte anders betrachtet). Diese Unterscheidung in der Konnotation kann aus regionalen, fachsprachlichen und/oder politischen Gründen entstehen (Roth 2006). Die letzte Sorte der Synonymie ist die Plesionymie (auch Beinahe-Synonymie genannt). Dies ist die Art der Synonymie, bei der die beiden in Betracht gezogenen Wörter zwar Synonyme sind, aber nicht vollkommen bedeutungsgleich sind. Genau aus diesem Grund müssen sie weiter spezifiziert werden. Um Missverständnisse oder Bedeutungsverschiebungen zu vermeiden, ist es notwendig, eine Spezifizierung wie zum Beispiel „x, oder besser ausgedrückt y“ anzugeben (Burstein 1997: 112). Es existiert ebenso die sogenannte Teilsynonymie (auch Quasisynonymie oder partielle Synonymie genannt). Unter Teilsynonymie versteht die Terminologiewissenschaft die Beziehung zwischen Benennungen mit identischen Begriffen, die nur in seltenen Fällen ausgetauscht werden können (Drewer et al. 2017b). Die Semantik sieht auch die partielle Synonymie als ein Zustand an, bei dem Begriffe manchmal austauschbar sind, aber nicht immer, und sie definiert diesen Prozess als Ergebnis einer Reihe von Bedingungen, Regionalismen, politischen und sozialen Entscheidungen (Drewer et al. 2017b). Beispiel dafür sind *Führer* und *Leiter*. Der erste trägt infolge historischer Ereignisse eine negative Konnotation und wird infolgedessen vom Teilsynonym *Leiter* ersetzt (Drewer et al. 2017b).

1.1.6.2 Antonymie

Unter dem Begriff Antonymie versteht man Gegensätzlichkeit und Gegensatzrelation zwischen lexikalischen Begriffen (Roth 2006). Wie für die Synonymie wurde auf einige Arten eingegangen und wird Antonymie in fünf Bedeutungsrelationen aufgeteilt: Inkompatibilität, Komplementarität, Kontrarität, Konversions-Relation und Reversivität (Roth 2006). Wenn zwei Begriffe auf den ersten Blick Synonyme zu sein scheinen, im Kontext eingebettet, aber nicht mehr Sinn ergeben, spricht man in diesem Fall von Inkompatibilität. Bei der Komplementarität handelt es sich immer um Gegensatzpaare. Aus diesem Grund wird die Komplementarität häufig auch Kontradiktion genannt. Im Gegensatz zur Komplementarität ist die Kontrarität nicht nur „weiß“ oder „schwarz“, sondern es wird ebenso betrachtet, was in der Mitte steht. Ein Beispiel dafür ist das Adjektivpaar gut-schlecht. Es handelt sich in diesem Fall nicht um die Komplementarität, weil es viele Grade und Schattierungen von Gut und Böse zwischen den zwei Extremitäten gibt (Roth 2006). In der Konversionsrelation werden dieselben

Fakten oder Aktivitäten aus zwei unterschiedlichen Perspektiven beschrieben. Diese *konversen* Paare beziehen sich auf Tätigkeiten, die im selben semantischen Feld zu finden sind, nur wird die Tätigkeit nicht aus demselben Blickwinkel betrachtet. Der Austausch: „x verkauft y das z“, kann auch so; „y kauft von x das z“, gesehen werden. Die Aktivität bleibt unverändert, nur wird sie anders betrachtet (Roth 2006). Zuletzt wird die Bedeutungsrelation der Reversivität erklärt. Es handelt sich dabei um zwei Verfahren, wo der Anfang der ersten Aktivität das Ende der zweiten Aktivität widerspiegelt und umgekehrt. Ein typisches Beispiel dafür sind einige Verben, die genau dieselbe Aktivität darstellen, wie „aufsperrn“ und „zusperren“ oder die englischen Verben „tie“ (eine Krawatte umbinden) und „untie“ (eine Krawatte losbinden) (Roth 2006). In diesem Fall spricht man in der lexikalischen Semantik von Reversiven (Roth 2006).

1.1.6.3 Ambiguität

Weitere Hindernisse, die bei der Modellierung semantischer Netze trotz der vorher aufgelisteten Klassifizierung entstehen könnten, sind Polysemie und Homonymie. Beide sind Unterarten der Ambiguität (Mehrdeutigkeit) (Drewer et al. 2017b). In diesem Fall repräsentiert eine Benennung mehrere Begriffe, wodurch Mehrdeutigkeit entsteht (Drewer et al. 2017b). Lexeme, die nur aus einer lexikalischen Einheit bestehen, versteht die lexikalische Semantik unter monoseme Lexeme (Monosemie). Verfügen Lexeme je nach Kontext über mehr als eine lexikalische Einheit bzw. mehr als eine Bedeutung, dann spricht man von Polysemie. Die Polysemie ist also in der Semantik die Fähigkeit eines Wortes, mehr als eine Bedeutung zu tragen (Roth 2006). Aus terminologischer Sicht versteht man unter Polysemie eine „Beziehung zwischen identischen Bezeichnungen mit einem erkennbaren gemeinsamen etymologischen Ursprung in derselben Sprache für unterschiedliche Begriffe“ (Drewer et al. 2017b: 17).

Im Gegensatz zur Polysemie steht die Homonymie. Unter Homonymie versteht die lexikalische Semantik ein Phänomen bei dem zwei Lexeme dieselbe orthographische und phonologische Form besitzen, ohne weder die Bedeutung noch die Etymologie zu teilen. Ein Beispiel der phonologischen, aber nicht orthographischen Ähnlichkeit sind die Wörter „Meer“ und „mehr“. Es ist selbstverständlich, dass diese zwei Lexeme zwei komplett unterschiedliche Bedeutungen tragen und nicht dieselben etymologischen Wurzeln teilen, aber aufgrund ihrer phonetischen Ähnlichkeiten können sie zu Missverständnissen führen, wenn sie nicht in einem spezifischen Kontext eingebettet sind. In Fällen, in denen zwei Lexeme gleich ausgesprochen aber anders geschrieben werden, spricht man insbesondere von Homophonie (Roth 2006). Die Homophonie stellt allerdings kein Problem bei der praktischen Terminologearbeit dar, da sich

die Terminologiarbeit mit der geschriebenen Form beschäftigt (Drewer et al. 2017b). Wenn zwei oder mehrere Lexeme gleich geschrieben werden, spricht man in der lexikalischen Semantik von Homographie (Roth 2006).

Aus terminologischer Sicht ist die Definition für Homonymie ähnlich wie die Definition für Polysemie, nur spielt in diesem Fall der etymologische Ursprung keine Rolle. Unter Homonymie versteht man die „Beziehung zwischen identischen Bezeichnungen in derselben Sprache für unterschiedliche Begriffe“ (Drewer et al. 2017b: 17). Genau die Tatsache, dass einigen Lexemen mehr als eine Bedeutung zugeordnet wird, kann bei der Wissensmodellierung ein Hindernis darstellen. Wie bereits erwähnt sind Polysemie und Homonymie bei der Fachkommunikation problematisch. Aus diesem Grund sollten Terminologinnen und Terminologen versuchen, mehrdeutige Benennungen innerhalb eines Fachgebiets zu entfernen (Drewer et al. 2017b).

1.1.6.4 Äquivalenz

Äquivalenz kann insbesondere bei mehrsprachigen Terminologien ein weiteres Problem sein (Drewer et al. 2017b). Unter Äquivalenz versteht man die Beziehung zwischen zwei Termini in unterschiedlichen Sprachen, die durch den Bezug auf denselben Begriff entsteht (Drewer et al. 2017b). Da Faktoren wie soziale und politische Einflüsse oder Regionalismen bei der Synonymie eine Rolle auf die Prägung der Termini spielen, ist eine völlige Äquivalenz selten der Fall. Durchaus kann es in Fachsprachen geschehen, dass sie Begriffe haben, die auf internationaler Ebene genormt worden sind (Drewer et al. 2017b). In diesem Fall, wenn alle Merkmale des einen und des anderen Begriffes übereinstimmen, spricht man von einer völligen Äquivalenz bzw. hundertprozentigen Übereinstimmung (Drewer et al. 2017b). Falls die Eigenschaften der Begriffe nicht hundertprozentig übereinstimmen, spricht man von Teiläquivalenz (Drewer et al. 2017b).

Obwohl die Übereinstimmung schwierig numerisch messbar ist, kann man von einer Übereinstimmung sprechen, die sich von 0% bis 99% erstreckt. Die 100-prozentige Übereinstimmung nennt man, wie gesagt, völlige Äquivalenz. Wird dieser Prozentsatz kleiner, dann spricht man von Überschneidung und Inklusion (Drewer et al. 2017b). Unter Überschneidung versteht man eine partielle Übereinstimmung zwischen zwei Termini, die allerdings einige abweichenden Merkmale aufweisen (Drewer et al. 2017b). Bei der Inklusion gibt es ebenso abweichende Merkmale, es handelt sich aber in diesem Fall von zusätzlichen Eigenschaften (Drewer et al. 2017b). Zuletzt ist die nullprozentige Übereinstimmung, die zum

Phänomen der sogenannten *faux amis* zurückzuführen ist. In diesem Fall teilen die zwei Begriffe keine gemeinsamen Begriffsmerkmale. Der Grund, weshalb zwei Begriffe als äquivalent betrachtet werden, hat mit ihrer Benennungsähnlichkeit zu tun (Drewer et al. 2017b).

Das Phänomen der Äquivalenz spielt vor allem bei mehrsprachiger Terminologearbeit eine Rolle. Bei mehrsprachiger Terminologearbeit sind auch die sogenannten terminologischen Lücken zu beachten. Terminologische Lücken können sowohl Benennungslücken als auch Begriffslücken sein (Drewer et al. 2017b). Bei Benennungslücken ist der Begriff bekannt, nur noch nicht benannt. In diesem Fall sind verschiedene Lösungen vorhanden. Terminologinnen und Terminologen können zu Entlehnungen und Lehnübersetzungen zurückgreifen, ein neues Wort in der Zielsprache prägen, das sogenannte Wortbildungsverfahren verwenden, den Begriff paraphrasieren oder, wenn das Bedürfnis nicht besteht, eine neue Benennung für einen bestimmten Begriff zu etablieren, den sogenannten lückenhaften Zustand beibehalten (Drewer et al. 2017b). Begriffslücke bedeutet, dass weder die Benennung noch der Begriff in einer bestimmten Sprache vorhanden sind. Aus diesem Grund sollen Terminologinnen und Terminologen eine Beschreibung des abwesenden Begriffes vornehmen bzw. die Bedeutung des Begriffes, der in einer anderen Sprache existiert, klären (Drewer et al. 2017b).

1.2 Wissensmanagement, Wissensmodellierung und Wissensrepräsentation

In diesem Unterkapitel werden die drei Gebiete vorgestellt, die in Bezug auf die vorliegende Masterarbeit eine wesentliche Rolle spielen. Das Hauptgebiet ist das Wissensmanagement (engl. Knowledge Management), innerhalb dessen die Wissensmodellierung (engl. Knowledge Engineering) zu finden ist und schlussendlich wird das bedeutsamste Teilgebiet der Wissensrepräsentation (engl. Knowledge Representation) behandelt, das der Kern dieser Arbeit ist.

Unter dem Begriff Wissensmanagement versteht man im Allgemeinen die Verwaltung und den Austausch von Wissen, das seit der Antike mit der Entwicklung von Bibliotheken und Kommunikationsmitteln fortgeführt wird (Frost 2020). Dieser Begriff beinhaltet alle praktischen, operativen, logistischen und strategischen Aktivitäten, die mit dem Wissen verbunden sind. Anders ausgedrückt zielt das Wissensmanagement darauf ab, zunächst das Wissen zu erwerben, und danach es zu entwickeln, zu speichern und anzuwenden (Frost 2020).

Diesbezüglich ist es wichtig, das Wissensmanagement vom Informationsmanagement zu unterscheiden, weil Wissen und Information keine Synonyme sind (Frost 2020).

Die Informationen bestehen aus Daten und legen das Fundament des Wissens. Sowohl die Informationen als auch das Wissen basieren auf reinen Fakten und Aussagen. Allerdings existieren Informationen unabhängig von jemandem oder etwas, der/die oder was sie verwenden kann (Frost 2020). Andererseits existiert das Wissen nur da, wo es einen Verstand - sei er in menschlicher oder in künstlicher Form - gibt, der es besitzen kann. Das reine Wissen findet deswegen nur statt, wenn eine Benutzerin oder ein Benutzer Informationen und persönliche Lebenserfahrungen bzw. erworbene Fähigkeiten bei einer Software zusammenfügt (Vassallo 2015). Grundsätzlich ist das Wissen nur dann vorhanden, wenn eine Intelligenz in der Lage ist, auf es zurückzugreifen und es zu benutzen. Bei einer Software sprechen wir in diesem Fall von Künstlicher Intelligenz (Vassallo 2015). Nur auf der Basis von bereits vorhandenen Informationen und vom Vorwissen kann neues Wissen erworben werden (Frost 2020).

Darüber hinaus kann Wissen in zwei Kategorien klassifiziert werden: implizites Wissen und explizites Wissen. Das explizite Wissen kann durch eine systematische und formale Sprache - Wörter oder Zahlen - vermittelt werden. Das implizite Wissen ist kontextspezifisch und basiert auf persönlichen Lebenserfahrungen, Handlungen und Intuitionen. Aus diesem Grund ist das implizite Wissen schwer vermittelbar und nicht aktualisierbar (Frost 2020). Infolgedessen kommt die Wissensmodellierung ins Spiel, die dabei hilft, das schwer explizierbare implizite Wissen zu auszudrücken und zu strukturieren (Pellegrini et al. 2006).

Unter Wissensmodellierung wird eine moderne Disziplin verstanden, die zur Künstlichen Intelligenz und zum Teilgebiet des Wissensmanagements gehört und die darauf abzielt, komplexe Probleme, die normalerweise eine hohe menschliche Spezialisierung erfordern würden, durch Expertensysteme zu lösen (Studer et al. 1998). Die Wissensmodellierung zielt darauf ab, Wissen (nicht die reinen Informationen) zu vermitteln (Studer et al. 1998). Wie oben erwähnt wird das implizite Wissen durch die Wissensmodellierung expliziert, deshalb fungiert die Wissensmodellierung als *Brücke* zwischen den mentalen Modellen bzw. implizitem Wissen und wissensbasierten Anwendungen bzw. Ontologien (Kienriech et al. 2006).

Wissensmodellierung wird meistens im textuellen Umfeld verwendet und als standardisierte semantische Modellierung verstanden (Kienriech et al. 2006). Auf diese Art und Weise werden die sogenannten semantischen Lücken und die menschentypische Ungenauigkeit überwunden, indem die typischen Ausdrücke der gesprochenen Sprache durch die formalisierte Repräsentation ersetzt werden (Hüsing 2016: 1'19"). Es wird von menschentypischer

Ungenauigkeit gesprochen, weil Menschen über unterschiedliche Hintergründe verfügen und die Welt entsprechend unterschiedlich interpretieren. Mit Hilfe von Ontologien, werden die Korrespondenzen zwischen Begriffen und Benennungen bzw. Gegenständen vereinheitlicht und lassen deshalb keinen Raum für Missverständnisse (Ehrig et al. 2006).

Nun wird versucht mit Hilfe eines konkreten Beispiels besser zu erläutern, was mit menschentypischer Ungenauigkeit gemeint wird. Person 1 und Person 2 unterhalten sich. Person 1 spricht von einem Jaguar und damit meint sie das Tier. Person 2, die einen anderen Hintergrund hat, referenziert damit das Auto. Hier scheitert plötzlich die Kommunikation. Ontologien dienen allerdings dazu, die Missverständnisswahrscheinlichkeit zu verringern (Ehrig et al. 2006).

Der Zweck ist aus diesem Grund nicht nur informativ, sondern das Wissen in ein Wissenssystem zu übertragen. Damit das Wissen dargestellt wird, sind einige Phasen notwendig. Zunächst muss das Wissen strukturiert sein, erst danach kann es repräsentiert werden und die Wissensdatenbank erstellt werden (Studer et al. 1998). Wissensmodellierungsmethoden sind von Software unabhängig. Maschinen erleichtern allerdings das gesamte Verfahren (Kienreich et al. 2006). Darüber hinaus ist das Wissen menschengebunden, weil, wie bereits gesehen, das Wissen ohne Vorkenntnisse, die auf früheren Erfahrungen basieren, reine Information wäre.

Die Wissensrepräsentation beschäftigt sich mit der Darstellung des Wissens und zielt darauf ab, in der Lage zu sein, von diesem Wissen zu Schlussfolgerungen zu kommen, neues Wissen zu schaffen oder eventuelle Probleme zu lösen (GITO 2021). Wissensrepräsentation untersucht die Art und Weise, wie menschliche Denkweise funktioniert und auf dieser Basis werden Sprachen definiert, die die Formalisierung des Wissens ermöglichen, um dasselbe Wissen in maschinenlesbarer Form darzustellen (GITO 2021). Auf diese Art und Weise wird die Maschine in die Lage versetzt, automatische Schlussfolgerungen zu ziehen und neues Wissen zu extrahieren (Minini 2021). Das Hauptziel der Wissensrepräsentation ist es, Informationsarchitekturen, die auf den Bedeutungszusammenhängen (Begriffen und Beziehungen) des Wissens basieren, zu gestalten. Verschiedene Systeme dienen zur Wissensordnung, wie zum Beispiel Glossare, Ontologien, Kataloge oder - wie in diesem Fall - terminologische Begriffssysteme (Stock et al. 2008). Ein Überblick über die Wissensrepräsentationsmodelle ist in Abb. 3 dargestellt.

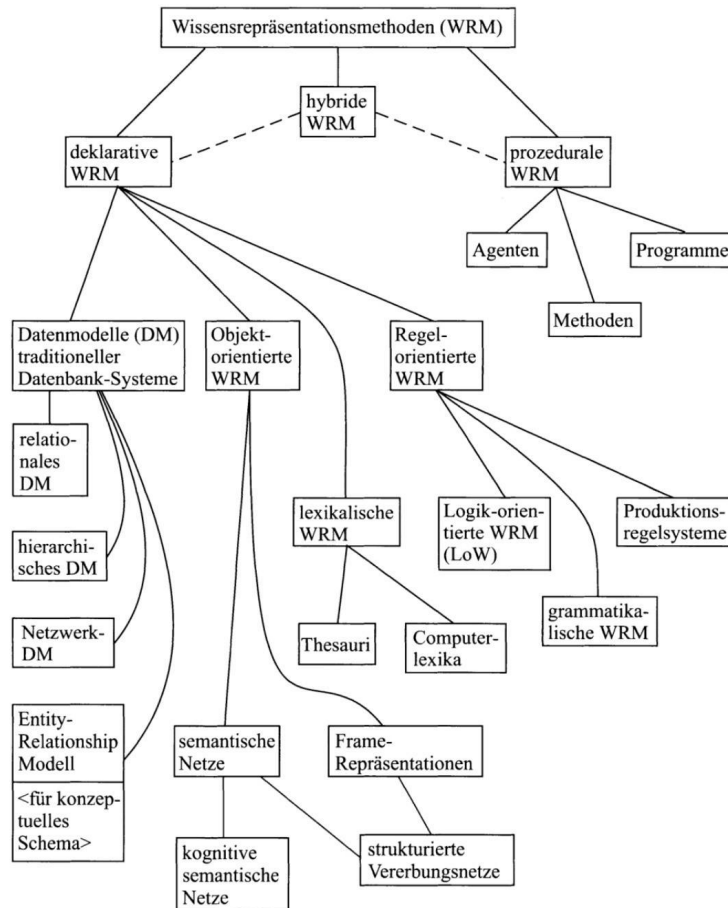


Abbildung 3. Wissensrepräsentationsmethoden (Helbig 2013)

Wie in Abb. 3 zu sehen ist, existieren zahlreiche Wissensrepräsentationsmodelle. Terminologische Begriffssysteme sind für den Zweck der vorliegenden Masterarbeit das richtige Modell. In den folgenden Unterkapiteln werden unterschiedliche semantische Netze vorgestellt und die verschiedenen Eigenschaften erklärt. Die Art und Weise, wie die Wissensvernetzung durch Begriffssysteme funktioniert wird in den nächsten Kapiteln erläutert. Auf der Basis deren werden dann im Kapitel der Methodik mögliche Repräsentationsmodelle bzw. -formate für eine multimodale Terminologie analysiert.

1.2.1 Semantische Netze

Semantische Netze haben ihren Ursprung in den frühen Jahren der zweiten Hälfte des 20. Jahrhunderts. Sie wurden zunächst von Richens im Jahr 1956 als Pivot-Sprache für die automatische Übersetzung natürlicher Sprachen konzipiert und erreichten später 1966 mit der Arbeit von Quillian ihren Erfolgshöhepunkt (Quillian 1967). Im Jahr 1969 entwickelte Quillian das erste Modell mit dem Namen Teachable Language Comprehender, das verwendet wurde, um das semantische Wissen zu beschreiben. Quillian (1967) untersuchte die Art und Weise,

wie das Wissen in dem menschlichen Gehirn gespeichert wird und bezeichnet die von uns gespeicherten semantischen Informationen mit dem Namen „semantisches Gedächtnis“ (Roth 2006). Semantische Netze sind wissensbasierte Repräsentationsmodelle. Helbig (2013) stellt eine detaillierte Definition von semantischen Netzen zur Verfügung. Laut Helbig sei ein semantisches Netz

das mathematische Modell einer Begriffsstruktur, die aus einer Menge von Begriffen und den zwischen diesen bestehenden kognitiven Relationen aufgebaut ist. Es wird in Form eines verallgemeinerten Graphen dargestellt, wobei den Knoten des Graphen die Begriffsrepräsentanten und den Kanten die Beziehungen zwischen den Begriffen entsprechen (Helbig 2013: 20).

Die Knoten stellen also die Begriffe dar, während die Kanten zwischen den Knoten die Funktionen und die Relationen darstellen. Dies ist möglich, weil der Wortschatz als hierarchisches Netz abbildbar ist (Roth 2006).

1.2.1.1 Taxonomie, Thesaurus und Ontologie

Unter den verschiedenen Arten von semantischen Netzen gibt es einige Unterkategorien, die sich je nach ihrer Funktion und ihrem Zweck unterscheiden.

Die Taxonomie beschäftigt sich meist mit der Klassifizierung von lebenden Wesen. In Taxonomien werden Begriffe hierarchisch in asymmetrischen Relationen angeordnet (Roth 2006). Genau wegen der Eigenschaft der Asymmetrie basieren Taxonomien auf semantischen Relationen wie Meronymie und Hyponymie (Roth 2006). Ein typisches Beispiel dafür ist die biologische Taxonomie. Mit Hilfe der Kriterien der biologischen Taxonomie, die sich um die Aufteilung der Organismen in Ordnung, Unterordnung, Familie, Unterfamilie, Gattung und Art kümmert, kann man zur Evolution der betroffenen lebenden Wesen zurückkehren.

Der Thesaurus, auch Wortnetz oder einsprachiges Fachwörterbuch genannt, ist die Gesamtheit der Termini in einem spezifischen Fachgebiet, die miteinander hierarchisch und nach Synonymie-Beziehungen verbunden sind (Vigini 1985). Da sich alle Termini auf dasselbe Gebiet beziehen, werden sie miteinander in einem sinnrelationalen Netz verbunden. Die lexikalisch-semantischen Relationen, auf denen die Thesauri basieren, sind Hyponymie, Hyponymie, Synonymie, Antonymie und teilweise sogar Meronymie (Vigini 1985).

Sowohl Taxonomien als auch Thesauri sind *Knowledge Organisation Schemes* (KOS). KOS ist ein allgemeiner Begriff, der sich auf eine Menge von Elementen bezieht, die

strukturiert und klassifiziert werden müssen. Mit Hilfe von KOS werden Objekte beschrieben oder Bücher in einem Regal angeordnet. In diesem Sinne weichen Taxonomien und Thesauri von Ontologien ab, weil KOS für Wissensmanagement verwendet werden, während Ontologien für die Wissensrepräsentation verwendet werden (Fernandes 2015). Die Themen Wissensmanagement und Wissensrepräsentation werden im nächsten Unterkapitel erläutert.

Der Begriff Ontologie mit der heutigen Bedeutung entstand im Zusammenhang mit der Untersuchung der Probleme der Semantik der Information und der Klassifizierung von Mengen von Begriffen in den 1980er Jahren. Ontologien bezeichnen die Terminologie, die ein bestimmtes Sachgebiet (engl. domain) repräsentiert, deren Darstellung in Klassen oder Konzepten sowie logische Verknüpfungen zwischen diesen Klassen oder Konzepten (Fernandes 2015). Um eine Ontologie zu erstellen, ist es notwendig, die Konzepte, aus denen die Ontologie besteht, auf eine hierarchische Art und Weise zu strukturieren und die semantischen Beziehungen zwischen den Konzepten zu definieren und beschreiben (Staab et al. 2010). Mit Hilfe von Ontologien werden in der Informatik komplexe Sachgebiete auf eine semantische Art und Weise repräsentiert (Staab 2010). Ontologien basieren auf Klassen und Relationen. Unter dem Begriff „Klasse“ wird in der Ontologie ein bestimmtes Fachgebiet mit den dazugehörigen Gegenständen bezeichnet, die Gemeinsamkeiten aufweisen. Während man unter dem Begriff „Relation“ bei Ontologien die Beziehungen zwischen Klassen versteht (Roth 2006). Bei Ontologien verfügen sowohl die Klassen als auch die Relationen über Eigenschaften (Drewer et al. 2017a).

Abb. 5 dient als Beispiel für eine Ontologie. Die Ellipsen stellen Klassen dar, während die Pfeile zwischen den Klassen Relationen darstellen. Die dicken Pfeile stellen die Vererbungen dar, wie zum Beispiel zwischen dem Hyperonym *Kunstwerk* und seinen Hyponymen *Skulptur* und *Bild*. Die Relationen, die durch die normalen Pfeile dargestellt werden, sind in einer inversen Beziehung zueinander verbunden. Durch diese nicht-hierarchischen Beziehungen ist es stets möglich einen Begriff zu einem anderen zurückzuverfolgen (Drewer et al. 2017a). Das heißt, die Richtung ist in beide Richtungen lesbar. In Abb. 4 ist ein *Bild gemalt Von Maler*, aber der *Maler malt* das *Bild*.

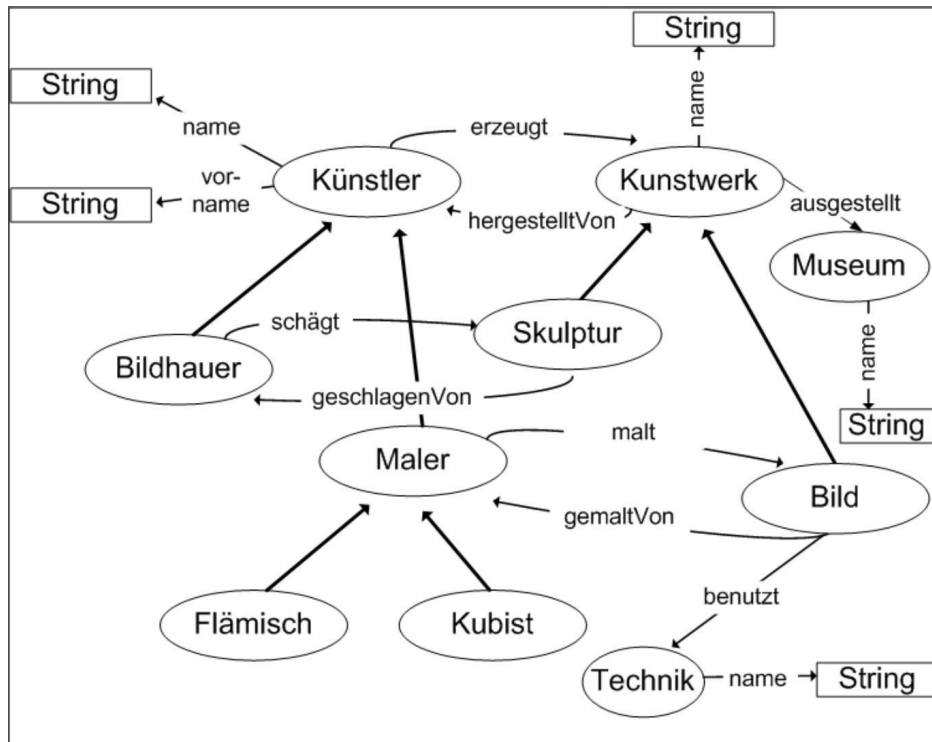


Abbildung 4. Eine Ontologie (Drewer et al. 2017a)

Ontologien können als das Herzstück des Semantic Web betrachtet werden, weil sie das Wissen eines bestimmten Sachgebiets formal repräsentieren können, indem ihre Modellierung auf logischen, formalen Sprachen basieren (Pellegrini et al. 2006). Wie bereits erwähnt, beschreiben Ontologien Konzepte und deren Beziehungen innerhalb eines Sachgebiets und unterstützen Maschinen dabei, Inhalte ohne Hilfe des menschlichen Eingriffes zu interpretieren (Pellegrini et al. 2006). Mit Hilfe logischer Axiome können bestimmte Anwendungen, sogenannte Reasoner, automatische Schlussfolgerungen aufgrund des dargestellten Wissens ziehen (Pellegrini et al. 2006).

Ontologien basieren auf semantischen Netzen so wie zum Beispiel die eben genannten Taxonomien und Thesauri und befinden sich aus diesem Grund an der Spitze der Skala der semantischen Reichhaltigkeit, die in Abb. 5 dargestellt ist (Pellegrini et al. 2006).

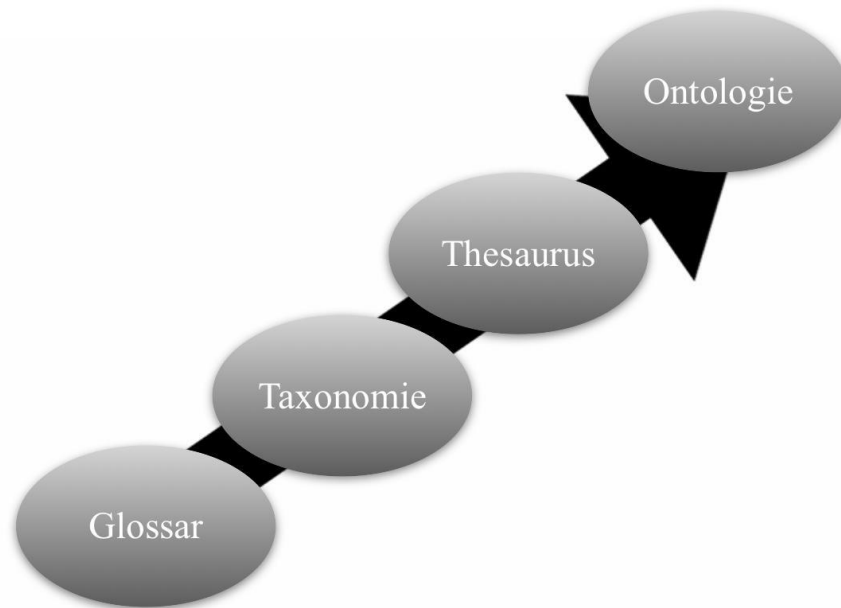


Abbildung 5. Semantische Reichhaltigkeit (Pellegrini et al. 2006: 19)

Die semantische Reichhaltigkeit steigt von links unten nach rechts oben. Bei einem Glossar wird eine Liste von Begriffen und der dazugehörigen Definitionen erstellt (Pellegrini et al. 2006). Bei der Taxonomie wird die hierarchische Struktur mit Hyperonymen und Hyponymen hinzugefügt. Thesauri verfügen noch zusätzlich über lexikalisch-semantische Relationen, wie zum Beispiel Synonymie, Antonymie, usw. (Vigini 1985). Wobei die Beschreibungen der Beziehungen zwischen Objekten bei Ontologien auf logischer Basis definiert werden. Die Konzeptionalisierung der Objekte ist die Haupteigenschaft der Ontologien. Das heißt, dass die Objekte Klassen zugeordnet werden. Es ist bei Ontologien sogar möglich, die Relationen zwischen Klassen zu beschreiben (Pellegrini et al. 2006).

Genau aufgrund ihrer semantischen Reichhaltigkeit werden Ontologien für unterschiedliche Zwecke verwendet (Pellegrini et al. 2006). Sie ermöglichen zum Beispiel:

- den Datenaustausch zwischen unterschiedlicher Software;
- die Bildung von Theorien;
- die Erleichterung der Kommunikation und des Wissensaustausches zwischen Menschen.

Nachdem nun erklärt wurde, warum Ontologien für die Wissensrepräsentation besser als andere semantische Netze sind, ist es wichtig zu verstehen, wie genau die Wissensvernetzung durch Ontologien funktioniert.

1.2.1.2 Wissensvernetzung durch Ontologien

Nach einer grundlegenden Erklärung des Begriffs der Ontologien, soll dieses Kapitel nun Einblicke in die Arten der Relationen einer Ontologie geben. In Abb. 5 wurde die Wichtigkeit der Relationen bei einer Ontologie hervorgehoben. In diesem Unterkapitel werden diese Relationen ausführlicher erklärt, um aus rein praktischer Sicht zu verstehen, wie eine Ontologie erstellt wird.

Zunächst einmal ist es notwendig, die Relationen vorzudefinieren (Ehrig et al. 2006). Die bekannteste und am häufigsten eingesetzte Relation ist die sogenannte *is-a-Beziehung*. Diese Art von Relation setzt Hyper- und Hyponyme in Verbindung (Ehrig et al. 2006). In Abb. 6 wird ein Beispiel für eine Ontologie gezeigt und diese Relation ist leicht erkennbar.

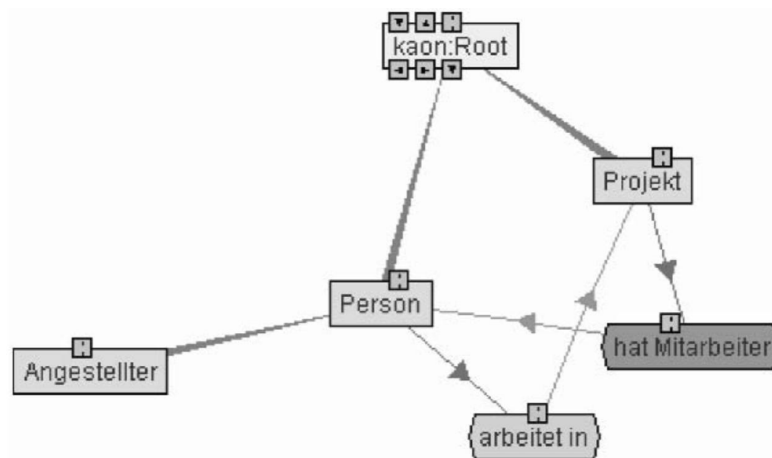


Abbildung 6. Eine andere Ontologie (Ehrig et al. 2006: 473)

In diesem Fall ist die *is-a-Beziehung* zwischen den Begriffen *Angestellter* und *Person* definiert, wo der Begriff *Person* das Hyperonym und der Begriff *Angestellter* das Hyponym ist (Ehrig et al. 2006). In diesem Fall handelt es sich um eine asymmetrische Relation und sie ist aus diesem Grund nicht invers (Roth 2006). Die Angestellten sind nämlich Personen, nicht alle Personen sind aber Angestellte. Es wird in diesem Fall vom Vererbungskonzept gesprochen (Ehrig et al. 2006). Es handelt sich um ein Konzept, das vorsieht, dass die Eigenschaften eines Hyperonyms seinem bzw. seinen Hyponymen vererbt werden. Mit Hilfe des Vererbungskonzeptes können

infolgedessen kompaktere Ontologien erstellt werden, weil die Beziehungen nur einmal für ein Hyperonym definiert werden und sind trotzdem für alle seine Hyponyme implizit spezifiziert (Ehrig et al. 2006).

In Abb. 7 ist ein weiteres Beispiel für eine Ontologie zu sehen. In diesem Fall werden die *is-a-Beziehungen* klar definiert.

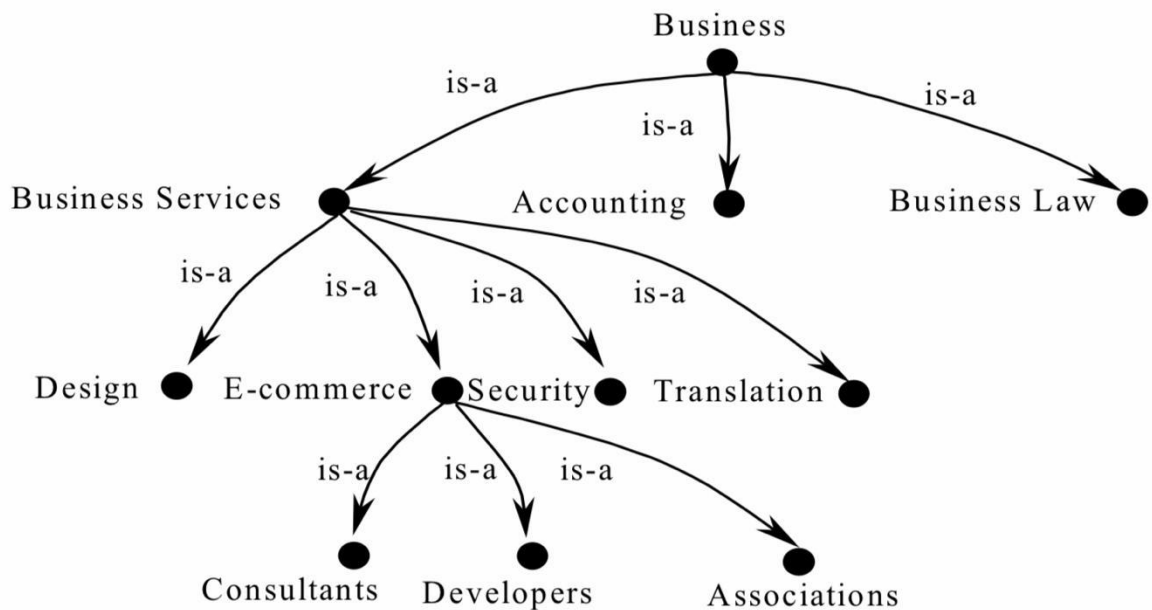


Abbildung 7. Ontologie mit is-a-Beziehung (Giunchiglia et al. 2003)

Ontologien erlauben es, unendliche Beziehungen zwischen Begriffen, und sogar zwischen Oberbegriffen, zu definieren. Man spricht in diesem Fall von anwendungsspezifischen Relationen. Ein Beispiel dafür in Abb. 6 ist die spezifizierte Relation zwischen *Projekt* und *Angestellter* mit Hilfe von *arbeitet in*. Es lässt sich also deutlich ableiten, dass Angestellte in Projekten tätig sind (Ehrig et al. 2006).

Invertiert ist diese Beziehung *hatMitarbeiter* und geht von *Projekt* zu *Angestellter*. In diesem Fall bedeutet es, dass Projekte Angestellte als Mitarbeiter haben. Das Vererbungskonzept ist auch bei dieser Relation gültig. Deswegen ist es nicht notwendig zu spezifizieren, welche Arten von Projekten Angestellte haben. Unabhängig davon, ob es sich um Sanierungs- oder Kunstprojekte handelt, es liegt auf jeden Fall ein Hyponym vom Hyperonym *Projekt* vor. Deshalb verfügen sie alle über *Mitarbeiter* (Ehrig et al. 2006).

Dies sind nur ein paar Beispiele, um zu verstehen, wie notwendig es ist, die Beziehungen innerhalb einer Ontologie a priori zu definieren. Für jede Ontologie ist es jedoch notwendig,

spezifische Relationen für die zu erstellende Ontologie zu definieren. Aus diesem Grund werden die spezifischen für diese Masterarbeit relevanten Relationen im Kapitel der Methodik geklärt.

1.2.1.3 Terminologische Begriffssysteme

Terminologische Begriffssysteme werden in der vorliegenden Masterarbeit als Repräsentationsmodell verwendet, obwohl sie zurzeit noch nicht an Relevanz gewonnen haben. Der Hauptgrund dafür liegt in der Tatsache, dass der Prozess der automatischen Extraktion von Ontologien aus einem Text immer noch eine Illusion ist und die manuelle Extraktion zu zeit- und kostenaufwändig ist (Gromann 2020).

Terminologische Begriffssysteme setzen sich aus Begriffen und ihren Beziehungen zu sprachlichen Definitionen zusammen. Der Zusammenhang von Informationen kann ebenso aus Begriffen aus mehreren Fachgebieten bestehen (Gromann 2020). Wie andere semantische Netze stellen auch terminologische Begriffssysteme ein Mittel zur Verwaltung und Erweiterung des Wissens in Fachgebieten dar (Gromann 2020).

Beispiele für eine Methode zur Erstellung eines multimodalen Begriffssystems werden von Rodríguez et al. (2013) vorgeschlagen. Folglich werden ihre Prinzipien aufgelistet:

1. Alle multimodalen Ressourcen sollen für den in die *Terminological Knowledge Base* TKB eingetragenen Konzept relevant sein. Eine Beziehung, obwohl unterschiedlicher Art, muss stattfinden. Sie kann zur Interpretation des Konzeptes entweder als *Entity* führen, als wäre es eine eigene Entität, die von einer Ursache ausgeht und zu einer Wirkung führt, oder als *Event*, das heißt durch Verfahren und Ergebnisse oder noch als *Property*, indem die mit dem Konzept verbundenen Beziehungen hervorgehoben werden (Rodríguez et al. 2013). Darüber hinaus sollten die Konzepte bzw. die eingetragenen Begriffe definiert werden (Rodríguez et al. 2013). Wenn man beispielsweise auf EcoLexicon auf den Begriff *Water* klickt, ist folgende Definition zu sehen: „colorless, tasteless, and odorless substance that is essential to all known forms of life and is the most universal solvent“ (EcoLexicon 2021).
2. Multimodale Ressourcen sollten ebenso dabei helfen, relativ abstrakte Konzepte zu konkretisieren (Rodríguez et al. 2013). Anders ausgedrückt sollte ein Verfahren stattfinden, das die Konzepte von *langue* in *parole* umwandelt, sodass das Wissen eindeutig verwaltet wird und die TKB keinen Interpretationsspielraum erlaubt (Rodríguez et al. 2013).

3. Multimodale Ressourcen sollen für jeden Begriff sowohl die Intension als auch die Extension definieren (Siehe 1.1.4.1) (Rodríguez et al. 2013).
4. Für die Zusammenfügung von textuellen und visuellen Ressourcen sind folgende Prinzipien ebenso zu berücksichtigen: das *principle of multimedia*, um verschiedene Ressourcen zu verwenden, das *principle of contiguity*, unterschiedliche Ressourcen sollten zeitlich und räumlich nebeneinander gestellt werden, das *principle of coherence*, Entzug irrelevanter Teile vom Ausgangstext und das *principle of background knowledge*, das vorsieht, dass multimodale Ressourcen an den Vorkenntnissen des Zielpublikums angepasst werden sollen (Rodríguez et al. 2013).

1.2.1.4 Darstellungsformate von semantischen Netzen

Zwei technologische Markup-Sprachen sind die für die Entwicklung des Semantic Web zentral: Extensible Markup Language (XML) und Resource Description Framework (RDF) (Dobratz 2008). XML ist eine Auszeichnungssprache, die es erlaubt, Daten auf eine hierarchische und strukturierte Weise bzw. mit den dazugehörigen Bedeutungen und Beziehungen darzustellen (Dobratz 2008). Solche Daten sind sowohl von Maschinen als auch von Menschen lesbar. Während XML eine strukturierte Gestaltung der Daten liefert, wird keine Information über die Struktur dieser Daten geboten (Dobratz 2008). Um Daten in einheitlicher Struktur darzustellen, ist RDF erforderlich. RDF drückt nämlich die Bedeutung bzw. die Semantik der Struktur durch lediglich drei Komponenten aus: Subjekt, Prädikat und Objekt. Deshalb wird RDF auch RDF Triples genannt. Diese Komponenten werden am Beispiel des folgenden Satzes erläutert: Bei „Alvin lives in New York“ wäre „Alvin“ das Subjekt, „livesIn“ das Prädikat bzw. die Beziehung und „New York“ das Objekt. Auf diese Art und Weise werden Datensammlung, Datenintegration und Interoperabilität ermöglicht (Dobratz 2008). Für Ontologien wurde allerdings die Web Ontology Language (OWL) entwickelt (Dobratz 2008). OWL dient dazu, die Informationen nicht nur vorzustellen, sondern auch, sie zu verarbeiten (Dobratz 2008). Im Gegensatz zu XML und RDF, ermöglicht OWL bessere Interpretationsmöglichkeiten der Daten mit Hilfe von einem Zusammenhang von Semantik und zusätzlichem Vokabular bzw. Beschreibung der Relationen (Dobratz 2008), sofern Daten in OWL mit Hilfe logischer Axiome korrekt dargestellt werden. Das ist ein Grund dafür, dass OWL für die Erstellung von Ontologien ausgezeichnet funktioniert, da OWL die Möglichkeiten der Darstellung aller notwendigen Informationen für die korrekte Interpretation derselben Ontologie zur Verfügung stellt sowie zum Beispiel die Beschreibungen von Klassen (Dobratz 2008). Wird OWL jedoch

nicht mit logischen Axiomen versehen, gibt es auch die Möglichkeit der automatischen Schlussfolgerung nicht.

Im Unterkapitel 1.2.1.1 wurden KOS vorgestellt bzw. Thesauri und Taxonomien. Nun wird SKOS vorgestellt. Das Simple Knowledge Organization System (SKOS) ist ein Projekt bezüglich der Darstellung von lexikalischen und besonders Thesaurus-bezogenen Informationen im breitesten Sinn im Semantic Web (Dobratz 2008). SKOS kann in RDF dargestellt werden, die zur maschinenverständlichen Darstellung von lexikalischem Wissen dient (Dobratz 2008). Darüber hinaus zielt SKOS darauf ab, das Vokabular für das semantische Web auf eine Art und Weise darzustellen, die die vorher aufgelisteten Voraussetzungen erfüllt bzw. das Wissen auf eine kontrollierte, strukturierte und maschinenlesbare Art und Weise darzustellen (Dobratz 2008). Was die Zusammensetzung vom SKOS-Datenmodell angeht, besteht es aus einer OWL-Ontologie, die aus zwei Elementen zusammengesetzt ist: Klasse und Eigenschaft (Dobratz 2008). SKOS ist also eine wertvolle Ressource zur Repräsentation unterschiedlicher Begriffsschemata (Dobratz 2008). Weder RDF noch SKOS sind aber formale Sprachen für die Wissensrepräsentation. Infolgedessen, da das Endprojekt der vorliegenden Masterarbeit ein terminologisches Begriffssystem ist, wird dafür TBX verwendet, um den Austausch der Terminologie zwischen CAT-Tools zu ermöglichen.

1.3 Multimodalität

In vielen Kontexten sind Text und Bild miteinander verbunden. Dies ist der natürlichste und älteste Zusammenhang auf dieser Welt, da der Mensch seit Jahrtausenden Darstellungen geschaffen hat, die sowohl Bilder bzw. visuelle Darstellungen als auch Textmittel, das heißt Wörter, Buchstabenfolgen, enthalten. Allerdings ist es nicht einfach, die Text-Bild-Beziehungen zu untersuchen, weil sich in der heutigen Welt eine recht deutliche Trennlinie zwischen den eher visuellen und den eher textlichen Disziplinen ziehen lässt. Um diese Untersuchung durchführen zu können, ist es jedoch notwendig, die Grenzlinie zu überschreiten und zu versuchen, Bilder als Text und Text als Bild zu sehen (Bateman 2014). Das heißt, dass die unterschiedlichen semiotischen Ressourcen, die in einem Text zu finden sind, als interaktive textuelle Elemente analysiert werden sollen (Rodríguez et al. 2013). Diese Untersuchung, die zu einem besseren Verständnis der Text-Bild-Beziehung führt, fließt in den Bereich der Multimodalität ein (Bateman 2014).

Unter dem Begriff Multimodalität versteht man ein Phänomen, bei dem verschiedene semiotischen Modi - das heißt verschiedene Zeichen, wie gesprochene Sprachen, visuelle

Darstellungen, Gesten - kombiniert und in kommunikative Situationen integriert werden (Jewitt 2011).

Laut Lemke (1998) könne die Zusammenstellung von Texten und Bildern als Bedeutungsverstärker (engl. *meaning multiplication*) betrachtet werden. Seiner Meinung nach ist die Zusammenstellung unterschiedlicher *Modi* mehr wert als die rein allein eingesetzte Information, die man von den *Modi* bekommen würde. Die Gegenüberstellung von Bildern und Textquellen dient nicht nur dazu, ein Konzept zu verdeutlichen, sondern die Zusammenfügung visueller Ressourcen und Textressourcen zielt darauf ab, sich gegenseitig zu *kompensieren* und daher ein Mehrwert zu schaffen (Lemke 1998). Das folgende Zitat von Benveniste bestätigt die Meinung von Lemke und von vielen anderen Forscherinnen und Forschern und Theoretikerinnen und Theoretikern unterschiedlicher Disziplinen:

Semiotic systems are not 'synonymous'; we are not able to say 'the same thing' with spoken words that we can with music, as they are systems with different bases. In other words, two semiotic systems of different types cannot be mutually interchangeable (Benveniste 1986: 235).

Infolgedessen ist es möglich zu behaupten, dass es zahlreiche *Modi* bzw. semiotische Systeme gibt, um eine Bedeutung auszudrücken. Allerdings kann man nicht davon ausgehen, dass die Zusammenstellung unterschiedlicher *Modi* in allen Fällen für die Übermittlung der Bedeutung wirksam ist. In manchen Situationen ist es notwendig, besonders vorsichtig zu sein, um erfolgreich kommunizieren zu können (Bateman 2014).

Es existieren mehrere Ebenen der Multimodalität. Die multimodalen Korpora reichen von Korpora, die nur Text und Bilder enthalten, bis zu Korpora, die reich an visuellen Darstellungen, Aufzeichnungen von Gesten, Haltungen, Blicken und emotionalen Zuständen sind (Allwood 2008). Da das menschliche Gehirn multimodal funktioniert, das heißt mit Hilfe unterschiedlicher sensomotorischer Eindrücke, ist es dank der Multimodalität möglich, externe Informationen leichter zu interpretieren. Multimodale Botschaften wirken wie „*schnelle Schüsse ins Gehirn*“ (Kroeber-Riel 1993: 53). Bei Sprache-Bild-Texten ist der Wahrnehmungsprozess aus diesem Grund schneller.

Kiros et al. (2014) beschäftigen sich mit dem Thema der Multimodalität und der Sprachmodelle. Laut den Autoren sind multimodale Modelle überall zu finden, zum Beispiel in der Werbung. Unter Multimodalität sollte man nicht nur die Zusammenstellung zwischen Text und Bildern verstehen, sondern auch Text und Ton, und alle Verknüpfungen, die Funktionsweise des menschlichen Gehirns nachahmen.

Der Kontext spielt nämlich eine wesentliche Rolle in Bezug auf die Zuschreibung von Bedeutung (Kiros et al. 2014). Folgende Abbildungen sollten dazu dienen, dieses Konzept besser zu erläutern.

In Abb. 8 wird die Darstellung eines Menschen gezeigt. Obwohl die Zeichnung sehr stilisiert ist, ist man trotzdem in der Lage, auf den Begriff im Kopf zurückzugreifen (Siehe 2.2.1) und Ähnlichkeiten mit realen Menschen zu finden. Dieser Entwurf verfügt über zwei Beine, einen Rumpf, zwei Arme und einen Kopf mit zwei Augen und einem Mund. Egal welchen kulturellen Hintergrund man besitzt, dieses Bild liefert allen Menschen in der Welt eine klare Bedeutung (Bateman 2014).

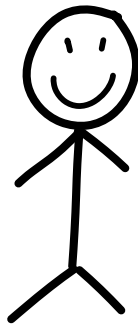


Abbildung 8. Stilisierter Mensch

Bilder können aber auch unvertraut erscheinen und nicht direkt interpretierbar sein, wie Abb. 9 illustrieren soll.



Abbildung 9. Ein Kreis

Abb. 9 erfordert zweifellos ein längeres Nachdenken und wahrscheinlich ist es nicht möglich, zu einer eindeutigen Schlussfolgerung zu kommen. Es könnte ein normaler Kreis sein, ein Rad, die Sonne oder ein Zeichen aus dem Maya-Kalender. Ohne weitere Informationen ist eine zuverlässige Interpretation jedoch schwierig. Im Kontext von Abb. 10 wird die Interpretation

von Abb. 9 jedoch deutlicher für all diejenigen, die mit diesem Fachbereich vertraut sind (Bateman 2014).



Abbildung 10. Ein Kreis im Konsektivdolmetschen

Es handelt sich um ein Zeichen, das in dem Bereich des Dolmetschens sehr bekannt ist. Viele Dolmetscherinnen und Dolmetscher verbinden dieses Zeichen mit der Vorstellung von Konferenz, Treffen, Gipfel usw.

Durch diese Beispiele sollte verdeutlicht werden, dass zwei Voraussetzungen erfüllt werden müssen. Zunächst ist es notwendig, mit einem bestimmten Thema vertraut zu sein. Zweitens müssen die Bilder bzw. die Informationen in einem bestimmten Kontext eingebettet sein, sodass die Interpretation ermöglicht wird (Schwarz-Friesel et al. 2014).

Wie bereits erwähnt, haben die oben vorgestellten Abbildungen dazu gedient, die Wichtigkeit der Kontextualisierung hervorzuheben. Umso wichtiger ist es bei Ontologien, die Bedeutung der einzelnen Begriffe zu umreißen, um Bedeutungsfehler zu vermeiden.

1.4 Medizinische Fachsprache

Unter medizinischer Fachsprache versteht man einen bestimmten Wortschatz, der verwendet wird, um den menschlichen Körper und die damit verbundenen Bedingungen zu beschreiben (Beddies et al. 2015). Wie alle Fachsprachen hat sich die medizinische Sprache im Laufe der Zeit entwickelt und wird heute in den Anwendungsbereichen der Medizin und der Pflege eingesetzt (Beddies et al. 2015). Die medizinische Fachsprache verfügt über eine regelmäßige Morphologie und verwendet einen systematischen Ansatz, um die Wörter zu bilden (Beddies et al. 2015). Der systematische Ansatz der medizinischen Fachsprache basiert auf Präfixen und Suffixen. Das Präfix wird vor den Terminus gestellt, um dem Terminus zusätzliche Informationen zuzuschreiben (Beddies et al. 2015). Es kann sich um jede Art von Information

handeln, wie beispielsweise die Lage eines Organs (Beddies et al. 2015). Suffixe werden am Ende des Terminus angehängt und fügen Informationen wie das Verfahren einer Krankheit, die Bedingungen oder den Zustand hinzu (Beddies et al. 2015). Obwohl Latein als auch Altgriechisch in der medizinischen Fachsprache verwendet werden, werden innerhalb eines Terminus nicht beide verwendet. Lateinische Präfixe sehen vor, dass es im Wort (falls notwendig) ein lateinisches Suffix geben wird und so dieselbe Grundregel gilt auch für die Einheitlichkeit von griechischen Affixen (Beddies et al. 2015). Besser ausgedrückt: Ein medizinischer Terminus hat in der Regel keine Zusammensetzung von beiden Sprachen (Beddies et al. 2015).

Dies ist lediglich für die Sprachen gültig, deren medizinische Fachsprache aus dem Latein und dem Altgriechischen stammt. In der deutschen Sprache gilt dies nicht für alle Wörter mit germanischem etymologischem Ursprung. Porep und Steudel (1974) haben die Anzahl der medizinischen Termini in der deutschen Sprache geschätzt. Laut ihrer Schätzungen betrage die Anzahl der Namen der Medikamente 80.000, 10.000 die Anzahl von Körperteilen und Organen(teilen), 20.000 die Namen der Organfunktionen und 60.000 die Untersuchungsmethoden, Krankheitsbezeichnungen und Operationsmethoden. Das Ergebnis beträgt rund 170.000 Fachbegriffe (Henne et al. 1978).

In der Medizin ist es wichtig, die Fachsprache und die Nomenklatur zu unterscheiden und sie zu definieren. Die Fachsprache ist offen für die externen Einflüsse bzw. Paradigmenwechsel, soziale Veränderungen usw. Darüber hinaus hängt jeder Fachbegriff von der Kultur ab, in der dieser Begriff geprägt wurde. Die Nomenklatur hingegen sieht eindeutige Termini vor und lässt keinen Raum für Synonyme und Doppelnamen (Henne et al. 1978).

Noch ein wichtiger Aspekt des medizinischen Bereichs ist die Tatsache, dass die medizinische Sprache im deutschsprachigen Raum zwei oder mehr Benennungen (die deutsche Version und die Latinismen) für jeden Gegenstand bzw. Krankheit, Therapie, Diagnose usw., verwendet. In dieser Hinsicht ist es wichtig, mit Vorsicht vorzugehen und eine eindeutige Methode zur Klassifizierung der Benennungen zu finden.

Infolge dieses Exkurses über die medizinische Sprache hat es sich wahrscheinlich erahnen lassen, dass die medizinische Sprache ein studierter Jargon ist, in dem Präfixe, Suffixe oder weitere Bestandteile eines Terminus kohärent sein müssen. In Wirklichkeit verhält sich die medizinische Sprache, wie die Umgangssprache. In beiden Fällen handelt es sich um natürliche Sprachen, die lebendig sind (Beddies et al. 2015). Es handelt sich aus diesem Grund um eine Sprache, die sich ständig verändert und die in ständiger Interaktion mit externen Einflüssen ist und in der, neue Termini geprägt werden (Beddies et al. 2015). Im Falle einer

Benennungslücke führen die Autorinnen und Autoren einer bestimmten Publikation einen neuen Terminus in die Publikation ein und dieser Terminus wird später von anderen Expertinnen und Experten oder Leserinnen und Lesern der Publikation abgerufen, wiederverwendet und, wahrscheinlich, sogar modifiziert werden (Beddies et al. 2015). Auf diese Art und Weise findet die natürliche Veränderung der Medizinsprache statt (Beddies et al. 2015).

An dieser Stelle ist es sinnvoll, die medizinische Terminologie aus einer terminologischen Perspektive zu analysieren. Es ist jedoch notwendig, die terminologischen Unterschiede zwischen Terminologie und Nomenklatur zu klären (Beddies et al. 2015). Unter Terminologie versteht man die Lehre eines wissenschaftlichen Jargons. Wobei man unter Nomenklatur ein Zeichensystem bzw. Bezeichnungssystem versteht, das darauf abzielt, eine breite Ressource an Termini zu ordnen bzw. zu benennen (Beddies et al. 2015). Diesbezüglich kann beispielsweise die historische Entwicklung der anatomischen Nomenklatur erwähnt werden (Beddies et al. 2015). Aufgrund der großen Menge an Rezeptionswegen des medizinischen Jargons, deren Wurzeln in der Renaissance stammen, ist heutzutage eine systematische und einheitliche Ordnung notwendig (Beddies et al. 2015). Um dieses Bedürfnis entgegenzukommen, wurde am Anfang des 19. Jahrhunderts die erste Kommission gegründet, die darauf abgezielt hat, die anatomische Nomenklatur zu vereinheitlichen (Beddies et al. 2015). Seit diesem Zeitpunkt wurden in den nächsten zwei Jahrhunderten weitere Kommissionen mit demselben Ziel gegründet bis zur heutig international anerkannten Internationalen Nomina Anatomica (INA) (Beddies et al. 2015).

Zur Prägung neuer Termini sind laut der Internationalen Nomina Anatomica einige Prinzipien zu befolgen bzw. einige Voraussetzungen zu erfüllen (Beddies et al. 2015). Der erste Grundsatz sieht vor, dass Eigennamen vermieden werden sollten (Beddies et al. 2015). In Bezug auf die Organe sollten sie eindeutig bezeichnet werden und die Organe, die durch eine topographische Beziehung verbunden sind, sollten über ähnliche Bezeichnungen verfügen wie beispielsweise *Vena femoralis* und *Arteria femoralis* (Beddies et al. 2015). Zuletzt sollten die Ausdrücke kurz, beschreibend, leicht merkbar und aus lateinischen Wörtern gebildet werden, sein (Beddies et al. 2015).

Für die Internationale Nomina Anatomica ist eine bestimmte Schreibweise in der Medizinsprache zu befolgen. Das erste Wort eines medizinischen Terminus auf Deutsch wird immer großgeschrieben, auch wenn es sich um ein Adjektiv handelt, beispielsweise *Arteria femoralis* (Beddies et al. 2015). Die zweite Regel bezüglich der Schreibweise der Internationalen Nomina Anatomica sieht vor, dass alle Diphthonge durch einzelne Vokale

ersetzt werden, beispielsweise wird *Oesophagus* durch *Esophagus* ersetzt (Beddies et al. 2015). Darüber hinaus gelten ebenso feste Regeln für die Abkürzungen. Entspricht beispielsweise die Abkürzung *A.* für *Arteria*, dann ist *Aa.* die Abkürzung für die Pluralform *Arteriae* (Beddies et al. 2015).

Allerdings hat sich diese Schreibweise in der deutschen Medizinsprache kaum durchgesetzt (Beddies et al. 2015). Im Deutschen gibt es Termini, die im Laufe der Zeit unverändert geblieben sind (Beddies et al. 2015). Dies stellt Hindernisse zur Verständigung solcher Termini auf internationaler Ebene dar (Beddies et al. 2015). Aus diesem Grund wird sich bemüht, Erkrankungen, Symptome und Syndrome zu Dokumentationszwecken in der deutschen Medizinsprache zu klassifizieren, nummerieren und ordnen (Beddies et al. 2015).

1.4.1 Medizinische Datenbanken

Als Basis für eine terminologische Ressource werden eine oder mehrere Quellen benötigt (KÜDES 2018). Als Informationsquellen werden in diesem Fall medizinische Datenbanken verwendet. Im Web wird eine bedeutsame Anzahl an medizinischen Datensätzen von Forschungsinstituten, Krankenhäusern usw. zur Verfügung gestellt. Allerdings ist es wichtig zu verstehen, welche Datenbanken bzw. Ressourcensammlungen die genügenden Bedingungen für die zu erstellende Terminologiearbeit erfüllen (KÜDES 2018). Einige Beispiele für multimodale Datenbanken, die also nicht nur Text enthalten, sondern auch non-verbale Informationen, sind die folgenden: ImageNet und MedPix®. Eine umfangreiche Sammlung von medizinischen Datenbanken wird von Beam (2020) bereitgestellt.

ImageNet ist ein Trainingsdatensatz für Machine Learning Programme, der reich an Bildern und Substantiven ist. Zurzeit beträgt die Anzahl der Substantive circa 20.000 und zu jedem Substantiv sind durchschnittlich 500 Bilder durch die WordNet-Hierarchie angeordnet. Das WordNet ist ein lexikalisch-semantisches Netz dessen Inhalt aus Wörtern besteht, die miteinander durch lexikalisch-semantische Beziehungen verknüpft sind (Fellbaum 2010).

MedPix® ist ein medizinischer, kostenloser und online-zugänglicher Datensatz. Diese Datenbank stellt denjenigen medizinische Informationen zur Verfügung, die bereits im Gesundheitswesen tätig sind, den Medizinstudierenden oder einfach denjenigen, die neugierig sind oder sich weiterbilden möchten. Das Material erstreckt sich von medizinischen Bildern bis zu klinischen Themen. Die Gesamtheit dieses Materials beläuft sich auf 9.000 Themen, 12.000 PatientInnenfälle und ungefähr 59.000 Bilder. Dazu enthält jedes medizinische Bild die notwendigen Daten der Patientinnen und Patienten und die Informationen bezüglich der

Diagnose. Die demographischen Daten, wie Geschlecht und Alter der anonymisierten Patientinnen und Patienten, Bildunterschrift, Untersuchungsmodalität, Anamnese, Diagnose, und Behandlung sind nur einige davon (MedPix® 2021).

1.5 ISO Normungen

Bevor zum zweiten Teil der vorliegenden Masterarbeit übergangen wird, ist es wichtig, den aktuellen Normungen im Bereich der Terminologie in diesem letzten Unterkapitel Raum zu geben. Die Internationale Organisation für Normung (ISO) ist die weltweit führende Organisation für die Festlegung technischer Standards bzw. Normen. Wie in allen Bereichen sind ebenso in der Terminologie einige Normen zu beachten. Laut ISO 10241-1 (ISO 1087:2019) müssen die Elemente eines Eintrags wie folgt erscheinen:

- entry number
- preferred term(s)
- admitted term(s)
- abbreviated form(s)
- definition
- example(s)
- note(s)

Laut ISO 704:2009 sind folgenden Tätigkeiten nachzugehen:

- Identifizierung der Begriffe und der Begriffsbeziehungen;
- Auf der Basis der identifizierten Begriffe und Begriffsbeziehungen, Begriffssysteme analysieren und modellieren;
- Erstellung von Begriffsdiagrammen, um die Begriffssysteme zu repräsentieren;
- Begriffe definieren;
- Den Begriffen Benennungen zuordnen;
- Daten aufnehmen.

2. Aktueller Stand der Forschung

Wie bereits erwähnt ist das Ziel dieser Arbeit, ein Modell für die Darstellung multimodaler Terminologien vorzuschlagen. Allerdings existieren derzeit solche Arbeiten noch nicht. Aus diesem Grund werden folglich in diesem Kapitel die Publikationen besprochen, die für das Thema der vorliegenden Masterarbeit relevant sind. Rodríguez et al. (2013) und Prieto Velasco (2012) stellen EcoLexicon vor, eine multimodale *Terminological Knowledge Base* über die Umwelt. Sánchez et al (2014) stellen VariMed vor, ein Projekt im Bereich der medizinischen Kommunikation. Die zuletzt vorgestellte Veröffentlichung von Drewer et al. (2017) hingegen befasst sich mit dem Thema der Wissensmodellierung und der Ontologie. Folglich werden die erwähnten Publikationen ausführlicher beschrieben.

Rodríguez et al. (2013) stellen in ihrer Publikation das Thema der Multimodalität am Beispiel von EcoLexicon unter besonderer Berücksichtigung der Rolle der Bilder bei der Darstellung der Fachsprache vor. EcoLexicon ist eine *Terminological Knowledge Base (TKB)*, die sich mit dem Thema Umwelt befasst. Es handelt sich um eine praktische Anwendung der Theorie der *Frame-based Terminology* (Faber 2015), in der die Wissensrepräsentation multimodal angereichert ist (Rodríguez et al. 2013). In der Praxis werden Begriffe in der Form einer Mindmap dargestellt. Diese Begriffe haben zwei Verwendungszwecke. Einerseits sind sie für Expertinnen und Experten des Umweltsektors nützlich, weil nicht nur der gesuchte Oberbegriff dargestellt wird, sondern auch die damit verbundenen Unterbegriffe. Andererseits sind auch Übersetzungsvorschläge für den Oberbegriff auf Deutsch, Englisch, Französisch, Griechisch, Niederländisch, Russisch und Spanisch vorhanden. Dadurch wird die TKB auch für Übersetzerinnen und Übersetzer und Dolmetscherinnen und Dolmetscher im Umweltbereich zugänglich gemacht.

Bezüglich der Berücksichtigung der Rolle der Bilder bei der Darstellung der Fachsprache ist die Auswahl eines bestimmten Bildes sehr wichtig, weil es einen großen Einfluss auf die künftigen Strategien in Bezug auf die Entwicklung der TKB ausüben wird (Rodríguez et al. 2013). Die meist relevanten konzeptuellen Beziehungen sollten an Relevanz gewinnen, indem die multimodalen Ressourcen für eine TKB ausgewählt werden (Rodríguez et al. 2013).

In EcoLexicon trägt jeder eingetragene Begriff relevante Informationen sowie sprachliche, konzeptuelle und graphische Informationen (Rodríguez et al. 2013). Sprachliche Informationen sind beispielsweise die Definitionen der Konzepte, ihre Synonyme oder äquivalente Begriffe in anderen Sprachen (Rodríguez et al. 2013). Unter konzeptuellen

Informationen verstehen Rodríguez et al. (2013) die konzeptuellen Relationen und die Struktur des Sachgebiets. Während man unter grafischen Informationen alle multimodalen Ressourcen versteht, wie beispielsweise Bilder, Videos oder Tonaufnahmen (Rodríguez et al. 2013).

Folglich werden die erwähnten Prinzipien bezüglich der Umsetzung der Bilder in EcoLexicon ausführlicher besprochen. Abb. 11 sollte dazu dienen, sich einen Überblick zu verschaffen, wie die Webseite von EcoLexicon aussieht und die folglich besprochenen Eigenschaften darzustellen.

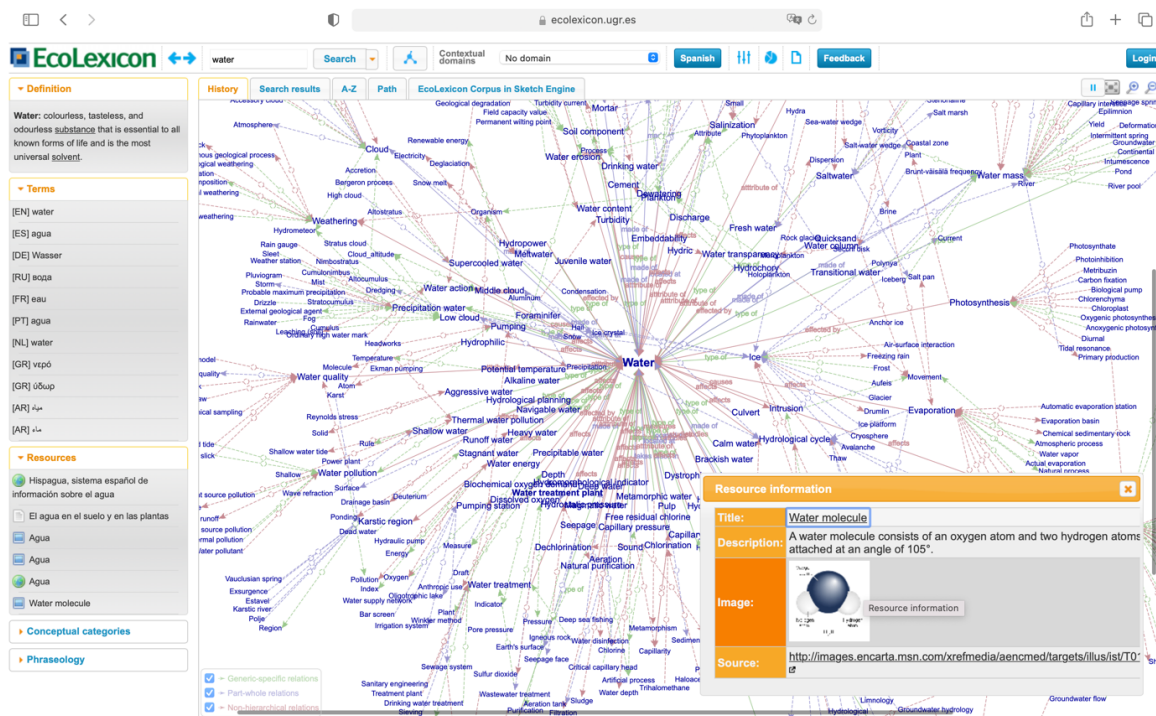


Abbildung 11. Begriff *Water* in EcoLexicon (EcoLexicon 2021)

Die Analyse der betroffenen TKB wird von der kleinsten Ebene ausgehen, das heißt vom einzelnen Begriff, in diesem Fall den Begriff *Water*, bis zu den Beziehungen zwischen den Begriffen und schließlich wird sich die Analyse auf die verwendeten multimodalen Ressourcen fokussieren.

Klickt man im Modell auf einen Begriff, so erscheint dessen Definition (EcoLexicon 2021). Von *Water* „is a colorless, tasteless, and odorless substance that is essential to all known forms of life and is the most universal solvent“ (EcoLexicon 2021) ist allerdings lediglich die *entity* präsent. Wasser ist nämlich nicht das Resultat eines Verfahrens (Rodríguez et al. 2013). Ein Begriff mit zusätzlichen Informationen ist beispielsweise *groyne* (Rodríguez et al. 2013). Der

englische Begriff *groyne* wird wie folgt definiert: “defense structure made up of wood, concrete, or rock, perpendicular to the coastline, which retards littoral drift and erosion. Its shape, height, and length are variable” (Rodríguez et al. 2013). Aus dieser Definition lassen sich zwei zusätzliche Arten von Beziehung erahnen. Als *defense structure* ist *groyne* das *result* eines *event*. Besser ausgedrückt ist *groyne* die *result-of* Beziehung eines Bauprozesses (Rodríguez et al. 2013). Darüber hinaus sind weitere Beziehungen erwähnenswert, sowie *located-at* oder *has-function* (Rodríguez et al. 2013).

Alle Begriffe werden durch unterschiedliche Beziehungen in Verbindung gesetzt. Die verschiedenen Farben: grün, blau und rot definieren die verschiedenen Beziehungen (EcoLexicon 2021). Unten links in Abb. 11 dient eine Legende zur richtigen Interpretation der betroffenen Relationen. *Generic-specific relations* steht für *type-of* Beziehungen bzw. *genericRelation*, *part-whole relations* steht für *part-of* Beziehungen bzw. *partiativeRelation* und *non-hierarchical relations* steht für weitere Beziehungen wie beispielsweise *effected-by*, *affects*, *has-function*, *attribute-of* usw. (EcoLexicon 2021).

Zuletzt sind weitere Informationen des Begriffes oben links aufgelistet. Diese Informationen sind die bereits erwähnte Definition, die Übersetzung des Begriffes in anderen Sprachen, konzeptuelle Kategorien mit der hierarchischen Zuordnung des Begriffes, die phraseologischen Einträge und die Ressourcen. Unter den Schirm dieser letzten Kategorie fallen ebenso die multimodalen Ressourcen bzw. die visuellen Repräsentationen des Begriffes (EcoLexicon 2021). Rodríguez et al. (2013) empfehlen, mehr als ein Bild pro Eintrag einzufügen, um den Begriff aus verschiedenen Perspektiven, in verschiedenen Zuständen und Umständen zu zeigen, um sicherzustellen, dass das Zielpublikum die richtige Bedeutung erfasst und ihm kein Interpretationsspielraum gewährt wird. Nicht nur die Oberbegriffe sollen über darstellende Ressourcen verfügen, sondern ebenfalls die damit verbundenen Unterbegriffe. In Abb. 11 links ist zu sehen, dass der Begriff *Water* über zwei Bilder verfügt namens *Agua* (Spanisch für Wasser) und ein Bild für *Water molecule*, das unten rechts aufgezeigt wird (EcoLexicon 2021).

Über 2.500 von 3.500 Einträgen in EcoLexicon verfügen über Bilder (Rodríguez et al. 2013). Da es aber keine perfekte Übereinstimmung zwischen konzeptuellen Kategorien und grammatikalischen Klassen gibt, ist es nicht leicht zu definieren, ob sich diese Bilder auf die *entities*, auf die *events* oder auf die *properties* beziehen (Rodríguez et al. 2013). Tendenziell beziehen sich die Substantiven auf die *entities*, die Verben auf die *events* und die Adjektiven auf die *properties*. Dies kann allerdings leider keine Regel sein, weil sie nicht in allen Fällen gültig ist. Trotz der Schwierigkeiten bezüglich einer präzisen Rechnung und Schätzung des

richtigen Ursprungs der Bilder, wird geschätzt, dass sich ungefähr 2.200 von 2.500 Bildern auf Substantive und sich lediglich circa 300 von 2.500 auf Verben und Adjektive beziehen (Rodríguez et al. 2013).

In dieser Publikation wurde das Konzept der Multimodalität in Bezug auf die Erstellung von Ontologien in EcoLexicon analysiert. Ontologiebasierte Wissensdatenbanken können mit Hilfe multimodaler Ressourcen angereichert werden, weil diese multimodalen Wissenrepräsentationen die Bedeutungsübertragung von Bildern auf Text und umgekehrt ermöglichen.

Eine weitere erwähnenswerte wissenschaftliche Arbeit ist jene von Prieto Velasco (2012), die sich mit den Themen Extraktion und Multimodalität aus einer noch konkreteren Perspektive beschäftigt. In dieser Publikation geht es ebenso um EcoLexicon, nur aus einer präziseren Sichtweise, und um Google Picasa als Extraktionstool von Bildern und Texten. Es werden der Extraktionsprozess der *Tags* bzw. der Beschreibungen, der Bilder und die folgende Verknüpfung von Bildern und Texten beschrieben.

Diese Forschung basiert auf einer neuen kognitiven Theorie, die sogenannte *Frame-based Terminology (FBT)*, die die Terminologie aus einer multimodalen Sichtweise betrachtet (Prieto Velasco 2012). FBT ist also ein ausgezeichneter Ansatz für die Extraktion von Ressourcen von Texten, die reich an visuellen Darstellungen sind (Prieto Velasco 2012).

Das Projekt besteht aus zwei Hauptteilen: Korpuszusammenstellung und -anmerkung. Die Korpuszusammenstellung besteht aus einer Auswahl von unterschiedlichen Texten: Texte für Laiinnen und Laien, semispezialisierte Texte und hochspezialisierte Texte für Expertinnen und Experten des Umweltbereiches. Die Zusammensetzung dieser Texte erfolgt nach einigen Eigenschaften, über die diese Texte verfügen sollten: Aus zuverlässigen Quellen extrahiert worden sein, im Web zur Verfügung stehen und auf Englisch verfasst worden sein. Aus diesen unterschiedlichen Textsorten entsteht der Corpus, der laut Prieto Velasco (2012) über folgende Eigenschaften verfügen soll: Synchron, schriftlich, maschinenlesbar und englischsprachig sein.

Der Teil der Korpusanmerkung besteht aus der Einfügung von Kommentaren im Korpus. Dies bedeutet, dass zusätzliche Informationen zu den existierenden Textteilen hinzugefügt werden. Implizite Informationen im Originaltext bzw. Korpus werden durch diese Anmerkungen explizit (Prieto Velasco 2012). Es existiert derzeit keine eindeutige Richtlinie weder zur Anmerkungsweise, noch zur Repräsentation des multimodalen Wissens (Prieto Velasco 2012). Prieto Velasco (2012) schlägt vor, die ursprüngliche Position von Bildern und Texten aufzufinden und den Kontext, in dem die Bilder eingebettet sind, zu analysieren.

Nachdem die richtige Bedeutung des Bildes erfasst wurde, sollte ein Tag gefunden werden, um das Bild kurz und bündig zu beschreiben (Prieto Velasco 2012).

Die konzeptuellen Tags werden manuell hinzugefügt, nachdem man das gesamte Konzept, in dem das Bild eingebettet ist, betrachtet hat und die Bedeutung des Bildes mit der tatsächlichen Definition des Begriffs verglichen hat. Erst danach werden schließlich diese Tags wie <WIND EROSION> anstelle des Bildes in den Text eingefügt (Prieto Velasco 2012).

Prieto Velasco (2012) ist, wie beschrieben, eine wertvolle Ressource in Bezug auf das *Tagging* der visuellen Ressourcen. Die Publikation bietet nämlich eine wirksame Methode an, um Bilder als Text zu betrachten und gemeinsam zu verwenden.

Sánchez et al. (2014) stellen in ihrer Publikation das Projekt VariMed vor. Es handelt sich um ein Forschungsprojekt im Bereich der medizinischen Kommunikation, das in eine multimodale und zweisprachige (Spanisch und Englisch) Ressource mit visuellen Repräsentationen und detaillierte Informationen zur Wissensverbreitung einfließt.

VariMed geht von der Annahme aus, dass schätzungsweise 80% der medizinischen Fehler einer unzureichenden Kommunikation zwischen den Patientinnen und Patienten und dem Personal im Gesundheitswesen zurückzuführen sind (Sánchez et al. 2014). Aus diesem Grund liegt der Fokus des Projektes auf der Untersuchung der Begriffe der medizinischen Sprache, insbesondere auf Krankheiten und Symptome (Sánchez et al. 2014). Aus diesem Projekt entsteht eine Online-Ressource, die als Informationsquelle für Recherchen im Bereich Kognitions- und Sprachwissenschaft oder für all diejenigen, die sich für die Medizin interessieren, dient (Sánchez et al. 2014).

Sánchez et al. (2014) unterscheiden zwischen verschiedenen Begriffsarten. Begriffe, die aus dem Latein und Begriffe die aus dem Altgriechischen stammen. Oder Begriffe wie *aerotitis* und *slapped face*, wobei der erste Begriff eine typische Ohrenentzündung für Fliegerinnen und Flieger bezeichnet, während der zweite den Fokus auf der Position der Erkrankung liegt (Sánchez et al. 2014). Durch die nicht-lexikalischen Darstellungen durch Bilder ist es möglich, die lexikalischen Definitionen zu ergänzen. Aus diesem Grund zielt VariMed darauf ab, eine Art und Weise zu finden, um die visuellen Darstellungen in eine medizinische Datenbank einzufügen (Sánchez et al. 2014).

Um diesen Austausch zwischen unterschiedlichen *Modi* zu ermöglichen, ist eine Beschreibung der Bilder erforderlich, die auf den Untersuchungen der Repräsentation visueller Darstellungen basiert. Die Bilder werden laut drei Kriterien ausgewählt: Repräsentativität, Spezialisierung und Ikonizität. Unter dem Prinzip der Repräsentativität versteht man die Art und Weise, wie die Realität wahrgenommen wird. Unter Spezialisierung versteht man den

Spezialisierungsgrad der Bilder, das heißt an welchem Zielpublikum sich die Ressourcen wenden können (Sánchez et al. 2014). Die Auswahl der Bilder in VariMed hat nach folgenden Kriterien erfolgt: Anmerkung wesentlicher Elemente des Begriffes und Hervorhebung visueller Eigenschaften und der Merkmale des Begriffes (Sánchez et al. 2014).

Drewer et al. (2017a) beschäftigen sich mit dem Thema der Wissensmodellierung mit Hilfe von Ontologien. Die Ontologien werden in den Bereichen der Informatik und der Künstlichen Intelligenz als maschinenlesbare Wissensmodellierungen verstanden, die Klassen enthalten, die miteinander in Verbindung stehen. Die Inhalte der Ontologien sind vom Einsatzzweck und vom Einsatzgebiet abhängig. Infolgedessen sind die Ontologien nicht für alle Nutzungszwecke sinnvoll. Ontologien sind nach Sachgebiet organisiert. Zwei oder mehrere Ontologien aus unterschiedlichen Gebieten können aber auch zusammengestellt werden, falls der Bedarf besteht. Das ursprüngliche Format für Ontologien ist XML-basierte Ressource Description Framework (RDF). Eine Weiterentwicklung dieses Formats ist Web Ontology Language (OWL), eine Spezifikation des World Wide Web Consortiums (W3C), dessen Ziel ist es, Termini und Beziehungen auf eine bestimmte Art und Weise zu beschreiben, sodass auch die Programme die Bedeutung erfassen können. Am relevantesten für den aktuellen Stand der Forschung ist Protégé, ein Open-Source-Tool zur praktischen Modellierung von Ontologien.

Als treffendes Beispiel für ein multilinguales Netzwerk und lexikalisierte Ontologie ist BabelNet nennenswert (BabelNet 2021). BabelNet entstand durch die automatische Integration von Wikipedia, die größte mehrsprachige Enzyklopädie, mit WordNet, dem bekanntesten Lexikon der englischen Sprache (BabelNet 2021). Die Termini, die auf anderen Sprachen nicht zur Verfügung stehen, werden mit Hilfe automatischer Übersetzungstechniken übersetzt. Das Ergebnis ist ein dynamisches, multimodales und multilinguales Wörterbuch, das Einträge in vielen Sprachen bietet, die durch semantische Beziehungen verknüpft sind. In Abb. 12 ist ein Beispiel für ein Relationssystem für Begriff Maus zu sehen. Oben links in der Abbildung werden alle verwendeten Beziehungen gezeigt (Drewer et al. 2017a).

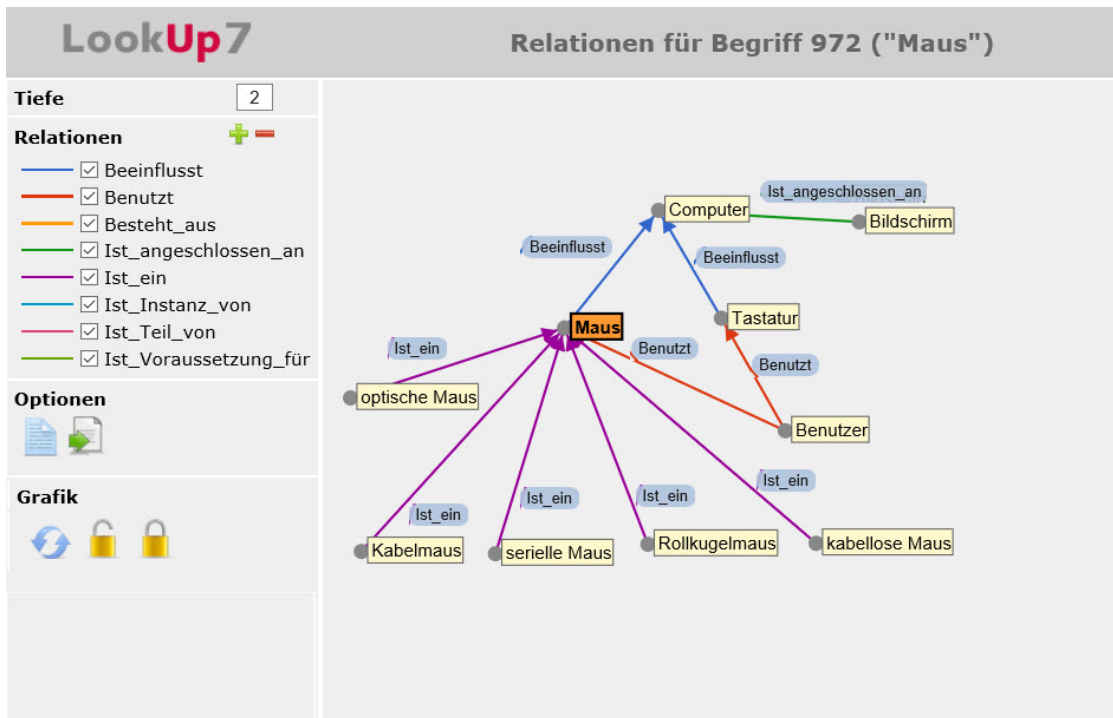


Abbildung 12. Beispiel für Relationssystem (Drewer et al. 2017a)

Diese Ressource ist nicht für ihre Menge an medizinischen Bildern relevant, sondern wegen ihrer Struktur, semiotischen Verbindungen und visuellen Darstellungen (BabelNet 2021).

Obwohl die Forschungsfelder der oben erwähnten wissenschaftlichen Veröffentlichungen von dem Forschungsfeld der vorliegenden Masterarbeit abweichen, leisten sie trotzdem einen bedeutsamen Beitrag zu der Gestaltung und der verwendbaren Methodik bei der Erstellung eines multimodalen Modellierungsansatzes für Terminologie.

3. Methode

Dieses Kapitel ist der Empirie gewidmet und zielt aus diesem Grund darauf ab, die angewendete Methode zu dem empirischen Teil vorzustellen mit besonderem Augenmerk auf die im ersten Kapitel erwähnte Forschungsfrage: „Welche Haupteigenschaften und zu betrachtende Kernaspekte sind bei der Modellierung einer multimodalen (Text-Bild-Beziehungen) Terminologie zu beachten, welche aus terminologischen Informationen aus Text und dazugehörigen Bildern zusammengestellt wird?“.

Wie im Unterkapitel 1.4.1 beschrieben wurde, wird MedPix® als Datensatz für das Endprojekt der vorliegenden Masterarbeit verwendet, weil er nicht nur eine breite Auswahl an medizinischen Bildern anbietet, sondern auch medizinische textuelle Informationen zu jedem Bild zur Verfügung stellt, welche notwendig für die richtige Interpretation der medizinischen Ressourcen sind. Folglich in diesem Kapitel werden die diversen Ansätze zur Extraktion der medizinischen Terminologie erläutert und ein potenzieller Leitfaden zur Erstellung eines multimodalen medizinischen Datensatzes vorgeschlagen.

3.1 Erster Ansatz zu MedPix®

Zunächst einmal ist es empfehlenswert, den ausgewählten Datensatz zu analysieren. In diesem Fall ist es notwendig, die Webseite von MedPix® (<https://medpix.nlm.nih.gov/home>) zu öffnen, oben rechts auf *Topics* und dann noch auf *Topic List* zu klicken und eine Kategorie zwischen den drei *ACR*, *Location* und *Category* auszuwählen. *ACR* ist die Abkürzung bzw. ein Code, die bzw. der für American College of Radiology steht. In dieser Kategorie kann man einen Körperteil bzw. einen Cluster von Körperteilen auswählen, wie beispielsweise lediglich *Breast* oder *Face, Sinuses & Neck*. Die Kategorie *Location* bezeichnet die Partie, wo die Erkrankung auftritt. Wobei die letzte *Category* die relevanteste Klassifizierung an medizinischen Ressourcen enthält und unterscheidet zwischen verschiedenen Kategorien, wie eine bestimmte Erkrankung klassifiziert wurde. Dies ist ein verflochtener Teil der Webseite, in dem sowohl Krankheiten als auch Medizinbranchen zu finden sind.

3.1.1 Die Ophthalmologie

Für die vorliegende Masterarbeit wurde die *Category Ophthalmology* von MedPix® (Siehe Abb. 13) ausgewählt. Hinter dieser Entscheidung sind verschiedene Gründe zu beachten. Die Ophthalmologie ist ein eingeschränktes Teilgebiet der Medizin, weil sie lediglich die Augen untersucht. Aus diesem Grund wird es in der vorliegenden Masterarbeit möglich, sich auf einen

hochspezifischen Teil der Medizin zu fokussieren. Darüber hinaus stellt die *Category* Ophthalmologie 15 Szenarien zur Verfügung, die eine ausgewogene Anzahl repräsentiert, um die konzipierte Methode vorzustellen, ohne die Leserinnen und Leser wegen des hohen Spezialisierungsgrad mit einer hohen Anzahl an Szenarien und Fällen zu verwirren. Allerdings ist es trotzdem möglich, Szenarien unterschiedlicher Art vorzustellen. Wegen des hohen Spezialisierungsgrads weicht das Endprojekt der vorliegenden Masterarbeit in diesem Sinne von VariMed ab. VariMed ist das einzige derzeit existierende Projekt bezüglich medizinischer Terminologie und Multimodalität und nährt sich am meisten an dem vorgestellten Projekt. Allerdings zielt VariMed darauf ab, die kommunikativen Missverständnisse zwischen dem Gesundheitspersonal und den Patientinnen und Patienten zu überwinden. Aus diesem Grund ist der Spezialisierungsgrad absichtlich relativ niedrig gehalten, sodass beide Parteien der kommunikativen Situation den Sinn erfassen können.

Die Ophthalmologie ist aus diesen Gründen ein geeigneter Medizinbereich, um als Beispiel für die Modellierung multimodaler und medizinischer Ressourcen zu fungieren.

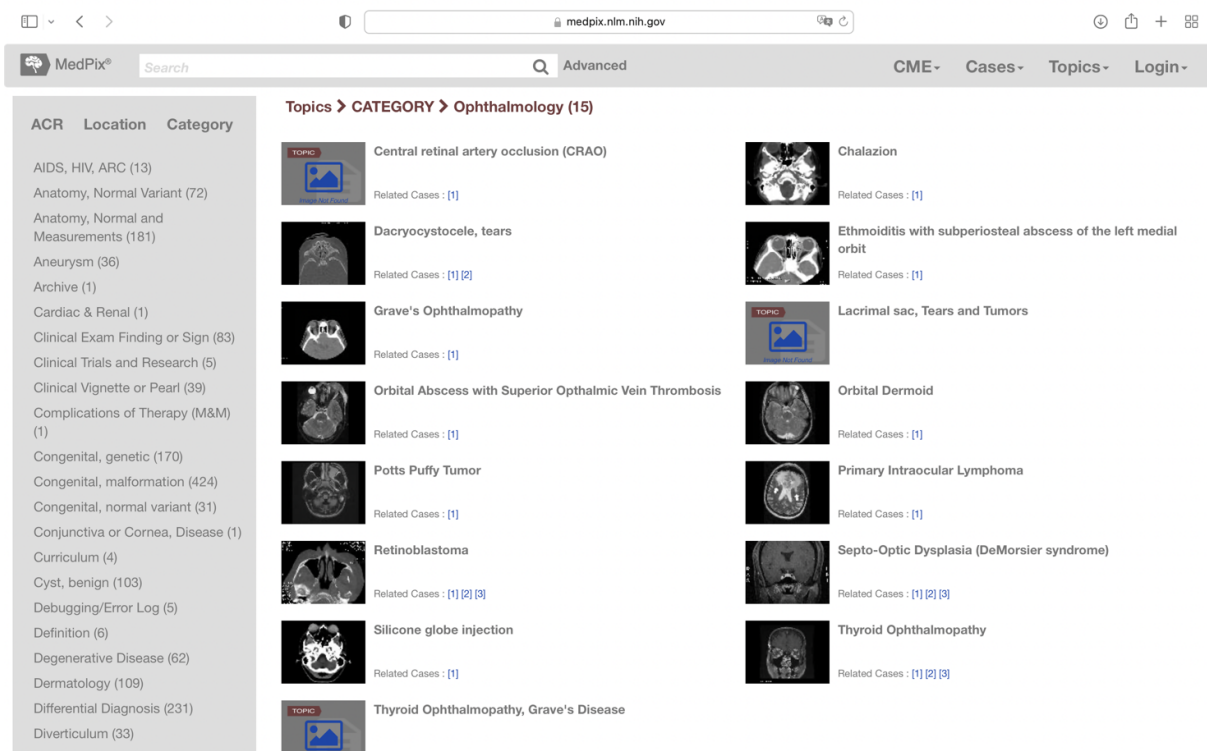


Abbildung 13. Seite Ophthalmology auf MedPix® (MedPix® 2021)

Die in Abb. 13 dargestellten Szenarien wurden in Schreibrichtung von oben links nach rechts nummeriert. Das erste Szenario, *Central retina artery occlusion (CRAO)*, wird als Szenario Nummer 1 verstanden, *Chalazion* als Nummer 2, *Dacryocystocele, tears* als Nummer 3 usw.

Diese einfache Kategorisierung wurde vorgenommen, um zu vermeiden, dass der medizinische Name des Szenarios jedes Mal wiederholt werden muss.

Zu diesem Zeitpunkt ist es möglich, entweder auf das Bild bzw. auf den Titel des Szenarios zu klicken, um die Seite des Szenarios zu öffnen, oder auf die kleinen blaufarbigen Nummern unten rechts, um den spezifischen Fall eines bestimmten Szenarios zu öffnen, falls mehrere Fälle pro Szenario vorhanden sind.

3.2 Analyse der 15 Szenarien

Als Vorbild für das Szenario des Datensatzes wird das Szenario 13 genommen, da es lediglich einen Fall und zwei Bilder enthält und knapp und bündig alle wichtigen Informationen zur Verfügung stellt. Fast alle Szenarien enthalten drei Teile: *Author*, *Topic* und *Case*. Relevant für die vorliegende Masterarbeit sind allerdings nur die letzten zwei, weil die Informationen über die Autorinnen und die Autoren keinen Mehrwert für die endgültige Arbeit darstellen, da sie weder das Szenario noch den Fall betreffen. In Abb. 14 ist *Topic* zu sehen, während *Case* in Abb. 15 angezeigt wird.

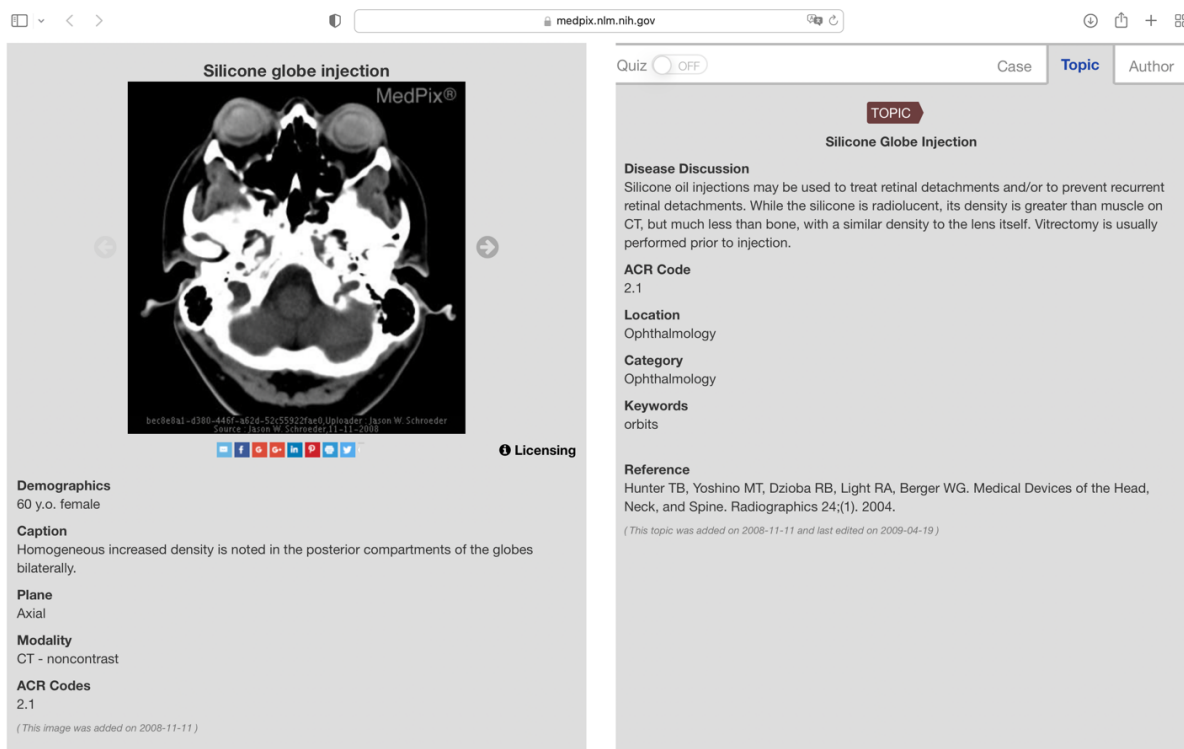


Abbildung 14. Szenario 13, Silicone globe injection, *Topic* (MedPix® 2021)

Innerhalb vom Teil *Topic* sind folgende weitere Aufteilungen zu finden: *Disease Discussion* (Beschreibung der Krankheit), *ACR Code* (American College of Radiology), *Location*

(Bereich), *Category* (Kategorie), *Keywords* (Schlüsselwörter) und *Reference* (Externe Links). In *Disease Discussion* wird eine Beschreibung der Krankheit angeboten, zusammen mit medizinischer Terminologie in Bezug auf die betrachtete Erkrankung und eventuelle Diagnosen. *ACR Code* ist die Abkürzung für American College of Radiology. Dieser Code dient dazu, die Daten in der Abteilung für Radiologie computerisieren zu können. Er wurde in der Foxbase Sprache geschrieben und wird seit 1992 in der radiologischen Abteilung des Wallace Memorial Baptiste Hospital benutzt (IAEA 2021). Dieser Code wird für die vorliegende Masterarbeit nicht benötigt, weil eine neue Methodologie zur Verwaltung der Ressourcen entwickelt wird. Dies ist ein Beispiel der Anwendung des *principle of coherence* laut Rodríguez et al. (2013), das der Entzug irrelevanter Teile vom Ausgangstext vorsieht. *Location* und *Category* zeigen die Verortung an. In diesem Fall ist es für alle 15 Szenarien *Ophthalmology*, weil lediglich diese Branche in der vorliegenden Masterarbeit untersucht wird. Die *Keywords* leisten einen wichtigen Beitrag, um die wesentlichen Konzepte der einzelnen Szenarien zu identifizieren. Dank der zur Verfügung gestellten Schlüsselwörter ist es leichter, sich einen Überblick über die behandelten Themenbereiche verschaffen zu können. Die *References* – in anderen Szenarien externe Links genannt – sind nützlich für eventuelle Vertiefungen des Themas.

Auf der linken Seite der Webseite bzw. unter dem Bild sind weitere Informationen zu finden wie *Demographics* (Demographische Daten), *Caption* (Bildunterschrift), *Plane* (Ebene), *Modality* (Modalität) und *ACR Codes*. Diese Informationen sind mit dem jeweils einzelnen Fall verbunden.

Silicone globe injection

MedPix®

VIEW CASE IMAGES

CASE

Silicone Globe Injection

History
60 year old woman with headaches.

Findings
There is homogenous increased attenuation of the posterior chamber of both globes. Incidentally, there were no abnormal intracranial findings.

Differential Diagnosis

- Silicone ocular injection
- Bilateral vitreal hemorrhages

Case Diagnosis
Silicone globe injection

Diagnosis By
History

Discussion
This patient received vitrectomy and silicone injections secondary to hemorrhagic retinitis and detachments.

Demographics
60 y.o. female

Caption
Coronal reconstruction again reveals homogenous increased density in the posterior compartments.

Plane
Coronal

Modality
CT - noncontrast

ACR Codes
2.1

(This image was added on 2008-11-11)

Abbildung 15. Szenario 13, Silicone globe injection, Case (MedPix® 2021)

Die demographischen Daten beinhalten das Geschlecht und das Alter der Patientinnen und Patienten. Die *Caption* (Bildunterschrift) ändert sich selbstverständlich je nach Fall (nicht nach Szenario), es sei denn es werden mehrere Bilder zur Verfügung gestellt, die dieselbe Krankheit bei derselben Patientin bzw. bei demselben Patienten darstellen. In mehreren Szenarien werden verschiedene Bilder aus unterschiedlichen Ebenen gezeigt. In diesem Fall ändert sich die Bildunterschrift nicht. Die unterschiedlichen Ebenen sind nötig, um die Erkrankung aus unterschiedlichen Perspektiven zu zeigen. Abb. 16 zeigt die drei verwendeten Arten von Ebene.

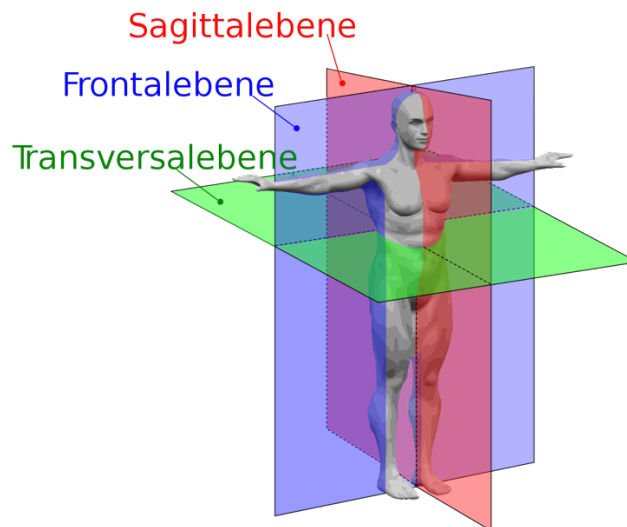


Abbildung 16. Drei Ebenen in der Anatomie (Guzmán et al. 2016: 3)

Wie in Abb. 18 zu sehen ist, spricht man in der Anatomie von drei Ebenen: Sagittal-, Transversal- und Frontalebene. Unter Sagittalebene versteht die Anatomie eine Ebene, die sich von oben nach unten und von vorne nach hinten erstreckt. Unter Transversalebene wird eine Ebene bezeichnet, die im rechten Winkel zur Längsachse quer liegt. Wobei unter Frontalebene eine Ebene gemeint wird, die sich von unten nach oben und von links nach rechts erstreckt. MedPix® verwendet ebenso drei weitere Ebenen, die normalerweise nicht als Ebenen der Anatomie betrachtet werden: *Gross Pathology*, *Histology* und *NOS*. Unter *Gross Pathology* versteht MedPix® ein reines Foto der Pathologie, als wäre sie mit bloßem Auge zu sehen. Unter *Histology* (Histologie) wird ein Foto verstanden, das unter dem Mikroskop aufgenommen wird. Wobei die Abkürzung *NOS* für *Not specified* (nicht angegeben) steht.

Zuletzt ist die Modalität, wie ein Bild aufgenommen worden ist. Die elf verwendeten Methoden werden folglich aufgelistet und mit Hilfe eines Bildes des Datensatzes dargestellt. Bei den ersten zwei Bildern handelt es sich um Computertomographie (engl. CT). Im ersten Fall wird ein Kontrastmittel verwendet, im zweiten nicht. Leider handelt es sich nicht um denselben Fall, was interessant gewesen wäre, um den Unterschied zwischen den zwei Methoden hervorzuheben.

- CT w/contrast (IV) (Siehe Abb. 17).



Abbildung 17. CT w/contrast (IV) (MedPix® 2021)

- CT – noncontrast (Siehe Abb. 18).



Abbildung 18. CT – noncontrast (MedPix® 2021)

In den nächsten Bildern handelt es sich nicht mehr um Computertomographie, sondern um Magnetresonanztomographie (engl. MR). Darüber hinaus wurde das nächste Bild ohne Kontrastmittel aufgenommen.

- MR – T1W – noncontrast (Siehe Abb. 19).



Abbildung 19. MR – T1W – noncontrast (MedPix® 2021)

Die im nächsten Bild verwendete Technik (weighted) funktioniert sehr gut bei Patientinnen und Patienten, die gerade eine Strahlentherapie begonnen haben (Bjørnbo et al. 2020).

- MR - T2 weighted (Siehe Abb. 22).

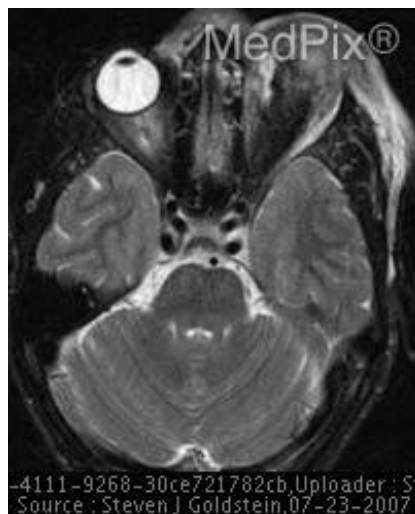


Abbildung 20. MR - T2 weighted (MedPix® 2021)

In den nächsten zwei Bildunterschriften steht die Abkürzung FLAIR für FLuid Attenuated Inversion Recovery.

- MR – FLAIR (Siehe Abb. 21).

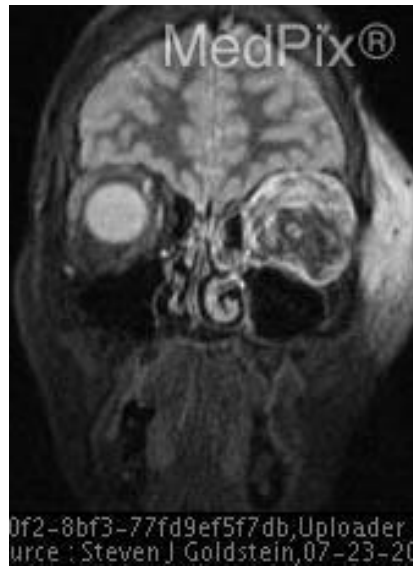


Abbildung 21. MR - FLAIR (MedPix® 2021)

- MR - T2 FLAIR w/Contrast (Siehe Abb. 22).

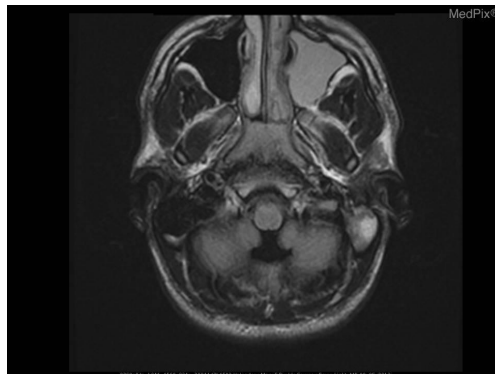


Abbildung 22. MR - T2 FLAIR w/Contrast (MedPix® 2021)

Die Abkürzung Gd in der nächsten Bildunterschrift steht für die chemische Abkürzung für Gadolinium (ein Kontrastmittel). Darüber hinaus wird im nächsten Bild das Fett entfernt (*fat suppressed* in Klammern). Dank dieser Technik soll der Tumor besser gesehen werden.

- MR - T1W w/Gd (fat suppressed) (Siehe Abb. 23).



Abbildung 23. MR - T1W w/Gd (fat suppressed) (MedPix® 2021)

Nun wird ein *gross photograph* angezeigt.

- GR - Gross photograph (Siehe Abb. 24).



Abbildung 24. Gross photograph (MedPix® 2021)

Die nächsten zwei Bilder sind, wie oben erwähnt, histologische Fotos mit unterschiedlichen Kräften des Mikroskops.

- HE - Med Power (~50-200x) (Siehe Abb. 25).

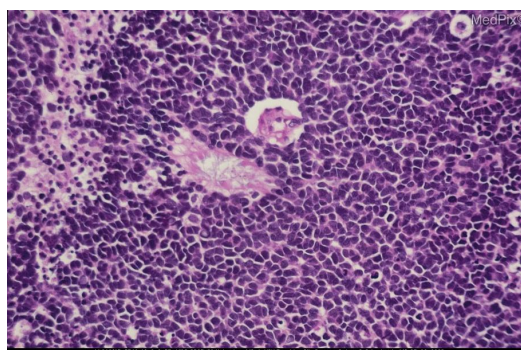


Abbildung 25. HE - Med Power (~50-200x) (MedPix® 2021)

- HE - High Power (>200x) (Siehe Abb. 26)

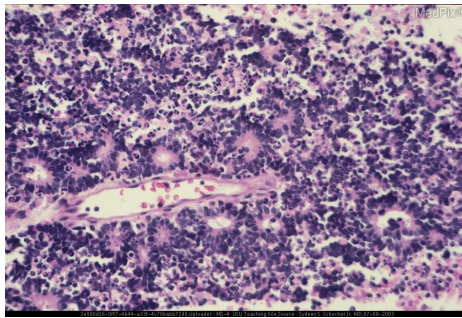


Abbildung 26. HE - High Power (>200x) (MedPix® 2021)

Im nächsten und letzten Bild steht die Abkürzung STIR für Short Inversion Recovery.

- MR – STIR (Siehe Abb. 27)

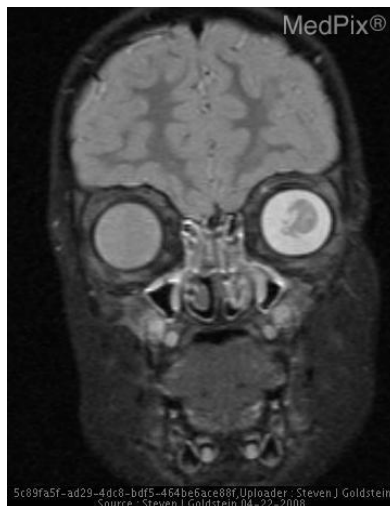


Abbildung 27. MR – STIR (MedPix® 2021)

3.3 Analyse der Schlüsselwörter

Infolge einer ersten Analyse des Datensatzes erschien die detaillierte Analyse der Schlüsselwörter als logischer nächster Schritt. Diese stellen nämlich die relevantesten Begriffe jedes Szenarios und jedes Falls dar. Darüber hinaus werden die ersten Beziehungen sichtbar. Folglich wird vom Szenario 1 begonnen und die Begriffe werden alphabetisch geordnet angezeigt, um ein Beispiel für die systematische Vorgehensweise zur Analyse der medizinischen Informationen zu verdeutlichen. Sobald sich die Notwendigkeit ergibt, einen Kommentar anzugeben, wird die Liste unterbrochen. Erst nach dem Kommentar wird die Auflistung der Begriffe der nächsten Szenarien fortgesetzt.

- Amaurosis fugax;
- Chalazion;
- Dacrocystocele;

- Hordeolum;
- Nasolacrimal duct cyst;
- Periorbital cellulitis;
- Timo cyst.

Sogar ohne über breites medizinisches Fachwissen zu verfügen ist es bereits möglich zu bemerken, dass der Begriff *cyst* zweimal vorkommt, obwohl es sich um zwei unterschiedliche Arten von Zysten handelt. Dies ist bereits ein Entwurf einer rudimentären Beziehung, wobei *cyst* der Oberbegriff ist und *nasolacrimal duct cyst* und *timo cyst* die Unterbegriffe sind bzw. eine type-of-Beziehung oder genericRelation (Nuopponen 2014). Ebenso das Schlüsselwort *dacrocystocele* ist eine Mehrwortbenennung, die *cyst* enthält. Es ist allerdings in diesem Fall notwendig, die Bedeutung des Begriffes nachzuprüfen, bevor eventuelle Relationen aufgebaut werden. Nun wird die Auflistung fortgesetzt, beginnend mit Szenario 4.

- Epiphora tearing;
- Ethmoiditis;
- Ethmoiditis with subperiosteal abscess of the left medial orbit subperiosteal abscess;
- Grave;
- Orbital cellulitis;
- Orbitopathy;
- Ophthalmopathy;
- Superior ophtalmic vein;
- Thrombosis.

In der eben vorgestellten Liste ergeben sich weitere interessante Beziehungen. Nicht nur sind Ober- und Unterbegriffsbeziehungen zwischen Benennungen der Pathologien zu sehen, wie bei *ethmoiditis* und *ethmoiditis with subperiosteal abscess of the left medial orbit subperiosteal abscess*, wo der zweite Begriff lediglich über eine weitere Spezifikation verfügt, sondern ebenfalls Körperteile wie *superior ophtalmic vein* werden ersichtlich, die eine located-at bzw. eine spatialRelation erfordern wird.

- Extraconal;
- Orbital Dermoid;
- Lacrimal gland lesion.

Die Beziehung ist nun nicht mehr so direkt und offensichtlich wie die bisher betrachteten. Allerdings ist es möglich, eine potenzielle Beziehung zwischen *lacrimal gland lesion* und *nasolacrymal duct cyst* in Betracht zu ziehen.

- CNS Lymphoma;
- MRI;
- Orbit;
- Potts Puffy Tumor;
- Proptosis;
- Sinusitis.

Zum ersten Mal sind in dieser Auflistung Abkürzungen aufgetaucht. Es ist aus diesem Grund notwendig zu verstehen, welche Bedeutung sie tragen. CNS steht für Central Nervous System (Zentralnervensystem), während MRI für Magnetic Resonance Imaging (Magnetresonanztomographie) steht.

- Intraocular;
- Retinoblastoma.

Hier ist ein Adjektiv zu sehen, das genau wie die anderen Schlüsselwörter betrachtet werden wird. Aus diesem Grund wird *intraocular* mit allen Begriffen verbunden werden, die intraokular sind oder sein können.

- Adipocytes;
- Autoimmune Thyrotoxicosis;
- Dyslapsia;
- Graves Disease;
- Ophthalmology;
- Orbits;
- Thyrotoxicosis;
- Thyrotoxicosis proptosis.

In dieser Gruppierung kommen sowohl die Begriffe *thyrotoxicosis* als auch *thyrotoxicosis proptosis* und *autoimmune thyrotoxicosis* vor. In diesem Fall wird *thyrotoxicosis* der Oberbegriff und die anderen zwei die Unterbegriffe sein. Mit dieser Liste wird die erste bündige Analyse zu den Schlüsselwörtern zu Ende. Folglich wird die vollständige Liste der betrachteten Terminologie alphabetisch geordnet:

- Adipocytes;
- Amaurosis fugax;
- Autoimmune Thyrotoxicosis;
- Chalazion;
- CNS Lymphoma;

- Dacrocystocele;
- Dyslapsia;
- Epiphora tearing;
- Ethmoiditis;
- Ethmoiditis with subperiosteal abscess of the left medial orbit subperiosteal abscess;
- Extraconal;
- Grave;
- Graves Disease;
- Hordeolum;
- Intraocular;
- Lacrimal gland lesion;
- MRI;
- Nasolacrimal duct cyst;
- Ophthalmology;
- Ophthalmopathy;
- Orbit;
- Orbits;
- Orbital cellulitis;
- Orbital Dermoid;
- Orbitopathy;
- Periorbital cellulitis;
- Potts Puffy Tumor;
- Proptosis;
- Retinoblastoma;
- Sinusitis;
- Superior opthalmic vein;
- Thrombosis;
- Thyrotoxicosis;
- Thyrotoxicosis proptosis;
- Timo cyst.

Nach der Erstellung dieses umfangreichen Begriffsverzeichnisses ergeben sich weitere interessante Aspekte bzw. neue Beziehungen. Das Adjektiv *periorbital* trägt teilweise die Bedeutung von *orbital*, wobei *orbits* die Pluralform von *orbit* ist. In diesem Fall wird nur die

Singularform betrachtet. Die Schlüsselwörter *orbit*, das zweimal vorgekommen ist, und *ophthalmology*, sind die einzigen zwei Wiederholungen der gesamten Datenbank. Insgesamt beträgt die Anzahl der zu betrachtenden Schlüsselwörter 34, nachdem die erwähnten Wiederholungen entfernt worden sind.

Im nächsten Kapitel werden die ersten Begriffsbeziehungen mit Hilfe von VariMed aufgebaut. VariMed, als einziger multimodaler medizinischer Datensatz, leistet einen wertvollen Beitrag zum ersten Verknüpfungsversuch der Terminologie.

3.4 Erste Begriffsbeziehungen mit Hilfe von VariMed

Nachdem die Liste der zu verknüpfenden Begriffe erstellt wurde, kann nun auf VariMed zurückgegriffen werden, indem die extrahierten Schlüsselwörter auf VariMed gesucht werden, um die von VariMed gebotenen Relationen zu verwenden. Wie bereits im aktuellen Forschungsstand erläutert wurde, ist VariMed eine wertvolle Ressource, obwohl die Originalsprache des Projektes Spanisch ist. Folglich wird gezeigt, wie die vorher von MedPix® extrahierten Begriffe zuerst in VariMed gesucht werden.

Zunächst wird empfohlen, VariMed auf Englisch zu benutzen, um die Ergebnisse bereits auf Englisch zu erhalten, ohne sie aus dem Spanischen übersetzen zu müssen. Die Sprache kann oben rechts umgestellt werden. Erst nach der Sprachumstellung können die Begriffe in der VariMed-Suchleiste auf Englisch gesucht werden (Siehe Abb. 28).

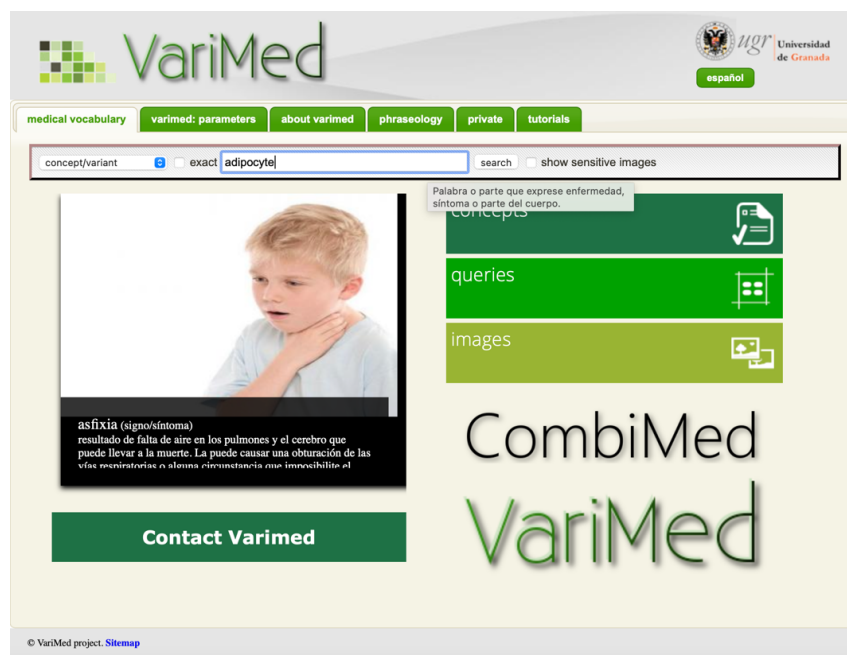


Abbildung 28. VariMed-Suchleiste (VariMed 2021)

Daraufhin wird die folgende Seite angezeigt, wo das Ergebnis auf beiden Sprachen zur Verfügung gestellt wird (Siehe Abb. 29).

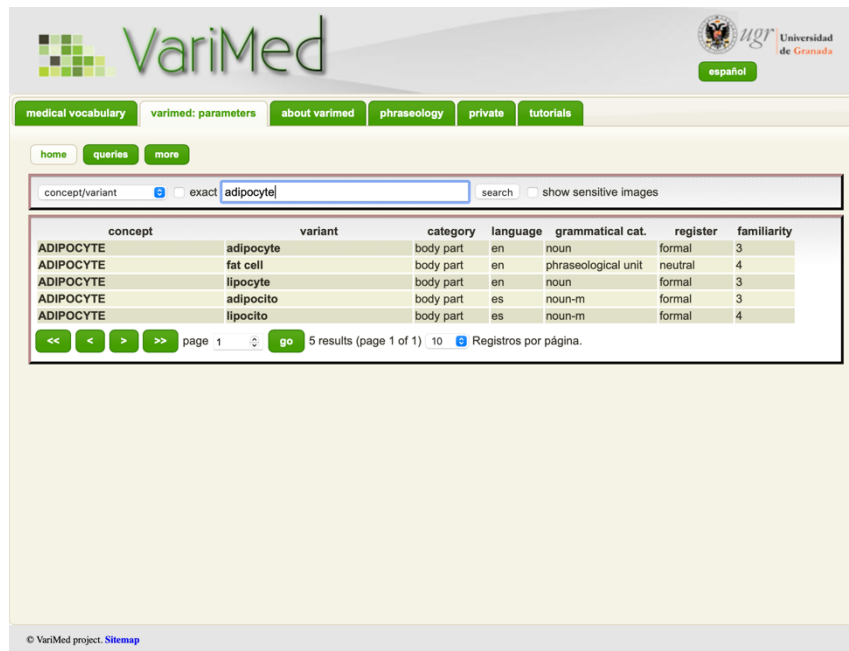


Abbildung 29. Begriff auf Spanisch und Englisch (VariMed 2021)

Klickt man nun auf *adipocyte* in der Spalte *concept*, dann öffnet sich folgende Seite (Abb. 30), wo die Beziehungen des gesuchten Begriffs angezeigt werden.

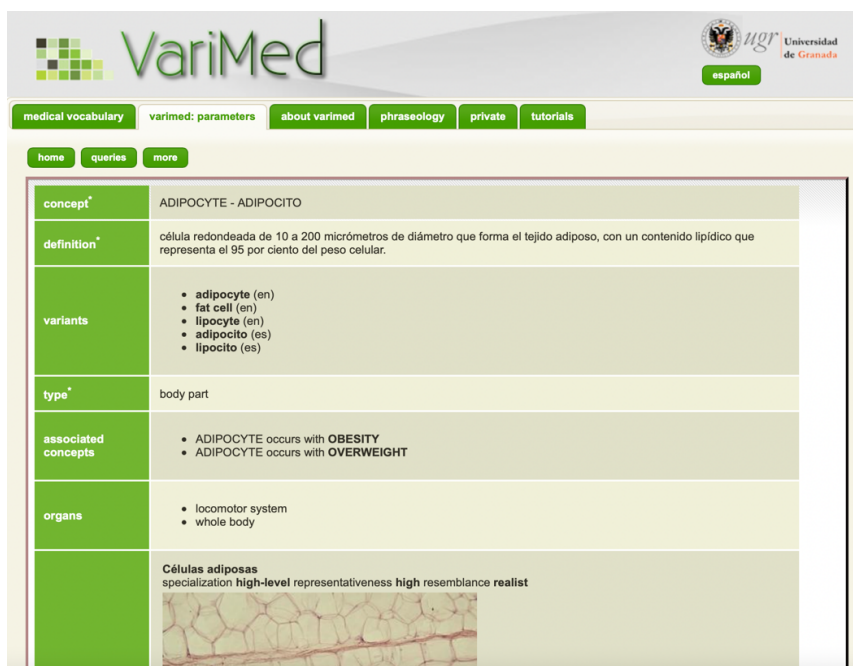


Abbildung 30. Ergebnisse für einen Begriff in VariMed (VariMed 2021)

Nun ist es notwendig die gefundenen Beziehungen niederzuschreiben. In diesem Fall sind zwei Beziehungen zu sehen: *adipocyte occurs with obesity* und *adipocyte occurs with overweight*. Zu diesem Zeitpunkt ist es notwendig, den Vorgang für alle ausgewählten Begriffe zu wiederholen.

Folgende Tabelle (Siehe Abb. 31) zeigt die Begriffe, für die Beziehungen in VariMed gefunden worden sind. In der folgenden Tabelle sind auch die spanischen Begriffe zu sehen, weil sie auch in Betracht gezogen worden sind, um die Relationen besser verstehen zu können. Diese Aufgabe verdeutlicht, dass es Begriffe gibt, die auf den ersten Blick für nicht Expertinnen und nicht Experten im Medizinbereich, keine gemeinsamen Merkmale teilen würden, eigentlich Synonyme sind. Diesbezüglich ist das Beispiel *thyrotoxicosis*, *Grave* und *Graves disease* erwähnenswert. All diese Benennungen bezeichnen nämlich den Begriff Hyperthyreose. Besser ausgedrückt: Es handelt sich um Synonyme. Diese Information wurde dank der spanischen Übersetzung der Begriffe gewonnen.

Der Vollständigkeit halber sind auch die Begriffe aufgeführt, für die keine Übersetzung gefunden wurde. Was die teilweise blauen Begriffe betrifft, wurde für den blauen Teil keine Äquivalenz im Spanischen gefunden (Siehe Abb. 31). Beispielsweise wurde für den von MedPix® extrahiertem Begriff *autoimmune thyrotoxicosis* nur eine spanische Äquivalenz für *thyrotoxicosis* und nicht für den gesamten Begriff *autoimmune thyrotoxicosis* gefunden. Aus diesem Grund wurde *autoimmune* in blau gefärbt. In diesen Fällen wurde jedoch beschlossen, Beziehungen für den Hauptbegriff zu suchen. *Hipertiroidismo*, die spanische Übersetzung für *thyrotoxicosis*, wurde nämlich nicht entfernt. In anderen Fällen wurde ebenso versucht, nur ein Wort des gesamten auf MedPix® gefundenen Begriffes zu suchen. Dies ist der Fall für *nasolacrimal duct cyst*, wo lediglich *cyst* gesucht wurde. Die Ergebnisse verweisen allerdings auf andere Arten von Zysten (Siehe Abb. 32) und können daher für diesen Begriff nicht verwendet werden, weil sie zu allgemein im Vergleich zum gesuchten Begriff sind.

EN	ES	Kommentare
Adipocytes	Adipocito	
Potts Puffy Tumor	Cáncer	
Amaurosis fugax	Ceguera fugaz	
Chalazion	Chalación	
Epiphora tearing	Epífora	
Autoimmune Thyrotoxicosis	Hipertiroidismo	
Grave	Hipertiroidismo	
Graves Disease	Hipertiroidismo	
Thyrotoxicosis	Hipertiroidismo	
Thyrotoxicosis proptosis	Hipertiroidismo	
Intraocular	Intraocular	Häufig mit dem Begriff „Hypertension“ verbunden.
CNS Lymphoma	Linfoma	
Hordeolum	Orzuelo	
Sinusitis	Sinusitis	
Thrombosis	Trombosis	
Superior ophthalmic vein	Vena	
Dacrocystocele		
Dyslapsia		
Ethmoiditis		
Ethmoiditis with subperiosteal abscess of the left medial orbit subperiosteal abscess		
Extraconal		
Lacrimal gland lesion		
MRI		
Nasolacrimal duct cyst		
Ophthalmology		
Ophthalmopathy		
Orbit		
Orbits		
Orbital cellulitis		
Orbital Dermoid		
Orbitopathy		
Periorbital cellulitis		
Proptosis		
Retinoblastoma		
Timo cyst		

Abbildung 31. Begriffe mit spanischer Äquivalenz in VariMed

concepto/variante exacto Palabra o parte que exprese enfermedad, sintc buscar mostrar imágenes sensibles

concepto	variante	categoria	idioma	cat. gramatical	registro	familiaridad
ABSCESO LAGRIMAL	dacryocystitis	enfermedad	en	sustantivo	formal	
CISTINOSIS	cystinosis	enfermedad	en	sustantivo	neutro	
COLECISTITIS	cholecystitis	enfermedad	en	sustantivo	formal	
COLECISTOLITIASIS	cholecystolithiasis	enfermedad	en	sustantivo	formal	
FIBROSIS QUISTICA	cystic fibrosis	enfermedad	en	unidad fraseológica	neutro	2
QUISTE	cyst	enfermedad	en	sustantivo	neutro	
QUISTE OVÁRICO	ovarian cyst	enfermedad	en	unidad fraseológica	neutro	
SÍNDROME DE OVARIO POLIQUISTICO	polycystic ovarian syndrome	enfermedad	en	unidad fraseológica	formal	
SÍNDROME DE OVARIO POLIQUISTICO	polycystic ovaries	enfermedad	en	sustantivo plural	neutro	
SÍNDROME DE OVARIO POLIQUISTICO	polycystic ovary disease	enfermedad	en	unidad fraseológica	neutro	

«<< < > >>» página 1 de 10 11 resultados (página 1 de 2) 10 Registros por página.

© Proyecto VariMed. Mapa del sitio

Abbildung 32. Ergebnisse für cyst in VariMed (VariMed 2021)

Nun wird vermerkt, wie VariMed vorschlägt, zu welcher Kategorie jeder Begriff gehört. VariMed unterteilt seine Begriffe in drei Kategorien: Erkrankung, Symptom und Körperteil.

Sobald alle Begriffsbeziehungen niedergeschrieben worden sind, ist es sinnvoll, ebenso alle Synonyme zu notieren, da diese als Benennungen desselben Begriffes eingetragen werden werden. Der Vollständigkeit halber werden für Begriffe wie *amaurosis fugax*, die eine Art von *amaurosis* ist, auch die Beziehungen von *amaurosis* dargestellt. Selbstverständlich wäre es möglich, unendlich viele Beziehungen zu finden, aber diese Arbeit soll nur eine Methode vorzeigen, wie multimodale medizinische Ressourcen am besten modelliert werden können.

3.5 Arten von Beziehungen für das Endprojekt

Wie bereits erläutert wurde ist es möglich, zwischen zwei Hauptkategorien von Relationen zu unterscheiden: hierarchischen und nicht-hierarchischen Beziehungen. Unter den hierarchischen Relationen sind Abstraktionsbeziehungen (type-of in VariMed bzw. genericRelation) und Bestandsbeziehungen (part-of bzw. partiativeRelation) zu zählen. Für das terminologische Begriffssystem der vorliegenden Masterarbeit wird das sogenannte Camel Case Mapping benutzt. Camel Case sieht vor, dass die Beziehungen anstatt mit Langstrich oder Unterstrich, mit großem Anfangsbuchstaben vor jedem Wort geschrieben werden (Microsoft 2021). Deswegen werden alle Wörter der Beziehung beigefügt. Dies ist das bevorzugte Mapping, weil

es auf diese Art und Weise für ein System leichter ist, diesen Schreibmodus zu erkennen und ihn in eine Beziehung umzuwandeln.

Die Abstraktionsbeziehung wird in der vorliegenden Masterarbeit zu einer *genericRelation*, wobei die Bestandsbeziehung *part-of* zu einer *partitiveRelation* wird (Nuopponen 2014). Die nicht-hierarchischen Relationen, die für diesen spezifischen medizinischen Fall notwendig sind, sind die folgenden: *located at*, *can cause*, *accompanied by*, *occurs at*, *diagnosed by* und *presents the symptoms*. *Located at* bezeichnet die Partie, in der sich ein Begriff befindet. Beispielsweise ist *lid sebaceous gland* *located at* *lid*. Diese Beziehung wird zu einer *spatialRelation* (Nuopponen 2014). Die bis jetzt verwendete *occurs at* Relation wird auch durch eine *spatialRelation* ersetzt. Der bis jetzt existierende Unterschied war, dass *located at* die Zuordnung eines Körperteils war, während *occurs at* immer mit einer Pathologie verbunden war. Allerdings ändert sich eine Partie, wo sich ein Körperteil befindet oder wo eine Erkrankung auftritt, nicht. Aus diesem Grund wurde die Entscheidung getroffen, bei beiden Fällen die *spatialRelation* zu benutzen. *Can cause* wird zuerst in *caused by* umgewandelt, um die Richtung umzukehren und wird danach zu einer *casualRelation*. Ausgehend von Beziehungen dargestellt in Nuopponen (2014), kann *accompanied by* als eine terminologische *correlationRelation* modelliert werden. Zuletzt sind zwei weitere Relationsarten: *diagnosed by* und *presents the symptoms*. Es handelt sich um sehr spezifische Beziehungen bezüglich des Medizinbereiches. Deshalb bietet Nuopponen (2014) keine passende Lösung an und diese Relationen mussten für diesen spezifischen Kontext neu definiert werden. Es wurde beschlossen, *diagnosed by* in eine *diagnosisRelation* und *presents the symptoms* in eine *symptomRelation* umzuwandeln. Folglich werden diese Beziehungen stichprobenartig aufgelistet:

- *type -of-relation* -> *genericRelation*
- *part-of-relation* -> *partitiveRelation*
- *located-at-relation* -> *spatialRelation*
- *occurs-at-relation* -> *spatialRelation*
- *caused-by-relation* -> *casualRelation*
- *accompanied-by-relation* -> *correlationRelation*
- *diagnosed-by-relation* -> *diagnosisRelation*
- *presents-the-symptoms-relation* -> *symptomRelation*

3.5.1 Ausgewähltes Darstellungsformat

Die bevorzugte Auszeichnungssprache für den Austausch der terminologischen Ressourcen ist TermBase eXchange (TBX). Aus diesem Grund werden für die vorliegende Masterarbeit die terminologischen Ressourcen vorerst in TBX gestaltet. In TBX sind die Modellierungselemente auf Englisch. In TBX bzw. Bei der Terminologieerstellung in TBX werden die Beziehungen auf Englisch aufgebaut, sodass sie nicht sprachspezifisch und deshalb international anwendbar sind (Wachowiak et al. 2021). Auf diese Art und Weise kann dieses Modell auch von terminologischen Expertinnen und Experten und Terminologieinteressierten verwendet werden (Wachowiak et al. 2021). Besser ausgedrückt ermöglicht TBX den Austausch der Terminologie zwischen CAT-Tools, aus diesem Grund ist TBX das bevorzugte Format innerhalb der translatorischen *Community*.

4. Ergebnisse

Infolge der Analyse der Datenbanken und des Aufbaus der Beziehungen kann nun mit der Erstellung von Begriffssystemen fortgesetzt werden. Zunächst wird eine Darstellung der mit VariMed erstellten Beziehungen gezeigt, dann wird das resultierende terminologische Begriffssystem mit seinen 100 Begriffen und acht Beziehungsarten vorgestellt.

4.1 Erstellung des Begriffssystems

Infolge der Analyse des Datensatzes und des VariMed-Projektes, der 15 Szenarien mit ihren 34 Schlüsselwörtern und 96 Bildern, der vier Arten von Beziehung, die von VariMed benutzt werden und der Terminologie auf Englisch ist es nun möglich mit der Erstellung des Begriffssystems fortzusetzen.

Wie im Folgenden durch Microsoft Word erstellten Diagramm zu sehen ist (Siehe Abb. 33¹), gibt es zwei Farben für die Ellipsen: blau und grün. In blau sind die Begriffe dargestellt, die aus der MedPix®-Datenbank extrahiert wurden, während in grün alle Wörter dargestellt sind, die ursprünglich nicht in der Datenbank vorhanden sind, aber durch Beziehungen in VariMed mit den ersten verbunden sind. Es gibt auch Pfeile, die die von VariMed vorgeschlagenen Beziehungen darstellen. Folgende Beziehungen werden in VariMed berücksichtigt: *accompanied by*, *type of*, *can cause* und *presents the symptom*. Sie sind an der kurzen Beschreibung auf demselben Pfeil zu erkennen. Andere Begriffe sind durch eine einzelne rote Linie ohne Pfeilspitze verbunden, welche alle Benennungen desselben Begriffes bezeichnet.

¹ Eine hochqualitative Version dieser Abbildung steht unter folgendem Link zur Verfügung: <https://ibb.co/phgSYw2>

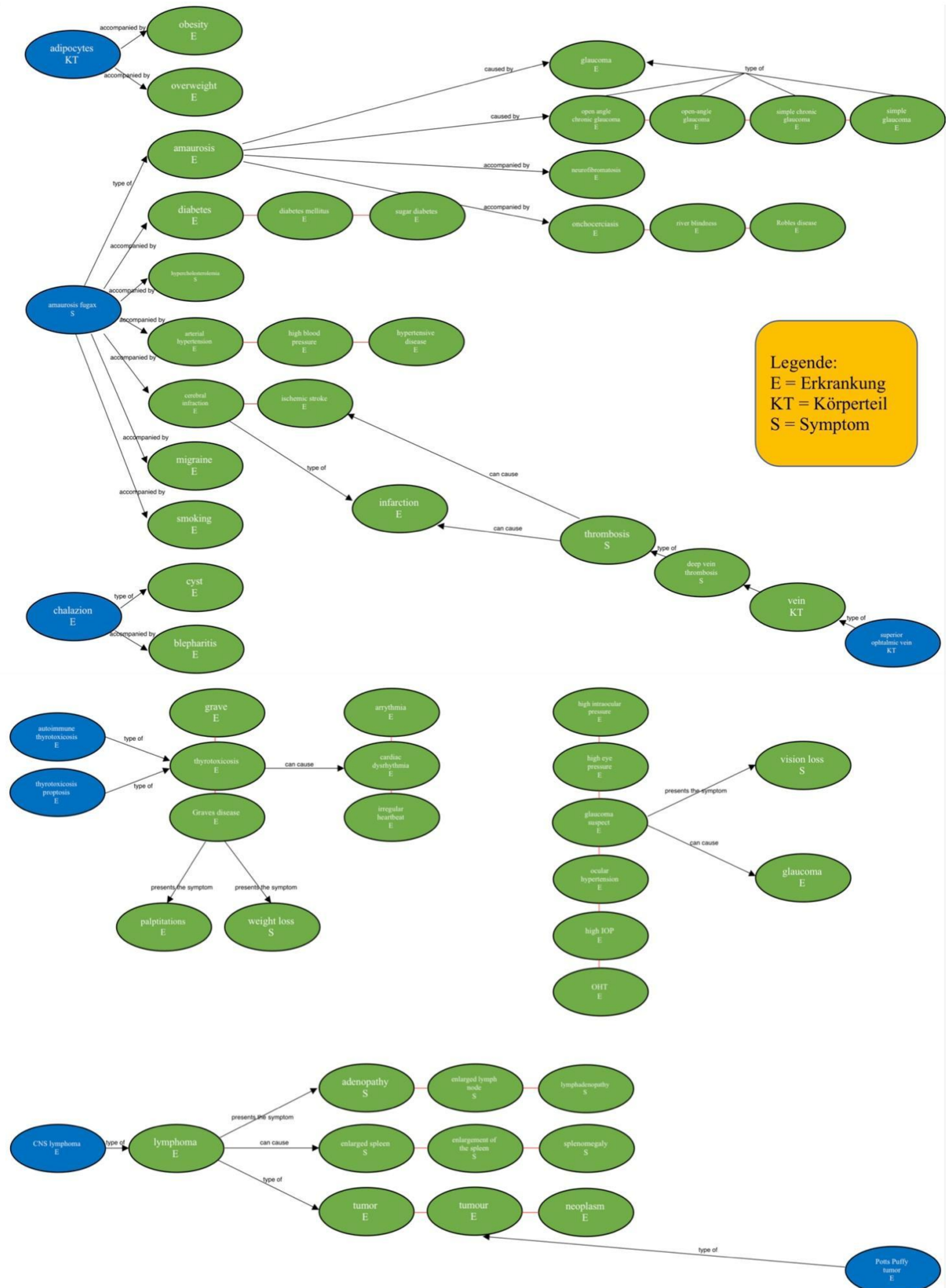


Abbildung 33. Begriffssystem mit von VariMed vorgeschlagenen Beziehungen

Wie in Abb. 33 zu sehen ist, gibt es noch viele Begriffe, für die keine Äquivalenz in VariMed gefunden wurde. Infolgedessen wurde nach Erstellung des Diagramms in Abb. 34 nach Beziehungen zwischen den Begriffen gesucht, die in VariMed nicht gefunden wurden. Erst danach wurden auch die Begriffe in Abb. 33 integriert. Hierfür war es zunächst notwendig, auf die Bedeutung aller Begriffe mit Hilfe von medizinischen Wörterbüchern, der Bildunterschrift der einzelnen Fälle und der Beschreibung der Krankheit der Szenarien zurückzugreifen. Es wurden nicht nur Beziehungen zwischen den neuen Begriffen, das heißt die, die noch keine Verbindungen hatten, hergestellt, sondern auch die Beziehungen zwischen diesen Begriffen und den bereits verknüpften Begriffen von Abb. 33 aufgebaut und so die endgültige Arbeit bereichert.

Interessanterweise wurde nach Fertigstellung der *Concept Map* ein Begriff mit seinen zuvor gefundenen Relationen, der in Abb. 33 ganz oben links angezeigt wurde, bzw. *adipocytes*, ausgeschlossen. Alle Begriffe in der *Concept Map* waren verbunden außer *adipocytes*. Eigentlich haben auf den ersten Blick Begriffe wie *obesity* und *overweight* kaum mit der Augenheilkunde zu tun. Aus diesem Grund wurde die Bedeutung dieses Begriffes weiter untersucht, um zu verstehen, weshalb dieses Schlüsselwort in diesem Bereich enthalten ist. Darüber hinaus wurde herausgefunden, dass *thyrotoxicosis/grave/graves disease/ophthalmopathy/orbitopathy* vorsieht, dass eine Art von Antikörpern die Schilddrüse und das Gewebe hinter dem Augapfel angreifen. In diesem Fall bindet sich der Antikörper an dem TSHR-Hormon, das in den Pre-Adipozyten vorhanden ist, die dann zu vollwertigen Adipozyten werden. Es ist jedoch klar, dass es sich hierbei um ein sehr spezifisches Szenario handelt. Aus diesem Grund wurde beschlossen, diesen Begriff aus der endgültigen Arbeit zu entfernen, um Verwirrung zu vermeiden und die Arbeit so verständlich wie möglich zu gestalten.

In Abb. 34² ist die Repräsentation des resultierenden Begriffssystems zu sehen. Diese Tabelle wurde, wie die Tabelle in Abb. 33, einfach mit Hilfe von Microsoft Word erstellt. Es wurde beschlossen, alle Begriffe in einem Bild bzw. auf einem Blatt darzustellen, um die Vorstellung des Ganzen besser zu vermitteln und um zu zeigen, wie alle Begriffe eigentlich miteinander verbunden sind. Da die Anzahl der Begriffe jedoch genau 100 beträgt, hat man sich für Farben und eine kleine Legende in der oberen Mitte der Abbildung entschieden, um zu vermeiden, dass die Anfangsbuchstaben (E, S oder KT) wie in Abb. 33 in allen einzigen

² Eine hochqualitative Version dieser Abbildung steht unter folgendem Link zur Verfügung:
<https://ibb.co/T8RX2Lt>

Ellipsen eingeschrieben werden müssen und, dass die Art der Relation auf jedem Pfeil ausgedrückt werden muss. Was alle Benennungen für jeden einzelnen Begriff betrifft, so wurde dieses Mal beschlossen, die Ellipsen zusammenzufügen, ohne die Verbindung durch eine Linie oder einen Pfeil herzustellen, im Gegensatz zu den kleinen roten Linien wie in Abb. 33.

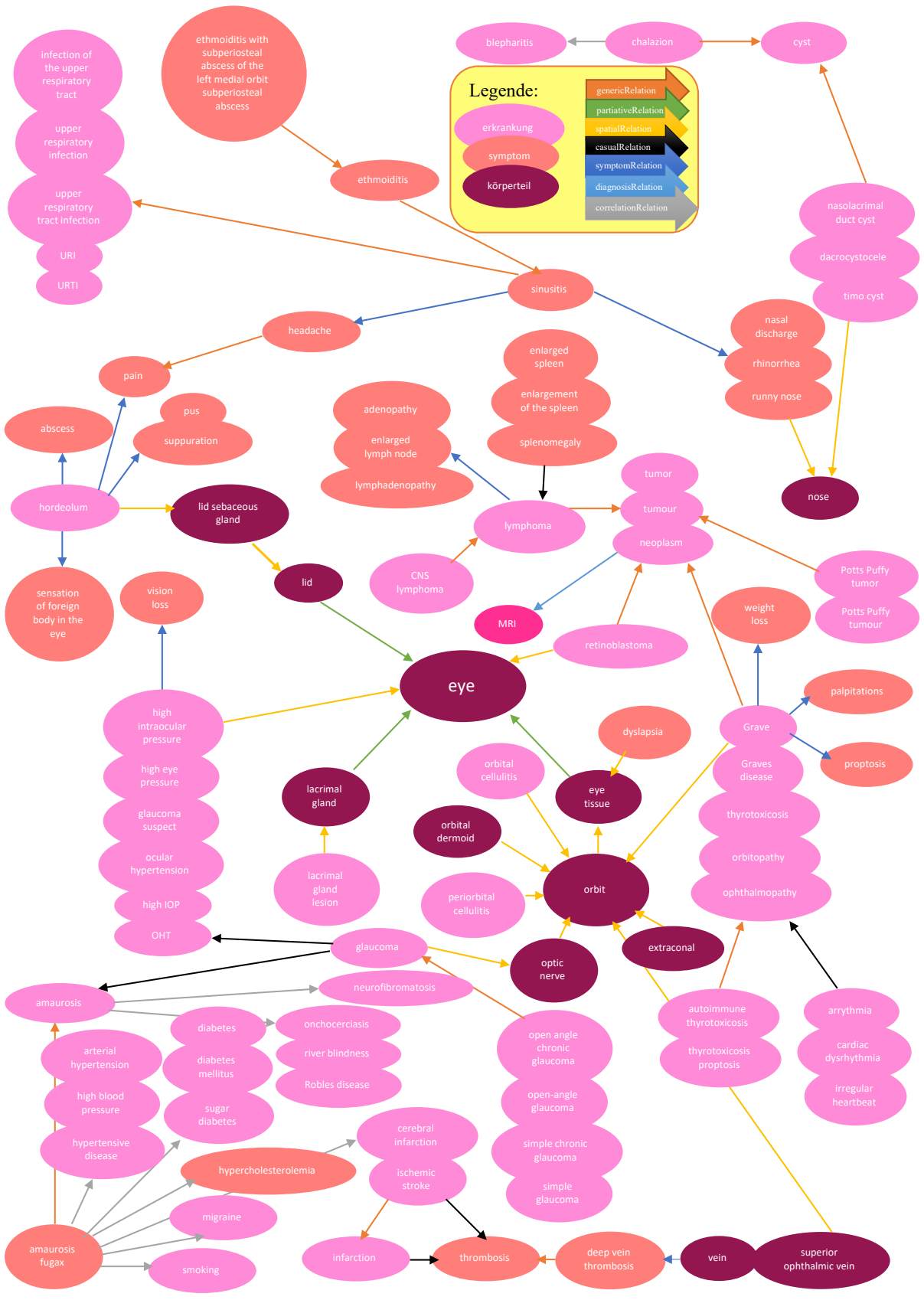


Abbildung 34. Resultierendes Begriffssystem

4.2 Aufbau von Text-Bild-Beziehungen

Auf Basis des zuvor erstellten Begriffssystems, kann nun die an und für sich interessante Fragestellung der Multimodalität adressiert werden, das heißt, die von der MedPix®-Datenbank extrahierten Bilder den Begriffen im Begriffssystem zugeordnet werden.

Die Anzahl der Bilder der Kategorie Ophthalmologie in MedPix® beläuft sich auf 96. Aus diesem Grund ist es schon vor der Erstellung der Text-Bild-Beziehungen verständlich, dass nicht allen Begriffen ein Bild zugeordnet werden kann, da die Anzahl der Begriffe 100 beträgt. Der Prozess der Klassifizierung der verschiedenen Bilder und der Untersuchung jedes einzelnen Falles jedes Szenarios mit der Zuordnung jedes Begriffs zu seinem eigenen Bild und den damit verbundenen Benennungen wird auch dieses Mal stattfinden, beginnend mit Szenario 1 und endend mit Szenario 15.

Szenario 1 liefert keine Bilder, deshalb wird gleich zu Szenario 2 übergegangen. Szenario 2 verfügt über einen Fall und zwei Bilder. Der Fall lautet: *Periorbital Cellulitis secondary to Infected Chalazion*. Beide vorhandenen Bilder verfügen über die gleichen Merkmale: gleiche Bildunterschrift, gleiche Ebene und gleiche Modalität. Aus diesem Grund wird nur ein Bild der Datenbank hinzugefügt. Mit Hilfe der Bildunterschrift wird es leichter zu verstehen sein, worum es geht. Fehlende Bildunterschrift kann für Unklarheit sorgen. Bei hochspezialisierten medizinischen Szenarien besteht die Schwierigkeit, nach verschiedenen Kategorien zu unterscheiden. Symptome, Erkrankungen und Körperteile sind Teile eines großen, koexistierenden Ganzen, weshalb die Schwierigkeit darin liegt, dass es praktisch unmöglich ist, eine Trennlinie zwischen diesen Kategorien zu ziehen. Dies ist nicht nur der Fall der Pathologien, sondern auch der Symptome, die von unterschiedlichen Erkrankungen verursacht werden oder der Körperteile, die ebenso nebeneinander existieren und sich einander deshalb ständig beeinflussen.

Bei *concept #2* und *concept #3* wurde beschlossen, dasselbe Bild mit derselben Bildunterschrift zweimal zu verwenden, weil die zwei Begriffe nicht unbedingt bei jedem Fall verbunden sind. Darüber hinaus handelt es sich nicht um zwei Benennungen desselben Begriffes. *Chalazion (concept #2)* und *periorbital cellulitis (concept #3)* können nämlich auch getrennt auftreten. Aus diesem Grund könnte eine Benutzerin oder ein Benutzer des Begriffssystems nur für einen der zwei Begriffe Interesse haben. Infolgedessen, falls sich die Benutzerin oder der Benutzer nur für *periorbital cellulitis* interessiert, würde es wenig Sinn ergeben, einen Verweis auf *chalazion* zu erstellen, oder umgekehrt.

Die ersten Synonyme kommen erst bei Szenario 3 vor. Es handelt sich um die Begriffe *dacrocystocele*, *nasolacrimal duct cyst* und *timo cyst*. Ebenso bei Szenario 9 wurde dasselbe Bild sowohl für *Potts Puffy Tumor* bzw. *Potts Puffy Tumour* als auch für *Proptosis* verwendet, weil beide im selben Bild zu sehen sind, und keine Bilder, wo sie getrennt gezeigt werden, vorhanden sind. Szenario 13 wurde entfernt, weil es nur das Schlüsselwort *orbits* zur Verfügung stellt, was bereits in der Singularform im Szenario 10 hinzugefügt wurde. Ein ähnlicher Fall kommt für Szenario 14 vor, wo die Schlüsselwörter *Graves disease* und *thyrotoxicosis* bereits als zweite und dritte Benennung im Szenario 5 hinzugefügt wurden.

Die vollständige in TBX konvertierte Fassung dieses Auftrags findet sich im Anhang am Ende der vorliegenden Masterarbeit. Daraus wird jedoch deutlich, dass die Anzahl der Links bzw. der Bilder lediglich vierzehn beträgt. Es handelt sich um eine sehr geringe Anzahl im Vergleich zu der Anzahl der Begriffe, die in Abb. 34 zu sehen sind. Dafür gibt es verschiedene Gründe: Mehrere Begriffe, die von VariMed extrahiert wurden, um Relationen zu erstellen, verfügen über kein Bild in MedPix®. Darüber hinaus wurden viele Begriffe hinzugefügt, die in keiner der beiden Datenbanken vorhanden waren, um die Vollständigkeit zu gewährleisten und die verschiedenen *Cluster* von Begriffen besser zu verknüpfen. Beispielsweise war der Begriff *cyst* weder in VariMed noch in MedPix® vorhanden, *nasolacrimal duct cyst*, *dacrocystocele* und *timo cyst* aber schon. Deshalb wurde ebenso der Begriff *cyst* hinzugefügt, um die *genericRelation* aufbauen zu können. Ein weiterer Grund, weshalb die Anzahl der Bilder von der Anzahl der Begriffe so erheblich abweicht, hat ebenso mit der Tatsache zu tun, dass eine bedeutsame Anzahl an Begriffen, die in Abb. 34 gezeigt wurde, Benennungen desselben Begriffes sind. In diesem Fall werden sie als *term2*, *term3* usw. desselben Begriffes dargestellt und einem einzigen Bild zugeordnet.

Folglich wird in Abb. 35 ein Beispiel bzw. ein Begriff vom TBX-File, der als Anhang zu finden ist, angezeigt. Es handelt sich um den Begriff *grave*, der als *conceptEntry* Nummer 8 nummeriert worden ist. Darüber hinaus wurden die Kategorie in rosa, die Benennungen in gelb, die Beziehungen in violett, den Link zum Bild in blau und die Beschreibung des Begriffes in grün hervorgehoben.

```

</conceptEntry>
<conceptEntry id="008">
  <type type="subjectField">disease</type>
  <transacGrp>
    <transac type="transactionType">origination</transac>
    <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
    <date>21-10-18_18h-49m</date>
  </transacGrp>
  <langSec xml:lang="en">
    <termSec id="008-en-t0">
      <term>grave</term>
    </termSec>
    <termSec id="008-en-t1">
      <term>Graves disease</term>
    </termSec>
    <termSec id="008-en-t2">
      <term>thyrotoxicosis</term>
    </termSec>
    <termSec id="008-en-t3">
      <term>orbitopathy</term>
    </termSec>
    <termSec id="008-en-t4">
      <term>ophthalmopathy</term>
    </termSec>
  </langSec>
  <descrip type="spatialRelation">020</descrip>
  <descrip type="genericRelation">039</descrip>
  <descrip type="symptomRelation">018</descrip>
  <descrip type="symptomRelation">040</descrip>
  <descrip type="symptomRelation">041</descrip>
  <descripGrp>
    <xref xmlns:target_file="https://medpix.nlm.nih.gov/topic?
id=3a1decd7-fe7e-47e6-9848-606a90a5461f" type="xGraphic "/>
    <descrip type="description">Grave's Ophthalmopathy.
Noncontrast axial CT image through the orbits. Soft tissue windows
reveal bilateral proptosis, periorbital tissue edema, intact but
enlarged inferior recti muscles, increased periorbital fat

```

Abbildung 35. Begriff *grave* im TBX-Format

5. Diskussion

Wie aus dem Kapitel über die Methodik abgeleitet wurde, bestehen Pathologien, Symptome und Körperteile in medizinischen hochspezialisierten Szenarien immer nebeneinander. In den meisten Fällen weist ein Bild mehr als eine dieser Kategorien auf. Infolgedessen ist es eine schwierige Aufgabe, eine Trennlinie zu ziehen, um diese Art von Informationen bestmöglich zu modellieren. Aufgrund der Schwierigkeit des beschriebenen Vorgangs ist es daher empfehlenswert, sich vor der Erstellung der multimodalen Datenbank genügend Zeit einzuplanen, um die verschiedenen Erkrankungen, Symptome und die Partie, wo sie auftreten, zu verstehen. Aus diesem Grund ist es notwendig zu verstehen, welche Kategorie (Pathologie, Symptom oder Körperteil) in dem Bild vorherrschend ist. Dieses Bild soll infolge der Analyse verwendet werden, um mit einem Begriff zusammengefügt zu werden. Wenn ein Bild mehrere Erkrankungen, Symptome und/oder Körperteile gleichermaßen aufweist, ist es ratsam, das Bild für jeden dieser Begriffe zu wiederholen, um Querverweise zu vermeiden, sodass jeder Begriff über alle Informationen verfügt, welche mit diesem zusammenhängen.

Die Annahmen in der Einleitung der vorliegenden Masterarbeit sahen vor, dass es gemeinsame Variablen geben wird, die auf gemeinsamen Eigenschaften basieren, so wie das Wortfeld (bedeutungsähnliche Wörter), den Bezugsgrad zwischen Bild und Text und die Beziehungen zwischen Nebeninformationen bzw. zusätzlichen Informationen der Patientinnen und Patienten, die den Kontext prägen, in dem die multimodale Terminologie angesiedelt ist.

In Wirklichkeit war von all diesen Annahmen nur die Hypothese bezüglich des Wortfeldes zwischen den verschiedenen Begriffen in der MedPix®-Datenbank in Kombination mit denen in der VariMed-Datenbank und den Bezugsgrad zwischen Bild und Text sinnvoll zu berücksichtigen. Damit wurden die bedeutungsähnlichen Wörter in TBX als Benennungen unter demselben Begriff aufgenommen. Informationen wie demografische Daten der Patientinnen und Patienten wurden nicht berücksichtigt, da sie für die finale terminologische Ressource keinen unmittelbaren Mehrwert dargestellt hätten.

In dieser Arbeit wurde ein hypothetisches Modell für die Modellierung der multimodalen Terminologie im medizinischen Bereich und insbesondere im spezifischen medizinischen Bereich der Ophthalmologie entworfen. Wenn es einerseits gelingt, Ordnung in ein Gebiet zu bringen, welches nach den Schätzungen von Porep und Steudel (1974) eine sehr hohe Anzahl von Synonymen für jede Pathologie, jedes Symptom und jedes Körperteil aufweist, so ist es andererseits, wie gezeigt wurde, schwierig, für jede Kategorie ein Bild zu finden. Aus diesem Grund kann ein Werk dieser Art in solch hochspezialisierten Szenarien

einen Beitrag für Menschen leisten, die sich dieser Welt nähern, wie beispielsweise Medizinstudierende, da es ihnen einen Überblick über ihr Studiengebiet in prägnanter Weise verschafft, ohne dabei wichtige Informationen auszulassen. Für das bereits spezialisierte Gesundheitspersonal ist diese Arbeit möglicherweise konfus und dürfte für eine bessere Archivierung von multimodalem medizinischem Material von geringerem Wert sein. Es muss berücksichtigt werden, dass es sich bei der vorliegenden Masterarbeit um eine Pionierarbeit handelt. Wie bereits erläutert wurde, ist VariMed das einzige Projekt, das dem aktuellen Forschungsstand entsprechend (Oktober 2021) verwirklicht wurde und zur Verfügung steht. Projekte, die sich mit der Modellierung der spezialisierten medizinischen bzw. ophthalmologischen Terminologie befassen, sind im aktuellen Forschungsstand noch nicht vorhanden.

Im Laufe der Verfassung der vorliegenden Masterarbeit wurde festgestellt, dass die Wiederholung von Bildern bei verschiedenen Begriffen nützlich sein kann, um Missverständnisse zu vermeiden und, um jedem Begriff Vollständigkeit zu gewährleisten. Außerdem hat sich gezeigt, dass es nicht immer einfach zu verstehen ist, welche Teile des Bildes sich auf welche Pathologie, welches Symptom oder welches Körperteil beziehen. Dies kann als das Hauptproblem des vorliegenden Projekts definiert werden, im Gegensatz zu VariMed, wo das Problem die Ärztinnen bzw. Ärzte sind, die die Patientin bzw. den Patienten als Krankheit und nicht als Patientin bzw. Patient sehen, und wo dies trotz der VariMed-Philosophie, die Kommunikation zwischen Patientinnen bzw. Patienten und dem Gesundheitspersonal gleichzuschalten, oft nicht möglich ist, weil die Krankheit die Dringlichkeit vorgibt.

Zu guter Letzt kann die vorliegende Masterarbeit als Sprungbrett für mögliche zukünftige Studien fungieren, die auf der Basis dieses Projektes die Forschung im Bereich der Modellierung multimodaler medizinischer Terminologie erweitern wollen.

6. Conclusio

In der vorliegenden Masterarbeit wurde ein multimodaler medizinischer Datensatz am Beispiel des ophthalmologischen Bereichs analysiert. Es wurden terminologische Informationen, sowie die von dem Datensatz vorgeschlagenen Schlüsselwörter, oder die Bilder der Erkrankungen unterschiedlicher Szenarien und Fälle analysiert. Infolge der Analyse des Datensatzes wurden solche Ressourcen modelliert und mit Hilfe von Beziehungen zusammengefügt. Zuletzt wurden Text-Bild-Beziehungen in TBX konvertiert und ein terminologisches Begriffssystem erstellt. Diese Arbeit soll aufzeigen, wie die multimodale medizinische spezialisierte Terminologie am besten modelliert werden kann. Da es sich um ein neues Projekt handelt, wurde lediglich ein eingeschränkter Bereich (die Ophthalmologie) betrachtet. Dies schließt allerdings die Möglichkeit nicht aus, dass weitere Forscherinnen und Forscher die gesamte Terminologie aus dem ganzen Datensatz extrahieren und modellieren können.

Darüber hinaus besteht die Hoffnung, dass sich terminologische Begriffssysteme durchsetzen werden, da andere semantische Netze in gleicher Weise zum Wissensmanagement beitragen.

7. Bibliographie

7.1 Literaturquellen

- Allwood, Jens (2008). *Multimodal corpora. An International Handbook*. Mouton de Gruyter. 207-225.
- Arntz, Reiner; Picht, Heribert & Schmitz, Klaus-Dirk (2014). *Einführung in die Terminologie arbeit*. Georg Olms Verlag.
- Bateman, John (2014). *Text and image: A critical introduction to the visual/verbal divide*. Routledge.
- Beddies, Thomas; Brinkschulte, Eva; Doetz, Susanne; Hess, Volker; Marcus, Benjamin; Marz, Ilona; Nabielek, Rainer; Hulverscheidt, Marion (2015). *Medizinische Fachsprache. Skript zum Kurs Modellstudiengang Humanmedizin*. Charité. Universitätsmedizin. Berlin.
- Benveniste, Émile (1986). *The Semiology of Language. Translated by Genette Ashby and Adelaide Russo. Semiotics: An Introductory Anthology*. 228-245.
- Biedermann, Hubert (2016). *Lean Smart Maintenance: Wertschöpfende, lernorientierte und ressourceneffiziente Instandhaltung*. Lean Smart Maintenance: Konzepte. Instandhaltungsforum. TÜV Media, Köln. 19–29.
- Bjørnbo Engel, Trine; Fliedner, Frederikke Petrine; El-Ali, Henrik; Elias Hansen, Anders & Kjaer, Andreas (2020). *Diffusion weighted magnetic resonance imaging (DW-MRI) as a non-invasive, tissue cellularity marker to monitor cancer treatment response*. *Bmc Cancer*. 1-9.
- Budin, Gerhard (2006). *Theoretische und methodische Grundlagen integrierter Wissens- und Lerntechnologien*. In: Mettinger, Arthur; Zwiauer, Charlotte & Oberhuemer, Pamela (eds.) *E-Learning an der Universität Wien. Forschung – Entwicklung – Einführung*. Münster: Waxmann, 43–56.
- Burstein, Marcos Herrera (1997). *Los conceptos de emociones y sus relaciones semánticas*. 107-129.

- Cruse, Alan (2011). *Meaning in language: An introduction to semantics and pragmatics*. Oxford University Press.
- Dobratz, Susanne (2008). *Einführung in Semantic Web*. HU Berlin.
- Drewer, Petra; Massion, François & Pulitano, Donatella (2017a). *Was haben Wissensmodellierung, Wissensstrukturierung, künstliche Intelligenz und Terminologie miteinander zu tun?*. Deutsches Institut für Terminologie e.V.
- Drewer, Petra & Schmitz, Klaus-Dirk (2017b). *Grundlagen der Terminologiewissenschaft*. In *Terminologiemanagement*. Springer Vieweg, Berlin, Heidelberg. 5-22.
- Drewer, Petra & Ziegler, Wolfgang (2011). *Technische Dokumentation: eine Einführung in die übersetzungsgerechte Texterstellung und in das Content-Management*. Vogel Buchverlag.
- Ehrig, Marc & Studer, Rudi (2006). *Wissensvernetzung durch Ontologien*. In *Semantic Web Wege zur vernetzten Wissensgesellschaft*. Springer, Berlin, Heidelberg. 469-484.
- El Naqa, Issam & Murphy, Martin (2015). *What is machine learning?*. In *machine learning in radiation oncology*. Springer, Cham. 3-11.
- Faber, Pamela (2015). *Frames as a Framework for Terminology*. *Handbook of terminology*. John Benjamins Publishing Company. 14-33.
- Fellbaum, Christiane (2010). *WordNet. Theory and applications of ontology. Computer applications*. Springer, Dordrecht. 231-243.
- Giunchiglia, Fausto & Shvaiko, Pavel (2003). *Semantic matching*. *The Knowledge Engineering Review*. 265-280
- Guzmán, Sergio & Peter Alken. (2016). *Operative Anatomie: Körperebenen und Prinzipien der Zugänge*. Die Urologie. Springer, Berlin, Heidelberg. 3-4.
- Helbig, Hermann (2013). *Die semantische Struktur natürlicher Sprache: Wissensrepräsentation mit MultiNet*. Springer-Verlag.
- Henne, Helmut; Mentrup, Wolfgang & Möhn, Dieter (1978). *Fachsprache Medizin*. Interdisziplinäres deutsches Wörterbuch in der Diskussion. Vol. 45. Pädagogischer Verlag Schwann.

- Jewitt, Carey Ed (2011). *The Routledge handbook of multimodal analysis*. Routledge/Taylor & Francis Group.
- Kienreich, Wolfgang & Strohmaier, Markus (2006). *Wissensmodellierung - Basis für die Anwendung semantischer Technologien*. In *Semantic Web Wege zur vernetzten Wissensgesellschaft*. Springer, Berlin, Heidelberg. 359-371.
- Kiros, Ryan; Salakhutdinov, Ruslan & Zemel, Richard (2014). *Multimodal neural language models*. International conference on machine learning.
- Klavans, Judith & Muresan, Smaranda (2000). *DEFINDER: Rule-based methods for the extraction of medical terminology and their associated definitions from on-line text*. Proceedings of the AMIA Symposium. American Medical Informatics Association.
- Kroeber-Riel, Werner (1993). *Bildkommunikation: Imagerystrategien für die Werbung*. Vahlen. 53.
- Lemke, Jay (1998). *Multiplying meaning. Reading science: Critical and functional perspectives on discourses of science*. 87-113.
- Louridas, Panos & Ebert, Christof (2016). *Machine Learning*. IEEE COMPUTER SOCIETY.
- Mayer, Felix (2019). *Terminologiewerk und Terminographie*. Terminologie: Epochen–Schwerpunkte–Umsetzungen. Springer Vieweg, Berlin, Heidelberg. 83-93.
- Menne, Albert (1991). *Einführung in die formale Logik*. Wissenschaftliche Buchgesellschaft Darmstadt.
- Muscariello, Marta. (2010). *Introduzione alla fonetica: con esercizi*. Arcipelago.
- Nora, Simon; Minc, Alain; Kalbhen, Uwe & Lohmar, Ulrich (1979). *Die Informatisierung der Gesellschaft*. Campus-Verlag.
- Nuopponen, Anita (2014) Tangled web of concept relations. concept relations for ISO 1087-1 and ISO 704. In *Terminology and Knowledge Engineering Berlin, Germany, 2014*. Association for Computational Linguistics.
- Pedersen, Susanne & Hasselbring, Wilhelm (2002). *Begriffssysteme für die medizinische Dokumentation*. 47-52.

- Pellegrini, Tassilo & Blumauer, Andreas (2006). *Semantic Web und semantische Technologien: Zentrale Begriffe und Unterscheidungen*. In *Semantic Web Wege zur vernetzten Wissensgesellschaft*. Springer Verlag. 9-26.
- Prieto Velasco, Juan Antonio (2012). *A corpus-based approach to the multimodal analysis of specialized knowledge*. *Language resources and evaluation*, 47(2), 399-423.
- Pulitano, Donatella (2006). *Il terminologo: cosa fa, cosa deve sapere, come si diventa*. *Mediazioni*, (3).
- Quillian, Ross (1967). Word concepts: A theory and simulation of some basic semantic capabilities. *Behavioral science*, 12(5), 410-430.
- Reutelshoefer, Jochen; Baumeister, Joachim & Puppe, Frank (2009). *A data structure for the refactoring of multimodal knowledge*. In *Proceedings of the 5th Workshop on Knowledge Engineering and Software Engineering*. 33-45.
- Rodríguez Lopez, Clara Ines; Prieto Velasco, Juan Antonio & Sánchez, Maribel Tercedor (2013) *Multimodal representation of specialised knowledge in ontology-based terminological databases: the case of EcoLexicon1*. *The Journal of Specialized Translation* 20, 49-67.
- Rosch, Eleanor (1973). *Natural Categories*. *Cognitive Psychology*. 7, 573-605.
- Roth, Sandra (2006). *Lexikalisch-semantische Netze: Anwendungsperspektiven für die Computerlinguistik*.
- Sánchez, Maribel Tercedor & Juan Antonio Prieto Velasco. (2014). *También los pacientes hacen terminología: retos del proyecto VariMed*. *Panace*. 95-102.
- Schwarz-Friesel, Monika & Consten, Manfred (2014). *Einführung in die Textlinguistik*. Darmstadt. 50-73.
- Srivastava, Nitish & Salakhutdinov, Ruslan (2012) *Multimodal learning with deep boltzmann machines*. *Advances in neural information processing systems*.
- Staab, Steffen & Studer, Rudi (2010). *Handbook on ontologies*. Springer Science & Business Media. 1-17.

- Stock, Wolfgang & Stock, Mechtild (2008). *Wissensrepräsentation: Informationen auswerten und bereitstellen* (Vol. 2). Oldenbourg Verlag.
- Stöckl, Hartmut (2011). *Sprache-Bild-Texte lesen. Bausteine zur Methodik einer Grundkompetenz*. Bildlinguistik. Theorien-Methoden-Fallbeispiele. 43-70.
- Studer, Rudi; Benjamins, Richard & Fensel, Dieter (1998). *Knowledge engineering: principles and methods*. *Data & knowledge engineering*. 25(1-2), 161-197.
- Tatievskaja, Elena (2003). *Einführung in die Aussagenlogik*. Berlin: Logos-Verlag.
- Vassallo, Nicla (2015). *Teoria della conoscenza*. Gius. Laterza & Figli Spa.
- Vigini, Giuliano (1985). *Glossario di biblioteconomia e scienza dell'informazione*. Lampi di stampa. 101.
- Wachowiak, Lennart; Lang, Christian; Heinisch, Barbara & Gromann, Dagmar (2021). Towards Learning Terminological Concept Systems from Multilingual Natural Language Text. In Gromann, Dagmar; Sérasset, Gilles ; Declerck, Thierry ; McCrae, John P. ; Gracia, Jorge ; Bosque-Gil, Julia ; Bobillo, Fernando ; & Heinisch, Barbara (Hg.) *3rd Conference on Language, Data and Knowledge (LDK 2021)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik.
- Wierzbicka, Anna (1984). *Apples are not a "Kind of Fruit": The Semantics of Human Categorization*. *American Ethnologist*. 313–328.
- Winston, Morton; Chaffin, Roger & Herrmann, Douglas (1987). *A taxonomy of part-whole relations*. *Cognitive science*. 417-444.
- Wolf, Thomas & Strohschen, Jacqueline-Helena (2018). Digitalisierung: Definition und Reife. *Informatik-Spektrum*. 41(1), 56-64.

7.2 Linksammlung

BabelNet (2021). <https://babelnet.org>

Badaraite, Zivile (2021). *Benvenuti nel futuro: l'Industria 4.0*.

<https://www.engusa.com/it/posts/welcome-to-the-future-industry-4-0>

- Beam, Andrew L. (2020). Medical Data for Machine Learning.
<https://github.com/beamandrew/medical-data>
- EcoLexicon. (2021). <http://ecolexicon.ugr.es/en/index.htm>
- Fernandes, Sheldon (2015). *Ontologies vs Taxonomies vs Thesauri , and its place on the Semantic Web*. <https://blog.thedigitalgroup.com/ontologies-vs-taxonomies-vs-thesauri-and-its-place-on-the-semantic-web>
- Frost, Jetta (2020). *Wissensmanagement Definition: Was ist „Wissensmanagement“?*. Springer Gabler. <https://wirtschaftslexikon.gabler.de/definition/wissensmanagement-47468>
- Gillis, Alexander (2019). *Internet of things (IoT)*:
<https://internetofthingsagenda.techtarget.com/definition/Internet-of-Things-IoT>
- IAEA (2021). https://inis.iaea.org/search/search.aspx?orig_q=RN:40031984
- ImageNet (2021). <http://www.image-net.org/>
- ISO 704:2009. <https://www.iso.org/obp/ui/#iso:std:iso:704:ed-3:v1:en>
- ISO 1087:2019. <https://www.iso.org/obp/ui/#iso:std:iso:1087:ed-2:v1:e>
- KÜDES (2018). *Empfehlungen für die Terminologearbeit*.
http://www.cotsoes.org/sites/default/files/CST_Empfehlungen_fuer_die_Terminologiearbeit.pdf
- MedPix® (2021). <https://medpix.nlm.nih.gov/home>
- Microsoft (2021) [https://docs.microsoft.com/en-us/previous-versions/dotnet/netframework-1.1/x2dbyw72\(v=vs.71\)](https://docs.microsoft.com/en-us/previous-versions/dotnet/netframework-1.1/x2dbyw72(v=vs.71))
- Minini, Andrea (2021). Online Personal Knowledge Base.
<http://www.andreaminini.com/ai/rappresentazione-conoscenza/>

7.3 Videoquellen

- Hüsing, Thomas (2016). Wissensmodellierung für Dummies (Thomas Hüsing). [Videoclip 1'19"] <https://www.youtube.com/watch?v=FPd9Y8UC9Rw> (Stand: 16.10.2020).

Anhänge

Anhang

1

```
<?xml version='1.0' encoding='UTF-8'?>
<?xml-model href="https://raw.githubusercontent.com/LTAC-Global/TBX-
Core_dialect/master/Schemas/TBXcoreStructV03_TBX-
Core_integrated.rng" type="application/xml" schematypens="http://
relaxng.org/ns/structure/1.0"?>
<?xml-model href="https://raw.githubusercontent.com/LTAC-Global/TBX-
Core_dialect/master/Schemas/TBX-Core.sch" type="application/xml"
schematypens="http://purl.oclc.org/dsdl/schematron"?>
<tbx type="TBX-Core" style="dca" xml:lang="en"
xmlns="urn:iso:std:iso:30042:ed-2">
  <tbxHeader>
    <fileDesc>
      <sourceDesc>
        <p>Data manually curated by Andrea Ferri and automatically
converted to TBX by Text2TCS</p>
      </sourceDesc>
    </fileDesc>
    <encodingDesc>
      <p type="XCSURI">TBXCSV02.xcs</p>
    </encodingDesc>
  </tbxHeader>
  <text>
    <body>
      <conceptEntry id="001">
        <type type="subjectField">symptom</type>
        <transacGrp>
          <transac type="transactionType">origination</transac>
          <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
          <date>21-10-18_18h-49m</date>
        </transacGrp>
        <langSec xml:lang="en">
          <termSec id="001-en-t0">
            <term>amaurosis fugax</term>
          </termSec>
        </langSec>
        <descrip type="genericRelation">025</descrip>
        <descrip type="correlationRelation">026</descrip>
        <descrip type="correlationRelation">027</descrip>
        <descrip type="correlationRelation">028</descrip>
        <descrip type="correlationRelation">029</descrip>
        <descrip type="correlationRelation">030</descrip>
        <descrip type="correlationRelation">060</descrip>
      </conceptEntry>
      <conceptEntry id="002">
        <type type="subjectField">disease</type>
        <transacGrp>
          <transac type="transactionType">origination</transac>
          <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
          <date>21-10-18_18h-49m</date>
        </transacGrp>
        <langSec xml:lang="en">
          <termSec id="002-en-t0">
```

```

        <term>chalazion</term>
      </termSec>
    </langSec>
    <descrip type="genericRelation">031</descrip>
    <descrip type="correlationRelation">032</descrip>
    <descripGrp>
      <xref xmlns:target_file="https://medpix.nlm.nih.gov/topic?
id=44943c5a-e1cf-4d6b-9171-87abd0598daa" type="xGraphic "/>
      <descrip type="description">Preseptal cellulitis is
present at the right orbit. No post septal extension is seen. A
small focal area of edema or fluis is noted along the medial aspect
of the anterior chamber of the globe.</descrip>
    </descripGrp>
  </conceptEntry>
  <conceptEntry id="003">
    <type type="subjectField">disease</type>
    <transacGrp>
      <transac type="transactionType">origination</transac>
      <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
      <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
      <termSec id="003-en-t0">
        <term>periorbital cellulitis</term>
      </termSec>
    </langSec>
    <descrip type="spatialRelation">020</descrip>
    <descripGrp>
      <xref xmlns:target_file="https://medpix.nlm.nih.gov/topic?
id=44943c5a-e1cf-4d6b-9171-87abd0598daa" type="xGraphic "/>
      <descrip type="description">Periorbital Cellulitis
secondary to Infected Chalazion. Preseptal cellulitis is present at
the right orbit. No post septal extension is seen. A small focal
area of edema or fluis is noted along the medial aspect of the
anterior chamber of the globe.</descrip>
    </descripGrp>
  </conceptEntry>
  <conceptEntry id="004">
    <type type="subjectField">disease</type>
    <transacGrp>
      <transac type="transactionType">origination</transac>
      <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
      <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
      <termSec id="004-en-t0">
        <term>hordeolum</term>
      </termSec>
    </langSec>
    <descrip type="spatialRelation">033</descrip>
    <descrip type="symptomRelation">034</descrip>
    <descrip type="symptomRelation">035</descrip>

```

```

    <descrip type="symptomRelation">036</descrip>
    <descrip type="symptomRelation">037</descrip>
  </conceptEntry>
  <conceptEntry id="005">
    <type type="subjectField">disease</type>
    <transacGrp>
      <transac type="transactionType">origination</transac>
      <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
      <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
      <termSec id="005-en-t0">
        <term>dacrocystocele</term>
      </termSec>
      <termSec id="005-en-t1">
        <term>nasolacrimal duct cyst</term>
      </termSec>
      <termSec id="005-en-t2">
        <term>timo cyst</term>
      </termSec>
    </langSec>
    <descrip type="genericRelation">031</descrip>
    <descrip type="spatialRelation">038</descrip>
    <descripGrp>
      <xref xmlns:target_file="https://medpix.nlm.nih.gov/topic?
id=f38413a0-5963-45a9-b6f0-6fa72d557254" type="xGraphic "/>
      <descrip type="description">Cystic mass with peripheral
enhancement at the left medial canthi.</descrip>
    </descripGrp>
  </conceptEntry>
  <conceptEntry id="006">
    <type type="subjectField">symptom</type>
    <transacGrp>
      <transac type="transactionType">origination</transac>
      <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
      <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
      <termSec id="006-en-t0">
        <term>ethmoiditis with subperiosteal abscess of the left
medial orbit subperiosteal abscess</term>
      </termSec>
    </langSec>
    <descrip type="genericRelation">007</descrip>
    <descripGrp>
      <xref xmlns:target_file="https://medpix.nlm.nih.gov/topic?
id=32fab30b-8e6c-48f2-aa4a-dc47c70938e2" type="xGraphic "/>
      <descrip type="description">Orbit - Superiosteal Abscess.
Ethmoiditis with subperiosteal abscess of the left medial orbit
subperiosteal abscess. Notice the fluid collection adjacent to the
medial rectus muscle, and the associated etmoid sinus disease.
Measurements.</descrip>

```

```

    </descripGrp>
  </conceptEntry>
  <conceptEntry id="007">
    <type type="subjectField">symptom</type>
    <transacGrp>
      <transac type="transactionType">origination</transac>
      <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
      <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
      <termSec id="007-en-t0">
        <term>ethmoiditis</term>
      </termSec>
    </langSec>
    <descrip type="genericRelation">017</descrip>
  </conceptEntry>
  <conceptEntry id="008">
    <type type="subjectField">disease</type>
    <transacGrp>
      <transac type="transactionType">origination</transac>
      <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
      <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
      <termSec id="008-en-t0">
        <term>grave</term>
      </termSec>
      <termSec id="008-en-t1">
        <term>Graves disease</term>
      </termSec>
      <termSec id="008-en-t2">
        <term>thyrotoxicosis</term>
      </termSec>
      <termSec id="008-en-t3">
        <term>orbitopathy</term>
      </termSec>
      <termSec id="008-en-t4">
        <term>ophthalmopathy</term>
      </termSec>
    </langSec>
    <descrip type="spatialRelation">020</descrip>
    <descrip type="genericRelation">039</descrip>
    <descrip type="symptomRelation">018</descrip>
    <descrip type="symptomRelation">040</descrip>
    <descrip type="symptomRelation">041</descrip>
  </descripGrp>
  <xref xmlns:target_file="https://medpix.nlm.nih.gov/topic?
id=3a1decd7-fe7e-47e6-9848-606a90a5461f" type="xGraphic "/>
  <descrip type="description">Grave's Ophthalmopathy.
Noncontrast axial CT image through the orbits. Soft tissue windows
reveal bilateral proptosis, periorbital tissue edema, intact but
enlarged inferior recti muscles, increased periorbital fat

```

```

distribution.</descrip>
  </descripGrp>
</conceptEntry>
<conceptEntry id="009">
  <type type="subjectField">disease</type>
  <transacGrp>
    <transac type="transactionType">origination</transac>
    <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
    <date>21-10-18_18h-49m</date>
  </transacGrp>
  <langSec xml:lang="en">
    <termSec id="009-en-t0">
      <term>epiphora tearing</term>
    </termSec>
  </langSec>
</conceptEntry>
<conceptEntry id="010">
  <type type="subjectField">body part</type>
  <transacGrp>
    <transac type="transactionType">origination</transac>
    <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
    <date>21-10-18_18h-49m</date>
  </transacGrp>
  <langSec xml:lang="en">
    <termSec id="010-en-t0">
      <term>superior ophthalmic vein</term>
    </termSec>
  </langSec>
  <descrip type="spatialRelation">020</descrip>
  <descrip type="symptomRelation">042</descrip>
  <descripGrp>
    <xref xmlns:target_file="https://medpix.nlm.nih.gov/topic?
id=0936415f-49bc-4363-8578-45377fead698" type="xGraphic "/>
    <descrip type="description">Superior ophthalmic vein
thrombosis left eye. The SOV does not enhance and is
indistinguishable from the adjacent superior rectus muscle. (Note
patent SOV on right side.) Severe proptosis. Marked swelling of
extra-ocular muscles. Preseptal cellulitis. Abscess formation medial
canthus.</descrip>
  </descripGrp>
</conceptEntry>
<conceptEntry id="011">
  <type type="subjectField">symptom</type>
  <transacGrp>
    <transac type="transactionType">origination</transac>
    <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
    <date>21-10-18_18h-49m</date>
  </transacGrp>
  <langSec xml:lang="en">
    <termSec id="011-en-t0">
      <term>thrombosis</term>
    </termSec>
  </langSec>
</conceptEntry>

```

```

        </termSec>
      </langSec>
    <descripGrp>
      <xref xmlns:target_file="https://medpix.nlm.nih.gov/topic?
id=0936415f-49bc-4363-8578-45377fead698" type="xGraphic "/>
      <descrip type="description">Superior ophthalmic vein
thrombosis left eye. The SOV does not enhance and is
indistinguishable from the adjacent superior rectus muscle. (Note
patent SOV on right side.) Severe proptosis. Marked swelling of
extra-ocular muscles. Preseptal cellulitis. Abscess formation medial
canthus.</descrip>
    </descripGrp>
  </conceptEntry>
  <conceptEntry id="012">
    <type type="subjectField">disease</type>
    <transacGrp>
      <transac type="transactionType">origination</transac>
      <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
      <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
      <termSec id="012-en-t0">
        <term>orbital cellulitis</term>
      </termSec>
    </langSec>
    <descrip type="spatialRelation">020</descrip>
  </conceptEntry>
  <conceptEntry id="013">
    <type type="subjectField">body part</type>
    <transacGrp>
      <transac type="transactionType">origination</transac>
      <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
      <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
      <termSec id="013-en-t0">
        <term>orbital dermoid</term>
      </termSec>
    </langSec>
    <descrip type="spatialRelation">020</descrip>
  </conceptEntry>
  <descripGrp>
    <xref xmlns:target_file="https://medpix.nlm.nih.gov/topic?
id=82f95cbe-5447-48f9-ad8e-f678cabedf9e" type="xGraphic "/>
    <descrip type="description">Orbital Dermoid. Coronal T1w
fat saturation sequence demonstrating loss of signal within left
lacrima fossa mass.</descrip>
  </descripGrp>
</conceptEntry>
<conceptEntry id="014">
  <type type="subjectField">body part</type>
  <transacGrp>
    <transac type="transactionType">origination</transac>

```

```

        <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
        <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
        <termSec id="014-en-t0">
            <term>extraconal</term>
        </termSec>
    </langSec>
    <descrip type="spatialRelation">020</descrip>
</conceptEntry>
<conceptEntry id="015">
    <type type="subjectField">disease</type>
    <transacGrp>
        <transac type="transactionType">origination</transac>
        <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
        <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
        <termSec id="015-en-t0">
            <term>lacrimal gland lesion</term>
        </termSec>
    </langSec>
    <descrip type="spatialRelation">043</descrip>
    <descripGrp>
        <xref xmlns:target_file="https://medpix.nlm.nih.gov/topic?
id=28d9aa4d-89b8-4bf8-8324-fe2472ce677b" type="xGraphic "/>
        <descrip type="description">T2 hyperintense mass within
the left lacrimal fossa.</descrip>
    </descripGrp>
</conceptEntry>
<conceptEntry id="016">
    <type type="subjectField">disease</type>
    <transacGrp>
        <transac type="transactionType">origination</transac>
        <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
        <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
        <termSec id="016-en-t0">
            <term>Potts Puffy tumor</term>
        </termSec>
        <termSec id="016-en-t1">
            <term>Potts Puffy tumour</term>
        </termSec>
    </langSec>
    <descrip type="genericRelation">039</descrip>
    <descripGrp>
        <xref xmlns:target_file="https://medpix.nlm.nih.gov/topic?
id=5a37a92b-8194-46c0-a21b-3439ae5cf442" type="xGraphic "/>
        <descrip type="description">Potts Puffy Tumor. These axial
images show soft tissue filling the left sphenoid sinus and

```

```

proptosis of the globe.</descrip>
  </descripGrp>
</conceptEntry>
<conceptEntry id="017">
  <type type="subjectField">symptom</type>
  <transacGrp>
    <transac type="transactionType">origination</transac>
    <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
    <date>21-10-18_18h-49m</date>
  </transacGrp>
  <langSec xml:lang="en">
    <termSec id="017-en-t0">
      <term>sinusitis</term>
    </termSec>
  </langSec>
  <descrip type="genericRelation">044</descrip>
  <descrip type="symptomRelation">045</descrip>
  <descrip type="symptomRelation">046</descrip>
</conceptEntry>
<conceptEntry id="018">
  <type type="subjectField">symptom</type>
  <transacGrp>
    <transac type="transactionType">origination</transac>
    <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
    <date>21-10-18_18h-49m</date>
  </transacGrp>
  <langSec xml:lang="en">
    <termSec id="018-en-t0">
      <term>proptosis</term>
    </termSec>
  </langSec>
  <descripGrp>
    <xref xmlns:target_file="https://medpix.nlm.nih.gov/topic?
id=5a37a92b-8194-46c0-a21b-3439ae5cf442" type="xGraphic "/>
    <descrip type="description">Potts Puffy Tumor. These axial
images show soft tissue filling the left sphenoid sinus and
proptosis of the globe.</descrip>
  </descripGrp>
</conceptEntry>
<conceptEntry id="019">
  <type type="subjectField">disease</type>
  <transacGrp>
    <transac type="transactionType">origination</transac>
    <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
    <date>21-10-18_18h-49m</date>
  </transacGrp>
  <langSec xml:lang="en">
    <termSec id="019-en-t0">
      <term>CNS lymphoma</term>
    </termSec>
  </langSec>

```



```

    <descrip type="genericRelation">047</descrip>
    <descripGrp>
      <xref xmlns:target_file="https://medpix.nlm.nih.gov/topic?
id=7c1a0431-c9ab-40d2-a088-540a3c4003b6" type="xGraphic "/>
      <descrip type="description">CNS lymphoma, large B-cell
type, retina and brain. Bilateral frontal lobe lesions, producing a
"butterfly" pattern, with secondary vasogenic edema extending into
the external capsule bilaterally (arrows).</descrip>
    </descripGrp>
  </conceptEntry>
  <conceptEntry id="020">
    <type type="subjectField">body part</type>
    <transacGrp>
      <transac type="transactionType">origination</transac>
      <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
      <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
      <termSec id="020-en-t0">
        <term>orbit</term>
      </termSec>
    </langSec>
    <descrip type="spatialRelation">048</descrip>
  </conceptEntry>
  <conceptEntry id="021">
    <type type="subjectField">disease</type>
    <transacGrp>
      <transac type="transactionType">origination</transac>
      <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
      <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
      <termSec id="021-en-t0">
        <term>retinoblastoma</term>
      </termSec>
    </langSec>
    <descrip type="genericRelation">039</descrip>
    <descrip type="spatialRelation">049</descrip>
    <descripGrp>
      <xref xmlns:target_file="https://medpix.nlm.nih.gov/topic?
id=28a6e306-b1b3-4a5d-adfc-057b4ed12a10" type="xGraphic "/>
      <descrip type="description">Retinoblastoma left eye. Non-
calcified mass left eye Abnormal signal in vitreous chamber on T1
and T2 sequences Mass enhances following contrast administration No
lesion in right eye or pineal gland.</descrip>
    </descripGrp>
  </conceptEntry>
  <conceptEntry id="022">
    <type type="subjectField">body part</type>
    <transacGrp>
      <transac type="transactionType">origination</transac>
      <transacNote type="responsibility">https://

```

```

medpix.nlm.nih.gov/</transacNote>
  <date>21-10-18_18h-49m</date>
</transacGrp>
  <langSec xml:lang="en">
    <termSec id="022-en-t0">
      <term>intraocular</term>
    </termSec>
  </langSec>
</conceptEntry>
<conceptEntry id="023">
  <type type="subjectField">symptom</type>
  <transacGrp>
    <transac type="transactionType">origination</transac>
    <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
  <date>21-10-18_18h-49m</date>
</transacGrp>
  <langSec xml:lang="en">
    <termSec id="023-en-t0">
      <term>dyslapsia</term>
    </termSec>
  </langSec>
  <descrip type="spatialRelation">048</descrip>
  <descripGrp>
    <xref xmlns:target_file="https://medpix.nlm.nih.gov/topic?
id=00e2990e-c234-4903-8049-7b64d8daf602" type="xGraphic "/>
    <descrip type="description">Septo-optic Dyslapsia. The
optic chiasm is asymmetric and the septum pellucidum is absent.</
descrip>
  </descripGrp>
</conceptEntry>
<conceptEntry id="024">
  <type type="subjectField">disease</type>
  <transacGrp>
    <transac type="transactionType">origination</transac>
    <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
  <date>21-10-18_18h-49m</date>
</transacGrp>
  <langSec xml:lang="en">
    <termSec id="024-en-t0">
      <term>autoimmune thyrotoxicosis</term>
    </termSec>
    <termSec id="024-en-t1">
      <term>thyrotoxicosis proptosis</term>
    </termSec>
  </langSec>
  <descrip type="genericRelation">008</descrip>
</conceptEntry>
<conceptEntry id="025">
  <type type="subjectField">disease</type>
  <transacGrp>
    <transac type="transactionType">origination</transac>
    <transacNote type="responsibility">https://

```

```

medpix.nlm.nih.gov/</transacNote>
  <date>21-10-18_18h-49m</date>
</transacGrp>
  <langSec xml:lang="en">
    <termSec id="025-en-t0">
      <term>amaurosis</term>
    </termSec>
  </langSec>
  <descrip type="correlationRelation">050</descrip>
  <descrip type="correlationRelation">051</descrip>
</conceptEntry>
<conceptEntry id="026">
  <type type="subjectField">disease</type>
  <transacGrp>
    <transac type="transactionType">origination</transac>
    <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
  <date>21-10-18_18h-49m</date>
</transacGrp>
  <langSec xml:lang="en">
    <termSec id="026-en-t0">
      <term>arterial hypertension</term>
    </termSec>
    <termSec id="026-en-t1">
      <term>high blood pressure</term>
    </termSec>
    <termSec id="026-en-t2">
      <term>hypertensive disease</term>
    </termSec>
  </langSec>
</conceptEntry>
<conceptEntry id="027">
  <type type="subjectField">disease</type>
  <transacGrp>
    <transac type="transactionType">origination</transac>
    <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
  <date>21-10-18_18h-49m</date>
</transacGrp>
  <langSec xml:lang="en">
    <termSec id="027-en-t0">
      <term>diabetes</term>
    </termSec>
    <termSec id="027-en-t1">
      <term>diabetes mellitus</term>
    </termSec>
    <termSec id="027-en-t2">
      <term>sugar diabetes</term>
    </termSec>
  </langSec>
</conceptEntry>
<conceptEntry id="028">
  <type type="subjectField">symptom</type>
  <transacGrp>

```

```

        <transac type="transactionType">origination</transac>
        <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
        <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
        <termSec id="028-en-t0">
            <term>hypercholesterolemia</term>
        </termSec>
    </langSec>
</conceptEntry>
<conceptEntry id="029">
    <type type="subjectField">disease</type>
    <transacGrp>
        <transac type="transactionType">origination</transac>
        <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
        <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
        <termSec id="029-en-t0">
            <term>migraine</term>
        </termSec>
    </langSec>
</conceptEntry>
<conceptEntry id="030">
    <type type="subjectField">disease</type>
    <transacGrp>
        <transac type="transactionType">origination</transac>
        <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
        <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
        <termSec id="030-en-t0">
            <term>smoking</term>
        </termSec>
    </langSec>
</conceptEntry>
<conceptEntry id="031">
    <type type="subjectField">disease</type>
    <transacGrp>
        <transac type="transactionType">origination</transac>
        <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
        <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
        <termSec id="031-en-t0">
            <term>cyst</term>
        </termSec>
    </langSec>
</conceptEntry>
<conceptEntry id="032">

```

```

        <type type="subjectField">disease</type>
        <transacGrp>
          <transac type="transactionType">origination</transac>
          <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
          <date>21-10-18_18h-49m</date>
        </transacGrp>
        <langSec xml:lang="en">
          <termSec id="032-en-t0">
            <term>blepharitis</term>
          </termSec>
        </langSec>
      </conceptEntry>
      <conceptEntry id="033">
        <type type="subjectField">body part</type>
        <transacGrp>
          <transac type="transactionType">origination</transac>
          <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
          <date>21-10-18_18h-49m</date>
        </transacGrp>
        <langSec xml:lang="en">
          <termSec id="033-en-t0">
            <term>lid sebaceous gland</term>
          </termSec>
        </langSec>
        <descrip type="spatialRelation">052</descrip>
      </conceptEntry>
      <conceptEntry id="034">
        <type type="subjectField">symptom</type>
        <transacGrp>
          <transac type="transactionType">origination</transac>
          <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
          <date>21-10-18_18h-49m</date>
        </transacGrp>
        <langSec xml:lang="en">
          <termSec id="034-en-t0">
            <term>abscess</term>
          </termSec>
        </langSec>
      </conceptEntry>
      <conceptEntry id="035">
        <type type="subjectField">symptom</type>
        <transacGrp>
          <transac type="transactionType">origination</transac>
          <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
          <date>21-10-18_18h-49m</date>
        </transacGrp>
        <langSec xml:lang="en">
          <termSec id="035-en-t0">
            <term>pain</term>
          </termSec>
        </langSec>
      </conceptEntry>

```

```

        </langSec>
      </conceptEntry>
      <conceptEntry id="036">
        <type type="subjectField">symptom</type>
        <transacGrp>
          <transac type="transactionType">origination</transac>
          <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
          <date>21-10-18_18h-49m</date>
        </transacGrp>
        <langSec xml:lang="en">
          <termSec id="036-en-t0">
            <term>pus</term>
          </termSec>
          <termSec id="036-en-t1">
            <term>suppuration</term>
          </termSec>
        </langSec>
      </conceptEntry>
      <conceptEntry id="037">
        <type type="subjectField">symptom</type>
        <transacGrp>
          <transac type="transactionType">origination</transac>
          <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
          <date>21-10-18_18h-49m</date>
        </transacGrp>
        <langSec xml:lang="en">
          <termSec id="037-en-t0">
            <term>sensation of foreign body in the eye</term>
          </termSec>
        </langSec>
      </conceptEntry>
      <conceptEntry id="038">
        <type type="subjectField">body part</type>
        <transacGrp>
          <transac type="transactionType">origination</transac>
          <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
          <date>21-10-18_18h-49m</date>
        </transacGrp>
        <langSec xml:lang="en">
          <termSec id="038-en-t0">
            <term>nose</term>
          </termSec>
        </langSec>
      </conceptEntry>
      <conceptEntry id="039">
        <type type="subjectField">disease</type>
        <transacGrp>
          <transac type="transactionType">origination</transac>
          <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
          <date>21-10-18_18h-49m</date>

```

```

</transacGrp>
<langSec xml:lang="en">
  <termSec id="039-en-t0">
    <term>tumor</term>
  </termSec>
  <termSec id="039-en-t1">
    <term>tumour</term>
  </termSec>
  <termSec id="039-en-t2">
    <term>neoplasm</term>
  </termSec>
</langSec>
<descrip type="diagnosisRelation">053</descrip>
</conceptEntry>
<conceptEntry id="040">
  <type type="subjectField">symptom</type>
  <transacGrp>
    <transac type="transactionType">origination</transac>
    <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
    <date>21-10-18_18h-49m</date>
  </transacGrp>
  <langSec xml:lang="en">
    <termSec id="040-en-t0">
      <term>palpitations</term>
    </termSec>
  </langSec>
</conceptEntry>
<conceptEntry id="041">
  <type type="subjectField">symptom</type>
  <transacGrp>
    <transac type="transactionType">origination</transac>
    <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
    <date>21-10-18_18h-49m</date>
  </transacGrp>
  <langSec xml:lang="en">
    <termSec id="041-en-t0">
      <term>weight loss</term>
    </termSec>
  </langSec>
</conceptEntry>
<conceptEntry id="042">
  <type type="subjectField">symptom</type>
  <transacGrp>
    <transac type="transactionType">origination</transac>
    <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
    <date>21-10-18_18h-49m</date>
  </transacGrp>
  <langSec xml:lang="en">
    <termSec id="042-en-t0">
      <term>deep vein thrombosis</term>
    </termSec>
  </langSec>
</conceptEntry>

```

```

        </langSec>
        <descrip type="genericRelation">058</descrip>
    </conceptEntry>
    <conceptEntry id="043">
        <type type="subjectField">body part</type>
        <transacGrp>
            <transac type="transactionType">origination</transac>
            <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
            <date>21-10-18_18h-49m</date>
        </transacGrp>
        <langSec xml:lang="en">
            <termSec id="043-en-t0">
                <term>lacrimal gland</term>
            </termSec>
        </langSec>
        <descrip type="partiativeRelation">049</descrip>
    </conceptEntry>
    <conceptEntry id="044">
        <type type="subjectField">symptom</type>
        <transacGrp>
            <transac type="transactionType">origination</transac>
            <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
            <date>21-10-18_18h-49m</date>
        </transacGrp>
        <langSec xml:lang="en">
            <termSec id="044-en-t0">
                <term>infection of the upper respiratory tract</term>
            </termSec>
            <termSec id="044-en-t1">
                <term>upper respiratory infection</term>
            </termSec>
            <termSec id="044-en-t2">
                <term>upper respiratory tract infection</term>
            </termSec>
            <termSec id="044-en-t3">
                <term>URI</term>
            </termSec>
            <termSec id="044-en-t4">
                <term>URTI</term>
            </termSec>
        </langSec>
    </conceptEntry>
    <conceptEntry id="045">
        <type type="subjectField">symptom</type>
        <transacGrp>
            <transac type="transactionType">origination</transac>
            <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
            <date>21-10-18_18h-49m</date>
        </transacGrp>
        <langSec xml:lang="en">
            <termSec id="045-en-t0">

```



```

        <term>headache</term>
      </termSec>
    </langSec>
    <descrip type="genericRelation">035</descrip>
  </conceptEntry>
  <conceptEntry id="046">
    <type type="subjectField">symptom</type>
    <transacGrp>
      <transac type="transactionType">origination</transac>
      <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
      <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
      <termSec id="046-en-t0">
        <term>nasal discharge</term>
      </termSec>
      <termSec id="046-en-t1">
        <term>rhinorrhea</term>
      </termSec>
      <termSec id="046-en-t2">
        <term>runny nose</term>
      </termSec>
    </langSec>
  </conceptEntry>
  <conceptEntry id="047">
    <type type="subjectField">disease</type>
    <transacGrp>
      <transac type="transactionType">origination</transac>
      <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
      <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
      <termSec id="047-en-t0">
        <term>lymphoma</term>
      </termSec>
    </langSec>
    <descrip type="genericRelation">039</descrip>
    <descrip type="symptomRelation">051</descrip>
  </conceptEntry>
  <conceptEntry id="048">
    <type type="subjectField">body part</type>
    <transacGrp>
      <transac type="transactionType">origination</transac>
      <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
      <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
      <termSec id="048-en-t0">
        <term>eye tissue</term>
      </termSec>
    </langSec>
  </conceptEntry>

```

```

    <descrip type="partiativeRelation">049</descrip>
  </conceptEntry>
  <conceptEntry id="049">
    <type type="subjectField">body part</type>
    <transacGrp>
      <transac type="transactionType">origination</transac>
      <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
      <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
      <termSec id="049-en-t0">
        <term>eye</term>
      </termSec>
    </langSec>
  </conceptEntry>
  <conceptEntry id="050">
    <type type="subjectField">disease</type>
    <transacGrp>
      <transac type="transactionType">origination</transac>
      <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
      <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
      <termSec id="050-en-t0">
        <term>neurofibromatosis</term>
      </termSec>
    </langSec>
  </conceptEntry>
  <conceptEntry id="051">
    <type type="subjectField">disease</type>
    <transacGrp>
      <transac type="transactionType">origination</transac>
      <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
      <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
      <termSec id="051-en-t0">
        <term>onchocerciasis</term>
      </termSec>
      <termSec id="051-en-t1">
        <term>river blindness</term>
      </termSec>
      <termSec id="051-en-t2">
        <term>Robles disease</term>
      </termSec>
    </langSec>
  </conceptEntry>
  <conceptEntry id="052">
    <type type="subjectField">body part</type>
    <transacGrp>
      <transac type="transactionType">origination</transac>

```

```

        <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
        <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
        <termSec id="052-en-t0">
            <term>lid</term>
        </termSec>
    </langSec>
    <descrip type="partiativeRelation">049</descrip>
</conceptEntry>
<conceptEntry id="053">
    <type type="subjectField">--</type>
    <transacGrp>
        <transac type="transactionType">origination</transac>
        <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
        <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
        <termSec id="053-en-t0">
            <term>MRI</term>
        </termSec>
    </langSec>
</conceptEntry>
<conceptEntry id="051">
    <type type="subjectField">symptom</type>
    <transacGrp>
        <transac type="transactionType">origination</transac>
        <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
        <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">
        <termSec id="051-en-t0">
            <term>adenopathy</term>
        </termSec>
        <termSec id="051-en-t1">
            <term>enlarged lymph node</term>
        </termSec>
        <termSec id="051-en-t2">
            <term>lymphadenopathy</term>
        </termSec>
    </langSec>
</conceptEntry>
<conceptEntry id="052">
    <type type="subjectField">symptom</type>
    <transacGrp>
        <transac type="transactionType">origination</transac>
        <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
        <date>21-10-18_18h-49m</date>
    </transacGrp>
    <langSec xml:lang="en">

```

```

    <termSec id="052-en-t0">
      <term>enlarged spleen</term>
    </termSec>
    <termSec id="052-en-t1">
      <term>enlargement of the spleen</term>
    </termSec>
    <termSec id="052-en-t2">
      <term>splenomegaly</term>
    </termSec>
  </langSec>
  <descrip type="casualRelation">047</descrip>
</conceptEntry>
<conceptEntry id="053">
  <type type="subjectField">disease</type>
  <transacGrp>
    <transac type="transactionType">origination</transac>
    <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
    <date>21-10-18_18h-49m</date>
  </transacGrp>
  <langSec xml:lang="en">
    <termSec id="053-en-t0">
      <term>high intraocular pressure</term>
    </termSec>
    <termSec id="053-en-t1">
      <term>high eye pressure</term>
    </termSec>
    <termSec id="053-en-t2">
      <term>glaucoma suspect</term>
    </termSec>
    <termSec id="053-en-t3">
      <term>ocular hypertension</term>
    </termSec>
    <termSec id="053-en-t4">
      <term>high IOP</term>
    </termSec>
    <termSec id="053-en-t5">
      <term>OHT</term>
    </termSec>
  </langSec>
  <descrip type="symptomRelation">054</descrip>
  <descrip type="spatialRelation">049</descrip>
</conceptEntry>
<conceptEntry id="054">
  <type type="subjectField">symptom</type>
  <transacGrp>
    <transac type="transactionType">origination</transac>
    <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
    <date>21-10-18_18h-49m</date>
  </transacGrp>
  <langSec xml:lang="en">
    <termSec id="054-en-t0">
      <term>vision loss</term>
    </termSec>
  </langSec>

```

```

        </termSec>
      </langSec>
    </conceptEntry>
    <conceptEntry id="055">
      <type type="subjectField">disease</type>
      <transacGrp>
        <transac type="transactionType">origination</transac>
        <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
        <date>21-10-18_18h-49m</date>
      </transacGrp>
      <langSec xml:lang="en">
        <termSec id="055-en-t0">
          <term>glaucoma</term>
        </termSec>
      </langSec>
      <descrip type="casualRelation">025</descrip>
      <descrip type="casualRelation">053</descrip>
      <descrip type="spatialRelation">056</descrip>
    </conceptEntry>
    <conceptEntry id="056">
      <type type="subjectField">body part</type>
      <transacGrp>
        <transac type="transactionType">origination</transac>
        <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
        <date>21-10-18_18h-49m</date>
      </transacGrp>
      <langSec xml:lang="en">
        <termSec id="056-en-t0">
          <term>optic nerve</term>
        </termSec>
      </langSec>
      <descrip type="spatialRelation">020</descrip>
    </conceptEntry>
    <conceptEntry id="057">
      <type type="subjectField">disease</type>
      <transacGrp>
        <transac type="transactionType">origination</transac>
        <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
        <date>21-10-18_18h-49m</date>
      </transacGrp>
      <langSec xml:lang="en">
        <termSec id="057-en-t0">
          <term>arrythmia</term>
        </termSec>
        <termSec id="057-en-t1">
          <term>cardiac dysrhythmia</term>
        </termSec>
        <termSec id="057-en-t2">
          <term>irregular heartbeat</term>
        </termSec>
      </langSec>
    </conceptEntry>
  </langSec>

```

```

        <descrip type="casualRelation">008</descrip>
    </conceptEntry>
    <conceptEntry id="058">
        <type type="subjectField">symptom</type>
        <transacGrp>
            <transac type="transactionType">origination</transac>
            <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
            <date>21-10-18_18h-49m</date>
        </transacGrp>
        <langSec xml:lang="en">
            <termSec id="058-en-t0">
                <term>thrombosis</term>
            </termSec>
        </langSec>
    </conceptEntry>
    <conceptEntry id="059">
        <type type="subjectField">disease</type>
        <transacGrp>
            <transac type="transactionType">origination</transac>
            <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
            <date>21-10-18_18h-49m</date>
        </transacGrp>
        <langSec xml:lang="en">
            <termSec id="059-en-t0">
                <term>infarction</term>
            </termSec>
        </langSec>
        <descrip type="casualRelation">058</descrip>
    </conceptEntry>
    <conceptEntry id="060">
        <type type="subjectField">disease</type>
        <transacGrp>
            <transac type="transactionType">origination</transac>
            <transacNote type="responsibility">https://
medpix.nlm.nih.gov/</transacNote>
            <date>21-10-18_18h-49m</date>
        </transacGrp>
        <langSec xml:lang="en">
            <termSec id="060-en-t0">
                <term>cerebral infarction</term>
            </termSec>
            <termSec id="060-en-t1">
                <term>ischemic stroke</term>
            </termSec>
        </langSec>
        <descrip type="casualRelation">058</descrip>
        <descrip type="genericRelation">059</descrip>
    </conceptEntry>
</body>
</text>
</tbx>

```

Anhang 2: Abstract (Deutsch)

Die vorliegende Masterarbeit hat den Anspruch, mit Hilfe der Wissensmodellierung, multimodale (Text-Bild) medizinische terminologische Ressourcen, insbesondere im Fachgebiet der Ophthalmologie, zu modellieren und ein terminologisches Begriffssystem zu erstellen. Im ersten theoretischen Teil der Arbeit werden die Grundlagen der Terminologie, der verschiedenen Arten von semantischen Netzen und der Wissensrepräsentation präsentiert. Im zweiten empirischen Teil wird die Datenbankanalyse durchgeführt: Mit Hilfe von MedPix®, einer multimodalen Datenbank, konnten zunächst Begriffe extrahiert und anschließend Beziehungen zwischen ihnen und den entsprechenden Bildern hergestellt werden, mit dem Ziel, ein terminologisches Begriffssystem zu erstellen. Das aus dieser Forschungsarbeit resultierende terminologische Begriffssystem ist in Form einer Concept Map mit Ellipsen und Beziehungen dargestellt.

Anhang 3: Abstract (English)

This Master's thesis aims at using Knowledge Engineering to model multimodal (text-image) medical terminological resources, especially in the field of ophthalmology, and to create a terminological concept system. In the first theoretical part of the thesis, the basics of terminology, different types of semantic networks and knowledge representation methods are presented. In the second empirical part, the database analysis is carried out. With the help of MedPix®, a multimodal database, terms could first be extracted and then relations between them and the corresponding images could be established, with the aim of creating a terminological term system. The resulting terminological concept system is displayed as a concept map with ellipses and relations.