

Begriffssysteme für die medizinische Dokumentation

Pedersen, S., Hasselbring, W.

Universität Oldenburg, Fachbereich Informatik, Abt. Software Engineering,

Postfach 2503, 26111 Oldenburg

<http://se.informatik.uni-oldenburg.de>

Zusammenfassung

Die uneinheitliche Verwendung von Begriffen sorgt in vielen Bereichen oft für Unklarheiten [1]. So werden häufig die gleichen Bezeichnungen für verschiedene Begriffe verwendet (Homonyme) und umgekehrt (Synonyme). Dadurch entstehen Probleme mit der verwendeten Terminologie für Diagnosen und Maßnahmen. Standardisierung von Nomenklaturen und Klassifikationen sind wichtige Hilfsmittel, um zu einer einheitlichen Terminologie in der Medizin zu kommen. Dazu existieren viele Standardisierungsinitiativen, wie z.B. das Unified Medical Language System (UMLS) [7][9].

Die standardisierte, objektorientierte Modellierungssprache Unified Modeling Language (UML) ist dabei ein gutes Mittel, um die Zusammenhänge zwischen den einzelnen Elementen von Nomenklaturen und Klassifikationen zu visualisieren [6]. In diesem Übersichtsvortrag werden wir die grundlegenden Unterschiede und Gemeinsamkeiten von aktuellen Standardisierungsinitiativen für Nomenklaturen und Klassifikationen aufzeigen. Ein typischer Anwendungsbereich für den Einsatz von Nomenklaturen und Klassifikationen ist die medizinische Dokumentation.

Grundelemente der Terminologielehre sind der Gegenstand, der Begriff und die Bezeichnung. Nach der semiotischen Triade gelten die Beziehungen: Eine Bezeichnung bezeichnet einen Begriff, eine Bezeichnung bedeutet einen Gegenstand und ein Begriff bezieht sich auf einen Gegenstand [2]. Bestimmte Begriffe fungieren als Merkmale und haben einen entscheidenden Einfluss auf die Struktur in Begriffssystemen. Begriffsrelationen, die genutzt werden, um Begriffe zu ordnen sind meist Spezialisierung/Generalisierung sowie die Teil-Ganzes-Beziehung (Komposition). Neben hierarchischen Begriffsbeziehungen gibt es auch nicht-hierarchische, zu denen sequentielle und pragmatische Beziehungen zählen. Eine Struktur der Begriffe nennt man Klassifikation. Nomenklaturen dagegen sind Zusammenstellungen von Bezeichnungen, die dahinterstehende Systematik basiert in der Regel auf einer Begriffsordnung. Daher können auch Nomenklaturen als Begriffssysteme angesehen werden.

Schlüsselwörter

Medizinische Dokumentation, Terminologie, Klassifikation, medizinische Begriffssysteme

1 Einleitung

Die uneinheitliche Verwendung von Begriffen sorgt in vielen Bereichen oft für Unklarheiten. So werden häufig die gleichen Bezeichnungen für verschiedene Begriffe verwendet, sogenannte Homonyme wie z.B. die Bezeichnung „Achse“, die zum einen ein Fahrzeugteil bezeichnet zum anderen eine Bedeutungsdimension eines Begriffssystems. Im umgekehrten Fall dazu, wenn also verschiedene Bezeichnungen für den gleichen Begriff verwendet werden, nennt man die Bezeichnungen Synonyme. Ein Beispiel wäre „Zunge“ und „Glossa“ als Bezeichnungen für den quergestreiften Muskel in der Mundhöhle, der Sprechen und anderes mehr ermöglicht. Das Problem der Homonyme ist also, daß sich deren Bedeutung aus dem Kontext ergibt. Ohne Kontext wie bei der Dokumentation ergeben sich zwangsläufig Missverständnisse. Eine Recherche liefert zu viele und nicht-relevante Dokumentationseinheiten [4]. Bei Synonymen verhält es sich umgekehrt, eine Recherche liefert zu wenig relevante Dokumentationseinheiten, da verschiedene Deskriptoren für Indexierung und Recherche verwendet werden [4]. Durch Homonyme und Synonyme entstehen damit Probleme mit der verwendeten Terminologie für Diagnosen und Maßnahmen. Auch der Austausch von Dokumenten zwischen verschiedenen Einrichtungen des Gesundheitswesens ist davon betroffen. Standardisierung von Nomenklaturen und Klassifikationen sind wichtige Hilfsmittel, um zu einer einheitlichen Terminologie in der Medizin zu kommen. Dazu existieren viele Standardisierungsinitiativen, wie z.B. das Unified Medical Language System (UMLS) [9].

Die standardisierte, objektorientierte Modellierungssprache Unified Modeling Language (UML) ist ein gutes Mittel, um die Zusammenhänge zwischen den einzelnen Elementen von Nomenklaturen und Klassifikationen zu visualisieren.

2 Medizinische Begriffssysteme

Grundelemente der Terminologielehre sind der Gegenstand, der Begriff und die Bezeichnung. Nach der semiotischen Triade gelten die Beziehungen: Eine Bezeichnung bezeichnet einen Begriff, eine Bezeichnung bedeutet einen Gegenstand und ein Begriff bezieht sich auf einen Gegenstand. Ein Gegenstand ist ein beliebiger Ausschnitt aus der wahrnehmbaren oder vorstellbaren Welt [2]. Sogenannte individuelle Gegenstände sind einmalig, allgemeine Gegenstände dagegen ohne Bezug zu Raum und Zeit. Entsprechend beziehen sich Individualbegriffe auf individuelle Gegenstände, z.B. Epidemiologisches Krebsregister Niedersachsen (EKN) und Allgemeinbegriffe auf allgemeine Gegenstände, z.B. Krebsregister. Abbildung 1 zeigt als UML-Diagramm eben genannte Zusammenhänge sowie weitere wichtige Beziehungen bei medizinischen Begriffssystemen.

Bestimmte Begriffe fungieren nun als Merkmale und haben einen entscheidenden Einfluss auf die Struktur in Begriffssystemen. Begriffsrelationen, die genutzt werden, um Begriffe zu ordnen sind meist Spezialisierung/Generalisierung sowie die Teil-Ganzes-Beziehung (Komposition). Neben hierarchischen Begriffsbeziehungen gibt es auch nicht-hierarchische, zu denen sequentielle und pragmatische Beziehungen zählen. Dazu später mehr. Eine Struktur der Begriffe nennt man Klassifikation. Nomenklaturen dagegen sind Zusammenstellungen von Bezeichnungen, die dahinterstehende Systematik basiert in der Regel auf einer Begriffsordnung. Daher können auch Nomenklaturen als Begriffssysteme angesehen werden.

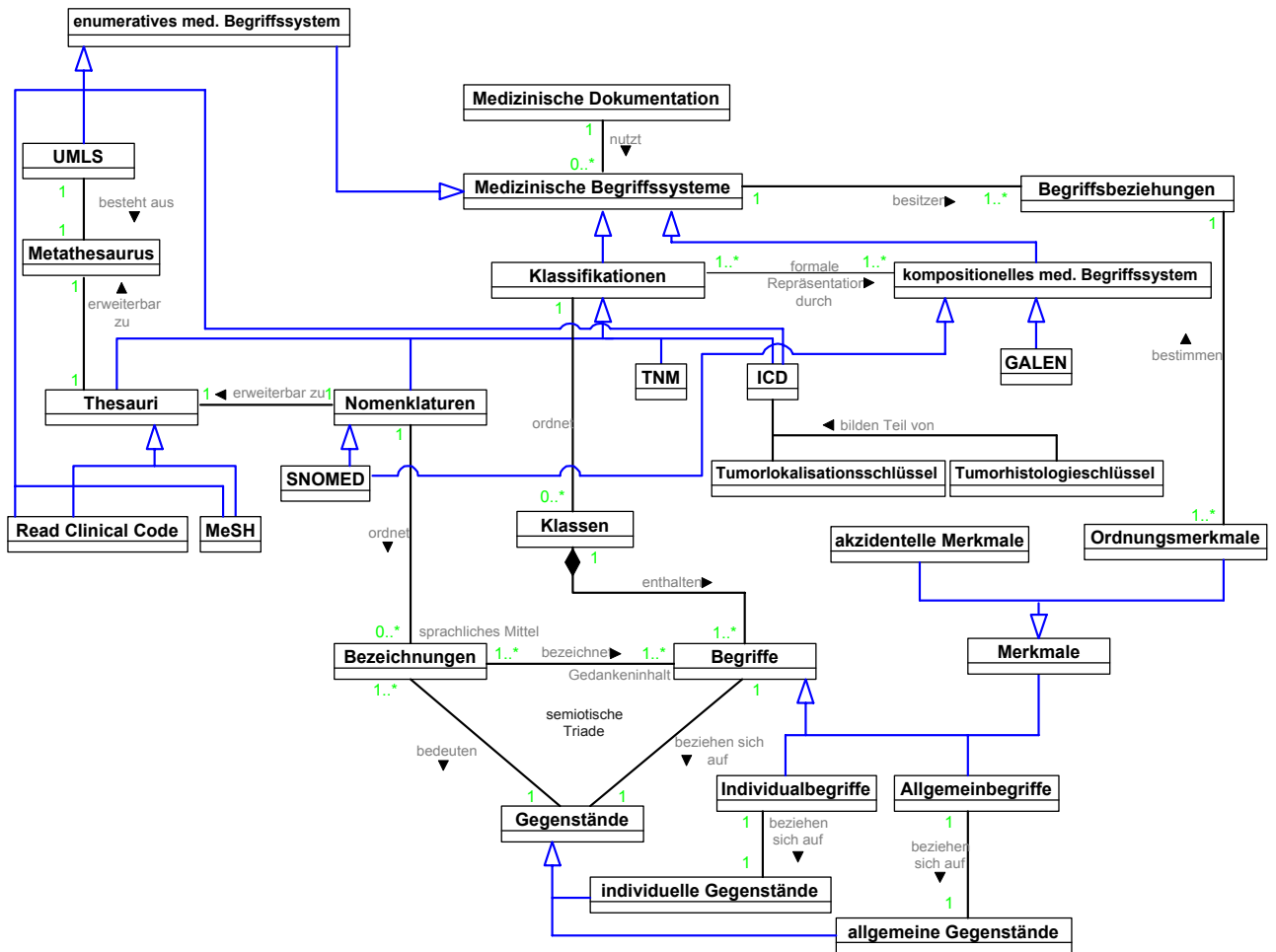


Abb. 1: Ausschnitt medizinische Begriffssysteme als UML-Diagramm

Mittels Definitionen, Angaben zur Bezeichnungsproblematik wie Homonymie und Synonymie, sprachliche Varianten und Wortformen kann eine Nomenklatur zu einem Thesaurus erweitert werden. Wird wiederum der Thesaurus erweitert um semantische und linguistische Informationen, so entsteht ein Metathesaurus, wie ihn der Kern von UMLS darstellt. UMLS gehört zu den enumerativen Begriffssystemen und versucht die „wichtigsten medizinischen Begriffssysteme in einem einzigen zu integrieren und dabei möglichst viele Begriffsbeziehungen explizit darzustellen“. Dazu abgegrenzt werden kompositionelle Begriffssysteme, wie z.B. die Generalized Architecture for Languages, Encyclopaedias and Nomenclatures in Medicine (GALEN) [2][9]. GALEN versucht im Gegensatz zu UMLS methodische Grundlagen für die formale Repräsentation medizinischer Begriffssysteme zu entwickeln [2]. GALEN erlaubt die formale Konstruktion von allen anderen fachrelevanten Begriffen durch die Komposition einer Anzahl primitiver Begriffe.

3 Begriffsbeziehungen und Codierprinzipien

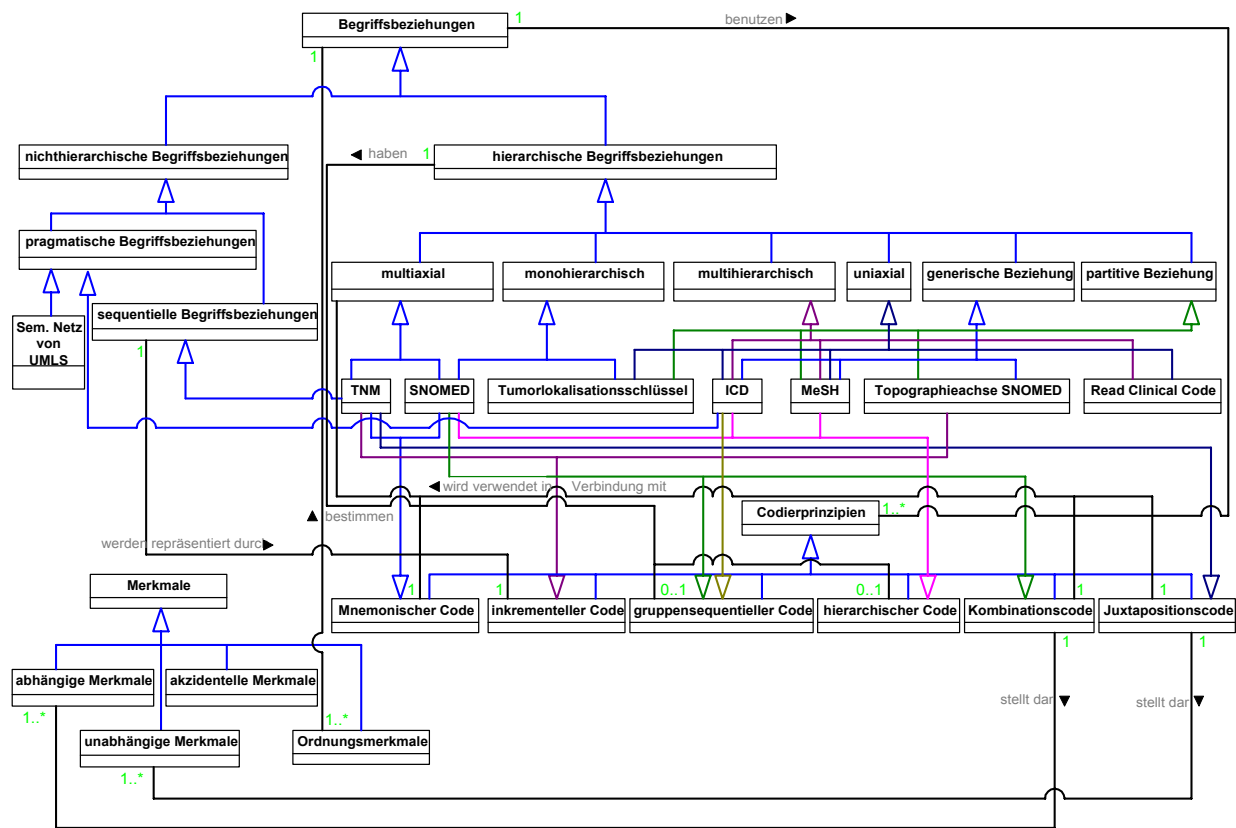


Abb. 2: Ausschnitt Begriffsbeziehungen und Codierprinzipien als UML-Diagramm

Abbildung 2 zeigt die möglichen Begriffsbeziehungen und die eingesetzten Codierprinzipien in medizinischen Begriffssystemen. Ferner sind in das Diagramm wichtige Begriffssysteme bezüglich ihrer Beziehungen und Codierprinzipien eingeordnet.

Wie bereits erwähnt unterscheidet man in medizinischen Begriffssystemen nicht-hierarchische und hierarchische Begriffsbeziehungen. Zu ersteren zählen pragmatische und sequentielle Begriffsbeziehungen, zu letzteren generische (Spezialisierung/Generalisierung) und partitive Beziehungen (Teil-Ganzes-Beziehungen). Ferner unterscheidet man bei hierarchischen Beziehungen multi- und uniaxiale Systeme sowie multi- und monohierarchische Begriffssysteme.

Bestimmt werden die Begriffsbeziehungen durch sogenannte Ordnungsmerkmale. Sie sind per definitionem diejenigen Merkmale, die in einem spezifischen Vorgang der Begriffsanalyse, zur Definition oder Festlegung von Begriffsbeziehungen, d.h. also zur Ordnung von Begriffen herangezogen werden [2]. Übrige Merkmale sind dann akzidentelle Merkmale, d.h. sie üben keinen Einfluss auf die Struktur des Begriffssystems aus. Ausserdem gilt: Ordnungsmerkmale sind Merkmale, die ein Begriff haben muss, akzidentielle Merkmale sind Merkmale, die ein Begriff haben kann. Merkmale sind voneinander abhängig, wenn das übergeordnete Merkmal gegeben sein muss, ehe das spezielle hinzutritt [3]. Die Struktur von Begriffssystemen wird also auch von der Abhängigkeit bzw. Unabhängigkeit der Merkmale tangiert. Abhängige Merkmale erzwingen eine bestimmte Hierarchie der Begriffe, unabhängige Merkmale werden oft auf verschiedenen Achsen dargestellt, da sie beliebig miteinander kombiniert

werden können. Der Kombinations- und der Juxtapositionscode sind beide zusammengesetzte Codes. Die einzelnen Segmente des Kombinationscodes stellen dabei voneinander abhängige Merkmale, die Segmente des Juxtapositionscodes voneinander unabhängige Merkmale dar.

Begriffsbeziehungen sind mit einem Codierschema verbunden, bei der Kodierung werden eindeutige Kürzel jedem Mitglied eines Begriffssystems zugeordnet. Die Beziehungen werden durch entsprechende Codes repräsentiert. In Begriffssystemen werden häufig mehrere Codierprinzipien miteinander kombiniert. Die wichtigsten Codierprinzipien sind mnemonischer Code, inkrementeller Code, gruppensequentieller Code, hierarchischer Code, Kombinations- und Juxtapositionscode. Hierarchische Begriffsbeziehungen werden mit gruppensequentiellem Code oder hierarchischem Code codiert, der Kombinationscode und der Juxtapositionscode werden verwendet in Verbindung mit multiaxialen Begriffssystemen. Sequentielle Begriffsbeziehungen werden durch inkrementelle Codes repräsentiert.

Die International Classification of Diseases (ICD) gehört zu den Klassifikationen und enumerativen Begriffssystemen [2][9]. Der ICD ist uniaxial, multihierarchisch mit generischen und pragmatischen Beziehungen. Seine Codierprinzipien sind der gruppensequentielle und der hierarchische Code. Speziell für die Onkologie gibt es den ICD-O, welcher sich in den Tumorlokalisations- und Tumorhistologieschlüssel aufteilt. Der Tumorlokalisationsschlüssel als topographischer Teil des ICD-O ist selbst wieder eine uniaxiale, monohierarchische Klassifikation mit partitiven Begriffsbeziehungen.

Die TNM-Klassifikation gehört ebenfalls zu den Klassifikationen, zusammen mit dem Tumorlokalisationsschlüssel gehört sie zu den kompositionellen medizinischen Begriffssystemen [2]. Die TNM-Klassifikation ist multiaxial mit sequentiellen Begriffsbeziehungen. Codierprinzipien sind der mnemonische, der inkrementelle und der Juxtapositionscode.

Die Systematized Nomenclature of Human and Veterinary Medicine (SNOMED) gehört zu den Nomenklaturen und wird zu den kompositionellen Begriffssystemen gezählt [2][9]. SNOMED ist multiaxial, monohierarchisch mit generischen und partitiven Begriffsbeziehungen. Der mnemonische, hierarchische, gruppensequentielle, inkrementelle und Kombinationscode bilden die Codierprinzipien für SNOMED. Problematisch ist, daß bei der beliebigen Komposition von Begriffen der einzelnen Achsen Ausdrücke entstehen können, die mehrdeutig sind, weil die Merkmale voneinander abhängig sind (siehe auch Kombinationscode).

Der Read Clinical Code (RCC) ist ein Thesaurus und gehört zu den enumerativen Begriffssystemen. Der RCC ist uniaxial und multihierarchisch aufgebaut [2]. Seit 1990 wird der RCC weiterentwickelt und gepflegt vom British National Health Service Centre for Coding and Classification (NHS CCC) [8].

Medical Subject Headings (MeSH) ist ein Thesaurus, der zur Erschließung der medizinischen Fachliteratur dient und dem INDEX MEDICUS und MEDLINE zugrundeliegt. Der INDEX MEDICUS ist die Druckversion, MEDLINE die Online-Version des weltweit größten Literaturdokumentationssystems der Medizin. MeSH, INDEX MEDICUS und MEDLINE werden von der National Library of Medicine (NLM), USA, erstellt und gepflegt [3][9].

4 Zusammenfassung und Ausblick

Medizinische Begriffssysteme spielen bei der medizinischen Dokumentation eine zentrale Rolle. Mit Hilfe der Unified Modeling Language (UML) ist es gelungen ein systematisches Begriffssystem für die Begriffssysteme zu modellieren.

Epidemiologische Krebsregister, wie zum Beispiel das Epidemiologische Krebsregister Niedersachsen (EKN) müssen eine flächendeckende Datenerfassung garantieren, um qualitativ gute Auswertungen leisten zu können. Die Melder besitzen eine Heterogenität in ihrer medizinischen Dokumentation, die es zu überwinden gilt. Medizinische Begriffssysteme sind ein wichtiges Mittel dazu, das wir in diesem Kontext zur Überwindung der Heterogenität zwischen verschiedenartigen Meldesystemen nutzen werden.

Literatur

- [1] HASSELBRING, W. (1999): On Defining Computer Science Terminology. Communications of the ACM, February 1999/Vol. 42, No. 2, S. 88-91.
- [2] BERNAUER, J. (1995): Zur semantischen Rekonstruktion medizinischer Begriffssysteme, Habilitationsschrift Universität Hildesheim.
- [3] BERNAUER, J. (2000): Dokumentations- und Ordnungslehre, Vorlesungsschrift Fachhochschule Ulm.
- [4] GAUS, W. (2000): Dokumentations- und Ordnungslehre. Springer-Verlag Berlin.
- [5] LEINER, F., GAUS W., HAUX R., KNAUP-GREGORI, P.: Medizinische Dokumentation: Lehrbuch und Leitfaden für die Praxis, 3. neubearbeitete und erweiterte Auflage. Stuttgart 1999. Schattauer Verlagsgesellschaft mbH.
- [6] OESTERREICH, B. (1999): Objektorientierte Softwareentwicklung; Analyse und Design mit der Unified Modeling Language, 4. aktualisierte Auflage. München 1999. R. Oldenbourg Verlag.
- [7] <http://www.nlm.nih.gov/pubs/factsheets/umls.html>
- [8] <http://www.ualberta.ca/dept/slis/cais/toews.htm>
- [9] INGENERF, J., REINER, J., SEIK, B.: Standardized Terminological Services enabling Semantic Interoperability between Distributed and Heterogeneous Systems. MIE 2000, August 2000