Auf dem Weg zum Web 3.0: Taxonomien und Ontologien für die medizinische Ausbildung - eine systematische Literaturrecherche

Wolf E. Blaum[1,2]
Anne Jarczweski[2]
Felix Balzer[1,2]
Philip Stötzner[2]
Olaf Ahlers[1,2]

1 Charité - Universitätsmedizin Berlin, Campus Charité Mitte und Campus Virchow-Kliniken, Klinik für Anästhesiologie mit Schwerpunkt operative Intensivmedizin, Berlin, Deutschland

2 Charité - Universitätsmedizin Berlin, Abteilung für Curriculumsorganisation, Lernzentrum, Berlin, Deutschland

Zusammenfassung

Einleitung: Sowohl für die Curriculumsentwicklung und –kartierung als auch zur Orientierung innerhalb der seit Jahren zunehmenden Fülle von Lernressourcen in der medizinischen Ausbildung stellt das semantische Web („Web 3.0“) ein niedrigschwelliges, effektives Hilfsmittel dar, das es ermöglicht, inhaltlich verwandte Elemente über Systemgrenzen hinweg zu identifizieren. Voraussetzung dafür ist der Einsatz eines geeigneten strukturierten Vokabulars zur maschinenlesbaren, inhaltlichen Beschreibung von Objekten, um die bisher notwendige manuelle Verknüpfung durch eine automatisch erzeugte, inhaltsbasierte bzw. semantische Verknüpfung zu ersetzen.
Ziel dieser Arbeit ist es, existierende Taxonomien und Ontologien zur Annotation medizinischer Lernressourcen und -inhalte zusammenzustellen und anhand ausgewählter Kriterien zu vergleichen sowie auf ihre Eignung zum Einsatz im eingangs beschriebenen Kontext zu überprüfen.
Methoden: Anhand einer systematischen Literaturrecherche wurden existierende Taxonomien und Ontologien zur Beschreibung medizinischer Lernressourcen identifiziert. Für jedes so identifizierte strukturierte Vokabular wurden mittels Websuche und/oder Kontakt zu den Herausgebern Thema, Struktur, Sprache, Umfang, Wartung und Technik der Taxonomie/Ontologie ermittelt und deren Eignung für den Einsatz im semantischen Web überprüft.
Ergebnisse: In 20 identifizierten Publikationen wurden 14 strukturierte Vokabulare identifiziert, die sich in Sprache, Umfang, Aktualität und Wartung zum Teil sehr stark unterschieden.
Keines der identifizierten Vokabulare erfüllte die erforderlichen Kriterien zur inhaltlichen Beschreibung medizinischer Ausbildungsinhalte und Lernressourcen im deutschsprachigen Raum.
Diskussion: Auf dem Weg zum Web 3.0 stellen Auswahl und Einsatz eines geeigneten, deutschsprachigen Vokabulars zur maschinenlesbaren, inhaltlichen Beschreibung von Objekten ein relevantes Problem dar. Mögliche Lösungsansätze umfassen die Neuentwicklung, die Übersetzung und/ oder die Kombination vorhandener Vokabulare, ggf. mit einer teilweisen Übersetzung englischsprachiger Vokabulare.

Schlüsselwörter: medizinische Ausbildung, semantisches Web, Web 3.0, Taxonomie, Ontologie, curriculares Mapping, Curriculumskartierung, e-learning, neue Medien

Autorenschaft

Die Autoren Blaum und Jarczewski haben zu gleichen Teilen beigetragen.

Einleitung

Die seit Jahren zunehmende Fülle von insbesondere online verfügbaren Lernressourcen in der medizinischen Ausbildung kann derzeitig kaum sinnvoll genutzt werden, weil es für Lehrende und Lernende an einem systematischen und einfachen Zugang zu deren Inhalten fehlt [1]. Die stetig steigende Zahl von verfügbaren Quellen stellt Studierende zunehmend vor die Schwierigkeit, thematisch passende, qualitativ hochwertige und didaktisch sinnvolle Angebote, welche zudem ihrem Ausbildungsniveau entsprechen, in einem Meer von Möglichkeiten zu identifizieren.
Eine beachtliche Zunahme von Ressourcen erlebt dabei insbesondere das computerbasierte Lernen. In einem kürzlich erschienenen Review zum Effekt von computer-

basierten virtuellen Patienten – einer Teilmenge der computerbasierten Ressourcen – haben Cook et al. allein zu diesem Thema 698 Publikationen identifiziert [2]. Gleichzeitig wird das PC-basierte Lernen mit der Immatrikulation der „Generation Y“ zunehmend wichtiger [3].
Lehrende und Planende stehen darüber hinaus vor der Herausforderung, die inhaltlichen Zusammenhänge in den von ihnen verantworteten Curricula darzustellen (zu kartieren) [4], was bisher eine manuelle Verknüpfung von curricularen Elementen voraussetzt.

Semantisches Web

Besonders die Entwicklung des Internets hin zum semantischen Web (Web 3.0) birgt ein erhebliches Potential zur inhaltlichen Verknüpfung von Lernressourcen ebenso wie zum Einsatz in der Curriculumskartierung [5], [6].
Während die Inhalte des Internets in seiner Entstehung im Wesentlichen bestimmt waren durch von wenigen Autoren bereit gestellte Informationen für ein breites, konsumierendes Publikum, ist das aktuelle „Web 2.0“ gekennzeichnet durch vom Nutzer erzeugte Inhalte [7]. Den Daten beider Versionen ist gemein, dass die darin enthaltenen Informationen ausschließlich von Menschen verstanden und interpretiert werden können. Jede inhaltliche Verknüpfung der Daten setzt demnach eine menschliche Aktion voraus.
Der Begriff „Web 3.0“ (auch „semantic Web“ genannt) bezeichnet ein Konzept zur Speicherung von Daten im Netz, welches durch Hinzufügen von maschinenlesbaren Metainformationen (z. B. mittels strukturierter Vokabulare) die Möglichkeit bietet, inhaltliche Verknüpfungen zwischen Datenobjekten von Maschinen automatisch generieren bzw. aufdecken zu lassen [8]. Voraussetzung dafür ist eine maschinell lesbare, inhaltliche Beschreibung der in den Daten enthaltenen Informationen [6], [9].
Eine geeignete Art der inhaltlichen Beschreibung ist der Einsatz strukturierter Vokabulare. Dabei wird jede Ressource mit definierten Begriffen eines strukturierten Vokabulars verschlagwortet.

Taxonomie und Ontologie

Taxonomien und Ontologien stellen formalisierte Vokabulare zur inhaltlichen Beschreibung (semantische Annotation) von Objekten, wie zum Beispiel Lernressourcen, dar [10]. Durch die darüber hinaus in Taxonomien hinterlegten Beziehungen der Vokabeln untereinander kann eine (manuell angestoßene) automatische Suche Lernressourcen mit inhaltlich ähnlichen Beschreibungen auch an voneinander unabhängigen Orten identifizieren [1], [11], [12].
Die Beziehungen der Vokabeln einer Taxonomie sind in der Regel monohierarchisch, jeder Vokabel sind also exakt eine (oder keine) übergeordnete Vokabel („ist ein“- oder „ist Teil von“ - Beziehung) und beliebig viele nachgeordnete Vokabeln zugeordnet. Ein klassisches Beispiel einer streng monohierarchischen Taxonomie ist die International Classification of Diseases http://www.dimdi.de/static/de/klassi/index.htm. Ontologien hingegen bilden auch polyhierarchische Strukturen mit der Zuordnung einer Vokabel zu mehr als einer übergeordneten Vokabel ab. Grundlegender Unterschied zu Taxonomien ist dabei, dass keine sprachlichen Konstrukte, sondern logisch definierte Formalismen dafür verwandt werden, Gegenstände und Zusammenhänge der Realwelt zu definieren. Durch diese Unabhängigkeit von sprachlichen Eigenschaften stellen Ontologien ein solides Instrumentarium für semantische Beschreibungen dar und können so eine große Hilfe in der Standardisierung von Begrifflichkeiten bieten [13].
Ein bekanntes Beispiel einer polyhierarchischen Ontologie mit Äquivalenzbeziehungen sind die Medical Subject Headings http://www.nlm.nih.gov/mesh/, die zur Beschreibung von in Medline indizierten Publikationen verwendet werden. Abbildung 1 verdeutlicht die Unterschiede zwischen Taxonomie und Ontologie.
Taxonomien und Ontologien sind prinzipiell geeignet, inhaltliche Zusammenhänge auch zwischen Objekten herzustellen, die in voneinander unabhängigen Systemen existieren. So können sie beispielsweise helfen, notwendige aber aufwendige manuelle Verknüpfungen zwischen verschiedenen Outcomeframeworks [14] oder Lernzielkatalogen [15], wie sie unter anderem in einer Minderheit der Ansätze zur Curriculumskartierung genutzt werden [5], [16], zumindest entscheidend vorzubereiten, wenn nicht sogar zu automatisieren.
In der Vorbereitung auf die hier dargestellte Arbeit konnten die Autoren keine allgemein akzeptierten Kriterien zur Beschreibung und zum Vergleich strukturierter Vokabulare identifizieren.

Fragestellung

Bei der Auswahl einer geeigneten Taxonomie bzw. Ontologie sehen sich medizinische Fakultäten einer Vielzahl von unterschiedlich umfangreichen, gewarteten und verfügbaren Angeboten gegenüber. Ziel dieser Arbeit ist es, existierende Taxonomien und Ontologien zur Annotation medizinischer Lernressourcen und -inhalte zusammenzustellen, anhand hierfür entwickelter Kriterien zu vergleichen und auf ihre Eignung für den Einsatz im semantischen Web zu untersuchen.

Methoden

Im ersten Schritt wurde eine Literaturrecherche in den Datenbanken der GMS (German Medical Sciences) http://www.egms.de/dynamic/de/index.htm im Juni 2011 durchgeführt, um Publikationen zu identifizieren, die entweder mindestens eine Taxonomie oder Ontologie für die medizinische Ausbildung oder deren dortigen Einsatz beschreiben. Die Suchbegriffe sind in Tabelle 1 links zusammengestellt.
Bei dieser Recherche konnten fünf strukturierte Vokabulare identifiziert werden.

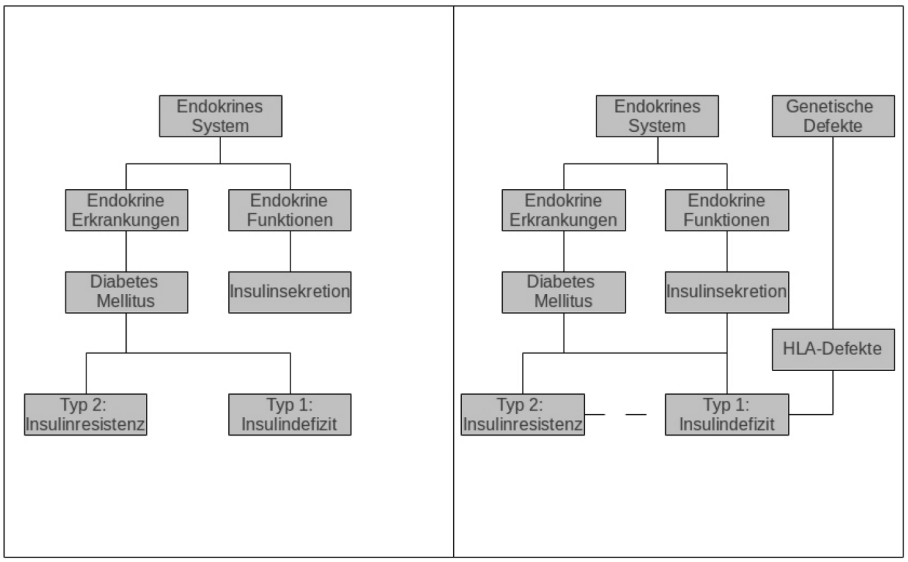

Abbildung 1: In einer Taxonomie (links) ist jedem Element genau ein übergeordnetes Element zugeordnet (Zuordnungen als durchgezogene Linien). In einer Ontologie (rechts) kann ein Element mehrere übergeordnete Elemente haben und zudem mit anderen Elementen assoziiert (gestrichelte Linie) sein.

Tabelle 1: Suchbegriffe der systematischen Literaturrecherche (Zahlenangaben: zur Analyse verwendete Publikationen/Gesamtzahl der Treffer)

Deutsch – vor allem GMS	5/37	English	15/564
medizinische Taxonomie	2/9	Medical education ontology	6/37
medizinische Ontologie	3/12	Taxonomy medical education german	0/93
deutschsprachige Taxonomie	0/2	Taxonomy medical education elearning	5/238
Taxonomie medizinische Lehre	0/2	taxonomy learning resources	1/127
Ontologie medizinische Lehre	0/6	medical education semantic network	2/15
Taxonomie medizinische Lernressourcen	0/3	medical education curriculum mapping taxonomy	0/18
Taxonomie medizinische Lerninhalte	0/3	taxonomy medical education learning objectives	1/36

Die Autoren extrahierten aus diesen Publikationen im Konsens Kriterien zur Beurteilung der Eignung eines Vokabulars für die Annotation von Elementen der medizinischen Ausbildung.
Diese extrahierten Kriterien umfassen:

- *Thema des Vokabulars:* es sollen Ausbildungsformate und -methoden und biomedizinische Inhalte beschreibbar sein.
- *Struktur:* Das Vokabular soll durch eine Vielzahl an unterschiedlichen Nutzern zur Beschreibung von Inhalten einsetzbar sein und deshalb Synonym- und Äquivalenzbeziehungen zwischen Termini unterstützen.
- *Sprache:* Die zu identifizierenden Vokabulare sollen von deutschsprachigen Lehrenden und Lernenden einsetzbar sein.
- *Umfang:* Das Vokabular soll einerseits fein granular genug sein, um Inhalte und Methoden des Medizinstudiums im deutschsprachigen Raum differenziert zu beschreiben, andererseits die annotierenden Nutzer aber nicht durch zu feine Granularität zu irrelevanten Unterscheidungen zwingen.
- *Wartung:* Um nachhaltig einsetzbar zu sein, sollte das Vokabular durch eine feste Institution oder Organisation wenigstens jährlich gewartet sein.
- *Technik:* Datenformat, Verfügbarkeit und Urheberrecht des Vokabulars sollten einen kostengünstigen Einsatz zur webbasierten Beschreibung von Ressourcen und Elementen medizinischer Curricula ermöglichen.

Im zweiten Schritt wurde die systematische Literaturrecherche im Juni und Juli 2011 auf die Datenbanken Medline http://www.ncbi.nlm.nih.gov/pubmed und GMS http://www.egms.de/dynamic/de/index.htm unter Einsatz von 7 zusätzlichen, englischsprachigen Suchbegriffen ausgeweitet. Die zusätzlichen Suchbegriffe sind in Tabelle 1 rechts zusammengestellt. Anhand der Abstracts der gefundenen Artikel wurden alle Publikationen mit Bezug zur medizinischen Ausbildung, die ein formalisiertes Vokabular beschreiben, zur weiteren Analyse ausgewählt und die in der Publikation beschriebene(n) Taxonomie(n) und Ontologie(n) identifiziert.
Zu jedem derart identifizierten strukturierten Vokabular wurde eine Websuche durchgeführt sowie die Herausgeber zu Aktualität, Nutzung, Größe und technischen Details wie Datenformat befragt.
Wo kein Herausgeber des Vokabulars zu eruieren war, wurden die Autoren der beschreibenden Publikationen angeschrieben.
Für jedes strukturierte Vokabular wurden so Sprache, Hierarchieform, Art der Querverweise, Thematik/Fachgebiet, Größe, Aktualität und Pflege, Format und Verfügbarkeit ermittelt.

Ergebnisse

Es konnten 601 Publikationen mit der beschriebenen Suchstrategie gefunden werden. Anhand der Abstracts dieser Publikationen wurden 20 Artikel identifiziert, die die Entwicklung oder den Einsatz eines strukturierten Vokabulars in der medizinischen Ausbildung beschreiben. In den 20 identifizierten Artikeln werden 14 verschiedene strukturierte Vokabulare beschrieben. Vier davon sind explizit für die Beschreibung von Ausbildungsformaten und -methoden entwickelt.
Mit dem „Unified Medical Language System“ (UMLS) existiert darüber hinaus eine Metaontologie, die ein semantisches Netz von über 130 einzelnen Terminologien umfasst.

Struktur, Anwendungsgebiete und Sprache

Die 14 identifizierten Vokabulare konnten einem von drei primären Anwendungsgebieten zugeordnet werden. Diese umfassen „Ausbildungsformate und -methoden“, „biomedizinische Inhalte“ sowie „Dokumentation und Verwaltung“ (siehe Tabelle 2). Im Wesentlichen unterscheiden sich die Vokabulare jedoch in ihrem Aufbau durch die Art der Verknüpfungen zwischen ihren Termini: Die polyhierarchischen Ontologien einerseits umfassen in aller Regel nicht nur Eltern-Kind-Beziehungen zwischen den Termini, sondern ermöglichen auch Äquivalenz- und Assoziationsverweise. Die hier beschriebenen monohierarchischen Taxonomien andererseits lassen regelhaft nur eine Form von Verknüpfung ihrer Termini zu und sind allenfalls teilweise um eine Verknüpfung vom Typ „Synonym“ ergänzt. Die einem Vokabular zugrunde liegende Struktur richtet sich bei den hier Betrachteten wesentlich nach dem Zweck und den durch das Vokabular zu beschreibenden Inhalten. Die eher zur Verwaltung und Dokumentation entwickelten Vokabulare folgen einer monohierarchischen Struktur, die Ontologien zur Beschreibung von Ausbildungsformaten und -methoden sind, ebenso wie jene zur Beschreibung biomedizinischer Inhalte, polyhierarchisch strukturiert.
Weitere zentrale Unterscheidungsmerkmale innerhalb dieser Anwendungsgebiete, wie Sprache und Umfang, sind ebenfalls in Tabelle 2 aufgeführt.
Drei der identifizierten Vokabulare zur Beschreibung von Ausbildungsformaten und -methoden liegen primär in Englisch, “Topics for Indexing Medical Education“ (TIME) darüber hinaus auch in Französisch vor. Die TIME-Ontologie ist zur inhaltlichen Beschreibung medizinischer Lehr- und Lerninhalte konzipiert, mit der Taxonomie der „Medical Education Taxonomy Research Organization“ (METRO) können vor allem Methoden und Prozesse der medizinischen Ausbildung annotiert werden. Der „British Education Thesaurus“ (BET) ist ein allgemeiner Ausbildungsthesaurus ohne spezifischen Bezug zur Medizin.
Mit der “Ontology of Bio-Medical Educational Objectives“ (OBEO) steht darüber hinaus eine deutschsprachige Ontologie zur semantischen Annotation vor allem von Lernzielen in der medizinischen Ausbildung zur Verfügung. Gleichzeitig sind die für die Annotation biomedizinischer Inhalte vorgesehenen Vokabulare teilweise sehr umfangreich und / oder stark spezialisiert. Lediglich vier von ihnen sind für die umfassende Beschreibung der Medizin entwickelt und fokussieren nicht auf ein Themengebiet wie beispielsweise die Anatomie, Onkologie oder Pharmazie (siehe Tabelle 2), sondern eignen sich zur Beschreibung eines umfassenden Teils der biomedizinischen Inhalte.

Umfang, Wartung und Technik

Während der BET die umfangreiche Annotation von allgemeinen Ausbildungsinhalten ermöglicht, sind die Möglichkeiten zur Beschreibung medizinspezifischer Inhalte mit diesem Thesaurus eingeschränkt. Wie aus Tabelle 2 ersichtlich, sind die anderen drei identifizierten Vokabulare (METRO, OBEO und TIME) zur Beschreibung von Ausbildungsformaten und -methoden weniger umfangreich als die Vokabulare zur Beschreibung von biomedizinischen Inhalten oder Inhalten der Verwaltung und Dokumentation.
Umgekehrt sind die vier identifizierten Vokabulare zur umfassenden Beschreibung biomedizinischer Inhalte teilweise sehr umfangreich, wie aus Tabelle 2 ersichtlich. Um aus diesen Vokabularen zur entweder inhaltlich-biomedizinischen oder methodisch-didaktischen Beschreibung von Objekten ein je nach Anwendungsgebiet geeignetes Vokabular wählen zu können, sind in Tabelle 3 Daten zu Aktualität und Pflege, Datenformat und Kosten für diese sieben Ontologien – drei zur Beschreibung von Ausbildungsformaten und -methoden, vier zur Beschreibung biomedizinischer Inhalte – zusammengestellt.

Diskussion

Die Nutzung strukturierter Vokabulare, wie sie Taxonomien und Ontologien darstellen, ermöglicht den maschinellen Vergleich und Austausch von Daten zwischen Institutionen und Systemen [17] und den Vergleich unterschiedlicher Inhalte mit automatischer Identifikation von inhaltlichen Schnittmengen zwischen zwei Informationsmengen: Dadurch wird beispielsweise der automatische Abgleich eines medizinischen Curriculums mit beliebigen Outcomeframeworks möglich [15] – eine Funktion, die nicht zuletzt im Rahmen von Akkreditierungen von besonderem Interesse sein dürfte [4].
Da die Vokabulare prinzipiell zur formalen inhaltlichen Beschreibung beliebiger Objekte einsetzbar sind, können neben klassischen Ressourcen wie Büchern, Medien und Programmen auch Kurse, Lernziele oder Prüfungsfragen und jede Form von Onlineressource annotiert und damit Ausgangspunkt oder Ergebnis einer Suche nach thematisch verwandten Objekten werden. Einige wenige Implementierungen curricularer Kartierung greifen diese Möglichkeit auf, um die inhaltlichen Zusammenhänge von

Tabelle 2: Anwendungsgebiete, Sprache und Umfang identifizierter Vokabulare. Angaben zum Umfang laut Beschreibungen des Herausgebers. Mit ° markierte Vokabulare sind in der Metaontologie UMLS enthalten.

Polyhierarchische Ontologien zur Beschreibung von Ausbildungsformaten und -methoden Verweise: Hierarchie, Assoziation (nicht *), Äquivalenz				
Kürzel	**Name**	**Thema**	**Sprache(n)**	**Umfang**
BET	British Education Thesaurus	Ausbildungsinhalte aller Art	Eng	Keine Angaben
METRO	Medical Education Taxonomy Research Organization	Prozesse & Methoden in medizinischer Ausbildung	Eng	ca. 180 Termini
OBEO	Ontology of Bio-Medical Educational Objectives *	Medizinische Ausbildung (v.a. Lernziele)	Deu	eigene: 72 Klassen, 12 Objekt-Relationen; Grundlage: Upper-Level Ontologie BioTop (375 Klassen, 76 Objekt-Relationen), Toplevel-Ontologie DOLCE-Lite (37 Klassen, 70 Objekt-Relationen)
TIME	Topics for indexing Medical education	Biomedizinische Inhalte, Medizinische Lehr-/ Lerninhalte	Eng, Frz	1.500 Termini, 2.400 Querverweise
Polyhierarchische Ontologien biomedizinischer Inhalte Verweise: Hierarchie, Assoziation (nicht *), Äquivalenz				
Kürzel	**Name**	**Thema**	**Sprache(n)**	**Umfang**
FMA	Foundational Model of Anatomy	Anatomie	Eng, (Deu), weitere	75.000 Klassen, 120.000 Termini, 2.100.000 Querverweise, 168 Beziehungsarten
-	Gene Ontology*	Biowissenschaften	Eng	34.439 Termini
MeSH	Medical Subject Heading°	Biomedizinische Inhalte	Eng, Deu	Englisch: 26.142 Hauptschlagwörter, 181.141 Synonyme Deutsch: 26.142 Hauptschlagwörter, 57.656 Synonyme
NCI Thesaurus	National Cancer Institute Thesaurus*	Onkologie	Eng	34.000 Konzepte, 200.000 Querverweise
RePORT	Research Portfolio Online Reporting Tools	Biomedizinische Inhalte	Eng	Bezug aus mehreren Datenbanken (eRA databases, Medline, PubMed Central, the NIH Intramural Database, iEdison)
-	RxNorm °	Pharmazeutika	Eng	keine genauen Angaben, jedoch alle verschreibungspflichtige Medikamente der USA enthalten
SNOMED-CT	Systematized Nomenclature of Medicine – Clincal Terms°	Biomedizinische Inhalte, Klinik	Eng, (Deu), weitere	311.000 Konzepte, 800.000 Termini, 1.360.000 Querverweise
UMLS	*Unified Medical Language System*	*Biomedizinische Inhalte*	*Eng, Deu, weitere*	*Semantisches Netz von >130 Einzelterminologien (unter anderem °); > 1.179.00 Konzepte, > 4.174.00 Termini*
Monohierarchische Taxonomien für Verwaltung und Dokumentation Verweise: Hierarchie, Synonyme (nur #)				
ICD	International Classification of dieseases°	Klassifikation der Krankheiten und verwandter Gesundheitsprobleme	Eng, Deu	70.000 Kodierungen
LOINC	Logical Observation Identifier Names and Codes#°	Untersuchungs-/ Testergebnisse aus Labor und Klinik	Eng, (Deu)	7 Kategorien, 100.000 Datensätzen (Deu: 3800)
OPS	Operationen- und Prozedurenschlüssel	Operationstechniken/ Prozeduren	Deu	6 Kapitel, 65 Bereiche, ca. 28.300 Prozedurenklassen

Tabelle 3: Aktualität und Pflege, Datenformat, Kosten und Weblinks von Vokabularen medizinspezifischer Inhalte.

Polyhierarchische Ontologien zur Beschreibung von Ausbildungsformaten und -methoden Verweise: Hierarchie, Assoziation (nicht *), Äquivalenz					
Kürzel	**Name**	**Aktualität & Pflege**	**Datenformat**	**Kosten**	**Weblinks**
METRO	Medical Education Taxonomy Research Organization	Seit 2005 keine Weiterentwicklung	Download – XML, Website (Blogspot)	Open Source	http://metro2.blogspot.de/
OBEO	Ontology of Bio-Medical Educational Objectives*	„work-in-progress“ Entwurf (Universitätsklinikum Freiburg, Dtl.)	Downoad – OWL	Open Source	http://www.imbi.uni-freiburg.de/ontology/obeo/obeo.owl
TIME	Topics for indexing Medical education	„work-in-progress“ (University of Ottawa, Kanada)	Download – XML, Textformat, Web based	Open Source	www.time-item.org (derzeit offline)
Polyhierarchische Ontologien biomedizinischer Inhalte Verweise: Hierarchie, Assoziation (nicht *), Äquivalenz					
Kürzel	**Name**	**Aktualität & Pflege**	**Datenformat**	**Kosten**	**Weblinks**
MeSH	Medical Subject Heading	jährlich (NLM,USA; DIMDI, Dtl.)	Download/ CD-ROM – XML, CSF, Textformat; Web based (MeSH-Browser)	Open Source (Englisch, nach Registrierung NLM), kostenpflichtig (Deutsch)	http://www.nlm.nih.gov/mesh/ (Englisch) http://www.dimdi.de/static/en/klassi/mesh_umls/mesh/index.htm (Deutsch)
RePORT	Research Portfolio Online Reporting Tools	wöchentlich (NIH, USA)	Download – XML, CSV; Web based (RePorter)	Open Source	http://report.nih.gov/
SNOMED-CT	Systematized Nomenclature of Medicine – Clincal Terms	halbjährlich (IHTSDO, Dänemark)	Download/ CD-ROM, MetamorphoSys (Java); Web based (UMLS Terminology Services (UTS))	kostenpflichtig für Nicht-Mitgliedsstaaten (Dtl.) (nach Registrierung UMLS/ IHTSDO)	http://www.ihtsdo.org/snomed-ct/
UMLS	*Unified Medical Language System*	*jährlich (NLM, USA)*	*Download/ CD-ROM, MetamorphoSys (Java); Web based (UMLS Terminology Services (UTS))*	*Open Source (nach Registrierung NLM, Copyright)*	*http://www.nlm.nih.gov/research/umls/*

Elementen eines Curriculums aufzuzeigen [5], [15]. Durch die maschinenlesbare inhaltliche Beschreibung von Lernmodulen in e-learning-Systemen mit Hilfe von Taxonomien könnten die Nutzer solcher Module automatisch auf thematisch verwandte Module in denselben (oder auch anderen) e-learning-Plattformen verwiesen werden. Formalisierte inhaltliche Beschreibungen von Ressourcen sind zudem flexibel zu durchsuchen und leicht zu warten [8].

Der Einsatz solcher Vokabulare setzt die konsequente Annotation aller neuen Objekte mittels des definierten Vokabulars voraus, was einen Mehraufwand beim Erstellen neuer Inhalte bedeutet. Gleichzeitig entfällt die Notwendigkeit der manuellen Verknüpfung neuer Inhalte. Das semantische Beschreiben von Objekten wird also erst mit zunehmendem Umfang an Inhaltselementen effizient.

Im deutschsprachigen Raum konnte keine ad hoc nutzbare Taxonomie oder Ontologie zur Annotation von Ressourcen und Inhalten der medizinischen Ausbildung identifiziert werden.

Die in deutscher Sprache verfügbaren strukturierten Vokabulare sind entweder - wie etwa OBEO - bisher nicht fein granular genug, um eine sinnvolle inhaltliche Unterscheidung zwischen Lernressourcen zu ermöglichen oder zwar sehr fein granular, aber eher für die inhaltliche Annotation von Patientendaten (SNOMED-CT) oder biomedizinischen Publikationen (MeSH) geeignet.

Die teilweise erheblichen Unterschiede im Umfang zwischen den identifizierten Vokabularen, die vor allem aus den sehr spezifischen Einsatzgebieten wie etwa der Beschreibung von Unterrichtsformaten (METRO), resultieren, werfen die Frage auf, ob die Granularität ausreichend fein zur differenzierten Annotation von Inhalten und Methoden der medizinischen Ausbildung ist.

Die in englischer Sprache verfügbaren, speziell für die Domäne der medizinischen Ausbildung entwickelten Ontologien, wie TIME oder METRO, müssten zudem zunächst ins Deutsche überführt werden. Dabei entstehen Probleme in der Auswahl inhaltlicher Entsprechungen. So findet der englischsprachige Begriff „medical education“ eine deutsche Entsprechung nur in der Trias aus medizinischer Aus-, Fort- und Weiterbildung, die im Englischen wiederum eher den Begriffen „graduate- “ , „continous- “ und „postgraduate- “ medical education entsprechen.

Ein weiteres Problem einer Übersetzung ist die Frage, ob und wie eine Übersetzung an die mögliche Weiterentwicklung und Pflege des originalsprachlichen Vokabulars gebunden werden kann.

Eine alternative Neuentwicklung einer primär deutschsprachigen Ontologie ist jedoch in jedem Fall mit erheblichem Aufwand verbunden.

Auffällig bei den identifizierten und in Tabelle 3 zusammengefassten Vokabularen ist insbesondere, dass alle wenigstens online verfügbar sind und teilweise direkt Technologien des Internets wie etwa XML als Datenformat einsetzten. Obwohl sie nicht primär für den Einsatz im Web entwickelt wurden, sind sie demnach jedoch zumindest dafür geeignet. Die technischen Grundlagen auf dem Weg zum Web 3.0 sind damit gelegt [9].

Limitationen

Wie jede Literaturrecherche ist diese Arbeit limitiert durch Suche und Auswahl der verwendeten Publikationen.

Darüber hinaus besteht die Möglichkeit, dass strukturierte Vokabulare existieren, zu denen keine in den verwendeten Datenbanken indizierten Publikationen vorliegen.

In der Vorbereitung auf die hier dargestellte Arbeit konnten die Autoren keine allgemein akzeptierten Kriterien zur Beschreibung und zum Vergleich strukturierter Vokabulare identifizieren. In einem ersten Schritt wurden deshalb zunächst in einer Teilmenge der durchsuchten Datenbanken Taxonomien identifiziert und aus diesen Publikationen die von den jeweiligen Autoren selbst verwendeten Kriterien zur Beschreibung ihres strukturierten Vokabulars ermittelt. Aus diesen Kriterien wurden für diese Arbeit im Konsens solche extrahiert, die den Autoren geeignet erscheinen, Taxonomien für die medizinische Ausbildung miteinander zu vergleichen. In der Auswahl dieser Kriterien haben die Autoren sich um die Generalisierbarkeit der Ergebnisse auf die Situation an den deutschsprachigen medizinischen Fakultäten bemüht. Lokal könnten andere oder weitere, hier nicht behandelte Kriterien von Bedeutung sein.

Fazit

Die inhaltliche Beschreibung (semantische Annotation) von medizinischen Lernressourcen bietet die Möglichkeit, inhaltlich verwandte Elemente über Systemgrenzen hinweg zu identifizieren. Bisher existiert keine geeignete deutschsprachige Taxonomie oder Ontologie, um eine systematische Ressourcenbeschreibung für Medizinstudierende in Deutschland zu realisieren und das Potential des Web 3.0 zur Unterstützung der Curriculumskartierung nutzen zu können. Mögliche Ansatzpunkte sind die Neuentwicklung eines geeigneten strukturierten Vokabulars, die Übersetzung existierender englischsprachiger Vokabulare oder eine Kombination aus bestehenden Teilvokabularen, ggf. mit einer teilweisen Übersetzung englischsprachiger Vokabulare.

Anmerkung

WB leitet das Lernzentrum der Charité und ist Mitglied der GMA und deren Ausschüsse „Praktische Fertigkeiten“ und „Methodik der Ausbildungsforschung“. AJ und PS sind studentische Tutor/-innen des Lernzentrums und Studierende der Charité. FB ist Co-Entwickler der OBEO. OA leitet die Abteilung für Curriculumsorganisation der Charité und ist Mitglied der GMA und deren Ausschuss „Methodik der Ausbildungsforschung“.

Danksagung

Die Autoren danken Dr. Ullrich Woermann, Bern sowie zwei anonymen Reviewern für Ihre kollegiale und konstruktive Kritik am Manuskript und Rudi Mörgli für die Übersetzung des Manuskripts ins Englische.

Interessenkonflikt

Die Autoren erklären, dass sie keine Interessenkonflikte im Zusammenhang mit diesem Artikel haben.

Literatur

1. Holzer M, Pfähler M, Hege I, Fischer M. Wer suchet, der soll finden! - Ein Überblick über Verschlagwortung und Suche medizinischer Lerninhalte. GMS Med Inform Biom Epidemiol. 2006;2(3):Doc20. Zugänglich unter/available from: http://www.egms.de/static/de/journals/mibe/2006-2/mibe000039.shtml
2. Cook DA, Erwin PJ, Triola MM. Computerized virtual patients in Health Professions Education: a systematic review and meta-analysis. Acad Med. 2010;85(10):1589-1602. DOI: 10.1097/ACM.0b013e3181edfe13
3. Junco R, Mastrodicasa J. Connecting to the Net Generation: What Higher Education Professionals Need to Know About Today's Students. Washington, DC: National Association of Student Personnel Administrators; 2007.
4. Harden RM. AMEE Guide No. 21: Curriculum mapping: a tool for transparent and authentic teaching and learning. Med Teach. 2001;23(2):123-137. DOI: 10.1080/01421590120036547
5. Ahlers O, Georg W, Blaum W, Stieg M, Hanfler S, Bubser F, Spies C. Der Einsatz einer interdisziplinären, webbasierten Lernzielplattform verbessert sowohl die Unterrichtsqualität als auch die Klausurergebnisse Studierender. Jahrestagung der Gesellschaft für Medizinische Ausbildung (GMA). Bochum, 23.-25.09.2010. Düsseldorf: German Medical Science GMS Publishing House; 2010. Doc10gma13. DOI: 10.3205/10gma013
6. Berners-Lee T, Hendler J, Lassila O. The Semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. Sci Am. 2001;284:34-43. DOI: 10.1038/scientificamerican0501-34
7. Alby T. Web 2.0. Konzepte, Anwendungen, Technologien. München: Hanser Verlag; 2007.
8. Berners-Lee T, Fischetti M. Weaving the web: the original design and ultimate destiny of the World Wide Web by its inventors. New York: HarperBusiness; 2006.
9. Segaran T, Evans C, Taylor J. Programming the semantic web. Sebastopol, CA: O'Reilly & Associates; 2009.
10. Willett TG, Marshall KC, Broudo M, Clarke M. TIME as a generic index for outcome-based medical education. Med Teach. 2007;29(7):655–659. DOI: 10.1080/01421590701615808
11. Willett TG, Marshall KC, Broudo M, Clarke M. It's about TIME: a general-purpose taxonomy of subjects in medical education. Med Educ. 2008;42(4):432-438. DOI: 10.1111/j.1365-2923.2008.03012.x
12. Boeker M, Schober D, Schulz S, Balzer F. Ontology of Bio-Medical Educational Objectives (OBEO): ein Vorschlag für eine Ontologie medizinischer Lernziele. GMS Med Inform Biom Epidemiol. 2010;6(2):Doc11. DOI: 10.3205/mibe000111
13. Stenzhorn H, Schulz S, Boeker M, Smith B. Adapting Clinical Ontologies in Real-World Environments. J Univers Comput Sci. 2008;14(22):3767-3780.
14. Ellaway R, Evans P, McKillop J, Cameron H, Morrison J, McKenzie H, Mires G, Pippard M, Simpson J, Cumming A, Harden R, Guild S. Cross-referencing the Scottish Doctor and Tomorrow's Doctors learning outcome frameworks. Med Teach. 2007;29(7):630-635. DOI: 10.1080/01421590701316548
15. Blaum WE, Dannenberg KA, Friedrich T, Jarczewski A, Reinsch AK, Ahlers O. Der praktische Nutzen des Konsensusstatements "praktische Fertigkeiten im Medizinstudium" – eine Validierungsstudie. GMS Z Med Ausbild. 2012;29(4):Doc58. DOI: 10.3205/zma000828
16. Willett TG. Current status of curriculum mapping in Canada and the UK. Med Educ. 2008;42(8):786-793. DOI: 10.1111/j.1365-2923.2008.03093.x
17. Uschold M, Gruninger M. Ontologies: Principles, Methods and Applications. Knowl Eng Rev. 1996;11:93–155. DOI: 10.1017/S0269888900007797

Korrespondenzadresse:
Wolf E. Blaum
Charité - Universitätsmedizin Berlin, Abteilung für Curriculumsorganisation, Lernzentrum, Charitéplatz 1, 10117 Berlin, Deutschland, Tel.: +49 (0)30/450-576003, Fax: +49 (0)30/450-576952
wolf.blaum@charite.de

Bitte zitieren als
Blaum WE, Jarczweski A, Balzer F, Stötzner P, Ahlers O. Auf dem Weg zum Web 3.0: Taxonomien und Ontologien für die medizinische Ausbildung - eine systematische Literaturrecherche. GMS Z Med Ausbild. 2013;30(1):Doc13.
DOI: 10.3205/zma000856, URN: urn:nbn:de:0183-zma0008568

Artikel online frei zugänglich unter
http://www.egms.de/en/journals/zma/2013-30/zma000856.shtml

Eingereicht: *29.06.2003*
Überarbeitet: *05.09.2012*
Angenommen: *12.10.2012*
Veröffentlicht: *21.02.2013*

Copyright
©2013 Blaum et al. Dieser Artikel ist ein Open Access-Artikel und steht unter den Creative Commons Lizenzbedingungen (http://creativecommons.org/licenses/by-nc-nd/3.0/deed.de). Er darf vervielfältigt, verbreitet und öffentlich zugänglich gemacht werden, vorausgesetzt dass Autor und Quelle genannt werden.

Towards Web 3.0: Taxonomies and ontologies for medical education - a systematic review

Abstract

Introduction: Both for curricular development and mapping, as well as for orientation within the mounting supply of learning resources in medical education, the Semantic Web ("Web 3.0") poses a low-threshold, effective tool that enables identification of content related items across system boundaries. Replacement of the currently required manual with an automatically generated link, which is based on content and semantics, requires the use of a suitably structured vocabulary for a machine-readable description of object content.
Aim of this study is to compile the existing taxonomies and ontologies used for the annotation of medical content and learning resources, to compare those using selected criteria, and to verify their suitability in the context described above.
Methods: Based on a systematic literature search, existing taxonomies and ontologies for the description of medical learning resources were identified. Through web searches and/or direct contact with the respective editors, each of the structured vocabularies thus identified were examined in regards to topic, structure, language, scope, maintenance, and technology of the taxonomy/ontology. In addition, suitability for use in the Semantic Web was verified.
Results: Among 20 identified publications, 14 structured vocabularies were identified, which differed rather strongly in regards to language, scope, currency, and maintenance.
None of the identified vocabularies fulfilled the necessary criteria for content description of medical curricula and learning resources in the German-speaking world.
Discussion: While moving towards Web 3.0, a significant problem lies in the selection and use of an appropriate German vocabulary for the machine-readable description of object content. Possible solutions include development, translation and/or combination of existing vocabularies, possibly including partial translations of English vocabularies.

Keywords: medical education, semantic web, web 3.0, taxonom, ontology, curricular mapping, curriculum charting, e-learning, new media

Wolf E. Blaum[1,2]
Anne Jarczweski[2]
Felix Balzer[1,2]
Philip Stötzner[2]
Olaf Ahlers[1,2]

1 Charité - University Medicine Berlin, Campus Charité Mitte and Campus Virchow-Klinikum, Department of Anesthesiology and Intensive Care Medicine Campus, Berlin, Germany

2 Charité - University Medicine Berlin, Department of Curriculum Management, Learning Center, Berlin, Germany

Authorship

The authors Blaum and Jarczewski contributed equally to the study.

Introduction

At the present time, mounting supply of available learning resources in medical education, particularly online, can hardly be used effectively by teachers and learners due to the lack of a simple and systematic mode of access to its contents [1]. Due to the constantly increasing number of available sources, students are confronted with the challenge of identifying, in a sea of possibilities, thematically appropriate, high-quality, and didactically significant material that corresponds to their level of education.
Computer-based learning in particular has experienced a substantial increase of resources. In a recent review in the topic of computer-based virtual patients - a mere subset of the computer-based resources - Cook et al. identified 698 publications on this single issue [2]. At the same time, the PC-based learning is becoming increasingly important with the arrival of the "Generation Y" [3]. Teachers and planners are also being challenged to present (to map) the contextual relationships in the curriculum for which they are responsible [4], a task that until now required only a manual linking of curricular elements.

Semantic Web

Notably, development of the Internet towards the Semantic Web (Web 3.0) presents considerable potential for the contextual linking of learning resources, as well as applications in curricular mapping [5], [6].
While at first content of the Internet was essentially determined by a few authors, who provided information to a broad, consuming audience, the current "Web 2.0" is characterized by user-generated content [7]. A common feature in data from both versions is that their information can only be understood and interpreted by humans. Hence, any contextual relation between data requires human interaction.
The term "Web 3.0" (also "Semantic Web") refers to a network data storage system which, through the addition of machine-readable meta-information (per example, via structured vocabularies), allows machines to automatically generate and discover contextual relations between data objects [8]. This innovation requires a machine-readable description of the information contained in data [6], [9].
A suitable type of content description is the use of structured vocabularies. In this case, each individual resource is tagged with the defined terms of a structured vocabulary.

Taxonomy und Ontology

Taxonomies and ontologies provide controlled vocabularies for content description (semantic annotation) of objects, such as, per example, learning resources [10]. Relationship between the terms of a vocabulary are stored within a taxonomy, allowing for an (manually initiated) automated search to identify similar content descriptions among a multitude of learning resources, independent of their location [1], [11], [12].
Relationships between vocabulary terms of a taxonomy are typically monohierarchical, thus each term has exactly one (or no) parent term ("is" - or "is a part of" - relationship), as well as an arbitrary number of subordinate (child) terms. A classic example of a strictly monohierarchical taxonomy is the International Classification of Diseases http://www.dimdi.de/static/de/klassi/index.htm. Conversely, ontologies may also utilize polyhierarchical structures, so that a single vocabulary term may have multiple parent terms. The fundamental difference to taxonomies is that, instead of language constructs, logically defined formalisms are used for the definition of objects and relationships in the real world. Due to this independence from linguistic properties, ontologies represent a robust toolkit for semantic descriptions, and can greatly facilitate the standardization of terminology [13].
A well known example of a polyhierarchical ontology with equivalence relationships are the Medical Subject Headings http://www.nlm.nih.gov/mesh/, which are used to describe publications indexed in Medline. Figure 1 clarifies the differences between taxonomy and ontology.
Taxonomies and ontologies are fundamentally suitable to create contextual relationships between objects, even if they are located in separate, independent systems. They may also further assist preparation, if not complete automation, of the necessary – but burdensome - manual linking between different outcome frameworks [14] or learning objective catalogs [15], as they are generally used in a minority of approaches to curriculum mapping.
To the best of our knowledge, generally accepted criteria for description and comparison of structured vocabularies are currently not available.

Research Question

When choosing an appropriate taxonomy or ontology, medical faculties are faced with a multitude of offers that differ in scope, maintenance, and availability. Aim of this study is to compile existing taxonomies and ontologies used for annotation of medical learning resources and content, to compare those using purpose-built criteria, and to examine their suitability for use in the Semantic Web.

Methods

A literature search in the databases of the GMS (German Medical Sciences) http://www.egms.de/dynamic/de/index.htm was initially carried out in June 2011, in order to identify publications describing at least one taxonomy/ontology with current or potential use in the medical field. Searched terms are summarized in Table 1, on the left column.
Five structured vocabularies were identified.
The authors extracted criteria used to assess the suitability of a vocabulary for annotation of elements of medical education from these publications in consensus.
These extracted criteria include:

- *Vocabulary theme:* training formats and methods, as well as biomedical content, should be described.
- *Structure:* The vocabulary should support content description by a variety of different users, and therefore support synonym and equivalence relationships between terms.
- *Language:* German-speaking teacher and students should be able to use the respective vocabulary.
- *Scope:* The vocabulary should be specific enough to differentiate the contents and methods used in medical studies in the German-speaking world; but conversely, not so detailed as to force the users to annotate using irrelevant distinctions.
- *Maintenance:* In order to remain current, the vocabulary should be updated by a specified institution or organization at least once a year.
- *Technology:* Data format, availability, and copyright of the vocabulary should allow for a cost-effective use in description of web-based resources and elements of medical curricula.

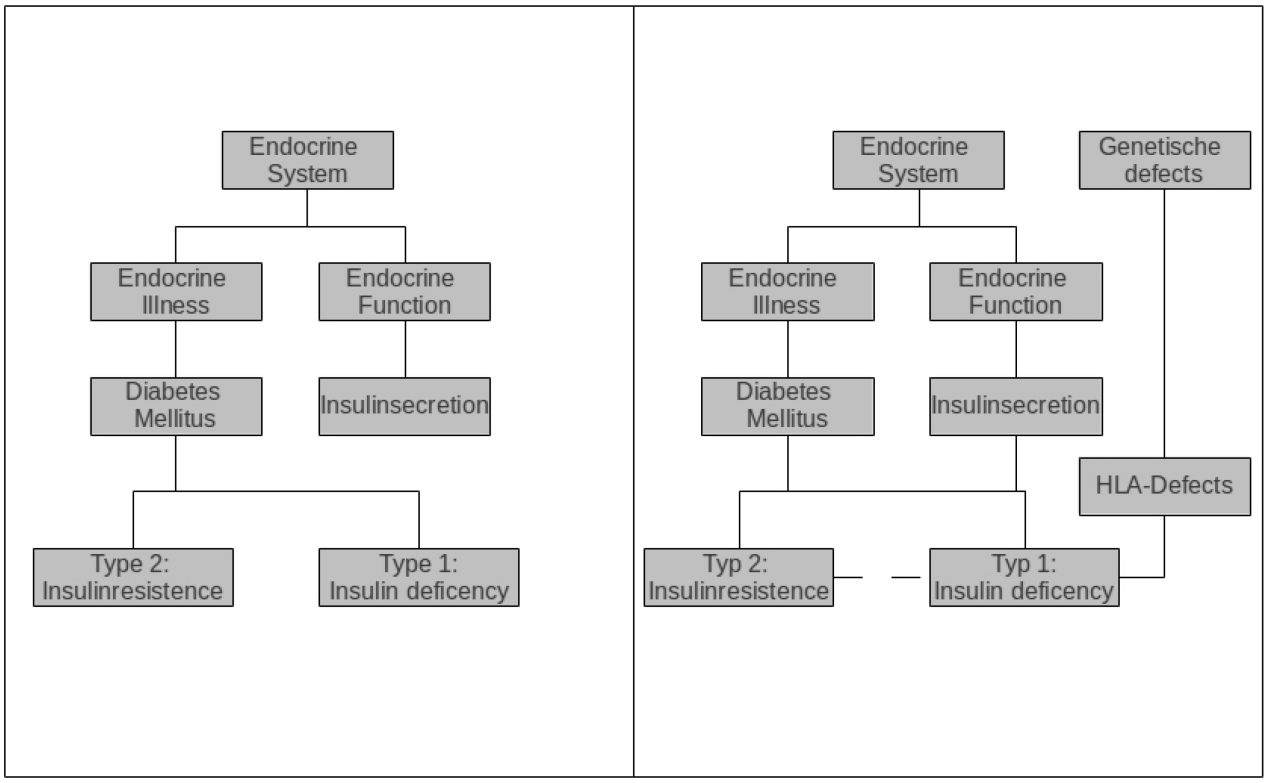

Figure 1: In a taxonomy (left), each element is assigned exactly one parent term (solid lines). In an ontology (right), an element may not only be assigned multiple parent terms, but may also be associated with other elements (dashed line).

Table 1: Terms used in the systematic literature search (Figures: publications used in the analysis / total number of hits)

German – mainly GMS	5/37	English	15/564
medizinische Taxonomie	2/9	Medical education ontology	6/37
(medical taxonomy)		Taxonomy medical education german	0/93
medizinische Ontologie	3/12	Taxonomy medical education elearning	5/238
(medical ontology)		taxonomy learning resources	1/127
deutschsprachige Taxonomie	0/2	medical education semantic network	2/15
(German language taxonomy)		medical education curriculum mapping taxonomy	0/18
Taxonomie medizinische Lehre	0/2	taxonomy medical education learning objectives	1/36
(taxonomy medical teachings)			
Ontologie medizinische Lehre	0/6		
(ontology medical teachings)			
Taxonomie medizinische Lernressourcen	0/3		
(taxonomy medical educational resources)			
Taxonomie medizinische Lerninhalte	0/3		
(taxonomy medical educational content)			

The second step took place in June and July 2011, and involved an expansion of the systematic literature search in Medline http://www.ncbi.nlm.nih.gov/pubmed and in GMS http://www.egms.de/dynamic/de/index.htm, using 7 supplementary English terms. The additional terms searched are summarized in Table 1, on the right column. Based on the abstracts of the search results, all publications describing a formal vocabulary related to medical education were selected for further analysis, and the cited taxonomies and ontologies were identified.

A web search was conducted for each of the structured vocabularies thus identified, and their respective editors were contacted and questioned regarding the actuality, use, size and technical details (such as data format) of their vocabularies.

If the editor of a vocabulary could not be identified, the authors of the publication describing the respective vocabulary were then contacted.

Each structured vocabulary was analyzed in regard to language, hierarchical design, type of cross-referencing, topic/subject area, size, currency, maintenance, format, and availability.

Results

The described search strategy identified 601 publications. Based on the abstracts of these publications, 20 articles were identified describing the development or use of a structured vocabulary in medical education.

Among the 20 articles found, 14 different structured vocabularies were described. Four of these were explicitly designed for description of training formats and methods. Furthermore, a Metaontology comprising a semantic network of more than 130 different terminologies is

available under the "Unified Medical Language System" (UMLS).

Structure, Applications and Language

The 14 identified vocabularies could be classified into one of three primary applications. These applications include "Training Formats and Methods", "Biomedical Content", and "Administration and Documentation" (see Table 2). In essence, vocabulary structures differ in the method by which their terms are connected: polyhierarchical ontologies include not only the parent-child relationships between terms, but also allow for analysis of equivalency and association references. Conversely, the described monohierarchical taxonomies allow only one type of relationship between its terms, and are at best supplemented by a "synonym" type link. The underlying structure of a vocabulary depends substantially on its purpose or on the content that the vocabulary is designed to describe. Vocabularies primarily developed for documentation and administration tend to follow a monohierarchical structure, while ontologies meant to describe training formats and methods, as well as those describing biomedical content, are usually polyhierarchically structured.
Further essential distinguishing features within these areas of application, such as language and scope, are listed in Table 2.
Three of the identified vocabularies designed to describe training formats and methods are mainly in English, such as "Topics for Indexing Medical Education" (TIME), and secondarily in French. The TIME-ontology is designed for description of medical teaching and learning content, while the "Medical Education Taxonomy Research Organization" (METRO) taxonomy is mainly suitable for the annotation of methods and processes of the medical education. The "British Education Thesaurus" (BET) is a general education thesaurus without any specific reference to medicine.
Additionally, an ontology in German language is available through the "Ontology of Bio-Medical Educational Objectives "(OBEO), which is used primarily for the semantic annotation of learning objectives in medical training.
Vocabularies intended for annotation of biomedical content are partially very extensive and/or specialized. Only four of them were designed for a comprehensive medical description, not focusing on a single thematic area, such as anatomy, oncology, or pharmacology (see Table 2), but rather fit to describe extensively a broad area of the biomedical content.

Scope, Maintenance, and Technology

While BET allows extensive annotation of general educational content, the possibility of describing specific medical content with this thesaurus is restricted. As seen in Table 2, the other three identified vocabularies (METRO, OBEO, and TIME), which have been designed for description of training formats and methods, are significantly less extensive than vocabularies used for describing biomedical content, or content administration and documentation.
Conversely, the four identified vocabularies for comprehensive description of biomedical content are sometimes very extensive (see Table 2).
Table 3 summarizes data on updates and support, data format, and costs for the seven identified ontologies - three for the description of training formats and methods, and four for the description of biomedical content.

Discussion

Use of structured vocabularies, such as taxonomies and ontologies, enables automatic comparison and exchange of data between institutions and systems [17]. It also allows automatic comparison of multiple sets of information, which automatically identifies overlapping content: a function that would, for instance, enable automatic adjustment of a medical curriculum with any possible outcome frameworks [15] – a particularly interesting feature in terms of accreditation.
Since the vocabularies can in principle be used for the formal content description of any object, they may be used - in addition to traditional resources such as books, media, and programs – to annotate courses, learning objectives, exam questions, or any other form of online resource. As such, they provide a starting point or search results for thematically related objects. A few curriculum mapping implementations utilize this opportunity to show contextual relationships between elements in a curriculum [5], [16]. Users of such modules can be automatically referred to related modules in the same (or another) e-learning platform by machine-readable description of learning module contents in e-learning systems, which is based on taxonomies.
In addition, formalized content description of resources is easy to browse and to maintain [8].
Use of such vocabularies requires consistent annotation of all new objects using a defined vocabulary, which signifies an additional effort when creating new content. Conversely, this eliminates the need for manually linking new content. Semantic description of objects, therefore, can only be effective with an increasing amount of content elements.
No German language based, purpose built taxonomy or ontology for the annotation of resources and medical education content could be identified.
The structured vocabularies available in German - such as OBEO - are not yet specific enough to allow for a meaningful distinction among learning resources, but rather appropriate for the content annotation of patient data (SNOMED-CT) or biomedical publications (MeSH).
Considerable differences in the scope of the identified vocabularies - which resulted primarily from very specific applications, such as the description of teaching formats (METRO) - raise the question of whether their granularity is fine enough for a differentiated annotation of content and methods in medical education.

Table 2: Applications, language and extent of the identified vocabularies. Details regarding scope were provided by the individual editors. Vocabularies marked with ° are included in the UMLS Metaontology.

Polyhierarchical Ontologies for the Description of Training Formats and Methods References: hierarchy, (no*) association, equivalency				
Acronym	**Name**	**Topic**	**Language**	**Scope**
BET	British Education Thesaurus	All forms of educational content	English	No details
METRO	Medical Education Taxonomy Research Organization	Processes & Methods in medical education	English	Approx.180 terms
OBEO	Ontology of Bio-Medical Educational Objectives *	Medical education (mainly learning objectives)	German	private: 72 classes, 12 object relations; Basis: Upper-Level Ontology BioTop (375 classes, 76 object relations), Toplevel- Ontology DOLCE-Lite (37 classes, 70 object relations)
TIME	Topics for Indexing Medical Education	Biomedizinische Inhalte, Medizinische Lehr-/ Lerninhalte	English, French	1,500 terms; 2,400 cross references
Polyhierarchical Ontologies for Biomedical Content References: hierarchy, (no*) association, equivalency				
Acronym	**Name**	**Topic**	**Language**	**Scope**
FMA	Foundational Model of Anatomy	Anatomy	Eng, (Ger), etc	75,000 classes; 120,000 terms; 2,100,000 cross references, 168 relationship forms
-	Gene Ontology*	Biosciences	English	34,439 terms
MeSH	Medical Subject Heading°	Biomedical Content	English, German	English: 26,142 keywords; 181,141 synonyms German: 26,142 keywords, 57,656 synonyms
NCI Thesaurus	National Cancer Institute Thesaurus*	Oncology	English	34,000 concepts, 200,000 cross references
RePORT	Research Portfolio Online Reporting Tools	Biomedical Content	English	Multiple database references (eRA databases, Medline, PubMed Central, the NIH Intramural Database, iEdison)
-	RxNorm °	Pharmaceutics	English	No specific details, but contains all US prescription medications
SNOMED-CT	Systematized Nomenclature of Medicine – Clinical Terms°	Biomedical Content, Clinics	Eng, (Ger), etc	311,000 concepts; 800,000 terms; 1,360,000 cross references
UMLS	*Unified Medical Language System*	*Biomedical Content*	*Eng, (Ger), etc*	*Semantic network with >130 single terminologies (among others°); > 1,179,00 concepts; > 4,174,00 terms*
Mono-hierarchical Taxonomies for Administration and Documentation References: hierarchy, synonyms (only #)				
Acronym	**Name**	**Topic**	**Language**	**Scope**
ICD	International Classification of Diseases [Check German Version]°	Classification of diseases and related health problems	English, German	70,000 encodings
LOINC	Logical Observation Identifier Names and Codes#°	Examination/test results from clinic/laboratory	English, (German)	7 categories; 100,000 records (German: 3,800)
OPS	Operationen- und Prozedurenschlüssel	Surgical techniques and procedures	German	6 chapters, 65 areas, approx. 28,300 classes of procedures

Table 3: Updates and support, data format, costs, and web links for vocabularies of medical content.

Polyhierarchical Ontologies for the Description of Training Formats and Methods References: hierarchy, (no*) association, equivalency					
Acronym	**Name**	**Updates & Support**	**Data Format**	**Cost**	**Web Links**
METRO	Medical Education Taxonomy Research Organization	No further development since 2005	Download – XML, Website (Blogspot)	Open Source	http://metro2.blogspot.de/
OBEO	Ontology of Bio-Medical Educational Objectives*	"work-in-progress" preliminary design (Freiburg University Hospital, Germany)	Download [Check German Version] – OWL	Open Source	http://www.imbi.uni-freiburg.de/ontology/obeo/obeo.owl
TIME	Topics for Indexing Medical Education	„work-in-progress" (University of Ottawa, Canada)	Download – XML, Textformat, Web based	Open Source	www.time-item.org (currently offline)
Polyhierarchical Ontologies for Biomedical Content References: hierarchy, (no*) association, equivalency					
Acronym	**Name**	**Updates & Support**	**Data Format**	**Cost**	**Web Links**
MeSH	Medical Subject Heading	annual (NLM,USA; DIMDI, Germany)	Download/CD-ROM – XML, CSF, Text format; Web based (MeSH-Browser)	Open Source (English, NLM registration required), fee required (German)	http://www.nlm.nih.gov/mesh/ (English) http://www.dimdi.de/static/en/klassi/mesh_umls/mesh/index.htm (German)
RePORT	Research Portfolio Online Reporting Tools	weekly (NIH, USA)	Download – XML, CSV; Web based (RePorter)	Open Source	http://report.nih.gov/
SNOMED-CT	Systematized Nomenclature of Medicine – Clinical Terms	biannual (IHTSDO, Denmark)	Download/ CD-ROM, MetamorphoSys (Java); Web based (UMLS Terminology Services (UTS))	Fee required for non-member states (Germany) (UMLS/ IHTSDO registration required)	http://www.ihtsdo.org/snomed-ct/
UMLS	*Unified Medical Language System*	*annual (NLM, USA)*	*Download/ CD-ROM, MetamorphoSys (Java); Web based (UMLS Terminology Services (UTS))*	*Open Source (NLM registration required, Copyright)*	*http://www.nlm.nih.gov/research/umls/*

The ontologies available in English that are purpose built for the domain of medical education, such as TIME and METRO, must first be converted into German. In this case, problems arise in choosing contextually equivalent translations. For instance, the English term "medical education" is equivalent simultaneously to three German terms, namely “medizinischer Ausbildung”, “Fortbildung”, and “Weiterbildung”, which individually refer to the English "graduate", "continuing” and "post-graduate" medical education.

Another problem posed by a translation is the question of whether, and how, a translation can be bound to the original version of the vocabulary in terms of potential developments and updates.

Alternatively, development of an ontology primarily based in the German language is associated with considerable time and effort.

Particularly interesting features among the identified vocabularies, summarized in Table 3, include the fact that all are available online, and at least in part, all utilize direct Internet technologies, such as the XML data format. Although they have not been primarily developed for use in the web, they are at any rate still suitable for this purpose. The technical foundations needed towards Web 3.0 are therefore established [9].

Limitations

As any literature research, this work is limited by the search and selection of the used publications.

Additionally, existence of structured vocabularies not listed in any of the publications of the searched databases cannot be ruled out.

During preparation of this work, the authors were unable to identify any generally accepted criteria for the description and comparison of structured vocabularies. Therefore, the first step was to identify any medical taxonomy described among the publications of the searched databases. This was followed by a publication analysis based on individual criteria - developed by the respective authors of each publication – used to describe the structured vocabularies. By consensus, the authors of this study then extracted the criteria that seemed most adequate for a comparison of taxonomies for medical training. During criteria selection, the authors have sought to generalize the results, making them suitable for use by German-speaking medical faculties. Alternate or additional criteria not considered in this study may be relevant for local use.

Conclusion

Content description (semantic annotation) of medical learning resources provides the ability to identify contextually related items across system boundaries. Thus far, there is no suitable taxonomy or ontology in German that may be used to implement a systematic description of resources for Germany’s medical students, thus limmiting the potential of the Web 3.0 to support curriculum mapping. Possible approaches include the new development of an appropriate structured vocabulary, the translation of existing English-language vocabularies, or a combination of existing vocabularies sections, possibly with a partial translation of English vocabularies.

Note

WB is the head of the Learning Center of the Charité, a member of the Medical Education Society (Gesellschaft für Medizinische Ausbildung - GMA), and of the GMA commitees “Practical Skills” and “Educational Research Methods.” AJ and PS are student tutors of the Learning Center, and students of the Charité. FB is a co-developer of the OBEO. OA is head of the Department of Curricular Management, a member of the GMA, as well as a member of the GMA committee “Educational Research Methods.”

Acknowledgement

The authors thank Dr. Ullrich Woermann, Bern, as well as two anonymous reviewers for their constructive and collegial critic of the manuscript and Rudi Mörgli for translating the manuscript into English.

Competing interests

The authors declare that they have no competing interests.

References

1. Holzer M, Pfähler M, Hege I, Fischer M. Wer suchet, der soll finden! - Ein Überblick über Verschlagwortung und Suche medizinischer Lerninhalte. GMS Med Inform Biom Epidemiol. 2006;2(3):Doc20. Zugänglich unter/available from: http://www.egms.de/static/de/journals/mibe/2006-2/mibe000039.shtml

2. Cook DA, Erwin PJ, Triola MM. Computerized virtual patients in Health Professions Education: a systematic review and meta-analysis. Acad Med. 2010;85(10):1589-1602. DOI: 10.1097/ACM.0b013e3181edfe13

3. Junco R, Mastrodicasa J. Connecting to the Net Generation: What Higher Education Professionals Need to Know About Today's Students. Washington, DC: National Association of Student Personnel Administrators; 2007.

4. Harden RM. AMEE Guide No. 21: Curriculum mapping: a tool for transparent and authentic teaching and learning. Med Teach. 2001;23(2):123-137. DOI: 10.1080/01421590120036547

5. Ahlers O, Georg W, Blaum W, Stieg M, Hanfler S, Bubser F, Spies C. Der Einsatz einer interdisziplinären, webbasierten Lernzielplattform verbessert sowohl die Unterrichtsqualität als auch die Klausurergebnisse Studierender. Jahrestagung der Gesellschaft für Medizinische Ausbildung (GMA). Bochum, 23.-25.09.2010. Düsseldorf: German Medical Science GMS Publishing House; 2010. Doc10gma13. DOI: 10.3205/10gma013

6. Berners-Lee T, Hendler J, Lassila O. The Semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. Sci Am. 2001;284:34-43. DOI: 10.1038/scientificamerican0501-34

7. Alby T. Web 2.0. Konzepte, Anwendungen, Technologien. München: Hanser Verlag; 2007.

8. Berners-Lee T, Fischetti M. Weaving the web: the original design and ultimate destiny of the World Wide Web by its inventors. New York: HarperBusiness; 2006.

9. Segaran T, Evans C, Taylor J. Programming the semantic web. Sebastopol, CA: O'Reilly & Associates; 2009.

10. Willett TG, Marshall KC, Broudo M, Clarke M. TIME as a generic index for outcome-based medical education. Med Teach. 2007;29(7):655–659. DOI: 10.1080/01421590701615808

11. Willett TG, Marshall KC, Broudo M, Clarke M. It's about TIME: a general-purpose taxonomy of subjects in medical education. Med Educ. 2008;42(4):432-438. DOI: 10.1111/j.1365-2923.2008.03012.x

12. Boeker M, Schober D, Schulz S, Balzer F. Ontology of Bio-Medical Educational Objectives (OBEO): ein Vorschlag für eine Ontologie medizinischer Lernziele. GMS Med Inform Biom Epidemiol. 2010;6(2):Doc11. DOI: 10.3205/mibe000111

13. Stenzhorn H, Schulz S, Boeker M, Smith B. Adapting Clinical Ontologies in Real-World Environments. J Univers Comput Sci. 2008;14(22):3767-3780.

14. Ellaway R, Evans P, McKillop J, Cameron H, Morrison J, McKenzie H, Mires G, Pippard M, Simpson J, Cumming A, Harden R, Guild S. Cross-referencing the Scottish Doctor and Tomorrow's Doctors learning outcome frameworks. Med Teach. 2007;29(7):630-635. DOI: 10.1080/01421590701316548

15. Blaum WE, Dannenberg KA, Friedrich T, Jarczewski A, Reinsch AK, Ahlers O. Der praktische Nutzen des Konsensusstatements "praktische Fertigkeiten im Medizinstudium" – eine Validierungsstudie. GMS Z Med Ausbild. 2012;29(4):Doc58. DOI: 10.3205/zma000828

16. Willett TG. Current status of curriculum mapping in Canada and the UK. Med Educ. 2008;42(8):786-793. DOI: 10.1111/j.1365-2923.2008.03093.x

17. Uschold M, Gruninger M. Ontologies: Principles, Methods and Applications. Knowl Eng Rev. 1996;11:93–155. DOI: 10.1017/S0269888900007797

Corresponding author:
Wolf E. Blaum
Charité - University Medicine Berlin, Department of Curriculum Management, Learning Center, Charitéplatz 1, 10117 Berlin, Germany, Phone: +49 (0)30/450-576003, Fax: +49 (0)30/450/576952
wolf.blaum@charite.de

Please cite as
Blaum WE, Jarczweski A, Balzer F, Stötzner P, Ahlers O. Auf dem Weg zum Web 3.0: Taxonomien und Ontologien für die medizinische Ausbildung - eine systematische Literaturrecherche. GMS Z Med Ausbild. 2013;30(1):Doc13.
DOI: 10.3205/zma000856, URN: urn:nbn:de:0183-zma0008568

This article is freely available from
http://www.egms.de/en/journals/zma/2013-30/zma000856.shtml

Received: *2003-06-29*
Revised: *2012-09-05*
Accepted: *2012-10-12*
Published: *2013-02-21*

Copyright
©2013 Blaum et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc-nd/3.0/deed.en). You are free: to Share — to copy, distribute and transmit the work, provided the original author and source are credited.