

# 4 Semantikmanagement in verteilten Systemen

Peter Haas und Robert Mützner

## 4.1 Management Summary

Gesundheitstelematische und darunter auch die telemedizinischen Lösungen basieren zum Großteil auf der semantischen Interoperabilität von Informationssystemen der beteiligten Akteure. Heute können diese Teilnehmersysteme in der Gesundheitstelematik von Enterprise-Systemen in Krankenhäusern (KIS), Medizinischen Versorgungszentren usw. über kleinere institutionelle Lösungen wie z.B. Arztpraxis- und Pflegeinformationssysteme im ambulanten Bereich bis hin zu Apps auf mobilen Geräten von Ärzten, Pflegekräften oder Patienten reichen. Kein regelrecht entworfenes und implementiertes Informationssystem kommt heute ohne eine Vielzahl von größeren und kleineren Vokabularen aus, die für die semantische Integrität und die effektivere Eingabe von Daten sorgen. Gerade auch für die medizinische Dokumentation ist eine Teilstandardisierung üblich, im Rahmen derer die Struktur und die Wertebereiche für darin enthaltene Dokumentationsattribute definiert werden. Sollen die vorgenannten Informationssysteme nun zur Unterstützung der oftmals über viele Institutionen hinweg durchzuführenden Behandlungsprozesse Daten bzw. Informationsobjekte austauschen, bedarf es neben einer sicheren technischen Infrastruktur auch der Festlegung der Struktur und Semantik dieser auszutauschenden Informationen bzw. Informationsobjekte. Da alle diese Systeme unabhängig voneinander entwickelt und in den Markt gebracht wurden, sind die strukturellen und semantischen Festlegungen darin über-

wiegend proprietär definiert bzw. werden oftmals von den einsetzenden Institutionen im Rahmen des sogenannten „Customizing“ selbst definiert. Sollen nun diese Systeme mit Systemen anderer Gesundheitsversorgungsinstitutionen Informationen austauschen, gelingt eine semantische Interoperabilität aufgrund des „Semantic Missmatches“ – also der darin festgelegten Semantik, die nicht kompatibel und auch inhaltlich nicht kongruent ist – nicht. Aus diesem Grund ist es für funktionierende gesundheitstelematische Lösungen bzw. eine nationale Gesundheitstelematikplattform mit darauf basierenden fachlogischen telematischen Anwendungen unabdingbar, dass ein systemweites Semantikmanagement betrieben wird, mittels dessen einerseits

- die Entwicklung bzw. Festlegung von attributbezogenen bzw. domänen-spezifischen Begriffssystemen, seien es nun Vokabulare, Terminologien oder Klassifikationen, in kollaborativer Weise innerhalb einer Gruppe von Fachexperten spezifiziert werden und
- andererseits das Deployment der vereinbarten Semantik für menschliche und maschinelle Akteure einfach und effektiv zu bewerkstelligen ist.

Hierzu bedarf es eines geeigneten und interoperablen Informationssystems innerhalb des verteilten Systems, das Semantik für die beteiligten Informationssysteme einerseits z.B. via Webservices maschinenles- und -verarbeitbar zur Verfügung stellt, aber auch menschlichen Benutzern die Recherche und das Browsen in der „Semantik-Welt“ ermöglicht. Dieses System bezeichnet man auch als Terminologieserver. Dieser muss eine Vielzahl von Services unterstützen, die im internationalen Standard CTS2 festgelegt sind. Ein systemweites Semantikmanagement kann somit als nationale Aufgabe angesehen werden, wenn es um nationale Anwendungen geht, aber auch bei regionalen systemübergreifenden Telematikprojekten wird dies notwendig.

## 4.2 Grundlegende Aspekte

### 4.2.1 Semantik in Informationssystemen

In Informationssystemen ist heute eine Vielzahl von größeren und kleineren attributbezogenen Vokabularen definiert, die der Sicherung der semantischen Integrität und der effektiven Eingabe von Daten dienen. Eine Darstellung der Bedeutung von Vokabularen und Begriffsordnungen für Medizinische Informationssysteme findet sich bei Haas [Haas, 2005, Kapitel 4.5 und 5.3]. Gerade auch für die medizinische Dokumentation ist eine Teilstandardisierung üblich, im Rahmen derer die Struktur der Dokumentation und die Wertebereiche für Dokumentationsattribute definiert werden.

Innerhalb von Informationssystemen und für die Betrachtung der Interoperabilität bezüglich der Definition von Kommunikationsstandards wird die Gesamtheit der Begriffe einer Anwendungsdomäne auf einzelne Attribut-bezo-

gene Vokabulare heruntergebrochen, man spricht dann von kontrollierten Vokabularen oder „(Value) Domains“.

Solche Vokabulare dienen innerhalb der Informationssysteme der Sicherstellung der semantischen Integrität der Datenhaltung, d.h. ein Benutzer darf nur erlaubte Werte eingeben bzw. das System darf nur solche im Rahmen der Kommunikation mit anderen Systemen erhaltene Angaben akzeptieren, die aus lokaler Sicht „korrekt“ sind. Auf der Richtigkeit dieser Angaben basiert dann auch die Weiterverarbeitung im lokalen Informationssystem z.B. für Abrechnungszwecke, Statistiken, Qualitätsmanagement und werteabhängige Selektionen und Sichten. „In my review ... I have pointed out that accurate representation of medical concepts or medical information is crucial to many functions of health information systems.“ [Moehr, 1998] Auch z.B. Cimino [Cimino, 1998], Ingenerf [Ingenerf, 1998] und Kuhn [Kuhn und Giuse, 2001] beschäftigten sich früh zu Beginn der Verbreitung von klinischen Informationssystemen mit diesem wichtigen Aspekt der Semantikbasierung medizinischer Informationssysteme.

Aber nicht nur bei der Verarbeitung, sondern auch bei der effektiven Erfassung von Werteausprägungen unterstützen solche Vokabulare. Sie präsentieren sich dem Benutzer an der Oberfläche als einfache Dropdown-Listen oder bei komplexeren Vokabularen innerhalb eigenständiger Auswahlfunktionen. Bei dynamisch generierten Oberflächen können sie auch als Checkboxes erscheinen, soweit die Zahl der Begriffe überschaubar ist.

Während man mit Blick auf die Summe aller dieser Vokabulare oftmals von einer Terminologie spricht – was die Gesamtheit aller Begriffe und Benennungen einer Fachsprache bezeichnet – sind die Wertebereiche in solchen Informationssystemen oftmals unabhängig voneinander und reichen von sehr kleinen Vokabularen mit wenigen Einträgen (z.B. für das Geschlecht) über mittelgroße mit bis zu hunderten von Einträgen (z.B. für die Nationalität) bis hin zu sehr umfangreichen (z.B. für die Diagnosen oder klinischen Maßnahmen). In klinischen Informationssystemen gibt es hunderte solcher kontrollierter Vokabulare.

In Anlehnung an und Erweiterung von Reiner [Reiner, 2003] können vor dem Hintergrund der Terminologienutzung in Informationssystemen folgende Klassen von Terminologien unterschieden werden:

- einfache lineare Wörterbücher mit 1 ... n charakterisierten Attributen,
- Taxonomien für die Klassifikation von Objekten,
- Wörterbücher ohne oder mit enthaltenen Taxonomien sowie
- mehrachsige Begriffssysteme.

Spezielle Ausprägungen sind Klassifikationen (Beispiel: ICD, ICPM etc.), die nur Einträge enthalten, die je eine Klasse von Begriffen repräsentieren. Klassifikationen sind zwar prominent im Gesundheitswesen präsent (ICD; ICPM), eignen sich aber für eine justiziable klinische Dokumentation in der Regel nicht, da sie zu ungenau sind.

Eine übersichtliche ontologische Darstellung der verschiedenen Begriffssystemtypen bzw. Ordnungsstrukturen in Form eines UML-Diagrammes findet sich bei Oemig in Kapitel 8 [Oemig, 2011].

Auch und vor allem im HL7-Standard gibt es z.B. eine ganze Reihe solcher Vokabulare [HL7, 2007–2014], dort werden für die Einträge folgende Attribute angegeben: Level, Type/Domain Name or Mnemonic Code, Concept ID, Mnemonic, Print Name, Definition/Description.

#### **4.2.2 Semantische Interoperabilität**

Zum Austausch von Daten und Dokumenten zwischen Informationssystemen ist die Festlegung der Syntax und Semantik der austauschbaren Informationsobjekte notwendig. Während die Syntax für Empfängersysteme die Möglichkeit schafft, die empfangenen Datensätze bzw. Informationsobjekte in ihre einzelnen Anteile zu zerlegen, um diese Teile in die Datenhaltung des Systems syntaktisch korrekt einfügen zu können, ermöglicht eine vereinbarte Semantik, diese Daten auch inhaltlich zu interpretieren und zu verarbeiten. In diesem Sinne bedeutet „semantische Interoperabilität“ nach IEEE die „Fähigkeit von zwei oder mehr Systemen oder Komponenten zum Informationsaustausch sowie zur adäquaten Nutzung der ausgetauschten Information“ [IEEE, 1991].

Was heißt nun „adäquate Nutzung“? Durch vereinbarte Semantik wird die „richtige“ Interpretation und Weiterverarbeitung im Sinne einer algorithmischen Bearbeitung – sei es zu Zwecken der Zuordnung, der Reaktion auf die empfangenen Daten usw. sowie der korrekten Speicherung der strukturell zerlegten Daten – erst möglich. Fehlen semantische Vereinbarungen, kommt es zu einem „Semantic-Mismatch“ zwischen den kommunizierenden Informationssystemen und beim Einfügen von Daten entweder zu einem Verlust der lokalen Datenintegrität oder bei Vorhandensein entsprechender Integritätsbedingungen zur Zurückweisung der Einfügung – die Kommunikation zwischen den Systemen scheitert also. Da die in den verschiedenen Systemen benutzte Semantik für bestimmte Attribute zumeist nicht kongruent ist, ist auch kein Cross-Mapping möglich. Behoben werden kann dieses Problem nur, indem systemübergreifend für die kommunikationsrelevanten Informationsobjekte auch für die bezüglich der Kommunikation relevanten Attribute entsprechende Wertebereiche in Form von Begriffsordnungen definiert und auch lokal in den Teilnehmersystemen implementiert werden. Eine einführende Präsentation zur semantischen Interoperabilität am Beispiel der Arztbriefkommunikation findet sich unter [eGesundheit.nrw, 2014].

### **4.3 Semantik in verteilten Systemen und Ist-Situation**

In verteilten Systemen – und dies sind gesundheits telematische und telemedizinische Anwendungen – arbeiten unabhängig entwickelte Informations-

systeme zur Unterstützung der einrichtungsübergreifenden Zusammenarbeit der diese Systeme betreibenden Akteure zusammen. Dies geschieht auf Basis einer sicheren Infrastruktur und vereinbarter Interoperabilitätsspezifikationen, die sodann die semantische Interoperabilität ermöglichen. Dabei kann es sich neben der direkten Punkt-zu-Punkt-Kommunikation zwischen zwei Systemen von Gesundheitsversorgungsinstitutionen – z.B. beim Austausch von Arztbriefen, Medikationsplänen etc. – auch um eine Kommunikation zwischen diesen Systemen und anderen telematischen Lösungen wie z.B. elektronischen Fall- oder Patientenaktensystemen handeln.

Die Realisierung verteilter Systeme erfordert daher ein hohes Maß an konsentierten Vereinbarungen bzw. Standards, deren Verwendung für die Teilnehmer die Investitionssicherheit ihrer Entwicklungen gewährleistet und gleichzeitig durch die Verlagerung von für alle Teilnehmer wichtigen Funktionen und Diensten auf die Infrastruktur den Herstellern von Anwendungssystemen ermöglicht, effektiv und ökonomisch die Anbindung ihres Systems an die zentrale Infrastruktur zu realisieren.

Mit Blick auf die obigen Ausführungen muss für eine verlässliche Kommunikation und Zusammenarbeit, bei der sich die Partner verstehen, vorausgesetzt werden, dass ein Begriff für alle Beteiligten auf das gleiche Objekt verweist – also die konzeptionelle Repräsentation des Begriffes gleich ist. Hierfür wird es also unabdingbar, den einzelnen Anwendungssystemen gemeinsame Begriffssysteme für kommunikationsrelevante Objektattribute zu Grunde zu legen. Benutzen alle Systeme für das gleiche Konzept auch die gleiche Bezeichnung, gelingt die semantische Interoperabilität und die semantische Integrität des verteilten Systems insgesamt kann gewährleistet werden.

Diese Vereinbarungen können auf zwei Wegen erfolgen:

- **Absprachen:** Es gibt in einer Anwendungsdomäne zwischen den Betreibern der verschiedenen interoperierenden Systeme Absprachen zur Benutzung definierter Begriffssysteme für bestimmte Sachverhalte. Jedes System implementiert die abgesprochenen Begriffssysteme lokal und sorgt für deren Aktualität. Diese Absprachen sind oftmals nur den Beteiligten bekannt bzw. zugänglich.
- **Referenz- bzw. Terminologieserver:** Im verteilten System wird ein Referenzserver für die Begriffssysteme, die für die Interoperabilität notwendig sind – oft auch als Terminologieserver bezeichnet – installiert, der die in der Anwendungsdomäne konsentierten Begriffssysteme maschinenles- und abrufbar bereitstellt. Lokale Semantik kann so automatisiert mit der globalen Semantik synchronisiert werden.

Zum jetzigen Zeitpunkt sind z.B. im Rahmen der gematik-Spezifikationen [gematik, 2014] für die medizinischen Anwendungen die in Tabelle 1 gezeigten Vokabulare festgelegt.

Tab. 1 Vokabulare in den gematik-Spezifikationen

e-Notfalldatensatz
Notfallrelevante Diagnosen
Notfallrelevante Medikamente
e-Verordnungsdatensatz
Applikationsformen (aus ABDA Stamm)
Einheiten
Sonderkennzeichen gemäß Arzneimittelvereinbarung nach § 300 SGB V
Kennzeichen für besondere Personengruppe der GKV
Versichertenart
DMP – Kennzeichnung
Rechtsgrundlage für Zuzahlungen/Selbstbeteiligungen
BtM-Merkmal
Berufsbezeichnung des Verordners
Kostenarten
Art der Verrechnungskosten des Bestandteils
Kostenträgertypen
Rezepttyp
Therapiegerechte Packungsgröße
Versichertenstatus RSA
Rechtskreis
Unfallkennzeichen

Durch die weiteren geplanten Anwendungen – wie z.B. den elektronischen Arztbrief, den elektronischen Medikationsplan, die klinischen Krebsregister oder Patientenaktensysteme – wird eine große Anzahl weiterer Vokabulare hinzukommen. Es kann aber auch über die gematik-Spezifikationen hinaus davon ausgegangen werden, dass die Gesamtzahl der heute schon für die Kommunikation im Gesundheitswesen benutzten Vokabulare durch andere von verschiedensten Akteuren und Projekten vereinbarte Datenübermittlungsverfahren (DMP, § 301, xDT-Standards, Krebsregistermeldungen usw.) bereits weit über einigen Hundert liegt. Dadurch wird es für die Softwarefirmen und die Anwender immer schwieriger, das für sie Relevante herauszufinden und zu überschauen sowie dann auch innerhalb der einzelnen Systeme korrekt zu implementieren und aktuell zu halten. Der Aufwand und damit der wirtschaftliche Schaden, aber auch die Innovationsbehinderung einer konventionellen manuellen Organisation der Herstellung von semantischer Interoperabilität, wie voran unter erstgenanntem Punkt dargestellt, in einem verteilten System mit über 100.000 beteiligten Informationssystemen in Arztpraxen,

Krankenhäusern, bei Krankenkassen usw. ist immens und steigt stetig und vermutlich exponentiell mit der Anzahl unkoordiniert eingeführter Vokabulare in verschiedenen Datenübermittlungsvereinbarungen. Schon heute zeigen sich vielfältige nicht kongruente semantische Festlegungen in den verschiedenen Interoperabilitätsspezifikationen.

Ein großes Problem besteht dabei perspektivisch auch in der Versionspflege und der Verteilung von Vokabularen und Ergänzungen/Änderungen dieser innerhalb einer Telematikplattform, die manuell bei der oben genannten Anzahl von Teilnehmersystemen kaum mehr zeitnah und adäquat ohne entsprechende Technologieunterstützung geleistet werden kann. Bei einer solchen Pflege sind auch die politischen und sachlichen Grenzen zu berücksichtigen, um eine hinreichende Stabilität der Vokabulare zu erreichen. Hinsichtlich der Versionisierung von Vokabularen sind dabei grundsätzlich zwei Arten von Änderungen zu berücksichtigen: Das Entfernen/Hinzufügen von Einträgen und das Ändern der Semantik wie z.B. durch Teilung eines Begriffes in mehrere differenziertere oder das Zusammenlegen mehrere Begriffe zu einem Begriff. Während Erstgenannte eher unproblematisch sind, bedürfen Zweitgenannte einer klaren Versionsstrategie auf Begriffsebene.

### 4.3.1 Semantikmanagement für verteilte Systeme

Unter Semantikmanagement soll hier die Entwicklung/Pflege, Bereitstellung und Verteilung von Begriffssystemen als Grundlage für die semantische Interoperabilität in verteilten Systemen verstanden werden. Während erstgenannter Aspekt zumeist unter dem Begriff „Terminologiearbeit“ subsumiert wird [FH-Köln, 2006; DIN, 2011], ist vor allem eine effektive und auch maschinenlesbare Zurverfügungstellung von Semantik in verteilten Systemen ein kritischer Erfolgsfaktor. Für eine Semantikkonsistenz im verteilten System gilt es also, durch kontrollierte Redundanz die lokale Semantik mit der Semantik im Terminologieserver zu synchronisieren, womit alle Aspekte der Integrität in Systemen, die verteilte Daten redundant halten, adressiert werden.

Heute arbeiten sehr viele Akteure in verschiedensten Kontexten und Projekten an der Spezifikation von Interoperabilitätsspezifikationen und legen im Rahmen dessen auch vielfältige Vokabulare fest. Dabei wird aufgrund fehlender Transparenz oder notwendiger aufwendiger Recherche zumeist auf die Wiederverwendung von bereits existierenden Vokabularen verzichtet, zumal damit auch die Möglichkeit der völlig freien und eigenen Semantikdefinition verloren geht.

Durch ein übergreifendes Semantikmanagement soll die Möglichkeit geschaffen werden, auf regionaler bzw. nationaler Ebene Semantikspezifikationen kohärent zu entwickeln und in verschiedensten Kontexten der Interoperabilität zu verwenden. Eine Besonderheit hierbei ist, dass der föderale Charakter dieser Terminologiearbeit nicht verloren gehen soll und darf, sondern durch eine tech-

nische Unterstützung die Transparenz existierender Definitionen sowie die Wiederverwendung gefördert werden soll. So können die einzelnen Projekte z.B. einfach recherchieren, ob es zu Attributen evtl. schon Festlegungen gibt, die benutzt werden können oder mittels Ergänzungen sinnvoll zu benutzen sind. Natürlich können auch die im Rahmen von gesetzlichen Vereinbarungen spezifizierten Kataloge enthalten sein, die dann auch zu verwenden sind.

Vor dem Hintergrund, dass sehr viele verschiedene Akteure an der Entwicklung und Nutzung beteiligt sind, bietet sich die Implementierung eines entsprechend unterstützenden IT-Systems in Form einer Webanwendung für menschliche Benutzer und in Form von Services für maschinelle Benutzer an. Dabei müssen dann auch die verschiedenen Rollen und die damit verbundenen Rechte berücksichtigt werden bzw. abgebildet werden können, denn die Semantikarbeit sollte in einem geordneten Prozess ablaufen. Detaillierte Ausführung hierzu finden sich bei Mütznier [Mütznier und Haas, 2013]. Je nach strategischer Bedeutung oder gesetzlichem Kontext oder zur operationalen Umsetzung einer Vorgabe durch die Selbstverwaltungsorgane kann ein Terminologieserver einerseits als Referenzstelle für verpflichtend zu nutzende Vokabulare, aber eben auch als Wissensbasis für die Fachszene eingesetzt werden, mittels der auch Vokabulare veröffentlicht werden können, die z.B. von Projekten entwickelt wurden und globale Bedeutung haben. Für die einzelnen Vokabulare bzw. Gruppen von Vokabularen sind domänenspezifische Expertengruppen einzurichten, die als Benutzer gemeinsam ein Vokabular entwickeln bzw. zu Änderungsanträgen Stellung nehmen. Ein ausgewiesener Benutzer sollte dabei der Vokabularverantwortliche sein. Natürlich kann eine solche Gruppe auch aus einem Projekt heraus bestehen. Für die nicht verpflichtenden Kataloge entsteht so quasi eine Art „Open Source Community“ für semantische Festlegungen im Gesundheitswesen, und für die Pflege und Weiterentwicklung können entsprechende Regelungen aus der Open Source Szene übernommen werden.

Ein wichtiger Aspekt ist dabei, dass sich auch Projekte, Hersteller etc. bezüglich der Nutzung von Vokabularen „akkreditieren“ können sollten, also darüber eine Cross-Reference möglich ist, wer wo für welchen Zweck bzw. Kontext ein Vokabular nutzt. Dies ist auch für die Pflege wichtig, um evtl. bei Neuerungen Seiteneffekte berücksichtigen, entsprechende Mitteilungen generieren oder die Nutzer sogar aktiv an der Weiterentwicklung beteiligen zu können. Auch eine Crossreference zwischen Vokabularen und den Interoperabilitätsvereinbarungen, in denen diese genutzt werden, ist wichtig zu führen, um bei der Weiterentwicklung auf Auswirkungen auf bestehende Interoperabilitätsvereinbarungen achten sowie Entwickler bei der Recherche effektiv unterstützen zu können. Möchte also ein Hersteller eine Interoperabilitätsvereinbarung umsetzen, erhält er auf Knopfdruck die dafür spezifizierte Semantik – auch direkt maschinenlesbar zum Import in das Produkt bzw. die Kundeninstanzen inklusive der laufenden Synchronisation.

## Lösungsansatz Überblick

Gefördert durch das Bundesministerium für Gesundheit und in der Folge durch die Europäische Union und das Land Nordrhein-Westfalen wurde ein auf CTS2 (Common Terminology Services) basierter **Terminologieserver** mit angeflanschter **Kollaborationsumgebung** realisiert. Den Gesamtansatz zeigt schematisch Abbildung 1.

Generelle Aufgabe des Terminologieservers ist die rechnergestützte Repräsentation bzw. Vorhaltung (medizinischer) Terminologien sowie das Bereitstellen von Diensten zum Recherchieren und Abrufen von Terminologien/Terminologieausschnitten bzw. attributbezogenen Vokabularen sowohl für menschliche als auch maschinelle Benutzer. Darüber hinaus kann dieser auch als Basis zur kollaborativen Entwicklung von Vokabularen/Terminologien eingesetzt werden.

Als Komponente einer nationalen Gesundheitstelematikplattform stellt er die semantische Homogenität aller attributbezogenen Vokabulare innerhalb der Gesundheitstelematikplattform sicher und ist damit eine wesentliche Voraussetzung für die Entwicklung und den wirtschaftlichen und integren Betrieb der gesundheitstelematischen Anwendungen. Durch eine telematikplattformweite Vereinbarung und Bereitstellung von attributbezogenen Vokabularen bzw. Kodierschemata, Nomenklaturen, Klassifikationen, Taxonomien kann die semantische Interoperabilität zwischen den einzelnen Anwendungssystemen sichergestellt werden. Ziel ist hierbei, dass für alle Inhaltskonzepte und Angaben in der Dokumentation und in den Nachrichten die gleichen Benennungen innerhalb des verteilten Systems benutzt werden.

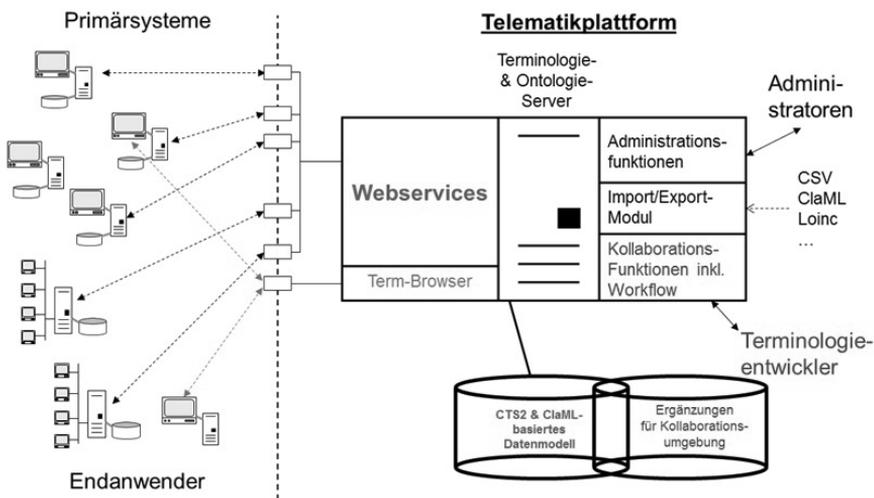


Abb. 1 Wesentliche Komponenten des Lösungsansatzes

Die Datenhaltung für den Terminologieserver wurde aus dem Standard CTS übernommen und arrondiert. Die Datenhaltung für die Kollaborationsumgebung wurde auf Basis einer umfassenden Anforderungsanalyse im Rahmen einer Master-Thesis entworfen und implementiert. Der Lösungsansatz besteht somit aus folgenden funktionalen Komponenten:

- **Ein Terminologiebrowser:** Für die Dialognutzer wurde ein webbasierter Terminologiebrowser implementiert, der für angemeldete Benutzer mit entsprechenden Rechten auch das direkte Eingeben, Ändern und Löschen von Inhalten ermöglicht. Um einen entsprechenden übersichtlichen Zugang zu den vielen enthaltenen Begriffssystemen zu haben, wurde eine frei parametrierbare Taxonomie implementiert, die quasi als hierarchisches Menü dient, um zu bestimmten Begriffssystemen zu navigieren. Der Browser nutzt die Services auch für den Zugriff auf die Daten.
- **Ein Importmodul:** Mittels des Importmoduls können initial Begriffssysteme in die Datenhaltung geladen werden. Neben einfachen CSV-Dateien wird auch der Standard ClAML unterstützt. Für LOINC musste eine spezielle Erweiterung des Importmoduls implementiert werden.
- **Ein Administrationsmodul:** Mit den Funktionen des Administrationsmoduls können neben der Verwaltung von Benutzern und Rechten sowie von Lizenzen auch eine ganze Reihe systemspezifischer Einstellungen verwaltet werden.
- **Ein Kollaborationsmodul:** Im Rahmen des angeflanschten Kollaborationsmoduls kann der Workflow der Terminologiearbeit rechnergestützt gesteuert werden. Integriert ist hier eine elektronische Diskussionsumgebung. Die zu beachtenden Workflows können vokabularspezifisch parametrisiert werden, da je nach Hintergrund verschiedene Abläufe zu beachten sind.

#### 4.3.2 IT-gestütztes Deployment – Terminologieserver

Zu Beginn der Entwicklung stand die Frage der nutzbaren Vorarbeiten. Hier zeigte sich schnell, dass die damals neu erschienene CTS2-Spezifikation eine gute Basis darstellt. CTS2 ist eine HL7-OMG-Spezifikation, im Rahmen derer ein Klassenmodell sowie die notwendigen Dienste spezifiziert sind, um Terminologien und Ontologien z.B. mittels eines Terminologieservers verwalten und diese in verteilten Umgebungen via Webservices verfügbar machen zu können. CTS2 ist eine Weiterentwicklung von CTS 1.0. Neu hinzugekommen ist beispielsweise ein Versionisierungskonzept. Aufgeteilt ist der Standard in zwei Modelle, die im Folgenden kurz beschrieben werden.

##### Conceptual Model

Das Conceptual Model ist ein Klassenmodell zur Verwaltung von Begriffssystemen mit hoher Generizität. Damit wird die Verwaltung von fast beliebigen

Vokabularen, Konzepten und deren Beziehungen, Cross-Mappings und Value Sets ermöglicht. Auf oberster Ebene aggregieren die einzelnen Vokabulare mehrere unterschiedliche Konzepte. Diese Konzepte können einzelne Begriffe einer Terminologie, Begriffskomplexe oder Beziehungen sein. Jedes Konzept kann in Beziehung mit anderen Konzepten stehen. Dabei ist es irrelevant, ob sich die in Beziehung gesetzten Konzepte im gleichen Ordnungssystem befinden oder nicht. U.a. ist damit auch ein Cross-Mapping zwischen Begriffssystemen möglich. Für alle wesentlichen Informationsobjekttypen (Vokabulare, Konzepte, Beziehungen, Value Sets) ist eine Versionierung möglich. Ein Vokabular kann so in Echtzeit weiterentwickelt werden, um zu einem bestimmten Zeitpunkt in eine komplett neue Vokabular-Version überführt zu werden.

Eine Besonderheit ist auch, dass Value Sets konzeptionell schon berücksichtigt wurden. Value Sets fassen Konzepte aus einem oder mehreren Ordnungssystemen zusammen, sie bilden quasi einen selektiven View auf ein Ordnungssystem bzw. eine Vereinigungsmenge vieler solcher Views. Die Komplexität eines Value Sets kann von einer flachen Liste von Konzeptcodes von einem Ordnungssystem bis hin zu einer unbegrenzten, hierarchischen Sammlung von Konzepten aus unterschiedlichen Ordnungssystemen reichen. Dabei gilt zu beachten, dass jede Sammlung von Konzepten in einem Value Set einzigartig bleiben sollte.

### Functional Model

Das Functional Model spezifiziert die Dienste, welche ein Terminologieserver anbieten sollte. Die verfügbaren Dienste werden in die vier Klassen bzw. Szenarien eingeteilt: Administrative Szenarien (Administration), Zugriff-Szenarien (Search), Pflege-Szenarien (Authoring) und Konzept-Szenarien (Association), was eine Übersicht zu der Vielzahl von Diensten erleichtert.

Die Aufteilung der Dienste in die vier Dienstklassen bietet den Vorteil der Modularisierbarkeit. Unterschiedliche Benutzer mit einzelnen Rollen müssen nur die für sie wichtigen Dienste verwenden. Die Last der Anfragen kann auf verschiedene Services verteilt werden, so dass der Betrieb des Terminologieservers performanter wird. Jede Dienstklasse hat dabei unterschiedliche Zugriffsrechte, so dass Benutzer auch nur für bestimmte Klassen zugelassen werden können.

Die Entwicklung des Terminologieservers erfolgte in den folgenden Schritten:

- Abbildung des Klassenmodells im Modellierungstool Enterprise Architect von Sparx Systems
- Arrondierung des Modells um zusätzliche wichtige Aspekte
- Generierung der MySQL-Datenbank von ORACLE
- Implementierung der Services mittels JAVA
- Implementierung des Import-Moduls mittels JAVA

- Implementierung des Browsers mittels Open Source Developer-Framework ZKOSS
- Implementierung der Administrationsumgebung

Dabei wurden auch Performance-Tests mit umfangreichen Ordnungssystemen durchgeführt, um das Zeitverhalten des Zusammenspiels zwischen Services und Datenbank zu überprüfen, zumal als Zugriffsschicht Hibernate zum Einsatz kommt. Hierbei ergaben sich keine kritischen Aspekte.

Das CTS-Datenmodell musste im Wesentlichen um Login-Informationen für eine Lizenzüberprüfung sowie die Lizenzinformationen selbst und um Beziehungstypen für die Unterscheidung zwischen ontologischen, taxonomischen oder cross-mapping-Beziehungen erweitert werden. Zusätzlich wurde eine einfache Möglichkeit der Verwaltung von Übersetzungen hinzugefügt, um diese nicht nur via Cross-Mapping zwischen Sprachversionen realisieren zu können. Damit ist der implementierte Terminologieserver auch multilingual. Die Dienste wurden um die beschriebenen Änderungen bzw. Erweiterungen modifiziert. Dazu entstand ein neues Szenario „Authorization“, welches die Anmeldung am System abbildet. Für das Qualitätsmanagement entstand ein weiteres Szenario: „Reporting“. So lassen sich in generischer Weise Statistiken erstellen.

Die servicebasierte Interoperabilität zwischen Terminologieserver und Primärsystemen wurde anhand zweier konkreter Anbindungen für Vokabulare der eGK-Notfalldaten an ein KIS und an ein Arztpraxisinformationssystem gezeigt. Es handelt sich dabei um das ClinicCentre von iSOFT sowie das Arztpraxisystem der Duria eG. Zudem wurde der Terminologieserver in die FH-eigene Web-Krankenakte ophEPA erfolgreich eingebaut. Ein beispielhafter Ablauf für die kaskadierend tiefergehende Recherche von Inhalt sieht wie folgt aufgelistet aus:

1. Die Methode „ListCodeSystems()“ wird aufgerufen, um eine Liste mit allen im Terminologieserver verfügbaren Vokabularen zu erhalten (ohne Anmeldung nur lizenzfreie)
  - Aufruf: es werden keine Parameter mitgegeben
  - Antwort: eine Liste der verfügbaren Ordnungssysteme
2. Die Methode „ListCodeSystemConcepts()“ wird aufgerufen, um alle Konzepte des Vokabulars „ICD 10-GM“ abzurufen
  - Aufruf: die Vokabular-ID 11 wird mitgegeben (ICD 10-GM, erhalten aus dem vorherigen Aufruf)
  - Antwort: eine Liste mit verfügbaren Konzepten des ICD 10-GM
3. Die Methode „ReturnConceptDetails()“ wird aufgerufen, um Details zu dem Begriff „Blutegelbefall o.n.A.“ abzurufen
  - Aufruf: die Konzept-ID 29603 wird mitgegeben (ID erhalten aus dem vorherigen Aufruf)
  - Antwort: alle Details zum Konzept „Blutegelbefall o.n.A.“

Das wichtigste Szenario für den alltäglichen Gebrauch bzw. die Sicherstellung der kontrollierten Redundanz der relevanten Semantik in den Primärsystemen ist jedoch die Synchronisation der lokalen mit der globalen Semantik bezogen auf ein oder mehrere Vokabulare. Ein sinnvolles Verfahren hierzu ist es, dass ein Primärsystem zu festgelegten Zeiten (z.B. immer nachts um 23 Uhr) oder zyklisch (z.B. jede Stunde) eine Liste der Konzepte aller lokal relevanten Ordnungssysteme abrufen, deren Änderungsdatum größer als der letzte Abrufzeitpunkt ist, und dann für die Einträge in dieser Liste die Details nachlädt. Aufgrund der geringen Dynamik solcher Veränderungen sind Rückgabemenge und resultierende Anfragelast für den Terminologieserver eher als gering einzuschätzen. So kann jedoch dann in einfacher Weise auch bei zig- bis hunderttausenden von Teilnehmersystemen neue oder zu ändernde Semantik einfach automatisch in die Fläche deployed werden. Im Sinne einer Lastreduktion des Terminologieservers effektiviert werden kann dieser Mechanismus, indem erst einmal nur je Codesystem abgefragt werden kann, wann die letzte Änderung darin stattgefunden hat.

In der Regel sollten die Primärsysteme in ihrer lokalen Vokabularverwaltung die Primärschlüssel der abgerufenen Objekte des Terminologieservers auf jeder Granularitätsstufe in der internen Datenhaltung mitführen.

Hinsichtlich der Implementierung der 32 Webservices mussten diese zuerst „geschnitten“ werden, da hierzu der Standard keine Festlegungen trifft. Dabei stellt für spätere Anwender bzw. Entwickler von Primärsystemen, die ihr System an den TS anbinden wollen, eine Dienste-Attribut-Matrix eine wichtige Hilfe dar. Sie enthält auf der Horizontalen die in CTS2 spezifizierten Dienste, auf der Vertikalen die Datentypen und den logischen Zusammenhang zwischen Attributen und Diensten. In dieser Matrix werden die Kardinalitäten eingetragen. So sieht ein Entwickler in übersichtlicher Weise, welcher einzelne Dienst welche Muss- und Kann-Angaben beim Request erwartet und welche Ergebnisse dieser im Rahmen der Response-Nachricht zurückliefert. So wurden alle CTS2-Dienste in diese Matrix übertragen, den Datenbankstrukturen gegenübergestellt und die Zellen dann mit Inhalten gefüllt. Diese Matrix war sodann die Basis für die Entwicklung der Programme für die Dienste. Um diesen Entwurfsprozess aufgrund der Größe der Matrix (ca. 32 \* 100 Zellen) zu vereinfachen, wurde eine entsprechende Anwendung implementiert (s. Abb. 2), mittels der durch Drag & Drop der Diensteszuschnitt vorgenommen wurde, mit dem Vorteil, dass aus der Datenhaltung dieser Anwendung auch die Dokumentation erzeugt werden konnte.

Auf Basis des Datenmodells, der CTS2-Dienste sowie der oben beschriebenen Matrix wurde der Terminologieserver in der Programmiersprache Java umgesetzt. Dabei gibt die Dienste-Attribut-Matrix vor, welche Attribute welchem Service mitgegeben werden. Folgende Services wurden implementiert (Anzahl der Methoden):

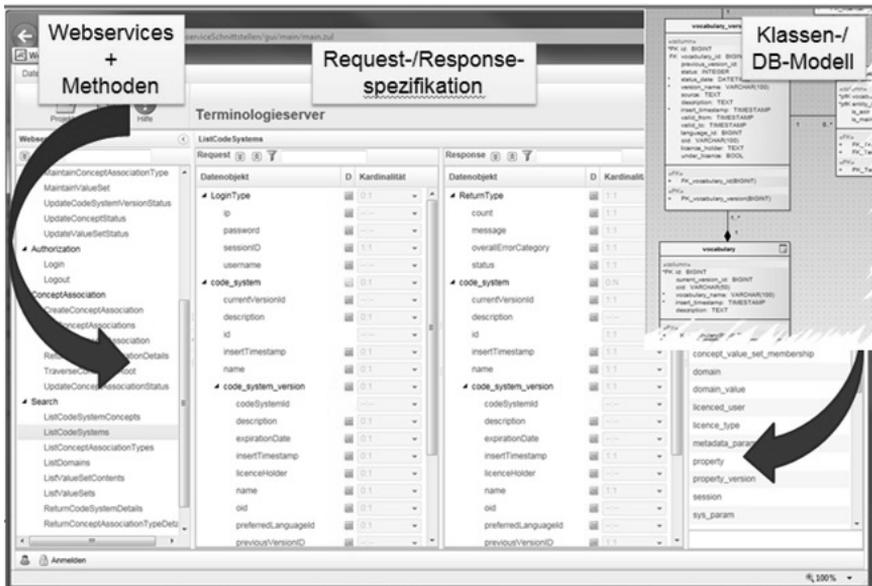


Abb. 2 Implementiertes Tool für den Dienstzuschnitt

- Administration (5)
- Authoring (12)
- Authorization (2)
- ConceptAssociation (6)
- Search (10)

Die detaillierte Dokumentation der Services ist öffentlich verfügbar und findet sich unter [FH-Dortmund, 2012–2014]. Auch die Services selbst und der Terminologiebrowser sind derzeit öffentlich erreichbar, für Änderungen der Inhalte ist aber ein Login notwendig.

Im Browser können interaktiv die Inhalte abgerufen bzw. durchsucht werden (s. Abb. 3).

Linksseitig erscheint die parametrisierte Taxonomie, wobei ein Ordnungssystem auch mehreren Klassen zugeordnet werden kann, rechts erscheinen dann die Inhalte zum links aktivierten Ordnungssystem/Vokabular. Mittels Kontextmenü können zu allen Einträgen die Detailinformationen abgerufen werden, ebenso ist eine Suche nach Begriffen in einem oder in allen Ordnungssystemen auch mittels Teilstring möglich. Bei hierarchischen Ordnungssystemen ist optional auch die Anzeige der darüber liegenden Hierarchieknoten bis hin zum Wurzelknoten möglich, um den Kontext des Begriffes in der Hierarchie erkennen zu können (s. Abb. 4).

Ist ein Benutzer angemeldet und hat die Rechte zum Einfügen, Ändern und Löschen von Einträgen, erscheinen zusätzlich die entsprechenden dazu notwendigen Interaktionselemente wie Buttons und Kontextmenü-Einträge.

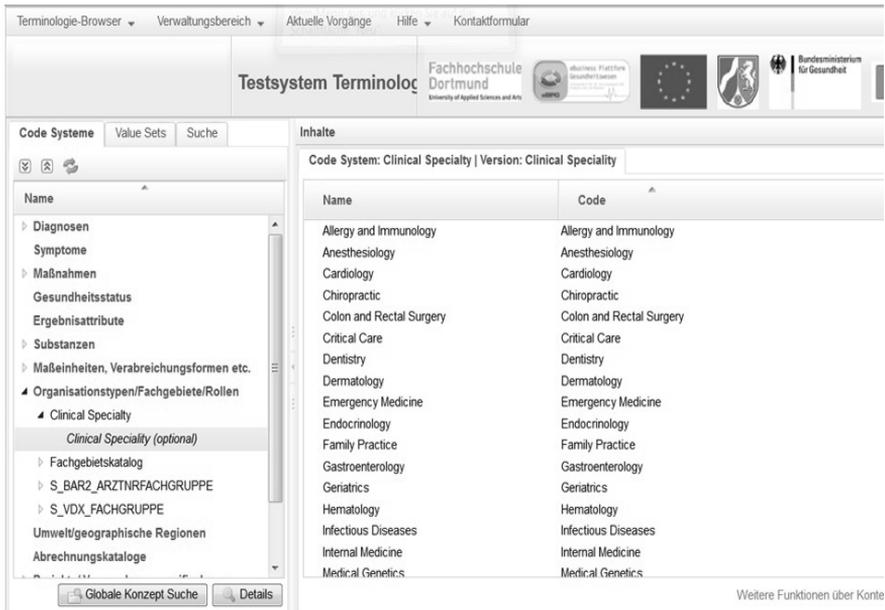


Abb. 3 Terminologiebrowser Bildschirmauszug

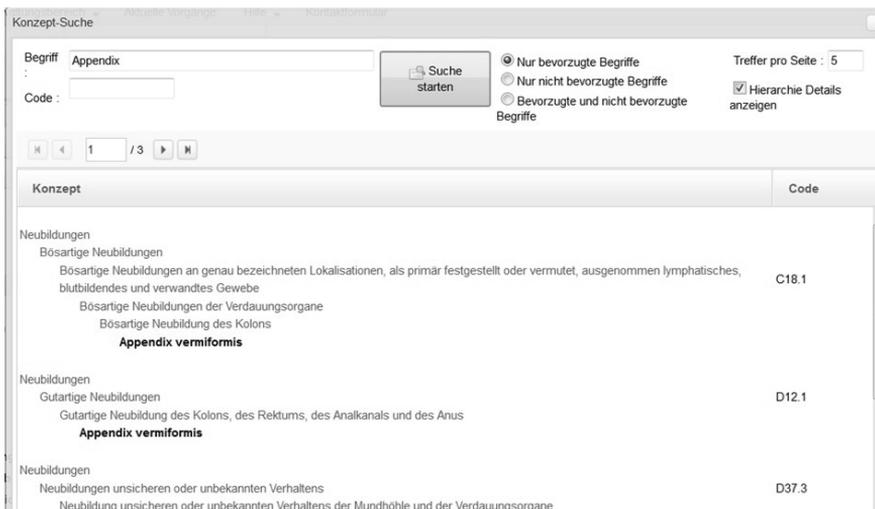


Abb. 4 Auszug Suchergebnis mit Hierarchietraversierungsanzeige

### 4.3.3 IT-gestützte Entwicklung und Pflege von Semantik, Entwicklung und Pflege-Kollaborationsanwendung

Wichtig für die Entwicklung von gesundheitstelematischen Anwendungen und deren Interoperabilität ist auch die Organisation und Optimierung der Terminologiearbeit. Die Organisation und Verwaltung von Begriffssystemen

sollte in der Regel durch domänenspezifische Gruppen von räumlich und zeitlich getrennt tätigen Experten erfolgen, was zu der Anforderung führt, dass es eine Plattform geben muss, mit deren Hilfe Terminologieentwickler orts- und zeitunabhängig (zusammen-)arbeiten, aber auch nicht zur Gruppe gehörende Branchenmitglieder Anträge auf Ergänzungen und Änderungen eines Vokabulars stellen können. Diese Forderung kann ebenfalls von einer webbasierten Anwendung erfüllt werden, wobei die Dienste des TS auch als Grundlage für eine solche Kollaborationsumgebung verwendet werden können. Dabei ist es notwendig, dass der Objektlebenszyklus von Vokabularen und zugehörigen Konzepten spezifisch abgebildet, also im System parametrisiert werden kann und zur Steuerung des Workflows dient. Den Gesamtzusammenhang zeigt Abbildung 5, aus [Mützner und Haas, 2013].

Der Statusgraph dient dabei zur Steuerung der für die einzelnen Benutzer je nach Rolle verfügbaren Funktionalitäten bezüglich der Weiterbearbeitung. Am Ende steht die Entscheidung, ob ein Vorschlag wie beantragt übernommen, modifiziert übernommen oder abgelehnt wird.

Die implementierte Kollaborationsumgebung bietet somit workflowbasierte Funktionalitäten zur Beantragung, Diskussion, Entscheidungsfindung und Freigabe von Neueinträgen oder Änderungen. Dazu wurde das CTS2-Datenmodell des TS um die notwendigen Objekttypen für diese Funktionalitäten erweitert, sodass z.B. für die Verwaltung von Vorschlägen, Diskussionslots,

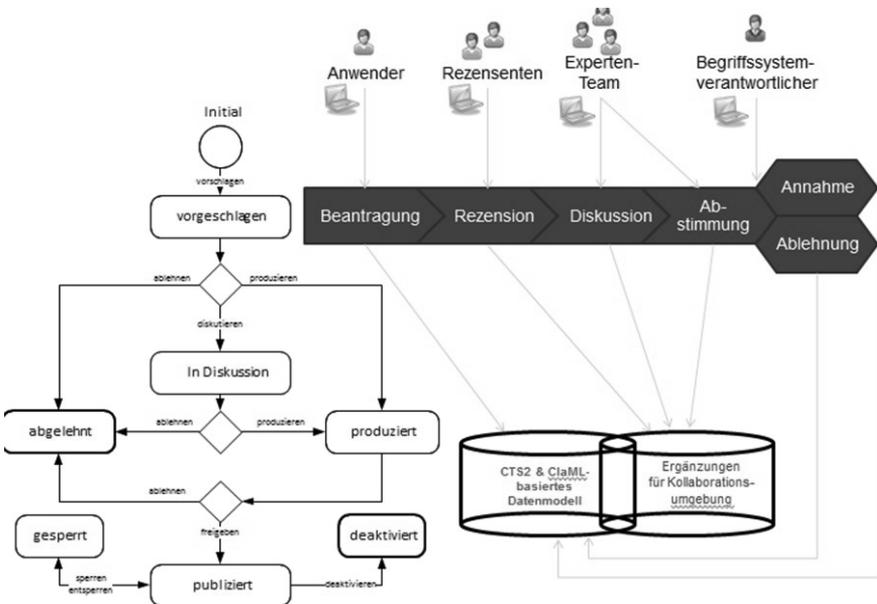


Abb. 5 Beispielhafter Objektlebenszyklus und Entwicklungsphasen eines Ergänzungs- oder Änderungsvorschlages, der die generellen Phasen durchläuft

Abstimmungen und die Workflowparametrierung entsprechende Objekttypen bzw. Datenbanktabellen existieren.

Entsprechende Vorschlags- und Bearbeitungsverfahren existieren auch beim DIMDI [DIMDI, 2014], das bei der Spezifikation der Kollaborationsumgebung unterstützt hat.

Eine weitergehende Darstellung der Kollaborationsumgebung findet sich bei Mützner [Mützner und Haas, 2013].

#### 4.4 Zusammenfassung und Ausblick

Die Zurverfügungstellung bzw. Verteilung von Semantik in verteilten Systemen ist für die semantische Interoperabilität und damit für einen wertschöpfenden Betrieb interoperierender Informationssysteme unabdingbar. Hierfür bieten sich Terminologieserver an, für die durch CTS2 ein internationaler Standard vorliegt. Zur Implementierung wurde im konkreten Fall das Klassenmodell des Standards in ein Datenbankmodell überführt und mittels einer Diensteschnittstellen-Matrix die Diensteschnittstellen spezifiziert und implementiert. Für einen professionellen Betrieb wurden kleinere Erweiterungen des CTS2-Klassenmodells notwendig. Der realisierte Lösungsansatz wurde anhand eines Szenarios (Nofalldatenvokabulare) mit zwei Primärsystemen evaluiert und kann als gut praktikabel bezeichnet werden. Aufbauend auf den Services wurde eine Webanwendung implementiert, mittels der ein Navigieren und das Suchen in enthaltenen Ordnungssystemen möglich sind. Zusätzlich können berechtigte Benutzer auch interaktiv Eintragungen und Änderungen direkt vornehmen. Ergänzend wurde eine Kollaborationsanwendung implementiert, mittels der Gruppen von Fachexperten die Begriffssysteme gemeinsam auf Basis parametrierbarer Workflows (weiter-)entwickeln und pflegen können, wobei mittels eines Diskussionsforums auch Vorschläge zuerst einmal diskutiert werden können, bevor es zu einer Abstimmung bzw. Entscheidung kommt. Damit stehen die zwei wesentlichen Anwendungen zur Verfügung, um ein Semantikmanagement für verteilte Systeme zu betreiben.

Die Entwicklungen wurden durch das Bundesministerium für Gesundheit, das Land Nordrhein-Westfalen und die Europäische Union gefördert.

#### Referenzen

- [Cimino, J.J., 1998] Cimino, J.J. (1998): Desiderata for controlled medical vocabularies in the twenty-first century. *Methods Inf Med* 37 (4-5): S. 394-403.
- [DIMDI, 2014] DIMDI. Vorschlagsverfahren zur Pflege von OPS und ICD-10-GM für Zwecke des G-DRG-Systems. <http://www.dimdi.de/static/de/klassi/ops/vorschlagsverfahren/index.htm> Letzter Zugang: 2014-08-12.
- [DIN, 2011] Deutsches Institut für Normung (2011): DIN 3242 – Begriffe der Terminologielehre.
- [eGesundheit.nrw, 2014] eGesundheit.nrw. Aktuelles + Trends eGesundheit. <http://egesundheit.nrw.de/aktuelles-trends/> Letzter Zugang: 2014-08-12.

- [FH-Dortmund, 2012–2014] FH-Dortmund. Termserver-CTS2. <http://www.wiki.mi.fh-dortmund.de/cts2/index.php?title=Hauptseite> Letzter Zugang: 2014-08-25.
- [FH-Köln, 2006] FH-Köln. Terminologiarbeit und Terminologierecherche. [http://www.termportal.de/eLearning/10\\_recherche.html](http://www.termportal.de/eLearning/10_recherche.html) Letzter Zugang: 2014-08-12.
- [gematik, 2014] gematik. gematik – Gesellschaft für Telematikanwendungen der Gesundheitsakte mbH. <http://www.gematik.de> Letzter Zugang: 2014-08-12.
- [Haas, P., 2005] Haas, P. (2005): Medizinische Informationssysteme und Elektronische Krankenakten. (1. Ed.). Springer.
- [HL7, 2007–2014] HL7. Health Level Seven International – Vocabulary. <http://www.hl7.org/special/committees/vocab/vocabresources.cfm> Letzter Zugang: 2014-08-12.
- [IEEE, 1991] Institute of Electrical and Electronics Engineers (1991): IEEE Standard Computer Dictionary: A Compilation of IEEE Standard Computer Glossaries.
- [Ingenerf, J., 1998] Ingenerf, J. (1998): Interoperabilität zwischen medizinischen Anwendungssystemen. *Informatik, Biometrie und Epidemiologie in Medizin und Biologie* 29 (1): S. 69–76.
- [Kuhn, K.A. und Giuse, D.A., 2001] Kuhn, K.A. und Giuse, D.A. (2001): From hospital information systems to health information systems. Problems, challenges, perspectives. *Methods Inf Med* 40 (4): S. 275–287.
- [Moehr, J.R., 1998] Moehr, J.R. (1998): Informatics in the service of health, a look to the future. *Methods Inf Med* 37 (2): S. 165–170.
- [Mützner, R. und Haas, P., 2013] Mützner, R. und Haas, P. (2013): Eine Kollaborationsumgebung zur Entwicklung und Pflege von semantischen Bezugssystemen In: Haas, P. et al., Hrsg. Nutzung, Nutzer, Nutzen von Telematik in der Gesundheitsversorgung – Eine Standortbestimmung. Tagungsband TELEMED 2013. S. 135–145. Akademische Verlagsgesellschaft – AKA.
- [Oemig, F., 2011] Oemig, F. (2011): Entwicklung einer ontologiebasierten Architektur zur Sicherung semantischer Interoperabilität zwischen Kommunikationsstandards im Gesundheitswesen. Hochschulschrift der Universität Regensburg (Dissertation) [http://epub.uni-regensburg.de/20076/1/Dissertation\\_v39\\_final.pdf](http://epub.uni-regensburg.de/20076/1/Dissertation_v39_final.pdf) Letzter Zugang: 2014-08-12.
- [Reiner, J.L., 2003] Reiner, J.L. (2003): Terminologiesysteme als Grundlagen für die semantische Interoperabilität von heterogenen Anwendungssystemen sowie deren fachsprachlicher Zugang. (1. Ed.). Akademische Verlagsgesellschaft AKA.