

Definition und Implementierung einer CTS2-standardisierten Abbildung von Terminologien aus dem Bereich des Infektionsschutzes

Franziska Krebs

Bachelorarbeit • Studiengang Medizininformatik • Fachbereich Informatik und Medien • 02.09.2013

Aufgabenstellung

Ziel der Arbeit ist es, Begriffssysteme aus dem Bereich des Infektionsschutzes gemäß des von OMG und HL7 geschaffenen Standards *Common Terminology Services (CTS2)* [1] abzubilden. Eine Anforderungsanalyse soll Fragen festlegen, welche eine Bewertung der entstandenen Terminologiestruktur ermöglichen.

Konzept

Die Datenmenge, welche vom Robert Koch-Institut zur Verfügung gestellt wird, wird zunächst inhaltlich und strukturell untersucht. Dabei soll herausgearbeitet werden, welche Konzepte beschrieben werden und ob diese gemäß des Standards repräsentiert werden können. Nach der Auswahl einer Untermenge, die die Beantwortung der festgelegten Fragen erlaubt, erfolgt die Definition der Abbildung. Diese wird mithilfe einer in Java programmierten Klasse implementiert.

Technologien und Systeme

Der CTS2-Standard bietet eine Art Schablone zur Strukturierung von terminologischen Inhalten jeglicher Art - von einfachen Codelisten bis hin zu komplexen Ontologien. Wird eine Terminologie dementsprechend umgesetzt, ist eine einheitliche Verwaltung und somit auch ein einheitlicher Zugriff ermöglicht. Um dies zu erreichen, muss ein sog. *Information Model* plattformspezifisch repräsentiert werden. Die Abbildung 1 visualisiert einen Ausschnitt des Modells. Dieses spezifiziert Elemente, die in allen Terminologien gemeinsam vorkommen, beispielsweise sog. *ClassDescriptions* und *PredicateDescriptions* zur Abbildung von Konzepten und Beziehungen zwischen diesen Konzepten.

Zentrales Werkzeug bei der Bewältigung der Aufgabenstellung ist das vom Fraunhofer FOKUS entwickelte System *CTS2-Le* [2], welches eine CTS2-konforme Implementation mithilfe semantischer Technologien präsentiert. Zur Abbildung des Information Models werden das *Resource Description Framework (RDF)* und das Schema-Vokabular *RDFS* verwendet. Verschiedene Komponenten erleichtern das Erstellen und Speichern von RDF-Graphen und ermöglichen eine Validierung des hinzugefügten Wissens gegen das durch das Information Model gegebene Schema.

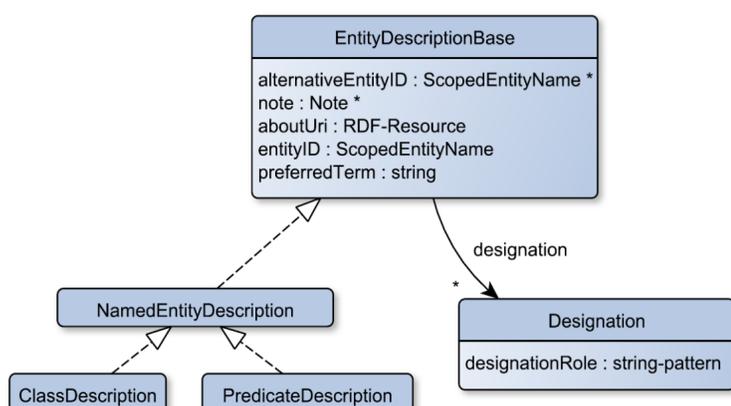


Abb. 1: Auszug aus dem Information Model der CTS2-Spezifikation

Definition und Implementierung der Abbildung

Die Terminologien beschreiben verschiedene Konzepte aus dem Bereich des Infektionsschutzes, so beispielsweise meldepflichtige Krankheiten, deren Erreger, Symptome, Nachweismethoden sowie Impfstoffe.

Es erfolgte eine Auswahl von geeigneten CTS2-Elementen zur Repräsentation der einzelnen Begriffe sowie die Zuweisung von Identifikationsnummern, Bezeichnern u.ä. zu durch den CTS2-Standard spezifizierten Attributen. Anschließend werden die Begriffe durch eigens definierte Relationen verbunden. Beispielsweise werden die Begriffe „Botulismus“, „Milzbrand“ und „Dyspnoe“ und die Relation „hat Symptom“ so zusammengesetzt, dass Aussagen der Form „Botulismus hat das Symptom Dyspnoe“ und „Milzbrand hat das Symptom Dyspnoe“ formuliert werden können. Diese Vernetzung zeigt die Abbildung 2.

Mithilfe einer Java-Klasse erfolgt die Verarbeitung der Datenmenge und die Erstellung des RDF-Graphen unter Zugriff auf die von CTS2-Le zur Verfügung gestellten Funktionen. Im Ergebnis ergibt das ein Wissensnetz, in dem jeder Knoten und jede Kante als Instanz einer CTS2-Entität angesehen werden kann.

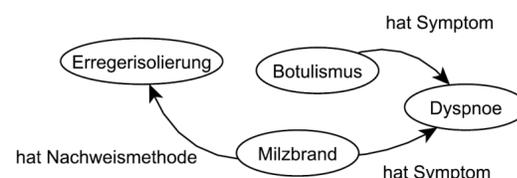


Abb. 2: Demonstration der Vernetzung mithilfe von Relationen

Ergebnisse

Zum Zwecke der Evaluierung der Ergebnisstruktur erfolgte die Formulierung der Fragen mithilfe von *SPARQL*. Diese überprüfen hauptsächlich die Korrektheit der Relationen und konnten erfolgreich beantwortet werden.

Ausgehend von den durch die Quelldaten gelieferten Informationen konnten ebenfalls Anfragen an benachbarte Terminologien erfolgreich realisiert werden. So wurden anhand des ICD-Codes einer Krankheit Inhalte des Diagnoseklassifikationssystems *ICD-10* erfragt.

Fazit

Die im Rahmen der Arbeit umgesetzte Abbildung zeigt deutlich den Nutzen des CTS2-Standards in Verbindung mit semantischen Technologien zur Entwicklung und Verwaltung einer Terminologie. Weiterhin wurde demonstriert, dass ohne aufwändige Mapping-Verfahren Wissen über Konzepte in benachbarten Codesystemen erfragt werden kann.

Quellen

[1] Mayo Clinic (Hrsg.): *CTS2 Common Terminology Services 2*, http://informatics.mayo.edu/cts2/index.php/Main_Page - letzter Zugriff am 05.09.13

[2] Billig, A.: *Utilizing Semantic Technologies for a CTS2 Store*. <http://semantik.fokus.fraunhofer.de/WebCts2LE/main3/ini.jsp> - Version: 2013. – Fraunhofer FOKUS, CC E-HEALTH; letzter Zugriff am 08.09.13