

# Konzeption und prototypische Implementierung eines Softwaretools zur dynamischen Erstellung von Befundberichten mit Hilfe ontologiebasierter Methoden

Sebastian Busse

Masterarbeit • Studiengang Informatik • Fachbereich Informatik und Medien • 02.03.2017

## Aufgabenstellung

Als Ergebnis dieser Arbeit soll ein Software-Prototyp stehen, mit dessen Hilfe pathologische Befundberichte im Diskursbereich von Endometrium-, Haut- und Prostatakrebs beispielhaft dynamisch, strukturiert und vollständig erstellt werden können. Dabei soll sowohl für den Aufbau der Berichte als auch für deren Vollständigkeits-Check eine speziell konzipierte Ontologie genutzt werden.

## Anforderungen

Motiviert durch die stetig wachsende Digitalisierung im klinischen Umfeld soll das Softwaretool die automatisierte Verarbeitung von Befundberichten erleichtern und zu diesem Zweck folgende Möglichkeiten bieten:

- Einheitliche Strukturierung der erstellten Befundberichte und einheitliches Speicherformat
- Maschinenverständliche Repräsentation der Berichtsinhalte
  - Globale dokumentenunabhängige Codierung durch eindeutige Identifikationsnummern (URI-Adressen)
  - Automatisierte Überprüfung der inhaltlichen Vollständigkeit von Befundberichten gemäß Vorgaben eines internationalen Expertenteams
  - Erleichterung des internationalen und interdisziplinären Austauschs durch Verknüpfung der Berichtsinhalte mit Entitäten anerkannter medizinischer Wissensbasen

## Konzept

Unter Verwendung der ontologiebasierten Methoden „Reasoning“, „Ontology Learning“ und „Mapping“ wird ein Programm entwickelt, welches dem folgenden Workflow folgt:

Zu Beginn der Erstellung eines neuen Berichts werden dem Pathologen auf einer grafischen Benutzeroberfläche alle Fragen angezeigt, die generell für eine vollständige Bearbeitung der gewählten Berichtsart beantwortet werden müssen (Abb. 1). Die Auswahl bestimmter Antwortmöglichkeiten kann weitere Fragen mit sich führen, welche dann dynamisch an entsprechender Stelle visualisiert werden.

Nach jeder gegebenen Antwort wird geprüft, ob der Report vollständig ist. Das Wissen darüber, welche Inhalte hierzu vorhanden sein müssen, erlangt die Software aus der Ontologie des entsprechenden Report-Typs. Diese wird zuvor im Zuge eines Ontology-Learning-Verfahrens aus einer Vorlage der ICCR (International Collaboration on Cancer Reporting) modelliert.

Anhand der Vorlagen ergeben sich zur Beantwortung jeder Fragestellung Strukturelemente aus den vier Kategorien Optionsfeld, Kontrollkästchen, Textfeld und Textbereich. Besonders bei ersteren beiden Arten wird durch das Bereitstellen von vordefinierten Antwortmöglichkeiten eine berichtsübergreifend einheitliche Codierung der gegebenen Antworten ermöglicht.

Die in der Web Ontology Language (OWL) formulierten Berichte können jederzeit gespeichert und zu einem späteren Zeitpunkt wieder geöffnet werden.

## Umsetzung

Jeder erstellte Report baut auf einer speziell strukturierten Basis-Ontologie auf und entspricht damit einer Wissensbasis mit formal definierter Semantik. Dadurch lassen sich unter Verwendung des Hermit-Reasoners Rückschlüsse auf die Vollständigkeit der Berichte ziehen. Das Ergebnis wird in Form einer Ampel visualisiert. Rot bedeutet, dass das Dokument aktuell wichtige Fragen offen lässt, während Gelb das Fehlen optionaler Angaben und Grün einen vollständigen Bericht kennzeichnen. Abb. 1 zeigt das implementierte Tool zu Beginn der Erstellung eines neuen Prostatakrebs-Reports.

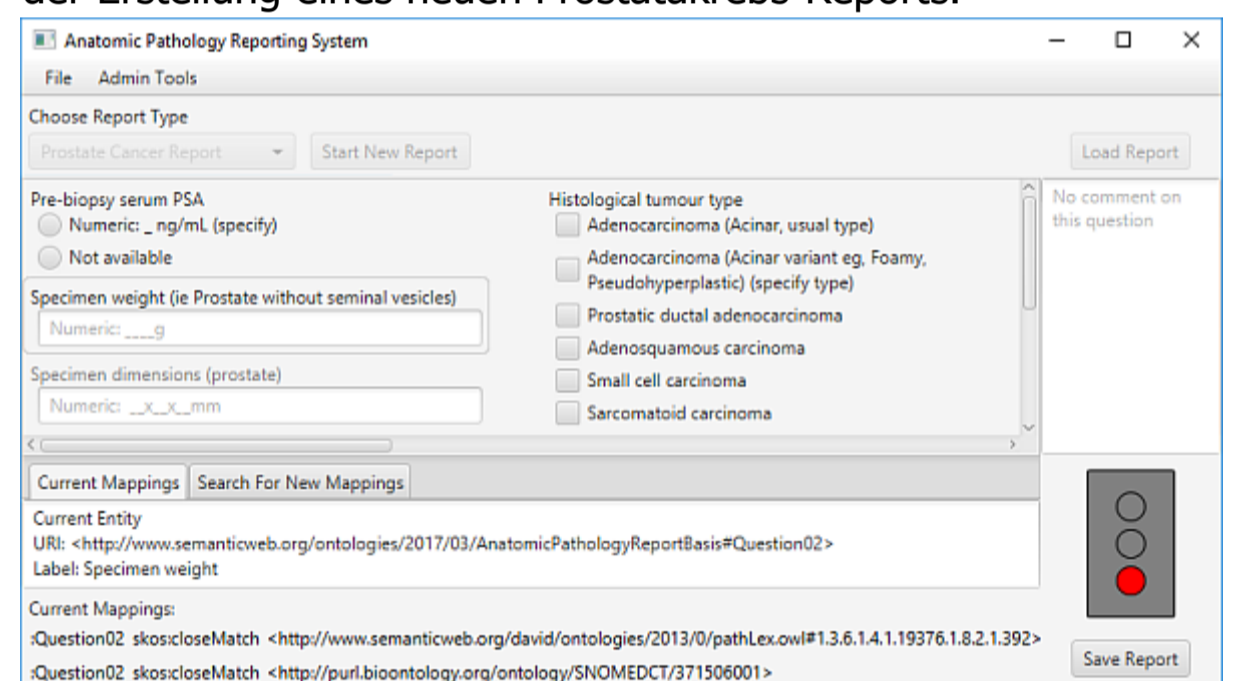


Abb. 1: Programm bei Neuerstellung eines Prostatakrebs-Reports

## Ergebnisse

Die drei in OWL modellierten Report-Vorlagen konnten softwaregesteuert im Zuge des Ontology-Learning-Prozesses erstellt werden. Der Vergleich der in Abb. 2 dargestellten Metriken der Basis-Ontologie und der erstellten Report-Vorlage für Prostatakrebs-Befunde verdeutlicht, dass der implementierte Lernprozess 1056 Axiome automatisch generiert hat. Dies erspart den manuellen Modellierungsaufwand und reduziert damit Zeit und Kosten bei der Erstellung der Vorlage-Ontologien. Weiterhin konnte beispielhaft ein Mapping von Report-Bestandteilen zu Entitäten der externen Wissensbasen SNOMED CT, NCIT und PathLex realisiert werden (vgl. Abb. 1).

Metrics		Metrics	
Axiom	84	Axiom	1140
Logical axiom count	33	Logical axiom count	434
Class count	10	Class count	143
Object property count	5	Object property count	5
Data property count	1	Data property count	1
Individual count	4	Individual count	137
DL expressivity	ALCHI(D)	DL expressivity	ALCHI(D)

Abb. 2: Metriken des Report-Grundgerüsts (links) und der erstellten Vorlage-Ontologie für Prostatakrebs-Befundberichte (rechts)

## Fazit und Ausblick

Die implementierte Software ist dazu in der Lage, pathologische Befundberichte im betrachteten Diskursbereich dynamisch zu erstellen und in einem maschinenverständlichen Modell zu repräsentieren. Dabei werden alle festgelegten Programmanforderungen erfüllt. Die Möglichkeit, Berichtsinhalte über URI-Adressen dokumentenübergreifend einheitlich anzusprechen, kann Grundlage für zukünftige Suchverfahren auf Basis der im Report codierten Konzepte sein.