

Unterstützung des Qualitätsmanagements klinischer Studien durch Integration eines semantisch fundierten Data Dictionary

Kühn K, Löbe M, Löffler M

*Institut für Medizinische Informatik, Statistik und Epidemiologie, Universität Leipzig, Deutschland
{katrin.kuehn | matthias.loebe | markus.loeffler}@imise.uni-leipzig.de*

Einleitung und Fragestellung

Im Qualitätsmanagement klinischer Studien spielt die korrekte und einheitliche Verwendung von Fachbegriffen eine wichtige Rolle, da sie für die Sicherstellung der Beobachtungsgleichheit und somit der Qualität der wissenschaftlichen Ergebnisse entscheidend ist. Besondere Probleme entstehen hierbei durch

- unscharfe Ausdrücke (z.B. Verwendung eines Hyperonyms anstatt des konkreten Fachbegriffs, Verwendung von Begriffen ohne eindeutige Definition),
- mehrdeutige Begriffe (Homonyme, Polyseme),
- unterschiedliche Verwendungskontexte,
- und vor allem auch verschiedene inkonsistente Definitionsvarianten aufgrund unterschiedlicher Expertenmeinungen.

Material und Methoden

Im Rahmen des Kompetenznetzes Maligne Lymphome [1] wurden verschiedene Werkzeuge zur Unterstützung der Qualitätssicherung in klinischen Studien entwickelt. Der *SOP-Creator* [2] ermöglicht die webbasierte Bearbeitung und zentrale Verwaltung von Standardarbeitsanweisungen (SOPs – engl. Abk. für ‚standard operation procedures‘). Der *Onto-Builder* [3] dient der Erstellung semantisch fundierter Data Dictionaries, in denen die Semantik von Fachtermini detailliert definiert werden kann. Insbesondere erlaubt der Onto-Builder die Abbildung verschiedener Definitionsalternativen zu einem Begriff, um unterschiedliche Verwendungskontexte abzubilden oder verschiedene Expertenmeinungen zu berücksichtigen [4,5].

Ein Data Dictionary bildet eine zentrale und wiederverwendbare Begriffsbasis, die jedoch von medizinischen Experten ohne spezielle Ontologiekennnisse nur sinnvoll genutzt werden kann, wenn die Inhalte über eine leicht zu bedienende Schnittstelle zur Verfügung gestellt werden. Des Weiteren sind sie nur dann motiviert, am Aufbau einer solchen Domänenontologie mitzuwirken, wenn sie einen Mehrwert durch die Verwendung der Inhalte erkennen.

Aus diesem Grund wurde eine Schnittstelle zwischen beiden Werkzeugen implementiert, die die Möglichkeit und den Nutzen einer engen Verzahnung der Redaktionsprozesse im *SOP-Creator* (Standardarbeitsanweisungen) und im *Onto-Builder* (Begriffsdefinitionen) demonstriert.

Die Inhalte des Data Dictionary können über eine Webservice-Schnittstelle von beliebigen Anwendungen abgefragt werden. Der SOP-Creator nutzt diese Schnittstelle im Prozess der Bearbeitung von Standardarbeitsanweisungen. Dabei hat der Redakteur zum einen die Möglichkeit, eine Begriffsdefinition aus dem Data Dictionary aktiv in die zu bearbeitende SOP einzubinden. Zum anderen wird von der in der SOP benutzten Fachterminologie automatisch ein Wortindex gebildet, der mit den im Data Dictionary definierten Termen verglichen wird. Sind für einen verwendeten Term bereits Definitionen vorhanden, werden sie dem Redakteur angezeigt. Er kann nun diejenige Definitionsalternative auswählen, die in der SOP gelten soll. Durch dieses Vorgehen wird dem Redakteur die aktive Suche nach bereits vorhandenen Definitionen zu verwendeten Begriffen weitgehend abgenommen.

Die ausgewählte Definition wird im Dokument eingebettet und kann maschinell ausgelesen und weiterverarbeitet werden. Ferner wird den SOP-Redakteuren die Möglichkeit gegeben, eigene Definition oder Kommentare in das Data Dictionary einfließen zu lassen.

Ändert sich eine Definition im Data Dictionary, die in einer SOP referenziert wird, werden die Änderungen nicht automatisch übernommen. Stattdessen wird der Redakteur über die Änderung informiert und um Überprüfung gebeten.

Ergebnisse

Die Autoren von studienbezogenen Dokumenten haben, wie hier am Beispiel von SOPs demonstriert, durch Nutzung des semantisch fundierten Data Dictionarys die Möglichkeit, sich (statt einer eigenen) auf eine vorgegebene, zentral verfügbare und von einem Expertengremium qualitätsgesicherte Definition eines Begriffs zu beziehen, was die Qualität des Dokuments deutlich erhöht. Des Weiteren werden durch die genaue Definition der verwendeten Begriffe Fehlinterpretationen reduziert und der Konsensprozess in der Qualitätssicherung beschleunigt.

Die Integration des Data Dictionary in das SOP-Verwaltungswerkzeug sorgt für eine intensivere Nutzung der erarbeiteten Begriffe und Definitionen. Durch die automatische Einbindung der Definitionen, die zentral verwaltet werden, wird außerdem die Fehlerquote durch manuelle Wiederverwendung deutlich gesenkt. Rückmeldungen der verwendeten Termini erlauben schließlich eine Abschätzung der Verwendungshäufigkeit einzelner Definitionen.

Diskussion und Ausblick

Die Annotation medizinischer Dokumente stellt einen wesentlichen Schritt im Hinblick auf das vernetzte Wissensmanagement in der Medizin dar. Maschinenlesbare, stark verlinkte und durch Metadaten angereicherte Informationseinheiten, die durch formale Modelle beschrieben werden, bilden die Basis für ein Semantic Web, welches sowohl formal-logische als auch auf Methoden künstlicher Intelligenz basierende Auswertungsverfahren ermöglicht. Langfristig ist dafür eine Anbindung weiterer Wissensbasen unabdingbar.

Als Weiterentwicklung der Integration des Data Dictionary in andere Anwendungen ist es denkbar, die Redakteure in der Erstellung und Bearbeitung von Dokumenten

noch stärker zu unterstützen, indem die verwendeten Begriffe automatisch mittels ontologischer Methoden auf ihre korrekte Verwendung hin überprüft werden. Im Hinblick auf dieses Ziel wird in der Forschungsgruppe Onto-Med (Ontologien in der Medizin) an der Entwicklung einer Top-Level-Ontologie gearbeitet, die u.a. eine grundlegende Kategorisierung von Fachbegriffen (z.B. Prozesse, Ereignisse, etc.) ermöglicht und Beschränkungen für die Verwendung definiert.

Danksagung

Die grundlegenden Ideen des Projekts stammen von unserer langjährigen Projektleiterin Barbara Heller, die leider nach schwerer Krankheit viel zu früh verstarb. Wir werden ihr ein ehrendes Andenken bewahren und ihre Ideen fortführen.

Literatur

- [1] Heller B, Löffler M. Telematics and Computer-Based Quality Management in a Communication Network for Malignant Lymphoma. Kompetenznetz Maligne Lymphome. Bundesministerium für Bildung und Forschung (BMBF), 2002.
- [2] <http://www.sop-creator.de/>
- [3] <http://www.onto-builder.de/>
- [4] Heller B, Kühn K, Löffler M. Onto-Builder: Software-Werkzeug für den Aufbau von Data Dictionaries. In: Steyer G, Tolxdorff T, Hrsg. bit for bit: Halbzeit auf dem Weg zur Telematikinfrastruktur. Berlin: Akademische V.-G. Aka; 2005: 108-114.
- [5] Heller B, Herre H, Lippoldt K, Löffler M. Standardized Terminology for Clinical Trial Protocols Based on Ontological Top-Level Categories. In: Kaiser K, Miksch S, Tu SW, Hrsg. Computer-based Support for Clinical Guidelines and Protocols. Amsterdam: IOS-Press; 2004: 46-60.