

SEMANTISCHE VALIDIERUNG VON ELGA-KONFORMEN LABORBEFUNDEN MIT XML-SCHEMA

Stadter M¹, Rinner C², Kohler M², Duftschmid G²

Kurzfassung

In dieser Arbeit wird ein auf den Technologien XML-Schema und Extensible Stylesheet Language Transformations (XSLT) basierender Ansatz zur Validierung von ELGA Laborbefunden vorgestellt. Ausgehend von dem Implementierungsleitfaden für ELGA Laborbefunde wurde ein XML-Schema erstellt. Der zu validierende ELGA Laborbefund wird über ein Web Service mit einem generischen XSLT Script transformiert und anschließend gegen das XML-Schema validiert.

Abstract

This work describes an approach for the validation of ELGA laboratory reports, which is based on the XML-Schema and Extensible Stylesheet Language Transformations (XSLT) technologies. Originating from the implementation guideline for ELGA laboratory reports an XML-Schema was created. By means of a web service the ELGA laboratory report, which has to be validated, is transformed via a generic XSLT script and then validated against the XML-Schema.

Keywords – ELGA, XML-Schema, Validierung, CDA

1. Einleitung

Für den österreichischen „Elektronischen Gesundheitsakt“ (ELGA) wurde von der Bundesgesundheitskommission die HL73 „Clinical Document Architecture“ (CDA) als Dokumentenstandard gewählt⁴. Von der ELGA GmbH⁵ wurden CDA Implementierungsleitfäden [3,4] für die ELGA-Dokumenttypen erstellt. Diese Leitfäden beinhalten detaillierte strukturelle und inhaltliche Vorgaben zur Erstellung der Dokumenttypen. Sie sind gegliedert in ein Basis Dokument und je ein Addendum für Laborbefund, Radiologiebefund und Entlassungsinformation. Damit wurde auch eine Voraussetzung geschaffen, medizinische Dokumente, die in ELGA eingestellt werden sollen, neben der Prüfung auf Konformanz mit dem allgemeinen CDA-Standard auch inhaltlich auf ELGA-Konformanz zu prüfen (semantische Validierung). Zur technischen Umsetzung der semantischen Validierung eignet sich laut [1,2] neben Schematron auch ein auf den Technologien XML-Schema

¹Student, Universität Wien

²Institut für Medizinisches Informationsmanagement und Bildverarbeitung, Medizinische Universität Wien

³<http://www.hl7.org>

⁴http://www.bmg.gv.at/home/Schwerpunkte/E_Health/ELGA_Die_Elektronische_Gesundheitsakte/Technisches_Framework_und_Standards_im_Gesundheitswesen

⁵<http://www.elga.gv.at>

und Extensible Stylesheet Language Transformations (XSLT) basierender Ansatz. Die vorliegende Arbeit beschreibt eine Implementierung dieses Ansatzes zur semantischen Validierung von ELGA-Laborbefunden.

2. Methoden

Zur semantischen Validierung von ELGA-Laborbefunden wurde die in [1,2] beschriebene XML-Schema / XSLT basierte Methodik angewendet (siehe *Abbildung 1*).

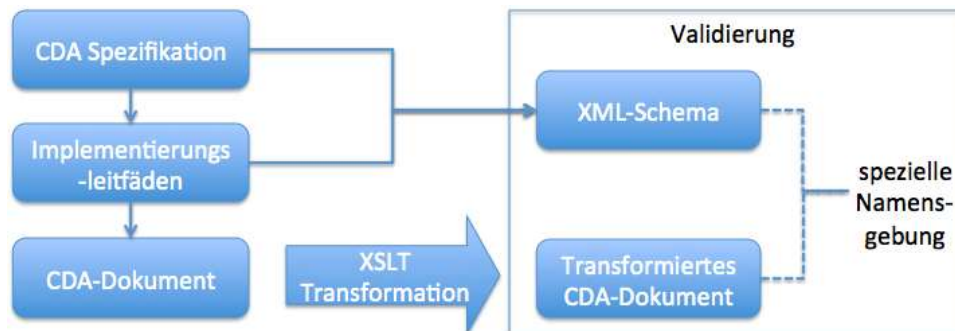


Abbildung 1: Semantische Validierung von CDA-Dokumenten mit XML-Schema / XSLT basierter Methodik.

In der in *Abbildung 1* beschriebenen Implementierung wurden die ELGA-Implementierungsleitfäden [3,4] herangezogen, die basierend auf der allgemeinen CDA Spezifikation das CDA-Dokument „ELGA-Laborbefund“ beschreiben. Basierend auf dieser Beschreibung wird ein XML-Schema erzeugt, welches zur Einhaltung der von der W3C für XML-Schemas verpflichtend vorgeschriebenen „Unique Particle Attribution Constraint Rule“¹ eine spezifische Namensgebung benötigt. Durch eine generische XSLT-Transformation wird das zu prüfende CDA-Dokument auf die gleiche Namensgebung gebracht und kann dann gegen das XML-Schema validiert werden.

Im Gegensatz zu den im EHR Standard ISO/EN 13606 verwendeten Archetypen sind CDA Implementierungsleitfäden nicht maschinell verarbeitbar. Aus diesem Grund kann das XML-Schema nicht wie im Fall der Archetypen automatisch aus dem Implementierungsleitfaden generiert werden, sondern muss manuell erstellt werden [1,2]. Bei der Erstellung muss darauf geachtet werden, dass die oben erwähnte „Unique Particle Attribution Constraint Rule“ nicht verletzt wird. Diese besagt, dass ein XML-Schema auf einer hierarchischen Ebene nicht mehrere, gleichnamige aber unterschiedlich strukturierte Elemente vorschreiben darf. Da sich ein Implementierungsleitfaden aber nur auf die vorgegeben Klassen des CDA Modells beziehen kann, kommt es bei einer „straight-forward“ Ableitung des XML-Schemas aus dem Implementierungsleitfaden typischerweise zu einer Verletzung der „Unique Particle Attribution Constraint Rule“ (z.B. könnte der Leitfaden zwei Instanzen der CDA-Klasse OBSERVATION auf derselben hierarchischen Ebene, mit unterschiedlicher Semantik und Struktur, vorschlagen). Um das Problem zu lösen wird in [1,2] ein Konzept zur Umbenennung der CDA-Elemente im Body-Teil des XML-Schemas vorgeschlagen. Dabei werden die im Implementierungsleitfaden für die verschiedenen Elemente vorgeschriebenen *templateIDs* verwendet, um eindeutige Bezeichnungen zu generieren.

¹ <http://www.w3.org/TR/xmlschema-1/#non-ambig>

3. Resultate

Zunächst wurde ausgehend von dem CDA XML-Schema der CDA R2 Normative Webedition1 und den ELGA Implementierungsleitfäden [3,4] ein XML-Schema erstellt. Um einen ersten Entwurf des CDA-Body-Teils des XML-Schemas zu erstellen wurde mit Hilfe eines XML Tools2 von einem fertigen Laborbefund ein XML-Schema automatisch abgeleitet, welches dann editiert wurde. Das XML-Schema spezifiziert auf Basis der Implementierungsleitfäden die Struktur eines ELGA Laborbefunds, inklusive Optionalitäten, Kardinalitäten und Befüllungsvorschriften von Elementen und Attributen.

Um die CDA-Dokumente gegen das XML-Schema validieren zu können, müssen sie dieselbe Konvention zur Namensgebung einhalten wie das XML-Schema. Das für die hierfür notwendige Transformation der CDA-Dokumente verwendete XSLT-Script sucht in den Act-Relationship Klassen *component*, *entryRelationship* und *entry* mittels XPath nach dem Element *templateID*. Wird dieses direkt in der CDA-Klasse oder deren Kindern gefunden, so wird der Name des CDA-Elements um ein aus dem Attribut *root* des Elements *templateID* gebildeten Suffix erweitert (z.B. *component* deren Kind bzw. Kindes-Kind mittels `templateId.root=1.3.6.1.4.1.19376.1.3.1.6` identifiziert ist, wird transformiert zu *component_1_3_6_1_4_1_19376_1_3_1_6*).

Weiters wurde mittels Eclipse3 ein Web Service erstellt, das einen ELGA Laborbefund im XML Format mittels des XSLT-Scripts gemäß der oben erwähnten Namensgebungskonvention transformiert. Die transformierte Version des Laborbefunds wird dann innerhalb des Web Services gegen das XML-Schema validiert, falls Fehler auftreten, werden diese zurückgesendet. Ansonsten wird gemeldet, dass das hochgeladene Dokument gültig ist. Ein webbasierter Testclient steht unter <https://mias.meduniwien.ac.at/ELGAVvalidateClient/> zur Verfügung.

Neben auf Basis der Implementierungsleitfäden selbst erstellten Laborbefunden wurde auch je ein vom Redaktionsteam „ELGA CDA Labor“ erstelltes Beispieldokument für einen *Klinisch-Chemischen Laborbefund* bzw. einen *Bakteriologischen Laborbefund mit Antibiogramm* validiert. Dabei konnten einzelne Fälle lokalisiert werden, in denen die Beispieldokumente nicht mit den Implementierungsleitfäden konform waren. Nach Bereinigung dieser Fehler (z.B. falscher Code für ein Code System oder Groß / Klein Schreibung bei Befüllungsvorschriften bzw. andere Schreibweise) validierten die Beispieldokumente erfolgreich.

4. Diskussion

Die semantische Validierung mit XML-Schema / XSLT gemäß der in [1,2] beschriebenen Methodik konnte erfolgreich zur Validierung von ELGA Laborbefunden angewendet werden. Wie erwartet, gestaltete sich die manuelle Erstellung des XML-Schemas jedoch aufwändig. Die fehlende maschinelle Verarbeitbarkeit der CDA Implementierungsleitfäden stellt hier einen klaren Nachteil gegenüber dem EHR-Standard ISO/EN 13606 dar, wo das XML-Schema automatisch aus Archetypen abgeleitet werden kann [6]. Bei der Umsetzung der Befüllungsvorschriften wurden im Implementierungsleitfaden für den Laborbefund [4] einige Fehler lokalisiert, die durch Zuhilfenahme des IHE Laboratory Technical Framework [5], auf welchem [4] basiert, gelöst werden konnten.

1 http://www.hl7.org/documentcenter/private/standards/cda/r2/cda_r2_normativewebedition.zip

2 z.B. Altova XML Spy oder Oxygen XML

3 <http://www.eclipse.org>

Nach der Transformation mittels XSLT Skript kann es vorkommen, dass die Fehlermeldungen und deren Zeilennummern nicht mit den Fehlern im Originaldokument übereinstimmen, da sie sich auf das transformierte CDA-Dokument beziehen. Befinden sich im Originaldokument mehrere XML-Elemente in der gleichen Zeile (Dokument nicht im „Pretty Print“ Format), kann es zu Verschiebungen der Zeilennummern in der Fehlermeldung kommen. Um die Fehlersuche zu erleichtern, wird daher das transformierte XML-Dokumente zusätzlich zurückgeben.

Das in dieser Arbeit erstellte XML-Schema könnte noch dahingehend optimiert werden, jene XML-Elemente nicht mehr zu prüfen, die bereits durch eine vorhergehende Validierung mit dem von HL7 zu Verfügung gestellten allgemeinen XML-Schema für die CDA geprüft werden könnten. Durch die initiale automatische Generierung unseres XML-Schemas werden einige XML-Elemente durch beide XML-Schemas validiert. Durch die Verwendung des XML-Schema Konstrukt `xs:anyType` ist es möglich, Unterbäume zu definieren ohne auf weitere Details eingehen zu müssen, sofern diese nicht weiter eingeschränkt werden. Eine doppelte Validierung könnte damit verhindert werden.

5. Literatur

[1] RINNER C., DUFTSCHMID G., Validieren von auf Zweimodell-Ansätzen basierenden, vollstrukturierten EHR-Daten am Beispiel EN/ISO13606 und HL7 CDA, Proceedings of eHealth 2009 & eHealth Benchmarking 2009 – Medical Informatics meets eHealth, pp. 215-221, 2009.

[2] RINNER C., JANZEK-HAWLAT S., SIBINOVIC S., DUFTSCHMID G., Semantic validation of standard based electronic health record documents with W3C XML Schema, Methods of Information in Medicine, 2010.

[3] BRANDSTÄTTER J., ELGA Kernanwendungen, CDA Dokumente für das österreichische Gesundheitswesen, Implementierungsleitfaden, ARGE ELGA, 2009.
http://www.elga.gv.at/fileadmin/user_upload/uploads/download_Papers/Harmonisierungsarbeit/Leitfaden/Leitfaden_ELGA_CDA_Dokumente_1.00.pdf

[4] MENSE A., SABUTSCH S., SAUERMAN S., ELGA Kernanwendungen, CDA Laborbefund für das österreichische Gesundheitswesen, Addendum, ARGE ELGA, 2009.
http://www.elga.gv.at/fileadmin/user_upload/uploads/download_Papers/Harmonisierungsarbeit/Labor/LeitfadenAddendum_ELGA_CDA_Laborbefund_Version_1.01.pdf

[5] IHE INTERNATIONAL, IHE Laboratory Technical Framework, 2008.
http://www.ihe.net/Technical_Framework/index.cfm

[6] JANZEK-HAWLAT S., KUTTIN O., SIBIOVIC S., DUFTSCHMID G., Automatisierte Generierung von XML-Schemata aus EN/ISO 13606 Archetypen, Proceedings of eHealth 2009 & eHealth Benchmarking 2009 – Medical Informatics meets eHealth, pp. 69-75, 2009.

Corresponding Author

Matthias Stadter

Student, Universität Wien

Email: matthias@stadter.at