

Hybride Dokumente auf der Basis internationaler Standards am Beispiel des CDA-Arztbriefs

Dr. Bernd Wild

intarsys consulting GmbH
Bahnhofplatz 8, D-76137 Karlsruhe
bwild@intarsys.de

Abstract: Die Verbesserung des Informationsaustauschs zwischen den Akteuren im Gesundheitswesen bedingt zwangsläufig die verstärkte Nutzung von strukturierten Daten und Nachrichten. Bereits frühzeitig hat sich dafür der internationale Standard HL7 und darunter insbesondere für Arztbriefe das Format CDA herausgebildet. Eine grundlegende Schwierigkeit ist die ambivalente Verwendung eines solchen Standards sowohl für den Transport der strukturierten Daten als auch für die visuelle Repräsentation der Informationen bei den Nutzern. Im Folgenden wird ein Ansatz beschrieben, wie durch die Nutzung des ISO-Standards PDF/A-3 als Trägerformat sowohl visuelle Repräsentationen als auch die maschinenlesbaren Daten gleichzeitig transportiert werden können. Darüber hinaus ermöglicht der Formatstandard sowohl die Langzeitarchivierbarkeit des visuellen Anteils wie auch verschiedene Signaturverfahren zur Absicherung von Integrität und Authentizität des Inhalts.

1 Der CDA-Arztbrief

Der Arztbrief wird verwendet, um Informationen über einen Patienten bezüglich eines Eingriffs, einer Therapie oder Diagnose zwischen medizinischen Instanzen im Gesundheitswesen auszutauschen. Der typische Fall ist die Erstellung eines Arztbriefs durch den behandelnden Arzt in einer Klinik oder den Facharzt zur Weiterleitung an den Hausarzt. Zusammen mit fachspezifischen Unterlagen wie Röntgen- oder MRT-Aufnahme soll gewährleistet werden, dass alle während einer Behandlung ermittelten diagnostischen Informationen zusammen mit der Behandlungsbeschreibung und der weiteren Therapie und Medikation an die weiter behandelnde Stelle weitergereicht wird.

Arztbriefe werden heute überwiegend in Textverarbeitungsprogrammen erstellt (z.B. MS Word), was zwar eine Lesbarkeit durch den Menschen gewährleistet, nicht jedoch eine automatische Extraktion der damit transportierten Informationen zur Weiterverarbeitung in Praxisinformationssystemen (PVS) oder Klinikinformationssystemen (KIS). Dadurch müssen von der empfangenden Stelle die Informationen in einem Arztbrief manuell im jeweiligen Informationssystem wieder erfasst werden, was neben dem Zeitaufwand auch Erfassungsfehler [Ba01] nach sich ziehen kann.

Hier setzt die Standardisierung von HL7 mit der Clinical Document Architecture (CDA) an [HL05]. Ähnlich wie in anderen Domänen wie z.B. dem Waren- und Rechnungsvverkehr definiert HL7 einen Satz an Nachrichten, die zwischen unterschiedlichsten Akteuren im Gesundheitswesen ausgetauscht werden können. Die CDA bildet mit ihrem Fokus auf die auszutauschenden Dokumente einen Teilbereich von HL7, der aber auf dem gemeinsamen Reference Information Model (RIM) beruht. Dieses Modell beschreibt alle grundlegenden Objekte, Aktivitäten und Relationen der Domäne „Gesundheitswesen“ und ist durch einen Satz an XML-Schemadefinitionen festgelegt [WW08]. Damit kann ein CDA-Dokument über XML-Schemavalidierung automatisiert auf syntaktische Korrektheit geprüft werden.

Erste Ansätze einer Standardisierung der medizinischen Dokumentation wie Arztbriefe spezifizierte die sogenannte Kona Architecture, die dann 1996 von der Patient Record Architecture abgelöst wurde [Al99]. Der Nachfolger, die CDA Version 1, wurde im Jahr 2000 auch als ANSI-Standard veröffentlicht, während als aktuellste Ausgabe nun CDA Version 2 im Jahr 2005 publiziert wurde [Do05].

Ein CDA-konformer Arztbrief besteht insgesamt aus 4 Teilen [Ob06]:

- Teil 1 enthält allgemeine Dokumentinformationen wie Version, verwendetes Schema etc.
- Teil 2 umfasst alle Angaben zum Erbringer einer medizinischen Maßnahme, also zu dem behandelnden Arzt oder der medizinischen Einrichtung.
- Teil 3 ist dem Empfänger einer medizinischen Maßnahme gewidmet, enthält also alle Angaben zu dem Patienten selbst inkl. Vorerkrankungen, chronische Krankheiten etc.
- Teil 4 beschreibt schlussendlich das eigentliche Ereignis, das zur Behandlung geführt hat, sowie die erfolgten Behandlungsschritte und Therapiemaßnahmen.

Eine Visualisierung des Inhalts eines CDA-Arztbriefs kann über übliche Maßnahmen und Werkzeuge zur XML-Bearbeitung erfolgen. Mit Hilfe eines XSL Stylesheets kann die XML-Struktur z.B. über einen XSL Prozessor wie Xalan oder Saxon in ein formatiertes und durch jeden Webbrowser darstellbares HTML überführt werden [WW10].

1.1 Anforderungen an einen CDA-Arztbrief

Die erfolgreiche Einführung eines Standards für den Arztbrief ist eng an die Erfüllung einer Reihe von Anforderungen geknüpft, die die aktuelle Praxis (Freiformat im Textverarbeitungsprogramm) nur teilweise abbildet. Entscheidend für die Durchsetzung bei den Anwendern im Markt sind eine einfache Handhabung und eine möglichst transparente Integration in die täglichen Abläufe. Jeder zusätzliche Arbeitsschritt stellt ein großes Hindernis dar. Im Folgenden sind diese Anforderungen zusammengefasst:

- standardisiert: Bei einer Neugestaltung des Arztbriefs sollen zumindest national anerkannte öffentliche Standards verwendet werden. Wenn möglich sollten CEN oder ISO Standards bzw. gleichrangige wie HL7 zur Anwendung kommen.

- **maschinell verarbeitbar:** Der automatisierten Verarbeitung von Arztbriefen wird ein großes Potential hinsichtlich Zeitersparnis und Vermeidung von Fehlerfassung zugerechnet. Demzufolge ist eine möglichst weitgehende maschinelle Verarbeitbarkeit wünschenswert. Das sollte sich sowohl auf die Erzeugung als auch die Erfassung eines Arztbriefes beziehen. Ein bewährter Ansatz ist hier die Nutzung von XML-basierten Formaten.
- **signierbar:** Obwohl die rechtliche Situation hinsichtlich der Frage, ob ein Arztbrief vom Arzt eigenhändig unterschrieben bzw. sein elektronisches Pendant dann qualifiziert elektronisch signiert werden muss, noch nicht abschließend geklärt ist, sollte die Möglichkeit bestehen, einen Arztbrief elektronisch entweder mit einer persönlichen, einem Organisations-Zertifikat oder einem Zeitstempel zu signieren.
- **leicht integrierbar:** Die Integration in existierende Patientenverwaltungssysteme sollte möglichst einfach durchführbar sein, um möglichst schnell eine hohe Marktdurchdringung zu erzielen. Dazu ist das Vorhandensein entsprechender standardkonformer Open Source oder kommerzieller Software-Module und -Bibliotheken eine wichtige Voraussetzung.
- **leicht verarbeitbar:** Die Interpretation des Standards und damit auch die Überprüfbarkeit auf Einhaltung des Standards sollten möglichst einfach möglich sein. Durch die XML/XSD-Basis von CDA sind hierfür wichtige Voraussetzungen geschaffen.
- **(langzeit-)archivierbar:** Da Arztbriefe zur klinischen Dokumentation einer Behandlung gehören, unterliegen sie auch den gesetzlich vorgeschriebenen Aufbewahrungspflichten. Im Fall von klinischen Behandlungen beträgt diese in Deutschland gegenwärtig 30 Jahre nach Abschluss der Behandlung. Im Gegensatz zum bislang noch vorherrschenden Archivierungsmedium Papier sollte das elektronische Medium zweifelsfrei über diesen Zeitraum hinweg interpretierbar und darstellbar sein. Damit ist ein Standard für langzeitarchivierbare Dokumente gefordert.
- **druckbar:** Der Arztbrief muss jederzeit druckbar sein und dabei eine der Darstellung auf dem Computer-Bildschirm adäquate Wiedergabequalität aufweisen. Da eine Kopie des Arztbriefs i.d.R. auch an den Patienten geht und dort nicht von einer maschinellen Nachverarbeitung ausgegangen werden kann, ist diese Eigenschaft besonders wichtig.
- **lesbar:** Anknüpfend an die Forderung »druckbar« spielt die visuelle Darstellungsmöglichkeit auf beliebigen Endgeräten eine wichtige Rolle. Für den Ersteller und Empfänger eines Arztbriefs ist es relevant, dass die Bildschirmdarstellung des Briefs mit der Druckfassung übereinstimmt. Der Arztbrief muss auch außerhalb einer wohl definierten Softwareumgebung mit Bordmitteln unterschiedlicher Endgeräte gleichartig und visuell eindeutig lesbar sein.

2 Das Dokumentenformat PDF/A

2.1 Historie

Die Internationale Organisation für Normung (ISO) hat am 28. September 2005 einen Standard für ein Dokumentenformat zur Langzeitarchivierung elektronischer Dokumente

definiert - die offizielle Bezeichnung lautet *ISO-19005-1 - Document management - Electronic document file format for long-term preservation - Part 1: Use of PDF 1.4 (PDF/A-1)* [IS05].

Begonnen hatten die Standardisierungsaktivitäten im Mai 2002 in den USA. Man wollte auf keinen Fall ein von Grund auf neues Dokumentenformat schaffen, das sicherlich erhebliche Schwierigkeiten in der weltweiten Verbreitung haben würde, sondern - wenn möglich - auf einem bestehenden Format aufsetzen und Regeln für eine langzeitarchivierungskonforme Erzeugung und Verwendung definieren. Die Wahl für das Ausgangsformat fiel auf PDF, damals in der Version 1.4. Beteiligt waren die AIIM (Association for Information and Image Management), die NPES (National Printing Equipment Association) und die Verwaltung der US-amerikanischen Gerichte. An den Standardisierungsarbeiten nahmen auch bekannte PDF-Hersteller wie Adobe Systems, Library of Congress, Surety Inc., Quality Associates Inc., Appligent, Merck, EMC, PDF Sages und die NARA (National Archives & Records Administration) und weitere Firmen wie Xerox, Honeywell, EDS und Glaxo Smith Kline teil.

Die Initiatoren des Standardisierungsprojekts stellten eine erste Version zusammen und reichten ihren Vorschlag bei der ISO ein, um diesen als internationalen Standard registrieren zu lassen. Das Projekt wurde von der ISO dem technischen Komitee TC 171 (Document Management Applications) zugeteilt. Der Standard wurde schließlich in mehreren Etappen verbessert, bis er im September 2005 als Norm ISO19005-1, oder landläufig einfach als PDF/A-1 bezeichnet, endgültig freigegeben wurde.

Doch die Normungsarbeit war damit nicht abgeschlossen. Beeinflusst durch die praktischen Erfahrungen bei der Einführung von PDF/A-1 im Markt, zu denen die PDF Association mit dem PDF/A Competence Center nicht unwesentlich beitrug, wurden weitere Anforderungen bei der Langzeitarchivierung elektronischer Dokumente analysiert und aufgenommen. Diese führten schließlich zur Veröffentlichung des zweiten Normteils ISO19005-2 bzw. PDF/A-2 im Herbst 2011 [IS11]. Mit dem Ende 2012 erwarteten dritten Normteil bzw. PDF/A-3 werden die Möglichkeiten, PDF/A als langzeitstabiles elektronisches Trägerformat für nahezu beliebige Informationstypen zu verwenden, nochmals erweitert.

2.2 PDF/A-1

PDF, vormals als Firmenstandard von Adobe Inc. und nun als ISO-Standard ISO32000 vorliegend, beschreibt ein generelles seitenorientiertes und plattformunabhängiges Dokumentenformat. Neben dem Anspruch, auflösungs- und ausgabegeräteunabhängige Repräsentationen der Dokumente zu liefern, besitzt PDF zwischenzeitlich auch etliche interaktive und dynamische Eigenschaften, so z.B. Formulare, JavaScript oder die Einbettung beliebiger Fremdformate. Für ein langzeitstabiles und rein auf Basis der Spezifikation reproduzierbares Dokumentenformat sind solche Fähigkeiten aber nicht geeignet. Vielmehr muss ein in sich abgeschlossenes Format definiert werden, bei dem jedes Dokument alle Ressourcen enthält, die zu seiner originalgetreuen Darstellung benötigt werden. Unter dieser Maxime beschreibt der PDF/A-1 Standard, welche Eigenschaften und Konstrukte eines PDF der Versionsstufe PDF 1.4 [Ad01] in einem

PDF/A-1-Dokument enthalten sein dürfen. Im Ergebnis hat man in jedem Fall ein gültiges PDF, das zudem noch die Auflagen des PDF/A-1-Standards erfüllt.

Die wesentlichen Einschränkungen und Auflagen an ein PDF, um als PDF/A-1 konform zu gelten, sind:

- Einbettung aller verwendeten Zeichensätze (TrueType, Type1, Type3,
- Kennzeichnung aller verwendeten Farbräume durch sog. OutputIntent,
- keine Verschlüsselung oder Zugriffsrechte,
- Verwendung von XMP-basierten Metadaten,
- keine interaktiven Inhalte wie JavaScript, 3D oder Multimedia,
- keine Einbettung von Fremdformaten (z.B. Office-Dokumente, Filme etc.),
- Beschränkung auf bestimmte Bildformate wie JPG, TIFF, JBIGoder JBIG2,
- keine transparenten Objekte.

Alle anderen Konstrukte aus PDF 1.4 wie z.B. eingebettete elektronische Signaturen sind hingegen erlaubt. Eine weitere Differenzierung des Standards bezieht sich auf die sog. Konformitätsstufe, die bei PDF/A-1 die Unterscheidung PDF/A-1a und PDF/A-1b annehmen kann. Während ein 1b-konformes Dokument keinerlei Informationen zu seiner Inhaltsstruktur beinhalten muss, wird diese Eigenschaft bei einem 1a-Dokument vorausgesetzt. Darunter versteht man Informationen zum semantischen Aufbau eines Dokuments wie "Überschrift", "Kapiteltext", "Fußzeile" etc., ähnlich den "Tags" in einem HTML-Dokument, mit denen einzelne Seitenelemente ausgezeichnet sind. Der Einsatz derartiger Strukturen erlaubt einen weitgehend barrierefreien Zugang zu dem Dokument, da nun sowohl Navigationen in den Dokumentstrukturen als auch multimediale Ausgaben wie das Vorlesen des Dokumenteninhalts möglich werden. Allerdings verlangt die Bereitstellung eines PDF/A-1a-konformen Dokuments bereits Vorkehrungen bei der Erzeugung aus dem Quellformat, weshalb eine nachträgliche Konvertierung eines PDFs nach PDF/A-1a nicht praktikabel ist. Die überwiegende Anzahl der gegenwärtig existierenden PDF/A-1 Dokumente entsprechen der Konformitätsstufe 1b, mit der eines der Hauptziele des Standards, quasi "elektronisches Papier" zu schaffen, vollauf erreicht werden kann. Insbesondere für die Digitalisierung von Papierdokumenten (Scanning, OCR-Anwendung) hat sich PDF/A-1b vollauf bewährt.

2.3 PDF/A-2

Eine Reihe von Beschränkungen bei PDF/A-1 war der zugrunde liegenden PDF 1.4-Spezifikation geschuldet. Dazu zählten u.a. die erlaubten Bildformate oder die einbindbaren Zeichensatzformate. Mit der Fortschreibung des PDF/A-Standards zu PDF/A-2, der seit Herbst 2011 veröffentlicht ist, ergeben sich nun im Vergleich zu PDF/A-1 u.a. folgende zusätzliche Möglichkeiten:

- Verwendung von JPEG2000 als Bildformat,
- Einbettung von OpenType Fonts,
- Unterstützung von Transparenzen bei Objekten (z.B. "weiche" Schatten),
- Einbettung von weiteren PDF/A-1- oder PDF/A-2-konformen Dateien (sog. File Collections).

Während für den Anwendungsbereich "Digitalisierung von Papier" die Freigabe des JPEG2000-Formats zur Bildspeicherung sehr begrüßt wird, ergeben sich durch die Möglichkeit, nun auch weitere (A-1- bzw. A-2-konforme) Dokumente in eine PDF/A-2-Datei einbetten zu können, ganz neue Einsatzgebiete. Lag bislang der Fokus ausschließlich auf Einzeldokumenten, die dann von einem bestehenden Archiv- oder DMS-System zu einer Art "Mappe" zusammengefasst wurden, so kann dies nun direkt auf Dokumentenebene unabhängig von der Dokumentenverwaltung erreicht werden. Es eröffnen sich Anwendungsfelder wie digitalisierte Patientenmappen, Versichertenordner oder die Archivierung von E-Mails, die inklusive der Anhänge nach PDF/A-1 bzw. PDF/A-2 konvertiert werden. PDF/A-2 kennt auch mehrere Konformitätsstufen (a, b, u), die inhaltlich weitgehend denen bei PDF/A-1 entsprechen.

2.4 PDF/A-3

Eine in der Praxis häufig gestellte Anforderung ist - neben der Aufbereitung eines Quelldokuments als PDF/A in einer unveränderlichen und visuell originalgetreu wiedergegebenen Form - auch die "Mitarchivierung" der Quelldatei oder des Quellformats, aus dem das PDF/A-Dokument erzeugt wurde. Dabei geht es nicht um die langzeitstabile Archivierung jeglicher Dateiformate, die grundsätzlich aufgrund fehlender Standards gar nicht möglich ist. Vielmehr steht im Vordergrund, dass die eingebetteten Quelldaten weiterverarbeitet werden können. Der Standard PDF/A-3, der voraussichtlich Ende 2012 veröffentlicht werden wird, unterscheidet sich von der Normstufe A-2 genau in diesem einen Punkt: der Einbettbarkeit beliebiger Dateien, sofern dafür ein MIME-Typ vorhanden ist. Die dadurch eröffneten Möglichkeiten sind weitreichend. So können nun elektronische Rechnungen in PDF/A-3 erstellt werden und gleichzeitig eine XML-Datei beinhalten, die die Nettodaten der Rechnung maschinenlesbar mitbringt. Oder Arztbriefe können visuell als Brief aufbereitet das sichtbare PDF/A-3-Dokument bilden, während sich die strukturierten Daten des Arztbriefes als CDA-XML-Dokument eingebettet im PDF/A-3-Dokument wiederfinden. Während für die eingebetteten Daten und Dateien keine Garantie auf Reproduzierbarkeit in der Zukunft abgegeben werden kann, existiert mit dem eigentlichen PDF/A-3-Dokument eine langzeitarchivierbare und stabile statische Repräsentation.

2.5 File Collections in PDF/A-3

Realisiert wird die Einbindung von Fremddateien durch die Nutzung des Tags */AF* (für *Associated Files*) im *Root*-Katalog der PDF-Datei. Darunter verbergen sich Verweise auf die in komprimierter Form eingebetteten Dateien und deren Basisattribute wie Dateiname, MIME-Typ und Dateigröße. Abbildung 1 veranschaulicht diese Einbindung. Sind im PDF-Standard bislang mehrere Varianten für die Einbindung externer Dateien möglich, so schränkt der A-3-Standard die Art der Einbindung genau auf das oben beschriebene Tag ein.

Als wichtige Metainformationen kann im Tag */AFRelationship* die Art der Beziehung zwischen der eingebundenen Datei und dem (visuellen) PDF/A-Teil abgelegt werden. Zur Wahl stehen nach ISO 19005-3 hier die Werte *Source*, *Data*, *Alternative*,

Supplement und *Unspecified*. Im Fall der Einbindung einer CDA-XML-Datei, die über eine XSL-Transformation in eine visuelle Repräsentation umgeformt wurde, würde hier z.B. der Wert *Source* stehen.

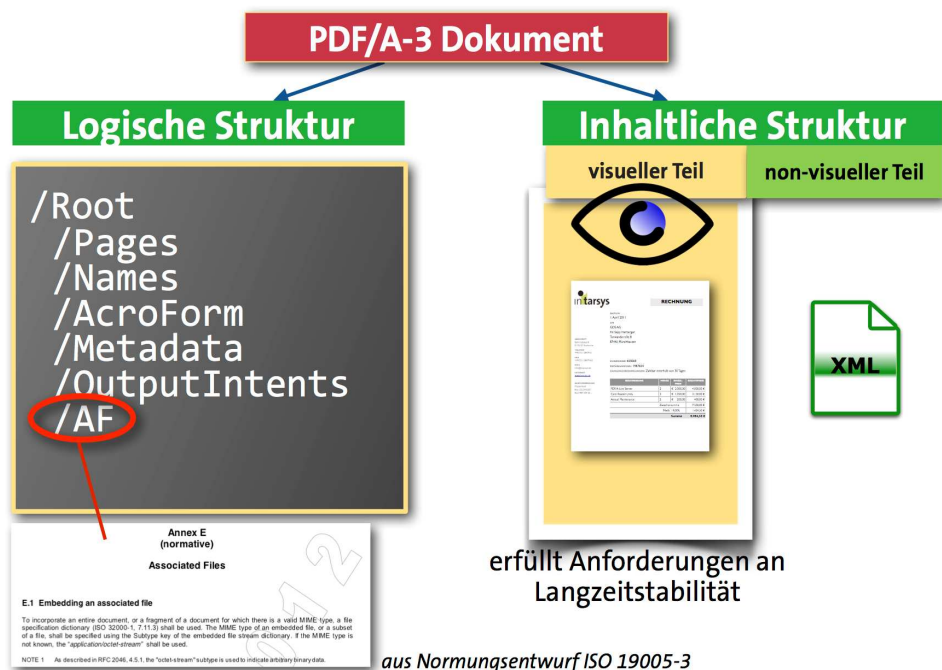


Abb. 1: Die beiden Teile eines PDF/A-3 Dokuments.

Ein PDF/A-3 konformer Viewer muss das Vorhandensein von eingebetteten Dateien analog einem Inhaltsverzeichnis entsprechend darstellen können. Dabei spielt es keine Rolle, wie er dies konkret umsetzt. Das Öffnen und die Darstellung von Fremddateien ist nicht die Aufgabe des PDF/(A-3)-Viewers. Vielmehr wird davon ausgegangen, dass anhand des MIME-Tags eine entsprechend geeignete Anwendung auf dem Rechner des Anwenders gestartet wird, die dann die Anzeige übernimmt.

3 PDF/A-3 als Trägerformat für Arztbriefe

Geht man von PDF/A-3 als Trägerformat für CDA-Arztbriefe aus, so stellt sich die Frage, wie derartige hybride Dokumente erzeugt und anschließend auch validiert werden können. Angesichts der Erfahrungen im Branchenverband PDF Association (vormals PDF/A Competence Center) mit der Einführung des Normteils A-1 nach der Veröffentlichung im Jahre 2005 stellen eine einfache Handhabung und ein ausreichendes Angebot an unterstützenden Softwarekomponenten kritische Faktoren für die Akzeptanz eines Dateiformats dar.

3.1 Die Erstellung des visuellen Anteils

Um aus der CDA-XML-Beschreibung eine wohlformatierte und lesbare Repräsentation zu gewinnen, die den bislang konventionell erstellten Arztbriefen ähnelt, kommen standardisierte XSL-Transformationstechniken zum Einsatz. So hat z.B. der VHitG ein Beispiel-XSL-Skript bereitgestellt, mit dem ein CDA-XML in eine HTML-Form überführt werden kann. Die Darstellung kann dann in jedem Browser erfolgen. Für die Überführung des CDA-XML-Codes in ein PDF/A-3 müssen ähnliche Schritte vorgenommen werden. Abbildung 2 veranschaulicht zwei Alternativen für diesen Transformationsprozess.

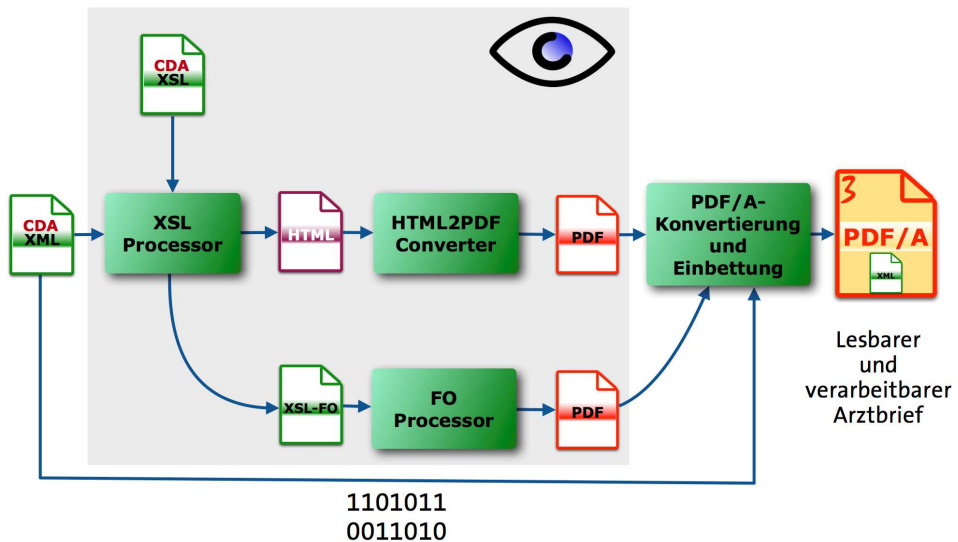


Abb. 2: Der Transformationsprozess von XML nach PDF/A-3.

So findet in der ersten Variante eine XSL-Transformation von XML nach HTML statt. Obwohl dieses HTML bereits in einem Webbrowser betrachtet werden kann, dient es hier nur als intermediäres Format. Daran schließt sich eine Konvertierung von HTML nach PDF/A-3 an. Abschließend erfolgt die Einbettung der Quell-XML-Datei über den /AF-Tag in das fertige PDF/A-3 Dokument. Für die Konvertierung von HTML nach PDF(A) existieren neben kommerziellen Werkzeugen auch etliche Open Source Anwendungen. Allerdings ist dem Autor kein frei verfügbares Werkzeug bekannt, das mit guter Qualität eine unmittelbare Konvertierung von HTML nach PDF/A-1 oder höher ermöglicht. Die Transformation in A-3 wird mit Veröffentlichung des ISO-Standards sicherlich noch eine gewisse Zeit den kommerziellen Produkten vorbehalten bleiben.

3.2 Signatur eines hybriden Arztbriefs

Die elektronische Signatur eines hybriden Arztbriefs auf PDF/A-3-Basis kann aufgrund der Flexibilität des Formats auf unterschiedliche Weise erfolgen. Im Wesentlichen können hierbei drei Varianten identifiziert werden:

- eingebettete elektronische Signatur im PDF/A-3-Dokument selbst,
- in die CDA-XML-Datei eingebettete XML-Signatur,
- zusätzliche abgesetzte Signatur der CDA-XML-Datei, die ebenfalls in das PDF/A-3 als Dateianhang eingebettet wird.

Alle drei Varianten sind standardkonform, weisen aber teils gravierende Unterschiede bezüglich der Unterstützung durch aktuelle Produkte zur Erstellung und Validierung auf. Variante 1 orientiert sich an der bisherigen Dokumentensignatur in PDF und PDF/A und entspricht den Standards PKCS#7 (für PDF/A-1) [IE04] und PAdES Part 2 (für PDF/A-2 und -3) [ET09a]. Dabei wird nach der Erzeugung des visuellen Anteils und der Einbettung der CDA-XML-Datei das gesamte Dokument über eine Dokumentensignatur signiert und damit versiegelt. Diese Signatur kann sowohl sichtbar, d.h. mit einer grafischen Repräsentation behaftet sein, als auch unsichtbar sein. Der Vorteil dieser Variante liegt darin, dass alle vorhandenen PDF/A-konformen Signaturanwendungskomponenten (SAK) diese Vorgehensweise ohne Einschränkung unterstützen.

Bei Variante 2 wird auf die Signatur des visuellen Teils verzichtet. Die CDA-XML-Datei wird vor der Einbettung selbst signiert, die Signaturcontainer ins XML mit eingebettet und danach erst das resultierende XML in das PDF/A-3 Dokument eingebracht. Dementsprechend weist der visuelle Anteil keine Anzeichen einer Signatur auf. Damit folgt diese Vorgehensweise dem ETSI Standard PAdES Part 5 [ET09b]. Für die PDF/A-3-Konformität hat dies keine Auswirkungen. Der Vorteil dieser Variante ist die Unabhängigkeit des visuellen Teils. D.h. es können auf dem PDF/A-3-Dokument noch sichtbare Änderungen durchgeführt werden (z.B. in Form von Annotationen, Anmerkungen etc.), ohne die Signatur der transportierten CDA-XML-Daten zu beeinträchtigen. Nachteile sind die aktuell nicht vorhandene Unterstützung durch SAKs, auch wenn dies keinen größeren Anpassungsaufwand darstellt, und die nicht mehr vorhandene Korrespondenz des visuellen Anteils mit den XML-Quelldaten. Für ein echtes Straight-Through-Processing (STP) ist dies allerdings ein geeigneter Ansatz.

Auf Variante 3 trifft im Wesentlichen das bei Variante 2 gesagte zu. Es wird lediglich keine XML-Signatur verwendet, sondern eine konventionelle Binärsignatur nach PKCS#7 [Ho05] oder CAdES BES [ET08]. Auch hier muss die Signaturerzeugung und -validierung bestehender SAKs entsprechend angepasst werden. Der Vorteil dieser Variante liegt darin, dass das CDA-XML-Datei zusammen mit der abgesetzten Signaturdatei auch unabhängig vom PDF/A-3-Container weiterverarbeitet und die Signatur validiert werden kann.

3.3 Validierung eines hybriden Arztbriefs

Die Überprüfung eines hybriden Arztbriefs auf Einhaltung aller Vorgaben der Standards PDF/A-3, CDA und elektronischer Signatur wird als Validierung bezeichnet. Im Wesentlichen kann dies in drei Schritten erfolgen:

- Validierung des PDF/A-3-Formats mit Hilfe eines PDF/A-Validierungsprogramms,
- Validierung der CDA-Struktur über eine XSD-Schema-Verifikation,
- Validierung der elektronischen Signatur über eine Signaturverifikation mit Hilfe einer SAK.

In einem optionalen letzten Schritt kann auch eine visuelle Inhaltskontrolle durch nochmalige Anwendung der Transformation des CDA-Dokuments nach PDF (vgl. Kap. 3.1) stattfinden. Eine Automatisierung dieses Vorgangs bedarf allerdings noch weiterer Untersuchungsarbeit.

4 Zusammenfassung

Es wurde untersucht, ob sich PDF/A-3 als Trägerformat für hybride Dokumente und insbesondere CDA-XML-Arztbriefe, d.h. Dokumente mit einem visuellen und einem Nettodatenanteil, eignet. Dies kann abschließend bestätigt werden. Alle Anforderungen aus Kapitel 1.1 können mit diesem Format befriedigt werden. Zudem wurde der Normteil A-3 genau vor diesem Hintergrund entworfen. Auch die verschiedensten Formen der elektronischen Signatur eines solchen Dokuments können über PDF/A-3 standardkonform abgebildet werden. Zwar ist die Softwareunterstützung zur Erzeugung und Validierung noch nicht durchgängig vorhanden, doch ist nach der finalen Veröffentlichung des ISO-Standards PDF/A-3 Ende 2012 mit entsprechenden Angeboten im Markt zu rechnen. Zudem ist PDF/A-3 kein grundlegend neuer Standard, sondern unterscheidet sich vom Normteil A-2 nur in einem wesentlichen Punkt, der Einbettbarkeit von Fremddaten.

Für die Anwendung bei CDA-Arztbriefen erscheint PDF/A-3 geeignet. Über die weitere Verbreitung und Akzeptanz dieser Kombination im Gesundheitsmarkt werden die nächsten 2 bis 5 Jahre entscheiden.

Literaturverzeichnis

- [Ad01] Adobe Systems Incorporated, PDF Reference: Adobe Portable Document Format, Version 1.4, 3rd ed., 2004.
- [Al99] Alschuler, L. et. al.: http://www.hl7.org/documentcenter/public_temp_276613DF-1C23-BA17-0C94DA782413B8B0/wg/sgml/sgml.htm, 1999.
- [Ba01] Bates, DW.; Cohen, M.; Leape, L.; Overhage, M.; Shabot, MM.; Sheridan, Th.: Reducing the Frequency of Errors in Medicine using Information Technology, Journal of the American Medical Informatics Association 200; 8(4):299-308.

- [Do05] Dolin, A.H. et al.: HL7 Clinical Document Architecture, Release 2. Journal of the American Medical Informatics Association, Epub 2005 Oct 12.
- [ET08] ETSI Technical Specification, Electronic Signatures and Infrastructures (ESI); CMS Advanced Electronic Signatures (CAAdES) v1.7.4, European Telecommunications Standards Institute, Sophia Antipolis, 2008.
- [ET09a] ETSI Technical Specification, Electronic Signatures and Infrastructures (ESI); PDF Advanced Electronic Signature Profiles v1.1.1, European Telecommunications Standards Institute, Sophia Antipolis, 2009.
- [ET09b] ETSI Technical Specification, Electronic Signatures and Infrastructures (ESI); PDF Advanced Electronic Signature Profiles; Part 5: PAdES for XML Content - Profiles for XAdES signatures v1.1.2, European Telecommunications Standards Institute, Sophia Antipolis, 2009.
- [HL05] Health Level Seven: HL7 Data Model Development, <http://www.hl7.org/Library/datamodel/index.cfm>, 2005.
- [Ho04] Housley, R.: Cryptographic Message Syntax (CMS), IETF RFC 3852, <http://www.ietf.org/rfc/rfc3852.txt>, Juli 2004.
- [ISO05] ISO, Document management - Electronic document file format for long-term preservation - Part 1: Use of PDF 1.4 (PDF/A-1), International Organization for Standardization, Genf, 2005.
- [ISO11] ISO, Document management - Electronic document file format for long-term preservation - Part 2: Use of ISO 32000-1 (PDF/A), International Organization for Standardization, Genf, 2011.
- [Ob06] Oberacher, F.: Die XML basierte „Clinical Document Architecture“(CDA) für medizinische Befundberichte und Arztbriefe, UMIT, Hall (Tirol), 2006.
- [WW08] World Wide Web Consortium, XML Standard des Word Wide Web Consortiums, <http://www.w3.org/TR/2008/REC-xml-20081126>, 2008.
- [WW10] World Wide Web Consortium. Website: XSL Transformations (XSLT) Version 2.0, www.w3.org/TR/2010/REC-xslt-xquery-serialization-20101214, 2012.