Informationspakete zur Informations- und Beweiswerterhaltung – ein Vergleich

Steffen Schwalm[1] Dr. Ulrike Korte[2] Dr. Detlef Hühnlein[3] Mike Prechtl[3] Tomasz Kusber[4] Dr. Bernd Wild[5]
[1]BearingPoint GmbH
Steffen.Schwalm@bearingpoint.com
[2]Bundesamt für Sicherheit in der Informationstechnik
Ulrike.Korte@bsi.bund.de
[3]ecsec GmbH
Detlef.Huehnlein@ecsec.de
[4]Fraunhofer FOKUS
Tomasz.Kusber@fokus.fraunhofer.de
[5]Intarsys GmbH
bwild@intarsys.de

Zusammenfassung

Die [eIDAS-VO] schuf in EU und EFTA einheitliche Vorgaben für elektronische Transaktionen im Binnenmarkt, wie sie zur vertrauenswürdigen Abwicklung und Nachweis elektronischer Geschäftsprozesse genutzt werden können. Dies umfasst insbesondere Anforderungen für die beweiswerterhaltende Aufbewahrung gemäß Artikel 34 [eIDAS-VO]. Ausgehend von diesen juristischen und technischen Vorgaben wird die Bedeutung selbsttragender Informationspakete zur Informations- und Beweiswerterhaltung erläutert und Anforderungen an die Informations- und Beweiswerterhaltung im Allgemeinen, sowie an selbsttragende Archivinformationspakete im Speziellen abgeleitet. Anschließend werden die drei aktuell repräsentativen Informationspakete zur Informations- und Beweiswerterhaltung, das XAIP-Paket (vgl. [TR03125] und [ISO13527]), der ASiC-Container (vgl. [EN 319 162]) und der PDF/A-3-Container [ISO19005-3] vorgestellt, mit einander vergleichen, und es wird eine Bewertung mit Ausblick gegeben.

1. Einleitung

1.1. Anforderungen an die Informations- und Beweiswerterhaltung

Im Zuge der Umsetzung der [eIDAS-VO] werden „beim Erlass von delegierten Rechtsakten bzw. Durchführungsrechtsakten die von europäischen und internationalen Normungsorganisationen und -einrichtungen, insbesondere dem Europäischen Komitee für Normung (CEN), dem Europäischen Institut für Telekommunikationsnormen (ETSI), der Internationalen Normungsorganisation (ISO) und der Internationalen Fernmeldeunion (ITU), festgelegten Normen und technischen Spezifikationen gebührend berücksichtigt“ [eIDAS-VO]. Diese technischen Standards werden zudem zunehmend weltweit angewandt, was deren Bedeutung hervorhebt. Diese regulatorischen Vorgaben werden mit weiteren Regularien sowie branchenspezifische Festlegungen untermauert [KuScDoV15].

Die Basis vertrauenswürdiger, elektronischer Geschäftsprozesse, sowohl in Behörden als auch in Unternehmen, bildet demgemäß die Möglichkeit zu deren Prüfbarkeit und so dem Nachweis hierauf basierender Entscheidungen gegenüber Prüfbehörden, Gerichten und Dritten. Hierzu gilt es, die Integrität, Authentizität, Verfügbarkeit sowie Verkehrsfähigkeit der geschäfts-relevanten Unterlagen lückenlos bis zum Ablauf der geltenden Aufbewahrungsfristen (bis zu 100 Jahren und länger), die teilweise erst Jahrzehnte nach der eigentlichen Entscheidung beginnen (z.B. Pharma, Luftfahrt, Bauwesen, Transportwesen etc.), zu gewährleisten. Hierzu muss es möglich sein, die entsprechenden Dokumente den Prüfbehörden (z. B. Rechnungshof), Gerichten, Dritten vorzulegen, um die o.g. Nachweise zu führen, was deren Verkehrsfähigkeit erfordert. Das bloße Speichern auf einem Datenträger ist also nicht ausreichend. Nicht zuletzt muss die Verfügbarkeit und Lesbarkeit der aufbewahrten Dokumente innerhalb der Aufbewahrungsfrist gewährleistet werden, um die Dokumente überhaupt visualisieren bzw. nutzen zu können [KoSH13].
Digitale Dokumente liefern aus sich heraus keine Hinweise auf die Integrität und Authentizität und gewährleisten angesichts der technischen Entwicklung ohne entsprechende Maßnahmen nicht ihre langfristige Nutzbarkeit. Um die o.g. Anforderungen erfüllen zu können, muss eine beweissichere Langzeitspeicherung die folgenden Aufgaben erfüllen:

- Informationserhaltung – die Aufbewahrung der Dokumente erfolgt unter Verwendung von sog. archivfähigen oder langzeitstabilen Formaten (z .B. PDF/A, JPEG 2000 etc.).
- Die aufbewahrten Primärinformationen (Inhaltsdaten), also die Dokumente selbst, werden einschließlich der zugehörigen beschreibenden Informationen zum Geschäftskontext oder zur technischen Umgebung (Formatidentifikation, Softwareumgebung etc.) (Metadaten) in zusammenhängenden Archivinformationspaketen (Container, Informationspakete) selbsttragend aufbewahrt und gegen Verlust geschützt.
- Beweiswerterhaltung – die Integrität und Authentizität wird durch den Einsatz von kryptographischen Sicherungsmittel hergestellt.

Zu den wesentlichen Standards, die die Umsetzung der o.g. Anforderungen unterstützen, gehören u. a. [KoSH13]:

- [ISO14721] – Beschreibung der notwendigen Funktionen und Informationspakete für die Implementierung dauerhafter Aufbewahrung elektronischen Unterlagen.
- [ISO14533]-{1,2,3} – Aufbau von Langzeitaufbewahrungsformate für CAdES, XAdES und PAdES Signaturen.
- [DIN31644] und [DIN31647] – Definition der Anforderungen an ein vertrauenswürdiges digitales Langzeitarchiv (dLZA), sowie fachliche und funktionale Anforderungen an ein generisches System zur Beweiswerterhaltung von kryptographisch signierten Unterlagen.
- [RFC4998], [RFC6283] sowie [TR03125] – insbesondere hier die TR (basierend auf den beiden RFCs) beschriebt eine Referenzarchitektur, notwendige Funktionen und Module zur Umsetzung von Beweiswerterhaltung der kryptographisch signierten Dokumente.

Dieses Vorgehen hat sich in der Praxis branchenübergreifend in zahlreichen Projekten bewährt, in denen sowohl Einzelinstallationen als auch komplexe Dienste zur beweissicheren Langzeitspeicherung im SOA-Sinne umgesetzt wurden, so z.B. im Digitalen Zwischenarchiv des Bundes bei der Bundesagentur für Arbeit, in Langzeitspeichersystemen verschiedener Bundesbehörden, Landesbehörden mehrerer Bundesländer sowie im Gesundheitswesen, in der Luftfahrt oder auch in Banken und Versicherungen.

Derzeit werden selbsttragende AIP[1] vielfach als XML-Informationspakete technisch realisiert, wie dies z.B. auch die [TR03125] vorsieht. Die Aufbewahrung eingebetteter Binärdaten in XML-Strukturen erfolgt dabei i.d.R. durch Base64-Codierung, was eine Steigerung der Daten-menge um ca. 33 % nach sich zieht. Die Verarbeitung solcher XML-basierter, selbsttragender Archivinformationspakete (AIP) hat sich in der Praxis bewährt, sofern keine Anforderungen an eine sehr hohe Performanz für sehr große Datenvolumina oder für hochfrequente Lese- wie Schreibvorgänge vorliegen. In einigen Fällen, so z.B. Geodaten, übersteigt das Datenvolumen bei einem selbsttragenden AIP in Form eines XML-Containers die Kapazitäten marktüblicher Schnittstellen, so dass weitere Lösungen in die Betrachtungen einzubeziehen sind. Eine mögliche Alternative ist der Einsatz standardisierter Container-formate, die binäre Daten ohne zusätzliche Kodierung aufnehmen können oder diese verlustfrei komprimieren können. Hier bieten sich derzeit das PDF/A-3 (vgl. [ISO19005-3]) oder das auf dem ZIP-Format basierende ASiC (vgl. [EN319162]) an. Um im Anwendungsfall eine gezielte Auswahl treffen zu können, gilt es, zunächst die Anforderungen an Formate für selbsttragende Archivinformationspakete zu benennen (Kap. 1.2). Im nächsten Schritt werden die verschiedenen Optionen möglicher Formate beschrieben (Kap. 2) und schlussendlich die vorgestellten Formate anhand der Anforderungen an AIPs bewertet sowie ein Fazit für die praktische Anwendung gezogen.

1.2. Anforderungen an Archivinformationspakete

Dabei haben sich folgende Anforderungen herausgestellt. Ein Archivinformationspaket soll

1. beliebig viele, verschiedene Inhaltsdaten, Metadaten und beweisrelevante Daten/ technische Beweisdaten enthalten;
2. signierte/ zeitgestempelte und un-signierte/ un-zeitgestempelte Daten parallel enthalten;
3. verschiedene digitale Signaturtechniken (z.B. (AdES) Signaturen, Zeitstempel, Evidence Records) unterstützen;
4. parallele und zusätzliche Signaturen unterstützen;
5. nachträgliches Einspielen von Sperrmaterial etc. ohne Zerstörung der voraus-gegangenen Signatur ermöglichen;
6. verschiedene Versionen von Inhalts-/ Meta- und Beweisdaten enthalten, die nach und nach in das Paket eingestellt werden können;
7. eine Referenzmöglichkeit (z.B. ein Manifest-Objekt oder ein Indexobjekt) enthalten können, dass entsprechende Nutzdaten-, Metadaten- und beweisrelevanten Daten/ Beweisdaten verknüpft, die durch ein spezielles Objekt der digitalen Signaturtechnik (z.B. digitale Signatur im AdES-Format, Zeitstempel, Evidence Record) geschützt werden;
8. unterstützen, dass Versionen oder spezielle Elemente des AIP gezielt ausgelesen werden können;
9. performant auch bei großen Datenmengen/ Dateien (z.B. Geodaten) sein;

[1] Unter einem selbsttragenden Archivinformationspaket im Sinne der [ISO-14721] i.V.m. der [DIN 31644] sowie der [DIN 31647] wird ein Container verstanden, der alle zur alle Informationen zur Interpretation, Lesbarkeit, Nutzbarkeit, Verständlichkeit, Recherche und zu den beweisrelevanten Nachweisen der Integrität und Authentizität der aufzubewahrenden Unterlagen in standardisierter und herstellerneutraler Form enthält, also Metadaten, Inhalts-/Primärdaten sowie die zum langfristigen Nachweis von Authentizität und Integrität notwendige Daten (beweisrelevante Daten und technische Beweisdaten) enthält. [KoSH13]

10. eine wirtschaftliche Realisierung (z.B. auf Basis von Open Source Softwarekomponenten, serviceorientierten Architekturen oder Softwarekomponenten mit kostengünstigen Lizenz- und Servicemodellen) ermöglichen;
11. selbsttragend sein, also alle zur Informations- und Beweiswerterhaltung, also dem Aufbewahrungszweck, notwendigen Daten in standardisierter Form enthalten;
12. auf weit verbreiteten, offenen Standards weithin anerkannter, unabhängiger Standardisierungsgremien beruhen.

2. AIP-Formate zur Informations- und Beweiswerterhaltung

2.1. XFDU, XAIP

Das XAIP-Paket (XML formatted Archival Information Package) ist ein XML-basiertes, in [TR03125-F] spezifiziertes, selbstbeschreibendes Archivinformationspaket (AIP). Es enthält die zu archivierenden Daten (Primärinformationen), die Metainformationen, die für eine „rechts- und revisionssichere Rekonstruktion von Geschäfts- oder Verwaltungsvorgängen bis zum Ablauf der gesetzlich vorgeschriebenen Aufbewahrungsfristen erforderlich sind“ sowie die zur Beweiswerterhaltung notwendigen Daten. Das XAIP-Paket kann gegen ein „gültiges und autorisiertes XML-Schema geprüft werden“ [TR03125-Schema]. Die Grundlage für die XAIP-Struktur bilden vor allem das OAIS Referenzmodell [ISO14721], die Victorian Electronic Records Strategy [VERS] und der XML Formatted Data Unit Standard [ISO13527].

Das XAIP-Paket gemäß [TR03125-F] beinhaltet neben dem Inhaltsdaten-Abschnitt (dataObjectsSection), dem Metadaten-Abschnitt (metaDataSection) und dem Beweisdaten-Abschnitt (credentialsSection) einen Kopfbereich (packageHeader), welcher Informationen zum AIP in seiner Gesamtheit und seinem Aufbau bereitstellt. Dabei kann es sich beispielsweise u.a. um einen eindeutigen Identifikator für das Archivdatenobjekt, um Angaben zum zugrundeliegenden XML-Schema, zur Aufbewahrungsfrist, zum Absender des XAIP und um versionsspezifische Inhaltsverzeichnisse mittels XML-basierter Manifeste handeln.

Weiterhin beinhaltet das XAIP-Paket einen Abschnitt für Metainformationen. Diese Metainformationen beschreiben den Inhalt des XAIP, der zur Interpretation des Geschäfts- und Archivierungskontextes des XAIP benötigt wird. Dabei ist es u.a. auch möglich, die in [ISO13527] definierten „category“- und „classification“-Attribute zu nutzen oder fachspezifische bzw. technische Metadaten gemäß [XBARCH], [XDOMEA] und [PREMIS] zu integrieren.

Der Original-Datenabschnitt kann beliebig viele Inhaltsdatenobjekte enthalten. Er „kann beispielsweise dafür genutzt werden, Inhaltsdaten in verschiedenen plattform- oder anwendungsspezifischen Datenformaten in einem XAIP-Container zu speichern oder ganze Akten mit vielen unterschiedlichen Dokumenten gemeinsam zu archivieren“ ([TR03125-F], S.9). Dabei wird für jedes Inhaltsdatenobjekt ein eindeutiges Identifikationsmerkmal vergeben [TR03125-F]. Die Inhaltsdaten selbst können als Base64-kodierte Binärdaten oder als XML-Struktur abgelegt werden. Außerdem ist es möglich, eine Prüfsumme und Informationen über Transformierungen eines Datenobjektes hinzuzufügen [TR03125-F].

Der letzte Abschnitt des XAIP-Paketes ist der Beweisdaten-Abschnitt, der beweisrelevante Daten, wie z.B. digitale Signaturen, elektronische Zeitstempel, Zertifikate, Zertifikatstatusinformationen, „Verification Reports“ gemäß [OASIS-VR], und/ oder technische

Beweisdaten in Form von einem oder mehreren Evidence Record(s) [RFC4998/RFC6283], beinhalten kann.. Diese beweisrelevanten Daten und technischen Beweisdaten dienen dem Nachweis der Unversehrtheit (Integrität), der Authentizität (sofern die Inhaltsdaten signiert sind) und des „Proof of Existence“ der im XAIP enthaltenden archivierten Datenobjekte.

Jeder dieser Bereiche enthält spezifische Datenelemente, die innerhalb des XAIP eindeutig über eine ID referenziert werden können (Bsp.: die *dataObjectID* eines *DataObjects* in der *DataObjectSection*). Zum Auslesen eines XAIP sind die *VersionManifest*-Elemente im *packageHeader* wichtig, die jeweils ein versionsspezifisches Inhaltsverzeichnis darstellen. Die *VersionManifest*-Elemente verweisen auf die zu einer Version gehörenden Datenelemente des *XAIPs* und legen fest, welche Datenelemente in die Bildung der Hashwerte einfließen.
Dabei darf nur auf die direkt in diesem XAIP enthaltenen Datenelemente verwiesen werden.
Der Verweis auf die einer Version zugehörigen Datenelemente im *VersionManifest* kann mittels zweier Pointer-Typen vorgenommen werden:

- protectedObjectPointers:
  „Durch die Menge der hier angegebenen <*protectedInfoPointer*>-Elemente wird definiert, welche Teile des Archivdatenobjektes in die Hashwertbildung einfließen und deshalb vom entsprechenden Evidence Record geschützt werden.“ ([TR03125-F], S. 14).
- unprotectedObjectPointers:
  “Durch einen hier angegebenen <*unprotectedInfoPointer*> wird klargestellt, dass das Objekt auf das hier verwiesen wird logisch zur angegebenen *XAIP*-Version gehört. Allerdings fließt dieses Objekt *nicht* in die Hashwertbildung ein und es ist deshalb nicht von einem Evidence Record umfasst und kryptographisch geschützt.“ ([TR03125-F], S. 14). Dies ermöglicht eine spätere Änderung der Daten dieser Version ohne vorhandene Evidence-Records zu zerstören.

Das Versionsmanifest ist in Abbildung 1 dargestellt.

2.2.ZIP, ASiC

Der in [EN319162] spezifizierte „Associated Signature Container“ (kurz ASiC) ist ein Daten-Container, der unbegrenzt viele Datenobjekt-Dateien und die dazugehörigen Signaturen und/ oder „time assertions“ (d.h. Zeitstempel oder Evidence Records) im ZIP-Format gemäß [ISO21320-1] enthalten kann. Ein ASiC-Container beinhaltet ein Stammverzeichnis, das wiederum Unterverzeichnisse enthält, die den Inhalt, d.h. die Datenobjekte-Dateien, repräsentieren. Weiterhin beinhaltet ein ASiC-Container ein Unterverzeichnis „META-INF“. Dieses Verzeichnis enthält die Metadaten zu den Datenobjekten und die dazugehörigen Signaturen, Zeitstempel und Evidence Records. ETSI definiert zwei Typen von ASiC-Containern.

Der **ASiC-S-Container** (**S**imple) beinhaltet nur eine Datenobjektedatei und eine Signaturdatei, die möglicherweise mehrere Signaturen beinhalten kann, oder eine „time assertion“-Datei. Die Signaturdatei kann eine oder mehrere CAdES- oder XAdES-Signaturen enthalten. Weiterhin kann sich auch eine „time assertion“-Datei im META-INF-Ordner befinden. Der ASiC-S-Container verwendet als Dateiendung „.asics“ oder „.scs“. Außerdem wird „application/vnd.etsi.asic-s+zip“ als MIME-Type verwendet.

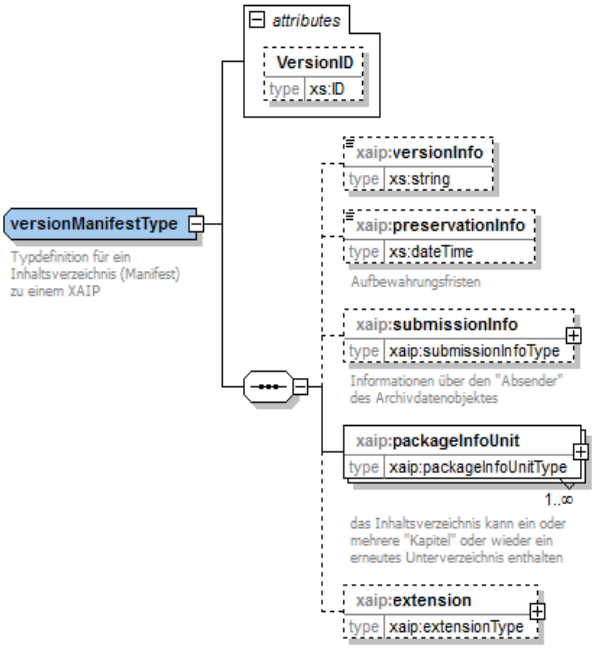

Abb. 1: Aufbau eines Elements vom Typ versionManifestType

Der zweite Typ, der **ASiC-E-Container** (**E**xtended), beinhaltet mehrere Datenobjekte, sowie eine oder mehrere Signatur-Datei(en) und/oder „time assertions"-Datei(en), die jeweils auf eine Untermenge der Datenobjekte anwendbar sind. Jedes Datenobjekt kann mit einer oder mehreren CAdES- oder XAdES-Signaturen oder Zeitstempeln oder Evidence Records assoziiert sein. Der ASiC-E-Container verwendet als Dateiendung „.asice" oder „.sce" und als MIME-Type soll „application/vnd.etsi.asic-e+zip" verwendet werden. Der MIME-Type des verwendeten Containers kann in einer MIME-Type-Datei, die sich im Stammverzeichnis befindet, angegeben sein. Es kann eine „container.xml"-Datei im Stammverzeichnis abgelegt sein. Diese Datei verweist auf eine Root-Datei. Beim ASiC-E-Container wird im Fall „ASiC-E mit XAdES" (siehe [EN319162], Abschnitt 4.4.3) zusätzlich die Datei „manifest.xml" im META-INF-Ordner abgelegt. Dieses Manifest beinhaltet Verweise auf die Datenobjekte und die dazugehörigen Signaturen und/oder Evidence Records. Im Fall ASiC-E mit CAdES-„time assertions" sind eine oder mehrere ASiCManifest- oder ASiCEvidenceRecordManifest-Dateien vorhanden, die jeweils eine Signatur- oder Zeitstempel- oder Evidence Record-Datei und mehrere Datenobjekt-Dateien referenzieren.

Alle ASiC-Typen erlauben eine Verschachtelung („nesting") von Containern, wobei die inneren Container wieder ASiC-Container oder andere Container-Typen darstellen können. Für die Sicherstellung der Langzeit-Verfügbarkeit und –Integrität gibt es in [EN319162] unterschiedliche Ansätze:

- Im Fall von ASiC-E-Containern mit XAdES-Signaturen mittels der in [EN319132] beschriebenen Mechanismen oder auf Basis der Evidence Record Spezifikationen [RFC4998] bzw. [RFC6283].
- Im Fall von ASiC-E-Containern mit CAdES-„time assertions" mittels einer oder mehreren ASiCArchiveManifest-Datei(en) und jeweils einem dazu gehörenden Zeitstempel oder einer oder mehreren ASiCEvidenceRecordManifest-Datei(en) und dazu gehörenden Evidence Records.

- Im Fall von ASiC-E-Containern mit Evidence Records mit den internen Mechanismen von [RFC4998] bzw. [RFC6283].

Ein Manifest in ASiC-E enthält jedoch nur den Verweis auf eine Signatur oder einen „time assertion“ (Zeitstempel, oder Evidence Record), weshalb unter Umständen mehrere Manifeste benötigt werden.

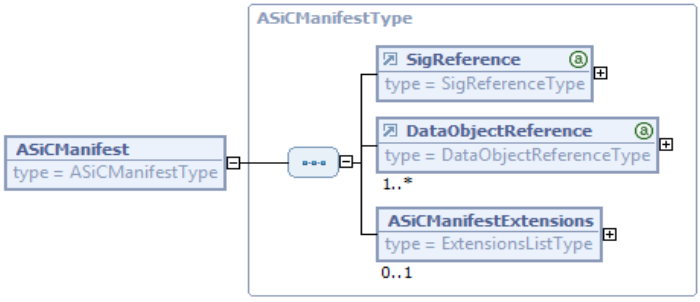

Abb. 2: Struktur des ASiCManifest-Elementes

Neben dem, im Vergleich zu XAIP, geringeren Speicherplatzbedarfs und der größeren Performanz, ist bei diesem Format positiv hervorzuheben, dass sowohl ASiC-S- als auch ASiC-E-Container im Grundsatz auf dem ZIP-Format gemäß [ISO21320-1] beruhen und somit von entsprechenden Werkzeugen grundsätzlich verarbeitet werden können. ASiC-E-Container gemäß [EN319162-1], Abschnitt 4.4.4 (ASiC-E with time assertions), sind grundsätzlich sehr gut zur beweiskräftigen Aufbewahrung geeignet, da durch die Manifest-Datei die Relation zwischen einem Evidence Record und den von ihm geschützten Dokumenten klar spezifiziert werden kann. Unglücklicherweise ist diese für die beweiskräftige Aufbewahrung ideal geeignete ASiC-E-Variante aktuell nicht unter den „Baseline-Varianten“ in EN319162-1] Abschnitt 5 aufgeführt, was im Zuge der weiteren Standardisierung vor einer Fortschreibung des Durchführungsbeschlusses [2015/1506/EU] korrigiert werden sollte

2.3.PDF/A-3

Unter PDF/A werden Konformitätsstufen des PDF Dokumentenformats [ISO320001] verstanden, die den besonderen Anforderungen an langzeitstabile Archivierbarkeit Rechnung tragen. Mit PDF/A-1, das 1995 als ISO-Standard 19005-1 [ISO19005-1] veröffentlicht wurde, wurde erstmalig ein klares und prüfbares Profil geschaffen, das Eigenschaften beschreibt, die ein PDF Dokument aufweisen muss, oder im Gegenteil nicht aufweisen darf, um als PDF/A-konform und damit langzeitstabil zu gelten. Im Wesentlichen sind dies der Verzicht auf jegliche aktiven und dynamischen Inhalte, wie z.B. JavaScripts, oder ausführbare Einbindungen, sowie die konsequente Ausprägung als selbsttragendes unverschlüsseltes Dokument ohne externen Referenzen, wie Zeichensätze und Farbräume, dass alle für die originalgetreue visuelle Reproduktion notwendigen Beschreibungen enthält.

Mit der 2012 erschienenen Konformitätsstufe PDF/A-3 [ISO19005-3] wurde eine entscheidende Einschränkung der niedrigeren Konformitätsstufen A-1 und A-2 aufgehoben, die Einbindbarkeit von Fremdformaten. Damit kann ein PDF/A-3 Dokument über eine genau spezifizierte Struktur beliebige Dateien und Formate als sogenannte File Attachments einbinden. Diese Struktur wird auch im kommenden ISO-Standard PDF 2.0 bzw. ISO 32000-2 die Grundlage der Dateieinbettung bilden. Über den Inhalt der eingebundenen Dateien macht der Standard mit Ausnahme dem Vorhandensein eines MIME-Typs keine Vorgaben.

Die eingebetteten Dateien können ZIP-komprimiert gespeichert werden, so dass ein PDF/A-3 Dokument auch als ZIP-Container mit einer visuellen Darstellung betrachtet werden kann.
Darüber hinaus ist die Beziehung, in welcher die eingebettete Datei zum PDF/A Dokument steht, zwingend anzugeben. Der Standard unterscheidet hier zwischen Alternative, Source, Data, oder Supplement. Ein weiterer Aspekt, der bei allen PDF/A-Konformitätsstufen eine zwingende Voraussetzung darstellt, ist das Vorhandensein entsprechender Metadaten im XMP-Format [ISO16684-1]. Neben der Verwendung von zehn, bereits im XMP-Standard enthaltenen Metadaten-Schemata (z.B. Dublin-Core), erlaubt PDF/A-3 die Verwendung beliebiger eigener Metadaten, sofern sie über ein PDF/A-Erweiterungsschema definiert wurden. Entscheidend ist dann die Einbettung der Metadaten zusammen mit dem zugrundeliegenden Erweiterungsschema im XMP-Datenblock des PDF/A-3-Dokuments.

Die Absicht hinter der Schaffung der Einbettungsmöglichkeit war primär nicht die Bereitstellung eines Archivcontainers, sondern die Anreicherung der visuellen Repräsentation um Metadaten, die prozessual und maschinell ausgewertet werden können. Ein Beispiel dafür ist das Format ZUGFeRD 1.0 für elektronische Rechnungen, bei dem neben der gewohnten visuellen und druckfähigen Darstellung der Rechnung im PDF/A-Teil auch eine maschinenlesbare XML-Datei mit den Rechnungsdaten als File Attachment in der Beziehung „Alternative“ beigefügt ist, um eine automatisierte Verarbeitung der Rechnung zu ermöglichen. Allerdings bieten die o.g. Eigenschaften des Formats grundsätzlich die Option, PDF/A-3 auch als Archivcontainer zu verwenden.

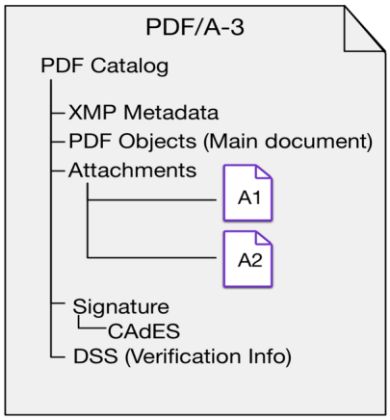

Abb. 3: Struktur eines signierten PDF/A-3 Dokuments mit 2 Dateianhängen

Abb. 3 stellt ein signiertes PDF/A-3-Dokument mit 2 Dateianhängen (A1 und A2) dar. Ein PDF-Viewer visualisiert nur die Objekte des Hauptdokuments (PDF Objects). Unabhängig von der Dateieinbettung erlaubt PDF/A-3 alle im PAdES-Standard von ETSI [EN319142] definierten Formen der eingebetteten elektronischen Signatur. Die Signatur bezieht sich dabei immer auf das gesamte PDF/A-Dokument und schließt auch sämtliche Dateieinbettungen mit ein. Im Baseline-Profil PAdES-B-LTA wird so auch explizit die auf Langzeitspeicherung ausgerichtete Versiegelung über wiederholt aufgebrachte Zeitstempel und Einbettung von Validierungsinformationen unterstützt. Abb. 3 zeigt beispielhaft die Struktur eines PDF/A-3 Dokuments mit 2 eingebetteten Dateien.

3. Bewertung der verschiedenen Optionen

Anforderung	**XAIP**	**ASiC**	**PDF/A-3**
Beliebige viele Nutzdaten, Metadaten, beweisrelevante Daten/ Beweisdaten,	**Ja;** als Objekte in der jeweiligen Sektion des XAIP	Ja, als Dateien	**Ja;** Nutzdaten im Hauptdokument oder über Dateianhänge Metadaten in Konformität zu XMP Beweisdaten werden als Datei-Anhänge behandelt.
Signierte/zeitgestempelte und unsignierte/nicht zeitgestempelte Daten	**Ja** **Herstellung** Zusammenhalt zwischen den Daten und genutzten Signaturtechniken mittels eines VersionManifest	**Ja** Herstellung Zusammenhalt zwischen den Daten und genutzten Signaturtechniken mittels ASiC-Manifest	**Ja;** Sowohl bei Hauptdokument als auch Attachments **Aber:** bei nachträglichen Signieren oder Zeitstempeln werden jeweils alle Objekte im Container signiert bzw. zeitgestempelt.
Verschiedene Signaturtechniken erzeugen und verifzieren können	**Ja,** AdES-Signaturen, - Zeitstempel, Evidence Recordes	**Ja,** AdES-Signaturen, -Zeitstempel, Evidence Records	**Eingeschränkt;** PAdES Signaturen und PAdES Zeitstempel Andere AdES-Signaturen möglich, jedoch nur via anwenderspezifischer Umsetzung
Parallele Signaturen	**Ja**	**Ja**	**Eingeschränkt;** Realisierbar über Einbettung von CAdES-Dateien; bislang keine Visualisierung oder Validierung im Standard definiert
nachträgliches Einspielen von Sperrmaterial ohne Verletzung Beweiswert	**Ja**	**Ja**	**Eingeschränkt:** Nur bei Verwendung des PAdES-LTA-Profils möglich; bedingt immer einen neuen Zeitstempel.

Anforderung	**XAIP**	**ASiC**	**PDF/A-3**
Versionierung Nutz-, Meta- und Beweisdaten	**Ja** Herstellung eines versionierten Zusammenhalts zwischen Daten und genutzten Signaturtechniken mittels jeweils einem VersionManifest pro Version. Unter Verwendung von mehreren VersionManifesten können mehrere Zusammenhalte mit einer erkennbaren Reihenfolge in einem XAIP Container abgelegt werden,	**Eingeschränkt:** Nicht-versionierte Herstellung eines Zusammenhalts zwischen Daten und genutzten Signaturtechniken mittels jeweils einem ASiC-Manifest. Unter Verwendung von mehreren ASiC-Manifesten können mehrere Zusammenhalte in einem ASiC-Container abgelegt werden, ohne dass eine Reihenfolge unter den Zusammenhalten erkennbar ist.	**Eingeschränkt:** Update-Mechanismus über Anhängen neuer bzw. geänderter Objekte ans Ende der Datei; das Ergebnis umfasst alle Bestandteile der PDF-Datei
Performant auch bei großen Datenmengen/Dateien z.B. Geodaten	**Eingeschränkt:** Einbettung in XML erfordert Nutzung Bas64-Codierung	**Ja**	**Ja** Speicherung von Daten komprimiert (ZIP). Schnelles Lesen durch direkte Sprünge auf die betreffende Seite
Wirtschaftlich	**Ja:** Keine zusätzlichen Lizenzpflichten	**Ja** Effiziente Speicherform durch Kompression; Keine zusätzlichen Lizenzpflichten	**Ja** Effiziente Speicherform durch Kompression; non-proprietäre Viewer; etabliertes Dateiformat **Aber:** Lizenzpflichtig
Selbsttragend	Ja	Ja	**Ja** Grundprinzip von PDF/A

Tabelle 1: Bewertung der verschiedenen Optionen

4 Fazit und Ausblick

Im vorstehenden Text wurde untersucht, ob neben XAIP sowohl ASiC als auch PDF/A-3 als Containerformat für AIPs zur beweissicheren elektronischen Langzeitspeicherung in Frage kommen können. Bei beiden Standards wurde geprüft, inwieweit die Formate äquivalent zu XAIP zur Abbildung selbsttragender Archivdatenobjekte genutzt werden können. Im Vergleich zu XAIP gibt es im ASiC- und PDF/A-3-Fall jedoch Einschränkungen. So lässt ASiC für den interoperablen ASiC-Baseline-Container keine Zeitstempel, EvidenceRecords und mehrere Manifeste zu.

PDF/A-3 erlaubt im Standard nur PAdES-Signaturen, nicht aber die anderen beiden Signatur-Formate gemäß [2015/1506/EU]. Grundsätzlich können auch weitere Signaturformate in PDF/A-3 eingebunden werden, ohne den PDF/A-3-Standard zu verletzen. In diesem Fall handelt es sich jedoch um anwenderspezifische Lösungen, bei denen zur Gewährleistung der Migrationsfähigkeit und Interoperabilität zusätzliche technische Beschreibungen zur Interpretation der im Container enthaltenen Daten anhand technischer Metadaten und die zusätzliche Bereitstellung spezifischer Softwarebausteine erforderlich sind. In XAIP dagegen sind alle in der [eIDAS-VO] definierten Signaturformate möglich, was eine breitere Einsatzfähigkeit verspricht. Eine PAdES-Document-Signatur bezieht sich zudem stets auf ein Dokument. Eine Möglichkeit, mittels einer Manifest-Struktur viele Nutz- und Meta-Daten mit einer Signatur, einem Zeitstempel oder einem Evidence Record in PDF/A-3 zu sichern, ist hier gegenwärtig nicht vorgesehen und nicht als Standard ausformuliert.

Daneben ist in PDF/A-3 derzeit nicht spezifiziert, wie in einem dLZA auf einzelne Datenelemente im Container zugegriffen werden kann, ohne in jedem Fall den kompletten Container abzurufen – im Gegensatz zu ASiC und AIP, wo dies in [TR03125-E] bzw. [ISO-21320-1] derzeit eindeutig spezifiziert ist.

Die Stärke des PDF/A-3 liegt in der Einbettungsmöglichkeit der visuellen Repräsentation von Metadaten, die bei einer automatisierten Dokumentenverarbeitung ausgewertet werden können. Insbesondere in einer homogenen PFD-Umgebung ist PDF/A-3 daher auch für die beweiswerterhaltende Langzeitspeicherung von einzelnen Dokumenten geeignet.

Um sowohl PDF/A-3 als auch ASiC als Archivinformationspakete zu nutzen und vor allem die Migration untereinander sowie mit XAIP zu gewährleisten, wären weitere detailliertere Betrachtungen und Standardisierungsarbeiten notwendig. Daneben wären die Interoperabilität und eine Migrationsfähigkeit zwischen XAIP-, ASiC- und PDF/A-Strukturen sowohl für Bestandsdaten als auch für neu zu archivierende Unterlagen in einem zweiten Schritt im technischen Detail zu untersuchen.

Literatur

[2015/1506/EU] DURCHFÜHRUNGSBESCHLUSS (EU) 2015/1506 DER KOMMISSION zur Festlegung von Spezifikationen für Formate fortgeschrittener elektronischer Signaturen und fortgeschrittener Siegel, die von öffentlichen Stellen gemäß Artikel 27 Absatz 5 und Artikel 37 Absatz 5 der Verordnung (EU) Nr. 910/2014 des Europäischen Parlaments und des Rates über elektronische Identifizierung und Vertrauensdienste für elektronische Transaktionen im Binnenmarkt anerkannt werden, 8. September 2015

[DIN31644] DIN 31644:2012 Information und Dokumentation – Kriterien für vertrauenswürdige digitale Langzeitarchive. 2012

[DIN31647] DIN 31647:2015 Information und Dokumentation – Beweiswerterhaltung kryptographisch signierter Dokumente. 2015

[eIDAS-VO] VERORDNUNG (EU) Nr. 910/2014 DES EUROPÄISCHEN PARLAMENTS UND DES RATES über elektronische Identifizierung und Vertrauensdienste für elektronische Transaktionen im Binnenmarkt und zur Aufhebung der Richtlinie 1999/93/EG vom 23.07.2014

[EN319122] ETSI EN 319 122 – {1,2}, Electronic Signatures and Infrastructers (ESI); CAdES digital signatures, ETSI V1.1.1, (2016-04

[EN319122-3] ETSI EN 319 122 – {1,2}, Electronic Signatures and Infrastructers (ESI); CAdES digital signatures, Part 3: Incorporation of ERS mechanisms in CAdES", ETSI V1.1.1, (2017-01)

[EN319132] ETSI EN 319 132 – {1,2}, Electronic Signatures and Infrastructures (ESI); XAdES digital signatures, ETSI V1.1.1, (2016-04

[EN319142] ETSI EN 319 142 – {1,2}, Electronic Signatures and Infrastructeres (ESI); PAdES digital Signatures, ETSI V1.1.1 (2016-04)

[EN319162] ETSI EN 319 162 – {1,2}, Electronic Signatures and Infrastrucutres (ESI); Associated Signature Containers (ASiC), ETSI V1.1.1 (2016-04)

[ISO13527] ISO 13527:2010, Space data and information transfer systems -- XML formatted data unit (XFDU) structure and construction rules, 2010

[ISO14533-1] ISO 14533-1, Processes, data elements and documents in commerce, industry and administration – Long term signature profiles, 2014

[ISO14533-2] ISO 14533-2, Processes, data elements and documents in commerce, industry and administration – Long term signature profiles, 2012

[ISO14721] ISO 14721, Space data and information transfer systems -- Open archival information system (OAIS) -- Reference model, 2012

[ISO16684-1] ISO 16684-1, Graphic technology – Extensible metadata platform (XMP) specification – Part 1: Data model, serialization and core properties, 2012

ISO19005-1] ISO19005-1, Document management — Electronic document file format for long-term preservation – Part 1: Use of PDF 1.4 (PDF/A-1), 2005

[ISO19005-3] ISO19005-3, Document management — Electronic document file format for long-term preservation – Part 3: Use of ISO 32000-1 (PDF/A-3), 2012

[ISO21320-1] ISO/IEC 21320, Information technology — Document Container File — Part 1: Core, 2015

[ISO32000-1] ISO 32000-1, Document management — Portable document format – Part 1: PDF 1.7, 2008

[KoSH13] U. Korte, S. Schwalm, D. Hühnlein, Vertrauenswürdige und beweiswerterhaltende Langzeitspeicherung auf Basis von DIN 31647 und BSI TR-03125, Informatik 2013, GI-LNI, P220, ISBN 978-3-88579-614-5, S. 550-566, 2013

[KoKuSH14] U. Korte, T. Kusber, S. Schwalm, D. Hühnlein, Standards und Lösungen zur langfristigen Beweiswerterhaltung, DACH-Security 2014, S. 46-58, 2014

[KuScDoV15] T. Kusber, S. Schwalm, A.Dörner, T.Vogt, Die Bedeutung der eIDAS-Verordnung für Unternehmen und Behörden. Neue Chancen und Herausforderungen für vertrauenswürdige elektronische Geschäftsprozesse in Europa, Berlin, 2015

[OASIS-VR] Hühnlein, D., OASIS DSS v1.0 Profile for Comprehensive Multi-Signature Verification Reports Version 1.0, Committee Specification 01, http://docs.oasis-open.org/dss-x/profiles/verificationreport/oasis-dssx-1.0-profiles-vr-cs01.pdf, 2010[PREMIS] Library of Congress: Preservation Metadata Maintenance Activity, http://www.loc.gov/standards/premis/

[RFC4998] IETF, T. Gondrom, R. Brandner, U. Pordesch, Evidence Record Syntax, 2007

[RFC6283] IETF, A. J. Blazic, S. Saljic, T. Gondrom, Extensible Markup Language Evidence Record Syntax (XMLERS), 2011

[SR019510] ETSI SR 019 510, Electronic Signatures and Infrastructures (ESI); Scoping Study and framework for standardization of long-term data preservation services, including preservation of/with digital signatures, ETSI V1.1.1 (2017-05)

[TR03125] BSI TR 03125, Beweiswerterhaltung kryptographisch signierter Dokumente (TR-ESOR), https://www.bsi.bund.de/DE/Publikationen/TechnischeRichtlinien/tr03125/index_htm.html, 2014

[TR03125-E] BSI TR 03125, Beweiswerterhaltung kryptographisch signierter Dokumente, Anlage TR-ESOR-E: Konkretisierung der Schnittstellen auf Basis des eCard-API-Frameworks, Version 1.2, 2015

[TR03125-F] BSI Technische Richtlinie 03125. Beweiswerterhaltung kryptographisch signierter Dokumente. Anlage TR-ESOR-F: Formate. Version 1.2, Stand: 31.01.2015

[TR03125-Schema]https://www.bsi.bund.de/DE/Publikationen/TechnischeRichtlinien/tr03125/index_htm.html#doc6617308bodyText11

[VERS] Victorian Electronic Records Strategy, siehe unter http://prov.vic.gov.au/government/vers

[XBARCH] https://www.bundesarchiv.de/fachinformationen/00895/index.html.de, Version 1.4.3

[XDOMEA] https://www.xrepository.de/Inhalt/urn:uuid:0e13664e -6df5-4d1f-8397-elee-d87a0d4a.xhtml, Version 2.2.0