# „E-ARK / eArchiving Building Block – auch für uns in Deutschland?“

Tagungsbericht des Bundesarchivs, 16. August 2019

**Beim Workshop des Bundesarchivs wurde das für die digitale Archivierung wichtige Thema „Informationspakete“ ausführlich diskutiert**

Am 25. April 2019 lud das Bundesarchiv in Berlin zum Workshop „E-ARK / eArchiving Building Block – auch für uns in Deutschland? Sind die im Projekt E-ARK entwickelten Informationspakete für archivierende Institutionen in Deutschland nutzbar – auch im Hinblick auf eIDAS und ETSI?“ ein. In seiner Begrüßung erläutert Sebastian Gleixner (Bundesarchiv), dass sich das Bundesarchiv schon seit längerem mit den Inhalten und Zielen des E-ARK-Projektes befasst (zeitweise im „E-ARK-Advisory Board“). Inzwischen liegen nun erste Ergebnisse vor, zum Beispiel frei nutzbare OAIS-konforme Informationspaket-Spezifikationen.

*Abbildung 1: Dr. Sebastian Gleixner eröffnet den Workshop*

Ziel des Workshops ist es, diese der deutschen Archiv- und Bibliothekslandschaft zu präsentieren und zeitgleich die Bekanntheit von E-ARK im deutschsprachigen Raum zu steigern. Der Workshop soll außerdem dabei helfen, E-ARK mit bereits bestehenden Standards und Modellen (z.B. ETSI) zu vergleichen und in Beziehung zu setzen. Außerdem soll die Frage behandelt werden, ob bzw. wie E-ARK auch bei deutschen Archiven und Bibliotheken einsetzbar ist.

In seinem Vortrag „Masse als Herausforderung – Ein Überblick zur digitalen Archivierung im Bundesarchiv“ gibt Sebastian Gleixner anschließend zunächst einen kurzen Abriss über den Umgang mit digitalen Unterlagen im Bundesarchiv, der verstärkt mit der Wiedervereinigung 1990 begann. Seit dieser Zeit wurden im Zuge der Übernahme von Daten aus ehemaligen DDR-Stellen Erfahrungen mit digitalen Daten und Systemen gesammelt (u.a. Robotron-Systeme). Diese werden im Bundesarchiv gesichert, aufbereitet und beauskunftet. Ein erstes OAIS-konformes digitales Archiv startete 2008. Aufgrund der

verzögerten Einführung von Dokumenten-Management-Systemen (DMS) in der Bundesverwaltung konnte die entwickelte Lösung jedoch nicht vollumfänglich eingesetzt werden. Man beschloss daher 2018, den Dienst nicht mehr zu nutzen, sondern auf das sogenannte Digitale Magazin zu setzen. Dieses war wenige Jahre zuvor eingeführt worden und zunächst lediglich für die Präsentation von Digitalisaten (Akten und Großformate) über „invenio“ [1] gedacht. In den Folgejahren kamen jedoch zusätzliche, wenn auch nicht OAIS-konforme Funktionen wie die Sicherung von Digitalisaten auf den Bandspeichersystemen des Bundesarchivs hinzu.

2016 wurde schließlich das Projekt „BASYS 3-Magazin“ zur Etablierung eines umfassenden Systems zur digitalen Archivierung auf Grundlage von OAIS im Bundesarchiv gestartet. Begleitet wird das Projekt von einem stetig ansteigenden Speicherbedarf, der vor allem durch Digitalisierungsvorhaben (z.B. in den Bereichen Film und Akte) bedingt ist.

Anschließend geht Sebastian Gleixner auf das bei der Bundesagentur für Arbeit gehostete Digitale Zwischenarchiv des Bundes (DZAB) und dessen Funktionen ein. Hierbei handelt es sich um eine „digitale Übersetzung“ des im Bundesarchiv und der Bundesverwaltung schon seit Jahrzehnten etablierten „analogen“ Zwischenarchivs, das die Aktenproduktion der Bundesbehörden während der Aufbewahrungsfrist verwahrt. Die Bundesbehörden profitieren durch diesen Service des Bundesarchivs etwa durch die Reduktion von Speicherkosten und eine richtlinienkonforme Beweiswerterhaltung.

Im Anschluss kommt Sebastian Gleixner auf die durch die Datenmengen erzeugten Probleme. Die Zuwächse aus der eigentlichen Verwaltung (40 GB pro Jahr und Behörde) fallen dabei weniger ins Gewicht als die prognostizierten 10 PB Zuwachs pro Jahr durch die interne und externe Digitalisierung (v.a. durch die Filmdigitalisierung). Die Übernahme nichtstaatlicher Unterlagen mit Bezug zur Bundesrepublik Deutschland (Nachlässe, Verbände) ist darüber hinaus mengentechnisch in der Regel nicht planbar. Die Masse an Daten stellt bei der Konzeptionierung der Speicher und der anzuwendenden Prozesse ein entscheidendes Kriterium dar. Derzeit wird deshalb ein neuer Bandspeicher mit einem erwarteten Aufnahmevolumen von ca. 120 PB (skalierbar auf den zehnfachen Wert (1,2

1 Siehe Internet-Auftritt des Bundesarchivs (https://www.bundesarchiv.de/DE/Navigation/Finden/Recherchesysteme/Invenio/invenio.html, aufgerufen am 22.05.2019)

ExaByte) bei Nutzung von LTO-8-Bändern) aufgebaut. Der Datendurchsatz beträgt im Moment 600 MB/s (bei Tests bis zu 1 GB/s; bei Ausbau auf 640 Bandlaufwerke skalierbar auf 790 TB/h). Für den Ingest-Prozess befindet sich derzeit die Ausschreibung für eine Ingest-Software zur OAIS-konformen Paketbildung (SIP, AIP, DIP) und zur Archivierung aller Archivaliengattungen (digitalisiert/genuin digital) in Vorbereitung.

Kurz geht Sebastian Gleixner auch auf die Lösungen zur Bereitstellung von digitalem Archivgut für die Benutzer ein. So erfolgt die Lagerung der DIPs zurzeit in einem Fedora-Repository, aus dem sie über die Rechercheplattform „invenio“ abrufbar sind. Originaldateien werden entweder über mobile Datenträger oder eine eigene Speicherlösung („BArchBox“) für die Benutzung herausgegeben.

Zusammenfassend schließt Sebastian Gleixner seinen Vortrag mit einem Ausblick auf die kommenden Jahre. In diesen werden die Steigerung der Performance der Speicherlösungen sowie der Aufbau OAIS-konformer Archivierungslösungen im Vordergrund stehen. Gleichzeitig ist eine Beschäftigung mit den aktuellen Möglichkeiten und Methoden der Datenverarbeitung und –verknüpfung (mit Bezügen zu den Themen „Künstliche Intelligenz (KI)“, Big Data, Digital Humanities) notwendig. Er betont, dass Standards die „Leitplanken“ für die digitale Archivierung bilden. Diese müssen jedoch flexibel und allgemeingültig sein. Hieraus leitet er auf den eigentlichen Ausgangspunkt des Workshops zurück: Diskussionen anstoßen, wie Standards sinnvoll eingesetzt werden können.

## „Wir wollen europaweit eine interoperable Schnittstelle haben.“ (Ulrike Korte über den ETSI-Standard)

Ulrike Korte (Bundesamt für Sicherheit in der Informationstechnik – BSI) beginnt ihren Beitrag mit einer Übersicht zum rechtlichen Rahmen für elektronische Signierung in Europa. Nach der Publikation der „eIDAS“-Verordnung[2] im Jahr 2014 ist diese seit September 2016 gültig. Sie soll elektronische Geschäftsprozesse europaweit vereinheitlichen und gilt in allen Mitgliedstaaten der EU unmittelbar. Sie löst die bis dato

2 Siehe „eIDAS-Verordnung über elektronische Identifizierung und Vertrauensdienste“ (https://www.bsi.bund.de/DE/Themen/DigitaleGesellschaft/eIDAS/eIDAS_node.html, aufgerufen am 22.05.2019)

geltenden nationalen Signaturgesetze ab. Die technische Umsetzung der in „eIDAS“ formulierten Grundsätze wird in einheitlichen Standards wie z.B. „ETSI“ festgeschrieben. „eIDAS“ definiert nach Aussage von Ulrike Korte verschiedene Stufen von Signaturen. Die höchste Stufe ist hierbei die qualifizierte elektronische Signatur, die in ihrer Wertigkeit mit einer klassischen, manuellen Unterschrift auf einem Dokument gleichzusetzen ist. Neu ist die Möglichkeit mit „eIDAS“ einen gesamteuropäischen Überblick über alle Anbieter von qualifizierten Signaturen zu erhalten. Des Weiteren betont Ulrike Korte, dass über „eIDAS“ und den „ETSI“-Standard die Interoperabilität zwischen den verschiedenen nationalen Anbietern von qualifizierten elektronischen Signaturen ermöglicht wird und kontinuierlich verbessert werden soll. Neu ist dabei auch die Möglichkeit, nicht nur qualifizierte elektronische Signaturen, sondern auch Siegel zu erstellen. Siegel beziehen sich im Unterschied zu Signaturen auf Organisationen und nicht einzelne Personen.

Der „ETSI“-Standard wiederum legt erstmals europaweit technische Standards fest, die es Softwareherstelllern und Anbietern von Vertrauensdiensten ermöglichen, ihre Produkte so zu gestalten, dass z.B. eine Signatur unabhängig von einem bestimmten Anbieter erstellt und geprüft (validiert) werden kann. Dazu werden in „ETSI“ Vorgaben zu folgenden Punkten gemacht:

- Hardwarekomponenten zur Signaturerstellung
- Prozessbeschreibungen (z.B. wie muss eine Signatur erzeugt werden)
- Prüfberichten
- Dateiformaten

Ferner wird über „ETSI“ die Liste aller zertifizierten Anbieter von Vertrauensdiensten (QTSP – Qualified Trust Service Provider) in Europa publiziert.

Ulrike Korte führt im Folgenden aus, wie die deutsche TR-ESOR[3] durch das BSI auf Basis des OAIS-Modells mit seiner Idee des selbsttragenden Archivs konzipiert wurde. Hierbei bildete TR-ESOR eine wesentliche Grundlage für die Erstellung von „ETSI“ und somit auch für die gesamteuropäische Standardisierung. Mit der deutschen TR-ESOR-Richtlinie wird u.a. festgelegt, wie qualifizierte elektronische Signaturen so gespeichert werden können, dass deren Beweiswert auch nach Jahren noch erhalten bleibt. Die grundlegende

3 Siehe auch „BSI TR-03125 Beweiswerterhaltung kryptographisch signierter Dokumente“ (https://www.bsi.bund.de/DE/Publikationen/TechnischeRichtlinien/tr03125/index_htm.html, aufgerufen am 22.05.2019)

Idee hierbei ist, dass jedes Dokument nur einmal mit einem Hashwert versehen wird, um dessen Beweiswert und Integrität sicherzustellen. Dann legt man diese in eine Struktur, die die Zusammengehörigkeit von mehreren Dokumenten (z.B. Vorgang oder Vertrag) dokumentiert und erzeugt über diese übergeordnete Struktur ebenfalls einen Hashwert. Die so entstandenen Baumstrukturen (sog. „Hashbäume“) werden dann zum Beispiel einmal täglich an oberster Stelle mit einem Zeitstempel versehen. Wird dieser Zeitstempel nach Jahren der Langzeitspeicherung schwach, weil zum Beispiel der verwendete Algorithmus veraltet ist, kann über diesen alten Zeitstempel ein neuer aufgebracht werden. Somit bleibt der Beweiswert aller darunter im Baum liegenden Dokumente erhalten, ohne dass diese einzeln erneut mit Hashwerten versehen werden müssen. Diese wirtschaftliche und performante Lösung wurde laut Ulrike Korte zuerst in RFC-Standards umgesetzt und fand über diese Standards Eingang in TR-ESOR.

Softwaretechnisch definiert TR-ESOR eine Middleware (sog. „Preservation Service“) zwischen Client (DMS) und dem Speichersystem mit definierten Schnittstellen (sog. S.4-Schnittstelle) für den über XML realisierten Datenaustausch. Das dazu verwendete Containerformat XAIP basiert ebenfalls auf einem Standard (XFDU nach ISO13527) und kann sowohl als physischer als auch als logischer Container (für große Datenmengen) realisiert werden. Ulrike Korte betont, dass innerhalb des Containers verschiedene Metadatenstandards für die Beschreibung der enthaltenen Daten verwendet werden können, u.a. auch die im E-ARK-Projekt definierten Informationspakete und Content Information Type-Spezifikationen.

Die Middleware kann verschiedene Anbieter von Hashwerten und Zeitstempeln nutzen (siehe Anbieterliste in „ETSI“) oder die entsprechenden Services auch integriert selbst durchführen. Um die Interoperabilität der TR-ESOR zum „ETSI“-Standard sicherzustellen, ist die S4-Schnittstelle so gestaltet, dass die verwendeten Funktionen mit „ETSI“ abgestimmt (gemappt) sind. In Zukunft soll laut Ulrike Korte auch das Wording angeglichen werden. Obwohl TR-ESOR primär für signierte Daten konzipiert wurde, können mit dem Verfahren auch nicht signierte Daten gespeichert werden. TR-ESOR zertifizierte Speicherlösungen erfüllen somit den im E-Government-Gesetz geforderten Anspruch auf Langzeitaufbewahrung nach „Stand der Technik“.

Abschließend erwähnt Ulrike Korte, dass das BSI eigenentwickelte Open Source-Tools zur Prüfung von Softwarelösungen auf Einhaltung von TR-ESOR sowie „ETSI“ bereitstellt.

„Unter Beweiswerterhaltung verstehen wir die Erhaltung der kryptographischen Sicherungsmechanismen.“ (Steffen Schwalm)

Steffen Schwalm (msg systems ag) widmet sich in seinem Beitrag, den er in Vertretung von Sandro Hardy (Bundesarchiv) hält, der Frage, wie AIPs strukturiert sein müssen, um eine langfriste Beweiswerterhaltung sicherstellen zu können.

Hierfür existieren aus seiner Sicht folgende Anforderungen:

- Beliebig viele verschiedene Inhaltsdaten in verschiedenen Versionen (nachträgliches Einfügen) müssen enthalten sein können
- Unterschiedliche Signaturtechniken (auch zukünftige) müssen enthalten sein können
- Signierte und unsignierte Daten müssen zusammen enthalten sein können
- Nachträgliches Einbringen von gesperrtem Material (z.B. Schwärzungen) ohne Zerstörung der existierenden Signatur muss möglich sein
- Das AIP muss definieren können, welche Teile von ihm für den Beweiswerterhalt relevant sind und welche nicht (Referenzen auf entsprechende Daten und Metadaten)
- Die Struktur muss auch bei großen Datenmengen eine hohe Performanz bieten (logisches Containerformat)
- Es müssen Open Source-Software und internationale Standards verwendet werden

*Abbildung 2: Steffen Schwalm*

Diese Anforderungen wurden vom BSI mit der Definition des Containerformats XAIP für die beweiswerterhaltende Langzeitspeicherung von Daten in TR-ESOR umgesetzt. Im Folgenden führt Steffen Schwalm aus, wie man sich in der Entwicklung zunächst auf den nationalen Metadatenstandard XDOMEA konzentrierte. XDOMEA wird in der E-Akte Bund (DMS-System für die Bundesverwaltung) verwendet und im XAIP-Containerformat für das DZAB sowohl in der „MetaDataSection“ oder auch der „dataObjectsSection“ abgelegt.

Dafür wurde die XDOMEA-Struktur auf die entsprechenden Elemente in XAIP gemappt. So bleiben die Daten selbst und auch die Struktur der Daten (Vorgang/Akte) im XAIP Container erhalten. In der regulären XAIP Struktur werden die einzelnen Datenobjekte (Dokumente/Dateien) komplett im Container abgelegt. Beim logischen XAIP (LXAIP)[4] befinden sich die Datenobjekte hingegen außerhalb des Containers und werden nur in der „dataObjectsSection“ des Containers mit „Protected Pointern“ referenziert. Die XOMDEA-Struktur zeigt auch bereits die Informationen an, die für den „evidence record“ im XAIP d.h. für den Beweiswerterhalt der Daten notwendig sind und entsprechend kryptographisch über Hashwerte gesichert werden müssen.

Abschließend schildert Steffen Schwalm, wie es mittels eines Adapters, welcher derzeit in Entwicklung ist, künftig möglich sein soll, auch andere Metadatenstandards als XDOMEA (z.B. solche aus dem E-ARK-Projekt) zu unterstützen. Somit würde z.B. realisiert werden, Daten aus Systemen, die den E-ARK-Standard verwenden, in ein valides XAIP zu wandeln und anschließend im DZAB beweiswerterhaltend zu speichern. Hier würden nach Ansicht von Steffen Schwalm auch Synergien geschaffen, da ein Beweiswerterhalt von qualifizierten elektronischen Signaturen und Siegeln in E-ARK bisher nicht vorgesehen sei.

## „Guter Wille ist nicht genug!” (Sophia Bünemann)

Sophia Bünemann (Europäische Kommission) stellt in ihrem Vortrag die „Connecting Europe Facility“ (CEF) sowie den Building Block „eArchiving“ vor. Der Building Block „eArchiving“ besteht unter anderem aus dem E-ARK Projekt und baut insbesondere auf den E-ARK Spezifikationen auf, die fortgehend überarbeitet und ergänzt werden. Sie beginnt mit einem Überblick über die Hintergrundüberlegungen und -strukturen.

Wesentliches Ziel der EU ist es, auf der Grundlage eines europäischen Binnenmarktes die wirtschaftliche Integration zu fördern. Inzwischen nimmt die Europäische Kommission ausgehend von dieser Überlegung auch den „Digitalen Binnenmarkt“ in den Blick. Dieser soll garantieren, dass Bürger, Firmen und öffentliche Verwaltungen die Vorteile der Digitalisierung effizient nutzen können. Eine enge Zusammenarbeit zwischen den EU-

4 Siehe auch: „Informations- und Beweiswerterhaltung in a Nutshell – Übersicht über Lösungsoptionen für Archivinformationspakete“ (https://www.sg.ch/content/dam/sgch/kultur/staatsarchiv/auds-2018/theorie/04-Korte_Schwalm_Kusber_Beweiserhaltung.pdf, aufgerufen am 22.05.2019)

Staaten ist in diesem Zusammenhang zentral, um dieses Ziel zu erreichen. Dabei verweist Sophia Bünemann auf die „Tallinn-Vision“. Diese sieht vor, dass öffentliche Verwaltungen hochqualitative digitale Dienste für die Bürger sowie möglichst nahtlose und grenzüberschreitende Dienste für Firmen bereitstellen. Sie betont, wie wichtig eine einheitliche Rechtsprechung und einheitliche Standards in allen Mitgliedsstaaten seien, um die gesetzten Ziele zu erreichen.

In Bezug auf „eIDAS“ (vgl. den Beitrag von Frau Korte) stellt sie heraus, wie die Europäische Kommission den öffentlichen Verwaltungen beim Einsatz von „eIDAS“ mitwirken kann. Auftrag der CEF sei es vor allem, Projekte zur Interoperabilität (bisher 120) in den einzelnen Mitgliedsstaaten zu unterstützen. Die CEF stellt zum Beispiel Dienste zur Verfügung, die bei der Einführung von Standards in den nationalen Systemen helfen, und leistet auch finanzielle Zuschüsse. Sophia Bünemann wirbt abschließend darum, dass ihr sowohl „Erfolgsgeschichten“ als auch interessante Ideen und Umsetzungen aus lokalen Archiven mitgeteilt werden, damit diese über die CEF publiziert und beworben werden können.

*Abbildung 3: Janet Anderson*

Janet Anderson (Reichsarchiv Dänemark) stellt in ihrem Vortrag das „E-ARK4ALL“-Projekt und seine organisatorischen Hintergründe, Projektteilnehmer, Ziele und aktuell angewandte Standards vor. Ziel des ursprünglichen Projekts (2014-2017) war es, die Interoperabilität zwischen den abgebenden Stellen und den Archiven bzw. archivierenden Institutionen zu verbessern. Projektergebnisse waren Spezifikationen zu den OAIS-Pakten, Leitlinien und Tools. Bei letzteren handelt es sich um Open-Source-Anwendungen zur digitalen Archivierung. Ferner wurden sieben Pilot-Projekte in sechs Ländern durchgeführt (z.B. die Archivierung relationaler Datenbanken im Dänischen Reichsarchiv). Zur besseren Vermittlung der Ergebnisse wurde Schulungsmaterialien erarbeitet und ein als Informationseinstieg zu E-ARK ein „Knowledge Center“ breitgestellt. Das aktuelle, dem „eArchiving“-Building Block der CEF zugrunde liegende Projekt hat sich v.a. den im o.g. Projekt entstandenen Spezifikationen gewidmet und diese überarbeitet und ergänzt.

Anschließend berichtet Janet Anderson über das „Digital Information LifeCycle Interoperability Standards“-Board (DILCIS)[5], welches die im E-ARK-Projekt erarbeiteten Spezifikationen pflegt. In diesem Zusammenhang betont sie, dass schneller Zugang einer der wichtigsten Faktoren für ein Archiv ist. In Bezug auf die durch E-ARK entwickelte Software hebt sie hervor, dass diese im Kern modular aufgebaut und damit für unterschiedliche Anwendungsbereiche nutzbar ist. E-ARK hat auf diesem Gebiet auch mit kommerziellen Herstellern zusammengearbeitet und stellt die Software als Open Source bereit. In ihrem Ausblick gibt Janet Anderson ihrer Hoffnung Ausdruck, dass der „eArchiving“-Building Block seine Arbeit bis mindestens 2021 fortsetzen kann. In dieser Zeit sollen die Spezifikationen erweitert und gegebenenfalls in Bestandserhaltungssoftware eingebracht werden.

## „Wir pflegen die Daten, aber wir möchten auch, dass sie genutzt werden.“ (Karin Bredenberg)

Karin Bredenberg (Reichsarchiv Schweden) gibt in Ihrem Vortrag einen Überblick über die Spezifikationen im „eArchiving“-Building Block.

Sie beginnt mit der der Darstellung der Zugangswege der Informationspakete in das Archiv / auf den Speicher. Im Speicher werden sie entweder als komplette Pakete gelagert (als Liefer-SIP) oder das SIP wird aufgeteilt und z.B. nach verschiedenen Datentypen im Speicher abgelegt.

Für Archive steht laut Karin Bredenberg die Benutzung im Vordergrund. Aus diesem Grund ist ihr die Betrachtung des DIPs sehr wichtig. Sie stellt zwei grundlegende Möglichkeiten vor, sie zu bilden: Zum einen kann ein DIP nach den konkreten Bedürfnissen und Wünschen des Nutzers zusammengestellt werden (nutzerorientiert). Zum anderen ist es dem Archiv möglich, ein

*Abbildung 4: Karin Bredenberg berichtet über Spezifikationen im „eArchiving“-Building Block.*

5 Siehe auch: Internet-Auftritt des DILCIS-Boards (www.dilcis.eu, aufgerufen am 22.05.2019)

DIP mit vorgefertigter Struktur bereitstellen (archivorientiert).

E-ARK hat für die Spezifikation seiner Informationspakete (SIP, AIP, DIP) auf bestehende Standards zurückgegriffen, um nicht noch ein „(europäisches) Rad“ neu zu erfinden. Als Metadatenschema wurde dabei METS gewählt. Die „Common Specification for Information Packages“ (CS IP) bildet die übergreifende Spezifikation für die Informationspakete SIP, AIP und DIP. Diesen sind jeweils optionale Inhaltsspezifikationen (Content Information Type Specification) für bestimmte Datentypen (Datenbanken, Geodaten, E-Akten, aber auch Finanzdaten, Gesundheitsdaten etc.) zugeordnet. Die unterschiedlichen Datentypen sind mit unterschiedlichen Nutzungsszenarien verknüpft, teilweise weichen auch die gesetzlichen Vorgaben ab. Wesentliche Elemente der Spezifikation sind:

- Dateiformat
- Datenstruktur innerhalb des Informationspakets
- PREMIS Pflichtmetadaten für das „Content Data Object“
- Liste notwendiger Dokumentation und Komponenten (Hard/Software) zur Darstellung und korrekten Interpretation der Daten

Um nationalen und regionalen Besonderheiten Rechnung zu tragen, wurden sogenannte „Userprofiles“ entworfen. Diese ermöglichen es im Rahmen einer vierstufigen Hierarchie ausgehend vom Internationalen Standard (z.B. „Metadata Encoding & Transmission Standard“ (METS))[6] über die Adaption eines EU-Standards (E-ARK CS IP) und die nationale Ausprägung des EU-Standards bis hinunter zur regionalen/lokalen Ebene, Anpassungen in den Spezifikationen vorzunehmen. Dies kann z.B. auf lokaler Ebene die Umbenennung ganzer Informationspakete beinhalten (d.h. SIP umbenennen in Abgabepaket o.ä.). Da die Datenpakt-Spezifikationen auf einer gemeinsamen Grundlage basieren, wird deren Interoperabilität zwischen Datenproduzenten und archivierenden Institutionen gewährleistet. So könnten z.B. ein Archiv und eine Bibliothek nach Aussage von Karin Bredenberg auf nationaler Ebene problemlos ihre Datenpakete komplett austauschen und verarbeiten, wenn beide auf den (nationalen) E-ARK-Standard setzen; international wäre es aber möglich, den Austausch mindestens über die verpflichtenden Teile der Spezifikation regeln.

6 Siehe Übersicht über den „Metadata Encoding & Transmission Standard“ (http://www.loc.gov/standards/mets/, aufgerufen am 22.05.2019)

Alle Spezifikationen werden im DILCIS-Board verwaltet und aktualisiert. Der Prozess wird, u.a. über die Verwendung von „Github“ für die Erstellung der Spezifikationen, so transparent wie möglich gestaltet. Hierbei legt E-ARK Wert auf die Mitarbeit von externen Experten. In Zukunft wird E-ARK zudem eine „endorsment procedure“ etablieren, über die bereits vorhandene Standards offiziell von E-ARK unterstützt und deren Nutzung empfohlen und beworben werden soll. Zum Abschluss betont Karin Bredenberg die Notwendigkeit vorhandene Standards wo immer möglich zu nutzen und ihre Verbreitung zu fördern, um gemeinsam Lösungen bei der digitalen Archivierung zu finden.

„Wir haben schon sehr viele Standards gesehen, aber wir haben nicht viele Erfahrungen in digitaler Archivierung. Ich bin mir nicht sicher ob wir digitale Archivierung wirklich umfassend verstehen.“ (Christian Keitel)

Nach der Kaffeepause folgt eine Podiumsdiskussion zwischen den Vortragenden Karin Bredenberg, Ulrike Korte, Steffen Schwalm sowie Christian Keitel (Landesarchiv Baden-Württemberg) und Rainer Jacobs (Bundesarchiv), die von Sina Westphal (Bundesarchiv) und Sebastian Gleixner moderiert wird. Die erste Frage thematisiert die Sinnhaftigkeit und Praktikabilität von Standards im Bereich der digitalen Archivierung allgemein und Informationspaketen im Speziellen. Nach Meinung von Karin Bredenberg werden derzeit viele „de facto“ Standards in der digitalen Archivierung verwendet. Vor diesem Hintergrund ist geplant, das CS IP von E-ARK zu einem einheitlichen Standard für Europa weiterzuentwickeln. Demgegenüber ist Christian Keitel der Ansicht, dass noch nicht genug Erfahrungen in der digitalen Archivierung gesammelt wurden, um normierte Standards in diesem Bereich zu etablieren. Er dankt dem Bundesarchiv deshalb, das Thema erstmals in Deutschland aufgegriffen zu haben und hofft, dass damit die Diskussion bei den archivierenden Institutionen in Gang kommt. Sowohl Steffen Schwalm als auch Ulrike Korte halten eine Standardisierung vor allem im Bereich des „Records Management“ – also dort wo die Unterlagen entstehen – für sinnvoll. Allerdings müssten auch hier vor der Entwicklung z.B. von ISO-Standards zunächst praktische Erfahrung gesammelt werden (z.B. bei E-Akte Bund).

Die Diskussion macht auf noch bestehende Unterschiede in der Definition von Begriffen wie „Digital Preservation“ sowie „Langzeitspeicherung“, die vor allem die Sicherung digitaler Objekte auf Produzentenseite betrifft, versus „digitaler Archivierung /

Langzeitarchivierung“, die anschließend bei den archivierenden Institutionen betrieben wird, aufmerksam. Den Bereich digitale Signaturen und Siegeln ordnet Ulrike Korte dem Beweiswerterhalt zu, der bei der Langzeitspeicherung notwendig ist.

*Abbildung 5: Fr. Dr. Sina Westphal, Hr. Dr. Sebastian Gleixner, Prof. Dr. Christian Keitel, Karin Bredenberg, Ulrike Korte, Steffen Schwalm (v.l.)*

Die Vertreter der archivierenden Institutionen weisen darauf hin, dass diese Maßnahmen im Rahmen der digitalen Archivierung berücksichtigt werden können, aber nicht müssen. Hier verschwimmen jedoch zunehmend die Grenzen, wie auch Christian Keitel feststellt und das Thema Nutzung aufgreift. Nutzer könnten sowohl nach rechtssicheren Informationen fragen, was digitale Signaturen oder zumindest deren Dokumentation erfordert, als auch aus rein geschichtswissenschaftlichem Interesse das Archiv besuchen, hier wären nicht zwingend digitale Signaturen erforderlich. Rainer Jacobs ergänzt, dass Benutzerinnen und Benutzer authentische Quellen einfordern und somit in jedem Fall Authentizität und Integrität des Archivguts gewahrt werden muss. Auch Karin Bredenberg ist der Meinung, dass digitale Langzeitspeicherung mit Beweiswerterhalt einerseits und digitale Archivierung andererseits zunehmend miteinander verschmelzen. Umso wichtiger ist für Christian Keitel, dass es statt normierter Standards, die irgendwann veralten, einheitliche Schnittstellen gibt, die von möglichst vielen Akteuren genutzt werden.

Bei der Frage, ob europäische E-ARK-Standards in Deutschland genutzt werden können, äußern sich die Teilnehmer eher zurückhaltend. Ulrike Korte und Steffen Schwalm

betonen zwar, dass E-ARK für den nationalen TR-ESOR-Standard genutzt werden kann und zu diesem kompatibel sein wird. Christian Keitel sieht jedoch die Gefahr, dass ein neuer Standard von der EU vorgegeben werden könne und somit die gerade erst beginnende Diskussion über die Ausgestaltung von Informationspaketen vorzeitig und nicht zum Besten aller Beteiligten beendet. Dem entgegnet Karin Bredenberg, dass sich viele Akteure in Europa gerade eine Standardisierung wünschen, um der digitalen Langzeitarchivierung mehr Gewicht zu geben. Allerdings sei es aufgrund der vielen verwendeten Systeme wahrscheinlich sinnvoll, zunächst auf standardisierte Schnittstellen zu setzen.

Einig sind sich die Diskussionsteilnehmer darüber, dass eine Interoperabilität zwischen E-ARK und in Deutschland verwendeten Standards sinnvoll ist und wo immer möglich ausgebaut werden sollte. Eine hohe Verbreitung z.B. des E-ARK-SIP würde nach Auffassung von Christian Keitel die Komplexität von möglichen Eingangswegen nach DIMAG (=Digitales Magazin, die beim Landesarchiv Baden-Württemberg entwickelte Lösung zur digitalen Archivierung, der sich eine Vielzahl von Bundesländern und Kommunen inzwischen angeschlossen haben) deutlich reduzieren. Eine Anpassung von AIPs in DIMAG an einen europäischen E-ARK-Standard hält er jedoch für unwahrscheinlich. Auch Sebastian Gleixner weist aus den Erfahrungen bei der digitalen Archivierung im Bundesarchiv darauf hin, dass eine Standardisierung der SIPs den Ingest-Prozess deutlich vereinfachen würde. Beide sind sich dabei einig, dass dies auch Auswirkungen auf die Struktur von AIPs und besonders auf DIPs haben wird, da v.a. DIPs oftmals ein Spiegelbild von SIPs darstellen.

Auf die Frage, ob es einen Bedarf für einen Beweiswerterhaltung digitaler Signaturen und Siegel gibt, obwohl dies in E-ARK derzeit nicht berücksichtigt wird, antwortet Karen Bredenberg auf der Grundlage ihrer Erfahrungen im schwedischen Reichsarchiv. Dort wird das Thema Signaturerhalt von der digitalen Archivierung getrennt behandelt. Zentrales Element seien dabei die Checksummen (Hashwerte), welche von einigen Institutionen sehr häufig (monatlich) überprüft werden. E-ARK sei offen für Erweiterungen in den Spezifikationen, die dem Beweiswerterhalt von Signaturen dienen. Wichtig dabei ist, dass man hierbei auf einen gemeinsamen „Nenner“ für ganz Europa komme und keine unnötigen verpflichtenden Erweiterungen zu den Spezifikationen hinzufüge, welche deren Flexibilität negativ beeinflussen. Steffen Schwalm unterstreicht diese Aussage. E-ARK und „ETSI“ seien zwei verschiedene Standards, die sich aber gut ergänzen. Während E-ARK sich mit dem Erhalt der Inhaltsinformationen befasst, kann „ETSI“ dazu dienen den

Beweiswerterhalt sicherzustellen, hier würden Checksummen allein nicht ausreichen, da sie keinen „proof of existence“ (d.h. Nachweis, dass ein Dokument so zu einem bestimmen Zeitpunkt erstellt wurde). ermöglichen. Deswegen müssen beide Standards nicht noch zusätzlich Elemente des jeweils anderen übernehmen (siehe dazu auch den Vortrag von Herrn Schwalm).

In der Schlussdiskussion mit dem Publikum wird Sebastian Gleixner gefragt, welchen Standard das Bundesarchiv bei der digitalen Archivierung bevorzugen würde. Er antwortet, dass es momentan noch keine konkrete Festlegung gebe und alle Optionen untersucht würden, darunter auch die Ergebnisse des E-ARK-Projektes. Das Bundesarchiv baue gerade sein Digitales Magazin komplett neu auf und begreife dies als Chance, auch neue Wege zu gehen. Dabei betont er, dass der kontinuierliche technische Wandel die Lösungsfindung besonders herausfordernd macht. Rainer Jacobs ergänzt, dass für das Digitale Zwischenarchiv die Standards XDOMEA und TR-ESOR genutzt werden. Christian Keitel merkt an, dass es durchaus positiv zu werten sei, wenn das Bundesarchiv einen eigenen Weg gehe. Der Einsatz unterschiedlicher Systeme in der deutschen Archivlandschaft könne als Schutz vor einem möglichen Totalverlust der Überlieferung durch Diversifizierung betrachtet werden, außerdem stärke dieses Vorgehen eine offene Diskussion.

Sophia Bünemann unterstreicht, dass in der EU die Mitgliedsstaaten die Initiative zur Standardisierung ergriffen hätten, nicht die Europäische Kommission. Dies zeige den Bedarf nach einer Lösung für die digitale Archivierung. Impulse könnten grundsätzlich von privater wie öffentlicher Seite kommen. Die EU möchte wiederum den Dialog zu diesem Thema fördern, um weitere Fortschritte zu ermöglichen. Janet Anderson ergänzt, dass in der jüngsten Vergangenheit eine zunehmende Flexibilisierung bei der Entwicklung der E-ARK-Standards eingesetzt hat, v.a. auch durch die Arbeit von Karin Bredenberg. E-ARK hat zudem nicht nur Archive als Nutzer im Blick. Sie stellt unter anderem die Nutzung des E-ARK-AIP durch das Amt für Veröffentlichungen der Europäischen Union („European Union Publications Office“) als Beispiel für die Nutzung einer AIP Spezifikation jenseits klassischer Archive vor und betont, dass der wesentliche Punkt für die Entwicklung von Standards ist, dass sie nur auf das absolut notwendige Minimum an Pflichtvorgaben setzen. Je starrer ein Standard ist und je mehr unbedingte Vorgaben er macht, desto weniger wird er in der Praxis genutzt werden können. Als Beispiel führt sie den

„MoReq2010“-Standard[7] an, der zwar sehr detailliert sei, durch sein starres Gerüst aber „perfekt unbrauchbar“ ist. Janet Anderson weist darauf hin, dass Pilotprojekte gezeigt haben, dass Entwicklungen von Standards Lösungen hervorbringen können, von denen die Allgemeinheit profitieren kann. Als Beispiel nannte sie die Entwicklung von Spezifikationen für das E-ARK-SIP, die in Österreich erstellt und anschließend in Ungarn erfolgreich verwendet wurden.

Sina Westphal unterstreicht im Schlusswort den positiven Gedanken hinter der Idee des E-ARK-Projekts, an dem viele europäische Archive und Institutionen mitgewirkt haben. Sie weist auch auf das – insbesondere für kleine Archive - außerordentlich attraktive Angebot des E-ARK-Projektes durch die Bereitstellung von Open Source-Tools hin. Mit diesen Tools wird nicht nur praktische Hilfe bei der Bewältigung digitaler Archivierung geboten, sondern sie helfen aus Sicht von Sina Westphal auch dabei, sich intellektuell mit digitaler Archivierung auseinanderzusetzen und die Expertise in der deutschen Archivlandschaft in diesem Bereich zu vertiefen.

Der Workshop hat deutlich gemacht, dass die Diskussion rund um (auch vergleichende) Ansätze zur Standardisierung im Rahmen der digitalen Archivierung längst noch nicht beendet ist. Zum einen sind aufgrund der bislang kurzen Phase, in der sich archivierende Institutionen mit diesem Thema beschäftigen, valide Erfahrungswerte, auf denen aufgebaut werden kann, gerade erst im Entstehen. Zum anderen müssen die gefundenen Antworten stetig entlang der technischen Weiterentwicklungen überprüft und gegebenenfalls angepasst werden. Von mehreren Seiten kam der Anstoß, ausgehend von diesem Workshop den gemeinsamen Austausch weiterzuführen.

*Tim Dahlmanns, Sebastian Gleixner, Sina Westphal (Koblenz), Maximilian Maede (Freiburg)*

7 Siehe MoReq-Internetauftritt (https://www.moreq.info/, aufgerufen am 22.05.2019)