

H. Neuroth, A. Oßwald, R. Scheffel, S. Strathmann, K. Huth (Hrsg.)

nestor Handbuch

Eine kleine Enzyklopädie
der digitalen Langzeitarchivierung

Version 2.3

Kapitel 9.2

Workflows für den Objektzugriff

nestor Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung
hg. v. H. Neuroth, A. Oßwald, R. Scheffel, S. Strathmann, K. Huth
im Rahmen des Projektes: nestor – Kompetenznetzwerk Langzeitarchivierung und
Langzeitverfügbarkeit digitaler Ressourcen für Deutschland
nestor – Network of Expertise in Long-Term Storage of Digital Resources
<http://www.langzeitarchivierung.de/>

Kontakt: editors@langzeitarchivierung.de
c/o Niedersächsische Staats- und Universitätsbibliothek Göttingen,
Dr. Heike Neuroth, Forschung und Entwicklung, Papendiek 14, 37073 Göttingen

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen
Nationalbibliografie; detaillierte bibliografische Daten sind im Internet unter
<http://www.d-nb.de/> abrufbar.

Neben der Online Version 2.3 ist eine Printversion 2.0 beim Verlag Werner Hülsbusch,
Boizenburg erschienen.

Die digitale Version 2.3 steht unter folgender Creative-Commons-Lizenz:
„Namensnennung-Keine kommerzielle Nutzung-Weitergabe unter gleichen Bedingungen 3.0
Deutschland“
<http://creativecommons.org/licenses/by-nc-sa/3.0/de/>



Markenerklärung: Die in diesem Werk wiedergegebenen Gebrauchsnamen, Handelsnamen,
Warenzeichen usw. können auch ohne besondere Kennzeichnung geschützte Marken sein und
als solche den gesetzlichen Bestimmungen unterliegen.

URL für Kapitel 9.2 „Workflows für den Objektzugriff“ (Version 2.3):
[urn:nbn:de:0008-20100305155](http://nbn-resolving.org/urn/resolver.pl?urn=urn:nbn:de:0008-20100305155)
<http://nbn-resolving.de/urn/resolver.pl?urn=urn:nbn:de:0008-20100305155>



Gewidmet der Erinnerung an Hans Liegmann (†), der als Mitinitiator und früherer Herausgeber des Handbuchs ganz wesentlich an dessen Entstehung beteiligt war.

9.2 Workflows für den Objektzugriff

Dirk von Suchodoletz

Es genügt nicht, lediglich ein digitales Objekt bit-getreu zu bewahren, sondern es sind Vorkehrungen zu treffen, um dieses Objekt zu einem späteren Zeitpunkt wieder darstellen zu können. Hierzu dienen bestimmte Workflows, die ein Langzeitarchiv implementieren sollte. Deshalb beschäftigt sich dieser Abschnitt mit der theoretischen Fundierung und Formalisierung technischer Abläufe, wie sie beispielsweise mit dem DIAS-Projekt zu Beginn des Jahrtausends an der Königlichen Bibliothek der Niederlande eingeführt wurden⁴.

Der zentrale Ausgangspunkt der Überlegungen liegt darin begründet, dass digitale Objekte nicht allein aus sich heraus genutzt oder betrachtet werden können. Stattdessen benötigen sie einen geeigneten Kontext, damit auf sie zugegriffen werden kann. Dieser Kontext, im Folgenden Erstellungs- oder Nutzungsumgebung genannt, muss geeignete Hardware- und Softwarekomponenten so zusammenfügen, dass je nach Typ des digitalen Objekts dessen Erstellungsumgebung oder ein geeignetes Äquivalent erzeugt wird. Für diese Schritte der Wiederherstellung sind seitens des Archivbetreibers geeignete Workflows (Arbeitsabläufe) vorzusehen. Um diese beschreiben zu können, sind sogenannte „View-Paths“ ein zentrales Konzept. Diese Darstellungspfade liefern die grundlegenden Anweisungen zur Konstruktion geeigneter technischer Workflows für das Sichtbarmachen oder Ablaufenlassen verschiedener digitaler Objekte.

Den Betreibern eines digitalen Langzeitarchivs wachsen aus diesen Überlegungen verschiedene Aufgaben zu. Hierzu zählen die Bestimmung des Typs eines Objekts bei der Einstellung ins Archiv (dem sog. Ingest in der OAIS-Terminologie) und die Beschaffung und Ablage der notwendigen Metadaten, auf die an anderer Stelle in diesem Handbuch ausführlich eingegangen wird.

Für den späteren Objektzugriff spielt die Überprüfung, inwieweit im Langzeitarchivierungssystem eine für diesen Objekttyp passende Nutzungsumgebung vorhanden ist, eine besondere Rolle. Deren technischer Workflow wird nachfolgend näher ausgeführt. Dabei können View-Path und Nutzungsumgebung je nach Art der betreibenden Organisation und ihrer spezifischen Anforderungen, die typischerweise durch „Significant Properties“⁵ beschrieben

4 Vgl. van Diessen; Steenbakkers 2002 und van Diessen 2002, S. 16f

5 Vgl. <http://www.significantproperties.org.uk/index.html> sowie <http://www.jisc.ac.uk/whatwedo/programmes/preservation/2008sigprops.aspx>

werden, unterschiedlich aussehen. Dies resultiert je nach Benutzergruppe oder Einrichtung in verschiedene Kriterien, nach denen Darstellungspfade bestimmt werden. Es lassen sich drei wesentliche Phasen voneinander unterscheiden (Abbildung 1):

- Notwendige Arbeitsschritte und Handlungen bei der Objektaufnahme in das OAIS-konforme Archiv.
- Workflows, die im Laufe des Archivbetriebs umzusetzen sind.
- Abläufe für den Objektzugriff nach der Objektausgabe an den Endbenutzer.

Der View-Path zum Zugriff auf die unterschiedlichen, im Langzeitarchiv abgelegten Objekttypen ist im Moment der Archivausgabe festzulegen. An dieser Stelle müssen diverse Workflows implementiert sein, die es erst erlauben, dass ein späterer Archivnutzer tatsächlich auf das gewünschte Objekt zugreifen kann. Hierbei spielt die auf das Objekt angewendete Langzeitstrategie, ob Migration oder Emulation, keine Rolle (Abbildung 2). In jedem Fall muss das Archivmanagement dafür sorgen, dass eine passende Nutzungsumgebung bereitgestellt wird.

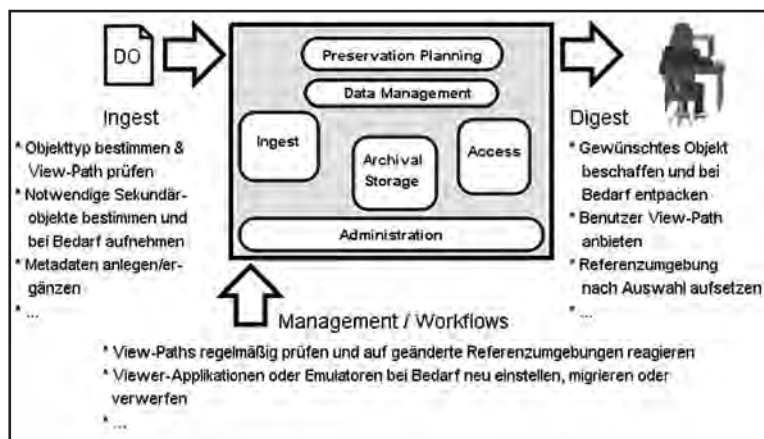


Abbildung 1: Bei der Unterhaltung eines digitalen Langzeitarchivs sind eine Reihe verschiedener technischer Workflows festzulegen und umzusetzen.

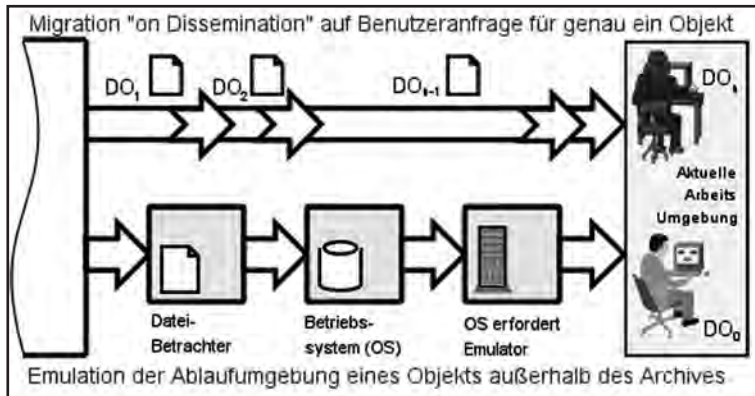


Abbildung 2: Unabhängig von der eingesetzten Archivierungsstrategie eines Objekts muss für einen geeigneten Zugriff gesorgt werden.

Formalisierter Zugriff auf archivierte Objekte

Die Wiederherstellung von Nutzungsumgebungen oder entsprechender Äquivalente lässt sich durch View-Paths formalisieren. Diese Wege starten vom darzustellenden oder auszuführenden digitalen Objekt. Sie reichen je nach angewandeter Archivierungsstrategie über verschiedene Zwischenschritte bis in die tatsächliche Arbeitsumgebung des Archivnutzers. Da digitale Archivalien, im Zusammenhang mit View-Path auch Primärobjekte genannt, nicht aus sich allein heraus genutzt werden können, werden weitere digitale Objekte benötigt. Hierzu zählen Viewer, Hilfsapplikationen, Betriebssysteme oder Emulatoren. Sie sind als Hilfsmittel, im Folgenden als Sekundärobjekte bezeichnet, nicht von primärem Archivierungsinteresse, jedoch zwingend ebenfalls zu berücksichtigen.

Das Konzept der Darstellungspfade wurde ursprünglich im Zuge des eingangs genannten DIAS-Projektes an der Königlichen Bibliothek der Niederlande entwickelt⁶. Die Abbildung 3 zeigt einen typischen View-Path-Verlauf ausgehend vom digitalen Objekt. Im Beispiel wurde es mittels einer bestimmten Software erzeugt, die ihrerseits wiederum auf einem Betriebssystem ausgeführt werden kann, das seinerseits wegen des Nicht-Vorhandenseins des originalen Rech-

6 Vgl. http://www.kb.nl/hrd/dd/dd_onderzoek/preservation_subsystem-en.html sowie van Diessen; Steenbakkens (2002).

ners einen Hardwareemulator erfordert. Dieser läuft als Applikation in der Arbeitsumgebung des Archivnutzers.



Abbildung 3: Das Beispiel zeigt einen generellen View-Path zum Rendering/ Ablaufenlassen eines digitalen Objekts eines bestimmten Typs unter Einsatz der Emulationsstrategie.

Weitere View-Path-Beispiele

Ein View-Path eines migrierten Objekts wäre entsprechend kürzer: Da eine in der aktuellen Arbeitsumgebung ablaufende Applikation genutzt werden kann, muss lediglich diese passend zum jeweiligen Objekttyp bereitgestellt werden. Umgekehrt kann sich ein View-Path bei geschachtelter Emulation verlängern: Steht der im obigen Beispiel genannte Hardwareemulator nur für ein älteres Betriebssystem statt der aktuellen Umgebung zur Verfügung, würde ein Zwischenschritt aus diesem obsoleten Betriebssystem mit passendem Emulator für die aktuelle Umgebung angehängt werden (Abbildung 4 unten).

Während der Ausgangspunkt des View-Paths durch das Primärobjekt fixiert ist, wird sich, erzwungen durch den technologischen Fortschritt und die sukzessive Obsoleszenz vorhandener Rechnerplattformen, der Endpunkt des View-Path im Zeitablauf verschieben. Weiterhin hängt die Länge eines Darstellungspfades vom Typ des Objekts ab: Ist eine Applikation wie z.B. eine alte Datenbank von primärem Interesse, so entfällt beispielsweise der Zwischenschritt der Erstellungs- oder Anzeigeapplikation, da sie direkt auf einem Betriebssystem ausgeführt werden kann.

View-Paths müssen nicht automatisch eindeutig bestimmt sein. Aus Sicht des Archivmanagements bieten sich generell folgende Szenarien für Darstellungspfade an:

- Es existiert zum gegebenen Zeitpunkt ein Weg vom Primärobjekt zu seiner Darstellung oder Ausführung.

- Es existieren mehrere verschiedene View-Paths. Diese sind mit geeigneten Metriken zu versehen. Diese erlauben die Bewertung zur Verfügung stehender Alternativen und werden weiter hinten besprochen.
- Es kann Archivobjekte geben, zu denen zu bestimmten Zeitpunkten kei-

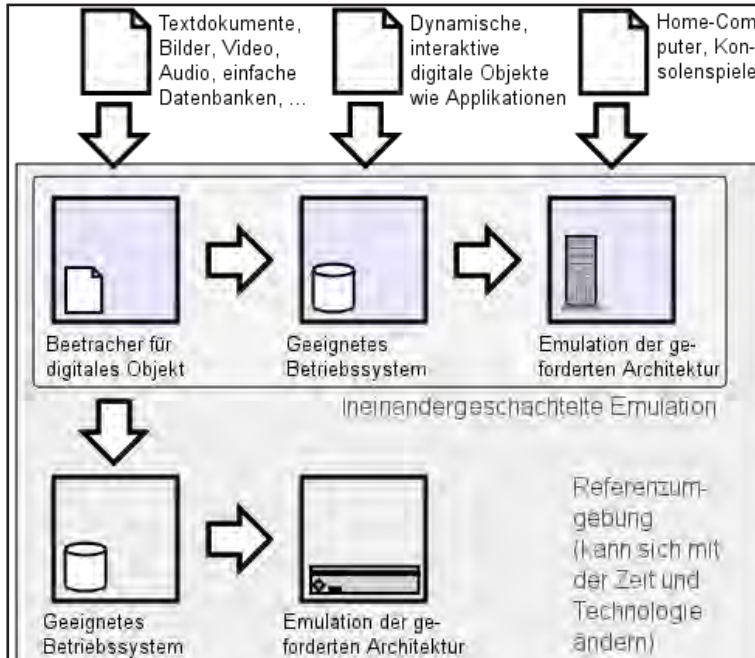


Abbildung 4: Die Länge eines View-Path hängt sowohl vom Typ des Objekts als auch von der eingesetzten Archivierungsstrategie ab.

ne View-Paths konstruierbar sind.

Referenzumgebung - Endpunkt von View-Paths

Referenzumgebungen haben jeweils festgelegte Eigenschaften, die sich aus der Definition ihrer Hard- und Softwareumgebung bestimmen. Dieses sind die aktuellen Arbeitsumgebungen, in denen sich Archivbenutzer bewegen. Sie ändern sich zwangsläufig im Laufe der Zeit und sind unter anderem durch die Beschaffungspolitik der jeweiligen Gedächtnisorganisation determiniert. View

Paths sollten sich auf wohldefinierte Referenzumgebungen⁷ beziehen.⁸ Geeignete Referenzumgebungen versuchen deshalb, in möglichst kompakter und gut bedienbarer Form ein ganzes Spektrum von Nutzungsumgebungen zur Verfügung zu stellen. Dabei sollte die Basisplattform möglichst der jeweils aktuellen Hardware mit jeweils üblichen Betriebssystemen entsprechen. Das verhindert einerseits das Entstehen eines Hardwaremuseums mit hohen Betriebskosten und andererseits findet sich der Benutzer zumindest für das Basissystem in gewohnter Umgebung wieder.

Eine Referenzumgebung sollte in der Lage sein, neben der jeweiligen Nutzungsumgebung zusätzlich die notwendigen Hinweise zum Einrichten und zur Bedienung bereitzustellen. Dies schließt den geeigneten Zugriff auf die Objektmetadaten mit ein. Weitere Kriterien liegen in der Güte der Darstellung der Nutzungsumgebung, was durch den jeweils eingesetzten Emulator und seine Benutzer-Interfaces in der jeweiligen Referenzumgebung mitbestimmt wird.

Andererseits können Referenzumgebungen als Endpunkte eines View-Paths

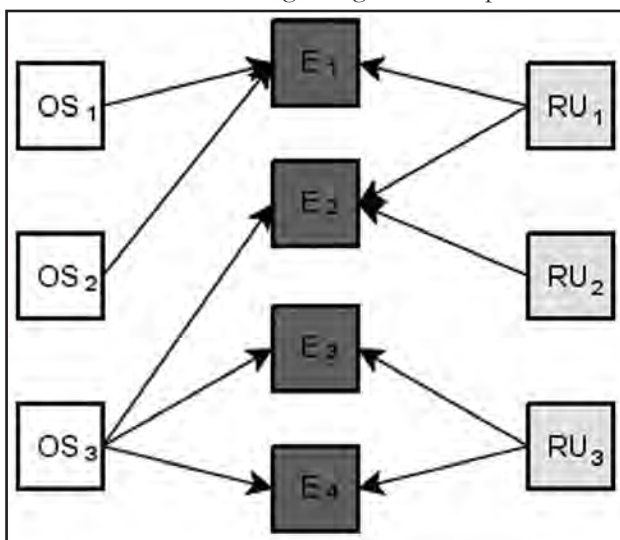


Abbildung 5: Die Auswahl der Emulatoren (E) wird in vielen Fällen nicht nur durch das Betriebssystem (OS) im View-Path, sondern auch von der anderen Seite durch die verfügbaren Referenzumgebungen (RU) beeinflusst.

7 Gedächtnisorganisationen haben typischerweise bestimmte Arbeitsumgebungen an ihren Arbeitsplätzen für ihre eigenen Archivare oder externe Nutzer (Abbildung 1). Diese Umgebungen können sie in einem gewissen Umfang selbst mitbestimmen. Vgl. a. van Diessen 2002.

8 van Diessen (2002b)

diesen umgekehrt mitbestimmen (Abbildung 5). Wenn technische, finanzielle oder personelle Restriktionen bestimmte Referenzumgebungen, die potenziell nachgefragt werden, nicht erlauben, kann dies die Wahl eines Emulators wesentlich beeinflussen. Die Rekursion kann sich unter Umständen noch weiter auf frühere Entscheidungsknoten auswirken.

Wegen ihrer spezielleren Anforderungen, die durch die eingesetzten Emulatoren und Viewer beeinflusst werden, ist es für die Betreiber von Langzeitarchiven vielfach sinnvoll, eine für die Objekte ihres Archivs relevante Referenzplattform selbst zu definieren und bereitzustellen. Diese unterscheidet sich je nach Anwendung und Gedächtnisorganisation: Bibliotheken und Archive brauchen in erster Linie Viewer für ihre migrierten statischen Objekte, die sie ihren Nutzern innerhalb ihrer Recherchesysteme anbieten wollen. Darüber hinaus können Emulatoren für nicht-migrierbare Archivinhalte und als Kontrollwerkzeug für mehrfach migrierte Objekte benötigt werden.⁹

Technische Museen oder Ausstellungen leben eher von interaktiven Objekten. Die Referenz-Workstation ist entsprechend den zu zeigenden Exponaten zu bestücken. Ähnliches gilt für multimediale Kunstobjekte. Hier könnten die Significant Properties jedoch sehr spezielle Anforderungen an eine möglichst authentische Präsentation stellen.

Für Firmen oder Institutionen werden in den meisten Fällen lediglich View-Paths erforderlich sein, die sich bereits mittels X86-Virtualisierern¹⁰ komplett erstellen lassen. Die erwarteten Objekte sind eher statischer Natur und wurden typischerweise auf PCs verschiedener Generationen erstellt. Generell muss es sich bei den eingesetzten Referenz-Workstations nicht um die jeweils allerneueste Hardwaregeneration handeln. Stattdessen sollte jene Technologie angestrebt werden, die einen optimalen Austausch erlaubt und den Anforderungen der jeweiligen Nutzer gerecht wird.

Je nach Art des Archivs kann ein Datenaustausch mit der Außenwelt erforderlich werden: Nicht alle Objekte müssen bereits von Anfang an im Archiv abgelegt sein. Es können später immer wieder Objekte, beispielsweise aus Nachlässen von Wissenschaftlern, Künstlern oder Politikern auftauchen. In solchen

9 Die Darstellung eines Objekts via Emulation wird typischerweise deutlich aufwändiger sein. Gerade für häufig nachgefragte, statische Objekte bietet sich deshalb die Migration an. Bei Zweifeln an der Authentizität kann mittels Emulation das Ergebnis des n-ten Migrationsschritts mit dem unveränderten Originalobjekt verglichen werden.

10 Softwareprodukte wie VMware oder VirtualBox erlauben das Nachbilden eines X86er PCs auf einer X86er 32 oder 64-bit Maschine. Solange die Treiberunterstützung besteht, können ältere Betriebssysteme, wie Windows95 oder 2000 bequem innerhalb einer Applikation (dem Virtualisierer) auf dem Standard-Desktop des Benutzers ausgeführt werden.

Fällen wird es zu Zwecken der Datenarchäologie¹¹ von Interesse sein, externe Objekte in bestehende Workflows einspeisen zu können. Umgekehrt sollen vielleicht Objekte für externe Nutzung speicher- oder ausdrückbar sein.

Für alle Schritte muss ein ausreichendes Bedienungswissen vorhanden sein. Hierzu werden sich neue Berufsfelder, wie das eines digitalen Archivars herausbilden müssen, um ausgebildetes Fachpersonal auch für recht alte Nutzungs-

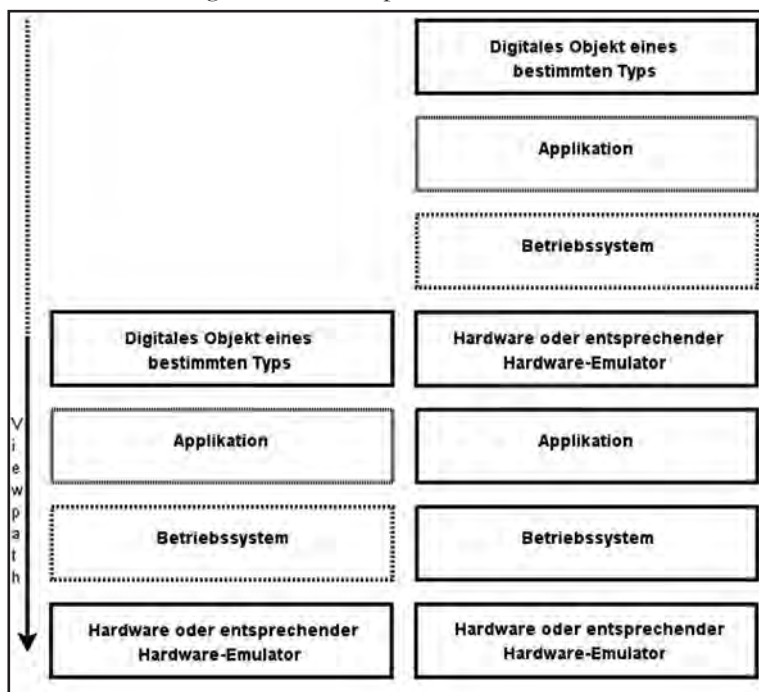


Abbildung 6: Ein View-Path in einer Darstellung als Schichtenmodell. Je nach Typ des Objekts muss nicht jede Schicht durch eine Softwarekomponente repräsentiert sein.

umgebungen vorhalten zu können. Selbst wenn diese nicht mehr direkt auf der Originalhardware ablaufen, so müssen sie innerhalb der Emulatoren bedient werden können.

View-Path-Varianten und -Auswahl

Einen View-Path kann man als Entscheidungsbaum interpretieren, an dessen Wurzel das interessierende Primärobjekt steht. Ein Blatt ohne weitere Verzwei-

¹¹ Im Sinne des Umgangs (Bestimmung) mit sehr alten, lange schon nicht mehr zeit-typischen digitalen Objekten.

gungen stellt das Ende des Pfades in Form einer gültigen Referenzumgebung dar. Zur Veranschaulichung des Aufbaus der geforderten Nutzungsumgebung sollte man sich am besten ein Schichtenmodell vorstellen, wie es in Abbildung 6 präsentiert wird.

Eine ganze Reihe digitaler Objekte können durch mehr als eine Applikation dargestellt oder ablauffähig gemacht werden. Dabei können die Ergebnisse in Authentizität, Komplexität oder Qualität differieren. Aus diesen Überlegungen folgen Pfadverzweigungen und auf der Schicht der Applikation eine Auswahlmöglichkeit. Ähnliches wiederholt sich für die Anforderung der Applikation nach einem Betriebssystem. Auf dieser Ebene kann erneut eine Verzweigung auftreten. Die Rekursion setzt sich mit dem Betriebssystem und einer Menge an geeigneten Hardwareemulatoren fort.

Da der technologische Fortschritt wesentlichen Einfluss auf die Referenzumgebung hat und nur bedingt durch den Archivbetreiber beeinflusst werden kann, bestimmen sich View-Path und Referenzumgebung gegenseitig. Auf den Zwischenschichten stehen Betriebssysteme und Hardwareemulatoren über Gerätetreiber in Abhängigkeit zueinander.

Die Modellierung des View-Paths in Schichten erfolgt nicht eng fixiert: So reduziert sich beispielsweise bei einem digitalen Primärobjekt in Form eines Programms die Zahl der Schichten. Ähnliches gilt für einfache Plattformen wie bei Home-Computern, wo keine Trennung zwischen Betriebssystem und Applikation vorliegt (Abbildung 4 Mitte). Darüber hinaus können Schichten wiederum gestapelt sein, wenn es für einen bestimmten Emulator erforderlich wird, seinerseits eine geeignete Nutzungsumgebung herzustellen, was im rechten Teil von Abbildung 6 gezeigt wird.

Metriken als Entscheidungskriterien

Eine sinnvolle Erweiterung des etwas starren Ansatzes im ursprünglichen DIAS-Preservation-Modell (van Diessen; Steenbakkers 2002) könnte in der Gewichtung der einzelnen View-Path-Varianten liegen. Dies ließe sich durch eine beschreibende Metrik formalisieren. Gerade wenn an einem Knoten mehr als eine Option zur Auswahl steht (Abbildung 7), erscheint es sinnvoll:

- Präferenzen der Archivnutzer beispielsweise in Form der Auswahl der Applikation, des Betriebssystems oder der Referenzplattform zuzulassen.
- Gewichtungen vorzunehmen, ob beispielsweise besonderer Wert auf die Authentizität der Darstellung (van Diessen; van der Werf-Davelaar 2002) oder eine besonders einfache Nutzung gelegt wird.
- Vergleiche zwischen verschiedenen Wegen zuzulassen, um die Sicherheit

und Qualität der Darstellung der Primärobjekte besser abzusichern.

- Den Aufwand abzuschätzen, der mit den verschiedenen Darstellungspfaden verbunden ist, um bei Bedarf eine zusätzliche ökonomische Bewertung zu erlauben.

Ein Ergebnis könnten mehrdimensionale Metriken sein, die mit den Objektmetadaten gespeichert und durch das Archivmanagement regelmäßig aktualisiert werden. Hierzu sollte eine Rückkopplung mit den Archivbenutzern erfolgen. So könnten in die Aktualisierungen Bewertungen seitens der Nutzer einfließen, die auf dem praktischen Einsatz bestimmter View-Paths sowie ihrer Handhabbarkeit und Vollständigkeit beruhen.

Aggregation von View-Paths

Wie erläutert, kann für bestimmte Objekttypen mehr als ein Darstellungspfad existieren. Dieses kann die Wahrscheinlichkeit des langfristig erfolgreichen Zugriffs verbessern – jedoch zum Preis potenziell höherer Kosten. Ausgehend vom Objekttyp und der eventuell notwendigen Applikation erhält man weitere View-Paths für andere Objekte, die automatisch anfallen: Ein einfaches Beispiel demonstrieren die sogenannten Office-Pakete, Zusammenstellungen verschiedener Applikationen. Sie können mit einer Vielfalt von Formaten umgehen – nicht nur mit denen der enthaltenen Teilkomponenten, sondern über Importfilter hinaus mit einer Reihe weiterer Dateiformate.¹²

Diese Betrachtungen können dazu dienen, eine potenziell überflüssige Referenzplattform zu identifizieren, die nur für ein bestimmtes Objekt vorgehalten wird, das aber auf alternativen Wegen ebenfalls darstellbar ist. So muss beispielsweise zur Betrachtung eines PDFs nicht eine besonders seltene Plattform genutzt werden, wenn ein gleichwertiger Viewer auf einer mehrfach genutzten anderen Plattform ebenfalls ablauffähig ist (Abbildung 7).

Schwieriger wird jedoch eine Zusammenlegung, wenn von den alternativen View-Paths nicht bekannt ist, ob sie ebenfalls zu 100% das gefragte Objekt rekonstruieren. Ein typisches Beispiel ist der Import von Word-Dokumenten in einer anderen als der Erstellungapplikation. An diesem Punkt könnten ebenfalls die im vorangegangenen Abschnitt vorgenommenen Überlegungen zu Benutzerrückmeldungen in Betracht gezogen werden.

12 Ein Paket wie OpenOffice kann bezogen auf seine Importfilter für die verschiedenen Teilanwendungen wie Textverarbeitung, Tabellenkalkulation oder Präsentation inklusive der Unterstützung verschiedener älterer Versionen schnell über 100 verschiedene Formate statischer Objekte lesen.

Mit jedem View-Path sind bestimmte Kosten verbunden, die vom Archivmanagement beachtet werden müssen. Diese lassen sich für jede Schicht eines

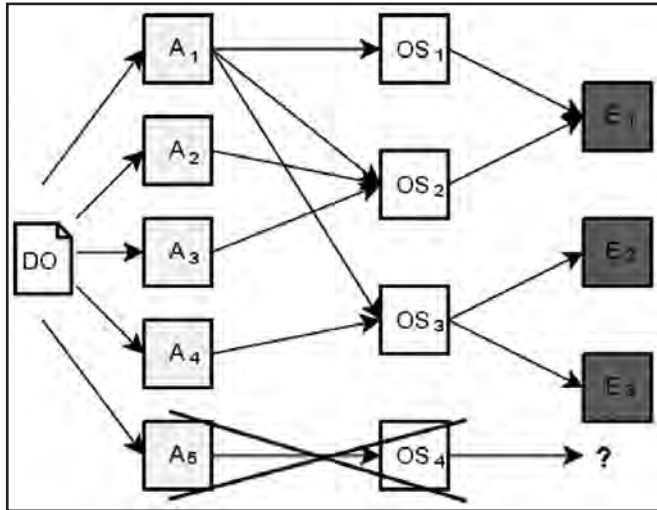


Abbildung 7: Redundante View-Paths zur Darstellung oder zum Abspielen eines digitalen Objekts (DO). Dieses kann im Beispiel von fünf verschiedenen Applikationen (A) geöffnet werden, die auf vier verschiedenen Betriebssystemen (OS) laufen können, wobei für ein OS kein geeigneter Emulator (E) bereitsteht.

View-Paths abschätzen. Übersteigen sie eine gewisse Schwelle, könnten ökonomische Erwägungen an die Aufnahme von Primärobjekten in das Archiv und das Archivmanagement geknüpft werden. In solchen Fällen könnten bestimmte Dateiformate abgelehnt werden, weil deren spätere Rekonstruktion zu hohe Aufwendungen erwarten lässt.

So sollte je nach Situation oder Archiv hinterfragt werden, einen Darstellungspfad aufrecht zu erhalten, wenn sinnvolle Alternativen existieren. Liegt beispielsweise ein Primärobjekt nach dem PDF 1.0 Standard vor, welches mit einem Werkzeug in einer Windows 3.11 Nutzungsumgebung erzeugt wurde, muss deshalb diese Umgebung nicht zwingend erhalten werden. Unter bestimmten Bedingungen kann auf einen View-Path verzichtet werden:

- Es existiert eine ausreichende Anzahl weiterer, gleichwertiger und dabei einfacherer Darstellungspfade.
- Wegen der guten und vollständigen Beschreibung des Formats ist es deutlich einfacher, einen Viewer für die jeweils aktuellen Arbeitsumge-

- bungen zu migrieren, als alte Nutzungsumgebungen durch Emulatoren zu sichern.
- Dieser Objekttyp ist der einzige, der eine Windows 3.11 Umgebung potenziell verlangt.
 - Es gibt kein spezielles Interesse, diese Nutzungsumgebung aufrecht zu erhalten, da sie nicht im Fokus der Institution liegt.

Ein solches Vorgehen ließe sich auf andere View-Paths ausdehnen, die für eine Reihe von Dateiformaten und Applikationen Apple- oder alternativ Microsoft-Betriebssysteme voraussetzen. Wenn beispielsweise kein gesondertes Bedürfnis speziell nach dedizierten Apple-Nutzungsumgebungen besteht, weil beispielsweise die Art der Benutzerinteraktion, das Aussehen der grafischen Oberfläche und der Applikation in dieser von eigenständigem Interesse sind, könnte ein solcher Zweig im Archiv geschlossen werden. Besonders gut lässt sich dies am OpenOffice veranschaulichen, welches für etliche kommerzielle und freie UNIX-Varianten, wie BSD, Solaris oder Linux, Mac-OS X und selbstredend für die Windows-Betriebssysteme angeboten wird.

Ähnlich liegt der Fall in der trivialen Vervielfachung der View-Paths durch unterschiedliche Anpassungen von Programmen und Betriebssystemen an nationale und regionale Besonderheiten. Hierzu zählen natürliche Sprachen, spezielle Zeichen oder auch Währungen und Einheiten. Während Open-Source-Betriebssysteme und Applikationen schon lange mehrere Sprachpakete in einer Installation erlauben, setzte dies sich bei kommerziellen Systemen erst recht spät durch. Das Bedürfnis von einem Betriebssystem wie Windows 2000 alle verschiedenen Landesversionen aufzubewahren, lässt sich vermutlich besser in Kooperationen ähnlicher Gedächtnisinstitutionen über Landesgrenzen hinweg erreichen. Dies würde zudem Kosten und Aufwand für redundante Lizensierungen reduzieren helfen.

Einfacher und damit oft günstiger zu pflegenden View-Paths kann der Vorrang vor anderen eingeräumt werden. Jedoch sind auch hier Voraussagen schwierig und es kann die Gefahr bestehen, dass sich mit dem Wandel der Referenzumgebungen die Kostenstrukturen erneut ändern. Andererseits lassen sich Vorkehrungen treffen, dass seltenere View-Paths zumindest an spezialisierten Institutionen mit besonderem Sammelauftrag und entsprechender Finanzierung weiter betreut werden.

Die verschiedenen Strategien der Langzeitarchivierung der Emulatoren – Migration oder Schachtelung generieren unterschiedliche Aufwendungen im Archivbetrieb:

- Die geschachtelte Emulation sorgt für eher längere Darstellungspfade bei geringem Migrationsbedarf. Der Aufwand entsteht beim zunehmend komplexer werdenden Zugriff auf das Primärobjekt.
- Die Migration von Emulatoren, Universal Virtual Machines und modulare Ansätze¹³ generieren eher kurze View-Paths bei einfacherem Zugriff auf das Primärobjekt. Jedoch liegt der Aufwand im regelmäßigen Update aller benötigter Emulatoren oder ihrer Teilkomponenten.

Der erste Ansatz ist aus diesem Grund eher für Objekte mit seltenem Zugriff oder Institutionen mit kleinen, speziell ausgebildeten Nutzerkreisen wie Archive, geeignet. Die zweite Strategie passt sicherlich besser auf viel genutzte

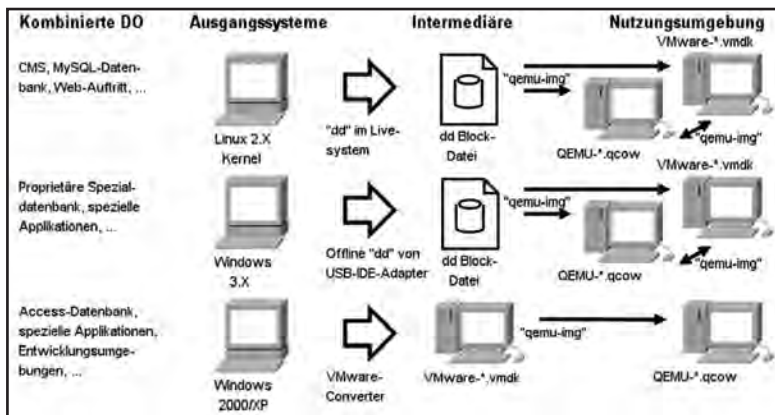


Abbildung 8: In Abhängigkeit vom zu archivierenden Gesamtsystem kombinierter digitaler Objekte existieren verschiedene Wege zur Archivierung.

Objekttypen mit größeren Anwenderkreisen. Eine Reihe von Kostenfaktoren kann durch gemeinschaftliche Anstrengungen und verteilte Aufgaben reduziert werden.

Kombinierte Archivobjekte

Primärobjekte mit bereits im Paket enthaltenen notwendigen Sekundärobjekten wie bestimmte Metadaten zur Beschreibung, Viewer und Hilfsprogramme, zusammengeführt in einem Archival Information Package (AIP), können bereits

13 Vgl. van der Hoeven; van Diessel; van der Meer (2005).

aufbereitet vorgehalten werden.

Solche kombinierten Objekte könnten aus der Überführung eines physischen Systems in ein virtuelles entstehen. Dieses Vorgehen lässt sich beispielsweise auf sehr spezielle dynamische Objekte, wie Datenbanken von Content Management Systemen (CMS) für WWW-Auftritte, Firmendaten aus Produktionsanlagen oder Entwicklungsabteilungen vorstellen (Abbildung 8): Statt Daten, Applikation und notwendiges Betriebssystem im Zuge der Archivierung voneinander zu lösen, könnte eine komplette Maschine in eine emulierte Umgebung überführt werden. Damit bleibt die Serverfunktionalität mit ihren gesamten Einstellungen als eine Einheit erhalten. Das Verfahren kann auf diese Weise sicherstellen, dass ein AIP über Archive hinweg ohne einschränkende Nebenbedingungen ausgetauscht werden kann.

Das Zusammenfassen bestimmter View-Paths in einer gemeinsamen Umgebung könnte einerseits für bestimmte Forschungsthemen und andererseits für eine Vereinfachung der View-Path-Erstellung für den Benutzerbetrieb sinnvoll sein.

Eine etwas anders gelagerte Zusammenfassung schlagen Reichherzer¹⁴ und Brown (2006) vor, um auf Datensätze der öffentlichen Administration zu späteren Zeitpunkten zugreifen zu können. Solche typischerweise in einem komprimierten Archiv zusammengefassten Daten von Erhebungen, Statistiken oder Auswertungen umfassen eine Reihe typischer Objektformate einer gewissen Epoche. Diese könnten in einem gemeinsamen Container untergebracht sein und Hilfsprogramme, wie den Dokumentenausdruck nach Postscript oder PDF enthalten. Passende virtuelle Drucker, die durch Postscript- oder PDF-Generatoren repräsentiert werden, können fehlende Exportfilter alter Applikationen ersetzen.

View-Paths als OAIS-Workflows

Im OAIS-Referenzmodell übernimmt das Management des digitalen Langzeitarchivs eine Reihe von Aufgaben, die sich mit dem Lifecycle-Management und damit verbundenen Workflows von Primärobjekten befassen. Hierfür wurden besonders die Arbeitsprozesse für die Objektausgabe betrachtet. Für das langfristige Management wird insbesondere die Aufgabe des Preservation Planning interessant.¹⁵

Ein zentrales Moment ist die regelmäßige Kontrolle der View-Paths bei

14 Die Idee hierzu findet sich bereits bei Rothenberg (1998) sowie Reichherzer und Brown (2006).

15 van Diessen (2002a)

einem Wechsel der Referenzumgebung als Bezugsgröße. Jeder Plattformwechsel stellt neue Anforderungen für die Wiederherstellung von Nutzungsumgebungen. Bei dieser Überprüfung handelt es sich um einen iterativen Prozess, der über alle registrierten Objekttypen des Archivs abläuft. Hierfür ist jeweils eine passende Strategie für den Übergang von einer Referenzumgebung auf eine neue zu suchen. Generell gilt: Neue Objekttypen und Dateiformate erfordern neue Darstellungspfade.

Es ergeben sich verschiedene Anforderungen an den Archivbetrieb:

- Erstellung eines Hintergrundarchivs - In diesem werden die einzelnen Elemente des View-Path dauerhaft abgelegt. Sie werden dann genauso behandelt wie Primärobjekte. An dieser Stelle kann überlegt werden, ob bestimmte Einzelobjekte, wie Emulatoren, spezifische Hilfsprogramme und Beschreibungen in einem AIP gebündelt oder einzeln abgelegt werden.
- Betrieb eines Online-Archivs für den Direktzugriff - Für häufig nachgefragte Sekundärobjekte kann es sinnvoll sein, diese zusätzlich zum Langzeitarchiv in einem speziellen Archiv, wie einem aktuellen Dateisystem einer Referenzumgebung, vorzuhalten. Das kann einerseits das Langzeitarchiv entlasten und andererseits zu einem beschleunigten Ablauf der View-Path-Erstellung führen.
- Anlage eines View-Path-Caches - Für häufiger nachgefragte und aufwändiger zu generierende Darstellungspfade kann die Vorhaltung vorbereiteter Nutzungsumgebungen den Aufwand für Nutzer und Archivbetreiber reduzieren. Diese Caches könnten als Teil des Online-Archivs oder direkt auf der Referenzplattform abgelegt sein.

Die aufgezeigten Überlegungen haben klare ökonomische Implikationen für die mit digitalen Objekten befassten Gedächtnisorganisationen. Sie werden sich im Zuge der technischen Workflows, die sich mit der Wiederherstellung von Nutzungsumgebungen befassen, einen Teil der notwendigen Entwicklungen der Emulatoren und Viewer selbst leisten oder diese Leistung am Markt einkaufen müssen. Entstehen bei den angestrebten Entwicklungen offene Standards und Werkzeuge, wie PRONOM¹⁶ oder Dioscuri,¹⁷ können sich einerseits die Belastungen der Einzelinstitution in Grenzen halten und andererseits verbindliche Verfahren entwickeln, die von einer breiten Anwendergemeinschaft unterstützt werden.

16 Vgl. <http://dioscuri.sourceforge.net/>

17 Vgl. <http://www.nationalarchives.gov.uk/PRONOM/Default.aspx>

Literatur

- The Preservation Manager for the e-Depot: http://www.kb.nl/hrd/dd/dd_onderzoek/preservation_subsystem-en.html
- van Diessen, Raymond; Steenbakkers, Johan F. (2002): *The Long-Term Preservation Study of the DNEP project - an overview of the results*, The Hague, The Netherlands; http://www.kb.nl/hrd/dd/dd_onderzoek/reports/1-overview.pdf
- van Diessen, Raymond (2002): *Preservation Requirements in a Deposit System*, The Hague, The Netherlands; http://www.kb.nl/hrd/dd/dd_onderzoek/reports/3-preservation.pdf
- van der Hoeven, Jeffrey; van Diessen, Raymond; van der Meer, K. (2005): *Development of a Universal Virtual Computer (UVC) for long-term preservation of digital objects*, Journal of Information Science, Vol. 31, No. 3, 196-208; DOI: 10.1177/0165551505052347
- van Diessen, Raymond; van der Werf-Davelaar, Titia (2002): *Authenticity in a Digital Environment*, The Hague, The Netherlands; http://www.kb.nl/hrd/dd/dd_onderzoek/reports/2-authenticity.pdf
- Rothenberg, Jeff (1998): *Avoiding Technological Quicksand: Finding a Viable Technical Foundation for Digital Preservation in „The State of Digital Preservation: An International Perspective“*, Washington; <http://www.clir.org/pubs/reports/rothenberg/contents.html>
- Reichherzer, Thomas; Geoffrey, Geoffrey (2006): *Quantifying software requirements for supporting archived office documents using emulation*, International Conference on Digital Libraries 06: Proceedings of the 6th ACM/IEEE-CS joint Conference on Digital Libraries; S. 86-94
- von Suchodoletz, Dirk (2009): *Funktionale Langzeitarchivierung digitaler Objekte - Erfolgsbedingungen des Einsatzes von Emulationsstrategien*, urn:nbn:de:0008-2008070219, ISBN 978-3-86727-979-6