

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

nestor Handbuch

Eine kleine Enzyklopädie
der digitalen Langzeitarchivierung

Version 2.3

Kapitel 14
Geschäftsmodelle

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 14 „Geschäftsmodelle“ (Version 2.3): [urn:nbn:de:0008-20100624138](http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:0008-20100624138)
<http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:0008-20100624138>



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

14 Geschäftsmodelle

14.1 Einführung

Achim Oßwald

Neben der vor dem Hintergrund neuerer Verfahren und Erfahrungen weiterhin relevanten Frage, auf welche Weise Langzeitarchivierung optimal realisiert werden könnte und sollte, drängt sich eine weitere Frage in den Vordergrund: Wie können die ausgewählten Verfahren finanziert und in Geschäftsmodelle eingebunden werden?

Nur in den LZA-Anfängen ist optimistisch über die Frage der Kosten spekuliert worden. Zu dieser Zeit bestand die Hoffnung, die Sicherung digitaler Objekte könne günstiger ausfallen als beispielsweise jene von Druckwerken. Schon bald jedoch wurde deutlich, dass die gewählten Maßnahmen zur Erhaltung bzw. Erneuerung von Daten, Datenträgern und Wiedergabeumgebungen relativ aufwändig und teuer sind. Dies aber bedeutet, dass die mit einem Verfahren der Langzeitarchivierung und Langzeitverfügbarkeit verbundenen Kosten – je nach gewähltem Archivierungskonzept – gänzlich oder zumindest in Teilen unregelmäßig wiederkehrend anfallen.

Zu ermitteln und zu analysieren, welche Kostenfaktoren überhaupt bei der LZA zum Tragen kommen und welche konkreten Kosten damit aus heutiger Sicht verbunden sein werden, ist Gegenstand von einigen wenigen Projekten. Bis auf weiteres stellt diese Frage für alle Einrichtungen, die mit der Aufgabe der Langzeitarchivierung aufgrund gesetzlicher Bestimmungen und sonstiger Vereinbarungen betraut sind, einen nur begrenzt abgesicherten Aspekt dar.

Kapitel 14.2 „Kosten“ gibt einen Einblick in die aktuelle Diskussion zum Thema und einen Überblick zu den Kostenfaktoren, die voraussichtlich mit den Aktivitäten für die Langzeitarchivierung und Langzeitverfügbarkeit verbunden sein werden.

Kapitel 14.3. „Service- und Lizenzmodelle“ zeigt auf, welche Optionen sich auf dieser Grundlage für dienstleistende Organisationen und Einrichtungen abzeichnen und welche Servicemodelle derzeit von den als Dienstleister aktiven Organisationen angeboten werden könnten.

Niemand weiß heute, ob die derzeit praktizierten Verfahren zur Langzeitarchivierung und Langzeitverfügbarkeit mit ihren z. T. sehr unterschiedlichen Geschäftsmodellen ihrerseits wiederum langzeitfähig sind. Dies wird sich zukünftig erweisen. Umso größere Sorgfalt und Professionalität ist notwendig, wenn Verfahren und Strategien für die Langzeitarchivierung und Langzeitverfügbarkeit ausgewählt und in öffentlichem oder privatwirtschaftlichem Auftrag realisiert werden. Dauerhafte Finanzierungskonzepte sind dabei eine unabdingbare Voraussetzung, um die methodischen und technischen Überlegungen dauerhaft zum Tragen kommen lassen zu können.

14.2 Kosten

Thomas Wollschläger und Frank Dickmann

In diesem Kapitel werden Kostenfaktoren benannt, die für den Betrieb eines digitalen Langzeitarchivs von Bedeutung sind. Des Weiteren werden Ansätze vorgestellt, wie die individuellen Kosten der LZA in einer Institution ermittelt werden können.

Kostenfaktoren bei Einrichtung und Unterhaltung eines Langzeitarchivs

Abhängig vom konkreten Langzeitarchivierungskonzept der jeweiligen Einrichtung werden folgende Kostenfaktoren zu berücksichtigen sein:

Initiale Kosten

- Informationsbeschaffung über LZA-Systeme
- Erhebung von Bestand, Zugang und gewünschten Zugriffsoptionen für digitale Materialien im eigenen Haus
- Erhebung von vorhandenen Personal- und Technikressourcen im eigenen Haus
- Projektplanung, ggf. Consulting, Ausschreibung(en)

Beschaffungskosten

- Hardware: Speichersysteme und *sämtliche* infrastrukturellen Einrichtungskosten (Serververbindungen, Datenleitungen, Mitarbeiterrechner usw.)
- Ggf. Lizenz(en) für Software-Systeme oder Beitrittskosten zu Konsortien
- Weitere Aufwendungen: z.B. Anpassungsentwicklungen von Open Source Software-Produkten, Entwicklung/Anpassung von Schnittstellen, Erstellung von physischen und digitalen Schutzmaßnahmen (auch solche aus rechtlichen Gründen)
- Ggf. Einstellung neuer Mitarbeiter und/oder Schulung vorhandener Mitarbeiter

Betriebskosten

- Dateningest des bisher vorhandenen Materials
- Dateningest des neu eingehenden Materials
- Laufende Storage-Kosten

- Sonstige Dauerbetriebskosten: z.B. Strom, Datenleitungskosten, *sämtliche* Sicherheitsmaßnahmen, Backups, regelmäßige Wartung(en) und Tests, Software-Upgrades
- Zukauf von weiteren Speichereinheiten
- Hard- und Software-Komplettersatz in Intervallen
- Ggf. laufende Lizenzkosten und/oder laufende Beitragszahlungen bei Konsortien

Die konkreten Kosten sind dabei jeweils abhängig von

- der Zahl und Komplexität der Workflows bei einer Institution
- der Menge, Heterogenität und Komplexität der zu archivierenden Objekte und ihrer Metadaten
- den gewünschten Zugriffsmöglichkeiten und Schnittstellen sowie ggf.
- den Anforderungen Dritter an die archivierende Institution bzw. Verpflichtungen der Institution gegenüber Dritten

Die Ermittlung von Kosten für die Langzeitarchivierung

Die tatsächliche Ermittlung der Kosten, die auf eine Einrichtung für die Langzeitarchivierung ihrer digitalen Dokumente zukommen, gestaltet sich in der Praxis noch relativ schwierig. Viele LZA-Unternehmungen befinden sich derzeit noch im Projektstatus oder haben noch nicht lange mit dem produktiven Betrieb begonnen. Daher liegen noch wenige Erfahrungswerte vor, wie sich insbesondere der laufende Betrieb eines solchen Archivs kostenmäßig erfassen lässt. Außerdem befindet sich nach wie vor die zunehmende Menge und Varianz insbesondere der Internet-Publikationen in einem Wettlauf mit den technischen Möglichkeiten, die von Gedächtnisorganisationen zur Einsammlung und Archivierung eingesetzt werden können.

Einen begrenzten Anhaltspunkt können die angesprochenen Unternehmungen zumindest in der Hinsicht liefern, was die Ersteinrichtung eines digitalen Langzeitarchivs betrifft. BMBF und DFG haben eine ganze Reihe von solchen Projekten gefördert und verschiedene Institutionen haben Projekte aus eigenen Mitteln finanziert.¹ Das bisher am umfangreichsten geförderte LZA-Vorhaben in Deutschland war das Projekt kopal mit einem Fördervolumen von 4,2 Mio. Euro.² Diese Kosten schließen die vollständige Entwicklung eines Ar-

1 Siehe dazu die Projektübersicht in der nestor-Informationsdatenbank http://nestor.sub.uni-goettingen.de/nestor_on/browse.php?zeitg=10

Alle hier aufgeführten URLs wurden im Mai 2010 auf Erreichbarkeit geprüft .

2 Vgl. Wollschläger (2007), S. 247.

chivsystems einschließlich Objektmodell, Aufbau von Hard- und Softwareumgebungen in mehreren Einrichtungen und mehrjährige Forschungsarbeiten ein. Zum Projektende hat kopal allerdings in einem Servicemodell konkrete Kosten für den Erwerb eines vollständigen Archivs zum Eigenbetrieb vorgelegt. Wenn das kopal-Archivsystem unter Zukauf von Beratung und ggf. Entwicklung eigenständig betrieben wird, soll ein Erstaufwand für Hard- und Software eines Systems mittlerer Größe von ca. 750.000 € anfallen. Hiervon entfielen 40% auf Softwarelizenzen und 60% auf Systembereitstellung und -betrieb.³ Wiewohl solche Angaben nur exemplarisch sein können, kann dennoch davon ausgegangen werden, dass die Kosten für die Ersteinrichtung eines LZA-Systems in einer Einrichtung einen gewissen Schwellenwert nicht unterschreiten werden.

Die Zahl der Ansätze, die bisher versucht haben, Modelle für die Betriebskostenermittlung digitaler LZA zu entwickeln, ist begrenzt. Nennenswert ist hierbei der Ansatz des LIFE-Projekts aus Großbritannien. „The LIFE Project“ war ein einjähriges Projekt (2005/2006) der British Library (BL) in Zusammenarbeit mit dem University College London (UCL) mit dem Hauptziel, ein Kostenmanagement für die Langzeiterhaltung elektronischer Ressourcen zu entwickeln. Es wurde eine Formel zur Ermittlung der Archivierungskosten entwickelt. Manche Fragen mussten noch offen bleiben, so war es z.B. bislang nicht adäquat möglich, im Rahmen des Projektes die Kosten der Langzeiterhaltung von gedruckten und elektronischen Veröffentlichungen zu vergleichen. Die Formel lautet: $L_T = Aq + I_T + M_T + Ac_T + S_T + P_T$. Dabei stehen die Werte für folgende Parameter.⁴

- L = complete lifecycle cost over time 0 to T.
- Aq = Acquisition
- I = Ingest
- M = Metadata
- Ac = Access
- S = Storage
- P = Preservation

Jeder der Parameter kann weiter in praktische Kategorien und Elemente aufgeteilt werden. Alle Schritte können entweder, wenn der Prozess direkt kalkulierbar ist, als Kostenfaktor berechnet werden oder, wenn nötig, jeweils auch noch

3 Siehe kopal (2007), S. 2.

4 Vgl. McLeod/Wheatley/Ayris (2006), S. 6. Das tiefergestellte (T) in der Formel bedeutet als Attribut der Parameter „over time“.

in beliebig viele Unterpunkte untergliedert werden. So kann die Berechnung für die jeweilige Institution individuell angepasst werden. Innerhalb des LIFE-Projekts wurden zum einen beispielhafte Berechnungen der LZA-Kosten des Projektmaterials vorgenommen und dabei Kosten für „the first year of a digital asset’s existence“ und „the tenth year of the same digital assets’ existence“ vergleichbar ermittelt⁵ und exemplarisch auch die Kosten pro Speichermenge. Zum anderen hat das Projekt die entwickelten Formelwerke zur Verfügung gestellt, so dass interessierte Institutionen selbst Berechnungen anhand der Individualparameter vornehmen können.

In Anbetracht der aktuellen Erkenntnisse fokussieren Kostenschätzungen für die LZA hauptsächlich auf die Kosten pro Speichermenge, wie z.B. in LIFE. Bei eingehender Betrachtung der Prozesse, die im Rahmen eines implementierten LZA-Systems anfallen, sind jedoch die Speicherkosten nicht der Hauptkostenfaktor. Vielmehr lässt sich aus den Prozessen Acquisition, Metadata und Preservation eine hohe Personalintensität ableiten, insbesondere aus dem Grund, da diese Prozesse nur sehr eingeschränkt automatisierbar sind. Daher sind es vielmehr die Personalkosten, die langfristig den höchsten Anteil an den LZA-Kosten haben werden.⁶

Ebenso von Bedeutung ist die Anzahl unterstützter Formate, da diese Anzahl und der Personaleinsatz eng miteinander verknüpft sind. Jedes zusätzliche Format erfordert zusätzlichen Aufwand durch qualifiziertes Personal. Demgegenüber steht aber der Nutzen, den das Angebot eines Formats liefert. Hierzu wurde festgestellt, dass weniger häufig verwendete Formate ein unterdurchschnittliches Kosten-Nutzen-Verhältnis aufweisen. Beispielsweise haben in LIFE die Formate PDF, TXT und HTML ca. 85% aller Dokumente abgedeckt, allerdings nur 7% der Kosten verantwortet, während die 12 am wenigsten verwendeten Formate 0,1% der Dateien abgedeckt, aber ca. 41% der Kosten hervorgerufen haben. Grundlage dabei sind die Gesamtkosten über einen Zeitraum von 20 Jahren.⁷ Aus diesem Grund ist eine Einschränkung der Formatvielfalt ein empfehlenswertes Kostensteuerungsinstrument für die LZA.

Eine bedeutende Frage für die Festlegung der Archivierungsstrategie – nämlich für das eigentliche „Preservation Planning“, die Erhaltungsmaßnahmen über die Lebenszeit eines digitalen Objekts – einer Institution ist, ob auf Dauer Migrationen oder Emulationen kostengünstiger sind. Hierzu sind noch keine abschließenden Aussagen möglich. Generell verbreitet ist die Auffassung, dass Migration der kostengünstigere Weg sei. Innerhalb von LIFE wurden dazu An-

5 Vgl. ebenda, S. 3.

6 Vgl. Ashley (1999), S. 123.

7 Vgl. Björk, B.-C. (2007), S. 23.

sätze formuliert, die jedoch hauptsächlich sehr exemplarische Migrationen behandeln und noch nicht repräsentativ sind.⁸ Andere Studien kommen dagegen zu dem Schluss, dass Emulationen auf längere Sicht kostengünstiger seien:

While migration applies to all objects in the collection repetitively, emulation applies to the entire collection as a whole. This makes emulation most cost-effective in cases of large collections, despite the relatively high initial costs for developing an emulation device. When considering the fact that only small fragments of digital archives need to be rendered in the long run, it may turn out that from a financial perspective emulation techniques will be more appropriate for maintaining larger archives.⁹

Da die bestehenden Langzeitarchive gerade erst dabei sind, die ersten „echten“ Maßnahmen von Preservation Planning umzusetzen, wird hier auf Erfahrungswerte zu warten sein, die entsprechende Ergebnisse unterstützen können.

Konsequenzen für die Gedächtnisorganisationen

Angesichts der zu erwartenden nicht unerheblichen Kosten für die *Ersteinrichtung* eines LZA-Systems dürften kleinere Einrichtungen nicht umhin kommen, zwecks Einrichtung eines solchen Systems mit anderen Institutionen zu kooperieren bzw. sich einem bestehenden System anzuschließen und/oder sich den Zugang dazu über Lizenzen zu sichern. Selbst größere Institutionen werden für die Einrichtung eines LZA-Systems oft kooperative Formen wählen, um hohe Ersteinrichtungskosten aufzuteilen, die sich sonst nicht auf mehrere Schultern verteilen lassen. Ebenso könnte angesichts der noch bestehenden Unsicherheit, wie sich künftig die Kosten für den Dauerbetrieb des Langzeitarchivs und das Preservation Planning entwickeln werden, die Entscheidung zugunsten der Variante ausfallen, sich in bestehende Systeme einzukaufen oder über kostenpflichtige Lizenzen Teilnehmer an einem kommerziell ausgerichteten System zu werden. Letzteres macht in der Regel Zugeständnisse an die gewünschte Preservation Policy notwendig, so dass eine Gedächtnisorganisation abwägen muss, welche Kosten – Lizenzen für ein kommerzielles System oder eigene Entwicklungskosten, z.B. für die Anpassung von Open Source Software – die jeweils lohnendere Investition ist.

Die Teilnahme an kooperativen Formen der Langzeitarchivierung ist unter Kostenaspekten in jedem Fall empfehlenswert. Hierbei können Institutionen über z.B. gemeinsame Speichernutzung bzw. gegenseitiges Backup, gegenseitige Nutzung von Entwicklungsergebnissen, gemeinsame Adressierung

8 Vgl. Ebenda, S. 10.

9 Zitiert nach Oltmans/Kol (2005), #5 – Conclusion.

übergreifender Herausforderungen oder kooperative Verwaltung von Open Source Software Synergien schaffen und erhebliche Ressourceneinsparungen ermöglichen.

Quellen und Literatur

- Ashley, K. (1999): *Digital Archive Costs: Facts and Fallacies*, in: *Proceedings of the DLM-Forum on Electronic Records (DLM '99)*, DLM-Forum, Brussels, 1999, S. 123, http://ec.europa.eu/archives/ISPO/dlm/fulltext/full_ashl_en.htm
- Björk, B.-C. (2007): *Economic evaluation of LIFE methodology*, LIFE Project, London, UK, 2007, URL: <http://eprints.ucl.ac.uk/7684/1/7684.pdf>
- Kopal (2007): *kopal: Ein Service für die Langzeitarchivierung digitaler Informationen*. Projekt kopal (Kooperativer Aufbau eines Langzeitarchivs digitaler Informationen), 2007, http://kopal.langzeitarchivierung.de/downloads/kopal_Services_2007.pdf
- McLeod, Rory; Wheatley, Paul; Ayris, Paul (2006): *Lifecycle Information for E-literature* : A summary from the LIFE project ; Report Produced for the LIFE conference 20 April 2006. LIFE Project, London (via <http://www.ucl.ac.uk/life/lifeproject/> or directly under <http://eprints.ucl.ac.uk/archive/00001855/01/LifeProjSummary.pdf>)
- Oltmans, Erik; Kol, Nanda (2005): *A Comparison Between Migration and Emulation in Terms of Costs*. In: RLG DigiNews, Volume 9, Number 2, 15.04.2005 (<http://worldcat.org/arcviewer/1/OCC/2007/07/10/0000068902/viewer/file1.html#article0>)
- Wollschläger, Thomas (2007): *kopal – ein digitales Archiv zur dauerhaften Erhaltung unserer kulturellen Überlieferung*. In: *Geschichte im Netz : Praxis, Chancen, Visionen ; Beiträge der Tagung .hist2006*, Berlin: Clio-online und Humboldt-Universität zu Berlin, 2007, S. 244 – 257 (Historisches Forum 10 (2007), Teilband I).
- Siehe außerdem die Einträge in der nestor-Informationsdatenbank zum Thema „Kosten“ unter http://nestor.sub.uni-goettingen.de/nestor_on/browse.php?show=8

14.3 Service- und Lizenzmodelle

Thomas Wollschläger und Frank Dickmann

In den wenigsten Fällen werden Langzeitarchivierungssysteme von einer einzigen Institution produziert und genutzt. Schon bei einer zusätzlichen Nutzer- oder Kundeninstitution für das hergestellte und/oder betriebene Archivsystem müssen Lizenz- oder Geschäftsmodelle aufgestellt sowie Servicemodelle für zu leistende Langzeitarchivierungs-Dienstleistungen definiert werden.

Lizenzmodelle

Lizenzkosten fallen in der Regel für die Nutzung kommerzieller Softwareprodukte an. Dabei gibt es unterschiedliche Möglichkeiten. Zum einen können solche Produkte lizenziert und eigenständig in der eigenen Institution eingesetzt werden. Dabei ist die Hersteller- oder Vertriebsfirma neben den (einmalig oder regelmäßig) zu zahlenden Lizenzgebühren zumeist durch Support- und Updateverträge mit der Nutzerinstitution verbunden. Beispiele hierfür sind etwa das System *Digitool* der Firma Exlibris¹⁰ oder das *DIAS*-System von IBM.¹¹

Zum anderen besteht bei einigen Produkten die Möglichkeit, dass eine Betreiberinstitution (die nicht identisch mit dem Hersteller oder Systemvertreiber sein muss) das Archivsystem hostet und eine Nutzung für Dritte anbietet. Hierbei werden Lizenzkosten meist vom Betreiber auf die Kunden umgelegt oder fließen in die Nutzungskosten für die Archivierung ein. Ein Beispiel hierfür ist das insbesondere auf die Archivierung von e-Journals ausgerichtete System *Portico*. Hierbei erfolgt eine zentrale, an geografisch auseinander liegenden Orten replizierte Archivierung. Die Kosten von Portico richten sich für eine Bibliothek nach dem verfügbaren Erwerbungssetat. Der jährliche Beitrag für die Nutzung des Systems kann daher je nach dessen Höhe zwischen 1% des Erwerbungssetats und maximal 24.000 US-\$ liegen.¹²

Neben den kommerziellen Produkten gibt es eine Reihe von Open Source-Lösungen im Bereich der Archivierungssysteme. Durch die Nutzung von Open Source-Lizenzen¹³ fallen oft keine Lizenzgebühren bzw. -kosten für die Nutzerinstitutionen an, sondern zumeist nur Aufwands- und Materialkosten. Zudem sind Archivinstitutionen, die eine Open Source-Software oder ein Open Sour-

10 Siehe <http://www.exlibrisgroup.com/digitool.htm>

11 Siehe <http://www-05.ibm.com/nl/dias/>

12 Vgl. http://www.portico.org/libraries/aas_payment.html

13 Siehe hierzu v.a. <http://www.opensource.org/licenses>

ce-Netzwerk nutzen, dahingehend gefordert, durch eigene Entwicklungsbeiträge das Produkt selbst mit weiterzuentwickeln.¹⁴ Beispiele für verbreitete Open Source-Lösungen sind das System *DSpace*¹⁵ und die *LOCKSS*- bzw. *CLOCKSS*-Initiative.¹⁶ Die *LOCKSS*-Technologie will die langfristige Sicherung des archivierten Materials dadurch sicherstellen, dass jedes Archivobjekt mit Hilfe des Peer-to-Peer-Prinzips bei allen Mitgliedern gleichzeitig gespeichert wird. Jedes Mitglied stellt einen einfachen Rechner exklusiv zur Verfügung, der im Netzwerk mit den anderen Mitgliedern verbunden ist und auf dem die *LOCKSS*-Software läuft.

Neben der Nutzung reiner kommerzieller Lösungen und reiner Open Source-Lösungen gibt es auch Mischformen. Dabei kann es von Vorteil sein, nur für Teile des eigenen LZA-Systems auf kommerzielle Produkte zurückzugreifen, wenn sich dadurch beispielsweise die Höhe der anfallenden Lizenzkosten begrenzen lässt. Andererseits erwirbt man mit vielen Lizenzen zumeist auch Supportansprüche, die etwa bei geringeren eigenen Entwicklungskapazitäten willkommen sein können. Ein Beispiel für eine solche LZA-Lösung ist das *kopal*-System. Hierbei wird das lizenz- und kostenpflichtige (modifizierte) Kernsystem *DIAS* verwendet, während für den Ingest und das Retrieval die kostenfreie Open Source-Software *koLibRI* zur Verfügung gestellt wird.¹⁷

Eine Institution muss somit abwägen, welches Lizenzmodell für sie am vorteilhaftesten ist. Kommerzielle Lizenzen setzen den Verwendungs- und Verbreitungsmöglichkeiten der Archivsysteme oft enge Grenzen. Open Source-Lizenzen bieten hier in der Regel breitere Möglichkeiten, verbieten aber ggf. die Exklusivität bestimmter Funktionalitäten für einzelne Institutionen. Hat sie ausreichende Entwicklungskapazitäten und Hard- bzw. Softwareausstattung, kann die Nutzung von Open Source-Lösungen ein guter und gangbarer Weg sein. Dies gilt beispielsweise auch, wenn sich die Institution als Vorreiter für leicht nachnutzbare Entwicklungen sieht oder im Verbund mit anderen Einrichtungen leicht konfigurierbare Lösungen erarbeiten will. Hat sie jedoch nur geringe Entwicklungsressourcen und decken die kommerziellen Lizenzen alle benötigten Services ab, so kann trotz ggf. höherer Lizenzkosten die Wahl kommerzieller Produkte bzw. von standardisierten Services seitens LZA-Dienstleistern angeraten sein.

14 Vgl. hierzu insbesondere das Kapitel „Kostenrelevante Eigenschaften einer ungewöhnlichen Organisationsform“, in: Lutterbeck/Bärwolff/Gehring (2007), S. 185 – 194.

15 Siehe <http://www.dspace.org/>

16 Siehe <http://www.lockss.org/>

17 Siehe http://kopal.langzeitarchivierung.de/index_koLibRI.php.de.

Servicemodelle

Wie bereits dargestellt, bestehen die wesentlichen Faktoren für die Entscheidung einer Institution für bestimmte Lizenz- und Geschäftsmodelle in den von ihr benötigten Services zur Langzeitarchivierung.¹⁸ Entscheidungskriterien für die Wahl der Einrichtung und/oder Nutzung bestimmter LZA-Services können sein:

Auftrag und Selbstverständnis

- Liegt ein (z.B. gesetzlicher) Auftrag vor, dass die Institution digitale Dokumente eines bestimmten Portfolios sammeln und (selbst) langzeitarchivieren muss?
- Gilt dieser Auftrag auch für Materialien Dritter (z.B. durch Pflichtexemplarregelung)?
- Hat die Institution den Anspruch oder das Selbstverständnis, LZA-Services selbst anbieten oder garantieren zu wollen?
- Liegt eine rechtliche Einschränkung vor, Materialien zwecks LZA Dritten zu übergeben?

Ausstattung und Ressourcen

- Hat die Institution die benötigte Hardware- und/oder Softwareausstattung bzw. kann sie sie bereitstellen, um LZA betreiben zu können?
- Tritt die Institution bereits als Datendienstleister auf oder ist sie selbst von Datendienstleistern (z.B. einem Rechenzentrum) abhängig?
- Stehen genügend personelle Ressourcen für den Betrieb, den Support (für externe Nutzer) und für nötige Entwicklungsarbeiten zur Verfügung?
- Lassen die Lizenzen des genutzten Archivsystems / der Archivsoftware eine Anbindung Dritter an die eigene Institution zwecks LZA zu?

18 Selbstverständlich spielen auch die technischen Möglichkeiten des eingesetzten Archivsystems selbst eine wesentliche Rolle. Einen Kriterienkatalog zur technischen Evaluierung von Archivsystemen bietet z.B. das Kapitel *Software Systems for Archiving* bei Borghoff (2003), S. 221 – 238.

Je nachdem, wie diese Fragen beantwortet werden, stehen für die Wahl des Servicemodells potentiell viele Varianten zur Verfügung. Diese drehen sich im Wesentlichen um die folgenden Konstellationen:

- Die Institution stellt einen LZA-Service (nur) für digitale Dokumente aus eigenem Besitz bereit.
- Die Institution stellt diesen LZA-Service auch für Dritte zur Verfügung.
- Die Institution stellt selbst keinen LZA-Service bereit, sondern nutzt die Services eines Dritten für die Archivierung der eigenen Daten.

Dabei ist jeweils zusätzlich und unabhängig von der Frage, welche Institution den *Service* an sich anbietet, relevant, ob die Daten bzw. respektive die Hardware-/Storage-Umgebung von der Service-Institution selbst oder von Dritten gehostet wird. Beispielsweise kann eine Institution verpflichtet sein, selbst einen LZA-Service anzubieten. Dennoch mag der Umfang des jährlich anfallenden Materials den aufwändigen Aufbau einer solchen Hardware-/Storage-Umgebung sowie entsprechender Betriebskompetenzen nicht rechtfertigen. Hier könnte die Institution entscheiden, zwar einen LZA-Service aufzubauen – und ggf. sogar Dritten gegenüber ein entsprechendes Geschäftsmodell anzubieten –, das Datenhosting jedoch an einen geeigneten Dienstleister abzugeben. Ein Beispiel für ein solches Servicekonzept ist das *kopal*-Projekt. Die Hauptmandanten betreiben zwar gemeinschaftlich das Archivsystem *kopal* und stellen ihre Dienstleistungen (zumeist kleineren) Nutzerinstitutionen zur Verfügung, die eigentliche Datenhaltung wird jedoch bei einem Rechenzentrum betrieben, wo die gemeinschaftlich genutzte Hardware zentral gehostet und per Fernzugriff genutzt wird.¹⁹

Zu den einzelnen Dienstleistungen, die im Rahmen eines LZA-Service-Modells von einer Institution angeboten werden können, gehören beispielsweise folgende:

- Der Betrieb des LZA-Systems und Annahme von Archivmaterial
- Durchführung von Erhaltungsmaßnahmen (von Bitstream-Preservation bis zur Migration von Material)
- Zurverfügungstellung von Datenkopien bei Datenverlusten seitens der Abliefererinstitution
- Bereitstellen eines Pseudonymisierungsdienstes²⁰, wobei die personen-spezifischen Daten und die inhaltlichen Daten (z.B. medizinische Daten)

19 Siehe Kopal (2007), S. 1-2.

20 Reng et. al. (2006), S. 49 f.

an jeweils anderen Standorten durch andere LZA-Services gespeichert werden (diese Service-Variante spielt im Hinblick auf die LZA von Forschungsdaten aus dem biomedizinischen Bereich eine entscheidende Rolle für die Akzeptanz der LZA)

- Installation des Systems bzw. von Zugangskomponenten für Remote Access vor Ort
- Beratungsleistungen, z.B. zum Geschäftsmodell, zum Einsatz der Archivsoftware, zur Speicherverwaltung etc.
- Support und Schulungen
- Weiterentwicklung des Archivsystems bzw. von gewünschten Komponenten

Handelt es sich bei dem Dienstleister, der von einer Archivinstitution in Anspruch genommen wird, um einen reinen Datenhost, könnten folgende Dienstleistungen relevant werden:

- Hardwarehosting und -betreuung
- Hosting und Betreuung von Standardsoftware
- Sichere Datenhaltung (z.B. durch Mehrfachbackups)
- Zurverfügungstellung von Datenkopien bei Datenverlusten seitens der Abliefererinstitution
- Notfall- und Katastrophenmanagement
- Beratungsleistungen, z.B. zur Speicherverwaltung

Gerade im Hinblick auf ein Commitment sind ebenso Service-Levels bezüglich der Aufbewahrungsdauer sinnvoll. Ebenso können mit derartigen Service-Levels die Wünsche von Nutzern feingranularer adressiert werden:

- Aufbewahrung bis zu 5 Jahre als Backup-Lösung
- Aufbewahrung bis zu 10 Jahre zur Realisierung guter wissenschaftlicher Praxis
- Aufbewahrung bis zu 30 Jahre zur Erfüllung gesetzlicher Anforderungen und langfristiger Speicherung
- Aufbewahrung für mehr als 30 Jahre als „richtige“ Langzeitarchivierung

Entsprechend einer Kostenkalkulation muss dann jeder LZA-Dienstleister Preise für die einzelnen Service-Levels definieren, die zum einen die eigenen Vollkosten decken und zum anderen die nachhaltige Entwicklung der LZA – im Sinne der Preservation – ermöglichen. Letzteres bezieht sich insbesondere auf

die Entwicklungskosten für technische Maßnahmen durch einen LZA-Dienstleister und auf Kostensteigerungen, wie z.B. durch höhere Personalkosten in Folge neuer Tarifverträge.

Jede Institution muss die eigenen Möglichkeiten bezüglich des Angebots von LZA-Services sorgfältig evaluieren. Hat sie einmal damit begonnen, insbesondere für Dritte solches Services anzubieten, werden dadurch Verpflichtungen eingegangen, die durch künftige technische Entwicklungen ggf. nur erschwert eingehalten werden können. Daher kann es ratsam sein, LZA-Services koordiniert oder kooperativ mit anderen Einrichtungen anzubieten bzw. zu nutzen. Lassen sich die Dienstleistungen von externen Anbietern nutzen und ist dies auch unter Kostengesichtspunkten der wirtschaftlichere Weg, kann es auch für Teile des digitalen Bestands einer Einrichtung sinnvoll sein, diese durch den Service eines solchen Anbieters archivieren zu lassen. Eine andere Möglichkeit bietet sich in dem beschriebenen Hardware-Hosting bzw. Storage-Betrieb durch einen ausgewiesenen Dienstleister.

Quellen und Literatur

- Borghoff, Uwe M. et al. (Hrsg.) (2003): *Long-Term Preservation of Digital Documents* : Principles and Practices. Heidelberg [u.a.] : Springer
- Kopal (2007): *kopal: Ein Service für die Langzeitarchivierung digitaler Informationen*. Projekt kopal (Kooperativer Aufbau eines Langzeitarchivs digitaler Informationen), 2007 (s. http://kopal.langzeitarchivierung.de/downloads/kopal_Services_2007.pdf)
- Lutterbeck, Bernd / Bärwolff, Matthias / Gehring, Robert A. (Hrsg.) (2007): *Open Source Jahrbuch 2007* : Zwischen freier Software und Gesellschaftsmodell. Berlin : Lehmanns Media, 2007 (s. <http://www.opensourcejahrbuch.de/download/jb2007>)
- Reng, C. et al.. (2006): *Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin*. Berlin : Medizinisch-Wissenschaftliche Verlagsgesellschaft.