

Reinhard Altenhöner, Stefan Hein, Karlheinz Schmitt

DP4lib – Digital Preservation for Libraries

Eine Herausforderung für Gedächtnisorganisationen ist die konsequente Ausrichtung aller Dienste auf die Bedürfnisse ihrer internen und externen Kunden. Hierbei stehen diese Organisationen und speziell deren IT-Abteilungen nicht nur vor der Aufgabe, ihren Betrieb, und insbesondere die von ihnen betreute Infrastruktur so performant und kostengünstig wie möglich zu betreiben, sondern auch die bereitgestellten Dienste anwenderbezogen zu verrechnen. Eine optimale Zuordnung der Infrastruktur und Prozesse auf die angebotenen Dienste muss somit stets das Ziel einer wettbewerbsfähigen Organisation sein. Zur Umsetzung sind innerhalb von IT-Bereichen oftmals zwei Herausforderungen zu meistern: Zum einen ist die etablierte Infrastruktur historisch gewachsen und somit nicht ohne weiteres kompatibel zu neuen Aufgaben. Zum anderen liegt der Schwerpunkt oftmals stark auf der Technik, wohingegen organisatorische und serviceorientierte Gesichtspunkte nachrangig behandelt werden. Beide Aspekte und deren Auswirkungen führen fast zwangsläufig zu Defiziten bei der Erfüllung der oben genannten Aufgabe. Diese Konstellation, eine technik- bzw. innovationsgetriebene Entwicklung einerseits und eine nur sehr begrenzt ausgeprägte Servicestruktur andererseits bilden die Ausgangspunkte für das Projekt.

Das Projekt DP4lib hat die Aufgabe, die im Rahmen des direkten Vorgängerprojekts kopal¹⁾ – Kooperativer Aufbau eines Langzeitarchivs digitaler Informationen – entstandenen Funktionalitäten zur digitalen Langzeitarchivierung elektronischer Publikationen sowohl organisatorisch als auch technisch den oben genannten Zielen anzunähern. Unter besonderer Berücksichtigung weiterer interner und externer Kundenanforderungen soll hierbei ein möglichst

umfassender integrierter Dienst konzipiert werden, welcher darüber hinaus prototypisch verfügbar gemacht werden soll.

DP4lib nahm im Dezember vergangenen Jahres seine Arbeit auf und wird dabei für die Dauer von zwei Jahren von der Deutschen Forschungsgemeinschaft (DFG) gefördert. Wie bereits im vorangegangenen Projekt kopal setzen die Niedersächsische Staats- und Universitätsbibliothek Göttingen (SUB) und die Deutsche Nationalbibliothek (DNB) auch in diesem Projekt ihre erfolgreiche Zusammenarbeit fort, wobei die DNB, organisatorisch vertreten durch die Abteilung Informationstechnik, für die Gesamtprojektleitung verantwortlich ist.

DNB und SUB übernehmen innerhalb dieses Projektes die Konzeptions- und Entwicklungsaufgaben und werden als Service Provider die organisatorischen, betriebswirtschaftlichen sowie technischen Anforderungen für die Bereitstellung der Dienstleistung: „Langzeitarchivierung elektronischer Publikationen“ evaluieren und in gemeinsamer Abwägung mit weiteren externen Partnern prototypisch umsetzen.

Externe Partner in diesem Projekt sind:

- das Bibliotheksservice-Zentrum Baden-Württemberg (BSZ), Konstanz,
- das Deutsche Institut für Internationale Pädagogische Forschung (DIPF), Frankfurt am Main,
- die Verbundzentrale des Gemeinsamen Bibliotheksverbundes (VZG), Göttingen,
- die Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden (SLUB),
- die Technische Informationsbibliothek (TIB), Hannover und

Projektorganisation und -umsetzung

Projektziele

□ die Thüringer Universitäts- und Landesbibliothek (ThULB), Jena.

Aufgabe von DP4lib ist es, gemeinsam mit allen Partnern die Umsetzbarkeit eines Dienstleistungsangebotes zur Langzeitarchivierung zu untersuchen.²⁾

Die Arbeiten in DP4lib profitieren von vielfältigen Vorarbeiten und Erfahrungen der beteiligten Partner. Zentraler Ausgangspunkt wird die Nutzung von Ergebnissen und Komponenten aus dem Langzeitarchivierungsprojekt kopal sein. Auf der Grundlage des Open Archival Information System (OAIS)³⁾ – ISO-Standards, dem anerkannten Referenzmodell für digitale Langzeitarchivierungssysteme, ist in Deutschland hierdurch erstmalig ein kooperativ betriebenes Langzeitarchivierungssystem entstanden. DP4lib wird mit der Öffnung des Systems für weitere externe Partner als Kunden die kooperative Ausrichtung des Projekts kopal weiterverfolgen.

Für die Konzeption der Dienstleistung werden weiterhin auch Ergebnisse des Kompetenznetzwerks nestor herangezogen. Zu nennen ist beispielsweise der „Kriterienkatalog vertrauenswürdige digitale Langzeitarchive Version 2“⁴⁾, der die Grundlage für den kürzlich veröffentlichten Norm-Entwurf DIN 31644 (Information und Dokumentation – Kriterien für vertrauenswürdige digitale Langzeitarchive) bildet. Es wird innerhalb von DP4lib zu prüfen sein, inwieweit dieser Katalog zur Qualitätssicherung der Dienste und zum Aufbau eines vertrauenswürdigen Langzeitarchivs für externe Kunden herangezogen und schrittweise umgesetzt werden kann.

Um das ehrgeizige Ziel des Aufbaus eines Service-Angebotes für Langzeitarchivierung zu erreichen, hat DP4lib eine Reihe von Einzelaufgaben zu erfüllen. Speziell zu nennen sind hier:

□ Die vollständige Ermittlung aller funktionalen und nicht-funktionalen Anforderungen, die von externen Kunden an die Langzeitarchivierung gestellt werden. Diese stellt die Grundvoraussetzung für eine wirtschaftliche und optimale,

qualitative und quantitative Angebotserstellung dar.

□ Die Identifikation, Spezifizierung und Anpassung sämtlicher notwendiger Geschäftsprozesse inklusive aller administrativen Funktionen. Speziell der Aufbau eines lückenlosen Berichtswesens und eines funktionsfähigen Controllings über die durchgeführten Aktivitäten ist eine Aufgabe, die von DP4lib an dieser Stelle erwartet wird.

□ Der Aufbau eines tragfähigen Betriebs- und Kostenmodells für die angebotenen Services. Letztendlich bedeuten der Aufbau, der Betrieb und endlich auch die Einstellung eines Service die Etablierung eines sozio-technischen Systems aus Hardware, Software, menschlicher Arbeitsleistung und vieles mehr. Den Aufbau und den Betrieb eines solchen Systemumfeldes muss ein Anbieter zumindest kostendeckend, wenn nicht sogar gewinnorientiert leisten. Hierbei sind alle Kostenfaktoren zu identifizieren und den verursachenden Diensten zuzuordnen, um eine vollständige Abrechenbarkeit zu erhalten.

□ Die Gewährleistung des Erhalts der Dateintegrität über einen unbegrenzten Zeitraum. Hierbei sind Verfahren und Prozesse zu beschreiben, die die Authentizität und die Integrität aller im Auftrag archivierten elektronischen Dokumente dauerhaft sicherstellen.

□ Die Ermittlung von Strategien, die die langfristige Sicherstellung der technischen Interpretierbarkeit der archivierten elektronischen Publikationen gewährleisten. Als mögliche Strategie wird vorerst die Dateiformatmigration umgesetzt. Zur deren Unterstützung ist die Möglichkeit der Normalisierung des sehr heterogenen Archivbestandes auf wenige Dateiformate zu prüfen. Hierdurch würde zwar eine erste Migration vor der Archivierung stattfinden, aber der so erhaltene Archivbestand würde nur aus wenigen Dateiformaten bestehen und so die komplexe Aufgabe des Preservation Planning (also der systematischen Vorplanung, Beobachtung und Aktion in einem Archiv) auf wenige Formate reduzieren.

Arbeitsgrundlagen

Erfüllung von Einzelaufgaben

□ Die Zusammenarbeit zwischen den Abnehmern einzelner Dienste und dem Dienstleister bedarf nicht nur der Abstimmung von Diensten, Schnittstellen und Workflows, sondern ebenso einer rechtlich-vertraglichen Regelung. Aus diesem Grunde werden in einem weiteren Arbeitsschritt die Erarbeitung von Service Level Agreements und die Definition entsprechender Service Levels evaluiert. Die Sicherung des Bitstreams kann beispielsweise durchgeführt werden, indem zwei, drei oder mehr Speichermedien ggf. parallel verwendet werden. Abhängig von der geforderten Medienanzahl werden sowohl der Service Level als auch die Kosten für diesen so vereinbarten Dienst kundenorientiert vertraglich festgelegt. Welche Erkenntnisse versprechen sich die Projektpartner aus dem Aufbau eines betriebsfähigen Langzeitarchivierungsdienstes für externe Kunden?

Zum einen ist die möglichst genaue und vollständige Kenntnis der Bedürfnisse von Institutionen im Sektor der öffentlichen finanzierten Informationsinfrastruktur ein wesentliches Ziel. Diese Erkenntnisse sollen nicht nur verwendet werden, um Services anforderungsgerecht zu designen. Dieses Wissen soll gleichzeitig so abstrahiert und dokumentiert werden, dass daraus Richtlinien und Handlungsempfehlungen abgeleitet werden können, die Institutionen einen Einstieg in die eigene Langzeitarchivierung erleichtern können. Die Ergebnisse der durchzuführenden Erhebungen und Interviews sollen u. a. in einen Fragenkatalog einfließen, der vollständig alle Rahmenbedingungen und Informationen bzgl. der jeweils vorliegenden Nutzungsszenarien eines neuen Kunden ermittelt. Seine Verwendung soll anderen Institutionen eine Hilfestellung geben, um einen vollständigen ersten Überblick über durchzuführende Maßnahmen zur Etablierung eines Langzeitarchivierungsdienstes zu erhalten.

Ableitung von Richtlinien und Handlungsempfehlung für die eigene Langzeitarchivierung

Ein weiterer wesentlicher Aspekt innerhalb von DP4lib ist der Versuch, ein funktionsfähiges Betriebs- und Kostenmodell für die Langzeitarchivierung aufzubauen. Derzeit existieren nur wenige Ansätze hierzu. Als Beispiel sei an dieser Stelle das LIFE-Projekt (Lifecycle Information for E-Literature) erwähnt, welches sich aktuell in seiner dritten Entwicklungsphase befindet.⁵⁾

Innerhalb von DP4lib ist es das Ziel, ein über den Zeitraum der Projektförderung hinaus tragfähiges Betriebs- und Kostenmodell aufzubauen, um die tatsächlichen Kosten der Service-Bereitstellung zu kennen und professionell zu verwalten. Zu diesem Zweck ist es im privatwirtschaftlichen Sektor üblich, IT-Rechnungswesen- und Budgetierungsprozesse und, sofern eine Weitergabe der Kosten an den Dienstnehmer gewünscht ist, auch Leistungsverrechnungsprozesse zu erstellen. Dieses Vorgehen ist auch im öffentlichen Umfeld hilfreich.

Betriebs- und Kostenmodell für die Langzeitarchivierung

Hierzu werden innerhalb von DP4lib zunächst Grundlagenarbeiten erforderlich sein, wozu prinzipiell die Erstellung eines Kostenmodells zu rechnen ist. Eine zu erwägende Herangehensweise ist die Erarbeitung einer Rahmenstruktur, in der alle bekannten Kosten erfasst und zu spezifischen Kunden, Aktivitäten oder anderen Kategorien zugeordnet werden. Solche Kostenmodelle beruhen meistens auf der Berechnung der Kosten für jeden einzelnen Kunden und bilden die Basis für die Einführung eines Leistungsverrechnungssystems. Es sind aber auch serviceorientierte oder standortorientierte Modelle denkbar.

Erarbeitung einer Rahmenstruktur

Auf dieser Grundlage können erste fundierte Entscheidungen über die Kosten und die Wirtschaftlichkeit digitaler Langzeitarchivierung getroffen werden.

Anmerkungen

1 Wollschläger, Thomas: Kopal goes live. In: Dialog mit Bibliotheken, 19 (2007) 2, S. 17 – 22.

2 Weitere Informationen zum DP4lib-Projekt finden Sie auf der offiziellen Webseite: <<http://dp4lib.langzeitarchivierung.de>>

3 OAIS - Reference Model for an Open Archival Information System. OAIS hat den Status eines ISO-Standards erreicht (ISO-Standard 14721: 2003). Öffentlich verfügbar unter:

<<http://public.ccsds.org/publications/archive/650x0b1.pdf>>

4 nector: Kriterienkatalog Vertrauenswürdige digitale Langzeitarchive Vers. 2 2008; siehe:

<<http://www.langzeitarchivierung.de/publikationen/expertisen/expertisen.htm>>

5 LIFE2-Projekt-Abschlussbericht: <<http://eprints.ucl.ac.uk/11758/1/11758.pdf>>