



Zentralbibliothek

Langzeitarchivierung am Beispiel LOCKSS

***Vergleich internationaler und nationaler Programme und
Initiativen zur Langzeitarchivierung und Test der „LOCKSS“
Software in der ZB***

Sonja Hendriks

Herausgeber und Vertrieb: Forschungszentrum Jülich GmbH
Zentralbibliothek, Verlag
D-52425 Jülich
Telefon (02461) 615368 - Telefax (02461) 61-6103
e-mail: zb-publikation@fz-juelich.de
Internet: <http://www.fz-juelich.de/zb/verlag/>

Redaktion: Zentralbibliothek

Copyright: Forschungszentrum Jülich 2007

Vollständig veröffentlicht im Internet

Persistent Identifier: [urn:nbn:de:0001-00404](https://nbn-resolving.org/urn:nbn:de:0001-00404)

Resolving URL: <http://nbn-resolving.de/urn/resolver.pl?urn=urn:nbn:de:0001-00404>

Alle Rechte vorbehalten. Kein Teil des Werks darf in irgendeiner Form ohne schriftliche Genehmigung des Verlags reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

INHALTSVERZEICHNIS

| | |
|---|------------------|
| <u>ABBILDUNGSVERZEICHNIS.....</u> | <u>3</u> |
| <u>LANGZEITARCHIVIERUNG ELEKTRONISCHER ZEITSCHRIFTEN.....</u> | <u>4</u> |
| <u>GEFAHR DES DATENVERLUSTES.....</u> | <u>4</u> |
| <u>GRÜNDE ZUR LANGZEITARCHIVIERUNG.....</u> | <u>4</u> |
| <u>DIE PROJEKTE.....</u> | <u>4</u> |
| <u>GRUNDLAGEN DER LANGZEITARCHIVIERUNG.....</u> | <u>5</u> |
| <u>MIGRATION.....</u> | <u>5</u> |
| <u>EMULATION.....</u> | <u>5</u> |
| <u>OAIS REFERENZ MODELL.....</u> | <u>5</u> |
| <u>PROGRAMME UND INITIATIVEN ZUR LANGZEITARCHIVIERUNG.....</u> | <u>6</u> |
| <u>NESTOR UND KOPAL.....</u> | <u>6</u> |
| <u>DISKUSSION – KOPAL.....</u> | <u>8</u> |
| <u>PORTICO.....</u> | <u>8</u> |
| <u>DISKUSSION – PORTICO.....</u> | <u>9</u> |
| <u>CLOCKSS.....</u> | <u>9</u> |
| <u>LOCKSS.....</u> | <u>10</u> |
| <u>FUNKTIONSWEISE VON LOCKSS.....</u> | <u>10</u> |
| <u>LOCKSS AUS SICHT DES NUTZERS.....</u> | <u>12</u> |
| <u>TEILNAHME AM PROJEKT LOCKSS.....</u> | <u>13</u> |
| <u>DISKUSSION – LOCKSS.....</u> | <u>14</u> |
| <u>LOCKSS-TEST IN DER ZENTRALBIBLIOTHEK.....</u> | <u>14</u> |
| <u>FAZIT.....</u> | <u>15</u> |
| <u>ANHANG A – LISTE DER TEILNEHMENDEN VERLAGE BEI LOCKSS.....</u> | <u>16</u> |
| <u>ANHANG 2 – LISTE DER TEILNEHMENDEN VERLAGE BEI PORTICO.....</u> | <u>17</u> |
| <u>ANHANG 3 – LISTE DER TEILNEHMENDEN VERLAGE BEI CLOCKSS.....</u> | <u>18</u> |

Abbildungsverzeichnis

Abb. 1 OAS-Modell

Abb. 2 KOPAL Workflow

Abb. 3 LOCKSS "How its work collection"

Abb. 4 LOCKSS "How its work auditing"

Abb. 5 Ablauf einer Benutzerabfrage mit LOCKSS

Langzeitarchivierung elektronischer Zeitschriften

Vergleich internationaler und nationaler Programme und Initiativen zur Langzeitarchivierung und Test der „LOCKSS“ Software in der ZB.

Die Bedeutung der Langzeitarchivierung rückt mit dem Schritt zur „e-only“ Strategie der Zentralbibliothek besonders in den Vordergrund. Die Entscheidung Zeitschriften so weit möglich nur noch in elektronischer Form zu abonnieren hat zur Folge, dass die Informationen und Zeitschriften nicht mehr physisch vor Ort verfügbar sind und im Magazin nicht mehr archiviert werden können. Um die nicht unerheblichen Investitionen zu sichern und die Informationsversorgung der Wissenschaftler im Forschungszentrum Jülich dauerhaft gewährleisten zu können muss nach digitalen Alternativen zur physikalischen Archivierung gesucht werden.

Gefahr des Datenverlustes

Die Gefahr des Datenverlustes entsteht dadurch, dass die Informationen in Zukunft in Formaten vorliegen können, die von modernen Computern nicht mehr interpretiert und gelesen werden können. Die Zeitschriften sind bei Verlagen auf Servern gespeichert. Server können ausfallen, Verlage können sich auflösen. Die Folge wäre immer, dass kein Zugriff mehr besteht und die Informationen nicht mehr genutzt werden können. Die Gefahr ist natürlich bei kleineren Verlagen größer, als bei großen Verlagen, bei denen die Zeitschriften teilweise durch die jeweiligen Nationalbibliotheken (wie z.B. Elsevier in den Niederlanden) geschützt sind. Auch die Frage was passiert, wenn man eine Zeitschrift abbestellt, die alten Jahrgänge jedoch weiterhin benutzen möchte, muss beantwortet werden und die bezahlten Informationen müssen weiterhin benutzbar sein.

Gründe zur Langzeitarchivierung

Was passiert bei großen unvorhergesehen Ereignissen, so genannten „trigger events“, wenn elektronische Zeitschriften auf gewohntem Weg nicht mehr verfügbar sind? Die kontinuierliche Literaturversorgung fällt aus und es stehen keine Alternativen zur Verfügung. Diese Problematik stellt sich vor allem bei Zeitschriften über die man keinerlei rechtliche Verfügung hat und mit deren Anbietern keine Verträge abgeschlossen werden können um dauerhafte Verfügbarkeit zu gewährleisten. Der Open Access-Bereich nimmt immer mehr Bedeutung bei Zeitschriften und Artikeln ein, wie z.B. mit dem JUWEL-Server in der Zentralbibliothek. Auch der Zugriff auf Open-Access Titel muss in ein paar Jahren noch gewährleistet sein.

Die Projekte

Für diese Problematik und den gesteckten Zielen sind vier Langzeitarchivierungsprojekte verglichen und unter dem Aspekt der Funktionalität und Effektivität überprüft worden.

- a) „kopal“, ein deutsches Projekt unter Beteiligung der deutschen Nationalbibliothek

- b) „PORTICO“, ein internationales Projekt in dem Bibliotheken und Verlagen zusammen Arbeiten
- c) „CLOCKSS“, ebenfalls ein internationales Projekt unter Beteiligung von Bibliotheken und Verlagen
- d) „LOCKSS“, eine internationale Kooperationen von Bibliotheken

Es wurde entschieden letztgenanntes Projekt, „LOCKSS – Lots of Copies Keep Stuff Save“ in der Zentralbibliothek des Forschungszentrums Jülich zu testen.

Grundlagen der Langzeitarchivierung

Um zu verstehen, wie Langzeitarchivierung funktionieren soll, muss man die Erhaltungsstrategien hinter diesen Projekten verstehen. Grundsätzlich gibt es für die Archivierung und Erhaltung von Texten und Zeitschriften zwei Strategien, Emulation und Migration.

Migration

Nahezu jede Zeitschrift verwendet andere Formate, Einstellungen, Programme usw., um die Zeitschriften herzustellen und zu publizieren. Die Daten, Inhalte und Informationen der Zeitschrift kommen also in den unterschiedlichsten Variationen in die Archivierungsstelle.

Als Erstes müssen also die eingehenden Datensätze vereinheitlicht werden. Dies ist nachzuvollziehen, da es natürlich einfacher ist, die einheitlichen Datensätze in andere Formate zu überführen, als jede andere Form weiter zu verwenden um dann die Daten in das jeweilige neue Format zu übertragen.

Die Übertragung in ein neues Format wird Migration genannt.

Sobald also eine einheitliche Grundlage, ein einheitliches Format, geschaffen wurde, können sie weiter in das aktuellste Format migriert werden.

Für die Entscheidung, wann migriert werden soll, gibt es mehrere Möglichkeiten. Entweder die Migration findet in einem bestimmten zeitlichen Rhythmus statt, oder die Migration wird nach Bedarf ausgeführt. Das heißt, wenn z.B. ein archivierter Artikel aufgerufen wird und die aktuellste Version nicht mehr mit der Version kompatibel ist, mit dem der Artikel gespeichert wurde, wird der Artikel in das neue Format migriert.

Emulation

Die zweite Erhaltungsstrategie ist die Emulation. Sie ist eine einmal durchzuführende Aktion. Bei der Emulation wird die aktuelle Umgebung, auf der das Dokument läuft, mitgespeichert, so dass wenn das Dokument aufgerufen wird, gleichzeitig die Software nachgebildet und simuliert wird. Für diese Strategie müssen die Quelltexte der Software zur Verfügung stehen, was allerdings nicht immer möglich ist.

OAIS Referenz Modell

In einem Standardmodell, dem OAIS-Modell (= Reference Model for an „Open Archival Information System“ – ISO 14721), wird das prinzipielle Verfahren zur Archivierung von elektronischen Dokumenten anschaulich

verdeutlicht indem die genauen Schnittstellen zwischen Ankommen und Herausgabe der Dokumente und Objekte definiert wird.

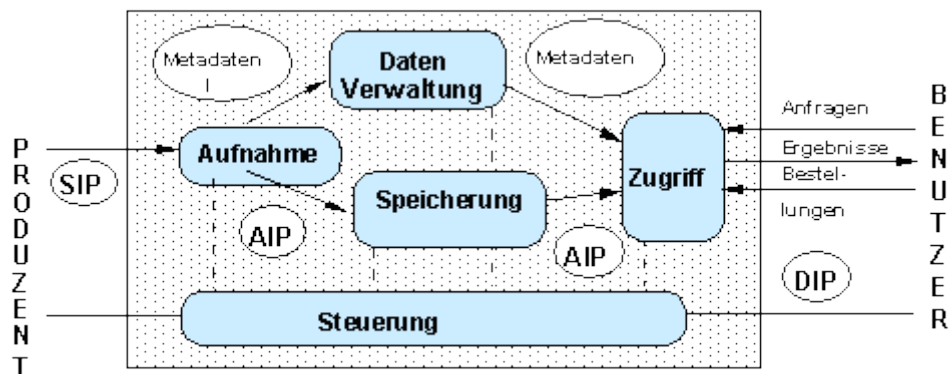


Abbildung 1 – OAIS-Modell

Die Produzenten liefern die zu archivierenden Dokumente und Objekte in ein, mit dem Archiv verabredetes standardisiertes Format, dem Submission Information Package (SIP).

Im Archivsystem selber wird jedes Dokument und Objekt in ein Archivpaket überführt (Archival Information Package, AIP). Die Funktionsmodule sind dem Datenfluss und den Arbeitsabläufen eines Archivs angepasst.

Eingangsbearbeitung, Metadatenverwaltung, Objektspeicherung, Bereitstellung und Administration. Diese Module haben das Ziel die Pakete zu speichern, zu erhalten und wiederauffindbar zu machen.

Die Bereitstellung für den Benutzer erfolgt über das Dissemination Information Package (DIP), dem Auslieferungsbehälter.

Programme und Initiativen zur Langzeitarchivierung

Zum Thema Langzeitarchivierung gibt es in Deutschland ein Kommunikationsforum mit dem Namen „Nestor“. Über dieses Projekt laufen Diskussionen und Projekte zum Thema Langzeitarchivierung und Langzeitverfügbarkeit in Deutschland.

Nestor und Kopal

Eines der Projekte, welches über „Nestor“ läuft ist das Projekt „Kopal“. „Kopal“ ist eine Abkürzung für das Ziel des Projektes, dem Kooperativen Aufbau eines Langzeitarchivs digitaler Informationen. In diesem Projekt arbeiten Wirtschaftseinrichtungen und Bibliotheken zusammen. Projektleiter und federführend ist die Deutsche Nationalbibliothek. Partner sind die niedersächsische Staats- und Universitätsbibliothek Göttingen (SUB), die Gesellschaft für wissenschaftliche Datenverarbeitung Göttingen mbH und die Firma International Business Machines Corporation Deutschland (IBM).

Das Projekt wurde 2004 ins Leben gerufen und vom Bundesministerium für Bildung und Forschung zuerst für drei Jahre, aktuell bis August 2009 gefördert. Allgemeines Ziel ist die praktische Implementierung eines Langzeit-Archivierungssystems für digitale Publikationen.

Am Ende des Projekts soll mit der „Kopal-Solution“ eine Möglichkeit entstanden sein, weitere Institutionen in das System mit aufzunehmen und nutzen zu lassen oder eine Nachnutzung durch die eigene Installation des DIAS-Core zusammen mit den kopal-Tools zu ermöglichen.

Das kopal-Archivsystem setzt sich aus dem DIAS Core und den kopal-Tools zusammen.

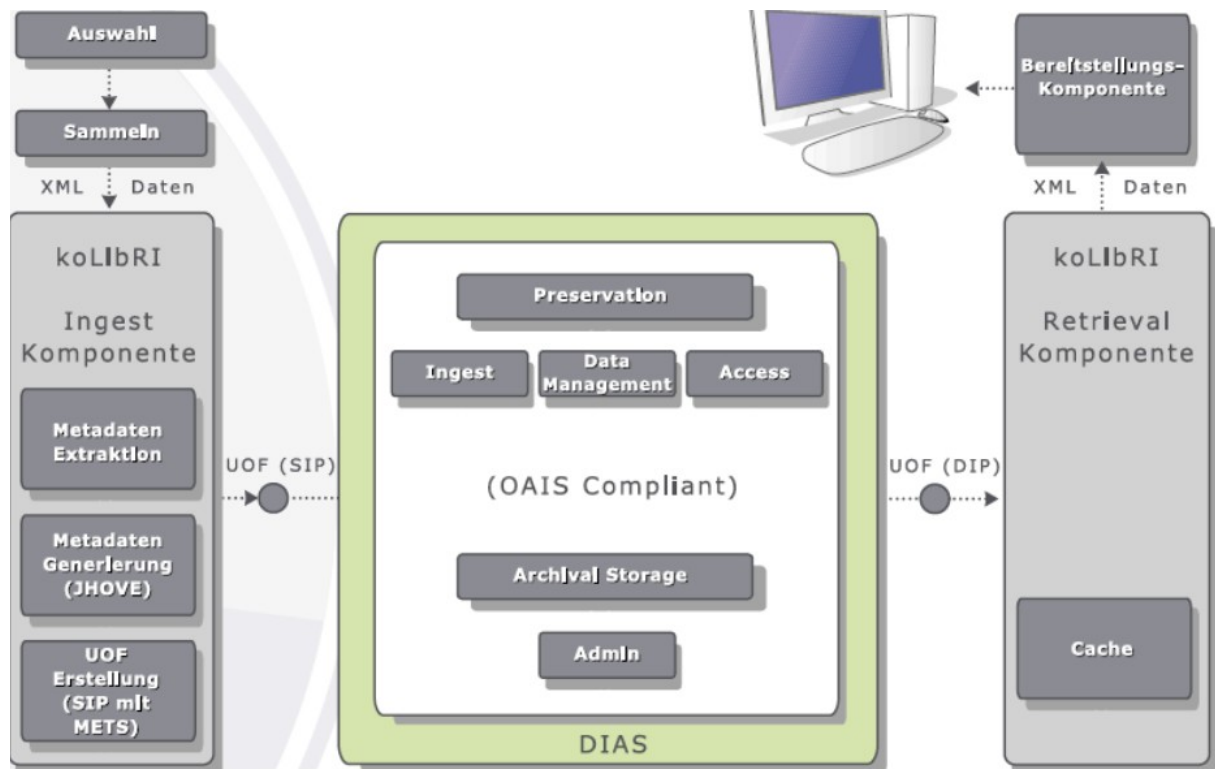


Abbildung 2 kopal - Workflow

Kern des „kopal-Archivsystem“ ist der DIAS-Core (Digital Information Archival System). Er beruht auf den Standardsoftwarekomponenten der IBM, die vor ein paar Jahren für ein Projekt der niederländischen Nationalbibliothek mit IBM entwickelt worden waren.

Für „kopal“ ist das DIAS jedoch weiterentwickelt worden. Es verwaltet die gespeicherten Daten und stellt eine Umgebung für unterschiedlichste Erhaltungsstrategien zur Verfügung. Durch präzise Import- und Exportschnittstellen wird eine Archivnutzung in verschiedensten Arbeitsabläufen unterschiedlichster Institutionen ermöglicht.

Über eine Open-Source Lizenz steht das leistungsfähige kopal-Tool „koLibRI“ (kopal Library for Retrieval and Ingest) zur Verfügung. „koLibRI“ organisiert das Einstellen und Einspielen von Archivpaketen und Objekten in das DIAS und stellt Funktionen zur Verfügung um diese abzurufen.

Anforderungen beim Datenimport (Ingest) wie automatisierter Import, Gewährleistung der Nutzbarkeit durch Dritte, Unterstützung des universellen Objektformats (UOF) usw. werden erfüllt.

Das universelle Objektformat beschreibt eine Paketstruktur inklusive Metadaten und ist gleichzeitig als Archiv und Austauschformat für Langzeitarchivierung geeignet. Es kann beliebige Dateistrukturen aufnehmen

und inhaltliche sowie technische Metadaten transportieren und bildet somit die Grundlage für Migration und Emulation.

Die Pakete werden bei „kopal“ im DIAS im Original gespeichert und wenn nötig migriert. Die älteren Versionen werden jedoch weiterhin gespeichert, so dass jeder Benutzer auch ältere bzw. die Originalversion anfordern kann.

Der Datenexport bei „koLibRI“ ist variable gestaltet, da jede Einrichtung andere Anforderungen an den Export stellt.

Diskussion – kopal

Ein Vorteil dieses Projektes ist, dass teilweise mit Open-Source Software gearbeitet wurde und daraus die Möglichkeit resultiert auf der entwickelten Grundlage ein eigenes Archivprogramm aufzubauen. Außerdem plant „kopal“, dass auch andere Bibliotheken an dem Projekt teilnehmen sollen. Insgesamt soll so auch die Zusammenarbeit der Bibliotheken untereinander gestärkt und gefördert werden.

Nachteilig an „kopal“ ist, dass die Testphase noch nicht beendet ist und eine Beteiligung an dem Projekt noch nicht möglich ist. Außerdem ist die Kostenstruktur noch völlig unklar.

Spezielle Vorteile, warum die Zentralbibliothek sich an diesem Projekt beteiligen sollte, gibt es im Moment nicht. Es handelt sich hierbei um ein nationales Projekt, weshalb viele deutsche wissenschaftliche Bibliotheken auf die Ergebnisse warten um sich an diesem Projekt zu beteiligen.

Ansonsten stehen bei „kopal“ zu viele Funktionen für die Ansprüche der ZB zur Verfügung. Es wird alles archiviert, was die deutsche Nationalbibliothek sammelt, also nicht nur Zeitschriften wie in den anderen Projekten. Die speziellen Ziele der ZB vor allem im Open Access Bereich werden bei diesem Projekt nicht verfolgt.

Portico

Ein zweites Projekt ist das Projekt PORTICO. „Portico's mission is to preserve scholarly literature published in an electronic form to ensure these materials remain accessible to future scholars, researchers and students“ (www.portico.org).

PORTICO ist eine not-for-profit Organisation mit dem Ziel wissenschaftliche Literatur dauerhaft zu erhalten. Das Programm ist ein zentralistisches Archiv, offen für alle peer-reviewed Zeitschriften.

Bei PORTICO werden die Quelldateien der elektronischen Zeitschriften (xml, pdf usw.) archiviert. Diese werden kurz nach der Veröffentlichung vom Verlag an PORTICO geschickt. Dort werden die Dateien in ein Archivformat konvertiert und ins Depot abgelegt. Die normierten Daten werden migriert, sobald neue Formate benötigt werden.

Durch die Speicherung der Quelldateien werden nur der intellektuelle Inhalt der Zeitschriften sowie Text, Bilder und limitierte Funktionen wie zum Beispiel interne Links gespeichert. Verlagszusätze und „look and feel“, also Aussehen und Handhabung und durch Hersteller und Konsortien standardisierte Design-Effekte, zum Beispiel Farben oder Layout, werden nicht mit archiviert.

Ziel des Projektes ist die langfristige Archivierung der Daten und nicht der kurzfristige Zugang dazu. Deswegen haben die teilnehmenden Bibliotheken auch im laufenden Betrieb keinen Zugang auf das Archiv.

Zugang gewährt PORTICO nur nach einem „trigger event“ und wenn die Zeitschriften auf dem Verlagsserver bzw. auf allen anderen Servern nicht mehr erreichbar ist.

Ein „trigger event“ ist ein unvorhergesehenes Ereignis, welches zur Folge hat, dass elektronische Zeitschriften auf gewohntem Weg nicht mehr verfügbar sind.

Beim Öffnen des Archivs haben alle Bibliotheken Zugang zu allen archivierten Zeitschriften, ungeachtet dessen, ob die Zeitschrift vorher von der Bibliothek abonniert worden war oder nicht.

Sobald das „trigger event“ behoben wurde und die Zeitschriften wieder über andere Wege zu erreichen sind, wird das Archiv wieder geschlossen.

Die Möglichkeit der Archivierung wird nur den Bibliotheken gewährt, die sich auch finanziell an dem Projekt beteiligen. Die Höhe der jährlichen Beiträge ist abhängig vom LME, Library Materials Expenditure (Ausgaben der Bibliothek, ohne Personalkosten), und liegt zwischen 1,500\$ und 24,000\$, mindestens jedoch 1% des LME.

Diskussion – Portico

Die allgemeinen Vorteile von PORTICO sind klar. Bei unvorhergesehenen großen Ereignissen, bleibt die Literaturversorgung gewährleistet und es entsteht die Sicherheit, dass die Inhalte der Zeitschrift nicht verloren gehen. Viele große Verlage beteiligen sich an dem Projekt, so dass sichergestellt ist, dass viele Zeitschriften die auch die Zentralbibliothek benutzt und abonniert hat darunter zu finden sind, so dass auch die Literaturversorgung gesichert bleibt (Siehe Anhang 2).

Nachteilig sind die jährlichen hohen Kosten und die Benutzbarkeit ausschließlich nach „trigger events“. Problembehebungen bei kurzzeitigen Ausfällen zählen nicht zu den Aufgabengebieten von PORTICO.

CLOCKSS

Ein drittes Programm heißt CLOCKSS, das für Controlled LOCKSS (Lots of Copies Keep Stuff Save) steht. Es agiert ähnlich wie PORTICO und funktioniert wie ein „Dark Archive“. Das bedeutet, dass während des normalen Betriebs kein Zugang auf die archivierten Dokumente erlaubt wird und nicht sichtbar ist, was sich im Archiv befindet.

CLOCKSS ist eine not-for-profit Partnerschaft zwischen Verlagen und Bibliotheken. Die Initiative will in zwei Jahren ein funktionierendes Archivsystem aufbauen mit dem Ziel der Entwicklung eines dezentralisierten, globalen Archivs mit Sicherstellung eines kontinuierlichen Zugangs zu den archivierten Inhalten im Falle eines „trigger events“.

Vorteil dieses Projekt ist, dass es kostenlos ist. Die Anmeldung erfolgt unverbindlich, weshalb auch hier in dieser Initiative viele große Verlage teilnehmen.

Das Projekt steht erst am Anfang und wie es sich im Hinblick auf die Kostenstruktur entwickelt ist offen. Wahrscheinlich ist jedoch, dass nach Ende

der Förderung durch die Library of Congress, Beiträge verlangt werden, ähnlich wie es bei LOCKSS auch ist. Genauso ungewiss ist die Frage, wer sich dann letztendlich verbindlich an diesem Projekt beteiligt (Siehe Anlage 3). Nachteilig ist, wie bei PORTICO, dass keine Problembehandlung bei kurzzeitigen Ausfällen erfolgt.

LOCKSS

Das vierte und letzte Programm ist das Projekt LOCKSS (Lots of Copies Keep Stuff Save). Je mehr Kopien es von den Daten gibt, desto sicherer ist der Erhalt in der Zukunft. Die LOCKSS Initiative geht davon aus, dass keine einzelne Einrichtung Langzeitarchivierung gewährleisten kann. Deshalb besteht dieses Projekt aus einer Kooperation von mehreren Bibliotheken.

Initiiert und weitergeführt wurde das Projekt von den Stanford University Libraries. Die LOCKSS Software ist Open Source und kann kostenlos verwendet werden.

Funktionsweise von LOCKSS

Die Software wandelt einen einfachen, kostengünstigen Computer in einen modernen Archivierungs-PC um. Neben dem LOCKSS Programm kann allerdings kein anderes Programm auf dem Rechner laufen.

Der Archivierungs-PC auch LOCKSS Box genannt, hat vier grundsätzliche Funktionen. Er sammelt neue publizierte Inhalte aus relevanten E-Journals mit Hilfe eines Web-Crawlers und vergleicht regelmäßig die vorhandenen Inhalte mit anderen LOCKSS Boxen. Außerdem agiert er wie ein Proxy-Server und unterstützt ein administratives web-basiertes Interface, welches den Bibliotheken erlaubt, neue Journals für die Archivierung aufzunehmen, archivierte Journals mit ihrem Status anzuzeigen und den Zugang zum Programm zu kontrollieren.

Collection via Web Crawler

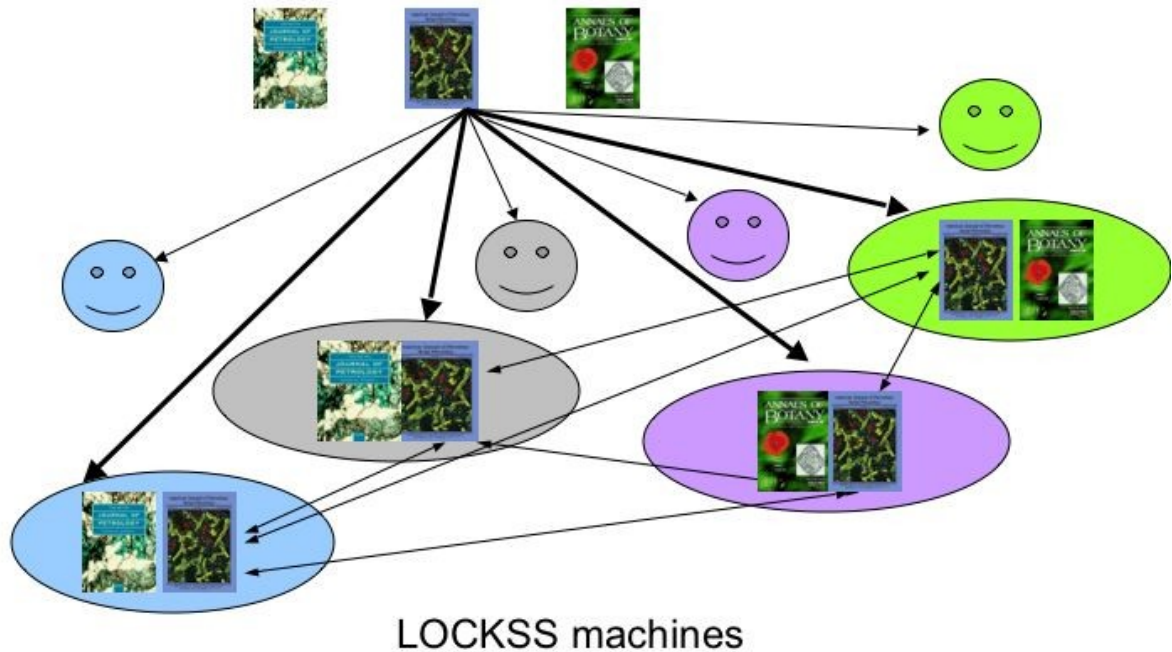


Abbildung 3, LOCKSS- "How its work collection"

Bevor LOCKSS eine Zeitschrift sammeln und archivieren kann, muss der Verlag seine Zustimmung an das LOCKSS System geben. Diese Zustimmung erfolgt durch hinzufügen einer Seite auf den Webseiten des Verlags, auf der dann die Zustimmungserklärung und der Link der Originalausgabe steht.

Die LOCKSS Box muss wissen, wo er die Seite finden kann und wie weit er dem Link folgen muss, um an sein Ziel zu gelangen. Hätte er diese Information nicht, wäre die Suche nicht eingeschränkt und der Crawler würde versuchen das gesamte Netz zu durchsuchen und zu erfassen. Die ganzen Informationen werden in einer Konfigurationsdatei, einem Publisher Plugin definiert.

Bevor die richtige Nutzung beginnt, werden mit Hilfe eines Kontrollprogramms diese Verlinkungen ausgetestet. Diese Kontrollprüfungen werden regelmäßig wiederholt. Die Ergebnisse der Prüfungen werden wiederum in eine Datei geschrieben und der Zugangsseite hinzugefügt.

Preserve and audit content integrity

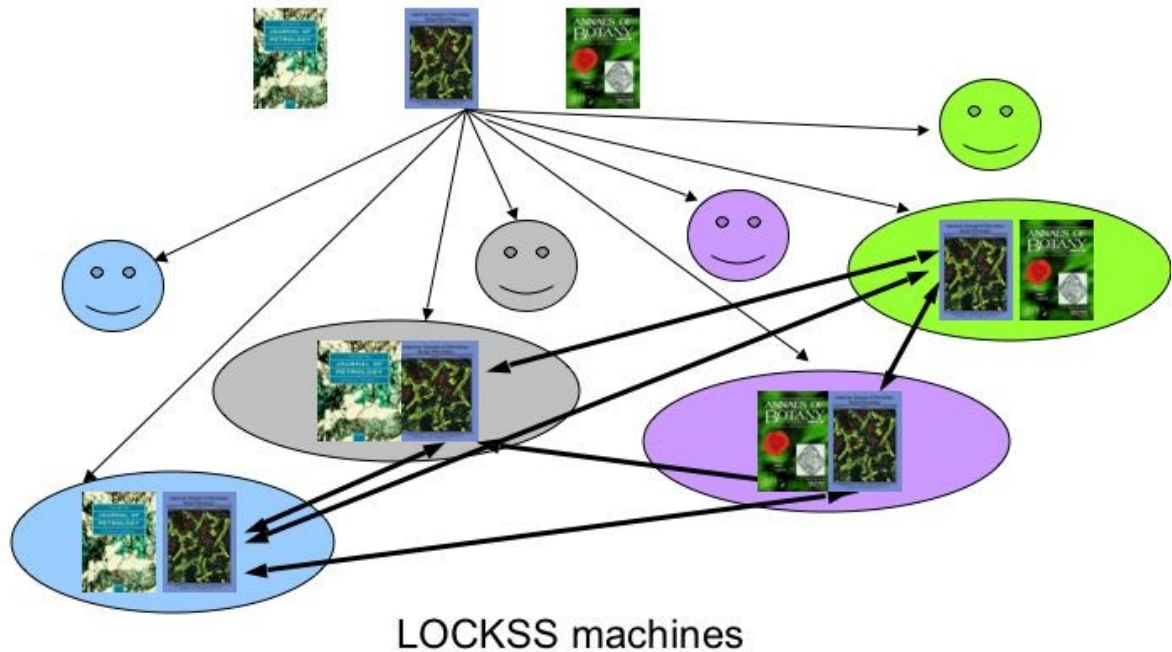


Abbildung 4 LOCKSS - "How its work auditing"

Sobald die Zeitschriften in das LOCKSS System aufgenommen wurden, überprüfen die LOCKSS Boxen untereinander regelmäßig die Informationen, die sie archivieren. Sollten Inhalte in einer LOCKSS Box beschädigt oder unvollständig sein, kann die Box mit den beschädigten Inhalten für diesen Teil nicht mehr aufgerufen werden. Über die anderen Boxen werden die beschädigten Informationen repariert, so dass diese dann auch wieder von der eigenen LOCKSS Box aufgerufen werden können.

Die automatische Zusammenarbeit der Boxen untereinander vermeidet individuelle Arbeits- und Aktualisierungsschritte jeder einzelnen Box.

So wird sichergestellt, dass jeder Leser auch die richtigen und aktuellsten Informationen zur Verfügung gestellt bekommt.

LOCKSS aus Sicht des Nutzers

Was passiert also, wenn ein Benutzer einen Artikel über z.B. Web of Science sucht und ihn über die SFX Funktion herunterladen möchte.

Ohne LOCKSS wird der Nutzer über SFX direkt auf die Verlagsseite geführt. Dort kann der Nutzer dann den Artikel lesen und gegebenenfalls herunterladen.

Wenn LOCKSS eingebunden ist, entscheidet die LOCKSS Software, welche Version genommen wird. Durch eine Abfrage von LOCKSS an den Verlag wird überprüft, ob die Version auf der LOCKSS Box die Aktuellste ist. Sollte dies der Fall sein, wird die Version vom LOCKSS Speicher genommen, andernfalls wird die aktuellere Version von den Verlagsseiten genommen. Gibt der Verlag

keine Rückmeldung wird automatisch die Version aus der LOCKSS Box genommen.

Es wird also sichergestellt, dass der Nutzer die aktuellste Version zur Verfügung gestellt bekommt. Der Verlag kann durch die Art der Anfrage von der LOCKSS Box unterscheiden, ob es sich um eine Anfrage eines Benutzers oder um eine regelmäßig stattfindende Überprüfung des Inhalts handelt, so dass die Statistiken, die der Verlag erstellt davon nicht beeinflusst werden.

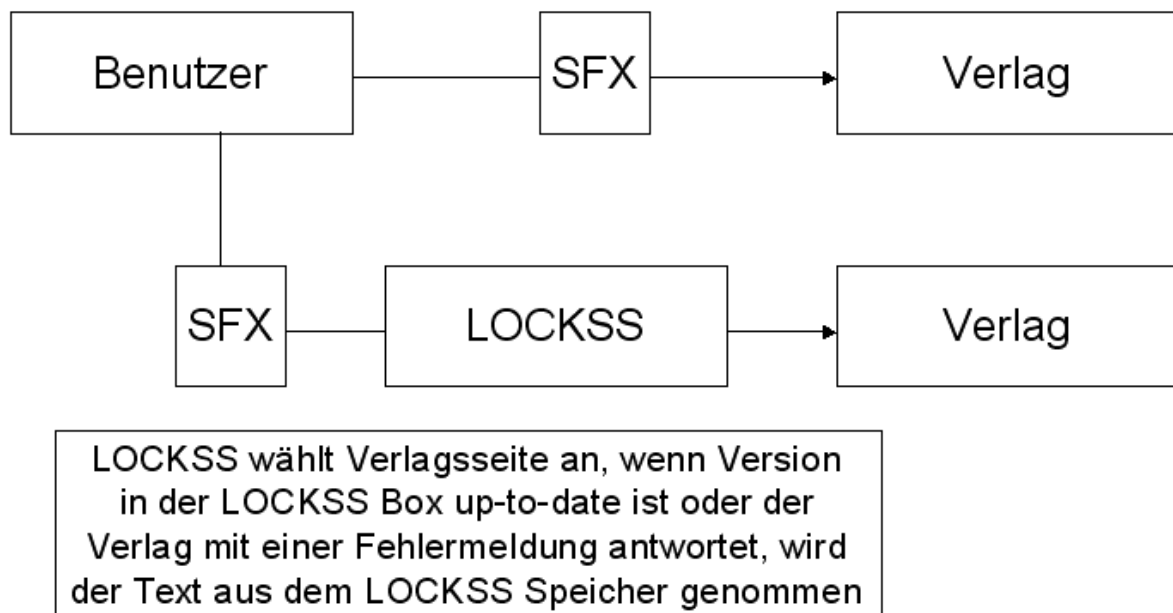


Abbildung 5 - Ablauf einer Benutzerabfrage mit LOCKSS

Teilnahme am Projekt LOCKSS

Es gibt zwei Möglichkeiten mit LOCKSS zu archivieren.

Als LOCKSS Member ist die Teilnahme am Projekt kostenlos. Abgesehen von der Administration dem Rechner und der Pflege, die aber bei LOCKSS immer eine Rolle spielen. Einige Zeitschriften sind kostenlos abrufbar und speicherbar, sofern sie Open Access sind oder die Bibliothek diese Zeitschrift abonniert hat. Im Moment ist keine Zeitschrift als kostenloses LOCKSS Mitglied zu bekommen, die die Bibliothek auch abonniert hat.

Alle anderen im LOCKSS Programm befindlichen Zeitschriften sind nur als LOCKSS Alliance Member zugänglich. Für die Zentralbibliothek würden das 11 Titel bedeuten, die im Moment abonniert und in LOCKSS aufgenommen sind (Stand 16.11.2006).

Um weitere Titel in das LOCKSS System aufgenommen zu bekommen müssen Publisher Plugins geschrieben werden. Dies wird von der Stanford University übernommen, wenn die Bibliothek als Alliance Member am Programm teilnimmt. Andernfalls muss die Bibliothek selber das Publisher Plugin für die Zeitschrift schreiben, nach Stanford senden wo es überprüft und beim nächsten Update der Software hinzugefügt wird.

Diskussion – LOCKSS

Für beide Teilnahmemöglichkeiten bleiben die Funktionen wie Fehlerbehebung und Archivierung bestehen. Es werden auch regelmäßig neue Zeitschriften in das LOCKSS Programm aufgenommen. In der Regel jedoch nur für Alliance Member und nur in seltenen Fällen für die LOCKSS Member.

Ein weiterer positiver Aspekt ist, dass durch die Open Source Software, die Möglichkeit besteht ein eigenes LOCKSS Netzwerk aufzubauen. Dies wäre überlegenswert, vor allem für die eigenen Publikationen. Die einzelnen Bibliotheken der Helmholtz-Forschungszentren könnten sich zum Beispiel zusammenschließen und für die eigenen Publikationen so eine Ausfall- und Archivierungssicherheit schaffen. Dieses Netzwerk würde dann auch unabhängig vom LOCKSS Programm laufen.

Einzigste Voraussetzung für die Gründung eines Netzwerkes ist die Anzahl. Es müssen mindestens sechs Boxen geschaltet werden, damit eine Sicherheit und Authentizität der Inhalte gewährleistet werden kann und die Software überhaupt läuft.

Zusammengefasst lässt sich also sagen, dass das LOCKSS Programm doppelte Sicherheit bietet. Es kann sichergestellt werden, dass die Inhalte immer zur Verfügung stehen und mögliche Probleme relativ zügig behoben werden.

Zum anderen wird durch die Migration auch die Langzeitarchivierung sichergestellt. Migriert wird übrigens immer, wenn ein Inhalt in der alten Form nicht mehr aufgerufen werden kann. Also quasi Migration auf Anfrage.

LOCKSS-Test in der Zentralbibliothek

Warum also LOCKSS testen? Die vielen eben genannten Vorteile überzeugen. Dazu kommt noch, dass die Software kostenlos ist und auch die Möglichkeit besteht in das LOCKSS Netzwerk Zeitschriften einzustellen. Das heißt, wenn man nicht bereit ist die Kosten als Alliance-Member zu zahlen, können auch manuell von der Bibliothek Zeitschriften eingestellt werden.

Diese stehen dann allen angemeldeten Bibliotheken (nicht nur Alliance-Member) zur Verfügung. Ein solcher Aspekt ist vor allem sinnvoll, wenn man an die angesprochene Problematik der Open-Access Zeitschriften denkt.

Beim Test in der ZB sollten vor allem die folgenden Fragen geklärt werden:

- Aufwand für Installation und Betrieb der Software (IT Aspekte)
- Aufwand zur Pflege und Administration des Systems (Bibliothekarische Aspekte)
- Überschneidungen zwischen Bestand der ZB und LOCKSS

Auf einem virtuellen Rechner wurde das LOCKSS Programm installiert. Bei der LOCKSS Software handelt es sich um eine Software Appliance, also eine Zusammenstellung von Betriebssystem (OpenBSD basierend) und der eigentlichen LOCKSS Software. Die Installation war nach Auskunft der DV problemlos. Der genaue Festplattenbedarf konnte nicht ermittelt werden, da der Platzbedarf pro Archival Unit (AU) unterschiedlich ist und nicht

verallgemeinert werden kann. Eine Archival Unit ist eine Jahresausgabe einer Zeitschrift. Eine komplette Zeitschrift besteht aus mehreren AUs.

Titel, die per LOCKSS archiviert werden sollen und die bereits im System vorhanden sind, werden direkt aus einer Drop Down Liste ausgewählt. So können einzelne Jahrgänge (AUs) einer Zeitschrift ausgewählt werden. Es gibt keine Möglichkeit in einer Art Paket alle Jahrgänge einer Zeitschrift auszuwählen. In LOCKSS sind zwar Pakete vorhanden, diese aggregieren die Pakete aber auf einer sehr hohen Ebene (Host).

Einzelne AUs oder vordefinierte Pakete lassen sich bequem hinzufügen und löschen, weitere Automatismen, die beispielsweise zu Beginn eines neuen Jahres den vorherigen Jahrgang zur Archivierung auswählen (Abofunktion) sind nicht vorhanden.

Fazit

Insgesamt lässt sich sagen, dass der Test des Programms problemlos und ohne viel Aufwand erfolgt ist. Weitere Tests zum Beispiel das Erstellen eines Publisher Plugins von einer Open Access Zeitschrift wären möglich.

Eine Mitgliedschaft als Alliance Member wäre jedoch unnötig. Die meisten Zeitschriften, die die Zentralbibliothek abonniert hat, sind durch andere Quellen geschützt. Bei außergewöhnlichen Zeitschriften, bei Zeitschriften mit häufigen Zugangsproblemen und bei Open Access Zeitschriften wäre die Aufnahme in das LOCKSS Programm sinnvoll, damit wäre die kontinuierliche Informationsversorgung aktuell, aber auch in ein paar Jahren noch gewährleistet werden kann.

Eine Frage, die jetzt noch beantwortet werden muss ist, inwieweit die Software in die Netzwerkstruktur des Forschungszentrums eingebunden werden kann, damit auch alle Nutzer die Vorteile von LOCKSS nutzen können.

Anhang A – Liste der teilnehmenden Verlage bei LOCKSS

- African Studies Association
- American Association for Clinical Chemistry
- American Folklore Society
- American Historical Association
- American Psychiatric Publishing Inc.
- American Society for Microbiology
- American Society of Hematology
- American Society of Plant Biologists
- Australian and New Zealand Association of Medieval and Early Modern Studies
- BePress
- BioOne
- Biophysical Society
- Brookings Institution Press
- Duke University Press
- Eastern Michigan University
- Edinburgh University Press
- Gallaudet University Press
- George Mason University Press
- George Washington University Institute of Ethnographic Research
- The Hastings Center
- Histochemical Society
- Indiana University Press
- International Vladimir Nabokov Society and Davidson College
- Johns Hopkins University Press
- Kent State University Press
- Massachusetts Medical Society
- Michigan University Press
- MIT Press
- Ohio State University Press
- Oxford University Press
- Penn State University Press
- Princeton University Press
- Radiological Society of North America
- SAGE Publikation
- Open Access Titles (LOCKSS Humanities Project)
- Hosted on High Wire Press
- Hosted on Project Muse

Anhang 2 – Liste der teilnehmenden Verlage bei PORTICO

- American Anthropological Association
- American Institute of Physics
- American Mathematical Society
- American Physical Society
- American Psychological Association (APA)
- Annual Reviews
- Association of Learned and Professional Society Publishers (ALPSP)
- Berkley Electronic Press
- Elsevier, including
 - Cell Press
 - The Lancet
- Haworth Press
- Institute of Physics Publishing (IOPP)
- Johns Hopkins University Press
- John Wiley & Sons
- The Journal of Bone and Joint Surgery (JBJS®)
- Now Publishers Inc.
- NRC Research Press
- Oxford University Press
- SAGE Publications Inc.
- SIAM (Society for Industrial and Applied Mathematics)
- Symposium Journals (Oxford UK)
- Taylor & Francis Group
- UK Serials Group (UKSG)
- United Kingdom Serials Group
- University of California Press (UC Press)
- University of Chicago Press

Anhang 3 – Liste der teilnehmenden Verlage bei CLOCKSS

- American Chemical Society
- American Medical Association
- American Physiological Society
- Blackwell Publishing
- Elsevier
- Institute of Physics
- Nature Publishing Group
- Oxford University Press
- SAGE Publications
- Springer
- Taylor and Francis
- John Wiley & Sons