

Sondergeld, Ute

Kooperative Langzeitarchivierung von erziehungswissenschaftlichen Open-Access-Publikationen im Fachrepositorium pedocs. Ein Praxisbericht

Potsdam : Fachhochschule 2010, 31, X S.

urn:nbn:de:0111-dipf-37885

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Kontakt:

Deutsches Institut für
Internationale Pädagogische Forschung
Mitglied der Leibniz-Gemeinschaft
Frankfurter Forschungsbibliothek
Schloßstraße 29
D-60486 Frankfurt am Main
publikationen@dipf.de
www.dipf.de/de/bildungsinformation/ffb

Fachhochschule Potsdam
Institut für Information und Dokumentation
Kurs 2010

**Kooperative Langzeitarchivierung von erziehungswissenschaftlichen
Open-Access-Publikationen im Fachrepositorium *pedocs*
- ein Praxisbericht**

Abschlussarbeit im Rahmen der Fortbildung zur
Wissenschaftlichen Dokumentarin / Information Specialist

vorgelegt von
Ute Sondergeld

Thematischer Schwerpunkt:
Management-Grundlagen für Informationsspezialisten
Dozentin: Valentina Engelhardt

Frankfurt/Main, August 2010

Abstract

Inhalt der Arbeit ist die Vorstellung des Forschungsprojektes *LZA Pädagogik*, das in Kooperation mit der *Deutschen Nationalbibliothek* im Rahmen des Projekts *DP4lib* die Erarbeitung eines nachnutzbaren Workflows zur Langzeitarchivierung digitaler Dokumente, die im fachlichen Open-Access-Repositoryum *pedocs* veröffentlicht sind, zum Ziel hat.

Nach der Vorstellung der beteiligten Partnerprojekte werden die Strategien, Modelle, Umsetzungen und Inhalte der digitalen Langzeitarchivierung aufgezeigt. Im Mittelpunkt der Arbeit stehen der technische Workflow, der die Bearbeitung der Objekte bis hin zu ihrem Transfer an die Deutsche Nationalbibliothek beinhaltet, und besondere Aufgaben des Managements der Langzeitarchivierung an der Schnittstelle von Repositoryum und Archiv. Abschließend werden besondere Handlungsfelder für den weiteren Projektverlauf identifiziert.

Inhalt

1. Einleitung.....	1
2. Ein kooperatives Geschäftsmodell der Langzeitarchivierung digitaler Objekte: die Projekte <i>kopal/DP4lib</i> und <i>LZA Pädagogik</i>	2
2.1 Die Projekte <i>kopal</i> und <i>DP4lib</i>	2
2.2 Das Projekt <i>LZA Pädagogik</i>	3
3. Die Langzeitarchivierung digitaler Objekte.....	5
3.1 Digitale Objekte und Metadaten.....	5
3.2 Das Referenzmodell OAIS.....	7
3.3 Zur Umsetzung der Langzeitarchivierung – die ‚mittlere Lage‘.....	9
4. Das fachliche Repositorium <i>pedocs</i>	11
4.1 Zertifizierung von Repositorien.....	12
5. Langzeitarchivierung der Inhalte von <i>pedocs</i> –Workflow.....	14
5.1 Auswahl der Objekte – Pflichtablieferung an die Deutsche Nationalbibliothek...	14
5.2 Vorbereitung der Objekte und Metadaten.....	15
5.2.1 Digitalisierung und OCR.....	15
Exkurs: OCR und Authentizität.....	16
5.2.2 Technische Prüfungen.....	16
5.2.3 Anreicherung mit Metadaten.....	17
5.2.4 Deckblatt generieren.....	17
5.2.5 Publikation in <i>pedocs</i>	18
5.3 Transferprozess.....	19
5.3.1 Generierung von Transferpaketen.....	19
5.3.2 Transfer und Weiterverarbeitung.....	19
6. Management der Langzeitarchivierung.....	21
6.1 Berichtswesen.....	21
6.2 Qualitätssicherung.....	21
6.3 Planung der Langzeitarchivierung.....	23
6.4 Rechtliches.....	24
7. Ausblick.....	26
8. Literaturverzeichnis.....	29
Anhang.....	i-x
Erklärung	

Abkürzungsverzeichnis

AIP	Archival Information Package
BASE	Bielefeld Academic Search Engine
CCSDS	Consultative Committee for Space Data Systems
CRC	Cyclic Redundancy Check
DC	Dublin Core
DDC	Dewey Dezimalklassifikation
DGfE	Deutsche Gesellschaft für Erziehungswissenschaft
DIAS	Digital Information Archiving Service
DINI	Deutsche Initiative für Netzwerkinformation
DIP	Dissemination Information Package
DIPF	Deutsches Institut für Internationale Pädagogische Forschung
DNB	Deutsche Nationalbibliothek
DP4lib	Digital Preservation for Libraries
DRM	Digital Rights Management
GWDG	Gesellschaft für Wissenschaftliche Datenverarbeitung mbH Göttingen
InSPECT	Investigating Significant Properties of Electronic Content over Time
JHOVE	JSTOR/Harvard Object Validation Environment
JISC	Joint Information Systems Committee
koLibRI	kopal Library for Retrieval and Ingest
kopal	Kooperativer Aufbau eines Langzeitarchivs digitaler Informationen
LIFE	Life Cycle Information for E-Literature
LMER	Langzeitarchivierungsmetadaten für elektronische Ressourcen
LZA	Langzeitarchivierung
MD5	Message-Digest Algorithm 5

nestor	Network of Expertise in long-term STOrage of digital Resources
OAI	Open Archives Initiative
OAIS	Open Archival Information System
OCR	Optical Character Recognition
ONIX	ONline Information eXchange
PAIMAS	Producer-Archive Interface Methodology Abstract Standard
PDF	Portable Document Format
pedocs	pedagogical documents
SHA	Secure Hash Algorithm
SIP	Submission Information Package
UOF	Universelles ObjektFormat
URN	Uniform Resource Name
XML	eXtensible Markup Language

1. Einleitung

Das beständige Anwachsen elektronisch verfügbarer wissenschaftlicher Publikationen und ihre damit steigende Relevanz für Wissenschaft und Forschung erfordern Strategien zur Sicherstellung ihrer langfristigen Erhaltung und Nutzbarkeit. Bedingt durch den hohen technischen, organisatorischen und finanziellen Aufwand, den die Umsetzung der Langzeitarchivierung erfordert, und aufgrund des notwendigen Spezialwissens verschiedener Fachdisziplinen ist die Entwicklung nachnutzbarer Lösungen ein Desiderat, dessen Realisierung sich in den letzten Jahren Gedächtnisorganisationen und Initiativen weltweit zur Aufgabe gemacht haben. Dabei liegt, gerade im Hinblick auf kleinere Datenproduzenten, für die der Eigenbetrieb eines Archivierungssystems nicht in Frage kommt, die Entwicklung kooperativer Geschäftsmodelle mit verteilten Kompetenzen und Aufgabenfeldern nahe.

Im Rahmen von *DP4lib*, einem von der Deutschen Forschungsgemeinschaft geförderten Projekt, erfolgt derzeit unter Beteiligung verschiedener Partner die Erarbeitung eines solchen Modells. Das *Deutsche Institut für Internationale Pädagogische Forschung (DIPF)* nimmt mit einem eigenen Forschungsprojekt an diesem Vorhaben teil und strebt in Zusammenarbeit mit dem *DP4lib*-Partner *Deutsche Nationalbibliothek* die Entwicklung eines exemplarischen Workflows zur Langzeitarchivierung erziehungswissenschaftlicher Volltexte an; Grundlage hierfür sind die Inhalte des am *DIPF* betriebenen Open-Access-Repositoryums *pedocs*. Der Abschluss eines Vertrages über die jeweils wahrzunehmenden Aufgaben innerhalb der angestrebten gemeinschaftlichen Lösung soll am Ende des Projekts die Aufnahme eines Routinebetriebs einleiten.

Vorliegende Arbeit bietet einen Einblick in die laufende Arbeit des Projekts. Neben einer Darstellung des operativen Workflows, der die Vorbereitung von digitalen Objekten bis zu deren Lieferung an das Archivierungssystem beinhaltet, werden Aufgabenbereiche des Managements der Langzeitarchivierung dargestellt, die aus dem Konzept des kooperativen Geschäftsmodells erwachsen.

Für ihre freundliche Unterstützung möchte ich allen Mitarbeitern der Projekte *LZA Pädagogik* und *pedocs*, insbesondere Henning Hinze, Dr. Julia Kreusch, Thomas Oerder und Dr. Ute Paulokat herzlich danken.

2. Ein kooperatives Geschäftsmodell der Langzeitarchivierung digitaler Objekte: die Projekte *kopal/DP4lib* und *LZA Pädagogik*

Langzeitarchivierung im digitalen Kontext bedeutet die Entwicklung von Strategien und Maßnahmen, um kontinuierlichen Zugang zu digitalen Materialien oder zumindest der in ihnen gespeicherten Informationen zu gewährleisten. "Langzeit" meint dabei einen indefiniten Zeitraum, in dem „wesentliche, nicht vorhersehbare technologische und soziokulturelle Veränderungen eintreten; Veränderungen, die sowohl die Gestalt als auch die Nutzungssituation digitaler Ressourcen vollständig umwälzen können“.¹ Dies macht die Langzeitarchivierung digitaler Objekte zu einer komplexen Aufgabe, zu deren Planung und Realisierung die Expertise verschiedener spezialisierter Disziplinen notwendig ist. Derzeit herrschen zwei methodische Ansätze zu ihrer Umsetzung vor: Migration, welche die Konvertierung obsoleter Dateiformate in aktuelle meint, und Emulation als Nachbildung einer früheren Systemumgebung auf einer aktuellen.

2.1 Die Projekte *kopal* und *DP4lib*

Kooperativ betriebene technische und organisatorische Systeme bilden einen aussichtsreichen Ansatz, die vielgestaltigen Aufgaben der Langzeitarchivierung zu lösen. Dieser Prämisse folgend hat das Forschungsprojekt *kopal - Kooperativer Aufbau eines Langzeitarchivs digitaler Informationen*² mit der ‚kopal-Solution‘ ein technisches System zur Langzeitarchivierung entwickelt, das in drei verschiedenen Anwendungsszenarien genutzt werden kann³: im ‚kopal-Eigenbetrieb‘ installiert und betreibt eine Institution das System vollständig selbst; als ‚Mandant‘ kann sie die Daten in einem abgeschlossenen Speicherbereich (Schließfach) des Archivsystems verwalten, ohne das System technisch selbst zu betreiben. Das dritte Nutzungsmodell als ‚kopal-Teilnehmer‘ sieht vor, Daten durch einen Mandanten archivieren zu lassen. Technisch verfolgt die ‚kopal-Solution‘ die Migration als Langzeitarchivierungsstrategie und setzt dazu das Archivsystem DIAS⁴ zusammen mit der im Projektverlauf ent-

¹ Neuroth (2010), Kap.1:2. http://nestor.sub.uni-goettingen.de/handbuch/nestor-handbuch_23.pdf
vgl. auch: CCSDS (2009), p.1-11. http://ddp.nist.gov/refs/650x0p11_OAIS_pink_book.pdf; Beagrie (2008a), S. 24f. http://www.dpconline.org/component/docman/doc_download/299-digital-preservation-handbook

² Laufzeit: 2004 – 2007, gefördert vom Bundesministerium für Bildung und Forschung (BMBF); Projektpartner: Deutsche Nationalbibliothek (DNB), Niedersächsische Staats- und Universitätsbibliothek Göttingen (SUB), Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG) und IBM Deutschland GmbH; Projektseite: <http://kopal.langzeitarchivierung.de/>

³ vgl. kopal (2007), S.2. http://kopal.langzeitarchivierung.de/downloads/kopal_Services_2007.pdf

⁴ Digital Information Archiving Service (DIAS); von IBM zusammen mit der niederländischen Nationalbibliothek entwickelt. http://www-935.ibm.com/services/nl/dias/is/implementation_services.html

wickelten Open-Source-Software koLibRI⁵ ein. Organisatorisch beruht die ‚kopal-Solution‘ auf Arbeitsteilung: während in DIAS die gespeicherten Daten verwaltet werden, unterstützt koLibRI das Erstellen, Einspielen (Ingest) und Abrufen von Archivpaketen und führt die Maßnahmen der Langzeitarchivierung wie Datenformatmigrationen durch.⁶ Das Archivsystem DIAS wird dabei von der *Gesellschaft für Wissenschaftliche Datenverarbeitung mbH* in Göttingen (*GWDG*) betrieben, die koLibRI-Software von den kopal-Mandanten *Deutsche Nationalbibliothek (DNB)* und *Niedersächsische Staats- und Landesbibliothek (SUB) Göttingen*, die per Fernzugriff Archivpakete einspeisen und abrufen.⁷

Während *kopal* verschiedene Nutzungsmodelle erarbeitet und mit DIAS und koLibRI ein technisches System zur Verfügung gestellt hat, zielt das Nachfolgeprojekt *Digital Preservation for Libraries (DP4lib)*⁸ mit der Entwicklung nachnutzbarer Vorgehensmodelle für die Langzeitarchivierung auf dessen Etablierung. Dafür erarbeiten die federführenden DP4lib-Projektpartner SUB Göttingen und Deutsche Nationalbibliothek zusammen mit externen Partnern⁹ Anforderungen an das technische System, auf deren Grundlage Spezifikationen zur Aufnahme eines routinemäßigen Archivbetriebs ausgearbeitet werden. Letztendlich sollen Handlungsempfehlungen und Richtlinien erarbeitet werden, so dass am Ende ein nachnutzbarer Service steht.

2.2 Das Projekt *LZA Pädagogik*

Das *Deutsche Institut für Internationale Pädagogische Forschung (DIPF)*¹⁰ ist an *DP4lib* im Rahmen seines Projekts *Kooperative Langzeitarchivierung erziehungswissenschaftlicher e-Ressourcen im Rahmen von kopal (LZA Pädagogik)*¹¹ beteiligt. Ziel des Projektes *LZA Pädagogik* ist es, einen nachnutzbaren Workflow zu Langzeitarchivierung elektronischer erziehungswissenschaftlicher Volltexte, die auf dem fachlichen Dokumentenserver *pedocs*¹²

⁵ koLibRI. kopal Library for Retrieval and Ingest, http://kopal.langzeitarchivierung.de/index_koLibRI.php.de

⁶ kopal (2007), S.1. http://kopal.langzeitarchivierung.de/downloads/kopal_Services_2007.pdf

⁷ Neuroth (2010), Kap.15:14 http://nestor.sub.uni-goettingen.de/handbuch/nestor-handbuch_23.pdf

⁸ Digital Preservation for Libraries - DP4lib. Laufzeit: 2009 – 2011, gefördert von der Deutschen Forschungsgemeinschaft; Projektseite: <http://dp4lib.langzeitarchivierung.de/>

⁹ Externe Partner von DP4lib sind: Bibliotheksservice-Zentrum Baden-Württemberg (BSZ), Deutsches Institut für Internationale Pädagogische Forschung (DIPF), Verbundzentrale des GBV (VZG), Sächsische Landesbibliothek - Staats- und Universitätsbibliothek Dresden (SLUB), Technische Informationsbibliothek und Universitätsbibliothek Hannover (TIB/UB), Thüringer Universitäts- und Landesbibliothek Jena (ThULB)

¹⁰ <http://www.dipf.de>

¹¹ Laufzeit: Dezember 2008 - November 2010; gefördert von der DFG; das Projekt wird im Folgenden mit dem internen Kurznamen *LZA Pädagogik* bezeichnet.

<http://www.dipf.de/de/projekte/foerderung-des-modellvorhabens-lza-paedagogik>

¹² *pedocs* ist Teil des Fachportal Pädagogik <http://www.fachportal-paedagogik.de/start.html> und wird betrieben vom Deutschen Institut für Internationale Pädagogische Forschung (DIPF) betrieben; <http://www.pedocs.de>

publiziert sind, zu erarbeiten. Das Projekt nutzt dabei die ‚kopal-Solution‘ im ‚Teilnehmer-Modell‘ und lässt die Daten durch die Deutsche Nationalbibliothek archivieren, die als *DP4lib*-Verhandlungs- und Vertragspartner fungiert. *pedocs* akquiriert die zu archivierenden digitalen Objekte, stellt die Rechtmäßigkeit ihrer Publikation und Langzeitarchivierung sicher, bereitet sie technisch auf und liefert sie zusammen mit bibliographischen Metadaten als Transferpakete an die Deutsche Nationalbibliothek. Dort werden diese in koLibRI eingespeist, für die Langzeitarchivierung relevante Metadaten ergänzt, Archivpakete erstellt und diese an das Archivsystem DIAS transferiert. Nach einer Migration werden die Archivpakete an das Repositorium zurückgeliefert und die digitalen Objekte über den Server von *pedocs* publiziert. Angestrebt wird der Abschluss eines Vertrags, der einen Routinebetrieb der Langzeitarchivierung einleiten soll, indem er die getrennt und kooperativ wahrzunehmenden Aufgaben der Projektpartner regelt.

Die Besonderheit von *LZA Pädagogik* ist die Implementierung einer Langzeitarchivierungsstrategie in ein Publikations- und Bereitstellungssystem, wodurch die Zitationssicherheit der in *pedocs* veröffentlichten wissenschaftlichen Publikationen gewährleistet ist. Für das Projekt wurde ein Antrag auf Verlängerung der Laufzeit bis Ende 2011 eingereicht, um exemplarische Workflows zur Langzeitarchivierung auch anderer für Erziehungswissenschaft und Bildungsforschung relevanter Materialtypen (Images, Forschungsdaten) erarbeiten zu können.

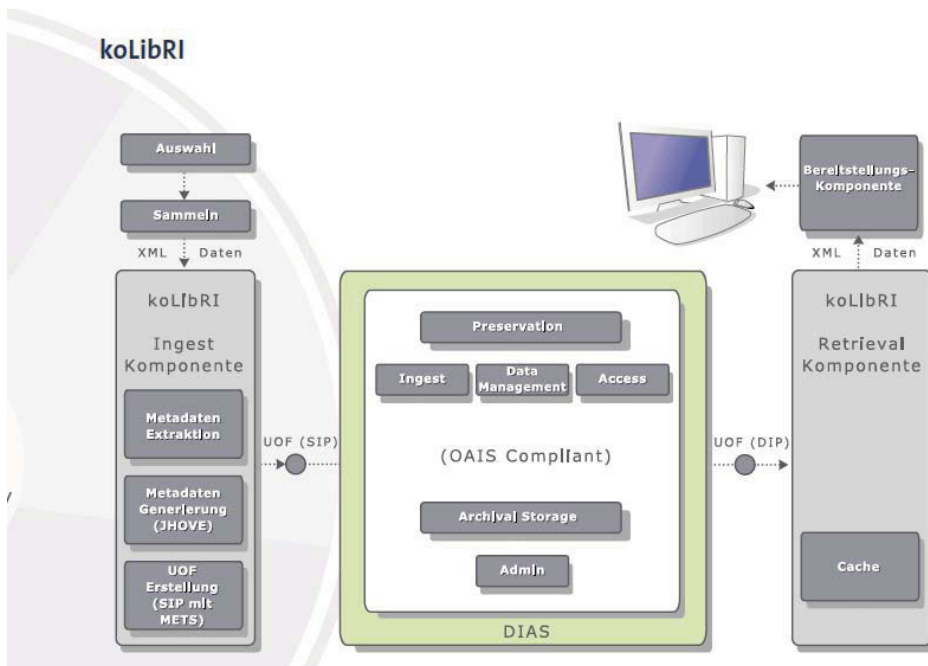


Abb.1: Das *kopal*-Archivsystem mit koLibRI und DIAS¹³

¹³ http://kopal.langzeitarchivierung.de/downloads/kopal_Broschuere_2006.pdf

3. Die Langzeitarchivierung digitaler Objekte

Das Veralten von Datenträgern und der für die Interpretation digitaler Objekte notwendigen technischen Infrastruktur erfordert es, Strategien zu ihrer langfristigen Erhaltung und Verfügbarkeit zu entwickeln.¹⁴ Auf welche Eigenschaften des digitalen Objekts die Langzeitarchivierung abzielt und wie Langzeitarchive konzipiert sein sollten, wird im Folgenden dargestellt.

3.1 Digitale Objekte und Metadaten

Ein digitales Objekt kann auf mehreren Ebenen betrachtet werden: auf der physischen Ebene als eine Menge gespeicherter Zeichen (Bits) auf einem Trägermedium. Auf logischer Ebene hat das Objekt ein definiertes Format, ist also ein von Software und Technik interpretierbares Objekt; die Strategien von Migration und Emulation setzen auf dieser Ebene an. Das für den Menschen verständliche und Bedeutung tragende Objekt bildet die konzeptuelle Ebene, die letztendlich Ziel der Erhaltungsbemühungen ist.¹⁵ Die Repräsentation des konzeptuellen Objekts ist abhängig von der korrekten technischen Interpretation der physischen und logischen Ebene:¹⁶ ein fehlerhafter Bitstrom oder ein unbekanntes Format bzw. nicht mehr existierende Software führen zu Informationsverlust. Daraus muss aber nicht zwangsläufig abgeleitet werden, dass digitale Objekte in ihrer ursprünglichen Form statisch erhalten bleiben müssen; es ist vielmehr möglich, die Art der Kodierung in logischen Objekten und das Speichern in physischen Objekten ohne negative Auswirkungen auf das konzeptuelle Objekt zu verändern: so kann etwa ein Zeitschriftenartikel in einem Bildformat oder Schriftzeichenformat gespeichert werden, ohne dass der intellektuelle Gehalt verlorengeht.¹⁷ Veränderungen sind also nicht grundsätzlich unvereinbar mit dem Gedanken der Langzeitarchivierung, sie müssen vielmehr kontrolliert durchgeführt werden, indem definiert wird, welche Veränderungen notwendig, zulässig oder nachteilig sind. Das bedeutet, dass bestimmt werden muss, welche Eigenschaften des konzeptuellen Objekts unbedingt erhalten bleiben müssen. Damit ist das Konzept der ‚significant properties‘ (erhaltenswerte Eigenschaften)

¹⁴ Exemplarisch wird in diesem Zusammenhang immer wieder auf den Report von Conyers im Auftrag des US-Repräsentantenhauses verwiesen, der die durch Obsoleszenz von Datenträgern verursachten Informationsverluste dokumentiert und die Entwicklung von Strategien zur Langzeitarchivierung digitaler Daten fordert. Conyers (1990), S.5f.

<http://www.bobgellman.com/rg-docs/Taking%20a%20Byte%20Out%20of%20History.PDF>

¹⁵ vgl. Thibodeau, Kenneth: Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years. In: CLIR (2002), S.6. <http://www.clir.org/pubs/reports/pub107/pub107.pdf>

¹⁶ vgl. van Niessen (2002), S. 5. http://www.kb.nl/hrd/dd/dd_onderzoek/reports/2-authenticity.pdf

¹⁷ vgl. Thibodeau, K. In: CLIR (2002), S.8

formuliert, die definiert sind als "the characteristics of digital objects that must be preserved over time in order to ensure the continued accessibility, usability, and meaning of the objects, and their capacity to be accepted as evidence of what they purport to record."¹⁸ Das Forschungsprojekt InSPECT¹⁹ hat sowohl eine Kategorisierung erhaltenswerter Eigenschaften vorgenommen wie auch Faktoren, die auf die Gewichtung dieser Kategorien Einfluss haben, identifiziert; hervorzuheben ist dabei das Informationsinteresse der avisierten Nutzergruppe. Das bedeutet, dass die Bestimmung wichtiger und erhaltenswerter Eigenschaften eines Objekts immer auch abhängig vom Kontext ist, in dem das Objekt genutzt wird.

Beim Umgang mit digitalen Objekten und der Durchführung von Maßnahmen zur Langzeitarchivierung sind besondere Sicherheitsanforderungen zu erfüllen, die aus der Informationstechnik abgeleitet werden können: dies betrifft die Schutzziele *Authentizität*, *Integrität*, *Vertraulichkeit* und *Verfügbarkeit* von digitalen Objekten und ihrer Metadaten. *Vertraulichkeit* meint dabei den Schutz vor unautorisierter Datenpreisgabe, *Verfügbarkeit* die Gewährleistung eines Zugriffs durch autorisierte Nutzer. *Integrität* bezeichnet sowohl die Vollständigkeit des digitalen Objekts wie auch den Ausschluss unautorisierter Modifikationen, beabsichtigter oder unbeabsichtigter.²⁰ Eine Maßnahme zur Gewährleistung der Integrität ist z.B. die Berechnung von Checksummen bei der Übertragung von Objekten, anhand derer der Empfänger ermitteln kann, ob die Datei bei der Übertragung verändert worden ist. *Authentizität* bedeutet schließlich, dass „das Objekt das darstellt, was es vorgibt darzustellen.“²¹ Das heißt, dass eine behauptete Identität eines Objekts (oder auch Subjekts) mit dessen charakterisierenden Eigenschaften übereinstimmt,²² seine Echtheit nachgewiesen ist. Die Authentizität wird unter anderem durch die Dokumentation aller vorgenommenen Veränderungen an den digitalen Objekten sichergestellt.²³

Die Nutzbarkeit digitaler Objekte hängt entscheidend von Metadaten ab. Ohne diese ist eine digitale Ressource möglicherweise unauffindbar oder unbrauchbar. Während zu einer thematischen Recherche in einem Katalogsystem die herkömmlichen bibliographischen Metadaten ausreichen, sind für die Langzeitarchivierung weitere, insbesondere technische Metadaten

¹⁸ Grace (2009), S.3 <http://www.significantproperties.org.uk/inspect-finalreport.pdf>

¹⁹ InSPECT. Investigating Significant Properties of Electronic Content. Vom JISC gefördertes Forschungsprojekt, Laufzeit 2007 - 2009. Homepage: <http://www.significantproperties.org.uk/>

²⁰ vgl. nestor8 (2008), S.15. <http://edoc.hu-berlin.de/series/nestor-materialien/2006-8/PDF/8.pdf>

²¹ nestor8 (2008), S.17. <http://edoc.hu-berlin.de/series/nestor-materialien/2006-8/PDF/8.pdf>

²² vgl. Eckert (2006), S. 7

²³ vgl. nestor8 (2008), S.17 <http://edoc.hu-berlin.de/series/nestor-materialien/2006-8/PDF/8.pdf>

notwendig. Daher hat die Deutsche Nationalbibliothek ein XML-basiertes Metadatenschema für die Langzeitarchivierung entwickelt, das in Ergänzung zu bibliographischen Metadaten technische Informationen zu einem Objekt wie auch seine technische Veränderungshistorie beschreibt. *LMER (Langzeitarchivierungsmetadaten für elektronische Ressourcen)*²⁴ wird im Rahmen von *DP4lib* zusammen mit dem Containerformat METS²⁵ eingesetzt, um Archivpakete zu erstellen.

3.2 Das Referenzmodell OAIS

Zur Formulierung von Anforderungen und Aufgaben, die ein Langzeitarchiv erfüllen muss, hat das *Reference Model for an Open Archival Information System (OAIS)*²⁶ herausragende Bedeutung gewonnen. Ursprünglich zur Archivierung von Weltraumdaten vom *Consultative Committee for Space Data Systems (CCSDS)*²⁷ entwickelt, beschreibt das Modell die Aufgaben und Funktionen eines Archivs (analogen wie digitalen), indem es Minimalanforderungen formuliert, ein funktionales Modell sowie ein Informationsmodell entwickelt. OAIS setzt dabei auf die Migration als Strategie der Langzeitarchivierung. Die Definition des CCSDS beschränkt sich dabei nicht auf das Langzeitarchiv als technisches System, sondern deutet es als eine Organisation von Menschen und Systemen, die die Verantwortung zur Erhaltung von Informationen und ihrer Verfügbarkeit für eine avisierte Nutzergruppe übernommen hat.²⁸

Nach den Mindestanforderungen von OAIS muss ein Archiv Kriterien über die Art von Informationen festlegen, die ins Archiv aufgenommen werden sollen und diese bei den Produzenten akquirieren; hinreichende Kontrolle über die gelieferten Informationen besitzen, um Maßnahmen der Langzeitarchivierung durchführen zu können. Außerdem muss es festlegen, welche Gruppe die avisierte Nutzergruppe sein soll, für die die archivierten Informationen erhalten werden sollen und sicherstellen, dass diese Gruppe diese aus sich selbst heraus verstehen kann, ohne die Hilfe von Experten in Anspruch nehmen zu müssen; die archivierten Informationen müssen zudem der Nutzergruppe unter Nachweis ihrer Authentizität zugänglich gemacht werden. Ein OAIS-konformes Archiv muss darüber hinaus etablierten Leitlinien

²⁴ LMER, Version 1.2 <http://www.d-nb.de/standards/pdf/lmer12.pdf>

²⁵ METS ist ein Containerformat in XML, das verschiedenartiger Metadaten aufnehmen kann. Version 1.4 <http://www.loc.gov/standards/mets/>

²⁶ ISO 14721:2003; die Revision des Standards ist eingeleitet; die überarbeitete Fassung liegt der ISO als "Pink book" vor. CCSDS (2009). http://ddp.nist.gov/refs/650x0p11_OAIS_pink_book.pdf

²⁷ Consultative Committee for Space Data Systems – CCSDS <http://public.ccsds.org/default.aspx>

²⁸ „An OAIS is an archive, consisting of an organization of people and systems, that has accepted the responsibility to preserve information and make it available for a Designated Community.“ CCSDS (2009), S.1-1. http://ddp.nist.gov/refs/650x0p11_OAIS_pink_book.pdf

und Prozeduren folgen, die sicherstellen, dass die archivierte Informationen auch in unvorhergesehenen Fällen geschützt ist und gewährleisten, dass sie außer als Teil einer genehmigten Strategie nie gelöscht werden.²⁹

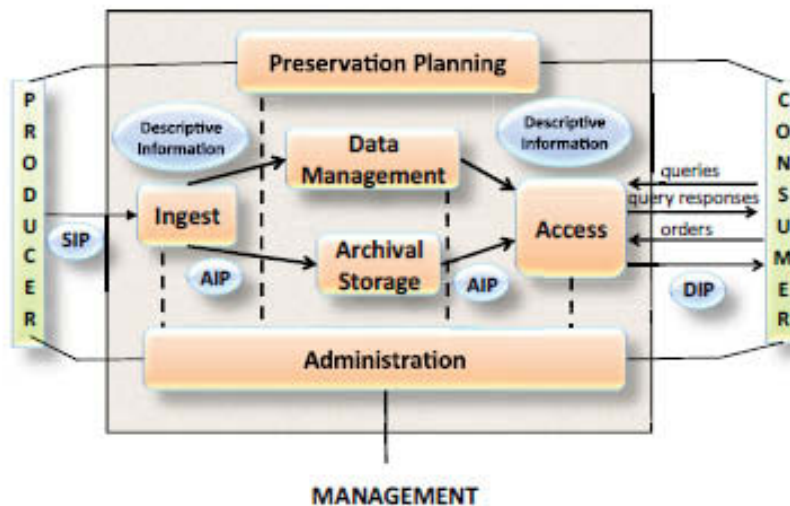


Abb.2: Funktionale Entitäten und Informationspakete eines OAIS³⁰

Das OAIS-Modell definiert drei Typen von Informationsobjekten: Das Eingabe- (*Submission Information Package, SIP*) und das Ausgabepaket (*Dissemination Information Package, DIP*) dienen dem Austausch mit der Umgebung (Produzent und Nutzer). Demgegenüber dient das Archivierungspaket (*Archival Information Package, AIP*) der langfristigen Speicherung der digitalen Informationen. Die Prozesse, die innerhalb des OAIS ablaufen, werden durch die funktionalen Entitäten modelliert. Die Entität *Ingest* beschreibt den Transfer von SIPs in das OAIS und die Erstellung der AIPs. Innerhalb des *Data Management* werden die aus dem Ingest stammenden, deskriptiven Informationen und die administrativen Informationen, die das Archiv betreffen, gespeichert. Die Entität *Archival Storage* umfasst Prozesse, die die Übernahme und Aufbewahrung der AIPs betreffen. Das Modul *Access* dient der Benutzerunterstützung in den Bereichen des Retrieval und der Ausgabe der DIPs. Das *Preservation Planning* übernimmt Aufgaben des technischen Monitoring und der Beobachtung von Veränderungen der Nutzergewohnheiten; daraus werden Erhaltungs- und Migrationsstrategien entwickelt. Das für das Management verantwortliche Modul *Administration* steuert die Gesamtabläufe im OAIS und koordiniert die Zuständigkeiten der anderen Funktionsmodule.

²⁹ vgl. CCSDS (2009), p. 3-1 http://ddp.nist.gov/refs/650x0p11_OAIS_pink_book.pdf

³⁰ CCSDS (2009), S.4-1. http://ddp.nist.gov/refs/650x0p11_OAIS_pink_book.pdf

Außerdem werden in diesem Bereich die Beziehungen mit der Umwelt (Producer/Consumer) geregelt.

Befindet sich ein Informationspaket im Bereich des *Ingest*, so sind bereits wesentliche und nicht mehr zu ändernde Merkmale definiert; schon vorher, im Bereich des *Pre-Ingest*, müssen also Nutzungsszenarien und Archivierungsprozesse berücksichtigt werden. Da die Qualität der digitalen Objekte kaum verbessert, sondern höchstens erhalten werden kann, sind bereits an der Schnittstelle von Datenproduzent und Archiv eine sorgfältige Planung der Abläufe, die Klärung rechtlicher Fragen und Vereinbarungen über die Beschaffenheit der zu archivierenden Objekte notwendig. Vor diesem Hintergrund ist es zu verstehen, dass eine JISC-Studie die Kosten für den Ingestprozess auf etwa 42% der Gesamtkosten für den Betrieb eines OAIS ansetzt.³¹ Da das OAIS-Modell diese Schnittstelle nicht differenziert, gab das CCSDS den *Producer-Archive Interface Methodology Abstract Standard (PAIMAS)*³² heraus, der den Ingest-Prozess in vier Phasen unterteilt: in der *Vorbereitungs-* und der *Definitionsphase* werden grundsätzliche Bedingungen zwischen Archiv und Produzent vereinbart und erprobt. Hierzu gehören etwa die Auswahl und Definition der zu übernehmenden Objekte, die Bestimmung der von Transferkonditionen und die Abklärung sämtlicher rechtlicher Aspekte. In der *Transferphase* findet die technische Umsetzung der Datenübertragung statt, in der *Validierungsphase* werden die digitalen Objekte auf ihre angenommenen Eigenschaften hin überprüft.

3.3 Zur Umsetzung der Langzeitarchivierung – die ‚mittlere Lage‘

OAIS wird weltweit als Referenzmodell zur Langzeitarchivierung digitaler Ressourcen anerkannt und als Grundlage zur Konzeption von Archivsystemen verwendet. Es bietet Orientierung bei der Ableitung und Strukturierung funktionaler Kriterien und schafft eine Grundlage für eine gemeinsame Begrifflichkeit. OAIS kann also als Kommunikationstool, das eine vereinheitlichende Terminologie bietet, genutzt werden.³³ Die Spezifikation einer Implementierung bietet das Modell explizit nicht, daher werden auch Randbedingungen wie Aufwand (Kosten) und Qualität vernachlässigt. Da das Modell auf abstrakter, konzeptueller Ebene formuliert ist und seine Mindestanforderungen nicht schwer zu erfüllen sind, bietet es eine hohe Flexibilität und ermöglicht es Archiven und Repositorien, ihre eigenen Modelle an das OAIS-Referenzmodell anzupassen und zu erweitern.³⁴

³¹ vgl. Beagrie (2008b) <http://www.jisc.ac.uk/media/documents/publications/keepingresearchdatasafe0408.pdf>

³² ISO 20652:2006. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=39577

³³ vgl. Allinson (2006), S.4. <http://www.ukoln.ac.uk/repositories/publications/oais-evaluation-200607/Drs-OAIS-evaluation-0.5.pdf>

³⁴ vgl. Allinson (2006), S.5

Zur Implementierung eines Langzeitarchivierungsbetriebes ist jedoch eine ‚mittlere‘ Ebene notwendig.³⁵ Auf dieser Stufe finden sich Technologien und zugehörige Standards, die abstrakt beschriebenen Modellelemente konkretisieren. Ihre Ausformung finden sie in Kriterienkatalogen, die der Ausarbeitung von Spezifikationen und den tatsächlichen Implementierungen vorangehen. In Deutschland hat das ‚Kompetenznetzwerk *nestor*³⁶ den Katalog ‚Kriterien für vertrauenswürdige digitale Langzeitarchive‘³⁷ herausgegeben, der als DIN 31644 standardisiert werden sollen.³⁸ In Ergänzung dazu hat die *nestor AG Standards* den auf den Modellen OAIS und PAIMAS beruhenden Praxisleitfaden ‚Wege ins Archiv‘,³⁹ für die Prozesse im Bereich der Informationsübernahme ins Langzeitarchiv publiziert. Unter Nutzung und in Übereinstimmung mit internationalen Standards sollen diese Publikationen den Transfer digitaler Objekte in die Langzeitarchive steuern und begleiten.

³⁵ vgl. Allinson (2006), S.14.

³⁶ *nestor* – Das Kompetenznetzwerk Langzeitarchivierung und Langzeitverfügbarkeit digitaler Ressourcen in Deutschland. <http://www.langzeitarchivierung.de/>

³⁷ *nestor8* (2008) http://files.d-nb.de/nestor/materialien/nestor_mat_08.pdf

³⁸ NABD (2010). <http://www.nabd.din.de/cmd?artid=133059734&contextid=nabd&bcrumblevel=1&subcommitteeid=112656173&level=tpl-art-detailansicht&committeeid=54738855&languageid=de>

³⁹ *nestor10* (2008) http://files.d-nb.de/nestor/materialien/nestor_mat_10.pdf

4. Das fachliche Repository *pedocs*

pedocs ist ein fachliches Open Access-Repository für wissenschaftliche Schriften und praxisbezogene Texte und Materialien der Bildungsforschung und Pädagogik. Von der Deutschen Forschungsgemeinschaft gefördert, wird *pedocs*⁴⁰ als Teil des *Fachportal Pädagogik* aufgebaut und ist am *Deutschen Institut für Internationale Pädagogische Forschung (DIPF)* angesiedelt.

Als fachliches Repository, das eine Open Access-Strategie verfolgt, ist bei *pedocs* der Dokumentenzufluss nicht automatisch gewährleistet; um Publikationen zu akquirieren, beschreitet *pedocs* daher den sogenannten „Grünen Weg“ des Open Access, der eine zeitgleiche oder nachträgliche Zweitveröffentlichung von Verlagspublikationen auf einem Dokumentenserver bedeutet,⁴¹ dies betrifft sowohl Pre- und Postprints⁴² wie auch Monographien, Forschungs- oder Konferenzberichte. Dazu hat *pedocs* eine Reihe von Kooperationsmodellen mit Verlagen ausgearbeitet⁴³ und bietet zudem die Möglichkeit einer Selbstarchivierung per Upload von Volltexten auf den Server. Die Identität des Autors wird dabei durch ein Double-Opt-In-Verfahren sichergestellt: die eintragende Person muss beim Eintrag ihre Email-Adresse angeben und erhält eine Benachrichtigung an diese Adresse, die sie bestätigen muss; so wird gesichert, dass die Person, die das Dokument einstellen möchte, auch diejenige ist, für die sie sich ausgibt.

Verträge mit Autoren und Verlagen bilden die rechtliche Grundlage der Veröffentlichung in *pedocs*. Diese Strategie ermöglicht es *pedocs*, begutachtete Verlagspublikationen zu veröffentlichen und den freien Zugang zu wissenschaftlicher Information, wie er von nationalen und internationalen Wissenschaftseinrichtungen gefordert wird, zu unterstützen.⁴⁴

Die Auffindbarkeit der Inhalte von *pedocs* ist durch die Anbindung an Suchmaschinen, die Weitergabe der Daten an OAI-Server wie BASE und OAIster, die Registrierung beim Directory of Open Access Repositories (OpenDOAR), die Dokumentation durch die FIS Bildung Literaturdatenbank und den Online-Katalog der Deutschen Nationalbibliothek gewährleistet.

⁴⁰ Förderzeitraum: 8/2008-12/2010. Gefördert durch die Deutsche Forschungsgemeinschaft.
<http://www.pedocs.de/>

⁴¹ vgl. Müller (2007), S.183

⁴² Preprint: Manuskriptfassung einer Publikation, nicht begutachtete wissenschaftliche Arbeit; Postprint: begutachtete Verlagsfassung (in anderem Layout). http://www.pedocs.de/info_openaccess.php?la=de

⁴³ Informationen zu den verschiedenen Kooperationsmodellen (Delayed Access, Selected Access, Open Access Direct etc.) unter http://www.pedocs.de/informationen_verlage.php?la=de

⁴⁴ vgl. Max Planck-Ges. (2003). http://www.mpg.de/pdf/openaccess/BerlinDeclaration_dt.pdf

In der *policy*⁴⁵ von *pedocs* sind das Profil des Dokumentenservers und die Kriterien dokumentiert, nach denen Dokumente aufgenommen werden. Die Redaktion von *pedocs* prüft die eingegangenen Publikationen gegen die Leitlinien, erschließt sie bibliographisch und inhaltlich, wobei auch der Begutachtungsstatus der Publikation angegeben wird und vergibt einen Persistent Identifier (URN)⁴⁶ zu ihrer eindeutigen Identifizierung.

Der von *pedocs* eingesetzte Dokumentenserver DIDoRe basiert auf einer den Anforderungen des DIPF angepassten Version der Software Opus (Version 3.1) und wird ständig weiterentwickelt. Die Datensicherung erfolgt täglich auf einem Bandlaufwerk, monatlich wird ein Vollbackup erstellt.

4.1 Zertifizierung von Repositorien

Dokumentenserver als Plattformen zur Publikation und Distribution elektronischer Dokumente können sich in Deutschland von der *Deutschen Initiative für Netzwerkinformation (DINI)*⁴⁷ als „Vertrauenswürdige Repository“ zertifizieren lassen. Die dafür zu erfüllenden Kriterien⁴⁸ umfassen die Sichtbarkeit des Gesamtangebots, die Ausarbeitung von Leitlinien für inhaltliche Kriterien und den Betrieb des Repositoriums, Autorenbetreuung, Rechtliche Aspekte, Informationssicherheit, Erschließung, Schnittstellen, Zugriffsstatistik und Langzeitverfügbarkeit. Für jedes Kriterium schreibt DINI Mindestanforderungen vor und gibt Empfehlungen, deren Erfüllung optional ist. *pedocs* möchte sich von DINI zertifizieren lassen und steht kurz vor der Eröffnung des Verfahrens.

DINI grenzt Dokumentenserver zwar in ihrem Aufgabenspektrum als Dokumenten- und Publikationsservices explizit von Langzeitarchiven ab, betont aber die Notwendigkeit, sich mit Fragen der Langzeitarchivierung auseinanderzusetzen. Vor diesem Hintergrund sind die Mindestanforderungen einer minimalen Verfügbarkeit der publizierten Dokumente und ihrer Metadaten für die Dauer von fünf Jahren sowie der Absenz jeglicher technischer Schutzmaßnahmen von Originaldateien und Archivkopien zu verstehen.⁴⁹ Maßnahmen zur Erfüllung dieser Kriterien sind nach DINI in Übereinstimmung mit den Empfehlungen des Kompetenznetzwerks *nestor* durchzuführen.⁵⁰ In ihren Empfehlungen legt DINI die Erarbeitung einer Strategie zur Sicherstellung der Langzeitverfügbarkeit der Dokumente nahe, etwa in Form

⁴⁵ vgl. <http://www.pedocs.de/leitlinien.html>

⁴⁶ Persistent Identifier identifizieren digitale Objekte eindeutig und standortunabhängig (siehe Kap.5)

⁴⁷ Deutsche Initiative für Netzwerkinformation e.V. <http://www.dini.de>

⁴⁸ DINI (2010): DINI-Zertifikat Dokumenten- und Publikationsservice 2010. <http://edoc.hu-berlin.de/series/dini-schriften/2010-3/PDF/3.pdf>

⁴⁹ vgl. DINI (2010), S. 31f.

⁵⁰ vgl. DINI (2010), S.45

einer Kooperation mit einer Archivierungsinstitution. Auch das Speichern der Dokumente in offenen und für die Langzeitarchivierung geeigneten Formaten sowie die Erarbeitung einer Regelung zum Umgang mit der Löschung von Dokumenten werden angeraten.⁵¹ Indem *pedocs* in seinen Workflow eine Lösung zur Langzeitarchivierung integriert, setzt das Repositorium also nicht nur die Mindestanforderungen, sondern auch die Empfehlungen von DINI um.

⁵¹ vgl. DINI (2010), S.32 <http://edoc.hu-berlin.de/series/dini-schriften/2010-3/PDF/3.pdf>

5. Langzeitarchivierung der Inhalte von *pedocs* – Workflow

Da das Projekt *LZA Pädagogik* an Organisation, Technik und Inhalten von *pedocs* anschließt, hat die Redaktion von *pedocs* den bestehenden, am Kriterienkatalog von DINI orientierten Workflow um die Erfordernisse, die sich aus dem Langzeitarchivierungsprojekt ergeben, erweitert. Grundlage hierfür ist der nestor-Leitfaden „Wege ins Archiv“.⁵² Die folgende Darstellung des Workflows von *pedocs* fokussiert auf die für die Langzeitarchivierung relevanten Aspekte; dabei soll in diesem Kapitel die technische Seite der Objektbearbeitung und des Transfers der zu archivierenden Dateien an die Nationalbibliothek im Mittelpunkt stehen, bevor der Blick sich im nächsten Kapitel auf Angelegenheiten des Managements im Prozess der Informationsübernahme durch das Archiv richtet.

Die Redaktion von *pedocs* bereitet die digitalen Objekte und Metadaten auf und liefert sie in Form eines Transferpaketes über einen Hotfolder⁵³ an die Nationalbibliothek, die als *DP4lib*-Partner die Archivierung der Objekte vornimmt. Die im Transferpaket enthaltenen Dateien werden in *koLibRI* so bearbeitet, dass sie als *Submission Information Package (SIP)* an das Archivierungssystem *DIAS* übertragen und dort als *Archival Information Package (AIP)* gespeichert werden. In jeden Arbeitsschritt integriert sind Maßnahmen zur Realisierung der Kernerhaltungsziele *Vertraulichkeit*, *Verfügbarkeit*, *Authentizität* und *Integrität* der digitalen Objekte.

5.1 Auswahl der Objekte – Pflichtablieferung an die Deutsche Nationalbibliothek

Infolge der Erweiterung des Sammelauftrages der Deutschen Nationalbibliothek unterliegen alle in *pedocs* publizierten Volltexte der Ablieferpflicht an die Deutsche Nationalbibliothek,⁵⁴ die diese ihrem Auftrag gemäß nach einem festgelegten Geschäftsgang zu archivieren hat. Da die Ansprüche der Nationalbibliothek an die Langzeitarchivierung der gelieferten Objekte aber nicht zwangsläufig mit den Anforderungen anderer *DP4lib*-Kooperationspartner übereinstimmen, steht die Deutsche Nationalbibliothek innerhalb des Kooperationsmodells vor einer zweifachen Aufgabe: zum Einen muss sie als zentrale Gedächtnisorganisation Deutschlands

⁵² nestor10 (2008). http://files.d-nb.de/nestor/materialien/nestor_mat_10.pdf

⁵³ Hotfolder: Ordner, um zwei oder mehr Programme, die über keine direkte Schnittstelle verfügen, miteinander kommunizieren zu lassen. Vgl. Der Siebdruck. <http://www.der-siebdruck.de/index.php?id=53>

⁵⁴ Nach den Sammelrichtlinien der DNB handelt es sich bei der Kollektion von *pedocs* um „parallel erscheinende Netzpublikationen“; DNB (2009), S.60. <http://files.d-nb.de/pdf/sammelrichtlinien.pdf>; vgl. auch: Pflichtablieferungsverordnung – PflAV, § 7: Beschaffenheit von Netzpublikationen und Umfang der Ablieferungspflicht sowie § 2: Beschaffenheit körperlicher Medienwerke und Umfang der Ablieferungspflicht; BMJ (2008a), o.S. <http://bundesrecht.juris.de/pflav/BJNR201300008.html>

die von *pedocs* im Rahmen des Sammelauftrags gelieferten Texte nach dem internen Geschäftsgang archivieren, zum Anderen muss sie als Dienstleister Kundenanforderungen erfüllen, die eine Modifikation des bestehenden Workflows erfordern können. Im Falle von *pedocs* ist eine dieser im Dienstleistungsbereich anzusiedelnden Aufgaben, die außerhalb des nationalbibliothekarischen Aufgabenspektrums liegen, die Rücklieferung von migrierten Objekten an das Repositorium. Um beiden Anforderungen gerecht zu werden, erstellt die Nationalbibliothek von jedem gelieferten Objekt eine Kopie, die dem internen Geschäftsgang unterzogen wird. Am „originalen“ Objekt werden die mit den Ablieferern vereinbarten Maßnahmen der Langzeitarchivierung vorgenommen, die momentan im Verbund der *DP4lib*-Partner prototypisch ausgearbeitet und im Folgenden dargestellt werden.

5.2 Vorbereitung der Objekte und Metadaten

Das Zielformat, in dem digitale Textobjekte an den Hotfolder übertragen werden sollen, ist das für die Langzeitarchivierung maßgebliche Format PDF/A.⁵⁵ Dokumente in diesem Format sind in sich abgeschlossen, d.h. sie verweisen weder direkt noch indirekt auf externe Quellen und enthalten alle notwendigen Informationen (sichtbare Inhalte wie auch Metadaten) in sich selbst.⁵⁶ Dies bringt allerdings auch eine Erhöhung des Datenvolumens mit sich. *pedocs* nutzt die Spezifikation PDF/A-1b, da das Format PDF/A-1a von der eingesetzten Validierungssoftware nicht unterstützt wird.

5.2.1 Digitalisierung und OCR

Wird ein Dokument nicht als „born digital“, sondern als gedruckte Publikation an *pedocs* geliefert, wird es zunächst digitalisiert⁵⁷ und anschließend mit einer OCR-Software behandelt, um es in einen durchsuchbaren Volltext zu konvertieren. Das Facsimile wird dabei mit einer unsichtbaren Hintergrundebene versehen, auf der das OCR-Ergebnis abgelegt ist, so dass der Nutzer das OCR-Ergebnis durchsucht, aber das Faksimile betrachtet. *pedocs* kann mit der eingesetzten Software eine Fehlerfreiheit von 98,5-99,8% erzielen, was leicht über der im Durchschnitt von automatischen OCR-Programmen erreichten Erkennungsgenauigkeit von 98-99% liegt.⁵⁸ Die als wirtschaftlich angesehene Erkennungsrate von 99,5%⁵⁹ könnte durchgängig nur durch nachträgliche manuelle Bearbeitung gewährleistet werden. Daher akzeptiert

⁵⁵ ISO19005-1:2005/Cor :2007. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=45613

⁵⁶ vgl. http://www.pdfa.org/doku.php?id=pdf_a_ein_standard_fuer_die_langzeit-archivierung

⁵⁷ Die Digitalisate werden in einer Auflösung von 300dpi und im verlustfreien Kompressionsformat TIFF erstellt. vgl. *pedocs* (2010)

⁵⁸ vgl. *pedocs* (2010)

⁵⁹ vgl. Parschik (2006), S. 1437 http://www.zlb.de/aktivitaeten/bd_neu/heftinhalte2006/Technik011206.pdf

pedocs solche "schmutzige OCRs". Für die Recherche bedeutet dies, dass nur eine Positivsuche durchgeführt wird: man kann weder sicher sein, dass mit den angezeigten Treffern alle theoretisch möglichen Ergebnisse gefunden wurden, noch bedeutet ein negatives Ergebnis, dass kein Treffer vorhanden ist.⁶⁰ Das OCR-behandelte Digitalisat wird als PDF abgespeichert und zur weiteren Bearbeitung in das Redaktionssystem eingespeist.

Exkurs: OCR und Authentizität

Die Fehleranfälligkeit der OCR-Behandlung löste in der Redaktion von *pedocs* eine Diskussion um die Authentizität des Objekts aus, die nicht mehr vollumfänglich gewährleistet werden kann. Dies führte schließlich zu einer Änderung des entsprechenden Passus im Autorenvertrag. In den Verträgen mit den Autoren und Verlagen wird nun darauf hingewiesen, „dass die Ergebnisse dieser Verfahren in Relation zur Qualität der Printvorlagen fehlerbehaftet sein können.“⁶¹ Die Einschränkung des Gewährleistungsumfangs ist eine Maßnahme, die auch im nestor-Kriterienkatalog nahegelegt wird, indem die Einführung und Definition eines „Erfüllungsgrades“ für jegliche Maßnahme der Langzeitarchivierung angeregt wird.⁶² Diese Empfehlung führt zurück zur Frage nach wichtigen und weniger wichtigen Eigenschaften der zu erhaltenden Objekte und damit zur Frage nach den erhaltenswerten Eigenschaften des digitalen Objekts.

5.2.2 Technische Prüfungen

Jede Datei wird nach ihrem Eingang lokal einer automatischen Virenprüfung unterzogen. Ob das Dokument frei von Schutzmaßnahmen (DRM) ist, wird von der Redaktion händisch in den Dokumenteigenschaften überprüft. Bei im Format PDF gelieferten Dokumenten wird über eine Umwandlung in ein Textformat geprüft, ob sie OCR-behandelt sind; automatisch wird dabei auch eine Überprüfung auf Schutzmaßnahmen vorgenommen. Sind diese vorhanden, kontaktiert die Redaktion den Rechteinhaber, um eine Aufhebung der Beschränkungen zu bewirken.⁶³ Vor der Freischaltung auf dem Dokumentenserver werden die Dokumente mit dem Open Source - Tool JHOVE (Version 1.5)⁶⁴ einer Authentizitätsprüfung unterzogen,

⁶⁰ vgl. DFG(2009), S.11 http://www.dfg.de/download/pdf/foerderung/programme/lis/praxisregeln_digitalisierung.pdf

⁶¹ *pedocsA*, § 4, 3. <http://www.pedocs.de/autorenvertrag.php?la=de>
vgl. auch: *pedocsV*, § 2, 2. http://www.pedocs.de/verlagsvertrag_muster.php?la=de

⁶² vgl. nestor (2008), S.13. http://files.d-nb.de/nestor/materialien/nestor_mat_10.pdf

⁶³ Nach dem Urheberrechtsgesetz dürfen technische Kopierschutzverfahren nicht ohne Weiteres umgangen werden; allerdings besteht der Anspruch, dass der Rechteinhaber der archivierenden Institution die zur Umgehung des Schutzes erforderlichen Mittel zur Verfügung stellt (§ 95a und §95b UrhG); BMJ (2008b) <http://www.gesetze-im-internet.de/urhg/BJNR012730965.html#BJNR012730965BJNG002501377>

⁶⁴ JSTOR/Harvard Object Validation Environment; Gemeinschaftsprodukt von JSTOR und der Harvard

indem sie auf Validität und Wohlgeformtheit geprüft werden; dabei extrahiert JHOVE auch elementare technische Metadaten.

5.2.3 Anreicherung mit Metadaten

Um die Recherchierbarkeit der digitalen Objekte in verschiedenen Informationssystemen zu gewährleisten, werden Metadaten in unterschiedlichen Formaten generiert: zur Publikation auf dem Dokumentenserver *pedocs* und in der *FIS Bildung Literaturdatenbank*⁶⁵ werden Metadaten in einem hausinternen Format erstellt. Nach dem Kriterienkatalog von DINI erfordert die Spezifikation des OAI-Protokolls Metadaten in Dublin Core;⁶⁶ zur Recherche im Katalog der Deutschen Nationalbibliothek sind Metadaten im Format ONIX⁶⁷ zu erstellen.

Die Redaktion von *pedocs* erfasst die Dokumente bibliographisch; die inhaltliche Erschließung erfolgt durch die Vergabe von Schlagwörtern nach kontrolliertem Vokabular und ein Abstract. Klassifiziert werden die Dokumente nach den Hauptgruppen der Dewey Dezimalklassifikation;⁶⁸ außerdem gemäß der Sektioneneinteilung der Deutschen Gesellschaft für Erziehungswissenschaft (DGfE) nach pädagogischer Teildisziplin. *pedocs* dokumentiert in den Metadaten, welchem Nutzungsrecht das Dokument unterliegt (dt. Urheberrecht, Creative Commons). Zur Erstellung elementarer technischer Metadaten (Dateiformat, Formatversion) wird JHOVE genutzt. Um die stabile Adressierung eines jeden Dokumentes sicherzustellen, vergibt *pedocs* neben einer individuellen URL einen „Uniform Resource Name“ (URN),⁶⁹ der digitale Objekte unabhängig vom Ort der Speicherung identifiziert und damit ein wichtiger Faktor zur Erhaltung der Langzeitverfügbarkeit ist.

5.2.4 Deckblatt generieren

Im nächsten Arbeitsschritt wird dem Objekt ein dokumentarisches Deckblatt, das wesentliche Metadaten enthält, vorangestellt; die Metadaten werden somit Teil des digitalen Objekts.⁷⁰ Das Deckblatt enthält das Logo von *pedocs* und des *DIPF*, wesentliche bibliographische Metadaten sowie den URN, ggf. Angaben zum Verlag der Erstpublikation, urheberrechtliche

University Library: <http://hul.harvard.edu/jhove/>

⁶⁵ http://www.fachportal-paedagogik.de/fis_bildung/fis_form.html

⁶⁶ DINI (2010), S.27. <http://edoc.hu-berlin.de/series/dini-schriften/2010-3/PDF/3.pdf>

⁶⁷ ONIX ist ein XML-basierter Metadatenstandard für verschiedene Medientypen, ursprünglich zum Austausch von bibliographischen Metadaten und Produktdaten im Buchhandel verwendet. <http://www.editeur.org/8/ONIX/>

⁶⁸ Anforderung von DINI. DINI (2010), S.53ff. <http://edoc.hu-berlin.de/series/dini-schriften/2010-3/PDF/3.pdf>

⁶⁹ In Deutschland werden URNs von der Deutschen Nationalbibliothek verwaltet und an Repositorienbetreiber vergeben. http://www.d-nb.de/netzpub/erschl_lza/np_urn.htm

⁷⁰ Dies hat zur Folge, dass jegliche Änderung an Metadaten, die im Deckblatt enthalten sind, zur Erstellung einer neuen Version des Objekts führen muss, obwohl sich der intellektuelle Gehalt des Dokuments nicht geändert hat.

Nutzungshinweise sowie Kontaktmöglichkeiten. Ist die Digitalisierung von *pedocs* vorgenommen worden, ist dies ebenfalls vermerkt.

5.2.5 Publikation in *pedocs*

Am Ende des Bearbeitungsprozesses steht die Konvertierung der PDF-Datei in das Format PDF/A und dessen anschließender Validierung durch JHOVE. Als eine Maßnahme der Qualitätssicherung hat *pedocs* im Redaktionssystem ‚Prüfpunkte‘ eingerichtet, die am Ende des Bearbeitungsprozesses erfolgreich durchlaufen sein müssen. Zur Sicherstellung der Datenintegrität werden schließlich bei der Freischaltung in *pedocs* die drei Checksummen (Hashwerte) CRC32, MD5 und SHA-1 berechnet. Bei Hashfunktionen handelt es sich um kryptologische Verfahren, die eindeutige, digitale Fingerabdrücke von Datenobjekten berechnen und entweder mit dem Dokument versandt oder gespeichert werden, damit der Empfänger oder Nutzer anhand dieses Wertes die Integrität des Objekts prüfen kann.⁷¹ Die Checksummen sind in der Vollanzeige des Datensatzes einsehbar.

The screenshot shows the metadata page for a document in the *pedocs* system. At the top, there is a navigation bar with links like 'Startseite', 'pedocs', and 'pedocs - Individuelle Wirkungen des Ganztags schulbesuchs in der Sekundarstufe...'. The main content area is titled 'pedocs - Individuelle Wirkungen des Ganztags schulbesuchs in der Sekundarstufe...' and includes the following information:

- Aufsatz (Zeitschrift) zugänglich unter:** URN: urn:nbn:de:hbz:5:1-111-opus-26653, URL: http://www.pedocs.de/volltexte/2010/26653/
- Individuelle Wirkungen des Ganztags schulbesuchs in der Sekundarstufe. Erkenntnisse aus der Studie zur Entwicklung von Ganztagschule (StEG)** by Fischer, Natalie; Brümmer, Felix; Kuhn, Hans Peter; Züchner, Ivo
- Originalveröffentlichung:** Schulverwaltung, Hessen, Rheinland-Platz 15 (2010) 2, S. 38-39
- pdf-Format:** Dokument 1.pdf (305 KB)
- Schlagwörter (Deutsch):** Ganztagschule, Sekundarstufe I, Sekundarstufe II, Außenunterrichtliche Aktivität, Einflussfaktor, Schulnote, Motivation, Soziales Verhalten, Freizeitverhalten, Familie, Empirische Untersuchung, Schulentwicklung, Deutschland
- Pädagogische Teildisziplin:** Empirische Bildungsforschung
- Pädagogische Teildisziplin:** Schulpädagogik
- Sonstige beteiligte Institution:** Deutsches Institut für internationale Pädagogische Forschung (DIPF) -Frankfurt, Mainz
- Dokumentart:** Aufsatz (Zeitschrift)
- ISSN:** 1432-9603; 1618-9159; 1865-2026
- Sprache:** Deutsch
- Erscheinungsjahr:** 2010
- Eintrag erfolgt am:** 04.03.2010
- Begutachtungszustand:** (Verlage-)Lektorat
- Abstract (Deutsch):** Der Beitrag konzentriert sich auf die ersten beiden Erhebungswellen von StEG. Zusammenfassend werden die Resultate zu individuellen Wirkungen der Teilnahme an Ganztags schulangeboten in der Sekundarstufe dargestellt. Insgesamt zeigen sich kleine positive Effekte des Ganztags schulbesuchs. (...) In Bezug auf Freizeitverhalten und Familienleben der SchülerInnen scheint der Ganztags schulbesuch keine grundlegenden Umgestaltungen zur Folge zu haben. Insbesondere beim Übertritt in die Sekundarstufe I können protektive Wirkungen der Teilnahme an Ganztags angebot im Hinblick auf die Entwicklung von Lernmotivation und Schulnoten nachgewiesen werden. Aber auch für ältere SchülerInnen gilt: Der dauerhafte Besuch von qualitativ hochwertigen Angeboten der Ganztagschule kann sich positiv auf das Sozialverhalten, die Motivation und schulische Leistungen auswirken. (DIPF/Orig.)
- Statistik:** Anzahl der Zugriffe auf dieses Dokument (Bar chart showing a small number of hits)
- Bestellmöglichkeit:** Recherche im Verzeichnis Lieferbarer Bücher (buchhandel.de)
- Prüfsummen:**

Dokument 1.pdf	CRC32	70e8154e
	MD5	0729e788f142882d1e07002010036b64
	SHA-1	0471e01e0e182e8212e426e740e088905d8ef

Abb.3: Vollanzeige eines Datensatzes in *pedocs*

⁷¹ Eckert (2006), S. 353

5.3 Transferprozess

5.3.1 Generierung von Transferpaketen

Für den Transfer der zu archivierenden Dateien an den Hotfolder der Deutschen Nationalbibliothek wird ein Transferpaket erstellt, das als ZIP-Datei übertragen wird.⁷² Das Transferpaket enthält ein Haupt- und ein Unterverzeichnis. Im Unterverzeichnis (Ordner „content“) befinden sich diejenigen Dateien, die archiviert werden sollen: das digitale Objekt (PDF/A) und die Metadaten im internen Format (.txt). Im Hauptverzeichnis befinden sich mit den Metadaten in Dublin Core (dc.xml) und ONIX (.onix.xml) diejenigen Dateien, die von der Nationalbibliothek zur Verarbeitung benötigt werden. Zu jeder gelieferten Datei wird eine Checksumme MD5 berechnet, so dass in einem Transferpaket acht Dateien enthalten sind (vier Dateien + vier Checksummen). Zur Absicherung der Übertragung wird für das Transferpaket ebenfalls eine Checksumme berechnet, die in einer Textdatei gespeichert wird; diese beiden Objekte (Transferpaket und Checksummen-Datei für das Transferpaket) werden an den Hotfolder übertragen.

File Name	File Type	Modified	Size	Ratio	Packed	CRC
content	Folder					
2665.dc.xml	XML-Dokument	25.08.2010 12:04:32	2,829	59%	1,166	CC1E8736
2665.onix.xml	XML-Dokument	25.08.2010 12:04:32	2,553	59%	1,044	3AC6658A
2665.dc.xml.md5	MD5-Datei	25.08.2010 12:04:32	32	0%	32	A6FE0687
2665.onix.xml.md5	MD5-Datei	25.08.2010 12:04:32	32	0%	32	077C2F31

Abb.4: Hauptverzeichnis Transferpaket

File Name	File Type	Modified	Size	Ratio	Packed	CRC
herp10_02_0002_D_A.pdf.md5	MD5-Datei	25.08.2010 12:04:32	32	0%	32	A5A93939
herp10_02_0002_D_A.pdf	Adobe Acrobat Doc...	25.08.2010 12:04:32	312,132	12%	274,743	70A51B4E
2665.txt.md5	MD5-Datei	25.08.2010 12:04:32	32	0%	32	BF0185CA
2665.txt	Text Document	25.08.2010 12:04:32	2,007	45%	1,106	E6311FA6

Abb.5: Unterverzeichnis Transferpaket

⁷² Zur Veranschaulichung befindet sich exemplarisch im Anhang ein Transferpaket mit seinen Dateien und Checksummen

5.3.2 Transfer und Weiterverarbeitung

Nach dem Harvesting des Hotfolders durch die Nationalbibliothek und den Validitätsprüfungen werden die Dateien in koLibRI eingespielt, wo im einfachen Fall aus den gelieferten deskriptiven und den von JHOVE automatisch extrahierten technischen Metadaten weitere für die Langzeitarchivierung relevante Metadaten generiert, im Containerformat METS kodiert und zusammen mit dem Archivierungsobjekt zu einem Submission Information Package (SIP) komprimiert werden. Dieses wird in das Archivierungssystem DIAS eingespielt und dort als Archival Information Package (AIP) gespeichert. Steht eine Migration an, werden Dissemination Information Packages (DIPs) generiert und nach koLibRI geliefert, wo die Migration durchgeführt wird.⁷³ Die migrierten Informationspakete werden dann zum Einen an *pedocs* rückgeliefert und zum Anderen wieder als SIPs in das Archivierungssystem eingespielt.

Der Beginn der Testphase zur Übermittlung der Transferpakete an den Hotfolder steht unmittelbar bevor (September 2010). Die Prozesse zur Erstellung des SIP sind mit Daten externer *DP4lib*-Partner noch nicht durchgeführt worden, getestet wurde die Erstellung und Einspeisung von SIPs in das Archivsystem mit Daten der Deutschen Nationalbibliothek und der SUB Göttingen.

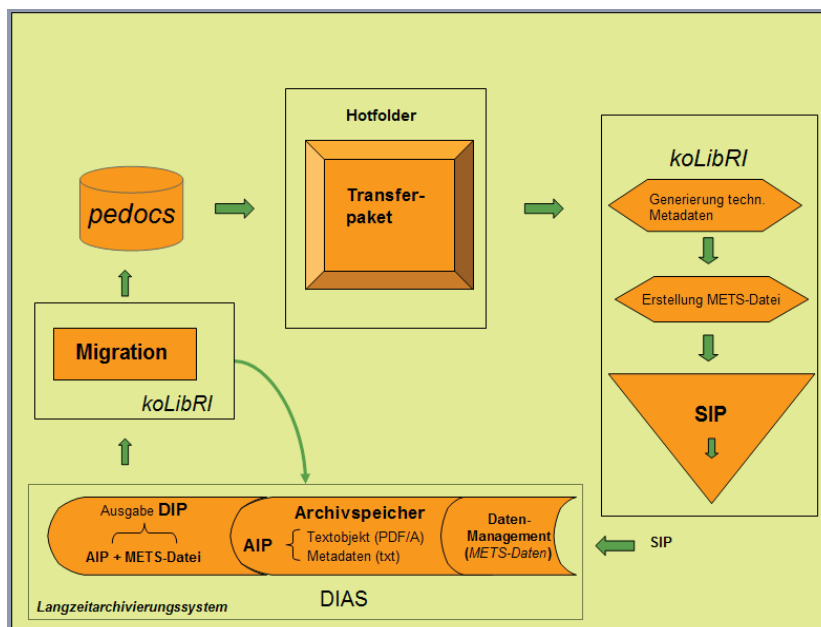


Abb.6: pedocs, koLibRI und DIAS⁷⁴

⁷³ vgl. Funk (2007), S.12 http://kopal.langzeitarchivierung.de/kolibri/koLibRI_v1_0_dokumentation.pdf

⁷⁴ Quelle: Dr. Julia Kreusch

6. Management der Langzeitarchivierung

Langzeitarchivierung ist nicht nur als technischer Prozess des Informationstransfers in das Archiv zu verstehen, sondern bringt besondere organisatorische und planerische Aufgaben mit sich. Im Folgenden sollen einige wichtige Managementaufgaben des Projekts LZA Pädagogik dargestellt werden, die hauptsächlich im Bereich der Schnittstelle Repositorium – Langzeitarchiv liegen.

6.1 Berichtswesen

Die Dokumentation aller durchgeführten Maßnahmen liefert die Datenbasis für fast alle Entscheidungen, die das Management in seinen verschiedenen Bereichen zu treffen hat. Indem sie jegliche Veränderung an den Objekten und Metadaten nachweist und damit nachvollziehbar macht, trägt sie zudem auf der Mikroebene entscheidend zur Sicherstellung der Authentizität des digitalen Objektes bei.

Die Datengrundlage des Berichtswesens erwächst aus den zu jedem Prozess erstellten Protokollen: für die Übertragung der Transferpakete an die Nationalbibliothek etwa enthalten diese spezifische Angaben über Zeit und Datum des Übernahmeprozesses, Ergebnisse der technischen Prüfungen und über valide und fehlerhafte Transferpakete. Die daraus erstellten Abnahmeberichte und müssen von beiden Parteien, Ablieferer und Empfänger, bestätigt werden. In Vorbereitung der ersten Testtransfers von Daten an die Nationalbibliothek werden derzeit zwischen *LZA Pädagogik* und der Deutschen Nationalbibliothek Form und Inhalte dieser Berichte, Übertragungswege und Aufbewahrungsdauern spezifiziert und als Übernahmevereinbarung einen Teil des abzuschließenden Vertrags bilden.

6.2 Qualitätsicherung

Qualitätsmanagement ist eine Querschnittsaufgabe, die alle Bereiche der Langzeitarchivierung umfasst und⁷⁵ die ständige Kontrolle von Prozessen, Standards und Zielen beinhaltet. Im Verbund der DP4lib-Partner werden Qualitätsziele und Maßnahmen der Qualitätsicherung zurzeit spezifiziert; unter anderem werden Kriterien ausgearbeitet, um die an den Archivierungsdienstleister gelieferten Objekte qualitativ bestimmen zu können.

LZA Pädagogik hat sich zum Ziel gesetzt, die Dokumente technisch so aufzubereiten, dass sie mit möglichst geringer Fehlerquote in das Langzeitarchiv eingespielt werden können. Dieser

⁷⁵ vgl. nestor8 (2008), S.13. <http://edoc.hu-berlin.de/series/nestor-materialien/2006-8/PDF/8.pdf>

Anspruch wird unter anderem dadurch realisiert, indem möglichst alle Objekte im Langzeitarchivierungsformat PDF/A an die Nationalbibliothek übertragen werden.

Dem Anforderungskatalog von *DP4lib* zufolge ist eine Ablieferung digitaler Textdokumente auch in anderen Formaten als PDF/A möglich.⁷⁶ Diese Formatvielfalt bedingt wiederum den Einsatz verschiedener Tools etwa zur Extraktion technischer Metadaten; folglich unterscheiden sich auch Funktionalitäten und Fehlermeldungen dieser Tools, da ihnen verschiedene Bewertungsschemata zugrunde liegen. Daher steht im Verbund der *DP4lib*-Partner eine Klassifikation der eingelieferten Objekte nach sogenannten ‚Ingest-Leveln‘ zur Diskussion: durch Zuweisung eines Levels wird der Bearbeitungsstand des Objekts bestimmt, und daraus kann der notwendige Archivierungsaufwand abgeleitet werden.⁷⁷ Damit würde ein einheitliches Schema zur Qualitätsbestimmung der gelieferten Objekte geschaffen.

Als weiteres Mittel der Qualitätsprüfung der Objekte dient die ‚Ingest-Policy‘. In dieser soll der Ablieferer dokumentieren, ab welchem Ingest-Level er die Aufnahme von Dateien in das Langzeitarchiv akzeptiert: der Dienstleister würde dazu den ermittelten Ingest-Level der Dateien gegen die in der policy festgelegten Werte prüfen. Den Ablieferern soll im Vorfeld eine Muster-Policy vorgelegt werden, die dann im Verlauf der Vertragsverhandlungen zwischen beiden Parteien spezifiziert und differenziert wird.

Eine Homogenisierung der Dateiformate als Maßnahme des Qualitätsmanagements, wie sie *LZA Pädagogik* bereits durchführt, ist auch von *DP4lib* vorgesehen. Um die anzuwendenden Maßnahmen der Langzeitarchivierung zu vereinfachen, soll nach der Bestimmung der Ingest-Level eine Konversion aller gelieferten Objekte in ein einheitliches Format stattfinden ("Migration vor der Migration").⁷⁸ Die Vereinheitlichung der Formate ist vor diesem Hintergrund auch als Teil des Kostenmanagements zu betrachten.⁷⁹

Eine weitere Aufgabe des Qualitätsmanagements ist es, für jeden Teilprozess Fehlerbehandlungen und Eskalationsverfahren⁸⁰ zu erarbeiten.

⁷⁶ DP4lib (2010c): Anforderungskatalog. 3.März 2010. Version 0.2 (Draft) (internes, unveröffentlichtes Papier)

⁷⁷ „Level 0: Objekt kann als Bitstream übernommen werden; Level 1: Level 0 + Objekt ist frei von Schutzmaßnahmen; Level 2: Level 1 + Generierung formatspezifischer Metadaten konnte durchgeführt werden; Level 3: Level 2 + Objekt ist valide und wohlgeformt hinsichtlich der Format-Spezifikation.“ DP4lib (2010b), S. 7

⁷⁸ DP4lib (2010b), S. 33f.

⁷⁹ vgl. Björk (2007), S.23 <http://eprints.ucl.ac.uk/7684/1/7684.pdf>

⁸⁰ „Delegation der Störungsbearbeitung an eine höhere Kompetenzstufe, wenn in der darunterliegenden Stufe

6.3 Planung der Langzeitarchivierung

Die fortwährende Beobachtung der Umgebung ist für die Sicherstellung der langfristigen Nutzbarkeit und Verfügbarkeit der gespeicherten Objekte unentbehrlich. Bestandteil des *Preservation Planning* ist daher die Ausarbeitung von Monitoring-Konzepten, um sogenannte Auslöser („Trigger-Events“) zu definieren, die dann Prozesse von Langzeiterhaltungsmaßnahmen einleiten. Während die Deutsche Nationalbibliothek das technische Monitoring übernimmt und seine Kooperationspartner über anstehende Migrationen berät, beobachtet *LZA Pädagogik* Veränderungen innerhalb der fachlichen Community, etwa hinsichtlich des Publikationsverhaltens und einer veränderten Nutzung von Publikationsformen. Das Monitoring dient somit als Grundlage der Erarbeitung von Migrationsplänen für obsoletere Formate und spielt auch eine Rolle bei der Definition der Inhalte von Einlieferungs-, Archiv- und Ausgabepaketen (SIP, AIP, DIP).

Zu den Aufgaben des *Preservation Planning* gehört zudem das Monitoring von Kosten und Risiken. *kopal* hat für die Implementierung seiner ‚*kopal*-Solution‘ eine Übersicht über Kostenfaktoren sowie eine exemplarische Kostenkalkulation für die verschiedenen Nutzungsmodelle vorgelegt.⁸¹ Da die Finanzierung der Projekte *DP4lib* und *LZA Pädagogik* durch die Deutsche Forschungsgemeinschaft noch gesichert ist, stehen mögliche Kostenmodelle derzeit noch im Hintergrund; eine Beschäftigung mit diesem Aufgabenfeld wird aber im Hinblick auf die Zeit nach Projektende relevant werden.

Einen Ansatz zum Kostenmanagement für die Langzeiterhaltung elektronischer Ressourcen erarbeitete das Forschungsprojekt *LIFE*;⁸² den Ergebnissen des Projekts zufolge spielt für die Langzeitarchivierung neben den Kosten für Archivspeicher und Personal der Faktor ‚Formatvielfalt‘ eine entscheidende Rolle: die Archivierung von Dokumenten in den gängigen Formaten *PDF*, *TXT* und *HTML*, die ca. 85% der Gesamtmenge aller Dateien umfassten, habe während der Versuchsphase 7% aller Kosten verursacht, während 41% der Kosten auf die zwölf am wenigsten verwendeten Formate, die 0,1% der Dateien abdeckten, zurückgingen.⁸³

Eine Analyse von Kosten der Langzeitarchivierung für den Bereich von Forschungsdaten, die auf einer Befragung von 13 hauptsächlich in Großbritannien ansässigen Institutionen beruhte, kam zu dem Ergebnis, dass hohes Potential zur Kosteneinsparung in der Automatisierung von

keine Korrektur möglich ist.“ <http://de.wikipedia.org/wiki/Eskalation>

⁸¹ vgl. *kopal* (2007), S.2 http://kopal.langzeitarchivierung.de/downloads/kopal_Services_2007.pdf

⁸² Life Cycle Information for E-Literature (LIFE) <http://www.life.ac.uk/>

⁸³ vgl. Björk (2007), S.23 <http://eprints.ucl.ac.uk/7684/1/7684.pdf>

Prozessen des Ingest und Access liege, also derjenigen Schnittstellen, an denen sich *pedocs* befindet.⁸⁴

Der *DP4lib*-Partner Deutsche Nationalbibliothek diskutiert zur Automatisierung des Bereichs *Preservation Planning* derzeit die Nachnutzung des Planungstools ‚Plato‘. Dabei handelt es sich um eine konkrete Ausarbeitung des Moduls *Preservation Planning* des Referenzmodells OAIS.⁸⁵ Plato steuert die Kernphasen von Anforderungserhebung, Auswahl von in Frage kommenden Strategien und deren Evaluierung, die Entscheidungen zur Auswahl der optimalen Strategie für Archivierungsmaßnahmen und unterstützt die Erstellung von Migrationsplänen.

6.4 Rechtliches

Zum Zeitpunkt des Transfers von Objekten an den *DP4lib*-Partner Deutsche Nationalbibliothek ist die Rechtmäßigkeit der Publikation der Dokumente und ihrer Weitergabe zum Zweck der Langzeitarchivierung gewährleistet.

In Verträgen mit den Autoren und Verlagen lässt sich *pedocs* gemäß dem Urheberrecht das einfache Nutzungsrecht an den Dokumenten übertragen. Dies beinhaltet, „das Werk auf die erlaubte Art zu nutzen, ohne dass eine Nutzung durch andere ausgeschlossen ist.“⁸⁶ Der Umfang des Nutzungsrechts wird in Autoren- und Verlagsverträgen festgelegt und umfasst die elektronische Speicherung zur internen Verwaltung, das Verfügbarmachen für die Öffentlichkeit zum individuellen Abruf, zur Wiedergabe auf dem Bildschirm und zum Ausdruck bei der Online-Nutzung sowie das Recht zur Konvertierung des Dateiformats zum Zwecke der Langzeitarchivierung bzw. der Visualisierung.⁸⁷ Mit der Sicherung des Rechts auf Formatkonversion ist eine der Grundanforderungen an ein Langzeitarchiv, nämlich die technische Kontrolle über das Objekt zu besitzen, erfüllt.⁸⁸ Möglichen daraus folgenden Änderungen des Erscheinungsbilds des Dokuments stimmt der Autor zu. Die inhaltliche Integrität des Werks wird von *pedocs* zugesichert,⁸⁹ der Gewährleistungsumfang ist jedoch eingeschränkt, wo es um die Volltexterkennung durch OCR geht.⁹⁰ Damit sind die rechtlichen Grundlagen für den Transfer zu archivierenden Objekte an den Archivierungsdienstleister geschaffen.

⁸⁴ vgl. Beagrie (2010), S.5 <http://www.jisc.ac.uk/media/documents/publications/reports/2010/keepingresearchdatasafe2.pdf>

⁸⁵ vgl. Neuroth (2010), Kap. 12:18ff. <http://nbn-resolving.de/urn:nbn:de:0008-2010071949>

⁸⁶ § 31 Abs. 2 UrhG. http://bundesrecht.juris.de/urhg/_31.html

⁸⁷ vgl. http://www.pedocs.de/informationen_autoren.html.

⁸⁸ vgl. nestor8 (2008), S.28. <http://edoc.hu-berlin.de/series/nestor-materialien/2006-8/PDF/8.pdf>

⁸⁹ *pedocsA*, §2b <http://www.pedocs.de/autorenvertrag.php?la=de>

⁹⁰ siehe Kap.5. OCR und Authentizität

Die im vorigen benannten Aufgaben, Prozesse und die Festlegung der Verantwortlichkeiten innerhalb des kooperativen Workflows werden nach Abschluss der Erprobungsphasen Bestandteile einer vertraglichen Regelung zwischen *LZA Pädagogik* und dem *DP4lib-Partner* Deutsche Nationalbibliothek sein. Derzeit werden Spezifikationen ausgearbeitet, die den Arbeitsablauf klären und Zuständigkeiten verteilen und später als Leistungsbeschreibungen dem angestrebten Vertrag beigelegt werden. Bis zum Ende des Projektes soll der Vertrag ausgearbeitet zur Unterzeichnung vorliegen und als Grundlage für die Aufnahme des Routinebetriebs dienen.

7. Ausblick

Mit der Integration einer Lösung für die Langzeitarchivierung digitaler Objekte in das Repositorium *pedocs* vollzieht das *Fachportal Pädagogik* einen wichtigen Schritt hin zur Gewährleistung der Nachhaltigkeit wissenschaftlichen Publizierens im virtuellen Raum. Da das Projekt den kooperativen Workflow derzeit ausgestaltet, können hier nur einige Handlungsfelder für den weiteren Projektverlauf umrissen werden.

So ist die Definition von ‚erhaltenswerten Eigenschaften‘ auf lange Sicht für den Betrieb eines Langzeitarchivs unerlässlich, da digitale Objekte möglicherweise nicht gesamtvolumfänglich erhalten werden können, sondern eine teilweise Obsoleszenz zu akzeptieren sein wird.⁹¹ Erst mit einer Bestimmung dieser Eigenschaften wird auch ein endgültiges Erhaltungsziel definiert werden können. Realisiert werden kann dies allerdings erst nach weiteren Tests und umfassenderen Erfahrungen mit der Migration. Da die Definition dessen, was erhaltenswert ist, auch von der Nutzergruppe abhängig ist,⁹² sollte *LZA Pädagogik* im Zuge seines fachlichen Monitoring auch diesen Aspekt einbeziehen und die Interessen seiner avisierten Nutzergruppe in die Diskussion mit dem Projektpartner einbringen.

In den Verträgen mit den Autoren und Verlagen lässt sich *pedocs* unter Zusicherung der inhaltlichen Integrität pauschal das Recht zur Veränderung des Objekts zusprechen. Perspektivisch wäre zu bedenken, Autoren ein Mitbestimmungsrecht bei der Definition derjenigen Eigenschaften, die an den zu archivierenden Objekten erhalten werden sollen, einzuräumen bzw. differenziert über die Maßnahmen der Langzeitarchivierung und mögliche Auswirkungen auf die Gestalt der Objekte zu informieren.

Erklärtes Ziel von *LZA Pädagogik* ist es, einen nachnutzbaren Workflow zu entwickeln; dazu sollen die Erfahrungen aus dem Projekt in einem Wiki veröffentlicht werden. Die besonderen Bedingungen, unter denen *LZA Pädagogik* agiert, müssen zu einer Einschätzung einer tatsächlichen Nachnutzbarkeit aber berücksichtigt werden: da das Repositorium *pedocs* sich noch immer im Aufbau befindet, hat es zum jetzigen Zeitpunkt noch kein allzu hohes Datenaufkommen und damit kein Mengenproblem; die Automatisierung von Arbeitsprozessen steht somit noch nicht dringlich zur Debatte. Außerdem stellt *LZA Pädagogik* hohe Ansprüche an

⁹¹ nestor10 (2008), S.8. http://files.d-nb.de/nestor/materialien/nestor_mat_10.pdf

⁹² Dappert, Angela et al. (2009).

http://www.planets-project.eu/docs/papers/Dappert_SignificantCharacteristics_ECDL2009.pdf

die Qualität der abzuliefernden Daten und schafft daher entsprechende Kapazitäten zur händischen Bearbeitung der Objekte.

Die Entwicklung eines tragfähigen Finanzierungskonzepts über das Ende des Projekts hinaus ist sicherlich eine der wichtigsten Aufgaben und eine Grundbedingung einer nachhaltigen Langzeitarchivierung. Da durch die Open-Access-Strategie eine Gebühr für den Zugang zu Informationen von vornherein ausgeschlossen ist, ist es umso dringlicher, ein tragfähiges Konzept zur Finanzierung über Mittelbeschaffung zu entwickeln.⁹³ Neben der Sicherstellung der Finanzierung sind auch die internen Arbeitsprozesse zu evaluieren und mögliche Einsparpotentiale durch die Automatisierung einzelner Elemente zu prüfen.

Ein Meilenstein für *LZA Pädagogik* und *DP4lib* wird die anstehende Migration der Datenobjekte sein, die eine erste Bewertung der bisherigen Arbeiten ermöglichen wird. Erst weitere Praxiserfahrungen mit der Archivierung und Rücklieferung an das Repository werden es dann erlauben, sich der Differenzierung des Workflows und der Weiterentwicklung kooperativer Aufgabenfelder zuzuwenden. Trotz aller noch zu lösenden Aufgaben und vielschichtigen Gelingensbedingungen kann das kooperative Modell von *LZA Pädagogik* und *DP4lib* diejenigen Synergien freisetzen, um der Herausforderung „Langzeitarchivierung“ erfolgreich zu begegnen.

⁹³ BRTF (2010), S.55 http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf

8. Literaturverzeichnis

Die verzeichneten Links wurden alle das letzte Mal am 30.08.2010 überprüft.

- Allinson (2006). Allinson, Julie: *OAIS as a reference model for repositories. An evaluation.* Version 0.5. Digital Repositories Programme Support Team. UKOLN, University of Bath, 2006.
<http://www.ukoln.ac.uk/repositories/publications/oais-evaluation-200607/Drs-OAIS-evaluation-0.5.pdf>
- Beagrie (2008a). Beagrie, Neil et al.: *Preservation Management of digital materials : the handbook.*
http://www.dpconline.org/component/docman/doc_download/299-digital-preservation-handbook
- Beagrie (2008b). Beagrie, Neil et al.: *Keeping Research Data Safe. A cost model and guidance for UK universities.* A study funded by JISC. Final Report, April 2008.
<http://www.jisc.ac.uk/media/documents/publications/keepingresearchdatasafe0408.pdf>
- Beagrie (2010). Beagrie, Neil et al.: *Keeping Research Data Safe 2.* A study funded by JISC. With support from OCLC Research and the UK Data Archive. Final Report, April 2010.
<http://www.jisc.ac.uk/media/documents/publications/reports/2010/keepingresearchdatasafe2.pdf>
- Björk (2007). Björk, Bo-Christer: *Evaluation of the costing activities and economic models for digital curation using the LIFE methodology.* Report commissioned by the LIFE-2 project consortium. London, 2007. <http://eprints.ucl.ac.uk/7684/1/7684.pdf>
- BRTF (2010). Blue Ribbon Task Force on Sustainable Digital Preservation and Access: *Sustainable economics for a digital planet : ensuring long-term access to digital information.* La Jolla, Kalifornien, Februar 2010. http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf
- Borghoff (2005). Borghoff, Uwe et al.: *Vergleich bestehender Archivierungssysteme.* nestor-Materialien 3. München, 2005. http://files.d-nb.de/nestor/materialien/nestor_mat_03.pdf
<http://nbn-resolving.de/urn:nbn:de:0008-20050117016>
- Consultative Committee for Space Data Systems (CCSDS). Homepage.
<http://public.ccsds.org/default.aspx>
- CCSDS (2009). Consultative Committee for Space Data Systems (CCSDS): *Reference Model for an Open Archival Information System (OAIS).* Pink Book, Issue I.1. Draft Recommended Standard. Washington DC, 2009. http://ddp.nist.gov/refs/650x0p11_OAIS_pink_book.pdf
- CLIR (2002). Council on Library and Information Resources Washington, D.C. (Hrsg.): *The state of digital preservation. An international perspective.* Conference proceedings. Documentation Abstracts, Inc. Institutes for Information Science, Washington D.C., April 24-25, 2002. (=CLIR Report 107). Washington, 2002.
<http://www.clir.org/pubs/reports/pub107/pub107.pdf>
- Conyers (1990). Conyers, John jr.: *Taking a byte out of history: the archival preservation of federal computer records.* Twenty-fifth report by the Committee of government operations. Washington, 1990 (House Report 101-978)
<http://www.bobgellman.com/rg-docs/Taking%20a%20Byte%20Out%20of%20History.PDF>
- Dappert (2009). Dappert, Angela et al.: *Significance is in the Eye of the Stakeholder.* Wetherby, 2009.
http://www.planets-project.eu/docs/papers/Dappert_SignificantCharacteristics_ECDL2009.pdf

- Der Siebdruck - Das Magazin der Druck- und Medienindustrie für Technik und Werbung.
Fachbegriffe. <http://www.der-siebdruck.de/index.php?id=53>
- DFG (2009). Deutsche Forschungsgemeinschaft: *DFG-Praxisregeln "Digitalisierung"*. Bonn, 2009.
http://www.dfg.de/download/pdf/foerderung/programme/lis/praxisregeln_digitalisierung.pdf
- Deutsches Institut für Internationale Pädagogische Forschung (DIPF). Homepage.
<http://www.dipf.de>
- Kooperative Langzeitarchivierung erziehungswissenschaftlicher e-Ressourcen im Rahmen von kopal. Projektseite.
<http://www.dipf.de/de/projekte/foerderung-des-modellvorhabens-lza-paedagogik>
- Deutsche Initiative für Netzwerkinformation e.V. (DINI) Homepage. <http://www.dini.de>
- DINI (2010). Deutsche Initiative für Netzwerkinformation e.V. (DINI): *DINI-Zertifikat. Dokumenten- und Publikationsservice 2010*. Version 3.0. <http://nbn-resolving.de/urn:nbn:de:kobv:11-100109986>.
- DNB (2009). Deutsche Nationalbibliothek: *Sammelrichtlinien*. Stand: 1.Juni 2009. Leipzig [u.a.], 2009. <http://files.d-nb.de/pdf/sammelrichtlinien.pdf>. <http://nbn-resolving.de/urn:nbn:de:101-2009033003>.
- Deutsche Nationalbibliothek. URN-Service. http://www.d-nb.de/netzpub/erschl_lza/np_urn.htm
- Digital Preservation for Libraries - DP4lib. Homepage. <http://dp4lib.langzeitarchivierung.de/>
- DP4lib (2010a). Digital Preservation for Libraries (DP4lib): *Spezifikation von Transferpaketen und deren Übertragung an einen LZA-Dienstleister mittels eines Hotfolders*. Version 1.0, 27.04.2010. (unveröffentlichtes, internes Papier)
- DP4lib (2010b). Digital Preservation for Libraries (DP4lib): *Spezifikation zur Übernahme von Transferpaketen in ein Langzeitarchiv*. Version 1.0, 19.07.2010. (unveröffentlichtes, internes Papier)
- DP4lib (2010c). Digital Preservation for Libraries (DP4lib): *Anforderungskatalog*. 3.März 2010. Version 0.2 (Draft). (unveröffentlichtes, internes Papier)
- Eckert (2006). Eckert, Claudia: *IT-Sicherheit. Konzepte – Verfahren – Protokolle*. 4., überarbeitete Auflage. München [u.a.], 2006.
- EDitEUR: ONIX Homepage. <http://www.editeur.org/8/ONIX/>
- Fachportal Pädagogik. Homepage. <http://www.fachportal-paedagogik.de/start.html>
- Funk (2007). Funk, Stefan et al.: *kopal Library for Retrieval and Ingest. Dokumentation*. Version 1.0. Göttingen, 2007. http://kopal.langzeitarchivierung.de/kolibri/koLibRI_v1_0_dokumentation.pdf
- Grace (2009). Grace, Stephen; Knight, Gareth; Montague, Lynne: *InSPECT. Investigating the Significant Properties of Electronic Content over Time*. Draft Final Report, Version 1.0. London, 2009. <http://www.significantproperties.org.uk/inspect-finalreport.pdf>.
- IBM Implementation Services. http://www-935.ibm.com/services/nl/dias/is/implementation_services.html
- InSPECT. Investigating Significant Properties of Electronic Content. Homepage.

<http://www.significantproperties.org.uk/>

ISO 19005-1:2005/Cor 1:2007. *Document management – Electronic document file format for long-term preservation – Part 1: Use of PDF 1.4 (PDF/A-1)*. Veröffentlicht am 28.03.2007.

http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=45613

ISO 20652:2006. *Space data and information transfer systems - Producer-archive interface - Methodology abstract standard*.

http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=39577

JSTOR; Harvard College: JHOVE - JSTOR/Harvard Object Validation Environment. Homepage.

<http://hul.harvard.edu/jhove/>

kopal: Kooperativer Aufbau eines Langzeitarchivs digitaler Informationen. Homepage.

<http://kopal.langzeitarchivierung.de/>

koLibRI. kopal Library for Retrieval and Ingest.

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

kopal (2007). kopal: *Ein Service für die Langzeitarchivierung digitaler Informationen*. 2007.

http://kopal.langzeitarchivierung.de/downloads/kopal_Services_2007.pdf

Life Cycle Information for E-Literature (LIFE). Homepage. <http://www.life.ac.uk/>

Max Planck-Ges. (2003). Max Planck-Gesellschaft: *Berliner Erklärung über den offenen Zugang zu wissenschaftlichem Wissen*. Berlin, 2003.

http://oa.mpg.de/openaccess-berlin/Berliner_Erklaerung_dt_Version_07-2006.pdf

METS – Metadata Encoding and Transmission Standard. Homepage <http://www.loc.gov/standards/mets/>

Müller (2007). Müller, Uwe; Schirnbacher, Peter: *Der „Grüne Weg zu Open Access“ in Deutschland*. In: Zeitschrift für Bibliothekswesen und Bibliographie, 4-5 (2007), S.183-193.

NABD (2010). NA 009 Normenausschuss Bibliotheks- und Dokumentationswesen (NABD): *DIN 31644. Information und Dokumentation – Kriterien für vertrauenswürdige digitale Langzeitarchive*. 2010.

<http://www.nabd.din.de/cmd?level=tpl-art-detailansicht&committeeid=54738855&artid=133059734&bcrumblevel=1&languageid=de>

nestor - Das Kompetenznetzwerk Langzeitarchivierung und Langzeitverfügbarkeit digitaler Ressourcen in Deutschland. Homepage. <http://www.langzeitarchivierung.de/index.htm>

nestor10 (2008). nestor Arbeitsgruppe Standards für Metadaten, Transfer von Objekten: *Wege ins Archiv. Ein Leitfaden für die Informationsübernahme in das digitale Langzeitarchiv*.

Version I: Entwurf zur öffentlichen Kommentierung. nestor-Materialien 10. Göttingen, 2008.

http://files.d-nb.de/nestor/materialien/nestor_mat_10.pdf

<http://nbn-resolving.de/urn:nbn:de:0008-2008103009>.

nestor8 (2008). nestor Arbeitsgruppe Vertrauenswürdige Archive – Zertifizierung: *Kriterienkatalog vertrauenswürdige digitale Langzeitarchive*. Version 2. nestor-Materialien 8). Frankfurt am Main, 2008. http://files.d-nb.de/nestor/materialien/nestor_mat_08.pdf

<http://nbn-resolving.de/urn:nbn:de:0008-2008021802>

Neuroth (2010). Neuroth, H. et al.: *nestor Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung*. Version 2.3. Göttingen, 2010.

http://nestor.sub.uni-goettingen.de/handbuch/nestor-handbuch_23.pdf

<http://nbn-resolving.de/urn:nbn:de:0008-2010071949>

Parschik (2006). Parschik, Thomas: *Durchführung von Digitalisierungsprojekten in Bibliotheken*. In: Bibliotheksdienst, 40.Jg. (2006), H.12, S. 1421-1443.

http://www.zlb.de/aktivitaeten/bd_neu/heftinhalte2006/Technik011206.pdf

pedocs Dokumentenserver. <http://www.pedocs.de/>

pedocs (2010). pedocs: *Digitalisierung*. Unveröffentlichtes, internes Papier.

pedocsA. pedocs: *Autorenvertrag*. <http://www.pedocs.de/autorenvertrag.php?la=de>

pedocsV. pedocs: *Verlagsvertrag*. http://www.pedocs.de/verlagsvertrag_muster.php?la=de

Wikipedia. Homepage. <http://www.wikipedia.de>

Gesetze

BMJ (2008a). Bundesministerium der Justiz: *Verordnung über die Pflichtablieferung von Medienwerken an die Deutsche Nationalbibliothek (Pflichtablieferungsverordnung - PflAV)*. Ausfertigungsdatum: 17.10.2008. <http://bundesrecht.juris.de/bundesrecht/pflav/gesamt.pdf>.

BMJ (2008b). Bundesministerium der Justiz: *Gesetz über Urheberrecht und verwandte Schutzrechte (Urheberrechtsgesetz)*. Ausfertigungsdatum: 17.12.2008. <http://www.gesetze-im-internet.de/urhg/BJNR012730965.html>.

Bildquellen

Abbildung 1: Das *kopal*-Archivsystem mit koLibRI und DIAS

Die *kopal*-Broschüre (Juni 2006), S.2.

http://kopal.langzeitarchivierung.de/downloads/kopal_Broschuere_2006.pdf

Abbildung 2: Funktionale Entitäten und Informationspakete eines OAIS

Consultative Committee for Space Data Systems (CCSDS): *Reference Model for an Open Archival Information System (OAIS)*. Pink Book. Draft Recommended Standard. Washington D.C., 2009.

http://ddp.nist.gov/refs/650x0p11_OAIS_pink_book.pdf.

Abbildung 3: Vollanzeige eines Datensatzes in *pedocs*.

http://www.pedocs.de/frontdoor.php?source_opus=2665&la=de

Abbildung 6: *pedocs*, koLibRI und DIAS

Quelle: Dr. Julia Kreuzsch

Fischer, Natalie; Brümmer, Felix; Kuhn, Hans Peter; Züchner, Ivo

**Individuelle Wirkungen des Ganztags schulbesuchs in der Sekundarstufe.
Erkenntnisse aus der Studie zur Entwicklung von Ganztags schule (StEG)**

Schulverwaltung. Hessen, Rheinland-Pfalz 15 (2010) 2, S. 38-39

urn:nbn:de:0111-opus-26658

Erstveröffentlichung bei:



Carl Link

eine Marke von Wolters Kluwer Deutschland

www.carllink.de

Nutzungsbedingungen

pedocs gewährt ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit dem Gebrauch von pedocs und der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Kontakt:

peDOCS

Deutsches Institut für Internationale Pädagogische Forschung (DIPF)

Informationszentrum (IZ) Bildung

Schloßstr. 29, D-60486 Frankfurt am Main

eMail: pedocs@dipf.de

Internet: www.pedocs.de

Individuelle Wirkungen des Ganztagsschulbesuchs in der Sekundarstufe

Erkenntnisse aus der Studie zur Entwicklung von Ganztagschule (StEG)

Der Beitrag konzentriert sich auf die ersten beiden Erhebungswellen von StEG. Zusammenfassend werden die Resultate zu individuellen Wirkungen der Teilnahme an Ganztagsschulangeboten in der Sekundarstufe dargestellt. Insgesamt zeigen sich kleine positive Effekte des Ganztagsschulbesuchs.

**Natalie Fischer
Felix Brümmer
Hans Peter Kuhn
Ivo Züchner**

**Deutsches Institut für Internationale
Pädagogische Forschung (DIPF), Frankfurt**

In den letzten Jahren wurden zahlreiche Unterstützungsprogramme ins Leben gerufen, um die Zahl der Ganztagschulen in Deutschland substanziell zu erhöhen und die Qualität ihrer Arbeit zu verbessern. Die Erwartungen sind hoch und beziehen sich unter anderem auf die Kompensation sozialer Benachteiligung und die Integration von Kindern mit Migrationshintergrund. Soziales Lernen, Motivation und das akademische Selbstkonzept sollen ebenso gefördert werden wie der Erwerb kognitiver Kompetenzen. Im folgenden Beitrag wird anhand des Vergleichs zwischen am Ganztagsbetrieb teilnehmenden und nicht-teilnehmenden Schülerinnen und Schülern überprüft, ob sich entsprechende Wirkungen nachweisen lassen. Berichtet werden Ergebnisse der Studie zur Entwicklung von Ganztagschulen (StEG), die den Auf- und Ausbau von Ganztagschulen seit 2005 evaluiert. Es wird analysiert, wie sich die befragten Sekundarstufenschüler/-innen in Abhängigkeit von der Teilnahme an Ganztagsangeboten von 2005 bis 2007 entwickelt haben.

Teilnahme an Ganztagsangeboten

Einleitend muss erläutert werden, was in Ganztagschulen der Sekundarstufe I unter einem »Teilnehmer«/einer »Teilnehmerin« verstanden werden darf. Im Gegensatz zur Grundschule ist die Teilnahmeintensität wesentlich geringer: Nur zwischen 40 und 50 % der Schüler/-innen nehmen an mindestens drei

Tagen in der Woche am Ganztagsbetrieb teil, viele »spielen im Ganztagsbetrieb ihrer Schule eine Gastrolle« (Klieme et al. 2007, S. 359). Betrachtet man die Teilnahme an Ganztagsangeboten über zwei Jahre, so nehmen nur 12 % aller befragten Schüler/-innen dauerhaft für mindestens drei Tage in der Woche an Ganztagsangeboten teil. Dieser Prozentsatz erhöht sich mit dem Bindungsgrad der Schule. Die StEG-Ergebnisse zeigen jedoch, dass selbst in voll gebundenen Ganztagschulen längst nicht alle Schüler an mindestens drei Tagen am Ganztagsbetrieb teilnehmen. Insofern sind Ganztagschüler/-innen im Sekundarbereich nur schwer von Halbtagschüler(n)/-innen, die außerunterrichtliche Zusatzangebote besuchen, zu unterscheiden. Was bedeutet das nun für die erhofften Wirkungen der Ganztagschule?

Betrachtet man US-amerikanische Ergebnisse zur Wirkung von extracurricularen Angeboten, so haben sich Dauer und »Dosierung« des Besuchs als wichtige Einflussgrößen gezeigt. Bei der in der Sekundarstufe vorgefundenen geringen Teilnahmeintensität darf man also nur mit geringen Auswirkungen rechnen. Wir gehen außerdem davon aus, dass Wirkungen der Angebote von spezifischen Qualitätsmerkmalen abhängen. Insbesondere Angebote, die von den Schüler(n)/-innen als gut strukturiert erlebt werden, in denen sie sich sozial gut aufgehoben fühlen, die ihre Interessen ansprechen und Partizipation zulassen sollten positive Auswirkungen auf die Entwicklung der Jugendlichen haben (vgl. Fischer, Radisch & Stecher 2009).

Im Folgenden werden Effekte der Teilnahme an Ganztagsangeboten sowie Einflüsse der Teilnahmeintensität und der Angebotsqualität berichtet.

Wirkungen auf Schulnoten

In StEG wurden – anders als in Schulleistungsstudien wie z.B. PISA – keine objektiven Leistungsdaten erhoben. Stattdessen wurden von den Schüler(n)/-innen die Zeugnisnoten in den Hauptfächern erfragt. Schulnoten sollten besonders sensitiv bezüglich der Wirkung einer Teilnahme an Ganztagsangeboten sein, da ein besonderes Potenzial der Angebote in der Förderung fachübergreifender Kompetenzen liegt, die neben den kognitiven Leistungen in die Benotung einfließen (vgl. Klieme 2003).

Anhand der StEG-Daten kann gezeigt werden, dass sich die reine Teilnahme an Ganztagsangeboten der Schule auf die Entwicklung der Noten in Deutsch und Mathematik nach dem Übergang in die Sekundarstufe auswirkt (Fischer, Kuhn & Klieme 2009). Die Noten der Schüler/-innen, die am Ganztagsbetrieb teilnehmen, entwickeln sich über zwei Jahre etwas positiver, als die ihrer nicht teilnehmenden Mitschüler/-innen. Weitere Analysen zeigen, dass im Hinblick auf die Notenentwicklung insbesondere der dauerhafte Besuch von Ganztagsangeboten (z.B. über zwei Jahre) von Vorteil ist, hingegen lässt sich kein Effekt der Teilnahmeintensität in Tagen pro Woche finden. Bei Gymnasiasten hat darüber hinaus die wahrgenommene Qualität des besuchten Ganztagsangebotes einen Einfluss auf die Entwicklung der Deutschnote.

Wirkungen auf die Motivation

Auch die Lernmotivation wird vom Besuch der Ganztagsangebote beeinflusst. Empirische Ergebnisse zeigen im Allgemeinen eine Verschlechterung der Lernmotivation nach dem Übergang in die Sekundarstufe, das kann auch in der StEG-Stichprobe festgestellt werden. Allerdings verläuft dieser Motivationsabfall bei den Schüler(n)/-innen, die das Ganztagsangebot ihrer Schule nutzen, weniger rasant als bei ihren Mitschüler(n)/-innen. Insofern wirken Ganztagsangebote protektiv im Hinblick auf den Anstieg der Schulunlust nach dem Über-



gang – als wesentliche Einflussgröße hat sich hier die Angebotsqualität erwiesen. Gut strukturierte Angebote, die sich am Vorwissen und den Bedürfnissen der Schüler/-innen orientieren und ein positives Sozialklima bieten, sind besonders geeignet, die Lernmotivation aufrecht zu erhalten (Fischer, Kuhn & Klieme 2009).

Wirkungen auf das Sozialverhalten

Bezogen auf das Sozialverhalten können ebenfalls Wirkungen nachgewiesen werden. Für alle Schüler/-innen gilt, dass die Teilnahme an einem qualitativ hochwertigen Angebot aggressives Verhalten vermindern kann. Auch ohne Berücksichtigung der Angebotsqualität lässt sich feststellen, dass Schüler/-innen, die dauerhaft an Ganztagsangeboten teilnehmen, weniger aggressives Verhalten berichten als ihre Mitschüler/-innen. Insbesondere scheint sich der Besuch von fachübergreifenden Arbeitsgemeinschaften und Freizeitangeboten positiv auf das Sozialverhalten auszuwirken.

Wirkungen spezifischer Angebote

Da es sich bei StEG um eine breit angelegte Schulentwicklungsstudie handelt, konnten hier keine detaillierten Informationen zu spezifischen Angeboten der Ganztagschulen eingeholt werden. Einzig die Hausaufgabenbetreuung als Angebot, das an fast allen Schulen durchgeführt wird, wurde näher untersucht. Die Schüler/-innen, die an diesem Angebot teilnehmen, entwickeln sich hinsichtlich Motivation und Schulnoten nicht besser aber auch nicht schlechter als ihre Mitschüler/-innen. Betrachtet man jedoch Qualitätsmerkmale genauer, so lässt sich feststellen, dass der Besuch einer gut strukturierten Hausaufgabenbetreuung, in der eine effiziente Zeitnutzung erfolgt, positive Auswirkungen hat. Hier profitieren insbesondere Schüler/-innen mit Migrationshintergrund hinsichtlich der Noten und des akademischen Selbstkonzepts (Kuhn u.a. in Vorbereitung). Außerdem zeigt eine angegliederte Studie, dass eine gut strukturierte Hausaufgabenbetreuung den Schüler(n)/-innen helfen kann, die Motivation aufrecht zu erhalten, was letztlich zu besseren Lernergebnissen bei den Hausaufgaben führen kann.

Wirkungen auf Freizeitverhalten und Familie

In StEG wurden auch Auswirkungen des Ganztagschulbesuchs auf Freizeitgestaltung und Familie betrachtet. Insgesamt muss konstatiert werden, dass sich Ganztags Teilnehmer/-innen hinsichtlich ihrer Freizeitgestaltung nicht wesentlich von ihren Mitschüler(n)/-innen unterscheiden. Sie gehören durchschnittlich etwa eben so vielen Vereinen an und unternehmen in ihrer Freizeit ähnliche Dinge. Auch hinsichtlich des Familienklimas und der Unternehmungen mit der Familie ergeben sich wenige Unterschiede. Insgesamt scheint die dauerhafte Teilnahme an Ganztagsangeboten eher dazu beizutragen, dass das Familienklima (welches sich über die Jugendzeit hinweg normalerweise verschlechtert) stabil gehalten wird. Eltern fühlen sich durch die Ganztagschule bei der Hausaufgabenunterstützung und in Bezug auf erzieherische Probleme entlastet. Zudem zeigt sich, dass sich Eltern von Ganztags Schüler(n)/-innen besonders aktiv an schulischen Aktivitäten und Gremien beteiligen.

Ausblick

Vor dem Hintergrund, dass viele Schüler/-innen nur wenige Angebote pro Woche wahrnehmen, sollten in Zukunft spezifische Angebote und ihre Wirkpotenziale detaillierter betrachtet werden. Die von den Schüler(n)/-innen wahrgenommene Angebotsqualität scheint ein wichtiger Faktor im Hinblick auf die Entwicklung der Motivation und des Sozialverhaltens zu sein – hier stellt sich die Frage, wie ein qualitativ hochwertiges Angebot gestaltet sein muss. Was können Lehrkräfte und Pädagogen dazu beitragen, dass sich die Schüler in der Schule wohl fühlen, die Lernmotivation und letztlich die Lernleistung gesteigert werden kann? Erste Analysen in Bezug auf die Hausaufgabenbetreuung zeigen die Wichtigkeit einer strukturierten Lernumgebung. Mutmaßlich sind für unterrichtsferne Angebote andere Qualitätskriterien bedeutsamer. Hierzu kann StEG nur relativ global Auskunft geben, weitere detaillierte Studien sind nötig. Mit der Auswertung der dritten Erhebungswelle, deren Daten Ende 2009 vorliegen werden, erwarten wir neue

Erkenntnisse über langfristige Wirkungen. Wir erhoffen uns unter anderem Hinweise darauf, was die Ganztagschule tun kann, um auch ältere Schüler/-innen dauerhaft anzusprechen

Fazit

In der breit angelegten Schulentwicklungsstudie StEG lassen sich erste Hinweise auf individuelle Wirkungen von Ganztagsangeboten finden. Das ist bemerkenswert, da diese Frage nicht im Vordergrund der Studie steht und Angebotsmerkmale deshalb nur relativ grob erhoben wurden. In Bezug auf Freizeitverhalten und Familienleben der Schüler/-innen scheint der Ganztagschulbesuch keine grundlegenden Umgestaltungen zur Folge zu haben. Insbesondere beim Übertritt in die Sekundarstufe I können protektive Wirkungen der Teilnahme an Ganztagsangeboten im Hinblick auf die Entwicklung von Lernmotivation und Schulnoten nachgewiesen werden. Aber auch für ältere Schüler/-innen gilt: Der dauerhafte Besuch von qualitativ hochwertigen Angeboten der Ganztagschule kann sich positiv auf das Sozialverhalten, die Motivation und schulische Leistungen auswirken.

Literatur

Fischer, Kuhn, Klieme (2009). Was kann die Ganztagschule leisten? Wirkungen ganztägiger Beschulung auf die Entwicklung von Lernmotivation und schulischer Performanz nach dem Übergang in die Sekundarstufe. In Stecher, Allemann-Ghionda u.a., Ganztägige Bildung und Betreuung. Weinheim, Beltz.

Fischer, Radisch, Stecher (2009). Halb- und Ganztagsbetrieb. In Blömeke, Bohl, u.a., Handbuch Schule. Bad Heilbrunn, Klinkhardt/UTB.

Klieme (2003). Benotungsmaßstäbe an Schulen: Pädagogische Praxis und institutionelle Bedingungen. Eine empirische Analyse auf der Basis der PISA-Studie. In: Döbert, von Kopp, u.a., Bildung vor neuen Herausforderungen. Neuwied, Luchterhand.

Klieme, Holtappels, u.a. (2007). Ganztagschule in Deutschland, Bilanz und Perspektiven. In Holtappels, Klieme, u.a., Ganztagschule in Deutschland. Ergebnisse der Ausgangserhebung der »Studie zur Entwicklung von Ganztagschulen« (StEG). Weinheim, Juventa.

Kuhn, Brümmer, Fischer, Züchner (in Vorbereitung). Differentielle Effekte der Hausaufgabenbetreuung in der Ganztagschule. ■



Anhang

Transferpaket

I.	Digitales Objekt.....	i-iii
II.	Metadatensatz im internen Format.....	iv-v
III.	Metadatensatz in Dublin Core.....	vi-vii
IV.	Metadatensatz in ONIX.....	viii-ix
V.	Checksummen.....	x

II. Metadatensatz im internen Format

@000==

SOP==2665

ACN==0

ABS==Der Beitrag konzentriert sich auf die ersten beiden Erhebungswellen von StEG. Zusammenfassend werden die Resultate zu individuellen Wirkungen der Teilnahme an Ganztagsangeboten in der Sekundarstufe dargestellt. Insgesamt zeigen sich kleine positive Effekte des Ganztagsschulbesuchs. (...) In Bezug auf Freizeitverhalten und Familienleben der Schüler/-innen scheint der Ganztagsschulbesuch keine grundlegenden Umgestaltungen zur Folge zu haben. Insbesondere beim Übertritt in die Sekundarstufe I können protektive Wirkungen der Teilnahme an Ganztagsangeboten im Hinblick auf die Entwicklung von Lernmotivation und Schulnoten nachgewiesen werden. Aber auch für ältere Schüler/-innen gilt: Der dauerhafte Besuch von qualitativ hochwertigen Angeboten der Ganztagschule kann sich positiv auf das Sozialverhalten, die Motivation und schulische Leistungen auswirken. (DIPF/Orig.)

AB2==

AUT==Fischer, Natalie; Brümmer, Felix; Kuhn, Hans Peter; Züchner, Ivo

DOK==Zeitschriftenaufsatz

CCP==

HRG==

INS==Deutsches Institut für Internationale Pädagogische Forschung (DIPF) <Frankfurt, Main>

ISB==

ISS==1432-9603; 1618-9159; 1865-2026

JAR==2010

LIF==DIPF_pedocs

SPR==ger

SWO==Ganztagschule; Sekundarstufe I; Sekundarstufe II; Außerunterrichtliche Aktivität; Einflussfaktor; Schulnote; Motivation; Soziales Verhalten; Freizeitverhalten; Familie; Empirische Untersuchung; Schulentwicklung; Deutschland;

TIT==Individuelle Wirkungen des Ganztagsschulbesuchs in der Sekundarstufe. Erkenntnisse aus der Studie zur Entwicklung von Ganztagschule (StEG)

TIS==

URL==

ZHE==2

ZJG==15

ZMP==1402062-2

ZST==Schulverwaltung. Hessen, Rheinland-Pfalz

SEI==S. 38 - 39

LIZ==0

PTD==Empirische Bildungsforschung; Schulpädagogik

QUE==Schulverwaltung. Hessen, Rheinland-Pfalz 15 (2010) 2, S. 38-39

PRE==verlag

DNB==1

BHS==

DGT==0

DLF==Carl Link

KNF==1

STA==o.k.

MGA==0
AFL==
WGL==1
BEX==
BIN==
LIC==Deutsches Urheberrecht
DOI==
URN==urn:nbn:de:0111-opus-26658
DCR==1267692845
DPB==1267692845
DMO==1277905481

III. Metadatensatz in Dublin Core

```
<?xml version="1.0" encoding="Latin1" ?>
- <record-list xmlns:dc="http://purl.org/dc/elements/1.1/">
- <dc-record>
- <dc:title>Individuelle Wirkungen des Ganztags schulbesuchs in der Sekundarstufe.
  Erkenntnisse aus der Studie zur Entwicklung von Ganztags schule
  (StEG)</dc:title>
<dc:subject>Ganztags schule</dc:subject>
<dc:subject>Sekundarstufe I</dc:subject>
<dc:subject>Sekundarstufe II</dc:subject>
<dc:subject>Außerunterrichtliche Aktivität</dc:subject>
<dc:subject>Einflussfaktor</dc:subject>
<dc:subject>Schulnote</dc:subject>
<dc:subject>Motivation</dc:subject>
<dc:subject>Soziales Verhalten</dc:subject>
<dc:subject>Freizeitverhalten</dc:subject>
<dc:subject>Familie</dc:subject>
<dc:subject>Empirische Untersuchung</dc:subject>
<dc:subject>Schulentwicklung</dc:subject>
<dc:subject>Deutschland</dc:subject>
<dc:description>Der Beitrag konzentriert sich auf die ersten beiden
  Erhebungswellen von StEG. Zusammenfassend werden die Resultate zu
  individuellen Wirkungen der Teilnahme an Ganztags schulangeboten in der
  Sekundarstufe dargestellt. Insgesamt zeigen sich kleine positive Effekte des
  Ganztags schulbesuchs. (...) In Bezug auf Freizeitverhalten und Familienleben
  der Schüler/-innen scheint der Ganztags schulbesuch keine grundlegenden
  Umgestaltungen zur Folge zu haben. Insbesondere beim Übertritt in die
  Sekundarstufe I können protektive Wirkungen der Teilnahme an
  Ganztags angebotenen im Hinblick auf die Entwicklung von Lernmotivation und
  Schulnoten nachgewiesen werden. Aber auch für ältere Schüler/-innen gilt: Der
  dauerhafte Besuch von qualitativ hochwertigen Angeboten der Ganztags schule
  kann sich positiv auf das Sozialverhalten, die Motivation und schulische
  Leistungen auswirken. (DIPF/Orig.)</dc:description>
<dc:type>Aufsatz (Zeitschrift)</dc:type>
<dc:type>article</dc:type>
<dc:type>Text</dc:type>
<dc:type>publishedVersion</dc:type>
<dc:source>Schulverwaltung. Hessen, Rheinland-Pfalz 15 (2010) 2, S. 38-
  39</dc:source>
<dc:creator>Fischer, Natalie</dc:creator>
<dc:creator>Brümmer, Felix</dc:creator>
<dc:creator>Kuhn, Hans Peter</dc:creator>
<dc:creator>Züchner, Ivo</dc:creator>
<dc:contributor>Deutsches Institut für Internationale Pädagogische Forschung
  (DIPF) <Frankfurt, Main></dc:contributor>
<dc:date type="created">2010-03-04</dc:date>
<dc:format>application/pdf</dc:format>
<dc:identifier>http://opus-hh/volltexte/2010/2665/</dc:identifier>
<dc:identifier>urn:nbn:de:0111-opus-26658</dc:identifier>
<dc:identifier type="ISSN">1432-9603</dc:identifier>
<dc:identifier type="ISSN">1618-9159</dc:identifier>
<dc:identifier type="ISSN">1865-2026</dc:identifier>
```

```
<dc:language>de</dc:language>  
<dc:rights>Deutsches Urheberrecht</dc:rights>  
</dc-record>  
</record-list>
```

IV. Metadatensatz in ONIX

```
<?xml version="1.0" encoding="Latin1" ?>
- <OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/
    http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
  <responseDate>2009-02-20T11:51:41Z</responseDate>
  <request verb="ListRecords" from="2002-05-01" until="2009-02-
    16">http://server.metadatenlieferant.de/oaipmh</request>
- <ListRecords>
- <record>
- <header>
- <identifier>0101</identifier>
- <datestamp>2010-08-25</datestamp>
- </header>
- <metadata>
- <ONIXmessage release="2.1" xmlns="http://www.editeur.org/onix/2.1/short"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.editeur.org/onix/2.1/short
    http://www.editeur.org/onix/2.1/short/ONIX_BookProduct_Release2.1_short
    .xsd">
- <header>
- <m174>DIPF</m174>
- <m182>20100825</m182>
- </header>
- <product>
- <a001>2665</a001>
- <a002>03</a002>
- <productidentifier>
- <b221>22</b221>
- <b244>urn:nbn:de:0111-opus-26658</b244>
- </productidentifier>
- <b012>DG</b012>
- <title>
- <b202>01</b202>
- <b203>Individuelle Wirkungen des Ganztagsschulbesuchs in der Sekundarstufe.
  Erkenntnisse aus der Studie zur Entwicklung von Ganztagschule
  (StEG)</b203>
- </title>
- <contributor>
- <b034>1</b034>
- <b035>A01</b035>
- <b037>Fischer, Natalie</b037>
- </contributor>
- <contributor>
- <b034>2</b034>
- <b035>A01</b035>
- <b037>Brümmer, Felix</b037>
- </contributor>
- <contributor>
- <b034>3</b034>
- <b035>A01</b035>
- <b037>Kuhn, Hans Peter</b037>
```

```

    </contributor>
- <contributor>
  <b034>4</b034>
  <b035>A01</b035>
  <b037>Züchner, Ivo</b037>
  </contributor>
  <b047>Deutsches Institut für Internationale Pädagogische Forschung (DIPF)
    <Frankfurt, Main></b047>
  <n386 />
- <mainsubject>
  <b191>18</b191>
  <b069>370</b069>
  </mainsubject>
- <othertext>
  <d102>32</d102>
  <d103>06</d103>
  <d104>Schulverwaltung. Hessen, Rheinland-Pfalz 15 (2010) 2, S. 38-39</d104>
  </othertext>
- <productwebsite>
  <b367>02</b367>
  <f123>http://www.pedocs.de/frontdoor.php?source_opus=2665</f123>
  </productwebsite>
- <publisher>
  <b291>01</b291>
  <b081 />
  </publisher>
  <b209 />
  <b003>2010</b003>
  </product>
  </ONIXmessage>
  </metadata>
  </record>
  </ListRecords>
  </OAI-PMH>

```


V. Checksummen

1. Transferpaket

Dateiname: 2665_2010-08-25.zip.md5

Checksumme: e804a7ae79a13cb31d59be57aed96a96

2. Unterverzeichnis (zu archivierende Objekte)

PDF

Dateiname: Herp10_02_0002_D_A.pdf.md5

Checksumme: 872b9c769fb48d92dfa37c023f80dbc4

Metadaten intern

Dateiname: 2665.txt.md5

Checksumme: e4b41a819cf578e02eb7be180c2f7137

3. Hauptverzeichnis

Dublin Core

Dateiname: 2665.dc.xml.md5

Checksumme: 3a1e1581b9273194dc0c4089712ea811

ONIX

Dateiname: 2665.onix.xml.md5

Checksumme: e96db098991b63d3e3fc602f6bb7b568

Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig verfasst habe und dass keine anderen als die angegebenen Hilfsmittel benutzt wurden.

Frankfurt, den 31.08.2010