

DIE PILOTPHASE DES PROJEKTS „DIGITALES ARCHIV“ AM BUNDESARCHIV IN KOBLENZ - ÜBERNAHME, BEWERTUNG UND ARCHIVIERUNG VON ELEKTRONISCHEN AKTEN

Karsten Huth*

UDC: 930.253(430):004

Karsten Huth: Pilot Stage of »Digital Archives« Project of Federal Archives in Koblenz - Acquisition, Appraisal and Archiving of Electronic Records. Technical and Field Related Problems of Traditional and Electronic Archiving. Conference Proceedings, Maribor 6/2007, pp. 24-31.

Izvirnik v nemščini, izvleček v angleščini in slovenščini, povzetek v angleščini.

This paper describes the automated workflow of the offering, appraisal, accession and storage of electronic files out of the agencies Document Management Systems. The described workflow was a test drive within the project "Digital Archive" carried out by a technical prototype at the Bundesarchiv in Koblenz, Germany.

-

UDK: 930.253(430):004

Karsten Huth: Pilotna faza projekta »Digitalni arhiv« v Zveznem arhivu v Koblenzu - prevzem, vrednotenje in arhiviranje elektronskih spisov. Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja. Zbornik referatov z dopolnilnega izobraževanja, Maribor 6/2007, str. 24-31.

Original in German, abstract in English and Slovenian, summary in English.

Prispevek opisuje avtomatizirano poslovanje postopkov ponudbe, vrednotenja ter prevzema in shranjevanja elektronskih spisov iz elektronskega sistema za upravljanje z dokumenti posameznih uradov. Opisano poslovanje je bilo kot poskus s pomočjo tehničnega prototipa izvedeno v okviru projekta »Digitalni arhiv« v Zveznem arhivu Koblez.

Das Bundesarchiv hat den gesetzlichen Auftrag das gesamte Schriftgut der Bundesbehörden auf seine Archivwürdigkeit zu prüfen und historisch Wertvolles dauerhaft zu erhalten. Da in den Bundesbehörden die Umstellung von einer papiergebundenen Aktenführung hin zum e-Government und somit hin zu elektronischen Vorgangsbearbeitungssystemen immer weiter fortschreitet, braucht das Bundesarchiv Konzepte und Lösungen für den Umgang mit elektronischem Schriftgut. Deshalb wurde das Projekt „Digitales Archiv“ ins Leben gerufen, an dessen Ende 2010 ein OAIS-konformes¹ digitales Archiv zur Erhaltung und Nutzung von digitalem Archivgut stehen soll.

DIE AUSGANGSLAGE

Zur Verwaltung von elektronischen Akten werden in den Behörden Vorgangsbearbeitungssysteme (VBS) eingesetzt. Die elektronische Akte ist jedoch eine physisch schwer zu erfassende Einheit. Einzelne Dokumente einer Akte werden durch Dateien repräsentiert und vom VBS technisch verwaltet. Die sachliche Einheit

* *Mag. Karsten Huth, Referat B2, Bundesarchiv Koblenz, 56064 Koblenz, Deutschland.*

¹ *ISO 14721:2003*

und Vollständigkeit einer elektronischen Akte bzw. eines elektronischen Vorgangs wird nur durch ein Referenzierungssystem und durch Metadaten innerhalb des VBS hergestellt. Um die Einheit und Vollständigkeit der elektronischen Akten auch außerhalb des VBS zu erhalten, müssen nicht nur die Dateien, welche die elektronischen Dokumente repräsentieren an das Archiv abgegeben werden, sondern auch die dazugehörigen Metadaten, die den Kontext des jeweiligen Dokuments wiedergeben.

Um die elektronische Akten langfristig verfügbar zu halten, müssen die Dateien im Archiv in einem Format gespeichert werden, dass auch über einen längeren Zeitraum hinweg in wechselnden technischen Umgebungen darstellbar ist. Es ist wesentlich, dass nur Dateiformate im Archiv verwendet werden, deren technischer Aufbau offen dokumentiert und frei verfügbar ist.

PROBLEMFELD AUSSONDERUNGSSCHNITTSTELLEN

Ein Konzept für die geregelte Aussonderung von elektronischen Akten liegt in Deutschland² als Teil des DOMEA Organisationskonzepts 2.0 bereits vor. Dennoch haben selbst DOMEA-zertifizierte VBS keine standardisierte Schnittstelle, über die Daten zwischen der Behörde und dem Bundesarchiv in einem einheitlichen Datenformat ausgetauscht werden könnten. Insbesondere die Metadaten zu den abgegebenen elektronischen Akten werden von jedem VBS in seiner eigener Art und Weise exportiert. Für die effiziente Verarbeitung und Speicherung der Metadaten im Bundesarchiv ist jedoch ein einheitliches Metadatenformat zwingend notwendig. Deshalb muss das digitale Archiv mit unterschiedlichen Metadatenformaten aus den Behörden umgehen können und diese in ein archiveigenes Metadatenformat umsetzen. Das Bundesarchiv entwickelt aus diesem Grund ein StandardArchivierungsModul (SAM) als Schnittstelle zwischen den Behörden und dem Archiv. Die Entwicklung von SAM wird als EfA-(Einer für Alle-)Projekt von der KBSt³ gefördert.

DER TECHNISCHE AUFBAU DES PROTOTYPEN

In der Pilotphase wurde ein vollständiger Prozess vom Anbieten und Bewerten bis zur Übernahme und Speicherung von elektronischen Akten technisch umgesetzt. Die Erkenntnisse aus dieser Phase des Projekts halfen die Anforderungen an ein voll funktionsfähiges digitales Archiv zu formulieren. Unbedingt notwendig ist eine Komponente, welche die archivischen Geschäftsgänge automatisiert steuert und zugleich dem Archiv die Möglichkeit bietet, neue Geschäftsprozesse einzufügen und bestehende Prozesse selbständig anzupassen. Während der Pilotphase wurden diese Anforderungen durch einen Business Process Management Server (BPMS) technisch realisiert. Die technische Installation während der Pilotphase, bestehend aus einem BPM-Server zur Steuerung des Übernahmeprozesses und einem simulierten Speichersystem inklusive der Software einer Archivdatenbank, stellt einen Prototypen dar, der noch nicht über die nötige Leistungskapazität eines digitalen Archivs im Produktivbetrieb verfügt. Dennoch konnten mit dem Prototypen alle wesentlichen

² DOMEA-Erweiterungssmodul „Aussonderung und Archivierung Elektronischer Akten“; im Internet: http://www.kbst.bund.de/cln_047/nn_836802/Content/Standards/Domea_Konzept/domea_node.html__nnn=true

³ <http://www.bundesarchiv.de/service/behoerdenberatung/01435/index.html>

Funktionen des Übernahmeprozesses und der Speicherung simuliert und durchgespielt werden.

Der BPM-Server steuert die automatischen Geschäftsprozesse für die archivischen Geschäftsgänge des Anbietens und Bewertens sowie der Übernahme. Die Software des BPM-Systems verfügt über eine grafische Entwicklungsoberfläche, mit der ein Anwender Prozesse selbständig erstellen und darstellen kann. Die Software übersetzt die grafische Darstellung des Prozesses in ein Script, welches einen neuen Geschäftsprozess auf dem BPM-Server implementiert. So könnte das Bundesarchiv seine Geschäftsprozesse bei Bedarf anpassen. Zudem verwaltet der BPM-Server alle Daten, die für die Steuerung der Geschäftsprozesse notwendig sind. Die Fachaufgaben innerhalb eines Geschäftsprozesses können den Archivmitarbeitern automatisiert zugewiesen werden.

	AEID	NAME	EMAIL	BPMSUSER
1	1	Mary Evaluator	ingestevaluator@b[REDACTED].de	062245158232116229727791000001
2	2	John Manager	ingestmanager@b[REDACTED].de	062245158232116229737721900002
3	3	Peter Mustermann	Mustermann @keep.com	192168001158116099777687500006
4	4	Kerstin Beispielfrau	Beispielfrau @barch.bund.de	062245158232116229382250100001
5	5	Michael Mustermann	Mustermann @barch.bund.de	062245158232116229390916500002
6	6	Antje Beispielfrau	Beispielfrau @barch.bund.de	062245158232116229406724100003
7	7	Elke Beispielfrau	Beispielfrau @barch.bund.de	062245158232116229418848400004
8	8	Barbara Beispielfrau	Beispielfrau @barch.bund.de	062245158232116229521410400005
9	9	Karsten Mustermann	Mustermann @barch.bund.de	062245158232116229537803000006
10	10	Kathrin Beispielfrau	Beispielfrau @barch.bund.de	062245158232116229606146100008

Abbildung 1: Datenbankeinträge für die Zuweisung der Sachbearbeiter (Bewerter) während des Geschäftsgangs

Die BPM-Software schreibt Protokolle über die technischen Vorgänge während des Prozesses und legt sie in einer Datenbank ab. Dadurch bleibt jeder Arbeitsschritt eines Geschäftsprozesses für das Archiv auch im Nachhinein nachvollziehbar.

Die Speicherlösung des Prototypen ist ein simuliertes Bandlaufwerk. Die Installation eines kompletten Speichersystems hätte den Rahmen der Testphase gesprengt.

DIE TESTDATEN

Zur Verfügung standen elektronische Akten aus dem Bundesministerium für Verkehr, Bau und Stadtentwicklung. Diese Akten wurden in der Behörde in Papierform angelegt und anschließend digitalisiert in ein VBS überführt.

Um einen Geschäftsgang mit Abgaben aus mehreren Behörden zu testen, wurden elektronische Dokumente aus einem Nachlassbestand als Datensatz zur Simulation einer zweiten Behörde verwendet. Beide Testdatensätze hatten

verschiedene Metadatenformate, die der BPM-Server identifizieren und in das archivinterne Format umsetzen musste.

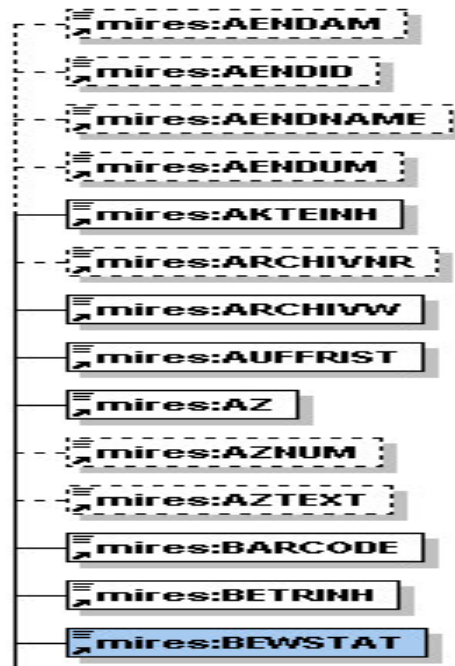


Abbildung 2: Auszug aus Schema von Metadatenformat einer Behörde

DER GESCHÄFTSPROZESS ANBIETEN UND BEWERTEN

Vor der ersten Abgabe einer Behörde an das Bundesarchiv steht eine Kontaktaufnahme. Bei dieser Kontaktaufnahme erhebt das Bundesarchiv wichtige Daten, mit denen ein Profil der Behörde erstellt werden kann. Dieses Behördenprofil wird in einer Datenbank des BPM-Servers gespeichert.

Der wichtigste Teil dieses Profils ist die Definition des Metadatenschemas für die Anbieterverzeichnisse, die später im Geschäftsprozess „Anbieten und Bewerten“ an das Bundesarchiv gesendet werden. Die Profildaten der Behörde werden im Bundesarchiv mit einer eindeutigen Behördenidentifikationsnummer versehen und in einer Datenbank gespeichert.

Der Prozess des „Anbietens und Bewertens“ wird ausgelöst, indem die Behörde dem Archiv per E-Mail ein Anbieterverzeichnis im behördeneigenen Format sowie eine spezielle Transaktionsdatei mit der korrekten Behördenidentifikationsnummer zusendet.

Stimmt die Behördenidentifikationsnummer nicht mit dem entsprechenden Profileintrag in der Datenbank des Archivs überein oder entspricht das Anbieterverzeichnis nicht dem im Behördenprofil hinterlegten Schema, so bricht der Prozess automatisch ab. Stimmen die eingehenden Daten mit dem Behördenprofil überein, wird das Anbieterverzeichnis durch das entsprechende Softwaremodul im BPM-Server in das XML-Metadatenformat des Bundesarchivs umgesetzt.

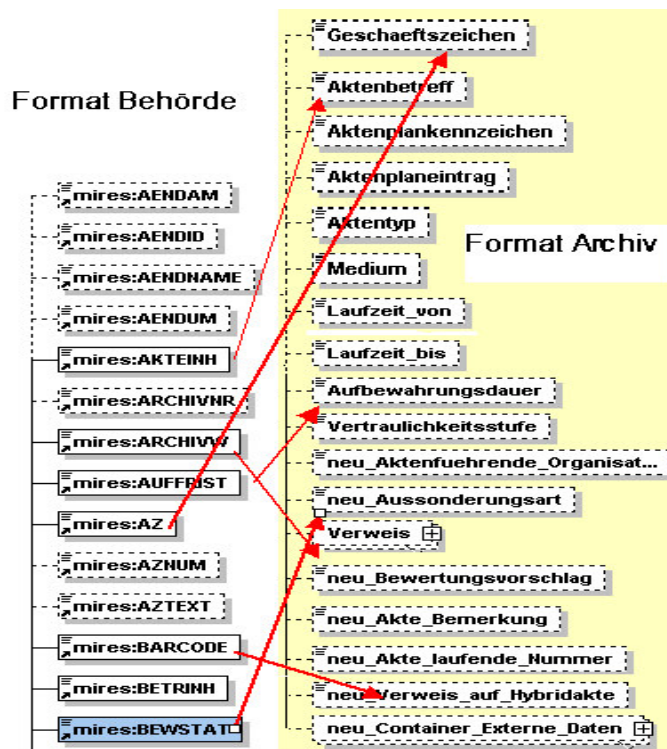


Abbildung 3: Auszug aus einem Mapping vom Behördenformat ins Archivformat

Der BPM-Server leitet das Anbieterverzeichnis automatisch an einen Archivmitarbeiter weiter, der jeden Eingang dem zuständigen Fachreferat zuweist. Die Sachbearbeiter des Fachreferats können nun über eigens entwickelte Eingabemasken ihre Bewertungsentscheidungen im Feld „Aussonderungsart“ vornehmen. Die Angabe „A“ steht für archivwürdig, „B“ für eine weitere notwendige Begutachtung und Bewertung und „V“ für vernichten.

Das Bundesarchiv OAIS-Pilotsystem

Ingest Process > Appraisal Evaluator

Angebot bewerten und Bewertung überprüfen

supervisor comment :

-ToDo-

Save Task Complete Task

Please wait while page is loading the data, it may take few minutes depending upon size of data

Offer List

Akte

Archiv-Signatur	Geschäftszeichen	neu_Aussonderungsart	Aktenbetri
2623.3/1-1/03	2623.3/1-1/03	B	V

Vorgang

Archiv-Signatur	Geschäftszeichen	neu_Aussonderungsart	Vorgangsbe
2623.3/1-1/03	2623.3/1-1/03	B	V

Abbildung 4: Bearbeitungsmaske für Anbieterverzeichnis

Nach Abschluss der Bewertung geht das Bewertungsverzeichnis zurück an die Behörde, wobei das archivinterne Datenformat wieder in das Format der Behörde umgesetzt wird. Derweil wurde der Behörde eine eindeutige Transaktionsnummer zugesendet.

DER GESCHÄFTSPROZESS „ÜBERNAHME UND SPEICHERN VON ELEKTRONISCHEN AKTEN“

Anhand des Bewertungsverzeichnisses werden in der Behörde die entsprechenden Dateien in eine Aussonderungsdatei exportiert. Dateien, die zu Akten und Vorgängen gehören, welche vom Sachbearbeiter des Archivs mit „V“ bewertet wurden, werden noch in der Behörde vernichtet. Die Aussonderungsdatei, die an das Bundesarchiv geht, besteht aus den Dateien der mit „A“ und „B“ bewerteten Vorgänge und aus den dazugehörigen Metadaten des VBS der Behörde.

Metadaten und Dateien werden in einen ZIP-Container gepackt. Zeitgleich wird in der Behörde eine Transaktionsdatei erstellt. Wichtigste Bestandteile der Transaktionsdatei sind die eindeutige Transaktionsnummer, die Behördenidentifikationsnummer und der Dateiname des Aussonderungsverzeichnisses, das alle Metadaten aus dem VBS enthält. Transaktionsdatei und ZIP-Container werden entweder auf einem Datenträger gespeichert und zum Archiv transportiert oder auf einem sicheren Datenkanal übermittelt. Der Weg der Datenübertragung wird von Archiv und Behörde festgelegt.

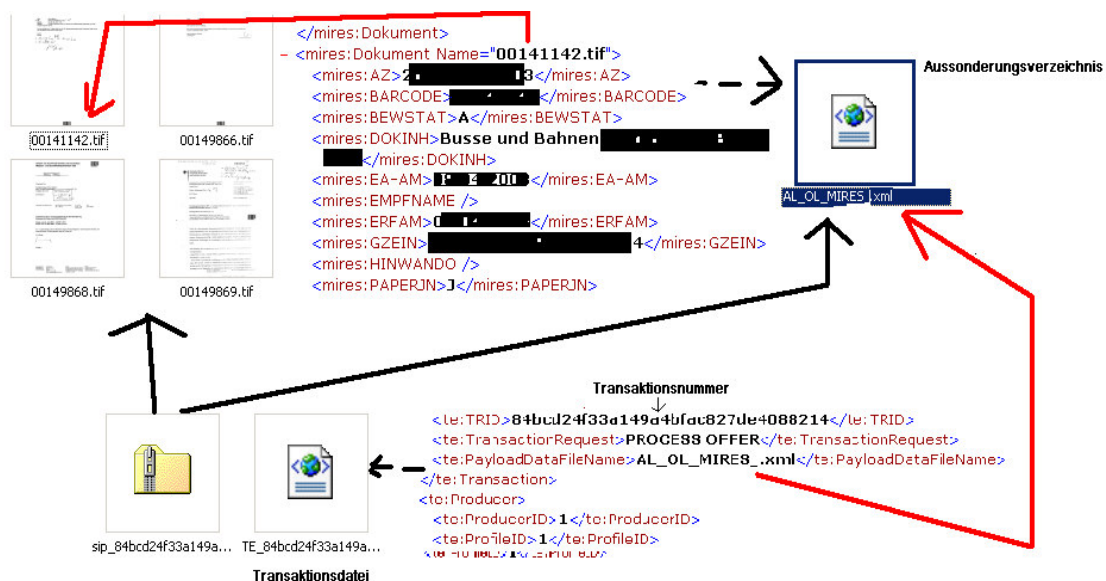


Abbildung 5: Aussonderung (SIP)

Beim Eingang der Aussonderung in das Archiv prüft der BPM-Server, ob die Transaktionsnummer und die Behördenidentifikationsnummer gültig sind. Erst dann wird der ZIP-Container entpackt und die Aussonderung auf ihre Vollständigkeit hin überprüft. Die Metadaten werden in das XML-Archivformat (XBARCH) umgesetzt und das Aussonderungsverzeichnis wird an das zuständige Fachreferat weitergeleitet.

Nach einer abschließenden Bewertung durch die Sachbearbeiter, bei der die Sachbearbeiter auch die Dokumente betrachten können, erfolgen die letzten technischen Prozesse vor dem Eingang der Daten in das Speichersystem.

Die eingehenden Dateien werden durch das Programm JHOVE⁴ technisch geprüft. Dabei werden spezifizierte Angaben über das Dateiformat protokolliert. Zusätzlich generiert JHOVE für jede Datei einen Hashwert, mit dem später geprüft werden kann, ob eine Datei auch nach längerer Speicherung noch unverändert und intakt ist. Diese technischen Angaben werden in einem aufwendigen Arbeitsschritt den XBARCH-Metadaten beigefügt. Hinzu kommen Angaben über die IT-Ausstattung der abgebenden Behörde, die unter dem entsprechenden Behördenprofil in der Datenbank des Bundesarchivs abgelegt sind. Nach diesem Arbeitsschritt enthält eine XBARCH-Datei die inhaltlichen Informationen zur Akte, zum Vorgang und zu den entsprechenden Dokumenten aus dem VBS der Behörde. Zusätzlich enthält XBARCH technische Informationen. Zum einen eine kurze technische Beschreibung des Vorgangsbearbeitungssystems der abgebenden Behörde und zum anderen enthält XBARCH technische Angaben zu den Dateien, die die Dokumente repräsentieren. Diese Informationen sind für eine dauerhafte Aufbewahrung notwendig. Die Dateien werden anschließend einzeln mit den dazugehörigen XBARCH-Metadaten in ein unkomprimiertes TAR-Archiv gepackt und an das Speichersystem übergeben. Zusätzlich werden ausgewählte Metadaten aus XBARCH in eine Archivdatenbank importiert, über die auf Akten- Vorgangs- und Dokumentebene im Archivbestand recherchiert werden kann. Der Prototyp simuliert ein Bandsystem, auf dem vollständige Archivinformationspakete (AIP) abgelegt werden. Wichtig ist, dass ein AIP nicht über mehrere Datenträger verteilt gespeichert wird, sondern auch physisch als Einheit auf einer Bandkassette verbleibt.

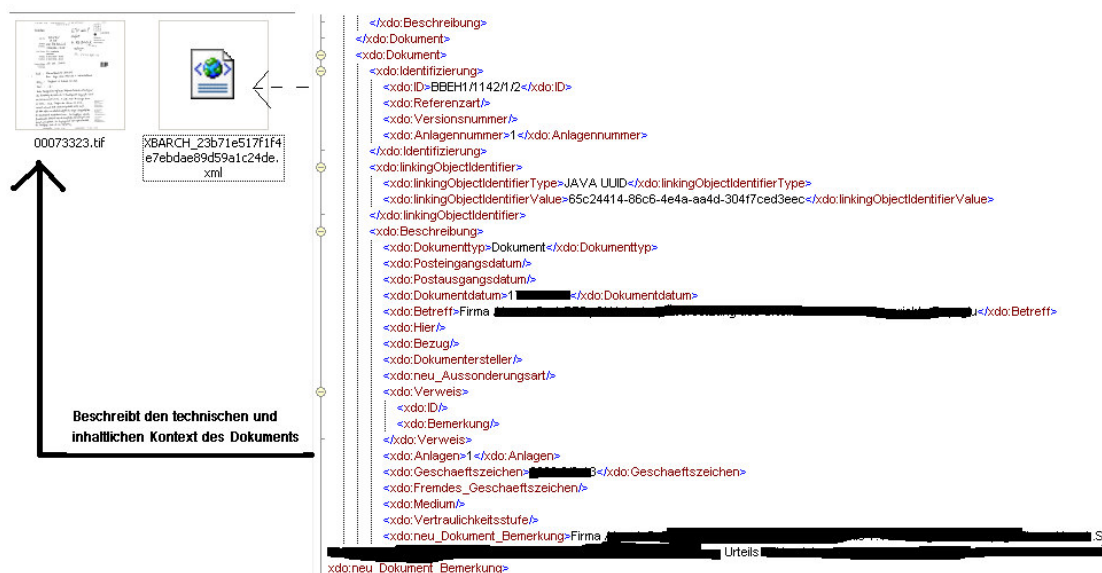


Abbildung 6: Archival Information Package

⁴ <http://hul.harvard.edu/jhove/>

Wenn alle ausgesonderten Dateien im Bandspeicher des Archivs abgelegt sind, wird die Behörde über den erfolgreichen Eingang ihrer Akten benachrichtigt und zum Löschen ihrer ausgesonderten Dateien und Metadaten aus dem VBS aufgefordert. Mit diesem Schritt ist der Geschäftsprozess der Übernahme abgeschlossen.

WEITERE AUSBLICKE

Die nächste Projektphase betrifft die Ausschreibung für die erste Ausbaustufe eines produktiven digitalen Archivs. Die Unterlagen für eine Ausschreibung werden zur Zeit am Bundesarchiv ausgefertigt.

Für die langfristige Speicherung digitalen Archivguts müssen Geschäftsprozesse entwickelt werden, die eine Migration von Dateiformaten ermöglichen. Das elektronische Archivgut kann nur dauerhaft verfügbar sein, wenn es an die kommenden technischen Umgebungen (Hardware und Software) angepasst wird. Dabei müssen alle vorgenommenen Veränderungen des Dateiformats und die Auswirkungen dieser Veränderungen auf das elektronische Dokument sorgfältig geprüft und dokumentiert werden. Nur so kann die Authentizität elektronischen Archivguts erhalten bleiben. Ein entsprechender Geschäftsprozess konnte während der Pilotphase noch nicht getestet werden.

In der ersten Ausbaustufe wird das Digitale Archiv elektronische Akten aus Bundesbehörden übernehmen und erhalten. In den nächsten Ausbaustufen wird der Bestand des Digitalen Archivs um weitere Arten von elektronischem Archivgut erweitert. So sollen auch Datenbanken aus den Behörden, Bilder und audiovisuelle Medien digital archiviert werden.

SUMMARY

PILOT STAGE OF »DIGITAL ARCHIVES« PROJECT OF FEDERAL ARCHIVES IN KOBLENZ - ACQUISITION, APPRAISAL AND ARCHIVING OF ELECTRONIC RECORDS

The Federal Archives has a legal duty to appraise the archival value of written records of federal offices and permanently keep historically important records. Since federal offices increasingly use electronic systems to manage records, the Federal Archives found itself in need for concepts and solutions for managing electronic records. Therefore it started the Digital Archives Project, whose aim is to establish a digital archives for keeping and use of digital archives, which will be in compliance with OAIS, till 2010.