

Die Verkündungsplattform Bayern – Services und vertrauenswürdige digitale Langzeitarchivierung

Anna Kugler, Markus Brantl und Astrid Schoger

1 Die Verkündungsplattform Bayern

Mit der offiziellen Eröffnungsveranstaltung im Kuppelsaal der Bayerischen Staatskanzlei am 22. Januar 2009 gaben der damalige Leiter der Bayerischen Staatskanzlei Siegfried Schneider und der Generaldirektor der Bayerischen Staatsbibliothek (BSB) Dr. Rolf Griebel den Startschuss für die Verkündungsplattform Bayern¹. Seit diesem Tag erfolgt die amtliche Verkündung der vier Amtsblätter der bayerischen Staatsministerien ausschließlich elektronisch. Bayern übernahm mit diesem Projekt eine wichtige Vorreiterrolle für ganz Deutschland und setzte einen weiteren Meilenstein in der eGovernment-Initiative² der Staatsregierung, die 2002 begonnen worden war. In einem gemeinsamen Projekt zwischen der BSB und der Staatskanzlei war die Verkündungsplattform im Jahr 2008 konzipiert und realisiert worden.

Die BSB hat den Auftrag der Staatskanzlei nicht zuletzt wegen ihrer umfangreichen Erfahrungen in den Bereichen Digitalisierung, digitale Langzeitarchivierung sowie dem Bereich elektronisches Publizieren erhalten. Das Münchener Digitalisierungszentrum/Digitale Bibliothek (MDZ)³ sowie das Zentrum für Elektronisches Publizieren (ZEP)⁴ der BSB waren die Partner des im Projekt federführenden Referats B II 4 der Staatskanzlei. Hinzu kam das Leibniz-Rechenzentrum (LRZ), das seit 2004 eng mit dem MDZ als technischer Serviceprovider zusammenarbeitet.

Die zentralen Anforderungen an die Verkündungsplattform Bayern umfassten die Sicherstellung der dauerhaften Bereitstellung sowie der Integrität und Authentizität der Amtsblätter, eine hohe Verfügbarkeit der Online-Plattform, die Präsentation der Verkündungsblätter sowohl in einer PDF/A- als auch in einer HTML-Fassung und auch die Verwaltung eines

Info-Diensts, mit welchem alle registrierten Nutzer auf das Neuerscheinen eines Amtsblatts hingewiesen werden.

Die Veröffentlichung des Bayerischen Gesetz- und Verordnungsblattes (GVBl) auf der Verkündungsplattform Bayern wurde bereits in der ersten Konzeptionsphase berücksichtigt miteinbezogen, jedoch erfolgte die Umsetzung erst ein Jahr später in einem Folgeprojekt. Eine besondere Herausforderung bei dieser Erweiterung war die Implementierung einer Schnittstelle zur Datenbank Bayern-Recht, die von der juris GmbH gehostet wird und vorher nicht öffentlich über das Internet zur Verfügung stand. Seit Abschluss dieses Folgeprojekts im Januar 2010 stehen nun auch die konsolidierten Fassungen aller bayerischen Gesetze und Verordnungen online und kostenfrei zur Verfügung⁵.

Die Maßnahmen zur Langzeitarchivierung bzw. Langzeitverfügbarkeit der bayerischen ministerialen Amtsblätter und des GVBl werden von Anfang an in die verschiedenen Verarbeitungsstufen der Amtsblätter miteinbezogen: Angefangen bei der Produktion der Amtsblätter im speziell für die Langzeitarchivierung ISO-standardisierten Dateiformat PDF/A über die Erschließung in verschiedenen Nachweissystemen (ZDB, Verbundkatalog) und die dauerhafte Identifizierung bis hin zur Langzeitspeicherung am LRZ. Im Folgenden werden zunächst die Produktionsprozesse der Amtsblätter sowie die Services, die dem Nutzer auf der Verkündungsplattform zur Verfügung stehen, beschrieben. In einem zweiten Teil erfolgt eine ausführliche Darstellung der umgesetzten Langzeitarchivierungsmaßnahmen.

2 Services der Verkündungsplattform Bayern

2.1 Elektronische Verkündung der vier Amtsblätter

Auf der Verkündungsplattform Bayern werden seit Januar 2009 die vier Amtsblätter der bayerischen Staatsministerien amtlich elektronisch verkündet:

1 <http://www.verkuendung-bayern.de> (letzter Aufruf, 25. April 2012).

2 <http://www.cio.bayern.de/internet/cio/4/19645/> (letzter Aufruf, 20. April 2012).

3 <http://www.muenchener-digitalisierungszentrum.de/> (Letzter Aufruf, 20. April 2012).

4 <http://www.bsb-muenchen.de/Zentrum-fuer-Elektronisches-Publizieren-ZEP.2349.0.html> (Letzter Aufruf, 20. April 2012).

5 <http://www.gesetze-bayern.de> (Letzter Aufruf, 25. April 2012).



Abbildung 1:
Startseite der
Verkundungs-
plattform
Bayern

- Allgemeines Ministerialblatt (AIMB),
- Bayerisches Justizministerialblatt (JMBl),
- Amtsblatt der Bayerischen Staatsministerien für Unterricht und Kultus und Wissenschaft, Forschung und Kunst (KWMBI),
- Amtsblatt des Bayerischen Staatsministeriums der Finanzen (FMBl).

Das Bayerische Gesetz- und Verordnungsblatt (GVBl) wurde in einem Folgeprojekt in die Verkündungsplattform Bayern integriert und steht seit Anfang 2010 ebenfalls in digitaler Form zur Verfügung (derzeit 53 Ausgaben). Jedoch handelt es sich dabei nicht um eine amtliche Verkündung, sondern – aufgrund der noch zu klärenden rechtlichen Rahmenbedingungen – nur um eine nachrichtliche Veröffentlichung (siehe 2.2). Ansonsten wird das GVBl mit den identischen Services wie die anderen vier Amtsblätter angeboten. Eine der zentralen Anforderungen an die elektronische Verkündung war die Vorgabe, dass das Layout und die optische Gestaltung sowie die inhaltliche Strukturierung der Amtsblätter der bisherigen gedruckten Ausgabe zu folgen hat. Das Layout der vormals gedruckten amtlichen Fassungen der Amtsblätter wird im nunmehr elektronischen PDF/A-Dokument authentisch abgebildet.

Auch in der digitalen Fassung stehen die Publikatio-

nen weiter als eigene Titel zur Verkündung von Verwaltungsvorschriften und weiteren Bekanntmachungen zur Verfügung. Die amtlich verkündete Fassung ist somit eine zweiseitige PDF/A-Ausgabe, die weiterhin in Form einzelner „Hefte“ dargestellt wird, die nach dem jeweiligen Erscheinungsbedarf des einzelnen Titels (z. B. zweiwöchentlich, monatlich oder vierteljährlich, ggf. auch mit Sonderausgaben) erscheinen. Die PDF/A-Fassung wird jahrgangswise fortlaufend paginiert, sodass die bisherige Zitierweise von Verwaltungsvorschriften und sonstigen Bekanntmachungen (Name des Verkündungsblattes, Jahrgang, Seitenzahl) beibehalten werden kann.

Auf der Verkündungsplattform stehen die digital verkündeten Amtsblätter jedoch nicht nur im PDF/A-Format zur Verfügung, sondern jede einzeln verkündete Verwaltungsvorschrift oder sonstige Bekanntmachung wird zusätzlich auch als einspaltige, nicht-paginierte und – aufgrund der fehlenden fortlaufenden Paginierung – nicht-amtliche HTML-Fassung bereitgestellt. Sie dient der leichteren Übernahme eines Textausschnitts in andere digitale Dokumente („Copy-and-Paste“) und ermöglicht auch den barrierefreien Zugang zu den jeweiligen Dokumenten, der in der PDF/A-Fassung nur eingeschränkt möglich ist. Die redaktionelle Verantwortung liegt bei den Amtsblatt-Redaktionen der zuständigen Ministerien, die eng mit dem ZEP zusammenarbeiten. Zu diesem Zweck wurde vom MDZ ein interner Arbeitsbereich aufgebaut, auf den die Amtsblatt-Redaktionen sowie die Mitarbeiter des ZEP gemeinsam zugreifen können und der so eine interne Qualitätskontrolle deutlich vereinfacht. Die Entwürfe der Amtsblätter werden hier sowohl in der PDF/A- als auch in der HTML-Fassung bereits in der originalgetreuen Umgebung der Verkündungsplattform angezeigt, sodass notwendige Korrekturen, wie etwa fehlerhafte Darstellungen in der Browser-Anzeige der HTML-Dokumente noch vor dem offiziellen Redaktionsschluss vorgenommen werden können.

Eine Volltextsuche ist über alle Verkündungsblätter sowie gezielt innerhalb eines Amtsblatts möglich. Die Trefferliste bietet sowohl den Zugriff auf das HTML-Dokument als auch auf die PDF/A-Ausgabe des jeweiligen Amtsblatts an. Eine inhaltliche Sortierung nach „Verwaltungsvorschriften“ und „Sonstiges“ (wie etwa Stellenanzeigen oder Literaturhinweise) wird bei der Anzeige der Trefferliste als zusätzlicher Service für den Nutzer angeboten. Innerhalb des GVBl liefert eine Suchabfrage überdies Ergebnisse aus der Datenbank Bayern-Recht zurück, die über einen Webservice mit der Verkündungsplattform Bayern verlinkt ist (siehe 2.2).

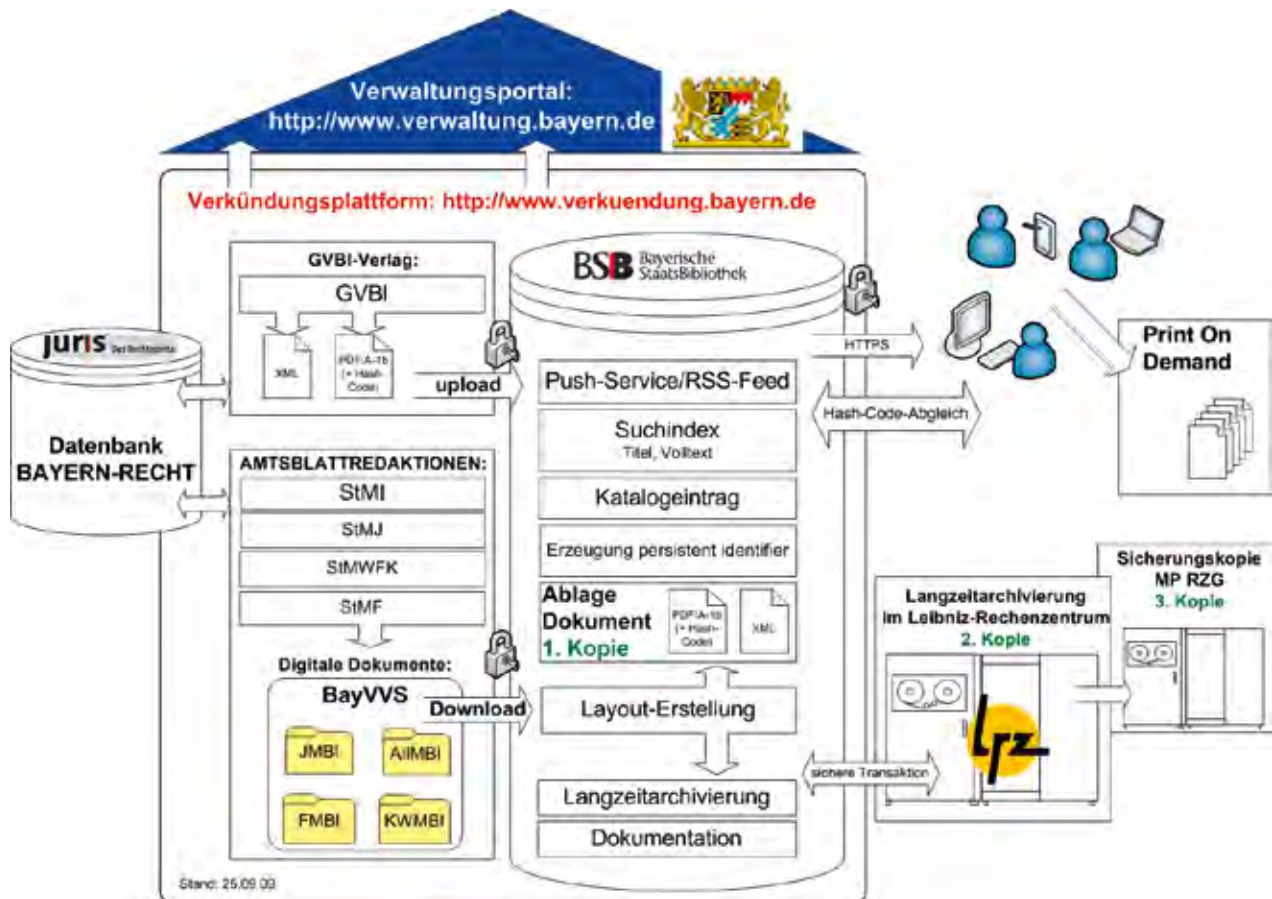


Abbildung 2: Graphische Darstellung aller Prozesse auf der Verkündungsplattform Bayern

Jeder Nutzer wird auf Wunsch per E-Mail darüber benachrichtigt, dass ein neues Amtsblatt erschienen ist („Info-Dienst“). Die E-Mail enthält Informationen darüber, um welches Amtsblatt es sich handelt (Titel, Jahrgang, Heftnummer), das genaue Erscheinungsdatum und den Link auf eine Vorschaltseite zum Amtsblatt, von der aus die entsprechenden Fassungen (PDF/A bzw. HTML) aufgerufen werden können. Seit 2009 (Stand Ende April) wurden auf der Verkündungsplattform Bayern 198 Amtsblätter amtlich verkündet. Insgesamt haben sich derzeit über 28.000 Abonnenten für den Info-Dienst registriert. Davon abonnierten fast 7.000 Nutzer das AlMBI, ca. 4.000 das JMBl, ca. 9.000 das KWMBI, ca. 4.500 das FMBl und ca. 4.000 das GVBI.

Über einen „Print-on-Demand-Service“ können sich Nutzer einzelne Amtsblätter oder auch ganze Amtsblattreihen im Jahresabonnement weiterhin in Papierform zusenden lassen. Dieses Angebot der Verkündungsplattform Bayern wird von der Justizvollzugsanstalt Landsberg am Lech betreut.

Eine sichere Übertragung der Amtsblätter wird durch ein https-Protokoll gewährleistet (= Hypertext Transfer Protocol Secure), wodurch die Kommunikation zwischen Webserver und Browser verschlüsselt er-

folgt. Die Authentizität⁶ und Integrität⁷ des heruntergeladenen PDF/A-Dokuments kann durch den Nutzer mittels einer Prüfsumme (Hash-Wert MD5, z.B. abf618382a771839a91fd5096f9cf03b) überprüft werden, der auf der Verkündungsplattform für jede Datei angegeben ist. Für die Überprüfung, ob es beim Herunterladen eines Dokuments einen Fehler gegeben hat, kann der Nutzer die Prüfsumme nach der Übertragung noch einmal selbst bilden (mittels entsprechender frei verfügbarer Software) und mit dem auf der Verkündungsplattform angezeigten Wert vergleichen. Durch die eindeutige elektronische Adresse der Verkündungsplattform, die gesicherte HTTPS-Übertragung der Dokumente sowie die Übermittlung eines Hash-Werts wird ausreichend gewährleistet, dass die von der Verkündungsplattform heruntergela-

6 „Authentizität: das Objekt stellt das dar, was es vorgibt darzustellen.“ Vgl. nestor-Kriterien – Kriterienkatalog vertrauenswürdige digitale Langzeitarchive, 2. Auflage, Frankfurt am Main 2008, S. 44, verfügbar unter: http://files.dnb.de/nestor/materialien/nestor_mat_08.pdf oder urn:nbn:de:0008-2008021802 (Letzter Aufruf: 24.04.2012).

7 „Integrität: 1. die Vollständigkeit der digitalen Objekte, 2. der Ausschluss unbeabsichtigter Modifikationen im Sinne der Erhaltungsregeln. Maßstab für die Integrität sind die als erhaltenswert definierten Eigenschaften eines digitalen Objekts.“ Vgl. nestor-Kriterien, S. 44 (Letzter Aufruf: 24.04.2012).

denen Dokumente – zumindest bis unmittelbar nach Ende des Übertragungswegs – mit den authentischen und unveränderbaren Dokumenten auf der Plattform übereinstimmen.

2.2 Veröffentlichung des GVBI mit Schnittstelle zur Datenbank Bayern-Recht

Auch für das GVBI stehen eine zweiseitige, jahrgangsweise fortlaufend paginierte PDF/A-Fassung sowie eine HTML-Version zur Verfügung. Das Besondere hierbei ist die Verlinkung aus der HTML-Fassung heraus zu den konsolidierten Fassungen der Gesetze und Verordnungen in der Datenbank BAYERN-RECHT. Somit gelangt jeder Interessierte mit nur wenigen Klicks von einem Änderungsgesetz oder einer Änderungsverordnung zur konsolidierten Gesamtausgabe eines Normentexts. Um diese Verlinkung zu ermöglichen, wurden zeitgleich mit der digitalen Veröffentlichung der Gesetzesblätter alle aktuellen Gesetzes- und Verordnungstexte in den Bürgerservice Datenbank BAYERN-RECHT aufgenommen. Zuvor standen die konsolidierten Fassungen der bayerischen Normen nur zu einem kleinen Bruchteil im Internet frei zur Verfügung. Technisch wird diese Verlinkung über einen Webservice realisiert: Bei jeder Neuveröffentlichung eines GVBI wird automatisch abgefragt, ob dazu konsolidierte Fassungen in der Datenbank BAYERN-RECHT zur Verfügung stehen: sollte dies der Fall sein, so wird in der HTML-Fassung des GVBI bei dem entsprechenden Dokument dynamisch ein Link zur konsolidierten Fassung in der Datenbank BAYERN-RECHT bereitgestellt. Im Rahmen des Projekts wurden auch alle Jahrgänge des GVBI von 1983 bis 2009 an der BSB retrodigitalisiert und auf der Verkündungsplattform in einer PDF/A-Version bereitgestellt.

3 Die Gewährleistung der langfristigen Verfügbarkeit durch vertrauenswürdige digitale Langzeitarchivierung

Durch die 2009 eingeführte, ausschließlich elektronische amtliche Verkündung der Amtsblätter kommt der vertrauenswürdigen digitalen Langzeitarchivierung besondere Bedeutung zu.

Unter dem Begriff der digitalen Langzeitarchivierung versteht man die Gesamtheit aller Maßnahmen zum Erhalt eines digitalen Bestandes und seiner Nutzbarkeit über lange Zeiträume hinweg. Dazu zählen eine Reihe von Maßnahmen, wie die geeignete Erschließung mit beschreibenden, technischen und strukturellen Metadaten sowie die eindeutige und dauerhafte Identifikation, die Archivspeicherung, die Bereitstellung der digitalen Daten für die Nutzung und insbesondere die Planung und Durchführung der

Langzeiterhaltungsstrategien: Die Migration (Konvertierung veralteter Dateiformate in aktuelle, standardisierte, offene oder offen gelegte Formate) bzw. die Emulation (Nachbildung alter Hardware sowie Betriebssysteme und Anwendungsprogramme auf neuen Rechnern). Beide Verfahren dienen dazu, die Interpretierbarkeit der Daten auch über die Lebensdauer der Dateiformate hinweg sicherzustellen, um letztlich die langfristige Nutzbarkeit der gespeicherten Informationen zu erhalten.

Das Konzept einer vertrauenswürdigen Langzeitarchivierung beinhaltet darüber hinaus eine Prüfung und Bewertung der organisatorischen und technischen Maßnahmen, die das digitale Archiv zum Erhalt der Information in digitaler Form vornimmt, um darzustellen, dass es gemäß seinen Zielen und Spezifikationen operiert.⁸

Das Münchener Digitalisierungszentrum der Bayerischen Staatsbibliothek engagiert sich seit über zehn Jahren im Forschungsfeld der digitalen Langzeitarchivierung und konnte durch mehrere DFG-Projekte sowie die Teilnahme an verschiedenen Arbeitsgruppen (nestor, DIN, ISO) umfassende Kompetenzen in diesem Bereich aufbauen.⁹ Seit 2004 besteht eine Kooperation mit dem LRZ, die immer weiter vertieft wurde. In den ersten Jahren wurde in gemeinsamer Leistung eine organisatorisch-technische Infrastruktur zur digitalen Langzeitarchivierung als Pilotsystem aufgebaut, das Bibliothekarische Archivierungs- und Bereitstellungssystem (BABS)¹⁰, welches sich am Referenzmodell Open Archival Information System (OAIS)¹¹ orientiert. Die BSB ist dabei für die Sammlung und Erschließung („Ingest“), die Verwaltung der Daten („Data Management“), die Bereitstellung und den Zugang („Access“) sowie die Erhaltungsmaßnahmen auf Software-Ebene („Preservation Planning and Action“) zuständig, während das LRZ die Langzeitspeicherung der Daten („Archival Storage“) verantwortet. Aktuell sind ca. 400 TB Daten (sowohl retrodigitalisierte als auch genuin erzeugte digitale Dokumente) im Archivsystem gespeichert.

In einem Folgeprojekt stand die Konsolidierung des Archivierungssystems im Vordergrund. Die gesamte Architektur wurde anhand des nestor-Kriterienkatalogs „Vertrauenswürdige digitale Langzeitarchive“¹² evaluiert und zu einem vertrauenswürdigen und skalierba-

8 Vgl. nestor-Kriterien, S. 46 (Letzter Aufruf: 24.04.2012).

9 SCHOGER, Astrid: Aufbau einer Infrastruktur zur digitalen Langzeitarchivierung. In: Bibliotheksforum Bayern 05/2011.

10 <http://www.babs-muenchen.de> (Letzter Aufruf: 04.04.2012).

11 ISO 14721: 2003; CCSDS: Reference Model for an Open Archival Information System (OAIS) (2002): <http://public.ccsds.org/publications/archive/650x0b1.pdf> (Letzter Aufruf: 25. April 2012).

12 Vgl. nestor-Kriterien (Letzter Aufruf: 24.04.2012).

ren System ausgebaut. Die Kriterien für eine vertrauenswürdige Langzeitarchivierung umfassen u. a. die Sicherstellung der Integrität, Authentizität und Nutzbarkeit der digitalen Objekte. Überdies sind eine Strategie zur Planung und Durchführung von Langzeiterhaltungsmaßnahmen sowie eine angemessene technische Infrastruktur für das digitale Archiv gefordert.

Die Verkündungsplattform Bayern bzw. alle darauf veröffentlichten Dokumente wurden von Beginn an in das Langzeitarchivierungs-Konzept der BSB integriert. Im Folgenden werden die einzelnen Maßnah-

gänge 2009 bis 2011 ist über die Auswahl in einem Dropdown-Menü realisiert.

Unabdingbare Voraussetzung für eine dauerhafte Bereitstellung und Verfügbarkeit ist die eindeutige und dauerhafte Identifizierung und damit Zitierbarkeit der digitalen Dokumente. Um dies garantieren zu können, sind stabile Referenzen, sogenannte Persistent Identifier zu schaffen, die auch über technische und organisatorische Veränderungen hinweg den langfristigen Zugang zu den Dokumenten sichern. Eine URL ist meistens nicht dauerhaft nutzbar, da sie sich, bei-

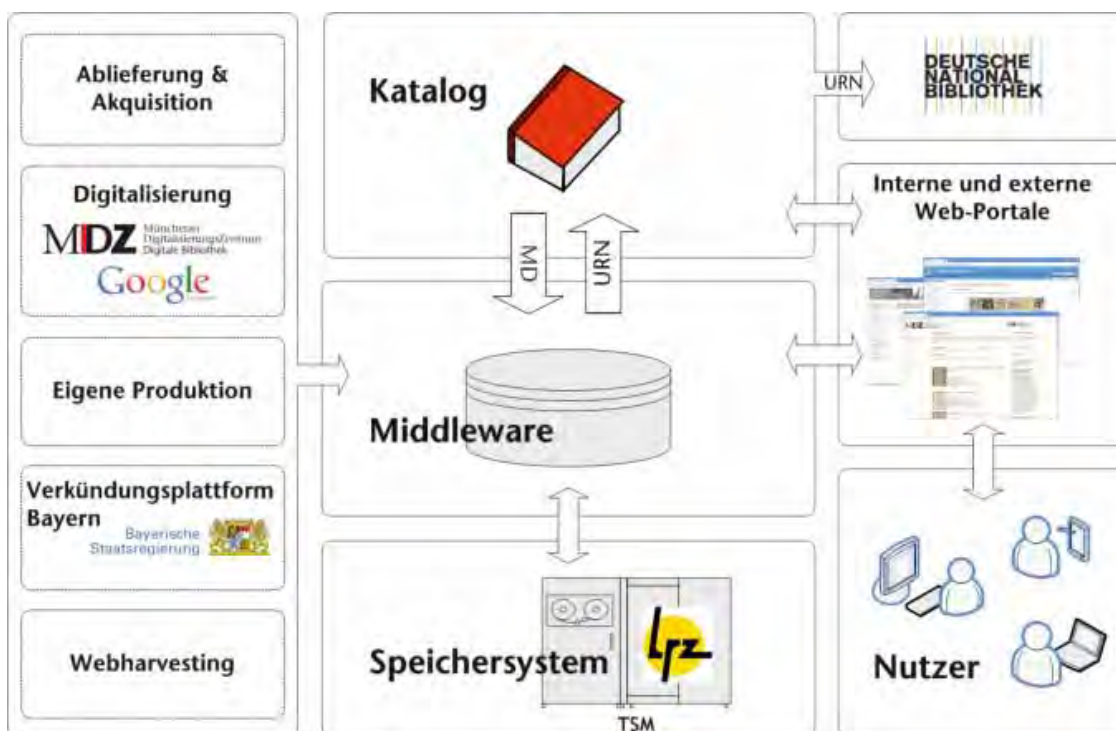


Abbildung 3:
LZA-Infrastruktur
an der BSB

men der vertrauenswürdigen Langzeitarchivierung im Falle der Verkündungsblätter detailliert dargestellt.

3.1 Die Erschließung im Katalog der BSB

Die dauerhafte Bereitstellung der Verkündungsblätter und deren langfristige Verfügbarkeit setzt – neben der klassischen Optimierung der Webangebote für die großen Suchmaschinen (z. B. Google) – auch eine entsprechende Verzeichnung, d. h. Erschließung, der einzelnen Amtsblatttitel sowie des GVBI in verschiedenen bibliothekarischen Nachweissystemen, wie dem bayerischen Verbundkatalog und der Zeitschriftendatenbank (ZDB), voraus. Bei der Katalogisierung wurde je Verkündungsblatt eine Gesamtaufnahme erstellt. Der Link im Katalog der BSB führt zur entsprechenden Amtsblatt-Seite des aktuellen Jahres auf der Verkündungsplattform. Hier sind alle bereits erschienenen Einzelausgaben des laufenden Jahrgangs gelistet und der Nutzer kann zwischen der PDF/A- und der HTML-Anzeige auswählen. Ein Browsing der Jahr-

spielsweise durch eine Serverumstellung oder den Einsatz einer neuen Bereitstellungssoftware, ändern kann.

Mit Hilfe der Persistent Identifier dagegen können digitale Objekte langfristig weltweit eindeutig identifiziert und adressiert werden, denn sie behalten ihre Gültigkeit, auch wenn sich der Speicherort der Netzpublikation ändert. Sie verweisen nicht selbst auf die Informationsressourcen, sondern auf einen zwischengeschalteten Resolvingdienst, der die Verwaltung der Adressen der tatsächlichen Speicherorte leistet. Die Deutsche Nationalbibliothek (DNB) stellt einen solchen Dienst zur Verfügung¹³ und sie verwendet als Schema für die Identifizierung digitaler Ressourcen den Uniform Resource Name (URN)¹⁴. URNs wurden bereits in den 1990er Jahren als weltweit gültige, ein-

13 http://www.dnb.de/DE/Netzpublikationen/URNService/urn-service_node.html (Letzter Aufruf: 16. April 2012).

14 Der URN ist ein Uniform Resource Identifier, siehe hierzu: <http://www.w3.org/Addressing/> (Letzter Aufruf 11. Mai 2012).

Abbildung 4:
Übersichtsseite
des KWMBI,
Jahrgang 2012

Startseite | Kontakt | E-Mail-Abo | RSS | Banner-Download | Inhaltsverzeichnis | Bayerische Staatsregierung

BURGER | UNTERNEHMEN | VERWALTUNG | SERVICE CENTER | AKTUELLES

Verwaltung auf einen Klick.

Service-Center > BAYERN | RECHT > Verkündungsplattform > KWMBI

AMTSBLATT DER BAYERISCHEN STAATSMINISTERIEN FÜR UNTERRICHT UND KULTUS UND WISSENSCHAFT, FORSCHUNG UND KUNST (KWMBI)

Die Verkündungsplattform Bayern genießt Datenbankschutz gemäß § 9 (7a) # UrhG. Dies bezieht sich auch auf die einzelnen elektronischen Ausgaben der Amtsblätter und das Gesetz- und Verordnungsblatt. Vervielfältigung, Verbreitung oder öffentliche Wiedergabe sind daher nur in den Grenzen der §§ 9 (7b) und 9 (7c) UrhG statthaft.

Jahresausgabe - KWMBI: 2012

KWMBI-Nr.	Datum	Seiten	Inhaltsübersicht	PDF-Format
Nummer 7	17. Apr. 2012	132 - 140	[Icon]	[Icon] (190 KB)
Beiblatt Nr. 7			[Icon]	[Icon] (270 KB)
Nummer 6	27. März 2012	85 - 132	[Icon]	[Icon] (1,10 MB)
Beiblatt zu Nr. 6			[Icon]	[Icon] (297 KB)
Nummer 5	13. März 2012	45 - 64	[Icon]	[Icon] (405 KB)
Beiblatt zu Nr. 5			[Icon]	[Icon] (125 KB)
Nummer 4	29. Feb. 2012	41 - 44	[Icon]	[Icon] (92 KB)
Beiblatt Nr. 04			[Icon]	[Icon] (403 KB)
Nummer 3	14. Feb. 2012	37 - 40	[Icon]	[Icon] (457 KB)
Beiblatt zu Nr. 3			[Icon]	[Icon] (233 KB)
Nummer 2	31. Jan. 2012	29 - 36	[Icon]	[Icon] (581 KB)
Beiblatt zu Nr. 2			[Icon]	[Icon] (327 KB)
Nummer 1	10. Jan. 2012	1 - 28	[Icon]	[Icon] (795 KB)
Beiblatt zu Nr. 1			[Icon]	[Icon] (294 KB)

deutige Namen für Informationsressourcen im WWW konzipiert. Ziel ist es, die Unabhängigkeit vom Ort der Speicherung und vom Zugriffsprotokoll zu erreichen. Das Schema der URNs ist wiederum unterteilt in verschiedene Namensräume, die sowohl den Geltungsbereich als auch spezifische Regeln festlegen. Für den Bibliotheks- und Archivbereich ist der Namensraum der National Bibliography Number (NBN) maßgeblich. Dieser Namensraum ist international gültig und für Deutschland, Österreich und die Schweiz hat die DNB die Koordinierungsfunktion übernommen. Seit der Umstellung auf die elektronische Fassung verfügt jeder Verkündungsblatttitel entsprechend dem Katalogeintrag über eine eigene URN, die bei der DNB verzeichnet ist:

- AllMBl: urn:nbn:de:bvb:12-bsb_ad000000011
- JMBl: urn:nbn:de:bvb:12-bsb_ad000000029
- KWMBI: urn:nbn:de:bvb:12-bsb_ad000000032
- FMBl: urn:nbn:de:bvb:12-bsb_ad000000047
- GVBl: urn:nbn:de:bvb:12-bsb_ad000000056

3.2 Der Einsatz von PDF/A

Jede Datei besteht zunächst aus einer Folge von Bits (0/1) bzw. Bytes (bestehend aus 8 Bits), die aufgrund der Spezifikation von Dateiformaten mittels entsprechender Software interpretiert, d. h. semantisch korrekt ausgewertet werden, und schließlich am Bildschirm dargestellt werden kann. Für die digitale Langzeitarchivierung ist die Auswahl des Dateiformats für die langfristige Interpretierbarkeit sehr wichtig. Es wird zwischen folgenden Typen von Dateiformaten unterschieden:

- Proprietäre Dateiformate, deren Spezifikation oft nicht hinreichend bekannt ist.
- Offene Dateiformate, deren Spezifikation offen und frei zugänglich sowie gut dokumentiert ist. Sollte die zur Interpretation notwendige Software nicht mehr zur Verfügung stehen, aber eine umfassende Spezifikation vorliegen, so kann die gespeicherte Information auf anderen Wegen extrahiert werden (z. B. durch Konvertierung in ein anderes Dateiformat).

Zu einem formalen Standard kann ein Dateiformat werden, indem es einen Normungsprozess durchläuft, etwa auf nationaler Ebene durch das Deutsche Institut für Normung e.V. (DIN) bzw. international durch die International Organization for Standardization (ISO). Die Vorteile einer Standardisierung sind die offene, umfassende Dokumentation, die die weite Verbreitung und die Austauschbarkeit zwischen verschiedenen Systemen begünstigt und auch langfristig die Interpretierbarkeit sicherstellt. Deshalb ist die Verwendung von standardisierten Dateiformaten für die digitale Langzeitarchivierung von Vorteil.

Ein solch standardisiertes Dateiformat wurde mit PDF/A geschaffen. PDF/A wurde bereits 2005 als Standard für die Archivierung digitaler Dokumente normiert und findet seither immer größere Verbreitung¹⁵. Der Standard liegt bisher in zwei Teilen vor, ein dritter ist bereits in Arbeit:

- ISO-Standard ISO 19005-1:2005, Document management – Electronic document file format for long-term preservation – Part 1: Use of PDF 1.4 (PDF/A-1)
- ISO-Standard ISO 19005-2:2011, Document management – Electronic document file format for long-term preservation – Part 2: Use of ISO 32000-1 (PDF/A-2)
- ISO-Standard ISO 19005-3 in Arbeit

Der Standard setzt auf dem PDF-Format („Portable Document Format“) auf, das seit seiner Einführung 1993 durch *Adobe*[®] den Datenaustausch zwischen verschiedenen Systemen enorm vereinfacht hat. Die Vorteile von PDF-Dokumenten sind, dass sie unter allen gängigen Betriebssystemen gelesen werden können, dass kostenfreie Anzeigeprogramme zur Verfügung stehen, Struktur und Layout können originalgetreu angezeigt werden und nicht zuletzt steht die Entwicklung von Software zur Nutzung des Formats auch anderen Firmen offen, da das Format durch eine entsprechende Dokumentation offen gelegt ist. Um jedoch eine Reproduzierbarkeit sowohl des visuellen Erscheinungsbildes als auch der logischen Struktur eines Dokuments auch noch nach mehreren Jahren garantieren zu können, müssen weitere Anforderungen erfüllt sein, die über das PDF-Format hinausgehen. Deshalb wurde das Dateiformat PDF/A entwickelt und standardisiert, welches „Access to Content“, d. h. die Zugänglichkeit der Dateiinhalte unabhängig von Software und System garantiert. Innerhalb des PDF/A-1- und PDF/A-2-Standards wird

noch einmal unterschieden zwischen zwei Konformitätsstufen:

- PDF/A-1b (= basic): Level B hat die visuelle Reproduzierbarkeit im Fokus und garantiert zwar eine korrekte Anzeige der Inhalte, aber nicht deren Struktur.
- PDF/A-1a (= accessible): Level A stellt sicher, dass die logische Dokumentenstruktur, der Unicode-Zeichensatz und die natürliche Leseordnung des Textes, also die semantische Korrektheit und Struktur erhalten bleibt.

Voraussetzung für eine Konvertierung in die erste Konformitätsstufe, PDF/A-1b, ist eine visuelle Reproduzierbarkeit des Dokuments, d. h. dass die in den Archivstandard zu konvertierenden Dokumente „komplett“ sind und alle Informationen zum Dokument, wie etwa die verwendeten Schriftarten oder Farben, direkt im Dokument selbst abgespeichert sind. Es dürfen keine externen Abhängigkeiten wie systemabhängige Schrifttypen bestehen. So ist sichergestellt, dass beispielsweise die verwendete Schrift innerhalb eines Dokuments auch dann noch korrekt dargestellt werden kann, wenn sie auf dem später verwendeten

Band 38

Katrin Gärtner

Recommendersysteme

Die Arbeit beschäftigt sich mit der Analyse von in Deutschland vorkommenden Empfehlungssystemen. Zum einen werden die Grundlagen zu den Empfehlungssystemen, ihre grundlegende Funktionsweise und die zugrundeliegenden Techniken von Recommendersystemen erläutert, um dann, anhand von ausgewählten Internetseiten, Empfehlungssysteme in Deutschland näher zu betrachten, sie zu analysieren und in die aktuell gültige Klassifikation einzuteilen.

Um die Relevanz von automatisch generierten Empfehlungen zu überprüfen, wird außerdem eine stichprobenartige Umfrage unter Wissenschaftlern mehrerer Fachgebiete durchgeführt.

ISBN 978-3-934997-43-1, 2012, Brosch., 174 Seiten, teilweise farbig, € 24,50*

* Preise zzgl. Versandkosten (Inland 1,30 €, Europa 3,50 €)



¹⁵ DRÜMMER, Olaf/ OETTLER, Alexandra, VON SEGGERN, Dietrich: PDF/A kompakt. Digitale Langzeitarchivierung mit PDF. Berlin 2007.

Rechner nicht mehr zur Verfügung steht. Daneben sind bestimmte PDF-Merkmale nicht erlaubt und verhindern eine Konvertierung nach PDF/A, so z. B. interaktive Elemente, Verschlüsselungen oder PDF-Ebenen. Eine Auswahl an administrativen sowie deskriptiven Metadaten innerhalb des Dokuments ist verpflichtend vorgeschrieben.

Die zweite Konformitätsstufe, PDF/A-1a, geht noch weiter und fordert darüber hinaus zusätzliche Informationen zur logischen Strukturierung des Dokuments sowie zur Anordnung der Seiten, einzelnen Abschnitte und Paragraphen.

Die 2011 verabschiedete Norm für PDF/A-2 ergänzt den ersten Normteil, sodass keine Migration von PDF/A-1 auf -2 notwendig wird. PDF/A-2 bietet einige Erweiterungen, so z.B. die in manchen Fällen notwendige Bildkompression nach JPEG2000. Ein dritter Normteil, PDF/A-3, zielt auf dynamische Elemente wie etwa Audio- oder Videodateien in PDF-Dokumenten ab und soll 2012 folgen.

Die Amtsblätter sowie das GVBl auf der Verkündungsplattform Bayern werden im PDF/A-1b-Standard abgespeichert, dadurch wird die originalgetreue Reproduzierbarkeit der Inhalte auf verschiedensten Systemen garantiert. Zusätzlich werden im Dokument Metadaten mit Informationen zu Titel (z. B. „Amtsblatt der Bayerischen Staatsministerien für Unterricht und Kultus und Wissenschaft, Forschung und Kunst 2012-05“) und Verfasser (z. B. „Bayerisches Staatsministerium für Unterricht und Kultus“) des jeweiligen Amtsblattes gespeichert. Auch wird noch einmal explizit festgehalten, dass das Amtsblatt frei von Urheberrechten ist.

3.3 Langzeitspeicherung am Leibniz-Rechenzentrum

Der sichere physische Erhalt der digitalen Daten über lange Zeiträume hinweg stellt den technischen Kern der digitalen Langzeitarchivierungsmaßnahmen dar. Für die Definition und Durchführung einer zuverlässigen Speicherstrategie („Archival Storage“ im Sinne des OAIS-Referenzmodells) der digitalen Objekte der BSB ist das Leibniz-Rechenzentrum, der Service-Provider der BSB für das Hosting der Serverinfrastruktur und besonders für Archivierung der großen Datenmengen (aus der Retrodigitalisierung), zuständig.¹⁶ Es setzt zur Langzeitspeicherung der Daten das Archiv- und Backupsystem unter dem Softwarepaket TSM (Tivoli Storage Manager) der Firma IBM ein. TSM ist

¹⁶ REINER, Bernd: Sicherung des Weltkulturerbes am Leibniz-Rechenzentrum, in: Akademie Aktuell, Zeitschrift der Bayerischen Akademie der Wissenschaften 3 (2007). Das MDZ ist mittlerweile der drittgrößte Kunde des LRZ bei der Datenarchivierung.

als Client-Server-Anwendung für heterogene Rechnerlandschaften konzipiert und basiert auf offenen Standards wie dem IEEE Mass Storage Reference Model¹⁷. Einige TSM-Funktionalitäten sind speziell auf die Langzeitarchivierung digitaler Objekte ausgelegt, so z. B. die gute Skalierung des Gesamtsystems, was bei einem jährlichen Datenzuwachs des BSB-Archivspeichers von ca. 100 TB essentiell ist. Auch bezüglich der Heterogenität von Dateiformaten liegt keine Beschränkung vor, d. h. das Archivsystem kann mit verschiedensten Dateiformaten umgehen. Alle gespeicherten Daten und ihre Metadaten werden automatisch und zentral durch ein Datenbankmanagementsystem verwaltet, worüber auch der Zugriff auf die Daten kontrolliert wird. Die Sicherung der digitalen Objekte passiert automatisch nach individuell definierten Zeitintervallen, sodass die Daten redundant und konsistent vorgehalten sind. Für das Datenbankmanagementsystem selbst sind Sicherungskonzepte vorhanden. Da die Übertragung der Daten vom Client zum Server oftmals eine Fehlerquelle darstellt, findet an dieser Stelle eine Validierung der digitalen Objekte durch „Cyclic Redundancy Checks“ (CRC) statt. Sollte bei der Datenübertragung oder auch durch andere Formen des Datenzugriffs ein Fehler festgestellt werden, können die Objekte „recycelt“ werden, was bedeutet, dass die fehlerhaften Objekte durch Sicherungskopien ersetzt werden. Im Hinblick auf die Langfristigkeit spielt die regelmäßige Migration aller Daten auf neue physische Datenträger eine besondere Rolle.

Wie alle digitalen Dokumente der BSB werden auch die Verkündungsblätter über das TSM in regelmäßigen Abständen in das Archivsystem des LRZ kopiert.

3.4 Langzeiterhaltungsstrategien

Um die langfristige Interpretierbarkeit digitaler Objekte garantieren zu können, ist es notwendig, geeignete Langzeiterhaltungsstrategien zu entwickeln, wie im OAIS-Referenzmodell unter dem Aspekt „Preservation Planning and Action“ ausgeführt wird. Sowohl eine Dateiformatmigration als auch eine Hard- und Software-Emulation sind mögliche Vorgehensweisen, die je nach Beschaffenheit des digitalen Objekts auszuwählen sind, wenn die Gefahr besteht, dass Dateiformate z. B. einer bestimmten Kollektion von Objekten aufgrund des Alters des Formats nicht mehr interpretiert werden können, da es keine geeignete Unterstützung durch entsprechende Software mehr gibt.

¹⁷ <http://cdsweb.cern.ch/record/281298/files/p44.pdf> (Letzter Aufruf: 21. Mai 2012).

Für eine Migration in ein neues Dateiformat sprechen neben dem Erhalt der Interpretierbarkeit auch andere Kriterien, wie etwa die Größe des Dateiformats oder die (fehlende) Standardisierung und damit Verbreitung des Formats. Als Beispiel sei die Evaluation des JPEG2000-Dateiformats an der BSB genannt, dessen Verwendung anstelle des Speicherformats TIFF im Rahmen eines „Preservation Planning“-Prozesses geprüft wurde. Ein Grund für den Test war die mögliche Kostenersparnis durch den Einsatz von JPEG2000.¹⁸ Unter Emulation versteht man die Nachbildung der originalen technischen Umgebung eines archivierten Objekts in einer aktuellen Hard-/Software-Umgebung zum Zwecke der Interpretierbarkeit ohne das originale Dokument zu verändern. Emulation kann auf der Ebene der Anwendungssoftware, des Betriebssystems oder auch der Hardware stattfinden. Dies kann beispielsweise bei komplexen, dynamischen Objekten wie Computerspielen oder interaktiven Multimediaanwendungen sinnvoll sein. Durch die Wahl des ISO-Standards PDF/A als Archivierungsformat für die Amtsblätter der bayerischen Staatsministerien wurde bereits ein Dateiformat gewählt, das speziell für die Langzeitarchivierung entwickelt wurde, weshalb in absehbarer Zeit keine Migration oder Emulation notwendig sein wird. Ein ständiger „Technology Watch“ gehört jedoch zu den Routineaufgaben der LZA-Mitarbeiter an der BSB, um neue Entwicklungen und Trends in der Formatlandschaft sofort in die aktuelle LZA-Strategie miteinbeziehen zu können.

4 Ausblick

Die BSB konnte aufgrund ihrer langjährigen Erfahrung und Kenntnisse im Bereich der digitalen Langzeitarchivierung die Verkündungsblätter der bayerischen Staatsministerien problemlos in ihren bestehenden Archivierungsworkflow integrieren. Durch die enge Kooperation mit dem LRZ kann der physische Datenerhalt, durch die eingeleiteten Langzeitarchivierungsmaßnahmen der Zugang und die Interpretierbarkeit der Amtsblätter auf Dauer garantiert werden. Derzeit führt die BSB mit Unterstützung des Baye-

rischen Staatsministeriums für Wissenschaft, Forschung und Kunst eine neue Langzeitarchivierungssoftware, Rosetta Digital Preservation System der Firma Ex Libris, für die Langzeitarchivierung ein¹⁹. Rosetta wird die neue Software-Infrastruktur für die gemeinsam von BSB, d. h. dem MDZ und der Verbundzentrale des Bibliotheksverbands Bayern, und dem LRZ betriebene technische Infrastruktur zur digitalen Langzeitarchivierung werden und so die seit dem ersten BABS-Projekt (2005–2007) bestehende technische Infrastruktur ergänzen. Die Einführung des neuen, vollständig OAIS-konformen Langzeitarchivierungssystems bietet auch die Chance, im Sinne der Vertrauenswürdigkeit und Skalierbarkeit deutliche Verbesserungen herbeizuführen, z. B. im Bereich der Qualitätssicherung, der Planung und Durchführung von Langzeitarchivierungsmaßnahmen („Preservation Planning and Action“) und bei der kooperativen Langzeitarchivierung.

Die Amtsblätter der bayerischen Staatsministerien sowie das GVBI werden mit hoher Priorität ebenfalls in den Routinebetrieb der neuen Langzeitarchivierungssoftware migriert werden. ■

.....
Anna Kugler, Markus Brantl, Astrid Schoger

Bayerische Staatsbibliothek
 Münchener Digitalisierungs-Zentrum (MDZ)/
 Digitale Bibliothek
 Munich Digitization Centre/
 Digital Library Bayerische Staatsbibliothek
 Ludwigstr.16
 80539 München
 Anna.Kugler@bsb-muenchen.de
 Markus.Brantl@bsb-muenchen.de
 Astrid.Schoger@bsb-muenchen.de

¹⁸ BEINERT, Tobias/ BRANTL, Markus/ KUGLER, Anna/ KULOVITS, Hannes/ RAUBER, Andreas/ SCHOGER, Astrid: From TIFF to JPEG 2000? Preservation Planning at the Bavarian State Library Using a Collection of Digitized 16th Century Printings. In: D-Lib Magazine Vol. 15, Nov/Dec 2009, verfügbar unter: <http://www.dlib.org/dlib/november09/kulovits/11kulovits.html> (Letzter Aufruf: 25.04.2012).

¹⁹ Brantl, Markus/ Ceynowa, Klaus/ GROSS, Mathias/ SCHOGER, Astrid: Digitale Langzeitarchivierung in Bayern. Vom explorativen Projekt zum nachhaltigen Modell. In: BIBLIOTHEK Forschung und Praxis. Band 35, Heft 1, S. 15–25 (April/2011).