Keine Angst vor der "Schlange" das Landesarchiv BadenWürttemberg als Partner des
Webarchivs BOA (BadenWürttembergisches Online-Archiv)

© Johannes Renz, Landesarchiv Baden-Württemberg, 2013



Warum betreibt das Landesarchiv Baden-Württemberg Webarchivierung?

- Gesetzlicher Auftrag: Das Landesarchiv Baden-Württemberg ist auch für "sonstige Informationsträger und maschinenlesbar auf diesen gespeicherte Informationen und Programme" zuständig (§ 2 LArchG Baden-Württemberg)
- Handlungsgrundlage: Landesarchivgesetz Baden-Württemberg (http://www.landesarchiv-bw.de/sixcms/media.php/120/49996/LArchG.29734.pdf)



Webarchivierung – Sonderform der Digitalen Archivierung

- Webarchivierung muss sich mit einer großen Bandbreite von Dateitypen auseinandersetzen, die durch Hyperlinks stark miteinander in Beziehung stehen
- Besondere Herausforderungen an die dauerhafte Haltbarkeit und spätere Migration (noch viele ungelöste Probleme)
- Es gilt der Grundsatz "Collect now ask later why" (schnelle Sicherung potenziell historisch wertvoller Überlieferung hat Priorität)



Webarchivierung – Sonderform der Digitalen Archivierung

- Webseiten sind sowohl Archivgut als auch Sammlungsgut
- Webseiten benötigen besondere Plattformen der Speicherung
- Es gilt grundsätzlich der in der Digitalen Archivierung verpflichtende OAIS-Standard
- -> prinzipielle Unterscheidung zwischen SIP, AIP und DIP



Webarchivierung – Sonderform der Digitalen Archivierung

- In der Webarchivierung verschwimmen diese Unterschiede Bereitstellung archivierter Webseiten erfolgt gleichzeitig mit der Sicherung ("SIP" gilt als vorläufiges "AIP" und kann auch als "DIP" verwendet werden).
- Mit Blick auf die Archivierung von (vertraulichen!)
 Intranetseiten muss diese Unterscheidung aber im Hinterkopf behalten werden!!!



Welche Webseiten kommen für eine archivische Überlieferung vorrangig in Frage?

- Homepages von Behörden und Institutionen:
- Wichtige Aushängeschilder nach außen (Selbstverständnis)



 Komprimierte Abbildung von Struktur und Arbeitsweise einer Behörde (verwaltungsgeschichtlicher Aspekt)



 Organigramme und Geschäftsverteilungspläne oftmals nur noch online verfügbar (Wegfall entsprechender analoger Überlieferung)



Welche Webseiten kommen für eine archivische Überlieferung vorrangig in Frage?

- Informationsportale:
- Allgemeines (z. B. service-bw)



- Bestimmte Verwaltungsgebiete (z. B. Kultusportal)
- Anlaufstellen für die Privatwirtschaft (z. B. gruendung-bw)



- Daseinsvorsorge (z. B. Ernährungsportal)







Welche Webseiten kommen für eine archivische Überlieferung vorrangig in Frage?

- Regierungskampagnen ohne speziell dafür erstellte Webseiten heute kaum mehr denkbar, z. B.
- Imagekampagne 2006 ("Wir können alles außer Hochdeutsch!")
- "50 80 90" (Kampagne zur Energiewende)

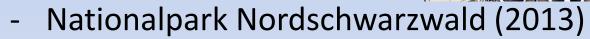






Welche Webseiten kommen für eine archivische Überlieferung vorrangig in Frage?

- Plattformen für Bürgerbeteiligung
- Volksabstimmung zu Stuttgart 21 (2011)
- Wiedereinführung der Verfassten Studentenschaft (2012)
- Landesjubiläum (2012)
- Beteiligungsportal (2013)













Bedeutung von Webseiten für die archivische Überlieferung (Zusammenfassung)

- Dokumentation der Informationsgesellschaft
- Das Archiv wird auch hier als "Gedächtnis der Verwaltung" tätig
- Bei amtlichen Webseiten zeitnahe Bereitstellung für die historische Forschung möglich (keine Sperrfristen, vgl. analoge und digitale Sachakten: 30 Jahre)



Wo und wie wird archiviert?

- Das Landesarchiv Baden-Württemberg betreibt seit 2006 ein Webarchiv für Internetseiten
- Es ist Verbundpartner im Projekt "Baden-Württembergisches Online-Archiv" (BOA) des Südwestdeutschen Bibliotheksverbunds beim Bibliotheksservicezentrum Baden-Württemberg (BSZ) in Konstanz (<u>www.bsz-bw.de</u>)
- Die in BOA archivierten Webseiten können öffentlich eingesehen werden (<u>www.boa-bw.de</u>)
- Eine Indexierung durch einschlägige Suchmaschinen wird jedoch abgelehnt

Wo und wie wird archiviert? (Fortsetzung)

- Aufgaben des BSZ
- stellt auf BOA ein Content-Management-System (SWBcontent) zur Verfügung
- entwickelt bzw. implementiert die notwendigen Tools
- allgemeiner technischer Support
- Jährlicher Austausch (Anwendertreffen) der SWBContent-Partner (u. a. Landesarchiv, Kommunalarchive, Dt. Literaturarchiv, Landes- und Hochschulbibliotheken)



Wo und wie wird archiviert? (Fortsetzung)

- Absprachen über den Inhalt von BOA (welche Webseiten werden gespiegelt?) mit den fachlich für die jeweiligen Provenienzbestände zuständigen Archivmitarbeitern, die auch für die Erschließung der Webseitenbestände zuständig sind
- Spiegelung der Webseiten durch das Mitarbeiterteam des "Digitalen Landesarchivs".
- Webcrawler: Heritrix (Standard ab Ende 2012 -> WARC-Format), Viewer: Wayback Machine) bzw. Httrack (Standard 2006-2012, seitdem nur noch in Ausnahmefällen, künftig: Konvertierung nach WARC)

Wo und wie wird archiviert? (Fortsetzung)

- Backend: Bildung "archivischer Bestände" nach dem Provenienzprinzip, welche in die Abteilungs- und Beständestruktur des Landesarchivs Baden-Württemberg eingebettet sind.
- Nachweis/Erschließung (inhaltliche Metadaten) in der Fachdatenbank ScopeArchiv
- Frontend: Bereitstellung der Erschließungsdaten über das Online-Findmittelsystem des Landesarchivs Baden-Württemberg (OLF 21), von dort wiederum Verlinkung auf BOA. Nutzung vom heimischen PC aus weltweit möglich.

Die bei der Webarchivierung eingesetzten Plattformen

ScopeArchiv

Erschließungsdatenbank, auf Oracle aufgesetzt -> inhaltliche Metadaten

OLF 21

Online-Findmittelsystem, mySQL-Datenbank -> inhaltliche Metadaten

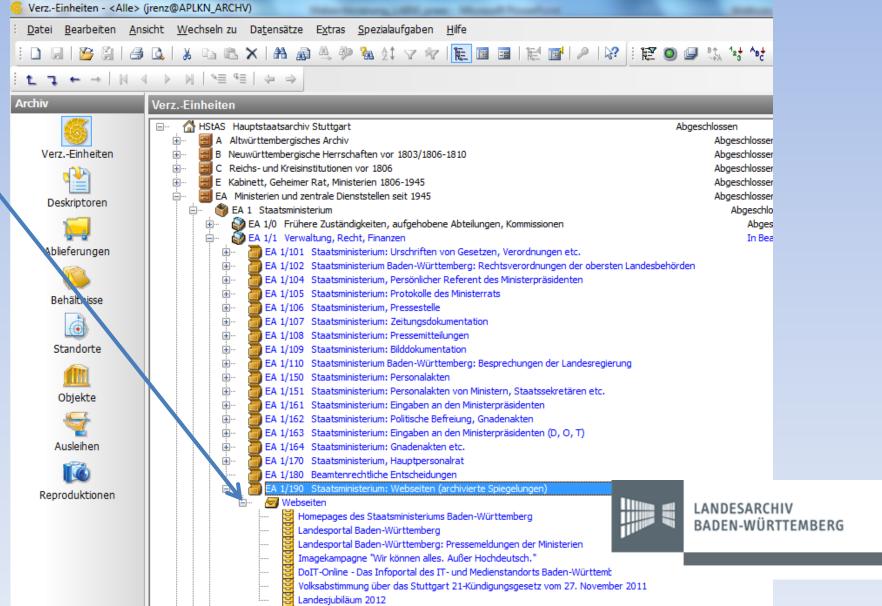
BOA/SWBContent

eingebunden: HTTRACK bzw. Heritrix/Wayback Machine) - > technische Metadaten Digitales Magazin (DIMAG)

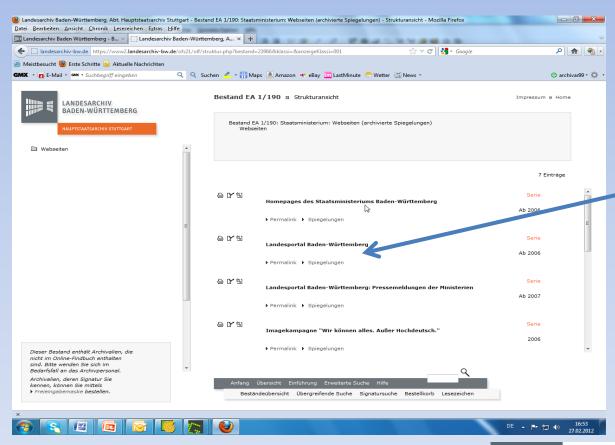
Archivierung von Intranetseiten u. a. (im Aufbau)



Einbindung von Webseiten in die Beständestruktur eines Archivs (hier: Hauptstaatsarchiv Stuttgart)



Blick in ein archivisches Online-Findmittel des Hauptstaatsarchivs Stuttgart mit Verlinkung auf BOA





Beschreibung einer Spiegelungsserie in BOA

BUA	BADISCHE LANDESBIBLIOTHEK 🖨 LANDESARCHIV BADEN-WÜRTTEMBERG 🖨 WÜRTTEMBERGISCHE LANDESBIBLIOTHEK 📑	Home Kontakt Hilfe
BOA Baden-Wür	ttembergisches Online-Archiv Landesarchiv	
EINFACHE SUCHE ERWEITERTE SUCHE BROWSEN EDITIEREN		
Detailansicht		
Zurück zu den Suchergebnissen		
Staatsministerium Baden-Württemberg Provenienz: Staatsministerium Baden-Württemberg ScopeID: 1-1007654		
Signatur: HStAS EA 1/1 Dokumentart: 22 Home		
DDC: 340 Recht DDC: 350 öffentliche Verwaltung		
Bemerkung: Archivierungszyklus 6 Monate. Darstellung mit Internet Explorer fehlerhaft. Spiegelungen unvollständig. Nächste Speicherung: 2013-04-25		
Archiv-URL: http://la.boa-bw.de/1-1007654.html Archiv-URN: urn:nbn:de:bsz:boa-9-1-10076545		
Zugehörige Dok	umente	



Einzelspiegelungen: Übergang von Httrack nach Heritrix

Stand 27.4.2012

Signatur: HStAS EA 1/190

Bemerkung: Vorgezogene Spiegelung anlässlich des Landesjubiläums und des Antrittsbesuchs von Bundespräsident Joachim Gauck. Youtube-Kanal der

Landesregierung mit eingeschränkte Funktionalität ist nur über das Youtube-Symbol

auf der Startseite erreichbar. Direktlinks auf die Videos funktioieren nicht.

ursprüngliche URL: http://www.stm.baden-wuerttemberg.de

Anzahl der Dateien: 4845

(rights): Free

Übernahmedatum: 27.04.2012

Einsteller: Landesarchiv

Archiv-URN: urn:nbn:de:bsz:boa-9-c24f5a8b-b327-4e7a-933e-3641f03a124e9

zum archivierten Objekt (Format: text/html , Gesamtgröße:2,32 GByte)

Stand 25.10.2012

Signatur: HStAS EA 1/190

Bemerkung: Vorgezogene Spiegelung auf Grund der Wahl von Ministerpräsident

Winfried Kretschmann zum Bundesratspräsidenten

ursprüngliche URL: http://www.stm.baden-wuerttemberg.de/

(rights): Free

Übernahmedatum: 26.10.2012

Einsteller: Landesarchiv Anzahl der Dateien: 1

<u>zum archivierten Objekt</u> (Format: text/html , Gesamtgröße:7,6 GByte)

Was wird archiviert? (Ist-Stand)

- Webseiten, die nach aktuellem Stand archiviert werden:
 - Landesministerien und Oberste Landesbehörden, incl.
 Webportale und eine Auswahl an (oftmals kurzfristig online geschalteten) Sonderseiten (z. B. Imagekampagne, Volksabstimmung Stuttgart 21, Landesjubiläum, Nationalpark...)
 - Regierungspräsidien
 - Oberlandesgerichte
 - Generalstaatsanwaltschaften
 - Pädagogische Hochschulen und Fachhochschulen
 - Wissenschaftliche und kulturelle Institutionen
 - Im Rahmen des Archivverbundes Main Tauber: Webseiten des Main-Tauber-Kreises, der Stadt Wertheim und kreisangehöriger Städte und Gemeinden

Beispiel aus der Anfangszeit von BOA: Homepage des Staatsministeriums Baden-Württemberg, Stand: 14.11.2006





Sonderseite: Volksabstimmung über das S21-Ausstiegsgesetz, Stand 22.11.2011



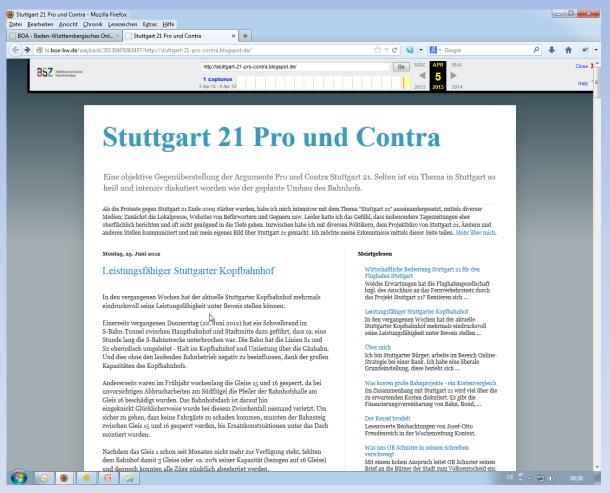


Besondere Einzelfälle

- In Einzelfällen: Archivierung der Webseiten von Körperschaften, Anstalten, Stiftungen, Verbänden und Vereinigungen nach § 2 Abs. 3 LArchG (bereits realisiert: zwei landeseigene Stiftungen).
- Aufbau zeitgeschichtlicher Sammlungen von Webseiten (aktuell: Blogs zum Thema Stuttgart 21)
- Archivierung ausgewählter Intranetseiten von Landesbehörden: Auf Grund der Vertraulichkeit der Informationen können diese Seiten nicht öffentlich über ein Portal wie BOA präsentiert werden, sondern müssen in den digitalen Langzeitspeicher (DIMAG) überführt werden.



Blog: Stuttgart 21 Pro und Contra, Stand 5.4.2013





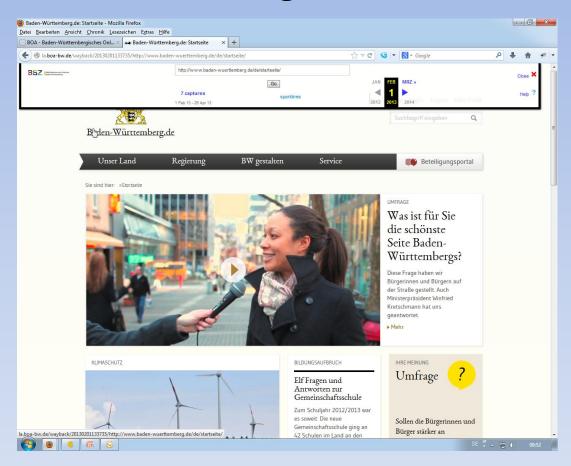
- Auch im Bereich der nichtstaatlichen Überlieferung werden Archive - wie bei Behördenüberlieferung zunehmend mit genuin digitalen Unterlagen zu tun haben.
- Private Homepages künftig auch Bestandteile von Nachlässen?
- Nachlassgeber sollten ggf. bereits zu Lebzeiten ihre (digitalen) Unterlagen einem Archiv überlassen, ansonsten Obsoleszenzgefahr
- Urheberrechte und Rechte Dritter müssen noch sorgfältiger überprüft werden

- Derzeit ist die Webarchivierung, insbesondere nichtstaatlicher Provenienzen, in Deutschland eine ziemlich große rechtliche Grauzone
- -> Das Landesarchiv Baden-Württemberg hat inzwischen eine Mustervereinbarung für die Archivierung privater Webseiten und Blogs entworfen
- -> Webseitenbetreiber der Landesverwaltung sind angehalten, ihren Webcontent ggf. einzukaufen, damit Inhalte dauerhaft genutzt werden können



- Endgültiger Umstieg von HTTRACK auf HERITRIX (derzeit "zweigleisig", User Interface auf BOA teilweise implementiert)
- Zunahme von Web 2.0-Elementen (z. B. Einbindung von Video-Podcasts aus Youtube oder anderen externen Portalen, Bildergalerien, Mediatheken)
- Zunehmende Einbindung externer Inhalte bzw. "Filesharing" durch mehrere Plattformen (sehr stark bei der Webüberlieferung zu Stuttgart 21, aber auch bei staatlicher Webüberlieferung zu beobachten)
- Folgen:
- Aus klassisch-archivischer Sicht: Verwischung der Provenienzen durch Verzahnung von Webinhalten
- Urheberrechtliche Aspekte werden u. U. komplizierter





Neues Landesportal mit Videoumfrage (Stand: 1.2.2013)



- Outsourcing der Erstellung und Pflege von Behördenwebseiten (Auswirkung auf Urheberrechte?)
- Weitere Herausforderungen bei nichtstaatlichen Webseiten:
- Abgrenzung der Zuständigkeiten mit anderen Institutionen
- Sorgfältige archivische Bewertung
- Erarbeitung einer künftigen Migrationsstrategie für Webinhalte (betrifft alle Webarchive gleichermaßen)



Fazit:

- Webarchivierung ist ein integraler Bestandteil im Auftrag des Landesarchivs Baden-Württemberg als kulturelle Gedächtnisinstitution
- Wie in allen Bereichen der digitalen (Langzeit-)
 Archivierung wächst die Bedeutung der Kooperation mit anderen Einrichtungen, auch länderübergreifend
- Bewusstsein für die Aufgabe wächst kontinuierlich
- -> Möglichkeiten zum Austausch/Support sind vorhanden, daher:
- -> keine Angst vor der Schlange!!!



Vielen Dank für Ihre Aufmerksamkeit!

Johannes Renz

Landesarchiv Baden-Württemberg – Hauptstaatsarchiv Stuttgart

Konrad-Adenauer-Str. 4

D-70173 Stuttgart

Tel.: +49711/212-4330

E-Mail: johannes.renz@la-bw.de

