

Archivierung von Webseiten

Der Archivierung von Inhalten des Internets, speziell von Webseiten, wird eine wachsende Relevanz zugemessen. Als Plattformen von Forschung und Wissenschaft, aber auch als Medium für Kultur und Bildung sind ihre Inhalte dauerhaft zu dokumentieren und zu sichern.

Die besondere Herausforderung liegt dabei in der inhaltlichen, graphischen und technischen Dynamik der präsentierten Inhalte, die den gewohnt schnellen Wechsel von Dateiformaten und Speichermedien der Informationstechnik deutlich übertrifft.

Die Gesamtaufgabe kann von Archiven, Bibliotheken, Rechenzentren und Museen nur gemeinsam und kooperativ unter Nutzung spezieller Software gelöst werden, die darüber hinaus noch laufend auf den spezifischen Bedarf angepasst werden muss.

Die Software SWBcontent

Zur Sicherung, Erschließung und Präsentation von Webseiten und digitalen Medienobjekten entwickelt das BSZ die Archivsoftware SWBcontent nach dem Bedarf von Bibliotheken und Archiven sowie verwandten Einrichtungen: In die Software können bibliographische Daten aus dem Südwestdeutschen Bibliotheksverbund in vollem Umfang geladen und aktualisiert werden; für die archivische Dokumentation steht ein Beschreibungsformat bereit, das unter Mitwirkung des Landesarchivs Baden-Württembergs festgelegt wurde. Einzeldokumente, z.B. im PDF-Format, können einfach hochgeladen werden. Mit Hilfe der Offline-Browser HTTrack bzw. Heritrix werden Webseiten übernommen.

Ein Alleinstellungsmerkmal der Software bildet die ausgeprägte Eignung, hierarchische Zusammenhänge darstellen zu können. Die automatische Vergabe und Registrierung persistenter Identifier (URN) gehört zum Leistungsumfang ebenso wie die Erhebung technischer Metadaten der archivierten Objekte. Die Einbettung in die Geschäftsgänge der betreibenden Institutionen wird von den Anwendern als ausgesprochen komfortabel gewertet.

Kooperation und Standardisierung

Das BSZ betreibt SWBcontent und entwickelt es laufend; die Auftraggeber wählen und erschließen das Archivgut und speichern es ab.

Diese Aufgabenteilung entspricht dem Schichtenmodell der digitalen Langzeitarchivierung des Ministeriums für Wissenschaft, Forschung und Kunst Baden-Württemberg (<http://opus.bsz-bw.de/swop/volltexte/2010/818/>).



Die Sammelgebiete werden z.B. in Baden-Württemberg zwischen Landesarchiv und -bibliotheken abgesprochen, um Überschneidungen zu vermeiden. Auch mit dem Programm der Deutschen Nationalbibliothek DNB zur Speicherung von Webseiten werden nicht alle regionalen und fachlichen Bedarfe abgedeckt werden, sondern eigene Anstrengungen erforderlich bleiben.

In Kooperation mit der DNB und der Bayerischen Staatsbibliothek veranstaltet das BSZ im Rahmen von nestor, dem deutschen Kompetenznetzwerk zur digitalen Langzeitarchivierung, Workshops, die die Projekte zur Webseitenarchivierung vernetzen.

Das BSZ beobachtet die technische Entwicklung, übernimmt Bausteine und nutzt Standards dort, wo sie existieren bzw. beteiligt sich an deren Entwicklung.

Installationen von SWBcontent

Der Ursprung von SWBcontent ist das **Baden-Württembergische Online-Archiv BOA** (<http://www.boa-bw.de/>), bei dem seit 2003 die Badische Landesbibliothek, die Württembergische Landesbibliothek, das BSZ und seit 2006 das Landesarchiv Baden-Württemberg kooperieren, um elektronische Publikationen zu archivieren.

Die Saarländische Universitäts- und Landesbibliothek nutzt **SaarDok** (<http://saardok.sulb.uni-saarland.de/>) seit 2003, um elektronische Publikationen und Websites mit Bezug zum Saarland zu erschließen, der Öffentlichkeit zur Verfügung zu stellen und zu archivieren.

In **Literatur im Netz** (<http://literatur-im-netz.dla-marbach.de>) sammelt das Deutsche Literaturarchiv Marbach seit 2007 literarische Online-Zeitschriften, Weblogs und Netzliteratur.

Mit Hilfe des **Publikationsservers im Informations- und Bibliotheksportal im Intranet des Bundes** werden Publikationen von Bibliotheken der obersten Bundesbehörden gegenseitig zur Verfügung gestellt und im Rahmen der Pflichtablieferung an die Deutsche Nationalbibliothek (DNB) übertragen. Der Publikationsserver ist nicht öffentlich zugänglich.

Seit 2001 sammelt das **Zentralarchiv zur Erforschung der Geschichte der Juden in Deutschland** Webseiten aus dem organisierten jüdischen Leben der Bundesrepublik. Dieses Webseitenarchiv wurde 2011 auf SWBcontent migriert.

Seit 2009 stellt das BSZ **SWBdok** (<http://swbdok.bsz-bw.de>) für alle Teilnehmerbibliotheken des SWB zur Verfügung. In SWBdok können freie Internetpublikationen eingestellt werden. SWBdok stellt ihre dauerhafte Verfügbarkeit und Zitierbarkeit sicher.

In **SWBregio** (<http://swbregio.de>) werden Webseiten von Kommunen und den mit ihnen verbundenen Körperschaften archiviert. SWBregio ist eine Kooperation des Landesarchivs Baden-Württemberg und des BSZ, die 2011 vereinbart wurde.

Bibliotheksservice-Zentrum Baden-Württemberg (BSZ)

Das Bibliotheksservice-Zentrum Baden-Württemberg (BSZ) ist eine Einrichtung des Landes Baden-Württemberg und gliedert sich in drei Bereiche:

- **SWB-Verbundsystem:**

Betrieb der Verbunddatenbank des Südwestdeutschen Bibliotheksverbunds (SWB) zur Katalogisierung, Literaturrecherche und Fernleihe.

- **Bibliothekssysteme:**

Betreuung von Bibliothekssystemen für Ausleihe, Nutzerkonten, Erwerbung etc. für über fünfzig Hochschulen.

- **Museen, Archive und Repositorien (MARE):**

Dokumentations- und Präsentationswerkzeuge für Archive und Museen; BAM- und weitere Portale zur Internetrecherche, Institutionelle und fachliche Repositorien, digitale Langzeitarchivierung und E-Mail-Auskunft.



Ihre Ansprechpartner für
SWBcontent im BSZ

Koordination

Stefan Wolf
Tel: 07531 / 88 3050
stefan.wolf@bsz-bw.de

Renate Hannemann
Tel: 07531 / 88 4951
renate.hannemann@bsz-bw.de

Weitere Informationen finden Sie unter
<http://www.bsz-bw.de/mare/lza/swbcontent.html>



Postanschrift

BSZ Baden-Württemberg
Universität Konstanz
78457 Konstanz

zentrale@bsz-bw.de
www.bsz-bw.de

Besuchsadresse

Felix-Wankel-Straße 4
78467 Konstanz

Tel: 07531 / 88 2929
Fax: 07531 / 88 3703

Leiterin Ltd. Bibl. Dir. Dr. Marion Mallmann-Biehler
Stellvertreter Dipl. Math. Volker Conradt

Mai 2012

Langzeitarchivierung
von Webinhalten am BSZ

SWBcontent

Bibliotheksservice-Zentrum
Baden-Württemberg

www.bsz-bw.de