Archivierung von Webseiten

Webseiten gelten als Quellengattung, die in öffentlichen Archiven zur Überlieferungsbildung herangezogen wird. Als Selbstdarstellung in der Öffentlichkeit, als Kommunikationsmittel der Verwaltung und als interaktives Medium sind sie Plattformen für bürgerschaftliche Information, Kultur und Bildung, Forschung und Wissenschaft.

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

Technische Hilfsmittel stehen zur Verfügung, doch ist es nicht rationell, regionale Webauftritte durch die Anbieter auf jeweils eigener Hard- und Software zu sichern; die Investitionen in Technik und Know-how sind zu groß. Die Bundeskonferenz der Kommunalarchive weist deshalb in zwei Empfehlungen zur Speicherung von kommunalen Webseiten auf die Möglichkeit hin, diese Inhalte mit Dienstleistern wie dem BSZ zu archivieren (Bundeskonferenz der Kommunalarchive beim Deutschen Städtetag: „Empfehlung zur Speicherung von kommunalen Webseiten“ http://www.bundeskonferenz-kommunalarchive.de/empfehlungen.html)

Kooperation und Standardisierung

Das BSZ beobachtet die technische Entwicklung, übernimmt Bausteine und nutzt Standards dort, wo sie existieren bzw. beteiligt sich an der Entwicklung. Dazu zählt die Integration von Heritrix mit der Wayback-Maschine und WARC als standardisiertem Archivformat in den Funktionsumfang der Software und Installation.

In Kooperation von nestor, dem Kompetenznetzwerk zur digitalen Langzeitarchivierung, insbesondere mit der Deutschen Nationalbibliothek und der Bayerischen Staatsbibliothek sowie dem Arbeitskreis Webarchivierung der Arbeitsgemeinschaft für Wirtschaftliche Verwaltung veranstaltet das BSZ Workshops, die die Projekte zur Webseitenarchivierung auf Experten- und Interessentenebene vernetzen.

SWBregio

Als Kooperation des Landesarchivs Baden-Württemberg und des Bibliotheksservice-Zentrums Baden-Württemberg bietet die Plattform SWBregio (http://www.swbregio.de) die Möglichkeit, Webauftritte von Kommunen, Städten und den mit ihnen verbundenen Körperschaften zu sammeln, zu erschließen, der Öffentlichkeit dauerhaft zur Verfügung zu stellen und zu archivieren. Eine Beschränkung des Angebots auf Baden-Württemberg ist nicht vorgesehen, Interessenten aus anderen Bundesländern sind willkommen.

Der Betrieb und die technische Sicherung der Inhalte liegen in den Händen des BSZ. Die Grundlage von SWBregio bildet die am BSZ entwickelte Archivierungssoftware SWBcontent.

Für die Nutzung von SWBregio wird ein Entgelt erhoben. Auskunft über die Teilnahme und Einweisung gibt das BSZ.

SWBcontent

Die Auftraggeber wählen die zu archivierenden Webauftritte aus, erschließen und harvesten sie. Das BSZ weist in die Bedienung der Software ein, in der Regel kann dies telefonisch geschehen. Auf Anfrage führt das BSZ die Erschließung und Speicherung im Auftrag des Auftraggebers unter Rückmeldung an ihn aus.

SWBregio nutzt die in Zusammenarbeit mit dem Landesarchiv Baden-Württemberg entwickelte Beschreibungsmaske für die archivische Dokumentation. Die erfassten Merkmale stehen für die Recherche zur Verfügung. Mit Hilfe der Offline-Browser Heritrix bzw. HTTrack werden die Webseiten übernommen.

Ein Alleinstellungsmerkmal der Basis-Software SWBcontent bildet die ausgeprägte Eignung, hierarchische Zusammenhänge darstellen zu können – also Spiegelungen eines Webauftritts zu unterschiedlichen Zeitpunkten zusammen zu halten.

Die automatische Vergabe und Registrierung persistenter Identifier (URN) gehört zum Leistungsumfang ebenso wie die Erhebung technischer Metadaten der Objekte. Die Einbettung in die Geschäftsgänge und die Verbindung mit den Findmitteln wird von den Anwendern als ausgesprochen komfortabel empfunden. Benutzungsbeschränkungen können verwaltet werden.

Das BSZ betreibt die auf der Grundlage der Software SWBcontent aufgebauten Installationen und gewährleistet ihre Entwicklung und ihren Betrieb sowie die Pflege der Daten und Software im Einvernehmen mit den beauftragenden Institutionen.

Weitere Installationen der Software SWBcontent sind in Betrieb – hingewiesen sei auf das Baden-Württembergische Online-Archiv BOA (http://www.boa-bw.de) als ursprüngliche Installation für die Badische Landesbibliothek Karlsruhe, das Landesarchiv Baden-Württemberg und die Württembergische Landesbibliothek Stuttgart oder den Archivserver für das Saarland Saardok (http://saardok.sulb.uni-saarland.de/).

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:**
Datenbank des Südwestdeutschen Bibliotheksverbunds (SWB) zur kooperativen Katalogisierung, Datenlieferung an Bibliothekssysteme, Recherchesysteme für Bibliotheken, Webservices und Fernleihzentrale.
- **Bibliothekssysteme:**
Bibliothekssysteme aDIS/BMS und Koha nach dem Modell Software-as-a-Service.
- **Museen, Archive und Repositorien (MARE):**
Dokumentations- und Präsentationswerkzeuge für Museen und Archive, Langzeitarchivierung und Repositorien.

Postanschrift
BSZ Baden-Württemberg
Universität Konstanz
78457 Konstanz

Besuchsadresse
Felix-Wankel-Straße 4
78467 Konstanz

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

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

Direktor Dr. Ralf Goebel
Stellvertreter Dipl. Math. Volker Conradt

Mai 2014

Ihre Ansprechpartnerin für SWBregio

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

Weitere Informationen finden Sie unter
http://www.swbregio.de

BSZ Bibliotheksservice-Zentrum Baden-Württemberg

Digitales Webseitenarchiv für Städte und Kommunen

SWBregio

Bibliotheksservice-Zentrum Baden-Württemberg

www.bsz-bw.de