arbido

2008/3 Informationseinrichtungen und Sport

Signori Barbara, Leiterin Dienst e-Helvetica

Tags: Archiv Bibliothek, Benutzung, Elektronisch,

Euro 08 für die Ewigkeit

Die Euro 08 ist vorüber, der Champion steht fest. Auf dem Schweizer Web findet Google über drei Millionen Einträge über die Meisterschaft. Die Schweizerische Nationalbibliothek archiviert die wichtigsten dieser Webseiten. Auf alle Zeit.

Heute ist die Stimmung während der Euro noch in frischer Erinnerung, morgen wird sie vergessen sein. In einer Generation aber werden Forschende wieder wissen wollen, wie das damals war während der Fussball-Europameisterschaft; wie die Schweiz damit umgegangen ist, wie sie darauf reagiert hat. Neben Büchern, Zeitungen und Zeitschriften werden die Webseiten eine wichtige Quelle sein. In den Sammlungen der Schweizerischen Nationalbibliothek wird dann nicht nur das zur Verfügung stehen, was heute gedruckt wurde, sondern auch eine Auswahl der interessantesten Webseiten.

Die Auswahl

Im gedruckten Bereich strebt die Schweizerische Nationalbibliothek eine nahezu vollständige und wertungsfreie Sammlung von Schweizer Helvetica an. Dieser Grundsatz gilt im Prinzip auch für die Sammlung von digitalen Helvetica[Mit Helvetica werden Publikationen bezeichnet, die in der Schweiz erscheinen, sich auf die Schweiz oder auf Personen mit schweizerischem Bürgerrecht oder Wohnsitz beziehen oder aber von schweizerischen oder mit der Schweiz verbundenen Autoren oder Autorinnen geschaffen oder mitgestaltet werden.]. Aufgrund der riesigen Datenmengen, die im Internet publiziert werden, drängt sich jedoch eine eingeschränkte Sammelpolitik auf.

Für die Sammlung der Euro-08-Seiten hat eine Informationsspezialistin der Nationalbibliothek eine Auswahl getroffen. Hauptkriterium ist, dass die Webseite die Euro 08 als zentrales Thema darstellt und nicht nur am Rande behandelt. Ziel ist es auch, ein möglichst breites Spektrum an Themen rund um die Euro 08 abzudecken. So enthält die Sammlung nun rund 40 Webseiten, die die Euro 08 aus verschiedenen Blickwinkeln dokumentiert:

- Die offizielle UEFA-Webseite zeigt zum Beispiel die Ergebnisse sowie die Berichterstattungen zu den einzelnen Spielen, Informationen zu den Highlights und Fotografien.

- Die Webseiten der offiziellen Host-Citys Basel, Bern, Genf und Zürich zeigen zum Beispiel die Sehenswürdigkeiten der einzelnen Städte und deren Übernachtungsmöglichkeiten sowie Schnellsprachkurse in Deutsch für alle ausländischen Fans.

- Auch die Webseite der SBB wirbt mit touristischen Attraktionen sowie mit Informationen zum öffentlichen Verkehr während der Euro 08 für alle Interessierten.

- Das Sportmuseum Schweiz stellt das Kulturprojekt «1924: Wir holen uns den Titel zurück» auf seiner Webseite vor.

- Auf der Webseite von ecosport.ch sind Umwelt-Tipps für die Veranstalter rund um die Fanzonen zu finden.

- Diverse Webseiten sind speziell von Fussballfans für Fussballfans erstellt worden und enthalten zum Beispiel Informationen zu Fanshops oder zu Campingmöglichkeiten.

Daneben gibt es aber auch Webseiten für die sogenannten Fussballmuffel. Von diesen Seiten wurde ebenfalls eine Auswahl gesammelt.

- Die Webseite der Post zeigt geeignete Marketingstrategien für einen Grossanlass wie die Euro 08 auf und nennt die wichtigsten Schritte, die ein Unternehmen durchlaufen muss, wenn es auf den EM-Zug aufspringen und von der Euro 08 profitieren möchte.
- Die Webseiten von Fedpol und der Schweizer Armee berichten über die Sicherheitsvorkehrungen während der Euro 08.
- Auch die «Kampagne Euro 08 gegen Frauenhandel» nutzt die Fussballmeisterschaft und informiert auf ihrer Webseite über die Anliegen der Kampagne.
- Daneben gibt es viele News-Webseiten, die täglich die wichtigsten Informationen rund um die Euro 08 zusammentragen und dokumentieren.

Die Verarbeitung

Bevor die ausgewählten Webseiten aus dem Internet heruntergeladen und eingesammelt wurden, wurden alle Webseiten-Betreiberinnen und -Betreiber schriftlich über das Vorhaben der Schweizerischen Nationalbibliothek informiert. Die Rückmeldungen waren durchwegs positiv.

Das Harvesting selber wird vom European Archive (http://www.europ-archive.org) ausgeführt. Das European Archive hat sich auf Webarchivierung spezialisiert und führt für diverse Gedächtnisinstitutionen in Europa Aufträge zu selektivem Harvesting durch.

Ein Harvester – ein Programm, das Dokumente im Internet systematisch einsammelt – lud während der Euro 08 die ausgewählten Webseiten auf einen Server herunter. Dabei werden alle Links auf einer Webseite angewählt, die auf Informationsangebote zeigen, welche auf der Webseite selber verfügbar sind. Die meisten Webseiten wurden nur einmal während der Euro 08 heruntergeladen. Die UEFA-Webseite wurde zweimal wöchentlich eingesammelt. Die Webseiten werden sozusagen zum Zeitpunkt des Herunterladens eingefroren und dokumentieren genau den Stand, an dem der Download stattfindet. Informationen, die nachher angefügt werden, sind nicht enthalten und werden somit auch nicht nachgewiesen. Die Flüchtigkeit der Informationen auf dem Internet ist eine Tatsache, mit der auch Gedächtnisinstitutionen umzugehen lernen müssen.

Ausgehend von den rund 40 ausgewählten Webseiten hat der Harvester an die 52 Millionen Dokumente eingesammelt. Das entspricht einem Datenvolumen von rund 470 Gigabyte und ist vergleichbar mit 100 einschichtig beschriebenen DVDs oder 700 CDs.

Zurzeit findet die Qualitätskontrolle der eingesammelten Webseiten statt. Diese Kontrolle ist mit grossem Aufwand verbunden, da sie manuell und visuell erfolgt. Bis heute gibt es keine Hilfsmittel, die eine Qualitätskontrolle automatisiert durchführen. Fehler, die bisher festgestellt wurden, betreffen zum Beispiel einzelne fehlende Grafiken (externe Logos) oder Funktionen wie Pulldowns, die nicht mehr korrekt ausgeführt werden können.

Sobald die Qualitätskontrolle abgeschlossen ist, werden die Webseiten erschlossen. Dadurch können sie in Helveticat (http://www.nb.admin.ch/helveticat), dem Bibliothekskatalog der Schweizerischen Nationalbibliothek, recherchiert und ausfindig gemacht werden. Für die eindeutige Identifizierung der digitalen Objekte verwendet die Nationalbibliothek die URN (Uniform Resource Name). Durch die URN wird auch der spätere Zugriff auf die archivierten Webseiten in der Nationalbibliothek gewährleistet. Die Webseiten werden alle im digitalen Langzeitarchiv der Schweizerischen Nationalbibliothek gespeichert und bleiben dauerhaft erhalten.

Webarchiv Schweiz

Die Euro-08-Seiten bilden einen Teil des Webarchivs Schweiz, das die Schweizerische Nationalbibliothek gemeinsam mit den Kantonsbibliotheken und weiteren interessierten Institutionen entwickelt. Neben den Internetseiten bedeutender Ereignisse (z. B. die eidgenössischen Wahlen vom Herbst 2007) enthält es die Webauftritte der Kantons- und Bezirkshauptorte sowie Seiten zu besonders wichtigen Schweizer Themen. Ziel von Webarchiv Schweiz ist es, eine breite Sammlung von landeskundlich relevanten Webseiten der Schweiz aufzubauen und deren Langzeiterhaltung und -verfügbarkeit sicherzustellen. Dabei wird das Synergiepotenzial zwischen den landeskundlich ausgerichteten Schweizer Archivbibliotheken optimal genutzt.

e-Helvetica

Die Webseiten dauerhaft zu erhalten und zugänglich zu machen, ist eine Herausforderung, die sich dem Gesamtprojekt e-Helvetica (http://www.nb.admin.ch/e-helve...) der Schweizerischen Nationalbibliothek stellt. Seit 2001 erarbeitet ein Team aus IT-Fachkräften und Bibliotheksspezialistinnen und -spezialisten die Grundlagen für die Sammlung, Erschliessung, Bereitstellung und Langzeiterhaltung elektronischer Publikationen aus der Schweiz oder mit einem starkem Bezug zur Schweiz. Die Sammlung umfasst aktuell neben Webseiten auch Online-Bücher und -Zeitschriften von Schweizer Verlagen, Online-Hochschulschriften der Schweizer Universitäts- und Hochschulbibliotheken sowie amtliche Online-Publikationen der Bundesverwaltung.

Die Plattform für die Archivierung der Inhalte ist bereits in Betrieb. Ein Teil der Verlagspublikationen, Hochschulschriften und Webseiten kann auch elektronisch eingeliefert werden. Zurzeit wird das Modul für die Abfrage der gespeicherten Inhalte entwickelt. Damit sollen sowohl den heutigen als auch den künftigen Benutzern die digitalen Sammlungen der Schweizerischen Nationalbibliothek zur Verfügung gestellt werden.

Barbara Signori

Barbara Signori ist Leiterin des Dienstes e-Helvetica an der Schweizerischen Nationalbibliothek (NB) in Bern. e-Helvetica umfasst die Sammlung original digitaler Inhalte mit Bezug zur Schweiz, darunter z.B. Websites, Wikipedia-Artikel, e-Books, e-Journals, Dissertationen und Normen. In dieser Funktion leitet und koordiniert Barbara Signori auch das Webarchiv Schweiz. Vor ihrer Tätigkeit als Dienstleiterin war sie mehrere Jahre als Projektmanagerin tätig und beteiligte sich an zahlreichen Projekten zum Aufbau der digitalen Sammlungen der NB. Zudem war sie an der technischen Entwicklung der Langzeitarchivierungssysteme beteiligt – sowohl der früheren als auch der heutigen Systeme. Vor ihrem Einstieg in die digitale Langzeiterhaltung im Jahr 2000 arbeitete Barbara Signori mehrere Jahre in der Erschliessung und Informationsvermittlung der NB.

www.nb.admin.ch

www.nb.admin.ch/de/webarchiv-schweiz