

2. DFG-Workshop 3.2. Erschließung/Ordnung/Verzeichnung

Protokoll des Vortrags von Christoph Becker-Schaum

Herr Becker-Schaum stellt, unterstützt von einer Powerpoint-Präsentation, eine Möglichkeit vor, auch passwortgeschützte Webbereiche und Intranets für die Nutzung verfügbar zu machen. Passwortgeschützte Seiten sind in der Regel mit Hilfe des Spiegelungsverfahrens nicht erreichbar, auch wenn die nötigen Passwörter zur Verfügung stehen. Intranetseiten können per se nicht gespiegelt werden. In beiden Fällen scheint deshalb die Übernahme der Daten direkt aus dem CMS-System ohne Alternative. Die Archivierung geschützter und Intranetseiten setzt aus diesem Grund eine enge Zusammenarbeit mit der Organisation voraus, die diese Seiten publiziert.

Die bündnisgrünen Internetauftritte werden aus diversen Content Management Systemen generiert. Eines von diesen ist für die Archivierung besonders bedeutend, weil es bei der Generierung von ca. 75 Internetauftritten eingesetzt wird, insbesondere auch von der Bundestagsfraktion und dem Bundesverband Bündnis 90/Die Grünen, aber auch von der Europafraktion, mehreren Landtagsfraktionen und Landesverbänden sowie Abgeordneten auf Europa-, Bundes- und Landesebene. Die folgenden Aussagen beziehen sich auf die inzwischen getroffenen Vereinbarungen zur Archivierung der Daten aus diesem Content Management System.

Das CMS der Bundestagsfraktion wird nicht nur zur Generierung von Internetseiten verwandt. Es bietet auch die Möglichkeit, Arbeitsabläufe zu steuern, und kann dazu eingesetzt werden, das „Papierlose Büro“ Wirklichkeit werden zu lassen. Das ist ein weiterer Aspekt, der die archivische Befassung mit dem CMS nahe legt. Erst einmal ist an dieser Stelle relevant, dass die Funktion der Steuerung der Arbeitsabläufe mit sich bringt, dass auch die Funktion „Archivierung“ nach miteinander vereinbarten Kriterien eingerichtet werden konnte.

Die Dokumente des CMS durchlaufen, nachdem die Freischaltung für den Internetauftritt beendet ist, drei „Archivstadien“, extern, intern und AGG:

- Extern: die Dokumente sind über die Suchfunktion im Internetangebot frei zugänglich,
- Intern: die Dokumente sind nur noch intern für die TeilnehmerInnen im CMS zugänglich,
- AGG: die Dokumente sind allein für die MitarbeiterInnen des Archiv Grünes Gedächtnis zu finden.

Beim Übergang zur Archivstufe intern werden sämtliche Links entfernt.

Die Archivfunktion wird relevant, nachdem die Dokumente aus der Präsentation im Internetauftritt herausgenommen sind.

Jetzt besteht für die ArchivmitarbeiterInnen die Möglichkeit, den Metadatensatz durch die archivspezifische Verschlagwortung und eine Körperschaftsindexierung zu erweitern. Im Rahmen der Archivfunktion wird das Dokument des MitarbeiterInnen des Archivs in der Form einer PDF-Datei angezeigt. Auch die – zu einem noch zu bestimmenden Zeitpunkt – erfolgende Archivierung (= Übernahme/Transfer) erfolgt

als PDF-Datei. Diese Vereinbarung ermöglicht, dass die einzelnen Inhalte ohne das Computerprogramm des CMS übernommen werden können.

Der im CMS vorhandene Metadatensatz, wer war die/der VerfasserIn, Datum der Erstellung, Beginn und Ende Freigabe für das Internet, Verschlagwortung, Volltextrecherche etc. wird auf Grundlage der Absprachen mit dem Archiv um vier Datensätze erweitert: wie schon erwähnt werden Felder eingerichtet für einen Körperschaftsindex und für den Schlagwortkatalog des Archivs. Außerdem werden die Rubrik und die Objektnummer automatisch erfasst.

Der Metadatensatz ist neben der Textrecherche die Grundlage für alle Recherchemöglichkeiten. Dadurch ergibt sich: In der so genannten „Dokument-Suche“ kann nach „klassischen“ Metadaten wie Titel, AutorIn, Erstellungsdatum usw. recherchiert werden. In der Archiv-Stufe ist der Metadatensatz um den Schlagwortkatalog und den Körperschaftsindex des Archiv Grünes Gedächtnis ergänzt. Das Feld Rubrik gibt an, wo das Dokument im Internetauftritt zuletzt zu finden war, z.B. die verschiedenen politischen Themenbereiche oder Pressemitteilungen, und die Objektnummer ist die unveränderliche Nummer, die jedes Dokument, Foto o.ä. im CMS automatisch erhält und die im Internet in der Browserzeile zu sehen ist. Diese Recherchemöglichkeit macht die Suche nach Dokumenten, deren damalige URL bekannt ist, äußerst komfortabel.

Diskussion

Da für den Großteil der Anwesenden diese Art des Zugriffs neu ist, gibt es verschiedene Verständnisfragen und erneut Fragen zur Sicherung und Pflege langfristig eventuell veralteter Dateiformate.

Was ist mit dem Zugriff auf ältere pdf-Dateien?

Wie sind die Erfahrungen, bzw. ist geplant, das neue Format pdf-A zu nutzen?

Herr Becker-Schaum weist darauf hin, dass es im gesamten Projekt-Bereich bisher wenig Erfahrungen mit der Klärung dieser Fragestellung gibt, ist aber optimistisch, dass es hierfür Lösungswege geben wird.