

4. Tagung des VdW-Arbeitskreises "Elektronische Archivierung" bei der Bertelsmann AG, Gütersloh, am 21. und 22. April 2005

"Von Standards, Metadaten und Formaten: Theorie und Praxis bei der Umsetzung digitaler Archivierungsstrategien"

Tagungsbericht

In den Räumen der Bertelsmann AG Gütersloh fanden sich am 21. und 22. April 2005 21 Teilnehmer und Teilnehmerinnen des Arbeitskreises "Elektronische Archivierung" ein, um Strategien und Lösungsansätze für die Archivierung elektronischer Unterlagen zu diskutieren. Ausgehend von den Ergebnissen der vergangenen Tagung standen insbesondere Fragen zu Standards, Formaten und Metadaten auf dem Programm.

Zu Beginn der Tagung wurden die Erwartungen der Teilnehmenden ermittelt. Dabei stellte sich heraus, dass vor allem Hinweise zum Thema Standards und konkrete Empfehlungen, sowie eine Vertiefung der Themen Formate und Metadaten gewünscht waren. Welche Formate eignen sich als Archivierungsformat? Welche unverzichtbaren Metadaten gibt es? Als Diskussionsgrundlage und Ergänzung dienten einige Kurzbeiträge von Referenten aus den Reihen des Arbeitskreises.

Zusammenfassung der Referate

Die folgende Zusammenfassung der Kurzbeiträge aus dem Kreis der Teilnehmenden zeigt einige Themen und Schwierigkeiten, mit denen sich Archive im Rahmen der digitalen Archivierung auseinandersetzen haben.

Antje Scheidung von der Bertelsmann AG eröffnete mit Ihrem Referat über die "Dateiformate für elektronische Aufzeichnungen in Archiven - Praxis und Standards" die Vortragsreihe. Sie stellte zunächst das Ergebnis Ihrer Umfrage bezüglich Speicherformaten in Archiven vor. Mehr als 50% Prozent der befragten Archive setzen sich mit dem Thema Archivierung elektronischer Aufzeichnungen auseinander, allerdings besitzt nur ein geringerer Teil der Archive ein Konzept oder eine Richtlinie für die Abgabe und Übernahme von elektronischem Schriftgut. Für die Dokumentenkategorien Daten/Datenbanken, strukturierter Text, Präsentationen, Bild und Officedokumente gab es übereinstimmende Aussagen bezogen auf die verwendeten Speicherformate, während für Audio, Video, Design, geographische Daten und Interaktionen in der Regel keine übereinstimmende Aussage getroffen wurde. In ihrem Beitrag behandelte Antje Scheidung anschließend die unterschiedlichen Standardformate für die verschiedenen Dokumentenkategorien und leitete daraus Hinweise für eine Empfehlung an die Archive ab.

Wolfgang Richter vom Archiv der Dresdner Bank AG in Frankfurt /M. stellte in seinem Beitrag "Qualität des elektronischen Archivguts - Was wollen / können wir unseren Kunden bieten" die Authentizität als Qualitätsmerkmal von elektronischem Archivgut in den Mittelpunkt. Er zeigte auf, dass die Nachweismöglichkeit für die Authentizität grundsätzlich anders ist als in der "Papierwelt". Durch eine sichere Verwahrung und durch die Verwendung einer digitalen Signatur als technisches Mittel zur Wahrung der Authentizität bei elektronischem Archivgut, sieht Wolfgang Richter Archive in der Lage, ihren Benutzern Unterlagen zu liefern, die hinsichtlich ihrer Authentizität der gewünschten Qualität entsprechen.

Anschließend berichtete Frau Ute Schiedermeier vom Siemens-Archiv in München von der Einführung einer neuen Archivsoftware. In Ihrem Beitrag "Moderne Archivsystemsoftware – Anforderungskriterien und Bewertungshinweise" stellte sie die Vorgehensweise am Beispiel ihres Auswahlprozesses vor. Ein Fragenkatalog, in dem alle Anforderungen an das neue Archivsystem gesammelt wurden, bildete dabei die Grundlage für die Ausschreibung. Neben Funktionen können auch Informationen über den Anbieter (Firmengröße, Umsatz, Sitz, Referenzen) von Interesse sein. Um

brauchbare und vergleichbare Informationen von den einzelnen Anbietern zu bekommen, ist es empfehlenswert, Vorgaben zur Beantwortung der Fragen zu machen. Abschließend erläuterte Ute Schiedermeier anhand eines Beispiels, wie man letztendlich die zurückgesandten Ausschreibungsunterlagen mit Hilfe eines Punktesystems bewertet, um das geeignete Produkt zu finden.

Den abschließenden Beitrag des ersten Tages lieferte Dr. Stefan Przigoda vom Deutschen Bergbau-Museum in Bochum. Er referierte über das Austauschformat des DFG-Projekts "Entwicklung von Werkzeugen zur Retrokonversion archivischer Findmittel" und die Langzeitspeicherung archivischer Metadaten. Ausgangspunkt für dieses DFG-Projekt war die forschungs- und archivpolitisch gewünschte Ausweitung des Internetangebots der Archive und die Einsicht, dass dies eine beschleunigte Retrokonversion bzw. Digitalisierung der überwiegend analog – zumeist in Form von Findbüchern – vorliegenden Erschließungsinformationen voraussetzt. Deshalb wurde eine Software entwickelt, mit deren Hilfe eine strukturierte, verlustfreie Übertragung von Erschließungsinformationen in ein digitales Recherchesystem möglich ist. Ein Standardformat für den archivübergreifenden Datenaustausch ist zur Nutzung der Software Voraussetzung. Das auf Basis von XML entwickelte Format ermöglicht ein Höchstmaß an Flexibilität, indem nur Kernstrukturen verbindlich festgelegt wurden, darüber hinaus aber vielfältige Differenzierungen und Spezifizierungen optional möglich sind. Eine Voraussetzung dafür, dass dieses Format auch über das Projekt hinaus Akzeptanz und Verbreitung findet, sieht Stefan Przigoda nicht nur in der Offenlegung der Dokumentation, sondern auch in der technischen Unterstützung des Formats. Weiterführende Informationen sind zu finden unter: <http://www.archive.nrw.de>.

Am 2. Tag der Tagung war Dr. Hans Jürgen Sproß vom Archiv der DaimlerChrysler AG in Stuttgart "Auf der Suche nach den Metadaten". Hans Jürgen Sproß stellte verschiedene Standards wie z.B. IPTC (für Bilder), Dublin Core (für Bilder und Schriftgut) und den Bayerischen Metadatenkatalog (für Schriftgut) einander gegenüber. Wichtige Aspekte im Hinblick auf Metadaten sind die Art ihrer Erfassung (automatisch oder manuell), die Verwendung der Metadaten und die Frage der Strukturierung. Die Tagungsteilnehmer waren übereinstimmend der Ansicht, dass man bei der Festlegung von Metadaten die Balance zwischen ausreichender Information aus archivarischer Sicht und Akzeptanz bei den Erfassenden finden muss.

Diskussion und Ergebnis

In der Diskussion einigten die Teilnehmenden sich darauf, dass zunächst die Anforderungen an Formate für die Archivierung von Text-Ausgangsformaten festgelegt werden sollten. Folgende unverzichtbare Anforderungen wurden ermittelt.

- lesbar (durch Maschinen, durch Menschen)
- verarbeitbar
- rechtsverbindlich
- nicht proprietär
- vollständig
- offen dokumentiert
- weit verbreitet
- zugänglich / migrierbar
- stabil / reif
- Layouterhaltung
- verknüpfbar
- explizite Struktur

Einige Teilnehmer wiesen sehr zu Recht auf die Bedeutung der Revisionsicherheit hin. Diese ist allerdings weniger eine Frage des Formats, als vielmehr der Datenhaltung und erscheint darum nicht in der Liste der Kriterien.

Ausblick und Themen für nächste Sitzung

Es hat sich herausgestellt, dass die Arbeit an den Themenkomplexen Metadaten und Formate auch in der nächsten Sitzung zentraler Themenschwerpunkt sein und weiter vertieft werden muss.

Anhand einer Matrix sollen abhängig von den Archivaliengattungen, die vom Arbeitskreis aufgestellten Anforderungen an ein Archivformat und die unterschiedlichen Speicherformaten einander gegenübergestellt werden.

Für jeden Archivalientyp wird eine Liste von Metadaten erstellt, die auf den gängigen Standards (siehe Vortrag Hr. Dr. Sproß) und einer Umfrage der am Arbeitskreis beteiligten Archive basieren.

Die nächste Tagung findet am 10. und 11. November 2005 bei der Daimler/Chrysler AG in Stuttgart statt.

Elke Schunk