

Herausforderung angenommen – zehn Jahre elektronische Archivierung im Siemens-Archiv

Ute Schiedermeier

Das Siemens-Archiv in München arbeitet seit 1969 mit Archiv-Informationssystemen. Zunächst wurde eine Schlitzlochkartei benutzt, die mit dem Siemens-Produkt GOLEM verbunden war. Dieses Lochkartensystem wurde Anfang der 1980er Jahre durch eine Rechnerversion ersetzt. 1993 wurde das neue, speziell auf das Siemens-Archiv zugeschnittene System InfoSim der Firma Sietec in Betrieb genommen.

Im Gegensatz zu GOLEM sollte InfoSim in der Lage sein, digitale Informationen verschiedenster Art zu verarbeiten. Diese waren zusätzlich zu den bereits mit GOLEM erfassten Attributen auch umfangreichere Volltexte, Abbildungen und Töne, die in elektronischer Form in den Archivbestand übernommen werden sollten. Bei den konzeptionellen Überlegungen und der Auswahl des Systems standen neben anderen Aspekten die Vermeidung von Medienbrüchen und die Gestaltung effizienter Arbeitsprozesse im Vordergrund. Was bereits in elektronischer Form vorlag, sollte ohne Nachbearbeitungsaufwand in das System übernommen, dort archiviert und bei Bedarf an die Kunden weitergegeben werden können. Ob und wie es uns gelungen ist, all diese Anforderungen zu realisieren, möchte ich im Folgenden ausführen.

Migration

Am Beginn der Arbeit mit dem neuen Archiv-Informationssystem standen die aus GOLEM migrierten 56 000 Datensätze sofort zur Verfügung. Hierbei handelte es sich ausschließlich um Attributinformationen von Schriftgut, das seit 1969 aus allen Bereichen in das Archiv eingepflegt wurde. Während des laufenden Betriebs wurden von Archivmitarbeitern die 56 000 Datensätze innerhalb von drei Wochen mittels Makros auf die neu entstandenen Themenarchive verteilt. Der zeitintensive Kopiervorgang fand nachts statt, dadurch standen tagsüber die Datensätze für Benutzer bereit. Morgens wurden Kontrollen durchgeführt, die fehlerhafte Übertragungen erkannten.

Retro-Digitalisierung – die ersten Schritte in eine neue Welt

Als Pilotprojekt wurden zunächst die wichtigsten Rundschreiben und Organisationsunterlagen von Siemens aus den Jahren 1945 bis 1993 digitalisiert. Die 150 000 gescannten Seiten durchliefen eine OCR-Erkennung und wurden zusätzlich im TIFF-Format im System abgelegt sowie mit Attributen versehen. Entsprechend wurde mit ca. 24 000 Volltextseiten aus den Provenienzen Aufsichtsrat und Vorstand verfahren. Erste Erfahrungen mit digitalen Bilddateien sammelten wir mit den 46 000 Fotos des sogenannten Handarchivs. Hierbei handelt es sich um vielbenutzte S/w-Fotos und Farbbilder jeglicher Größe. Diese wurden grob verzeichnet und in "Vorschauqualität" (50 KB) als JPEG gespeichert. Das Startvolumen nach Inbetriebnahme von InfoSim im Jahr 1993 war demnach:

150 000 Seiten Volltext-Rundschreiben, Organisations-Unterlagen

24 000 Seiten Volltext-Unterlagen der Firmenleitung

46 000 Fotos in geringer Qualität

56 000 migrierte Datensätze aus Vorgängersystem

Die erwähnten 24 000 Seiten Aufsichtsrats- und Vorstandsunterlagen leisteten einen wichtigen Beitrag zur Finanzierung des Gesamtvorhabens, das aus Restrukturierungsgeldern bestritten wurde. Geleitet vom Gedanken eines "vorbeugenden Krisenmanagements" sollte das Archiv in der Lage sein, der Unternehmensleitung zu aktuellen Fragestellungen die jeweils relevanten, Rechtssicherheit herstellenden

Dokumente – idealerweise in Kombination mit Zusatzinformationen – in kürzester Zeit zu liefern. Die Praxis hat gezeigt, dass wir diesen Anspruch vollständig befriedigen können.

Diesen positiven Erfahrungen folgten weitere Retro-Digitalisierungs-Projekte, sodass unser Datenbestand innerhalb von zehn Jahren auf 240 000 Datensätze anwuchs. Der Umfang der gescannten Volltexte kann von einer Seite bis zu einigen hundert Seiten betragen. Die Größe der digitalen Objekte beeinflusst jedoch die Antwortzeiten beachtlich, sodass wir Objekte mit mehr als 100 Seiten auf mehrere Datensätze verteilen. Insgesamt sind wir mit den Antwortzeiten sehr zufrieden, was auf die Systemstruktur und die eingesetzte Software zurückzuführen ist. Wir arbeiten mit einem Unix-Server, Windows-Clients, Speicherung auf Juke-Box und der Datenbank FULCROM.

Die Digitalisierung und OCR-Erfassung sowie die zugehörige Attributvergabe von Standarddrucksachen und Fotoalben erfolgten durch die hauseigene Kopierstelle. Unter Einsatz von einer eigens für diese Stelle entwickelten InfoSim-Komponente konnten bis zum Jahresende 2003 nachstehende Bestände digitalisiert und als TIFF- bzw. JPEG-Dateien in das System eingepflegt werden:

Aktuelle Rundschreiben

Aktuelle Vorstands- und Aufsichtsratsunterlagen

Bücher

Chroniken

Zeitschriften

Fotoalben

Publikationen

Presseinformationen

Die Daten wurden mittels Dat-Kassette nachts über eine Software-Schnittstelle an die Archiv-Datenbank übergeben. Die Auswahl der zu digitalisierenden Dokumente erfolgte nach festgelegten Kriterien, die heute noch unverändert Gültigkeit haben:

Inhaltliche Wichtigkeit (z. B. Unterlagen der Firmenleitung; häufig nachgefragte Daten)

Notwendigkeit (Massendaten, die nicht einzeln erfasst werden können und ohne OCR nicht komplett erschließbar sind, wie z. B. Rundschreiben oder Zeitschriften)

vertretbares Kosten/Nutzen-Verhältnis

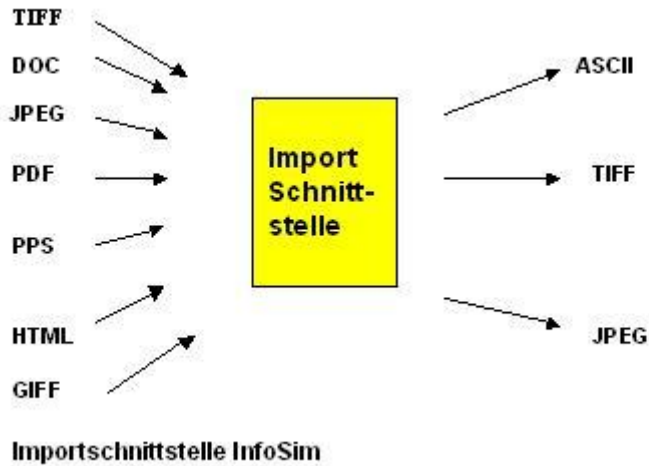
Möglichkeit einer Volltextgenerierung

Übernahme digitaler Daten – Arbeitsabläufe

Parallel zu den Digitalisierungsprojekten, die von uns in Auftrag gegeben wurden, wuchs der Bestand an Dokumenten, die zusätzlich zur gedruckten auch in elektronischer Form in das Archiv gelangten. Anfänglich hatten wir drei Standardformate für die Übernahme und Archivierung definiert: das ASCII-Format, das TIFF-Format und das komprimierte JPEG-Format. Über eine Import-Funktion konnten die Dateien mit den entsprechenden Attribut-Datensätzen verknüpft werden.

Ab Mitte der 1990er Jahre sahen wir uns zunehmend mit der Ankündigung konfrontiert, dass Publikationen demnächst ausschließlich in elektronischer Form erscheinen würden

("Das erscheint demnächst nur noch elektronisch!"). Die aktenabgebenden Stellen produzierten in einer Übergangsphase gedruckt und digital und dann nur noch digital. Um diesen erweiterten Anforderungen gerecht zu werden, programmierte die Firma Sietec eine komfortable Eingabe-Schnittstelle für das Archivsystem InfoSim. Diese Schnittstelle ermöglichte es, dass die Dateien – in den gängigsten Formaten angeliefert – automatisch in die drei genannten Standardformate konvertiert wurden. Bei Texten wird eine Volltextgenerierung durchgeführt.



Über diese Schnittstelle ist es möglich, digitale Informationen in einem Arbeitsgang konventionell zu verzeichnen, zu konvertieren und zu speichern.

Status quo und Perspektiven für die Zukunft

Heute gehen konventionelles Verzeichnen und elektronische Archivierung Hand in Hand. Elektronische Dokumente werden genauso verzeichnet wie Papier-Dokumente. Ein Bestand kann folglich aus gedruckten Texten, elektronischen Informationen, Bildern, Büchern, Filmen, Exponaten und anderen Objekten bestehen. Räumlich getrennt gelagerte oder zeitversetzt akquirierte Unterlagen und Objekte bilden durch Vergabe von Attributen und Volltextinformation wieder eine Einheit.

Im Laufe der vergangenen zehn Jahre haben wir an vielen Stellen Neuland betreten. Die Anforderungen an Archive generell haben sich durch die technischen Entwicklungen grundlegend verändert. Ich bin sicher, dass diese Entwicklung auch in den kommenden Jahren voranschreiten wird und zahlreiche neue Aufgaben mit sich bringt. Heute bereits absehbar ist die zuverlässige Langzeitarchivierung von:

Web-Seiten

CD-ROMs oder andere Medien

kompletten Datenbanken

Was diese Themen angeht, befinden wir uns im Experimentierstadium. Erste Erfahrungen bestätigen uns jedoch, dass wir bei der Langzeitarchivierung digitaler Dokumente auf dem richtigen Weg sind; auch wenn das Ziel noch nicht zum Greifen nah ist. Die Situation zu Beginn des Jahres 2004 ist dadurch gekennzeichnet, dass

einheitliche Dateiformate bei Siemens durch CIO-Rundschreiben definiert wurden

unsere Standardformate um das PDF-Format erweitert wurden

der Umstieg auf ein Standard-Archivinformationssystem vorbereitet wird

Absprachen mit Zentralstellen im Bezug auf Aktenabgaben getroffen wurden

Aktenzugänge der letzten drei Jahre auf zukünftige Abgabeformate untersucht wurden

Arbeitsprozesse im Archiv den neuen Anforderungen angepasst wurden, indem Mitarbeiter neben herkömmlicher Verzeichnungsarbeit auch digitale Daten erzeugen bzw. übernehmen und erschließen

Mit steigender Anzahl nimmt auch die Verantwortung für die Langzeitarchivierung ausschließlich elektronisch vorhandener Objekte zu. Bei den grundsätzlichen Überlegungen müssen neben den Investitionskosten für leistungsfähigere Systeme auch Folgekosten für das Kopieren und Migrieren von Daten aus Sicherheits- und Sicherungsgründen bedacht werden. Auch die aktive Teilnahme am Arbeitskreis "Elektronische Archivierung" sehe ich als einen Beitrag zur Auseinandersetzung mit den neuen Anforderungen an das Siemens-Archiv. Das Wissen, die Ideen und vor allem die Erfahrungen der Kolleginnen und Kollegen sollen gebündelt und nachnutzbar gemacht werden. Nur so lassen sich unter Umständen kostspielige Fehler sowie Überlieferungslücken vermeiden und die eigenen Arbeitsabläufe so effektiv und effizient wie möglich gestalten.

Anschrift: Ute Schiedermeier, Siemens-Archiv, Oskar-von-Miller-Ring 20, 80333 München; E-Mail: ute.schiedermeier@siemens.com