

Schöne neue Datenwelt

Vom Nutzen und Schaden sogenannter Archivierungssysteme

Von FRIEDER KUHN

Der Beitrag erschien erstmals in: Bestandserhaltung. Herausforderung und Chancen, hrsg. von Hartmut Weber (Veröff. der Staatlichen Archivverwaltung Baden-Württemberg Bd. 47), Stuttgart 1997, S. 355-360.

Es vergeht wohl kaum ein Tag, an dem nicht jeder von uns auf irgendeine Form von Werbung stößt, die den Beruf des Archivars überflüssig zu machen verspricht: Ein Klick mit der Computer-Maus, ein virtueller Aktendeckel öffnet oder schließt sich auf dem Bildschirm - und fertig ist das digitale *Archiv*. Man möchte kalauern: Je bunter das Bild, desto archivierter die Datei.

Was aber hat dies mit den Archiven, mit den Archivarinnen und Archivaren zu tun? Es sind hierzu drei grundsätzliche Fragen zu stellen:

Erstens: Digitalisierung, Dokumentenverwaltung, Archivierungssystem - was versteht man darunter?

Zweitens: Welche Probleme ergeben sich für die archivischen Kernaufgaben der Überlieferungsbildung und der Bestandserhaltung?

Drittens: Wie reagieren die Archive auf diese neuen Herausforderungen?

Im folgenden sollen diese Grundfragen wenigstens näherungsweise beantwortet werden. Allerdings - dies sei gleich zu Beginn eingeräumt - stellen sich bei der Beantwortung dieser Fragen weitere Probleme, die mit Fragezeichen zu versehen sind.

Zunächst also: Was ist oder was will ein sogenanntes *Archivierungssystem*? Der im EDV-Bereich inflationär gebrauchte Begriff *System* beschreibt hier eine Kombination aus Geräteausstattung - Hardware - und Programm - Software - zur bildlichen Speicherung und Wiederauffindung von Schriftgut. Ein Abbild der Vorlage soll auf dem Bildschirm erscheinen und auf einem Drucker wiederum auf Papier zu reproduzieren sein. Hierfür wird die Vorlage zunächst *digitalisiert*: Ein Abtastgerät - Scanner - zerlegt das Bild in Bildpunkte - Pixel - und vergibt im einfachsten Fall des bitonalen Digitalisierens für jeden dieser Bildpunkte einen binären Wert - 0 oder 1 -, der anschließend magnetisch oder optoelektronisch abgespeichert wird. Wenige größere Punkte erscheinen als holzschnittartiges Schwarzweißbild, je kleiner die Bildpunkte und je differenzierter ihre Verteilung, um so eher erscheinen sie dem menschlichen Auge als Grauton.

Um auf das einzelne abgelichtete Blatt wieder zugreifen zu können, müssen entweder Informationen über dessen Inhalt manuell eingegeben oder automatisch generiert werden, indem einzelne Wörter oder Begriffe aus dem Text herausgelesen und in einer Indexdatei abgelegt werden, so daß durch Eingabe eines Suchbegriffs das gewünschte Dokument gefunden wird. Dieses Verfahren ermöglicht wohlgemerkt nur Zugriffe auf einzelne Dokumente; sollen Akten oder Aktengruppen formiert werden, müssen weitere Zusatzinformationen eingegeben werden. So viel - in groben Zügen - zum Verfahren.

Die EDV-Branche unterscheidet gern zwischen *Archiv-Systemen* mit dem Schwerpunkt der Mengenspeicherung und *Retrieval-Systemen* mit dem Schwerpunkt auf raschen Zugriff. Was die Archivarin oder den Archivar am unangenehmsten berührt, ist die unqualifizierte, ja irreführende Verwendung des Begriffs *Archiv*, die mit dem fachsprachlichen Archivbegriff nichts zu tun hat. Schließlich handelt es sich bei allen diesen Programmen bestenfalls um elektronische Ablagen für Unterlagen, die zur Aufgabenerfüllung noch benötigt werden und die auch nur so lange interessant sind.

Was also leisten diese Systeme tatsächlich?

Beim Einsatz technischer Apparaturen hat das Wort *Leistungsfähigkeit* stets quantitative und qualitative Aspekte:

Folgende Faktoren bestimmen im wesentlichen die technische Leistungsfähigkeit von Computern: Der Prozessortyp, dessen Taktrate, die Größe des Hauptspeichers, die Zugriffsgeschwindigkeit des Massenspeichers, die Speicherverwaltung des Betriebssystems. Dank ständiger Verbesserung aller dieser Komponenten sind die quantitativen Fähigkeiten handelsüblicher Personalcomputer sprunghaft gestiegen, und zwar auf ein Vielfaches der ursprünglichen Rechenleistung. Erst diese Leistungszuwächse ermöglichen überhaupt eine bildliche Speicherung schriftlicher Vorlagen. Eine Schreibmaschinenseite Text braucht zirka zwei KB Speicherplatz als ASCII-Zeichen, 500 KB bei einer Auflösung von 200 dpi (Bildpunkt pro Zoll), und zwei MB bei einer Auflösung von 400 dpi, die vor allem zur Darstellung von Grautönen und graphischen Elementen erforderlich werden. Die Qualität der Darstellung ist im Grunde eine Funktion der Rechenleistung, mit anderen Worten: eine Geldfrage, denn grundsätzlich kann man jede beliebige Arbeitsgeschwindigkeit und jedes Speichervolumen kaufen - falls dies insgesamt wirtschaftlich erscheint.

Die Frage nach der Qualität ist damit bereits fast beantwortet: Bei bildlicher Speicherung kann ein vorhandener Speicherplatz X entweder für eine höhere Auflösung - mehr Bildpunkte pro Fläche - oder für eine höhere Speicherdichte - mehr Bilder in geringerer Auflösung - genutzt werden. Bei farbigen Abbildungen ist die Anzahl der darstellbaren Farben, bei Tondateien der Frequenzumfang bei der Kapazitätsberechnung einzuplanen. Diese Abwägungen beeinflussen wesentlich die Wirtschaftlichkeitsrechnung solcher Systeme: Je größer die Anzahl der gespeicherten Seiten auf einem Datenträger vorgegebener Größe, um so wirtschaftlicher die Beschaffungsmaßnahme. Hierbei erfreuen sich Verfahren der Datenkompression oder -reduktion außerordentlicher Beliebtheit, da sie den Speicherbedarf etwa um den Faktor 4 -15 reduzieren. Es handelt sich hierbei um Rechenoperationen, die nur noch Informationen über die abzubildenden Vorlagen enthalten, nicht mehr deren vollständige Datenmenge. Eine Konsequenz ist klar: Nur wer im Besitz der richtigen Systemkomponenten ist, kann das gewünschte Bild erhalten.

Die Konsequenzen für Überlieferungsbildung und Bestandserhaltung

Der Versuch, mögliche Anwendungsgebiete der Digitalisierung vollständig aufzuzählen, soll an dieser Stelle unterbleiben. Ganz allgemein empfehlen die Anbieter solcher Systeme ihre Produkte für alle Anwendungsbereiche, in denen große Datenmengen anfallen: Von Bibliothekskatalogen über audiovisuelle Dokumente bis zu Geschäftsunterlagen, auf die besonders rasch zurückgegriffen werden soll. Gerade die Digitalisierung schriftlicher Unterlagen erscheint als Annehmlichkeit in einer Arbeitsumgebung, in der aus Wirtschaftlichkeitsgründen keine Arbeitszeit auf die Strukturierung der Unterlagen, gar auf Einhaltung eines Aktenplans oder ähnliches *verschwendet* werden soll. Als wichtiger Gesichtspunkt für die Wirtschaftlichkeit des gesamten EDV-Einsatzes gilt vielen Anwendern die Häufigkeit des Rückgriffs auf ein abgelegtes Dokument.

In die Begründungen für die angebliche Notwendigkeit solcher Systeme für den raschen und reibungslosen Betriebsablauf fließen harte betriebswirtschaftliche Daten ebenso ein wie unbestimmte gesamtgesellschaftliche Veränderungen, die mit dem Einsatz der neuen Medien einhergehen.

Zu diesen Phänomenen gehört unzweifelhaft auch die Attraktivität sogenannter *Oberflächen*, das sind Bildschirmgestaltungen zum systemgerechten Bedienen der Programme ohne unmittelbare Funktion in der Anwendung selbst. Angesichts des Markterfolgs der digitalen Schriftgutspeicherung mag der Verdacht aufkommen, daß manche Entscheidungsträger entweder selbst mit Problemen der ordnungsgemäßen

Aktenführung zu kämpfen haben oder ihren Mitarbeitern in diesem Punkt nicht viel zutrauen; jedenfalls erscheinen diese von der EDV-Branche stets als *Lösung* angepriesenen Produkte als attraktiver Ausweg aus der ordnungsgemäßen Akten- und Registraturführung. Insbesondere mit dem Argument des angeblich unkomplizierten Zugriffs auf die Speicherinhalte werden täglich Unternehmen der Wirtschaft von den Vorteilen dieser Verfahren überzeugt - und zunehmend auch Entscheidungsträger in öffentlichen Verwaltungen.

Wie oft aber wird im Vorfeld von Beschaffungsentscheidungen die Frage nach der Lebensdauer des Systems gestellt, das heißt für welchen Zeitraum die neuen digitalen Dokumente in dem zu beschaffenden System aufzubewahren sind - und wie oft wird hierüber mit den zuständigen Archiven gesprochen? Private und öffentliche Archive, aber auch Archivbibliotheken als Verwahrstellen von Nachweisexemplaren, stehen immer dann vor existentiellen Herausforderungen, wenn nur noch digitale Originale oder Exemplare zu bewerten und zu übernehmen sind, weil keine realen mehr produziert wurden. Wir müssen dann mit dem Paradoxon leben, daß das Original - falls dieser Begriff noch verwendet werden soll - virtuell, sein Abbild hingegen real ist. Bereits vor der Übernahme stellt sich somit die Frage nach der künftigen Aufbewahrungsform, aber weniger nach der physischen Erhaltung eines bestimmten Trägers, sondern nach der Erhaltung der Information, unabhängig von dem einst übernommenen Träger.

Magnetspeicher gelten zu Recht als flüchtig: Ändert sich die Magnetisierung der Metallpartikel, ist der Inhalt zerstört oder nicht mehr reproduzierbar. Als üblicher Weg, magnetisch gespeicherte Informationen zu erhalten, gilt daher die Umkopierung auf neue Träger. Bei analogen Signalen wie Ton- oder Videobändern war dies traditionell mit Informationsverlusten verbunden; die digitale Speicherung erfolgt hingegen ganz oder gar nicht: Ist eine digitale Kopie vollständig, so ist sie mit dem Original identisch. Der durch viele Debatten geisternde *ewige Datensatz* bedeutet also nichts anderes als ständiges Umkopieren auf neue, jeweils zeitgemäße Speichermedien - *continuous file conversion* -. Dies setzt aber - siehe oben - voraus, daß die entsprechenden Hard- und Software-Komponenten vorhanden sind. Was aber, wenn die hierfür erforderlichen Systeme, das heißt die passende Hardware und das Betriebssystem, für das die Anwendungssoftware konzipiert wurde, eines Tages nicht mehr zur Verfügung stehen, weil keine Ersatzteile mehr geliefert werden und die nächste Gerätegeneration das alte Betriebssystem nicht versteht? Wie ist dann die Kernaufgabe der Archive, das verwahrte Archivgut zu erhalten und einer Nutzung zugänglich zu machen - *preservation and access* -, noch zu bewerkstelligen?

Seriöse Vorhersagen über den dauerhaften Markterfolg oder die technische Lebenserwartung von magnetischen oder optischen Platten heutiger Bauart oder der zum Lesen und Beschreiben notwendigen Software sind nicht möglich. Dies gilt übrigens auch für populäre, inzwischen massenhaft verbreitete Speichermedien wie Audio-CD und CD-ROM, die ebenfalls keinen Ewigkeitswert haben. Es geht wohl gemerkt nicht nur um die physische Widerstandsfähigkeit der silbrigen Scheiben: Ein Blick in die Bedienungs- und Pflegeanleitung eines CD-Spielers wird jeden von deren Empfindlichkeit überzeugen. Es geht um den fraglichen Lebenszyklus des Systems: In wenigen Jahren wird der Aufwand für Produktion und Vertrieb solch einzelner Scheiben als nicht mehr gerechtfertigt erscheinen, weil die Musikproduktionen der Phonoindustrie online auf der sagenumwobenen *Datenautobahn* angeboten werden: Im Zeitalter von Multimedia werden die Hörer eine Datei namens *Beethoven/5.Symphonie/ Karajan/Deutsche Grammophon 1979* in ihr Heim-Terminal laden und über Kopfhörer oder Lautsprecher wie gewohnt genießen. Entsprechendes gilt für Computer-Programme, die aus dem Netz heruntergeladen werden. Wozu also noch diese - ökologisch bedenklichen - Glitzerscheiben? Wenn aber diese enormen Stückzahlen fehlen, wie viele CD-Fabriken werden überleben, um Speicherplatten für Archivierungssysteme herzustellen? Und wie viele Hardware-Anbieter werden noch Abspielgeräte für aussterbende Scheiben reparieren?

Deren gespeicherte Informationen können nur bewahrt werden, indem sie auf künftige neue Systeme übertragen werden, auch auf solche, die wir heute noch nicht kennen. Diese Übertragung wird mit dem Begriff *Migration* umschrieben. Wenn nun Unterlagen keine physische Gestalt mehr haben, die es zu erhalten gilt, sondern nur noch als Daten vorliegen, die - hoffentlich - in spätere Systeme hinüberzuretten sein werden, wird das zentrale Kriterium für die Auswahl eines Speichersystems dessen Migrationsfähigkeit. Spätestens bei der übernächsten oder drittnächsten Generation endet jedoch jede Prognose: sie wäre, wie jede Aussage über künftige technische Entwicklungen, sehr fragwürdig. Kein Zweifel aber besteht daran, daß analog dem Zuwachs des Speicherinhalts zwangsläufig ein größer werdender Teil der Kapazität für die Übertragung auf die nächste Systemgeneration reserviert werden muß - Kapazität, die, da nur einmal vorhanden, der eigentlichen Anwendung verlorengelht. Wenn aber bei jedem Generationswechsel ein immer größerer Teil des Aufwands in die Übertragung auf die nächste Systemgeneration investiert werden muß, kann man die Frage stellen, ob die ursprüngliche Kosten-Nutzen-Rechnung noch Bestand hat. Auf jeden Fall ist es eine profitable Aussicht für Hersteller von Hard- und Software oder für Dienstleister, denen sich ein neuer Markt erschließt.

Alternativen aufzeigen!

Damit soll nun abschließend versucht werden, die Frage nach der Reaktion der Archive auf diese Herausforderungen zu beantworten. Die Antwort kann nur lauten: Alternativen aufzeigen! Ähnlich wie beim Zerfall saurer und/oder holzschliffhaltiger Papiere, wo ebenfalls keine Aussicht besteht, die Originale - und schon gar nicht in großer Zahl - dauerhaft zu erhalten, wird ein Ersatzträger benötigt. Und wie beim Papier gilt auch bei modernen Informationsträgern die banale Forderung der langfristigen Zugänglichkeit zu den gespeicherten Informationen: Werden Unterlagen aller Art, insbesondere solche von bleibendem Wert, auf einen Ersatzträger übertragen, muß dessen langfristige Lesbarkeit gewährleistet sein. Daneben muß der Ersatzträger den Überlieferungszusammenhang der einzelnen Schriftstücke wiedergeben.

Für eine Konversion ist daher nicht das rasche Einlesen am Büroarbeitsplatz, nicht der komfortable Zugriff oder eine vordergründige, weil stets zeitgebundene Wirtschaftlichkeitsrechnung das entscheidende Kriterium, sondern die Fähigkeit des Ersatzmediums, den inneren und äußeren Informationsgehalt der Vorlage zu bewahren, und zwar auf Dauer.

Was aber ist nun die Alternative? Die Suche danach ist selbstverständlich an den Anforderungen auszurichten, die Archive und Archivare an die bildliche Speicherung für Archivgut oder Unterlagen von bleibendem Wert zu stellen haben:

das Verfahren zur Speicherung muß vorlagenschonend sein,
die Abbildungsqualität muß so hoch wie möglich sein,
die Lesbarkeit muß auf Dauer gewährleistet sein,
der Ordnungszusammenhang muß zweifelsfrei wiedergegeben werden,
der Ersatzträger muß selbst alterungsbeständig sein und
das Verfahren muß langfristig wirtschaftlich und berechenbar sein.

Es liegt nahe, an dieser Stelle den Vergleich zum Mikrofilm zu ziehen, über den jahrzehntelange Erfahrungen vorliegen, und den abzulösen die zuvor erwähnten Systeme angetreten sind. Eine Matrix der genannten Anforderungen sähe wie folgt aus:

Anforderung	Mikrofilm	Digitalverfahren
Bildqualität (Auflösung)	sehr hoch	niedrig

Alterungsbeständigkeit	ja	nein
Systemlebensdauer	langfristig	unvorhersehbar
Wirtschaftlichkeit	berechenbar	fraglich

Der Vergleich geht eindeutig zugunsten des Mikrofilms aus. Mikrofilm, rechtzeitig eingesetzt und richtig verarbeitet, ist und bleibt der einzige alterungsbeständige und dauerhaft wirtschaftliche Massenspeicher für verlorene, beschädigte oder in ihrer Erhaltung gefährdete Papiere. Er bietet genau jene Langzeitperspektive, die allen digitalen Verfahren fehlt. Im übrigen ermöglicht auch er eine vielfache Nutzung an unterschiedlichen Orten ohne Rückgriff auf das Original. Dies soll die Annehmlichkeiten der Elektronik bei der Nutzung nicht schmälern: Eine schnelle Auswahl einzelner Seiten unter der genannten Voraussetzung einer funktionierenden Indizierung, die Weiterverarbeitung von Kopien in herkömmlichen Textverarbeitungen und vieles mehr sind rosige Aussichten für manchen Nutzer. Aber: Diese Nutzungsdateien sind wohlgerne auch vom Mikrofilm zu gewinnen, und zwar ebenfalls mittels Scanner. Wenn Archive und Bibliotheken eines Tages zu der Auffassung gelangen, daß es nützlicher und wirtschaftlicher ist, zunächst die Findmittel, später vielleicht auch die Bestände über Datenbildschirm statt über Mikrofilmbildschirm nutzen zu lassen, so steht dem - läßt man einmal die Kostenfrage außen vor - grundsätzlich nichts im Wege: Der Mikrofilm verbaut keine Nutzungsform der Zukunft. Umgekehrt aber gleiche der Versuch, elektronische Systeme heutigen Typs als Langzeitspeicher einzusetzen, der sprichwörtlichen Reise ins Ungewisse.

Anbieter von digitalen Speicherungssystemen äußern sich nicht über Aufbewahrungszeiten - wie sollten sie auch. Sie sind vollauf damit beschäftigt, alle bisherigen Aufbewahrungsformen - Papier, Fotos, Tonträger, Bücher, Pläne und Zeichnungen - abzulösen. Der gute alte Mikrofilm hat in der breiten Öffentlichkeit keine Lobby mehr, in der täglichen bunten Werbung kommt er nicht mehr vor. Es ist ja bekannt, daß das Bild einer Sache in der Öffentlichkeit, das *Image*, oft wirkungsmächtiger ist als die Sache selbst. Um so wichtiger erscheint es, daß die Archive und die Bibliotheken selbst über eigene Parameter verfügen, um die mitunter verwirrenden Informationen über die Informationstechnologie einzuordnen. Es gilt also vernünftig zu entscheiden, auf welche Weise und mit welchen Mitteln die zur Verfügung stehenden Ressourcen einzusetzen sind, um das Beste für die Zukunft unserer Zeugnisse der Vergangenheit zu erreichen.