Dokumenten-Management (DMS) und Archivierung für KMU

Ein Leitfaden für die Praxis, Version 9.1.0

Inhaltsverzeichnis

1 Einleitung 2
- 1.1 Weshalb gibt es diesen DMS-Leitfaden? 2
- 1.2 Tipps für den Einstieg 3

2 Rechtslage bei der elektronischen Archivierung 5
- 2.1 Rechtslage in der Schweiz . . 5
- 2.2 Rechtslage in Deutschland . . 5
- 2.3 Elektronische Signaturlösungen 6
- 2.4 Frage der Zertifizierung . . . 7
- 2.5 Spezialgesetzgebung 7

3 Konzepte für ein DMS 8
- 3.1 Datenbank- versus Ordnerstruktur 8
- 3.2 OpenSource versus Closed-Source 8
- 3.3 Integration in bestehende IT-Landschaft 11
- 3.4 Original- versus Archivformat 12
- 3.5 Bilddateien als Archivformate 13
- 3.6 Die wichtigsten Formate . . . 14
- 3.7 Haltbarkeit der Medien . . . 15
- 3.8 Backup-Konzept 16

4 Wahl der Hardware 19
- 4.1 Archiv-Server 19
- 4.2 WORM-Datenträger 20
- 4.3 Scanner-Wahl 21

5 Aufbau der Archive 24
- 5.1 Digitale Archive auf dem Server 24
- 5.2 Internet basierende Archive . 24
- 5.3 Selbsttragende Archive . . . 25
- 5.4 Ausgelagerte Archive (ASP) . 26
- 5.5 Boxbasierende Lösungen . . . 26
- 5.6 Ins Betriebssystem integrierte Lösungen 26

6 Art der Beschlagwortung 28
- 6.1 Manuelle Beschlagwortung . . 28
- 6.2 Automation mit Barcodes . . 28
- 6.3 Texterkennung (OCR) 29
- 6.4 Formularerkennung und weiteres 29

7 Evaluation einer Lösung 31
- 7.1 Was benötigen wir? 31
- 7.2 Sichten der Unterlagen . . . 31
- 7.3 Produkte prüfen 31
- 7.4 Demo-Termin, Referenzen, Offerten und Pflichtenhefte . 32
- 7.5 Projektabwicklung 33

8 Eckpunkte einer Lösung 34
- 8.1 Dokumente erfassen 34
- 8.2 Papierbelege 34
- 8.3 Einzel-Dateien 35
- 8.4 Office-Dateien 35
- 8.5 E-Mails 36
- 8.6 ERP-Daten 36
- 8.7 Workflow 38
- 8.8 Suchen von Informationen . . 39
- 8.9 Navigation 40
- 8.10 Import- und Exportfunktionen 40
- 8.11 Administration und Sicherheit 41

9 Copyright-Hinweise 42

© 23.1.2009 by Archivista GmbH, Homepage: www.archivista.ch

1 Einleitung

1.1 Weshalb gibt es diesen DMS-Leitfaden?

Was ist ein Dokumenten-Management-System (DMS)? Gibt es das papierfreie Büro? Was ist ein digitales Archiv? Und was um Himmels willen ist ein ECM-System?

Grob vereinfacht kann gesagt werden: Archivierung bedeutet mehr das Lagern und DMS mehr das Verwalten (Managen) von Belegen. ECM (Enterprise Content Management) ist der Versuch, sämtliche Informationen eines Unternehmens in einer Anwendung (eben der ECM-Lösung) zu bündeln.

Und lassen Sie mich noch etwas sagen. Vergessen Sie diese Schlagwörter am besten gleich wieder. Die Grenzen sind derart fliessend, dass einmal jeder Anbieter von DMS spircht und ein Jahr später reden dann alle (weil es trendy ist) von ECM.

Gelingt es nur schon, die Ablagekopie einer jeden Rechnung in digitaler Form über 10 Jahre aufzubewahren (hier ist dann mehr von digitalier Archivierung die Rede), dann ist schon viel gewonnen.

Und noch etwas: Gerade bei KME-Unternehmen dürfte für den Einsatz einer DMS-Lösung in erster Linie matchentscheidend sein, wie einfach sich die Lösung in die bestehende Umgebung einpflegen lässt. Hier dürfte vorallem die Integration in die im Einsatz stehende ERP-Lösung (Enterprise Ressourcing Planing) zentral sein.

Meist kommt aber der Anstoss zu einem digitalen Archiv oder DMS nicht direkt aus dem Effizienzgedanken heraus, sondern ist das Resultat von einem oder mehreren Fragestellungen. Diese können sein:

- Es muss ein neuer Kopierer angeschafft werden; man spielt mit dem Gedanken, gleich einen Scankopierer zu kaufen und das **Ablagewesen generell zu modernisieren.**
- Die Mail-Konten quellen über. Niemand weiss, was wann, wo und wie archiviert werden soll.
- Die neue Version der ERP-Lösung arbeitet vorzugsweise mit Laserdruckern, sodass Archivkopien wesentlich umständlicher werden. Würden Sie digital abgelegt, könnten erhebliche Kosten gespart werden.
- Plötzlich fehlt der **Platz für die Ablageschränke,** weil zwei Abteilungen zusammengelegt werden.

- Und wenn mehrere Standorte bestehen, dann stellt sich immer die Frage, von wo und wie aus auf die Archivdaten zurückgegriffen werden kann.

Dies sind nur einige Beispiele. Es lohnt sich, eine Ist-/Soll-Analyse auf Papier zu machen, ganz egal, welche Überlegungen einen zu einer DMS- und oder Archivlösung hinführen. Die Vorteile der digitalen Variante sind zwar verbrieft. Das Fraunhofer Institut fand heraus, dass man beim Ablegen mehr als die Hälfte schneller ist und beim Suchen sogar zehn bis zwanzigmal schneller. So richtig unheimlich wird das Potential der elektronischen Ablage also erst beim Suchen. Doch selbst wenn Sie wenig nach Belegen suchen müssen, aber diese einfach gerne digital vorrätig haben, kann sich die Anschaffung nach einer bestimmten Zeit lohnen, ganz abgesehen vom Komfort, der Ihnen einen Sofort-Nutzen bringt.

➲ *Dieser Leitfaden richtet sich in erster Linie an kleinere und mittlere Unternehmungen (KMU). Dies deshalb, weil gerade hier die wichtigsten Fragen bei der Evaluation oft zu kurz kommen. Praktisch heisst dies, dass ein entsprechendes Projekt meist ohne externe Beratung bzw. Pflichtenheft angegangen wird. Dies ist bei grösseren Betrieben seltener der Fall, aber selbstverständlich dürfte dieser Leitfaden in jeder Unternehmung wertvolle Dienste leisten.*

1.2 Tipps für den Einstieg

Bevor wir die einzelnen Fragestellungen vertieft behandeln werden, möchten wir Ihnen einige Fragen mit auf den Weg geben, die Ihnen später bei der Evaluation immer wieder helfen werden, den Überblick zu bewahren:

- Fragen Sie sich konkret, wie oft denn nach bestimmten Belegen und/oder Informationen gesucht werden muss?
- Seien Sie sich über ihr **Beleggut** im Klaren: Legen Sie in einem Tabellenkalkulationsprogramm ein Datenblatt an und listen Sie auf, wieviele Dokumente in welchem Formaten Sie haben. Schwarz/Weiss? Farbe? A4? Andere Grössen? Einseitig oder zweiseitig Bedrucktes? Gut leserliche Originale oder Durchschlagpapiere in allen Regenbogenfarben?
- Seien Sie sich auch über Ihre Visionen im Klaren: sind es "nur" die **laufenden Belege,** die im DMS auffindbar sein sollen oder gibts auch **Altbestände?** Wieviele Belege sollen ausserdem direkt aus dem Computer ins Archiv überführt werden? Von welchen Applikationen? Auch in Farbe? Nach welchen Kriterien sollen die Belege später suchbar sein?
- Versuchen Sie die **Kosten der bisherigen Ablage** möglichst genau zu erfassen. Seien Sie dabei ehrlich mit sich selbst. Es kann nicht sein, dass Sie nur die Miete des Ar-

chivraumes bzw. die Materialkosten erfassen. Denken Sie daran, dass jeder Gang ins Archiv Personalkosten verursacht.

2 Rechtslage bei der elektronischen Archivierung

Leider ist es nicht möglich, die Rechtslage im Rahmen dieses Leitfadens abschliessend zu behandeln. Trotzdem möchten wir in einigen Sätzen die derzeitigen Regelungen für die Länder Deutschland und Schweiz darlegen.

2.1 Rechtslage in der Schweiz

Am 1. Juni 2002 trat die Geschäftsbücherverordnung (GeBüV) des Bundesrates in Kraft. Damit hat die digitale Archivierung in der Schweiz eine Rechtsgrundlage erhalten.

Zwei Punkte sind zentral: Einmal sind die **Daten geordnet** zu archivieren, das heisst jede Information muss innert **kurzer Zeit auffindbar** sein. Weiter müssen die Belege so erfasst und aufbewahrt werden, dass sie **nicht geändert** werden können, ohne dass sich dies feststellen lässt. Obwohl der Gesetzgeber auch eine Sicherung auf einem wiederbeschreibbaren Datenträger (Beispiel Festplatte) erlaubt, sofern eine Signierung erfolgt, ist es aus unserer Sicht sicherer, auf einmal beschreibbare Medien zu setzen. Wie eine Signatur auszusehen hat, darüber schweigt das Gesetz nämlich. Das wichtigste Ziel, die Sichtbarmachung der Originaldatei, erreicht man unkompliziert mit der CD-ROM bzw. einer DVD.

2.2 Rechtslage in Deutschland

Seit dem 1. Januar 2002 sind in Deutschland alle Unternehmungen archivierungspflichtig, sofern Sie die Buchhaltung elektronisch führen.

Es geht dabei um die Buchhaltung an sich, d.h. die Buchungsvorgänge, Bilanzen, Erfolgsrechnungen, etc. Die herkömmlich klassisch in Ordnern abgelegten Papierbelege müssen nicht elektronisch abgelegt werden.

Um dem Gesetz zu 100% nachzukommen, müssten Sie (unabhängig der zu archivierenden Belege) die Quelldateien der Buchhaltungssoftware sowie die Software selber über die gesamte Dauer von 10 Jahren lauffähig halten – oder aber die Buchhaltung gar nicht elektronisch führen (wie grosszügig der Gesetzgeber mal wieder war). Das Risiko besteht nun darin, dass Sie das mit vertretbarem Aufwand nicht hinkriegen.

Zentral für den Gesetzgeber in Deutschland ist die **"maschinelle Auswertbarkeit"** der Daten. Diese dürfen in einem gebräuchlicheren Datenbankformat (XBase, Access, ODBC) oder auch

als ANSI-Dateien vorliegen. Nicht zulässig sind reine Bitmap-Dateien oder auch PDF-Dateien, weil diese eben nicht maschinell auswertbar sind.

Ein DMS- bzw. Archivsystem 'schlägt zwei Fliegen mit einer Klappe': einerseits archiviert sie die Daten, die aus der Buchhaltungssoftware stammen. Andererseits ist sie natürlich auch für die Archivierung der Belege (z.B. Debitoren- und Kreditorenrechnungen, Kassenquittungen, etc.) geeignet, d.h. sie vermag die klassischen Ordner zu ersetzen.

Bei der elektronischen Archivierung (ohne Quellsoftware) ist entscheidend, dass Sie die Daten direkt ab Ihrer Buchhaltungs- in eine Archivierungslösung überführen können. Normalerweise benötigen Sie dazu einen virtuellen Druckertreiber, das heisst Sie drucken ins Archiv anstatt auf einen richtigen Drucker. Entscheidend dabei ist, dass nicht nur ein Bitmap erstellt wird, sondern dass Sie ebenfalls die **ANSI-Rohdaten** übernehmen können.

Was "steuerlich relevant" für Ihr Unternehmen bedeutet, müssen Sie selber entscheiden. Wenn jedoch alle steuerlich relevanten Daten im ANSI-Format – und für die Belege in der Form von Bitmaps – in Ihre Archivierungslösung überführt werden, genügen Sie den gesetzlichen Vorschriften vollauf.

2.3 Elektronische Signaturlösungen

Derzeit haben die Marketing-Abteilungen die digitalen Signaturen entdeckt. Dabei werden elektronische erstellte Belege (und nur um diese geht es) gar nicht mehr zu Papier gebracht, sondern mit einem Siegel versehen, der sicherstellen soll, das später festgetellt werden kann, dass das Dokument von einem bestimmten Empfänger stammt. Leider müssen diese Siegel alle zwei bis drei Jahre erneuert werden, weil die Computertechnologie derzeit nach wie vor derart rasant weiterentwickelt wird, dass die Siegel schon innerhalb einiger Jahre gefälscht werden könn(t)en. Dabei werden dann Nachsignaturen erstellt.

KMU-Unternehmen sollten sich den Einsatz von Signatur-Lösungen derzeit gut überlegen. Notwendig sind sie dort, wo aufgrund digitaler Signaturen ein Vorsteuerabzug geltend gemacht werden soll. Für ein KMU-Unternehmen bedeutet dies aber nur, dass (zur Hauptsache geht es um PDF-Dateien) die entsprechenden Dateien von grösseren Unternehmen im Originalformat aufzubewahren sind. Erst wenn das Unternehmen elektronische Rechnungen versenden will, ist eine Signatur-Lösung einzuführen. Die derzeitigen Einführungskosten übersteigen im Moment aber gut und gerne die Druck- und Portokosten, sodass sich ein Einsatz aus marktwirtschaftlichen Gründen kaum lohnt.

Im übrigen tragen Signaturlösungen nicht zur Rechtssicherheit bei. Die Signaturen bezugen nur, dass eine bestimmte Datei von einem Absender mit einem bestimmten Zertifikat kommt,

nicht aber, ob die generierten Daten auch frei von inhaltlichen Widersprüchen geschweige denn lesbar sind.

2.4 Frage der Zertifizierung

Einige Anbieter bescheinigen ihren Lösungen eine Zertifizierung entweder durch einen externen Treuhänder oder aber direkt durch das Finanzamt. Eine **Zertifizierung durch die Finanzämter gibt es nicht,** wie das Bundesministerium für Finanzen festhält. Letztlich sind alleine die gesetzlichen Anforderungen massgebend[1].

2.5 Spezialgesetzgebung

Die Archivierungs- und Aufbewahrungspflicht wird meist von der Buchführungspflicht (Buchhaltung) her abgeleitet. Es gilt zu beachten, dass diese primär auf privatrechtlich organisierte Unternehmungen zugeschnitten ist. Die öffentliche Hand kennt meist ganz andere Anforderungen, genauso wie dies zum Beispiel für Patientendossiers von Ärzten und Spitälern der Fall ist.

[1] siehe Homepage www.bundesfinanzministerium.de

3 Konzepte für ein DMS

Welche Konzepte verfolgen die verschiedenen Hersteller von DMS-Lösungen? Vieles ist gleich (ein Scan-Modul oder 'Capturing-Modul' haben fast alle), vieles ist aber auch anders. In der Folge einige Bereiche, wo die Philosophien auseinander gehen.

3.1 Datenbank- versus Ordnerstruktur

Die Ordnerstruktur kennen Windows-Benützer vom Windows-Explorer her. Sie bildet die Papierablage sehr gut ab. Sie hat aber auch Nachteile. Hier sind Vor- und Nachteile aus unserer Sicht zusammengefasst.

Ablageart	Vorteile	Nachteile
Datenbankstruktur Jahr | Gebiet | Firma 1999 | Ausland | Müller 1999 | Ausland; Inland | Peter 1999 | Inland | Baumann 2000 | Ausland; Inland | Peter 2000 | Inland | Müller	Geordnete Ablage mit eindeutigen Kriterien, einfaches Erfassen, flexibles Suchen und Sortieren der Akten	Benützer brauchen ein gewisses Abstraktionsvermögen
Ordnerstruktur archiv 1999 Ausland Müller Peter Inland Baumann Peter 2000 Ausland Peter Inland Müller Peter	Anschaulich; der Benutzer "versorgt" seine Daten im entsprechenden Ordner	Belege, die kategorienübergreifend sind, müssen mehrfach abgelegt werden, Ablegen und Suchen benötigt mehr Zeit (Aufklicken der Ordner), Sortieren nur nach bestimmten Kriterien

➲ *Für beide Strukturen – wie auch für die Papierablage – gilt, dass ohne ein gewisses Minimum an Disziplin das Gesuchte zu wenig zuverlässig wieder auffindbar ist.*

3.2 OpenSource versus ClosedSource

3.2.1 ClosedSource

Gehört die Lizenzierungsform eines Produktes in das Kapitel "Konzepte für ein DMS"? Oder bedarf es dazu eines eigenen Kapitels? Ist das überhaupt matchentscheidend für mein Projekt?

Diese Fragen sollten gestellt werden, denn ein DMS-System sollte langfristig ausgelegt sein und hier spielt die Lizenzierungsform (dies ist generell bei Softwareprodukten der Fall) eine wichtige Rolle.

ClosedSource-Software bedeutet, dass Sie eine ausführbare Applikation erhalten. Der Hersteller/Verkäufer sichert bestimmte Eigenschaften zu. Bei Problemen sichert der Hersteller bzw. Lieferant Hilfestellung zu, d.h. Sie erhalten entsprechende Updates. Dies entweder (meist) kostenfrei im Rahmen eines Wartungsvertrages (nicht kostenfrei) bzw. gegen eine Updategebühr (i.d.R. alle eins bis zwei Jahre).

Der Kunde erhält dabei keinen Einblick in den Quellcode, d.h. die Art und Weise, wie das Produkt technologisch aufgebaut ist, bleibt das "Firmengeheimnis" des Anbieters. Bei Problemen ist der Kunde dringend auf die Hilfe des Herstellers angewiesen. Aus diesem Grunde wird der Quellcode bzw. die Applikation auch ClosedSource genannt, weil die Quellen der Software nicht offengelegt sind.

3.2.2 ClosedSource mit Hinterlegung des Sourcecodes

Bei grösseren Projekten wird deshalb oft ein Softwarehinterlegungsvertrag abgeschlossen. Dabei erhält der Kunde, im Falle, dass die Firma nicht mehr in der Lage ist zu liefern (z.B. bei Konkurs), Einblick in die Sourcen. Ob und inwiefern er auch Rechte erhält, die Applikation weiterzuwarten, ist detailliert in einem entsprechenden Vertrag vertraglich zu regeln.

Auch wenn die Softwarehinterlegung einen gewissen Schutz bieten kann, so hilft sie in vielen Fällen trotzdem nicht weiter. Ein Einsichtsrecht mag zwar helfen, den grössten Schaden zu vermeiden. Viel problematischer erscheint heute aber die Tatsache, dass es selbst mit einer Einsichtnahme praktischerweise kaum möglich sein wird, eine Applikation zu warten. Denn viele ClosedSource-Projekte basieren auf weiteren kommerziellen Software-Module, und für diese bestehen i.d.R. keine Lizenzrechte.

Offen gestanden hat der Autor in mehr als 20 Jahren bei sämtlichen KMU-Firmen keinen Fall angetroffen, wo die Hinterlegung der Sourcen geregelt worden wäre. Das Hinterlegen der Sourcen spielt vornehmlich bei grösseren Firmen oder bei der öffentlichen Hand eine Rolle. Es ist aber in jedem Falle von Kosten im vier- bis fünfstelligen Breich pro Jahr für entsprechende Verträge zu rechnen.

3.2.3 OpenSource

Dies ist bei OpenSource-Programmen anders. Hier muss der Quellcode des Programmes ausgeliefert werden. Eine unter OpenSource-Lizenz publizierte Anwendung kann und darf jederzeit angepasst werden. Allerdings gibt es für die Anwender/innen i.d.R. keinen Rechtsanspruch, falls eine OpenSource-Applikation nicht ordnungsgemäss arbeitet.

Zwar wird ein Anbieter (das gilt wohl allgemein für Open- und ClosedSource) kaum Interesse haben, eine fehlerbehaftete Applikation auszuliefern. Allerdings muss eine OpenSource-Applikation nicht zwingend à jour gehalten werden, d.h. es kann vorkommen, dass Sie zwar einen Quellcode erhalten, dieser aber unter Umständen seit Jahren nicht mehr gewartet wurde.

3.2.4 OpenSource mit Supportvertrag

Aus diesem Grunde war OpenSource eine lange Zeit bei Firmenkunden eher verpönt bzw. es herrschte eine Unsicherheit, inwiefern es denn überhaupt Support bei OpenSource-Produkten gebe.

Heute gibt es für OpenSource-Software, die aktuell gewartet wird, diese Problematik kaum mehr. Support und Wartung werden i.d.R. ausdrücklich in einem Softwarewartungsvertrag geregelt. Der Kunde erhält einmal die Sourcen unter einer OpenSource-Lizenz und weiter werden diese durch einen Vertragspartner vertraglich im Rahmen eines Softwarewartungsvertrages bindend gepflegt, d.h. der Kunde erhält Anspurch auf Support und Wartung der Software. Updates werden i.d.R. laufend zur Verfügung gestellt und der Kunde kann (muss aber nicht!) jederzeit Einblick in die Arbeitsweise des Anbieters nehmen.

Dies führt i.d.R. zu einer erhöhten Transparenz seitens des Kunden, kann dieser doch jederzeit "nachprüfen", ob und wie die Software gewartet wird. Er kann diesen Job aber auch durch ein Drittunternehmen ausführen lassen, was bei KMU-Unternehmungen oft gemacht wird, weil die entsprechenden Ressourcen inhouse nicht zur Verfügung stehen.

Einfache Checks sind selbst für einen Laien möglich, denn Software wird heute praktisch ausschliesslich über Softwareverwaltungssysteme gewartet, und für diese Systeme gibt es meist webbasierte Lösungen, um sich die Änderungen in grafischer Form darstellen zu lassen. Dies hilft mitunter dann, wenn plötzlich ein Feature, das bisher tatellos funktionierte, in einem neuen Release plötzlich nicht mehr funktioniert.

3.2.5 Und warum ist die Lizenzierung nun zentral?

Richtig, diese Frage wurde noch nicht explizit beantwortet. Stellen Sie sich vor, die DMS-Lösung steht und alle sind zufrieden. Nun wir ihr Anbieter aufgekauft (das soll vorkommen), oder er geht bankrott (das sollte nicht vorkommen, kommt aber zuweilen leider auch vor). Was passiert dann? Wie gut stehen die Chancen, dass Sie ihr DMS-Projekt im gewohnten Rahmen weiterbetreiben werden können?

Nehmen wir einmal den Fall eines Firmenaufkaufs. Dem Autor sind mehrere Fälle bekannt, wo ein grösserer Anbieter einen kleineren Anbieter "geschluckt" hat. Oder ein grosser Anbieter zieht sich zurück vom Markt, oft kommt es in diesem Falle zu einem Management-Byout. In beiden Fällen kann es passieren, dass zwar die Lösung weiterlebt, nicht aber das Preisgefüge.

Der Autor hat mehrere Firmen betreut, die Schlag auf Fall die Lösung wechseln mussten, weil das Update-Angebot plötzlich eine oder zwei Nullen mehr (vor dem Komma!) enthielt als dies bisher der Fall war.

Im Falle der Firmenliquidation sieht es noch ärger aus. Es gibt in diesem Falle bei ClosedSource-Lösungen (Ausnahme Software-Hinterlegungsvertrag, der aber bei KMU-Betrieben kaum je zur Anwendung gelangt) praktisch keine Chance, die Lösung im gewohnten Rahmen weiter zu betreiben. Der Migrations- und Kostendruck wird schnell hoch, es fallen ausserordentliche Kosten und vorallem ein hoher eigener Aufwand an, weil ja niemand (ausser vielleicht die alten Anweder/innen) noch wissen, wie die Lösung funktioniert.

Bei OpenSource-Produkten ist das einfacher. Die Sourcen stehen öffentlich zur Verfügung. Entweder kann der Kunde in einem solchen Falle selber Hand anlegen, oder aber er sucht einen Dienstleister, der diesen Job für ihn übernimmt. Dies ist bei ClosedSource-Projekten praktisch nicht möglich und gerade deshalb ist der Unterschied zentral und sollte auch bei der Anschaffung bereits berücksichtigt werden.

Und noch etwas, auch wenn der Anbieter von sich aus plötzlich Updates "absahnen" möchte, gibt es bie OpenSource deutlich bessere Chancen, dieser "Kostenfalle" elegant zu entrinnen, denn letztlich darf ich mit dem einmal publzierten Code ja machen, was ich möchte. Und dies wiederum wird sich ein jeder Anbieter von OpenSource-Lösungen bewusst sein, wenn er die Update- bzw. Wartungskosten festlegt.

So gesehen ist das Abhängigkeitsverhältnis bei OpenSource-Lösungen i.d.R. kleiner. I.d.R. heisst, dass es natürlich auch hier Ausnahmen geben kann, so etwa dann, wenn der Anbieter zwar angibt, OpenSource-Produkte auszuliefern, die Sourcen aber nicht ausliefert oder in einer Form ausliefert, die nicht dem entsprechen, was produktiv im Einsatz steht. Ebenfalls vorgekommen sind Fälle, wo die Dokumentation nicht zur Verfügung gestellt wird oder aber sämtliche Kommentare bewusst entfernt oder verfremdet wurden. Gut hinschauen lohnt sich daher in jedem Falle, auch dies gilt sowohl für Open- als auch ClosedSource-Lösungen gleichermassen.

3.3 Integration in bestehende IT-Landschaft

Bis vor kurzem gab es zwei Hauptrichtungen, mit welchen die Anbieter um die Gunst der Kunden buhlten.

Integration ins Betriebssystem: Wozu eine eigene Software für ein DMS, wenn wir doch alles direkt im 'Explorer' abbilden können? Die einfachste Lösung müsste doch diejenige sein, bei der sich die Benutzer/innen gar nicht erst bewusst sind, dass sie mit einer DMS-Lösung arbeiten. Als Vorteil fällt eine geringe Einarbeitungszeit für die Benutzer an, als Nachteil

müssen aber hohe technische Hürden erklommen werden, weil zum Beispiel jede Änderung auf der Stufe des Betriebssystems das DMS-System zum Erliegen bringen kann.

Plattformunabhängigkeit: Manche Hersteller verlassen sich auf Technologien, die auf verschiedenen Plattformen (z.B. Unix und Windows) eingesetzt werden können. Die meisten Desktop-Anwender arbeiten heute nach wie vor mit irgendeiner Windows-Version. Was aber, wenn ich migrieren möchte? Damit wir uns richtig verstehen, es geht hier nicht darum, ob eine Migration heute und jetzt einen Sinn ergibt, sondern nur darum, sich darüber Gedanken zu machen, was passiert, wenn ich wollte oder müsste. Vorteil der Plattformunabhängigkeit ist, dass eine Migration später viel einfacher ist.

Die beiden Richtungen – Integration ins Betriebssystem und Plattformunabhängigkeit – schliessen sich in der Regel (aber nicht immer) aus, weil je mehr eine Lösung mit dem Betriebssystem verzahnt ist, es für den Anbieter desto schwieriger wird, die gleiche Lösung auf mehreren Plattformen anzubieten. Zudem muss zwischen Unabhängigkeit zwischen Client und Server unterschieden werden. Eine Lösung kann zum Beispiel bei der Wahl des Servers plattformunabhängig sein, nicht aber bei den Desktops.

Seit jüngerer Zeit gibt es auch das **Prinzip des digitalen Aktenschranks:** Hardware und Software sind zu einem Archiv-Server verwoben, der Zugriff erfolgt meist über den Web-Browser. Damit ist die Lösung in Bezug auf den Desktop unabhängig. Der Archiv-Server wird als Black-Box betrachtet, die als Plug- and Play-Komponennte in die bestehende Systemlandschaft eingebunden werden kann. Eine Installation von Clients fällt dahin. Das Zusammenwirken mit dem Rest der IT erfolgt über offengelegte Schnittstellen und die Verantwortlichkeiten sind klar. Dies erspart manchen Supportaufwand. Andererseits muss auch hier wieder die Migrationsfähigkeit der Lösung abgeklärt werden.

3.4 Original- versus Archivformat

Die Schwerpunkte der Produkte im Bereich digitale Archivierung und Dokumenten-Management sind oft verschieden. Es gibt solche, bei denen wird der Wert auf die klare (schlanke und kostengünstige) Organisation aller im Unternehmen anfallenden Dokumente gelegt. Hier gibts wieder Subtypen; die einen sehen alles weborientiert, da heisst das Zauberwort dann ECM oder Enterprise Content Management, bei den andern steht die Produktion von Papierdokumenten im Vordergrund, da heisst die Losung 'Total Output Management'. Die wirkliche, langfristige Archivierung ist in einem solchen Zusammenhang oft nur ein Nebeneffekt. Diese Art von Produkten arbeiten häufig mit Originaldateien, auch wenn die Rede von digitaler Archivierung ist.

Andere Produkte setzen auf die **Langlebigkeit von digital vorhandenen Daten.** Diese ist natürlich relativ, gerade wenn wir die Geschichte menschlicher Zeugnisse betrachten: In Stein Gehauenes kann der Witterung ausgesetzt sein und hält trotzdem Tausende von Jahren, auf Papier Vorhandenes mag auch Tausende von Jahren halten, aber nur wenn die Dokumente entsprechend geschützt vor Witterung, Licht und extremen Temperaturunterschieden werden. Digitale Daten sind insofern heikler, weil neben der Aufbewahrung des Datenträgers eben auch die Daten so abgelegt werden müssen, dass sie später wieder gelesen werden können. Darum lohnt es sich, auf Formate zu setzen, die sich in der kurzen Geschichte der Computertechnologie bereits bewährt haben und zu einem Standard geworden sind. Zusätzlich spart man Geld, wenn man patentfreie Formate nimmt.

3.5 Bilddateien als Archivformate

Der Unterschied zwischen Original- und Archivformat bedeutet bis heute in der Praxis leider meist, dass wir das Originalformat (zum Beispiel eine Text-Datei) nicht als Archivdatei verwenden können, da eine Text-Datei beim Öffnen mit einem Textverarbeitungsprogramm mit jeder Software bzw. nur schon mit jeder Version einer Software ganz anders aussehen kann. Dazu ein Beispiel: Eine Winword6-Datei kann von Microsoft Word 2007 nicht in exakt der Weise dargestellt werden, wie sie vor zehn Jahren mit der damaligen Winword-Version des gleichen Herstellers ausgesehen hat.

Damit eine **(veränderbare) Text-Datei** einfach und sicher archiviert werden kann, wird eine **Bilddatei**[1] daraus hergestellt. Nicht mehr einzelne Wörter sowie die Formatierung werden gespeichert, sondern nur noch eine Bild-Datei.

Abbildung 3.1: Links: Gescannter Ausschnitt, das heisst Darstellung ist im Bild enthalten; Rechts: vektorbasierte Version, korrekte Anzeige muss durch Software realisiert werden. Hinweis: beide Ausschnitte sind stark vergrössert, damit der Unterschied ersichtlich wird.

[1] Die Bilddatei wird oft auch als Bitmap-Datei bezeichnet.

Dies kann zum Beispiel dadurch erfolgen, dass wir eine gedruckte Word-Seite erneut scannen und die Bilddatei speichern. Raffinierter ist das Arbeiten mit einem Druckertreiber, mit dem wir eine Bilddatei direkt ins Archiv drucken können.

Ebenso unsinnig wäre es, Papierbelege am Computer mit einem Textverarbeitungs- bzw. Layoutprogramm "nachbauen" zu wollen; letztlich geht es ja nur um die Archivierung des Dokumentes. In diesem Falle belassen wir die gescannte Datei direkt als Bilddatei.

3.5.1 PDF als Archivformat

In den letzten Jahren hat das PDF-Format[2] eine grosse Verbreitung erfahren. Einige Archiv- bzw. DMS-Anbieter haben es zum Standardformat ihrer Archivierungslösung erkoren. Der Vorteil dieses Verfahrens ist, dass die Dokumente leicht weitergegeben werden können. Es gibt aber auch Nachteile, und diese sind nicht zu unterschätzen:

- Es gibt mittlerweile unzählige Varianten des PDF-Formates. Jede PDF-fähige Applikation erzeugt leicht andere PDF-Dateien. Dies bedeutet, dass die Anzeige in (zugegeben) eher seltenen Fällen nicht zu 100% gewährleistet ist.
- Normalerweise sind weder die Schriften noch das komplette Bild in der Datei selber enthalten, dies kann zu einer späteren Zeit zu Darstellungsproblemen führen.
- Obwohl der Acrobat-Viewer kostenlos ist, fallen beim Erstellen von PDF-Dateien (zum Beispiel Acrobat-Capture oder -Writer von Adobe) zum Teil recht hohe Erstellungskosten an.
- Es gibt einen Archivansatz beim PDF/A-Format. Allerdings besteht hier das Problem, dass sämtliche Schriften eingebunden werden müssen und die Dokumente deswegen wesentlich mehr Platz in Anspruch nehmen als gerasterte Daten. Die Komplexität der Dokumente bzw. der Dokumentenherkunft bleibt dabei allerdings bestehen.

➲ *Zusammengefasst kann gesagt werden, dass sich das PDF-Format dann hervorrangend eignet, wenn es darum geht, Dokumente auszutauschen, nicht aber dann, wenn diese langfristig verfügbar gemacht werden müssen. Ob wir PDF-Dokumente, die wir heute erstellen, in zehn Jahren noch ohne Probleme lesen können, kann nicht schlüssig beantwortet werden.*

3.6 Die wichtigsten Formate

Betrachten wir nun die wichtigsten Bildformate, die wir heute in der Archivierung antreffen anhand einiger Merkmale:

[2] Portable Document Format von Adobe

Format	SW	Farbe	Patente	Verlustfrei	Platzbedarf
TIFF G4	ja	nein	nein	ja	klein
BMP	ja	ja	nein	ja	hoch
PNG	ja	ja	nein	ja	mittel
JPG	bedingt	ja	nein	nein	klein
JPG2000	bedingt	ja	bedingt	nein	klein
GIF	ja	bedingt	nein 5.2004	ja	klein/mittel
TIFF LZW	nein	ja	nein 5.2004	ja	klein/mittel

Die Frage nach der Wahl des richtigen Formates kann nicht pauschal beantwortet werden, da dies von der Art des Archives abhängig ist. Oft dürfte eine Kombination der verschiedenen Formate zum Einsatz kommen. Tendenziell können aber die folgenden Aussagen gemacht werden:

- **Bilddateien (Rasterformate)** sind in der Regel die besseren Archivformate als vektorbasierende Dateien. Bei einer "Archivierung" mit Originalformaten ist die originalgetreue Darstellung (Anzeige) nicht gewährleistet. Ebenso besteht ein Konvertierungsrisiko.
- **Freie (das heisst nicht patentbehaftete)** und weitverbreitete Formate sind besser als proprietäre.
- **Verlustfrei komprimierbare Formate** sind besser als solche mit Verlust. Platzsparende Formate sind besser als solche, die viel Platz benötigen.

3.7 Haltbarkeit der Medien

Bevor Sie für die Langzeit-Archivierung einen Datenträger auswählen, sollten Sie sich kurz mit den Eigenschaften der einzelnen Datenträger auseinandersetzen. Die nachfolgende Tabelle gibt einen kurzen Überblick.

➲ *Fazit: Derzeit dürfte der CDR-Datenträger die kostengünstigste und sicherste Methode für die Langzeit-Archivierung darstellen. Wichtig ist aber (und das gilt für alle Medien) eine optimale Lagerung. Datenträger sollten trocken und vor Sonneneinstrahlung geschützt gelagert werden.*

Die Daten zur Haltbarkeit der Datenträger entstammen einer Studie des Fraunhofer Instituts aus dem Jahre 1999. Betreffend der Preise beachten Sie bitte, dass diese in der Informatik immer einem Wandel unterworfen sind.

	CD-ROM	DVD	WORM	Magnetbänder
Standard-Medium	ja	beschränkt (DVD+R/-R)	nein	nein
Haltbarkeit	20-50 Jahre	20-50 Jahre	10-60 Jahre	2-20 Jahre
Schutz nötig vor	Sonne, Temperatur-schwankungen, Schmutz und Staub	Sonne, Temperatur-schwankungen, Schmutz und Staub	Sonne, Temperatur-schwankungen, Schmutz, Staub und Magnetfeldern	Sonne, Temperatur-schwankungen, Schmutz, Staub und Magnetfeldern
Fazit Haltbarkeit	relativ geringe chemische und physikalische Anfälligkeit; magnetisch nicht veränderbar	relativ geringe chemische und physikalische Anfälligkeit, magnetisch nicht veränder-bar	relativ geringe chemische und physikalische Anfälligkeit, magnetisch veränderbar bei hohen Feldstärken	relativ hohe chemische und physikalische Anfälligkeit, magnetisch veränderbar
Kosten 200 GB	CHF 600	CHF 600	CHF 8000	CHF 3600
Kosten 1 GB	CHF 3	CHF 3	CHF 40	CHF 18

3.8 Backup-Konzept

Als überlebenswichtige Faustregel gilt: Keine Archivierungslösung ohne Backup-Konzept. Wer Daten ohne Backup-Kopie hält, spielt nicht nur sprichwörtlich mit dem Feuer (im Brandfall), sondern mit dem Leben der Firma. Beim eigentlichen Backup müssen wir unterscheiden zwischen Redundanz-, Kurz- und Langzeit-Archivierung.

3.8.1 Redundanz auf Stufe Hardware

Alle wissen, dass eine Hardware dann und wann ausfallen kann. Daher werden Lösungen, bei denen längere Ausfallzeiten nicht verkraftbar sind, meist in irgendeiner Weise doppelt aufgebaut. Wir können dabei unterschieden zwischen Redundanz betr. Festplatte/Netzteil sowie kompletter Redundanz sowie der Zeit, bis ein neues System komplett in Betrieb genommen werden kann.

Eine Lösung, die zu 100 Prozent redundant aufgebaut werden kann, lässt sich i.d.R. sicherer betreiben als dies der Fall ist, wenn nur die Festplatte oder das Netzteil redundant gehalten

wird. Ein guter Server bringt heute zwar Festplaten- und Netzteil-Redundanz mit, ist aber nicht automatisch gefeilt (z.B. gegen Fehlmanipulationen auf dem Server direkt).

Zwei Rechner, die miteinander über ein Netzwerkkabel verbunden sind, und die Daten automatisiert austauschen, bieten hier deutlich mehr Sicherheit. Einmal, weil die Rechner nicht am gleichen Standort stehen müssen (Problematik eines Brandfalles) und weiter weil jederzeit mit dem zweiten Rechner normal weitergearbeitet werden kann, sollte einer der beiden Rechner ausfallen.

Redundanz ist sicher zu empfehlen, allerdings kann in kleinen Unternehmungen oft auch darauf verzichtet werden. Redundanz bietet zwar mehr Sicherheit, bedingt aber auch mehr Ressourcen beim Betreiben der Lösung. Oft und viel wichtiger (gerade bei kleinen Unternehmungen mit einigen Mitarbeiter/innen ist die Datensicherung (Backup). Denn mit Redundanz haben Sie vielleicht eine zweite Lösung im Betrieb, ein Datensicherungskonzept haben Sie damit allerdings noch nicht.

3.8.2 Kurzzeit-Backup

Ein **tägliches Backup** auf ein Band-Laufwerk oder zumindest auf eine andere Festplatte (mit Lagerung ausserhalb des Hauses) ist das Mindeste, was Sie benötigen. Falls die Datenmengen relativ klein bleiben, könnten auch DVD- oder CDR-Datenträger zum Einsatz kommen. Sinnvoll sind selbstverständlich auch gespiegelte Platten (RAID).

In jüngerer Zeit haben sich zunehmend auch protable USB-Festplatten als Backup-Medium etabliert. Diese sollten den Strombedarf direkt über den USB-Port beziehen können, da es ansonsten recht mühsam und umständlich ist, den Datenträger korrekt anzuschliessen. Und ja, um Himmels willen verwenden Sie midestens drei oder mehr Festplatten für die Datensicherung. Eine Festplatte geht nun mal erheblich schneller in die Brüche als Tape- oder CD/DVD-Datenträger (z.B. reicht unter Umständen bereits ein Runterfallen).

3.8.3 Langzeit-Backup

Für die Langzeit-Archivierung benötigen Sie **Datenträger, bei denen das Überschreiben nicht möglich** ist. Deshalb sind Band-Laufwerke für die Langzeit-Archivierung kaum geeignet, vielmehr sollten Sie CDR- und neuerdings auch DVD-Datenträger einsetzen. Bei den DVD-Datenträgern besteht leider noch immer die Problematik, dass DVD+R als auch DVD-R auf dem Markt sind. Das ist unbefriedigend, obwohl im Moment ein Qualitätslaufwerk beide Formate lesen bzw. schreiben kann. Sollten Sie aber einige Dutzend DVD-R erstellt haben und in einigen Jahren feststellen müssen, DVD-R hat sich nicht durchgesetzt, so werden Sie einige Mühe haben, diese dannzumal noch zu konvertieren, vom Aufwand ganz abgesehen.

Nicht verwendet werden sollten BlueRay-Datenträger, da diese im Moment (Stand 2009) viel zu wenig lange am Markt etabliert sind.

➲ *Auch wenn das Gesetz Sie dazu nicht verpflichtet, erstellen Sie bitte bei jedem Backupvorgang mindestens zwei Exemplare. Sinnvollerweise verwenden Sie dazu Datenträger zweier verschiedener Qualitätsmarken. Ebenso sollten Sie nach dem Brennvorgang die Integrität der Daten mit einem Prüfvorgang festhalten.*

4 Wahl der Hardware

Archivierungsprojekte sprengen oft und gerne Budgets, weil zuviele Ressourcen in die Hardware gesteckt werden. Dabei geht es nicht nur um den Anschaffungspreis der Hardware, sondern auch um deren Pflege und Wartung.

4.1 Archiv-Server

Fragen Sie nie einen Hardware-Lieferanten, ob Sie einen neuen Server benötigen oder nicht. Die Antwort wird bestimmt lauten, dass ein neuer Server das System stabiler und sicherer macht. Dass ein leistungsfähigerer Computer die Performance Ihrer EDV erhöht, gilt jedoch nicht unbedingt.

Abgesehen von dieser prinzipiellen Vorsicht gegenüber der Beschaffung von Hardware um der Harware willen gibt es doch Gründe, warum der Kauf eines separaten Servers für das Archiv-System Sinn machen kann:

- die vorhandene Festplattenkapazität reicht für die benötigten Datenmengen tatsächlich nicht aus
- im Supportfall ist oft schneller klar, welcher Lieferant sich dem aufgetretenen Problem widmen muss
- das zu beschaffende Archiv-System ist bereits vorkonfiguriert, so dass die durch die Hardware bedingten Mehrkosten durch Einsparungen bei der Implementation aufgewogen werden (siehe auch 3.3).

Was für einen Rechner Sie benötigen, hängt in erster Linie von der verwendeten Datenbank der Archivlösung ab. Je nach System (Oracle, IBM DB, MS SQL, MySQL, Access bzw. Btrieve) wird etwas mehr oder weniger Leistung vorausgesetzt. Die nötige 'Power' hängt neben der verwendeten Datenbank primär von der Häufigkeit der Archiv-Abfragen ab. Viele Benutzer/innen (Concurrent User) in einer Workflow-Umgebung erfordern mehr Rechenkapazität als ein paar Abfragen in einer reinen Archivlösung.

4.1.1 Skalierbarkeit der Hardware

Es lohnt sich zu vergleichen, für welche Produkte welche Hardwareanforderungen gelten. Wenn System X 256 MByte an Ram erfordert und System Y 4 GByte, dann werden Sie von der Preisgestaltung her wohl beide Produkte bezahlen können. Die Hardwareanforderungen sagen aber viel aus, wie hoch die Anforderungen an sich sind. Bei niedrigen Hardwareanforderungen

wird ein Ausbau verhältnismässig einfach sein, bei hohen Ressourcen können Sie nur noch darauf zählen, dass die Hardware beim ersten notwendigen Upgrade in ihr bestehendes Budget passt.

4.1.2 Berechnung der Harddiskgrösse

Um eine Vorstellung von den Datenmengen zu erhalten, die bei der Archivierung anfallen können, hier ein kleines Beispiel: Gehen wir von einem täglichen Belegvolumen von 500 Seiten aus und rechnen wir mit 250 Arbeitstagen, so fallen pro Jahr 125'000 Seiten an.

Gehen wir weiter davon aus, dass wir unseren Server im **Idealfall über fünf Jahre** im Betrieb haben, so ergibt das 625'000 Seiten. Damit wir eine gewisse Reserve haben, multiplizieren wir diese Zahl mit 1.5, damit erreichen wir nach 5 Jahren maximal 937'500 Seiten. Nun runden wir auf, das heisst wir rechnen mit einer Million Seiten. Der benötigte Speicherplatz je Seite hängt vom Archivgut ab. Bei einem hohen Scananteil dürften wir mehr Kapazität benötigen als wenn die Daten bereits im Computer vorliegen. 50 KByte je Seite haben sich in der Vergangenheit gut bewährt und so rechnen wir:

$$\frac{250\ Arbeitstage * 500\ Seiten * 5\ Jahre * 50\ KByte * 1.5}{1'000'000} = \text{ca. 50 GByte} \simeq \text{1 Mio. Seiten}$$

Wir sehen aus dieser Kalkulation, dass selbst im hohen Millionenbereich an Seiten herkömmliche Festplatten für die Zugriffspeicherung absolut ausreichen.

➲ *Und damit wir uns an dieser Stelle richtig verstehen: Die Festplatte alleine reicht für eine Archivierung der Daten nicht aus, Sie benötigen zusätzlich zur Zugriffsspeicherung ein Backup-Konzept. Es lohnt sich aber, Zugriff und Backup zu trennen, weil dadurch (siehe nachfolgend) der Einsatz teurer Hardware vermieden werden kann.*

4.2 WORM-Datenträger

Einige Archivanbieter bundeln die Lösungen mit sogenannten WORM-Jukeboxen[1]. Einstiegsmodelle mit ca. 100 bis 200 GByte kosten mehrere tausend Franken, wobei die Preise nach oben praktisch offen sind. Wo liegt nun der Nutzen dieser Geräte, mal abgesehen davon, dass sie (zusammen mit den Spezial-Platten) ein Vielfaches teurer sind? Die Hersteller nennen eine lange Lebensdauer und höhere Sicherheit.

Im Prinzip könnte das ein Vorteil sein, würden die entsprechenden Laufwerke nicht alle ein bis zwei Jahre in neuen Modellen auf den Markt gebracht, wobei man froh sein muss, wenn die neuen Modelle die alten Datenträger noch lesen können. Sie sollten sich bei der Evaluation auch bewusst sein, dass der **Zugriff im Vergleich zur Festplatte weit langsamer** ist und es

[1] Write Once, Read Many times Datenträger

keine Standards beim Zugriff auf die Datenträger über das Betriebssystem gibt. Schön für die Treiberlieferanten, Pech für Sie, falls Sie eine Migration in Angriff nehmen möchten/müssen.

Per 30. Juni 2006 stellte Hewlett Packard, ein namhafter Hersteller von WORM-Jukeboxen, die Produktion in diesem Bereich ein. Als Grund wurde die EU-Richtlinie 2002/95/EG 'Restriction of Hazardous Substances' (RoHS) genannt.

4.3 Scanner-Wahl

Welches Gerät für welche Fälle? Die Scanner-Evaluation ist eine Sache für sich und sollte ebenfalls seriös durchgeführt werden, besonders wenn spezielles Beleggut vorhanden ist. Fragen, die man sich hierzu stellen sollte:

- **Beschaffenheit der Belege:** Wie sehen die Dokumente aus, die gescannt werden sollen? Neben Grösse, einseitig bzw. doppelseitig bedruckten Belegen (Simplex/Duplex) spielt auch die Beschaffenheit des Papiers eine Rolle. Wird der Einzug zum Beispiel mit sehr vielen dünnen Papieren fertig?
- **Anzahl der Belege:** Wieviele Seiten werden es täglich sein? Die meisten Scanner-Hersteller geben an, für welche durchschnittliche Menge ihr Gerät ausgelegt ist. Es ergibt zum Beispiel keinen Sinn, für 90 Scans pro Tag einen 16'000-fränkigen Scanner zu kaufen. Ein Einzug dagegen lohnt sich bereits ab 20 Seiten pro Tag.
- **Scangeschwindigkeit:** Wie schnell soll Ihr Scanner sein? Beachten Sie, dass Scanner-Hersteller die Geschwindigkeit ihrer Produkte oft für 200dpi-Scans mit Schwarz/Weiss-Seiten angeben. Dies ist üblicherweise nicht das, was Sie in der Praxis benötigen. Denn entweder werden Sie zumindest mit 300dpi Schwarz/Weiss scannen oder Sie arbeiten mit Graustufe/Farbe, was eine erhebliche Einbusse bei der Scangeschwindigkeit zur Folge hat.

Die Wahl der Auflösung ist deshalb zentral, weil einmal das Scannen bei höherer Auflösung viel länger dauert und Sie zweitens grössere Dateien erhalten. Dazu einige Beispiele: Bei doppelter Auflösung erhalten Sie viermal mehr Daten, bei einem Wechsel zwischen Schwarz/Weiss zu Graustufe bzw. gar zu Farbe die 8- bzw. 24-fache Datenmenge. Eine A4-Seite bei 300dpi benötigt (unkomprimiert!) ziemlich genau 1 MByte Speicher. Falls Sie die gleiche Seite in Farbe scannen, fallen 24 MByte Rohdaten an, die allerdings meist komprimiert abgelegt werden, sodass mit ca. 300 bis 500 KByte pro Seite zu rechnen ist.

➲ *In der Praxis haben sich 300dpi sowohl bei Schwarz/Weiss als auch bei Farbe bewährt.*

Um das optimale Gerät zu ermitteln, empfiehlt es sich, zunächst vom **täglichen Belegvolumen** auszugehen. Meist finden Sie in den technischen Unterlagen eine empfohlene maximale Seitenanzahl pro Tag. Diese Zahlen sind grosszügig ausgelegt und sollten als Richtwerte angesehen werden. Es sollte aber nicht sein, dass (selbst eine Hilfskraft) den ganzen Tag beschäftigt ist, nur um das Volumen knapp bewältigen zu können.

In Verkaufsgesprächen wird oft herausgestrichen, dass Highend-Geräte bessere Scans liefern würden. Dies ist dann der Fall, wenn Sie mit Schwarz/Weiss scannen und das Gerät über eine Optimierungseinheit verfügt. Entsprechend ausgerüstete Scanner liefern in der Regel bessere Schwarz/Weiss-Scans. Aber gerade bei **kritischen Belegen** kann die Optimierung auch scheitern, zum Beispiel wenn ein Text hinterlegt ist. In der Praxis dürfte es daher einfacher sein, entsprechende Belege in **Graustufe bzw. Farbe** zu scannen, weil Sie dadurch eine fotografische Abbildung und nicht nur eine "aufgepeppte" Schwarz/Weiss-Kopie erhalten.

Nachfolgend stellen wir kurz die verschiedenen Scanner-Klassen vor. Wir unterscheiden Einstiegs- (bis ca. 200 Seiten am Tag), Mittelklasse- (bis ca. 1500 Seiten pro Tag) sowie Highend-Scanner (ab mehreren Tausend Seiten je Tag).

Abbildung 4.1: Gute Einstiegs-Geräte wie zum Beispiel Einzugs-Scanner mit USB-Schnittstelle sind für unter 1'000 Franken erhältlich. Ideal da, wo wenig Platz zur Verfügung steht. Nachteil: Kein Flachbett-Modus; sperrige Belege oder Bücher können nicht verarbeitet werden.

Abbildung 4.2: Mittelklasse mit Komfort: Flachbett-Scanner mit (Duplex-)Einzug. Ideal für Belegvolumen bis zu ca 1500 Seiten am Tag, Kostenpunkt ab ca. 2'000 Franken.

Abbildung 4.3: Die Anschaffung eines Hochleistungs- bzw. Spezial-Scanners (hier A0-Grossformat) bedingt Investitionskosten ab 5'000 Franken.

➲ *Eine Faustregel bei der Scanner-Evaluation lautet: Kaufen Sie kein Gerät, ohne das betreffende Modell mit ihren Belegen getestet zu haben. Achten Sie bei einem Test insbesondere darauf, ob und wie gut die Belege mit dem Einzug verarbeitet werden.*

4.3.1 Vernetzungsfähigkeit der Scanner

Es lohnt sich zu überprüfen, wie die Scanner ins System eingebunden werden. Teurere Scanner bieten meist schon eine Netzwerkschnittstelle an. Damit ist zwar die Vernetzung möglich, sie hängt aber oft davon ab, ob das Gerät später von der Lösung unterstützt wird.

Nicht zu empfehlen sind Low-Cost-Lösungen, die mit Windows-Treibern arbeiten. Der Aufwand, diese Geräte als Insel-Lösung in der Informatik-Landschaft am Leben zu erhalten, ist ungemein höher, als wenn die Lösung mit einer IP-Adresse betrieben werden kann. Im zweiten Falle werden Sie fast immer auch die Option finden, auf eine andere IP-Adresse in eine Datei zu scannen (FTP-Scannen).

Weiter stellt sich die Frage, wie einfach und zu welchen Kosten weitere Geräte angeschlossen werden können. Falls mehrere Standorte eingebunden werden sollen, ist diese Anforderung zwingend abzudecken. Und zwar bereits in der Lösung, mit der Sie starten möchten und nicht erst mit der Super-Mega-High-End-Lösung, die derzeit ihr Budget sprengt. Denn was derzeit ein Budget-Problem darstellt, wird auch morgen kaum anders aussehen.

4.3.2 Kopierer mit Scan-Funktion

Neuere Kopierer verfügen über Scanning-Funktionalitäten, welche für Dokumenten-Management und digitale Archivierung ebenfalls verwendet werden können. Je nachdem, wie die Scans vom Kopierer aufbereitet werden, kann es jedoch sein, dass zusätzlicher Aufwand entsteht, um Kopierer und DMS- oder Archiv-System optimal aufeinander abzustimmen.

5 Aufbau der Archive

Die Frage dieses Kapitels lautet, wie können wir eine zukünftige Archivlösung in unsere EDV-Umgebung einbinden, das heisst wer muss wann, wie und wo auf die Daten zugreifen können und wie soll das realisiert werden?

5.1 Digitale Archive auf dem Server

Das Archiv wird **zentral auf einem Firmenrechner** geführt. Der Zugriff erfolgt von den einzelnen Arbeitsrechnern über Zugriffsberechtigungen innerhalb des Netzwerkes. Dieser Zugriff wird entweder nur über die Datenbank (komplexer im Aufbau, dafür hohe Sicherheit) oder aber über Datenbank und Dateisystem (einfacher Aufbau, tiefere Sicherheit) realisiert.

Aussenstellen werden entweder über eine webfähige Applikation (am besten verschlüsselt mit HTTPS), Terminal-Server-Dienste oder über VPN-Verbindungen realisiert. In Einzelfällen kann auch Fernwartungssoftware (zum Beispiel VNC) zum Einsatz kommen. Die Lösungen im einzelnen:

- **HTTPS:** Verbindungsaufbau direkt über Web-Browser, der gesamte Datenverkehr wird verschlüsselt abgewickelt. Zentral dabei ist, dass die zuverwendende Lösung eine Verschlüsselung von Haus aus mitbringt.
- **Terminal-Server/Citrix:** Komplex(er) im Aufbau (Server), relativ hohe Kosten, Arbeiten mit mehreren Benutzern und tiefen Bandbreiten möglich.
- **VPN:**[1] Mittlerweile tiefe Implementierungskosten, keine Anpassungen auf Stufe Betriebssystem notwendig.
- **VNC:**[2] OpenSource, unter Windows nur als 1:1-Verbindung machbar, zum Teil etwas langsam.

➲ *Sofern mehr als 30 Prozent der Arbeitsplätze nicht direkt am internen Netz angeschlossen sind, sollte eine Lösung auf Stufe VPN oder HTTPS in Erwägung gezogen werden.*

5.2 Internet basierende Archive

Web-basierende Lösungen haben den Vorteil, dass auf den einzelnen Arbeitsplatzrechnern in der Regel **keine zusätzliche Software** installiert werden muss. Für einen Zugriff auf die

[1] Virtual Private Network, das heisst verschlüsselter Zugriff vom Extra- ins Intranet.
[2] Virtual Network Computing

Archivdaten benötigen wir lediglich einen Web-Browser (zum Beispiel Mozilla oder Internet Explorer).

Allerdings gilt es die einzelnen Anforderungen an den Browser zu beachten. Erfordert eine Web-Lösung Java-Script, so muss diese Option aktiviert sein, da ansonsten kein Zugriff auf das Archiv möglich ist. Problematisch sind Lösungen, die zum Beispiel nur mit einem bestimmten Browser laufen (z.B. Internet Explorer 7.x), weil damit sämtliche anderen Browser (Betriebssysteme) aussen vor bleiben.

Weiter kann das Aufsetzen eines Internet-Servers eine kostenintensive und von der Sicherheit her relativ heikle Sache darstellen (Patientendossiers z.B. sollten nicht öffentlich zugänglich im Netz stehen).

➲ *Ein weiterer Nachteil kann im Funktionsumfang von web-basierten Applikationen liegen. Browser-basierende Lösungen sind oft nicht ganz so effizient zu bedienen, wie dies bei installierten Applikationen der Fall ist. Modernere Web-Lösungen (Stichwort Web 2.0 oder Ajax) unterscheiden sich aber kaum mehr von herkömmtlichen Lösungen.*

5.3 Selbsttragende Archive

Bei der Archivierung fallen in der Regel die Belege in einem definitiven Zustand an, das heisst die zu archivierenden Belege müssen nachträglich nicht mehr geändert werden. Damit die Belege rechtlich gesehen als 'Archivbeleg' Gültigkeit erlangen, dürfen sie ja auch nicht mehr verändert werden.

Aus dieser Überlegung heraus kann der Wunsch entstehen, das was bereits abgeschlossen ist, möglichst so zu archivieren, dass später autonom auf diese Belege zugegriffen werden kann. Das heisst im Unterschied zu einer Netzwerkinstallation, wo die ständige Verfügbarkeit dadurch sichergestellt wird, dass gewisse Softwarekomponenten installiert sind, wird die **Software (Viewer) bei den selbsttragenden Archiven direkt bei den Archivdaten** abgelegt, so dass später autonom und ohne zusätzliche Software das Archiv durchforstet werden kann.

Vorteile: Die Archive bleiben überschaubar und der Zugriff kann jederzeit unabhängig von einer Softwareinstallation erfolgen. Die Archive können einfach transportiert werden und ebenso einfach an externe Stellen (zum Beispiel Revision) weitergereicht werden.

Nachteile: Die benötigte Viewersoftware befindet sich direkt bei den Daten. Die Software kann zu einem späteren Zeitpunkt (zum Beispiel bei Erscheinen von Windows 2022) unter Umständen nicht mehr ohne Probleme bzw. gar nicht mehr gestartet werden.

5.4 Ausgelagerte Archive (ASP)

Sofern Sie bis hierhin durchgedrungen und zur Überzeugung gelangt sind, dass Ihnen das alles viel zu kompliziert ist, so könnte die komplette Auslagerung der gesamten Archivierung für Sie eine Option darstellen. Je nach Ausbaustufe beim ASP[3] kann der Dienstleister (Service Provider) den gesamten Ablauf oder nur einen Teil des Ablaufes übernehmen. Meist wird dabei in monatlichen Gebühren bzw. fixen Seitenpreisen abgerechnet.

Allgemeine Aussagen zu ASP-Diensten sind derzeit kaum machbar. Es kann aber gesagt werden, dass a) die **Abhängigkeit vom Dienstleister** relativ hoch ist und dass b) die **Kosten** (so tief der Seitenpreis auch liegen mag) in fast jedem Fall **weit höher** sind, als wenn Sie die Archivierung in-house erledigen. Ebenfalls nicht vernachlässigt werden sollte die Problematik, dass die Daten physikalisch nicht bei Ihnen liegen. Läuft zum Beispiel der ADSL-Anschluss einmal nicht, so können Sie nicht mehr auf die Informationen zugreifen.

5.5 Boxbasierende Lösungen

In den letzten Jahren sind vermehrt Lösungen auf den Markt gekommen, die eine bestimmte Funktionalität auf einer bestimmten Hardware anbieten. Als Beispiel sei hier eine Firewall-Lösung angeführt, die heute meist fixfertig auf einer kleinen Hardware-Box geliefert und über einen Browser verwaltet wird.

Im Prinzip stellt eine boxbasierende Lösung eine Kombination aus einem Archiv-Server und einer webfähigen Lösung dar. Hier sollte unbedingt darauf geachtet werden, dass Box-Lösungen möglichst flexibel auf- und ausgebaut werden können. Auch die Wartung der Lösung sollte sichergestellt sein, weil die hauseigene Informatik kaum bereit sein wird, zusätzliche Wartungspflichten zu übernehmen, zumal boxbasierende Lösungen meist nicht mit einem handelsüblichen Betriebssystem ausgeliefert werden.

5.6 Ins Betriebssystem integrierte Lösungen

Bei der Einführung einer DMS-Lösung spielt die Akzeptanz einer Lösung eine zentrale Rolle. Jede zusätzliche Applikation bedeutet einen Lernaufwand für die Mitarbeiter/innen. Im Prinzip darf heute nur davon ausgegangen werden, dass ein/e Mitarbeiter/in gängige Office-Pakete kennt und mit einem Web-Browser zu arbeiten weiss. Daher sollten Sie eher davon ausgehen, dass die Akzeptanz dann hoch ist, wenn die Lösung entweder webbasierend laufen wird oder aber nathlos ins Betriebssystem integriert ist.

[3] Application Service Providing

Bei voll ins Betriebssystem integrierten Lösungen finden die Benutzer/innen meist zwei drei Menü-Einträge im Dateiexplorer vor, mit denen sich die Dokumente entsprechend verwalten lassen. Dadurch ist der Lernaufwand für die Mitarbeiter/innen minimal. Allerdings sollte bedacht werden, dass derartige Lösungen eng an bestimmte Betriebssysteme bzw. Versionen gebunden sind. Steht im Betrieb kein Update-Konzept für die Rechner zur Verfügung, kann das Aufsetzen eines später gekauften Computers einen überproportional hohen Aufwand verursachen. Es kann gar unmöglich werden, Rechner verschiedener Generationen nebeneinander laufen zu lassen, weil das DMS schlicht und einfach die Betriebssystem nicht unterstützt.

Als kleines Beispiel sei hier angefügt, dass Sie selbt auf einem Windws 9x/W2K-Rechner noch heute einen Mozilla-Browser installieren können, wohl aber nicht das aktuelle Relase einer DMS-Lösung, die ins Betriebssystem integriert ist. Spätestens wenn der offizielle Support eines Herstellers für eine Betriebssystemversion ausläuft, dürfte auch kein Support mehr für die DMS-Applikatioen zur Verfügung stehen. Eine weitere Problematik besteht darin, dass in einer Unternehmung unter Umständen verschiedene Betriebssysteme bzw. Varianten im Einsatz stehen. Bei ins Betriebssystem integrierten Lösungen sollte a) sichergestellt sein, dass alle gängigen Betriebssystem (Windows, Mac, Linux) unterstützt werden und b) dass die verwendeten Protokolle komplett offengelegt sind.

6 Art der Beschlagwortung

Bevor wir zur Evaluation schreiten (siehe Kapitel 7), sollten wir uns an dieser Stelle einige Gedanken betreffend der Beschlagwortung machen. Wichtig zu diesem Zeitpunkt ist weniger, dass wir eine definitive Entscheidung treffen können, sondern vielmehr, dass wir die verschiedenen Möglichkeiten beim Ablegen kennenlernen. Jeder "gute" Verkäufer wird Ihnen gerne bestätigen, dass das Ablegen von Dokumenten ganz von alleine abläuft; ob sich das letztlich für Sie auch lohnt, ist allerdings eine andere Frage.

6.1 Manuelle Beschlagwortung

Bei der manuellen Beschlagwortung werden die Belege von Hand kategorisiert. Sofern Hilfsmittel wie automatisch verwaltete **Schlagwortkataloge** (zum Beispiel mit einem Drop-Down-Feld) zur Verfügung stehen, kann diese Art der Beschlagwortung durchaus sehr effizient sein. Problematischer wird die ganze Sache, wenn keine Schlagwortkataloge vorliegen und mehrere Personen kategorisieren müssen. Dann ist entweder Disziplin angesagt oder das spätere Auffinden von Dokumenten endet im Chaos.

Kritisch wird es ebenfalls, wenn die Beschlagwortung in der Hektik des Arbeitstages zu kurz kommt oder im Extremfall ganz ausbleibt. Letztlich verhält sich hier ein digitales Archiv wie ein herkömmliches. Wer nicht diszipliniert arbeitet, wird die Quittung beim späteren Suchen erhalten.

6.2 Automation mit Barcodes

Barcodes werden bei der digitalen Archivierung vornehmlich für zwei Zwecke verwendet. Zum Einen kann das **Scanning vereinfacht** werden, indem man dem Scanner mittels Barcode mitteilt, dass er die Einstellungen anpassen muss: So können ganze Stapel verschiedenartigster Blätter (zum Beispiel A4 Hochformat Schwarz/Weiss, A4 Querformat in Graustufen oder A5 hoch in Farbe) ohne menschliche Intervention mit den richtigen Scaneinstellungen verarbeitet werden.

Zum Anderen – und dies ist wohl das aufregendere Einsatzgebiet – lässt sich die **Beschlagwortung von Datensätzen automatisieren.** Wenn immer wieder ähnliche Belege anfallen, zum Beispiel Kreditorenrechnungen oder Kundenrapporte, kann die wichtige Information, in

unserem Fall also Kreditor und Belegart (Rechnung) oder Kundenname und Belegart (Rapport) im Barcode codiert und nach dem Scanning automatisiert in die entsprechenden Felder abgefüllt werden. Auch wenn das korrekte Datum nicht in allen Fällen automatisiert eingetragen werden kann, kann die Verwendung von Barcodes sehr zeitsparend sein. Zudem ist die Fehlerquote geringer als beim Menschen.

Falls Sie immer wieder dieselben Belege nach denselben Kategorien ablegen müssen, können Sie ad hoc Barcode-Kleber (ausdrucken, auf Beleg kleben) oder vorbereitete Barcodeblätter (zum Beispiel für immer wiederkehrende externe Belege) verwenden. Als dritte Möglichkeit kann der Barcode auf die vom eigenen ERP-System kreierten Belege gedruckt werden (im oben erwähnten Fall der Kundenrapporte sinnvoll).

➲ *Barcodes sind erst ab einer gewissen Masse effizient. Wenn Sie weniger als 50 Belege pro Tag scannen wollen, bringt Ihnen der Barcode keine Vorteile.*

Wenn Sie denken, dass Barcodes für Sie in Frage kommen, dann erstellen Sie ein Mengengerüst. Wieviele Belege pro Tag? Nach welchen Kategorien sollen sie abgelegt werden?

6.3 Texterkennung (OCR)

Wenn wir unseren Computer mit dem gescannten Bild eines Beleges füttern, dann ist ja noch kein Text da. Es ist die OCR-Software, welche aus einem Scan den Text extrahieren kann. Sie macht das konkret, indem sie alle schwarzen Striche, Rundungen und Pünktchen abtastet und unter Zuhilfenahme von linguistischen Regeln als Buchstaben und Wörter interpretiert. Ein OCR-Modul ist da sinnvoll, wo ein **Volltextarchiv** benötigt wird. Als Beispiel mag eine Knowhow-Datenbank dienen, in die eine Treuhänderin alle aus Fachzeitschriften gesammelten Aufsätze sowie Vortragsskripts scannt. Wenn irgendwann eine knifflige Frage auftaucht, ist Wühlen in jahrelang angesammelten Papieren nicht mehr nötig, denn dank Volltextrecherche findet sie die benötigte Referenz in Sekunden. Auch ein E-Mail-Archiv macht als Volltextarchiv Sinn.

➲ *Wie gut ist die OCR? Lassen Sie sich vom Hersteller den Namen der eingesetzten Software sagen und Beispiele zeigen. Testen Sie die Qualität der Texterkennungssoftware mit eigenen Belegen selbst.*

6.4 Formularerkennung und weiteres

Oft werben Anbieter mit automatischer Beschlagwortung und/oder Belegkategorisierung. Oft wird dabei auf Praxisfälle verwiesen, wo ein hohes Volumen an Belegen automatisiert verarbeitet wird.

Meistens geht es dabei um die eingehenden Kreditoren-Rechnungen, die es zu erfassen gilt. Eine Formularerkennung kann eine wunderbare Sache sein, und ja, wenn eine Lösung dies zu erträglichen Konditionen mitbringt, dann kann ein Einsatz sicher geprüft werden.

Allerdings sind dem Autor keine Lösungen bekannt, die komplett ohne Lernphasen funktionieren. Meist werden die einzelnen Formularetypen einmal erfasst und mit einem Editor so bearbeitet, dass weitere Dokumente nunmehr automatisiert verarbeitet werden können.

Bei einer klassischen KMU-Unternehmung fallen einmal nicht allzu viele Kreditoren-Belege an und weiter gibt es aber doch eine stattliche Anazhl von Lieferanten (letztlich wird dort bestellt, wo der Preis stimmt). Daher scheitert das Einführen einer Formularverarbeitung i.d.R. schlicht und einfach am fehlenden Kosten-/Nutzenverhältnis. Es gibt pro Tag zuwenige Kreditoren-Rechnungen, als dass sich der Trainingsaufwand innerhalb nützlicher Frist (z.B. einem oder zwei Jahr/e) rentieren würde. Es muss auch daran gedacht werden, dass Briefformulare heute relativ schnell neuen Verhältnissen angepasst werden, und bei jeder Anapssung seitens eines Lieferanten bedarf es Anpassungen bei den erfassten Formulartypen.

Weiter kommt dann und wann auch Handschriftenerkennung zum Einsatz. Auch hier gilt, dass Lösungen in klassischen KMU-Unternehmungen am fehlenden Preis-/Leistungsverhältnis scheitern.

7 Evaluation einer Lösung

Nachdem nun schon wesentliche Themen und Möglichkeiten von DMS-Lösungen besprochen und konkretisiert wurden, soll der Rahmen nochmals geöffnet werden: wie bringt man eine Evaluation zum guten Ende? Auch hier gilt, methodisches Vorgehen hilft mehr als überstürztes Agieren[1]

7.1 Was benötigen wir?

Falls Sie eine reine Archivierungslösung benötigen, wird Ihnen das beste Workflow-Modul keine Hilfestellung sein. Am Anfang einer Lösung sollten Sie sich zunächst darüber im Klaren sein, was durch die Lösung abgedeckt werden soll. Hilfreich ist das Aufnehmen des Ist-Zustandes sowie eine kritische Würdigung. Anhand dieser können Sie in groben Zügen erkennen, auf welche Punkte Sie achten müssen.

7.2 Sichten der Unterlagen

Nachdem wir wissen, was wir in etwa suchen, sollten wir uns einen **Überblick über den Markt verschaffen.** Welche Anbieter kommen in Frage? Besuchen Sie deren Internet-Seiten[2]. Daneben kann auch der Besuch von Messen, die Teilnahme an Infotagen der Anbieter oder das Lesen von Fachzeitschriften eine Hilfestellung sein. Auf jeden Fall sollten Sie die Unterlagen einiger Anbieter in einer ruhigen Minute durcharbeiten. Informieren Sie sich stets zielgerichtet; von einer Produkte-Demonstration ist zu diesem Zeitpunkt noch eher abzuraten.

7.3 Produkte prüfen

Sie haben die Unterlagen geprüft. Nun können Sie anhand dieser die einzelnen Produkte einem ersten Test unterwerfen. Versuchen Sie herauszufinden, was die verschiedenen Produkte und Module der Anbieter für Sie konkret bringen. Zu diesem Zeitpunkt sollten Sie ebenfalls in etwa die **Kostenfolgen abschätzen** können. Wo liegen die Lizenzkosten? Jeder Anbieter sollte zumindest angeben können, in welcher Grössenordnung sich der Preis seines Produktes bewegt. Machen Sie dabei auch eine Schätzung bezüglich Hardware und Implementierung.

[1] Die nachfolgenden Punkte sind an sich trivial. Betrachten Sie das Kapitel ganz einfach "pro memoria" und vielleicht finden Sie ja dennoch ein paar Tipps – oder überspringen Sie es einfach!

[2] Falls Sie die entsprechenden Seiten noch nicht kennen, finden Sie eine neutrale Übersicht unter www.dmoz.org. Auf der Hauptmaske geben Sie am besten Suchwörter wie 'Archivierung' oder 'DMS' ein.

Dies hilft Ihnen, in der nachfolgenden Offertphase die Anbieter untereinander und in Bezug auf Ihre Erwartungen noch besser zu vergleichen.

➲ *Und noch etwas, vielleicht veröffentlicht der Hersteller ja die Dokumentation oder die Software selber (OpenSource), oder Sie erhalten eine Testversion gegen ein kleines Entgelt oder gar kostenlos zum Download. Benützen Sie diese Angebote, je offener die Informationspolitik sein wird, je besser Sie das Produkt testen können, desto eher können Sie abschätzen, ob ein Produkt für Ihre Bedürfnisse geeignet ist.*

Sollten Sie beim Testen fesstellen, dass ihnen die Materie zu komplex ist, dann kann das entweder am Produkt liegen oder daran, dass die nötige Zeit fehlt. Je nach dem ob Sie entweder zu a) oder b) neigen, sollten Sie ein Produkt weiter evaluieren oder von der Wunschliste streichen.

7.4 Demo-Termin, Referenzen, Offerten und Pflichtenhefte

In der engeren Auswahl sollten nicht mehr als eine Handvoll Anbieter übrig bleiben. Falls Sie bis jetzt Ihre Hausaufgaben gemacht haben, so können Sie nun ganz gezielte Fragen zu den Produkten stellen. Ebenso lohnt sich die Installation einer Demo-Version beziehungsweise die Vereinbarung eines Präsentationstermins. In bestimmten Fällen kann auch eine Referenz hilfreich sein. Am Ende dieses Prozesses sollte es möglich sein, entsprechende Offerten einzuholen. Ob diese ein Pflichtenheft enthalten oder nicht, hängt primär von Ihnen beziehungsweise von Ihrem Projekt ab. Genauso kann es sinnvoll sein, Pflichtenheft und Offerte strikte voneinander zu trennen.

Abbildung 7.1: Ein gutes Pflichtenheft bedeutet zwar meist Aufwand, den Sie aber spätestens bei der Projektabwicklung schätzen werden, weil unliebsame Überraschungen ausbleiben.

7.5 Projektabwicklung

Sobald Sie sich für eine Lösung entschieden haben, stellt sich die Frage der Projektabwicklung. Diese gehört zwar nicht mehr zum Kern der Evaluation, dennoch seien an dieser Stelle einige Punkte erwähnt:

- Der **Zeitrahmen** sollte nicht zu eng gewählt werden.
- **Mitarbeiter/innen** möglichst früh ins Projekt einbeziehen.
- Ein **schrittweises Vorgehen** ist einfacher als eine Komplettumstellung.
- **Genügend Zeit** für Testläufe einplanen.
- **Schulung der Anwender/innen** erst durchführen, wenn das System produktiv ist.

8 Eckpunkte einer Lösung

Bei der Evaluation werden Sie sich früher oder später mit Produkten konfrontriert sehen, die Sie bewerten müssen. Technische Eckpunkte (läuft Produkt X auf System Y) oder benötigte Funktionen dürften auf den ersten Blick ersichtlich sein. Andere Punkte sind nicht so einfach fassbar.

Ein sicherer Wert, um verschiedene Produkte testen und miteinander vergleichen zu können, stellt die Methode dar, bei der Sie den **Ablauf theoretisch (oder praktisch mit einer Demo-Version) durchspielen.** Am einfachsten beginnen Sie beim Erfassen der Dokumente. Weiter geht es zum Suchen, Bearbeiten und Verwalten der Dokumente bis hin zum späteren Export. Letztlich sollten auch die Administration und das spätere "Entsorgen" der Dokumente erörtert werden.

8.1 Dokumente erfassen

Beim Erfassen von Dokumenten müssen wir zwischen zwei Arten unterscheiden: Einmal den Dokumenten, die noch nicht digital erfasst sind (Papierbelege) und weiter jenen, die bereits in irgendeiner Form elektronisch verfügbar sind, hier sprechen wir nachfolgend von Dateien und Daten.

8.2 Papierbelege

Über die Wahl eines optimalen Scanners haben wir bereits gesprochen. An dieser Stelle geht es darum, wie einfach die Belege ins Archiv gelangen. Sicher muss eine Software das rationelle Einstellen von Helligkeit, Kontrast, Farbtiefe, Auflösung, Blattabmessungen und anderen wichtigen Werten erlauben. Ebenso sollte es möglich sein, die Werte so abzulegen, dass wir die immer gleichen Werte (Vorlagen!) einfach abrufen können.

Für die professionelle digitale Archivierung ist die Geschwindigkeit beim Erfassen von Belegen essentiell: Je schneller der Scanner scannt und die Daten an den Computer übermittelt umso besser. Falls Sie eine **Texterkennung** wünschen, sollte diese **später und automatisiert** ablaufen, weil Sie ansonsten zuviel Zeit beim Erfassen der Seiten benötigen. Die auf dem Markt vorhandenen Lösungen sind hier recht unterschiedlich. Wichtig bei der Evaluation dürfte zudem sein, dass die **Lösung aus einem Guss** ist. Was bringt Ihnen der beste Scanner bzw. die tollste Software, wenn beide nicht optimal zusammenarbeiten?

8.3 Einzel-Dateien

Wie können einzelne Dateien (zum Beispiel ein Word-Dokument) archiviert werden? Hier gelten unsere Empfehlungen aus dem Abschnitt "Original- versus Archivformat". Mit einem Standard-Rasterformat wird man seine Dateien auch noch in 15 Jahren lesen können, ohne dass in der Zwischenzeit ein grosser und teurer Konvertierungsaufwand geleistet werden muss.

Grundsätzlich gilt: **Alles, was aus Applikationen heraus gedruckt werden kann, kann auch ins digitale Archiv überführt werden.** In der Regel wählen Sie dazu ganz einfach einen (virtuellen) Drucker aus, der die Datei abfängt und zur Archivlösung weiterleitet.

Selbverständlich gibt es dabei verschiedene Automationsgrade je nach Volumen und Komplexität: Reicht es, wenn alle Mitarbeiter/innen ihre jeweiligen Word-, Excel-, PDF- oder anderen Dokumente eigenhändig ins Archiv überführen? Reicht die Beschlagwortung von Hand? Oder lohnt es sich, diese automatisiert ablaufen zu lassen, indem das Archivierungssystem relevante Daten aus den Eigenschaften-Feldern herausholt? Ist Volltext nötig oder genügt die Suche über die Stichwortfelder?

➲ *Lassen Sie sich diesen Vorgang vom Anbieter der DMS- oder Archivierungs-Software erklären und vorführen.*

8.4 Office-Dateien

In den letzten Jahren ist die Auswahl an Office-Paketen (Textverarbeitung, Tabellenkalkulation und Präsentation sowie weitere Applikationen) wieder etwas grösser geworden. Derzeit verwendet die Mehrheit nach wie vor Microsoft-Produkte, aber es gibt gute Alternativen (OpenOffice und weitere). Zentral für Sie sein sollte, dass die Lösung a) ihre derzeitige Lösung unterstützt und b) ob und wie allenfalls andere Pakete (bei einem späteren Wechsel) unterstützt werden können.

➲ *Als Faustregel gilt hier, dass je stärker ein DMS-Produkt X genau auf eine Office-Applikation Y zugeschnitten ist, je schwieriger es wohl werden wird, die DMS-Software später unabhängig von der Office-Applikation zu migrieren.*

Entscheidend für eine gute DMS-Lösung ist, ob und wie die Dokumente ins Archiv gelegt werden können. Bedarf es dazu Software auf den Clients, können archivierte Dokumente ohne Plugins bzw. die Quellapplikation betrachtet werden. Ebenso sollte es möglich sein, archivierte Dokumente wieder zu reaktivieren.

8.5 E-Mails

E-Mails waren früher vielleicht Beigemüse zur Korrespondenz auf Papier. In den letzten Jahren sind sie immer mehr der Ort von geschäftskritischen Mitteilungen geworden. Deshalb müssen sie archiviert werden. Auch sie liegen der Natur der Sache gemäss bereits elektronisch vor und können grundsätzlich relativ einfach ins digitale Archiv überführt werden. Ein paar Dinge gibt es zu bedenken:

Es lohnt sich kaum, alle E-Mails zu archivieren. Zu hoch ist der Anteil von Werbemails, privaten Nachrichten, Abwesenheitsnotizen oder wenig aussagekräftigen Kürzestbotschaften. In den meisten Fällen ist eine Installation sinnvoll, die den einzelnen Benützern erlaubt, selber zu entscheiden, welche ihrer eingehenden E-Mails archiviert werden sollen. Selbstverständlich kann der Versand von E-Mails automatisch ins Archiv gelegt werden, allerdings sollte hier eine Möglichkeit vorgesehen sein, fälschlicherweise verschickte E-Mails nachträglich zu löschen (falls denn mal ein Wurm zuschlagen sollte). Auch stellt sich die **Frage des Zeitpunktes.** Der Archivvorgang kann durchaus in bestimmten Zeitintervallen (zum Beispiel täglich, wöchentlich, monatlich oder gar nur jährlich) erfolgen.

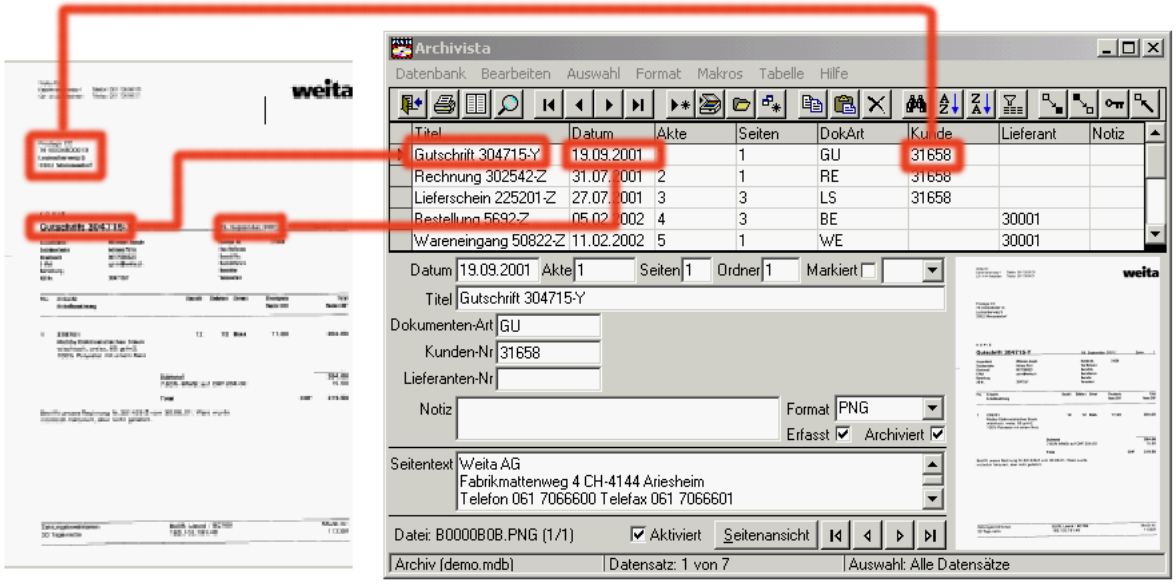
➲ *Überlegen Sie, wie wichtig E-Mails für Sie sind. Genügen die Kategorien Sender, Empfänger, Titel und Datum für das Wiederauffinden gesuchter Korrespondenz oder wünschen Sie eine Volltextsuche?*

➲ *Bei der Wahl einer Lösung sollte darauf geachtet werden, dass archivierte Mails auch wieder aktiv gestellt werden können.*

8.6 ERP-Daten

Ihr ERP-System[1] generiert Offerten, Lieferscheine, Auftragsbestätigungen und Rechnungen. Diese Belege werden ohne Archiv-Lösung meistens doppelt ausgedruckt: das Original wird verschickt, die Kopie wird abgelegt. Bei einer Umstellung auf digitale Archivierung wäre es nun nicht sinnvoll, die bisherige Papierkopie mittels Scanner wieder einzulesen. Eine Rechnung, die bereits im eigenen Computer vorhanden ist, wird besser gleich ins elektronische Archiv überführt. Das physische Ablegen in Ordner und Hängeregister kann in diesem Fall ersatzlos gestrichen werden.

[1] Enterprise resource planning system



Konkret funktioniert dies so: von allen Belegen, die am Tag verschickt werden, wird im Computer eine Kopie abgelegt, und zwar im selben Format, in dem die Rechnung oder der Lieferschein normalerweise zum physischen Drucker geht. Üblicherweise ist dies Postscript oder die weit verbreitete HP-Druckersprache PCL. Mit diesen Dateien, die in einer Art Korb warten, bis die Archiv-Applikation sie abholt, passieren in der Folge zwei Dinge:

- **Relevante Index-Informationen** wie "Kundennummer", "Kundenname", "Rechnungsnummer", "Lieferadresse" werden aus dem Druck-Spool oder direkt aus der ERP-Lösung extrahiert und in entsprechende Felder der Archiv-Datenbank abgefüllt.
- **Umwandlung der Druckdaten** in ein für die Archivierung geeignetes Format (Bild-Datei), Herausfiltern des Textes (Volltext) und Zusammenführen der Index-Informationen.

Das Verfahren heisst COLD[2] und damit alles reibungslos abläuft, muss vorher einiges an Arbeit geleistet werden: Analyse der Belege, Parametrisieren der zu extrahierenden Informationen, ausführliche Tests, damit die Belege im Computer genau so aussehen, wie wenn sie frisch aus dem Drucker kämen.

Die Machbarkeit ist praktisch immer gegeben. Weit schwieriger zu beantworten ist jedoch die Kostenfrage bei der Realisierung. Findet eine aktive Zusammenarbeit zwischen ERP- und Archivierungs-Anbieter statt, so dürften in der Regel tiefere Kosten anfallen. In der Werbung werden Archiv-Lösungen oft mit einer Vielzahl von Schnittstellen angepriesen. Die Anzahl der Schnittstellen sagt aber leider noch nichts über deren Qualität aus. Ein Referenz-Besuch bei einer vergleichbaren Systemumgebung bzw. ausführliche Tests sind weit mehr wert als Hochglanzprospekte.

[2] Computer Output on Laser Disc

➲ *Es ist möglich, ERP-System und Archivierungs-Applikation zu einem Ganzen zu verbinden, wenn im ERP-System ein "Button" eingebaut wird, der mit einem Klick vom Geschäftsvorgang direkt zum entsprechenden Beleg im digitalen Archiv hinführt. Das Potential eines solchen "Buttons" wird oft unterschätzt, bedeutet aber letztlich in der Praxis, dass der Suchvorgang im Archiv wegfällt.*

Abschliessend einige Tipps für die Anbindung einer Archivierungs- an eine ERP-Lösung:

- Der Hersteller Ihres ERP-Systems arbeitet vielleicht bereits mit einem Anbieter eines digitalen Archivs zusammen. Lassen Sie trotzdem eine Konkurrenzofferte erstellen. Wer vergleichen kann, hat die Wahl.
- Denken Sie bereits beim Einführen der ERP-Software daran, dass das Layout der produzierten Belege die spätere elektronische Archivierung erleichtern oder erschweren kann.
- Mit COLD kann ein beträchtliches Rationalisierungspotential realisiert werden. Auf der anderen Seite stehen die Kosten, die bei Anpassungen des ERP-Systems archivseitig anfallen: wird ein Formular oder Logo geändert, kann das Folgen für die Archivierung haben. Es lohnt sich abzuklären ob und wie einfach sich diese Änderungen im Archiv abbliden lassen.
- Der Aufbau eines Volltextarchivs ist einfach, da Daten bereits "in Buchstabenform" vorliegen, eine Texterkennung ist nicht nötig. Dieser Volltextindex leistet bei der Recherche wertvolle Hilfe, stellt aber keine geordnete gesetzeskonforme Ablage dar.
- Nach welchen Kriterien sollen wir Belege ablegen? Auftragsnummer? Rechnungsnummer? Lieferadresse? Es besteht die Gefahr, dass das ERP-System in der Archivierung nachgebaut wird. Empfehlung: nicht zuviele Kriterien wählen!

8.7 Workflow

Was ist Workflow im Zusammenhang mit Dokumenten-Management oder digitaler Archivierung? Vereinfacht gesagt hilft Ihnen Workflow, die Stationen, die ein Dokument durchlaufen soll, zu verfolgen. Oder mit anderen Worten: mittels **Workflow kann ein Dokument einer Reihe von Personen zugewiesen werden, die damit etwas machen müssen.**

Dazu ein Beispiel: Zu Zeiten manueller Ablage gingen eingehende Kreditorenrechnungen jeweils zum Kostenverursacher, dieser stempelte den Beleg, trug für die Buchhaltung das zu belastende Konto ein und unterzeichnete. Bei Sammelrechnungen konnte dies bedeuten, dass eine Rechnung bei vier, fünf Kostenstellenleiter/innen vorbeigeschleust werden musste, bevor der Beleg in die Buchhaltung kam.

Ein nicht immer befriedigend verlaufendes Verfahren, gerade bei Ferienabwesenheiten. Mit Workflow wird dieses Problem folgendermassen gelöst. Der Beleg (= die Kreditorenrechnung) wird zentral in eine Datenbank eingescannt, in der für jede/n Kostenstellenleiter/in Felder bestehen, für welche nur er oder sie die Berechtigung hat.

Die Kostenstellenleiter/innen erhalten nun regelmässig diejenigen Belege, die noch nicht abgezeichnet sind, für die der Buchhaltung also noch die Zahlungserlaubnis gegeben werden muss. Sie tragen im Feld "Wo verbuchen?" die Kostenstelle ein und geben auf diese Weise die Zahlung frei. Da Berechtigungen in der Datenbank auf Feldstufe vergeben werden, kann niemand anders Zahlungen freigeben.

Im Dokumenten-Management-Bereich sind ausgefeilte Workflow-Lösungen bei Krankenkassen, Versicherungen oder auch bei der öffentlichen Hand zu finden, da dort über einen relativ langen Zeitraum hinweg die Dokumente von verschiedenen Stellen (Abteilungen) eingesehen und bearbeitet werden müssen.

➲ *Im Zusammenhang mit Workflow stellt sich die Frage, ob das digitale Archiv eher für die Endablage oder die Zwischenablage der Dokumente gedacht ist und ob eher im Originalformat oder im Archivformat gespeichert werden soll. Im eingangs geschilderten Beispiel ist der Beleg für die Endablage prädestiniert. Wenn man sich einen Fall denkt, bei dem 14 verschiedene Personen in einer bestimmten Reihenfolge an einem Bericht arbeiten müssen, dann kann sogar ein reines Workflow-System sinnvoller sein; die Archivierung würde dabei erst später zum Einsatz kommen.*

8.8 Suchen von Informationen

Zunächst müssen wir zwei verschiedene Ansätze beachten. Bei der **klassischen Datenbankabfrage** können wir mit genau jenen Informationen arbeiten, die zuvor erfasst wurden, das heisst hier arbeiten wir mit strukturierten Informationen, die uns aufgrund der Beschlagwortung zur Verfügung stehen. Diese Form der Ablage muss zwingend gewählt werden, um ein gesetzeskonformes Archiv zu erhalten.

Daneben gibt es auch **Volltextabfragen,** bei denen wir das gesamte digital abgelegte Beleggut nach bestimmten Wörtern oder Satzteilen durchsuchen, daher der Name 'Volltext'. Entscheidend ist dabei, wie leistungsfähig wir die Abfragen gestalten können; wobei wir uns bewusst sein sollten, dass die Komplexität beherrschbar bleiben muss.

Ganz allgemein können wir uns folgende Fragen stellen: Wie einfach gelangen wir zu unseren Dokumenten? Können zum Beispiel mehrere Dokumente gleichzeitig betrachtet werden? Können wir den Suchvorgang nachträglich bei Bedarf erweitern oder noch enger eingrenzen? Ist

es möglich, die Treffer beliebig zu sortieren? Letztlich spielt auch die Suchgeschwindigkeit eine Rolle, wobei hier Demo-Installationen kaum aussagekräftig sind, wohl aber Referenz-Lösungen.

Im übrigen sollte es natürlich auch möglich sein, **wiederkehrende Suchabfragen abzuspeichern,** genauso wie zu den Abfragen gehört, dass niemand Dokumente betrachten kann, die nicht für ihn bestimmt sind (Stichwort Zugriffsberechtigungen).

8.9 Navigation

Die gute Navigierbarkeit ist bei der Evaluation oft ein etwas vernachlässigter Punkt. Es lohnt sich, sich folgende Fragen zu stellen beziehungsweise den Hersteller nach einer Demoversion zu fragen und die untenstehenden Punkte selber zu testen:

- Wie gut kann man vom einen zum nächsten Datensatz oder von der einen zur nächsten Seite springen?
- Wie bequem sind die Abfragen und Suchmodalitäten gestaltet?
- Wie schnell erscheint die gesuchte Seite auf dem Bildschirm? Erscheint sie direkt oder ist ein weiterer Doppelklick nötig? (zum Beispiel bei Webapplikationen)
- Wie gut leserlich ist die A4-Seite, wenn sie auf dem Bildschirm erscheint? Wie gut lässt sich zoomen?
- Wie bequem ist die Beschlagwortung? Gibt es Abkürzungsfunktionen, Pulldown-Menüs oder ähnliches?

8.10 Import- und Exportfunktionen

Wie gut lassen sich Datensätze importieren oder exportieren? Einfache **Import**funktionen schützen vor hohen Anfangskosten, wenn man vom einen Dokumenten-Management-System zum andern wechselt und das bestehende digitale Archiv auf ein neues Produkt übertragen will. **Export**funktionen sind ein Investitionsschutz für die Zukunft. Dies schon alleine deshalb, weil Sie mit leistungsfähigen Exportfunktionen gar nicht erst in eine Abhängigkeit zu einem Hersteller gelangen und die Dokumente (sollte es notwendig sein) jederzeit weiterverarbeiten können.

➲ *Abgesehen von diesen mehr strategischen Überlegungen, ist es auch wichtig, wie einfach Dokumente für den täglichen Austausch aus dem System herausgebracht werden können. Kann man die Dokumente ohne viel Aufhebens entweder im Originalformat oder im PDF an Geschäftspartner weiterschicken?*

8.11 Administration und Sicherheit

Wie einfach können Sie zum Beispiel neue Benutzer hinzufügen? Wie lange benötigen Sie für die Installation auf einem neuen Rechner? (Kriegen Sie das überhaupt hin?) Können Sie die gesamte Administration direkt über Hilfsprogramme erledigen und besteht daneben die Möglichkeit, direkt über ein Drittprogramm mit der Datenbank Kontakt aufzunehmen? Wie gut ist die Dokumentation? Welche Art von Support erhalten Sie mit welchen Kostenfolgen?

Gerade bei der Administration sollten Sie sich bewusst sein, dass eine einfach aufgebaute Lösung meist auch weniger Support-Aufwand bedeutet.

Bliebe noch die Datensicherheit, das heisst die Frage, wie einfach Sie Ihre Dokumente schützen können. Wie gut ist die Lösung gegen Drittzugriffe abgesichert (zum Beispiel Passwortschutz)? Dazu gehört ebenfalls die Fragestellung, wie einfach bzw. schwierig es ist, Dokumente unbeabsichtigt zu löschen. Letztlich sollte es aber schon möglich sein, seine Dokumente (zum Beispiel nach Ablauf der Aufbewahrungspflicht) definitiv und endgültig zu löschen.

9 Copyright-Hinweise

Copyright (c) 2009 Archivista GmbH, Urs Pfister. Dieses Dokument untersteht der Open Publication Lizenz, v1.0 (8. Juni 1999) oder später (siehe www.opencontent.org/openpub für die letzte Version).

Die Weitergabe ist ausdrücklich unter diesen Konditionen erlaubt und erwünscht. Für Rückfragen erreichen Sie uns am einfachsten unter uwucher@archivista.ch oder über die untenstehende Kontaktadresse.

Kontakt: Archivista GmbH, Zürichstr. 80, CH-8118 Pfaffhausen
Tel: +41 (0)44 254 54 00, Fax: +41 (0)44 254 54 02, Web: www.archivista.ch