

Aufgabenstellung:

Grobentwurf, Document Image Processing System

Profil: 10'000 A4-Seiten/Tag, Indexierung zur Mittagszeit abgeschlossen

a) Entwurf Systeminfrastruktur für die Posteingangsstelle und die Sachbearbeitung

Der Entwurf besteht aus dem Bereitstellen einer passenden technischen Infrastruktur und der Definition eines Workflows, der den Ablauf der Dokumentenverarbeitung innerhalb der Organisation beschreibt.

Nach Eingang der Dokumente in die Poststelle werden sie eingescannt, damit die weitere Verarbeitung papierlos geschehen kann. Dazu werden leistungsfähige „Double Sheet“ Scanner benutzt, welche die 10'000 Seiten innerhalb weniger Stunden bewältigen können. Falls die Dokumente mit einem Barcode versehen sind, kann eine spezifische Vorindexierung automatisch vorgenommen werden; z.B. bei Vorlagen zu einer Adressberichtung, einer Vertragsverlängerung oder einem Zahlungsauftrag. Dies ermöglicht eine Kategorisierung und erleichtert das Sichten der Dokumente, weil sie direkt an die verantwortlichen Sachbearbeiter weitergeleitet werden können. Ansonsten werden die Dokumente einfach als Bild digitalisiert, also ohne Unterscheidung der Semantik des Inhalts. Danach werden die Dokumente komprimiert und auf einem Server zwischengespeichert.

Von einem PC werden jetzt alle Dokumente einzeln von dem Server nach dem FIFO-Prinzip aufgerufen. Vorindexierte Dokumente können dabei gezielt von spezialisierten Sachbearbeitern ausgewählt werden. Um Zeit zu sparen, kann das Sichten wie auch das Scannen parallel an mehreren Maschinen gleichzeitig durchgeführt werden. Für jeden Arbeitsplatz kommen dafür hochauflösende Monitore zur Anwendung. Die Dokumente werden indexiert, indem zum Beispiel ein Zeitstempel und ein Zuständigkeitsbereich hinzugefügt werden. Dieser Vorgang ermöglicht auch das Erkennen von Fehlern beim Scanningprozess. Durch geeignete Software und standardisierte Dokumente, kann der Inhalt in „Coded Information“ also in einzeln codierte Text- und Formatierungszeichen transformiert werden. So kann mittels Metadaten und einem Schlüssel eine Zuordnung gemacht werden. Als Beispiel können in der Folge alle Dokumente einer Kundennummer zusammen verarbeitet werden.

Nach dem ersten Sichten und Indexieren, was eigentlich einer Vorselektion entspricht, werden die Dokumente den zuständigen Sachbearbeitern übergeben. Hier beginnt der eigentliche Workflow, der bestimmt, welche Aktivitäten in welcher Reihenfolge durchgeführt werden müssen. Diese Aktivitäten können je nach Prozessstruktur parallel oder sequentiell ausgeführt werden. In beiden Fällen kommen mehrere Sachbearbeiter zum Einsatz, die alle ihre eigene Workstation samt Monitor bedienen. Wichtig dabei ist, dass alle Geräte an einem Netzwerk angeschlossen sind, welches Zuverlässigkeit und Sicherheit gewährleistet. Erst dies ermöglicht eine Arbeitsteilung, die ohne Papier auskommt und die Effizienz steigern kann.

Eine effiziente Systeminfrastruktur benötigt eine sorgfältige Evaluation und Vorabklärungen. Bei der Hardware ist dabei bedeutend, dass sie kompatibel mit den anderen Komponenten des Systems ist; proprietäre Lösungen sind zu vermeiden, damit die Geräte unabhängig voneinander ersetzt werden können. Die Bedürfnisse der Benutzer dürfen ebenfalls nicht ausser Acht gelassen werden, da sie die Schnittstelle zum System bilden. Das Management muss Zeitpläne und Arbeitsabläufe so planen, dass sie mit der natürlichen Arbeitsweise der Mitarbeiter übereinstimmen und nicht als

störend empfunden werden. Gleichzeitig sind die Kernkompetenzen der Unternehmung miteinzubeziehen. Ein Unternehmen, das Kundenfreundlichkeit als Ziel vorgibt, muss Anfragen und Aufträge von Kunden so schnell wie möglich bearbeiten. Folglich sollte ein vorrangiger Prozess für solche Fälle definiert sein. Im Allgemeinen kann ein maximaler „Return on Investment“ nur dann generiert werden, wenn die beabsichtigte Nutzung der Hardware gut mit realistischen Zeitplänen übereinstimmt

b) Spezifikation des Schwebe und Archivspeichers

Das Ziel einer elektronischen Dokumentenverarbeitung, ist das Eliminieren der „hardcopy“ Exemplare. Um den ganzen Workflow elektronisch zu gestalten sind mehrere Formen von Speicher notwendig.

Die eingescannten Dokumente werden direkt vom Scanner zu einem Server weitergeleitet. Von diesem Schwebespeicher können sie jetzt von einer Workstation gesichtet, indexiert und danach wieder auf den Server geladen werden. Dies ermöglicht eine zentrale Datenhaltung. Die Daten sind so für alle Sachbearbeitern zugänglich und erlauben die Arbeitsteilung optimal zu gestalten, weil nach jedem Arbeitsschritt eine andere Person auf den Server zugreifen kann. Durch ein gutes Konsistenzmanagement kann zudem sichergestellt werden, dass nie zwei Leute gleichzeitig am selben Dokument arbeiten.

Nach der Bearbeitung eines Falles ist es meist notwendig, die involvierten Dokumente zu komprimieren und zu archivieren. Beispiele für die Notwendigkeit einer Archivierung sind rechtliche Vorschriften, als Beweismittel und für das Controlling. Der Schwebe- und Archivspeicher sind wie die anderen Geräte an das LAN angeschlossen. Insbesondere beim Archivspeicher ist zu beachten, dass nur berechtigte Personen Zugriff haben und vor allem auch dass niemand von aussen eindringen kann. Die operativen Daten einer Unternehmung können je nach Branche einen sehr hohen strategischen Wert haben und müssen sorgfältig geschützt werden.

Für die Auswahl der konkreten Speichermedien müssen verschiedenen Evaluationskriterien beachtet werden:

- Geschwindigkeit (lese / schreibe, Datentransfer)
- Kapazität
- Zuverlässigkeit (Stabilität, Redunanz)
- Kosten

Übereinstimmend mit diesen Kriterien kann man für den Schwebespeicher schnelle Festplatten wählen, deren Kapazität flexibel wählbar ist. Um die Sicherheit zu erhöhen, können die Daten parallel auf zwei Platten gespeichert werden, die sich immer abgleichen („mirroring“). Beim Archivspeicher ist die Geschwindigkeit weniger wichtig. Es sollte vor allem auf eine hohe Kapazität zu tiefen Kosten und eine lange Beständigkeit geachtet werden. Hier sind Magnetbänder oder Zip-Disketten zu erwähnen.

Es können also verschiedene Technologien für die Speicherung der Daten eingesetzt werden. Das heisst jedoch nicht, dass die Architektur regelmässig den neuesten Trends angepasst werden soll. Eine bewährte Technologie gewährleistet den besten Support. Zudem sollte im Hinblick auf die schnell sinkenden Preise und die Kosteneffektivität keine Kapazität auf Vorrat angeschafft werden.