

Compliance zum (Fast-) Null-Tarif

Dokumentenmanagement mit Open Source

Die Menge aufbewahrungspflichtiger Dokumente wächst ständig. Weithin bekannt und beachtet ist die Aufbewahrungspflicht für Geschäftsbücher. Daneben entstehen jedoch ständig neue Regularien. Diese betreffen die gesamte Geschäftskorrespondenz einschließlich des E-Mail-Verkehrs und weitere Geschäftsdokumente wie z.B. Dokumentationen zu Produkten oder Projekten, die zur Sicherstellung von Wartung und Pflege aufbewahrt werden müssen. Die klassische Aufbewahrung in Papierform erfüllt immer seltener die regulatorischen Auflagen und macht eine gezielte Suche mit steigendem Volumen nahezu unmöglich, erst recht dann nicht, wenn noch Referenzen auf weitere Dokumente ermittelt werden müssen.

Dokumentenmanagement - die Lösung (?)

In vielen Fällen könnte ein Dokumentenmanagement helfen. Viele Unternehmen, große wie auch mittlere und kleinere, scheuen jedoch die Einführung einer solchen Lösung. Die Gründe dafür sind vielfältig:

- Befürchtungen, dass langjährig praktizierte Arbeitsabläufe verändert werden müssen,
- der Initialaufwand, um den vorhandenen Dokumentenbestand zu bereinigen und aufzunehmen,
- Mehraufwand im Betrieb,
- Anschaffungs- und Wartungskosten der Lösung.

Betrachtet man jedoch die möglichen Einsparungen durch Reduktion von Fehlerquellen, kürzere Suchzeiten, leichtere Erfüllung rechtlicher Auflagen zur Aufbewahrung, dann stellt man schnell fest, dass sich Dokumentenmanagement rechnen. Hinzu kommt, dass es inzwischen leistungsfähige Dokumentenmanagementsysteme im Open Source Bereich gibt, womit zunächst schon einmal die Anschaffungskosten entfallen.

Metadaten - Kernstück des Dokumentenmanagements

Was sind Metadaten? Metadaten beschreiben den Inhalt eines Dokuments. Nun kann man sagen, dass das zumindest bei Textdokumenten doch eigentlich überflüssig ist. Schließlich gibt es die Volltextrecherche. Das würde aber bedingen, dass alle relevanten Begriffe immer in der gleichen Weise geschrieben werden und unterschiedliche Beugungsformen immer zum selben Stichwort führen. Außerdem müssten auch Synonyme berücksichtigt werden.

Manchmal findet man aber das richtige Stichwort gar nicht, weil der Sachverhalt umschrieben wurde. Seit es Textsuchverfahren gibt, ist bekannt, dass selbst bei klaren Vorgaben zur Schreibweise immer wieder Abweichungen auftreten, die ein Auffinden erschweren oder verhindern, sei es aufgrund von Tippfehlern oder weil das entsprechende Dokument von

außerhalb der Organisation kam und damit die eigenen Regeln nicht berücksichtigt.

Metadaten sind aber noch mehr als die Liste der möglichen Suchwörter in einem Dokument oder der Stichwörter zu einem Bild oder Media-Clip. Selten gibt ein Dokument Auskunft darüber, wie lange es aufzubewahren ist, oder unter welchen Bedingungen seine Gültigkeit überprüft werden muss. Versionierung und Bearbeitungsdatum, wie sie vom Betriebssystem oder der entsprechenden Anwendung gesetzt werden, sind in keiner Weise zuverlässig.

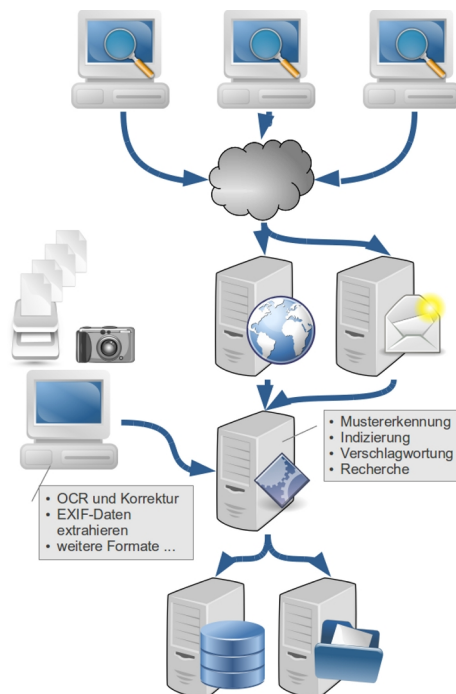
Durch Speicherung von Stichworten, einem kurzen beschreibenden Text und Kenndaten wie Ersteller, Erstellungsdatum, Kontext, Aufbewahrungsfristen und Gültigkeitsvermerken kann gezielt nach solchen Zusatzinformationen gesucht werden, sowohl bei Textdokumenten wie auch bei so genanntem Rich Content, also Bildern, Tonaufnahmen oder Video-Clips. Metadaten eröffnen die Möglichkeit, den gesamten Lebenszyklus eines Dokuments zu verwalten.

Dokumentenformate und Metadaten

Während Office-Formate gleich welcher Office-Suite grundsätzlich die Erfassung von Metadaten erlauben, auch wenn davon selten Gebrauch gemacht wird, ist die Integration von Metadaten in anderen

Formaten zumindest nicht immer direkt möglich. Zwar gibt es für diverse weitere Formate Editoren zum Bearbeiten der Metadaten wie z.B. der ID3-Tags von Audiodateien oder der EXIF-Daten von Bilddateien, aber nicht alle Datenformate unterstützen solche Metadaten und die Bearbeitung setzt oft zusätzliche Programme voraus.

Zudem werden oft bei der Bearbeitung je nach verwendeter Software Teile dieser Metadaten wieder gelöscht oder überschrieben, so dass die ursprünglichen Informationen verloren gehen. Eine praktische Alternative in diesen Fällen ist die Verwendung einer so genannten SideCar-Datei, die die Metadaten enthält und die bei der Ablage in einem Dokumentenmanagementsystem ausgewertet wird. Allerdings birgt eine solche Datei die Gefahr, dass sie nicht mit aktualisiert wird und so ggf. einen veralteten Stand reflektiert.



Prinzipielle Architektur einer DMS-Lösung¹



Unternehmensvokabular – Basis guter Metadaten

Prinzipiell geht man davon aus, dass Wörter die selten vorkommen eine höhere Relevanz haben, als solche die sehr häufig auftreten. Bei der automatischen Generierung von Suchindices aus Texten werden deshalb alle Wörter, die für die Suche irrelevant sind, die so genannten Stoppwörter, ausgeschlossen. Im jeweiligen Kontext können aber auch branchen- oder fachspezifische Termini oder andere Zeichenkombinationen, die im sprachlichen Sinn keine Wörter sind, relevante Suchbegriffe darstellen. Um diesem Umstand Rechnung zu tragen, sollte ein unternehmensweit einheitliches Vokabular zusammen gestellt werden.

Vorlagen-Management

Ein weiterer wichtiger Baustein für die Einführung eines Dokumentenmanagements sind gute Vorlagen für alle Dokumententypen im Unternehmen. Die Anwendung dieser Vorlagen muss geschult und bei Aktualisierungen aufgefrischt werden.

Grundsätzlich sollte jeder Benutzer von Office-Anwendungen mit der Anwendung von Dokumenten- und Formatvorlagen vertraut sein. Leider sieht man immer wieder Dokumente die keine Formatvorlagen nutzen. Hinzu kommt das manche Office-Suiten für jede direkte Formatierung ein neues Absatz oder Zeichenformat erstellen, so dass ein Dokument, das mehrfach zwischen verschiedenen Anwendern hin und her gesendet wurde, am Ende ein unüberschaubare Liste von so genannten akzidenziellen Formaten enthält.

Von Metadaten zur Dokumentenverwaltung

Wenn diese grundlegenden Regeln und Prozesse definiert sind, ist es nur noch ein kleiner Schritt zur Einführung einer Dokumentenverwaltung. Dazu stehen verschiedene kommerzielle oder Open Source Lösungen zur Verfügung. Die meisten davon sind Web gestützt und können bei entsprechender Konfiguration auch von mobilen Anwendern oder dem Home-Office aus genutzt werden.

Auch eine mehr oder weniger direkte Einbettung in Office-Programme ist in den meisten Fällen möglich und sollte als Auswahlkriterium berücksichtigt werden. Eine gute Integration in die Office-Umgebung ermöglicht es, die Arbeitsabläufe nahezu unverändert zu lassen. Dies erhöht in aller Regel die Akzeptanz und ermöglicht so einen zeitnahen Nutzen der Investition.

Altbestände integrieren

Wenn Dokumentenmanagement eingeführt wird, startet man im allgemeinen nicht auf der grünen Wiese, sondern hat

einen umfangreichen Altbestand an Dokumenten, der noch weiter gepflegt werden muss. In vielen Fällen fehlen dazu die Metadaten. Nun ist es kaum wirtschaftlich, alle Dokumente von Hand nachträglich mit passenden Metadaten zu versehen. In diesem Fall kann man versuchen, aus dem Kontext, in dem die Dokumente verwendet werden, und den Inhalten passende Metadaten zu generieren. Hat man einmal ein Unternehmensvokabular geschaffen, kann man gezielt nach solchen Begriffen in den Dokumenten suchen. Dies funktioniert mit allen Dokumenten, die Texte enthalten, auch wenn diese zunächst durch Extraktion oder OCR-Bearbeitung wieder nutzbar gemacht werden müssen.

Hilfreich sind auch Sammlungen von unternehmensweiten Mustern z.B. für interne Standards. Durch Mustersuche können auch abweichende Schreibweisen erkannt werden und so kann man anhand der so erkennbaren Begriffe die Dokumente oftmals schon bestimmten Themenbereichen zuordnen.

Parallel dazu sollte man das ausgewählte Dokumentenmanagementsystem von Beginn an in die Arbeitsumgebung integrieren, um den „Altbestand“ nicht noch weiter wachsen zu lassen.

Kosten und Nutzen

Die Gesamtkosten einer DMS-Lösung werden im Wesentlichen von den laufenden Betriebskosten bestimmt. Bei richtiger Einführung ist der Betrieb eines DMS jedoch nicht aufwändiger als der Betrieb einer klassischen Datenablage auf Dateiservern. Bei größeren Dokumentenbeständen ist zusätzlich die Anschaffung eines Speichersystems sinnvoll auf dem die Datenbank und je nach gewähltem System die Dateiverzeichnisse eingerichtet werden. Ein solches System bietet höhere Kapazitäten, eine bessere Ausfallsicherheit und verbessert damit die Zuverlässigkeit der Lösung. Die auf der Vorderseite abgebildeten Server kann man dann in vielen Fällen auf einer Hardware virtualisieren.

Mit der Einführung eines Dokumentenmanagements hat man einige Vorteile wie

- Nachvollziehbarkeit
- Versionierung
- schnelle Recherche
- Kontrolle der Aufbewahrungsfristen
- (halb-)automatische Bereinigung des Datenbestands

Damit lassen sich in der Regel alle rechtlichen Auflagen erfüllen. Bestehende Lösungen zeigen, dass bei planmäßigem Betrieb und ordnungsgemäßer Dokumentation eine Abnahme durch einen externen Audit kein ernsthaftes Problem darstellt.

ZPK Moser UG (haftungsbeschränkt)
Schlossstraße 7
D-78244 Gottmadingen

+49 (0)7734 395 494
info@zpk-moser.de
www.zpk-moser.de



Ulrich Moser ist Geschäftsführer der ZPK Moser UG (haftungsbeschränkt) und als Berater, Trainer und Coach tätig. Zuvor war er tätig als Professor für Informationsmanagement an der Berner Fachhochschule und als Senior Consultant für Informationsmanagement, Informationssicherheit und Geschäftsprozesse.

