

Frank Linde, Ina van Kann

Ordnung in der Dokumentenflut schaffen

Wie oft passiert dies im Arbeitsalltag eines Unternehmens: Ein Mitarbeiter sucht ein bestimmtes Dokument. Er weiß, dass es irgendwo im Computernetzwerk gespeichert sein muss, findet es aber nicht. Möglicherweise hat er es sogar selbst irgendwann abgelegt, doch er kennt den Dateinamen nicht mehr. Die Ordnerbereiche, die er sich als Ablageplatz vorstellen könnte, hat er bereits überprüft. Er überlegt, einfach seinen Kollegen zu fragen, ob dieser vielleicht eine Idee hat, wo er noch suchen könnte...

Mit den wachsenden Möglichkeiten zum Erstellen und Weiterverarbeiten von Dokumenten steigt auch die Menge der elektronisch vorliegenden Informationen in Unternehmen an. Häufig werden große Bestände immer noch über die herkömmlichen Ordnerstrukturen des Betriebssystems verwaltet, doch diese können den Anforderungen im Arbeitsalltag kaum gerecht werden. Meist handelt es sich um ein kontinuierlich gewachsenes Ordnungssystem, welches die abgedeckten Sachverhalte nicht eindeutig voneinander abgrenzt und sich nur selten durch eine übergeordnet nachvollziehbare Struktur auszeichnet. Dadurch kommt es zu Unsicherheiten bezüglich der richtigen Zuordnung von Informationen, denn Dokumente können inhaltlich häufig aus mehreren Blickwinkeln betrachtet und daher verschiedenen Themen zugeordnet werden. Dies führt dazu, dass zusammengehörige Dokumente teilweise an völlig unterschiedlichen Stellen gespeichert werden. Neben diesen Faktoren erschweren auch mangelnde Recherchemöglichkeiten den zielgerichteten Zugriff. Wenn bereits vorhandene Dokumente nicht gefunden werden können, müssen

sie neu erstellt werden – was die Unübersichtlichkeit und Inkonsistenzen im Dokumentenbestand weiter erhöht.

Ordnungsprinzipien für einen besseren Informationszugang

Für die erfolgreiche Nutzung jeder Dokumentenmenge ist es notwendig, die einzelnen Dokumente über festgelegte Kriterien wie beispielsweise eine Kundennummer oder Schlagworte zugänglich zu machen. Eine sinnvolle Ordnung verhilft dazu, sie unter zweckmäßigen Merkmalen zusammenzuführen. Dadurch werden die erforderlichen Maßnahmen für den Zugriff auf einzelne Dokumente, z. B. bei der Recherche über Schlagworte, klar vorhersehbar. Entsprechend muss dann bei der Suche nach einer Information nur noch eine überschaubare Teilmenge des Dokumentenbestands auf deren Vorhandensein überprüft werden.

Im Bibliotheksbereich haben sich verschiedene Ordnungsprinzipien entwickelt, mit deren Hilfe die Zugänglichkeit von Informationen verbessert wird:

- Die Klassifikation teilt bestehende Sachverhalte in Klassen ein, so dass jedem Dokument ein Ablageplatz im System zugewiesen werden kann. So wären beim Thema „Sport“ z. B. Klassen für „Wassersport“ oder „Ball sport“ denkbar. In diesen Klassen sollen inhaltlich ähnliche Dokumente zusammengeführt werden. Es treten jedoch immer dann Probleme auf, wenn keine eindeutige Zuordnung vorgenommen werden kann: Gehört z. B. ein Dokument über Wasserball zu Wassersport oder zu Ball sport? In-

Kurz gefasst:

- Dokumentenmanagement-Systeme (DMS) setzen unterschiedliche Ordnungsprinzipien ein, um den Zugang zu den in einem Dokumentenbestand enthaltenen Informationen zu eröffnen: die Erschließung über Klassen, Kategorien oder die Deskriptoren eines Thesaurus sowie die Volltextindexierung.
- Grundsätzlich sind zwei Gruppen von DMS zu unterscheiden: DMS, welche die konventionellen Ablageverfahren aus der gegenständlichen Welt abbilden, sowie DMS, welche sich an den Inhalten der Dokumente orientieren.
- Welches der verschiedenen Verfahren das geeignete ist, hängt von den Eigenschaften des Dokumentenbestandes und vom überwiegenden Anwendungszusammenhang ab.

nerhalb der Klassifikation lässt sich diese Schwierigkeit nicht beheben.

- Erst wenn mehrere Zuordnungen möglich sind, tritt diese Problematik nicht mehr auf, so bei dem Verfahren der Erschließung über Kategorien. Hier wird die Gesamtinformation über ein Dokument zerlegt und es werden verschiedene Merkmale in getrennten Kategorienfeldern wie z. B. „Autor“, „Erstellungsdatum“, „Schlagworte“ etc. aufgenommen.
- Ein weiteres Ordnungsprinzip mit mehrfacher Zuordnungsmöglichkeit ist der Thesaurus. Dieser enthält normiertes Vokabular, so genannte Deskriptoren, welche zur inhaltlichen Beschreibung der Dokumente verwendet werden können, und macht damit Begriffsbeziehungen deutlich. So würde z. B. der Wassersport als Unterbegriff zu Sport gekennzeichnet werden.
- Neben diesen Ordnungsprinzipien existiert die Volltextindexierung.

Diese macht alle in Dokumenten enthaltenen Wörter suchbar und stellt sie für die Volltextsuche zur Verfügung. So wird zwar keine Ordnung erzeugt, das Verfahren ist aber dennoch zur Aufbereitung einer Dokumentenmenge für die Recherche geeignet.

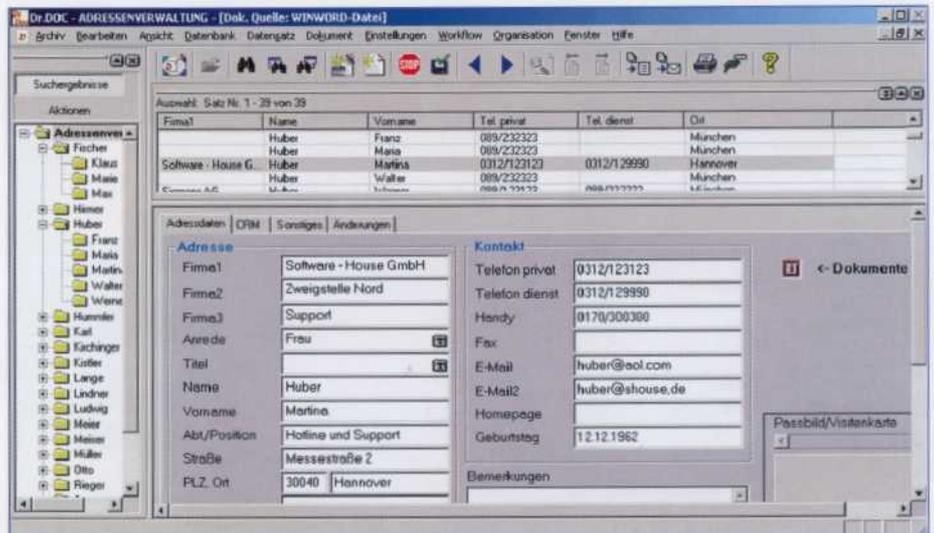
Die genannten Ordnungsprinzipien werden in unterschiedlichen Anwendungszusammenhängen eingesetzt und stellen auch einen essenziellen Bestandteil von Dokumentenmanagement-Systemen (DMS) dar. Erst ihre Verwendung ermöglicht einen angemessenen Zugriff auf die im DMS abgelegten Informationen. Grundsätzlich werden in diesem Zusammenhang zwei Gruppen von DMS unterschieden:

1. DMS, die eher konventionelle Ablageverfahren aus der gegenständlichen Welt auf die elektronische Umgebung abbilden.
2. DMS, die sich an den Dokumenteninhalten orientieren und auf dieser Grundlage Gruppierungen vornehmen.

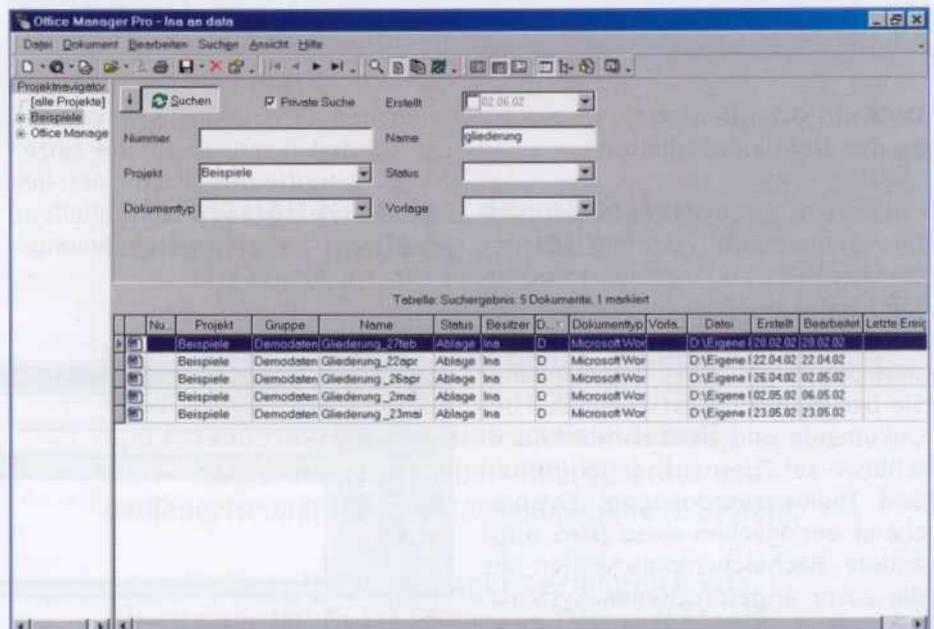
DMS mit konventionellen Ablageverfahren

Bei gegenständlichen Ablageverfahren, beispielsweise beim Einsortieren von Rechnungen in einen Ordner, wird ein Ordnungskriterium, etwa die Rechnungsnummer, bestimmt. Der spätere Zugriff kann ausschließlich über dieses Kriterium vorgenommen werden; ein Zugang über andere Merkmale wird nicht unterstützt. Um weitere Zugriffsmöglichkeiten zu schaffen, müssten zusätzliche Ordner angelegt werden, in denen nach anderen Kriterien abgelegt wird. Ein solches Verfahren ist also von einer Eindimensionalität geprägt, die nur durch Redundanzen behoben werden kann.

Bei der elektronischen Ablage in einem DMS wird diese Problematik gemildert, indem jedes Dokument nur einmal gespeichert, aber mit mehreren Merkmalen verknüpft und hierüber suchbar gemacht wird. Bei der Recherche stellt die jeweilige



Eingabemaske eines Adressenarchivs bei Dr. Doc mit angezeigtem Datensatz

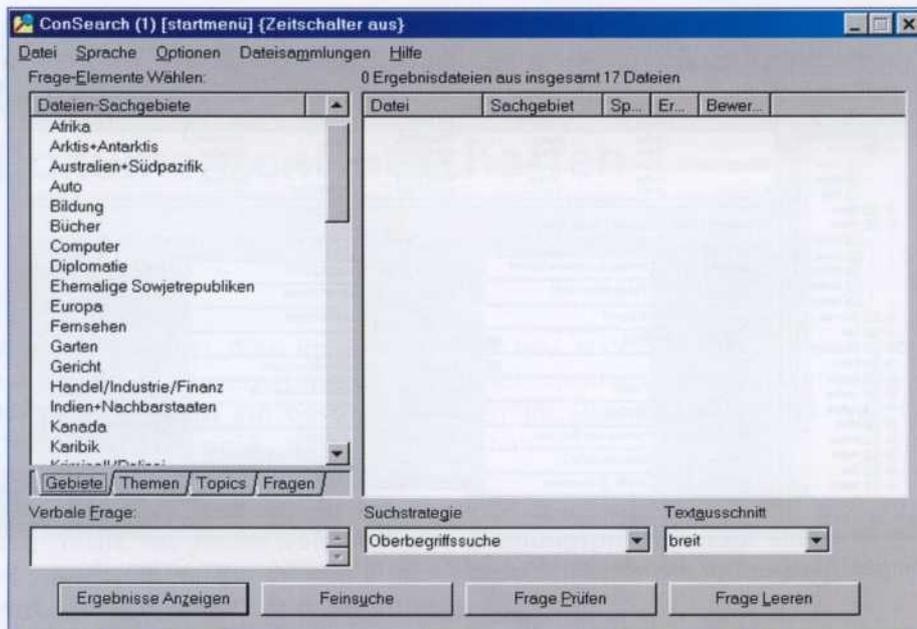


Suchmaske des Office Manager Pro mit Suche „Name: Gliederung“

Software Treffermengen zu den verschiedensten Frageformulierungen zusammen. Ein solches Erschließen der Dokumente über vorgegebene Kategorien erweitert also die Möglichkeiten der physischen Ablage bedeutend, dennoch handelt es sich um ein Prinzip, welches von der gegenständlichen Welt einfach auf die elektronische Umgebung übertragen wird.

Mit einem solchen konventionellen Verfahren des Verfügbarmachens von Informationen arbeiten beispielsweise die Dokumentenmanagement-Systeme Dr.Doc und Office Manager Pro.

Es können formale Merkmale (z.B. Dokumenttyp) und inhaltliche Merkmale (z.B. Schlagworte) einzelner Dokumente in Kategorien unterteilt über eine Erfassungsmaske aufgenommen werden. Für die spätere Suche nach einem Dokument steht eine entsprechende Suchmaske zur Verfügung, mit der über die zuvor eingegebenen Eigenschaften der Zugriff auf die abgelegten Dokumente ermöglicht wird. So wird im Gegensatz zur gegenständlichen Ablage ohne redundante Lagerung ein Zugreifen über zahlreiche Kriterien unterstützt.



Benutzeroberfläche von ConSearch mit enthaltenen Sachgebieten

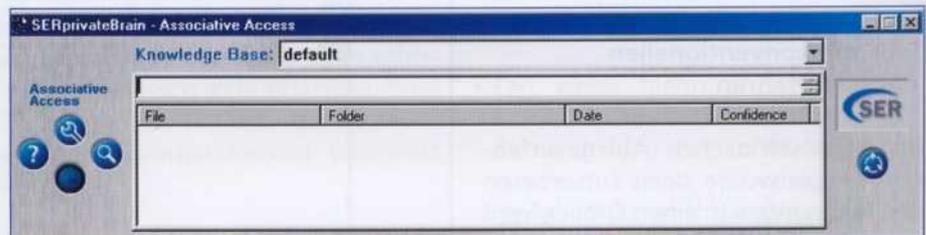
DMS mit Orientierung an den Dokumentinhalten

Die zweite Gruppe von Dokumentenmanagement-Systemen arbeitet weniger mit den einzelnen formalen Merkmalen der Dokumente, wie z.B. Autor oder Datum, sondern organisiert den Zugriff über deren Inhalte. Sie berücksichtigt das Vokabular der Dokumente und zieht darüber Rückschlüsse auf Zusammengehörigkeiten und Themenzuordnungen. Entsprechend ermöglichen diese DMS auch andere Recherchemöglichkeiten als die zuvor angesprochenen Systeme. Die Softwarehersteller gehen in diesem Bereich sehr unterschiedliche Wege, so dass eine allgemeine Vergleichbarkeit nicht wie bei den Systemen des ersten Typs gegeben ist. Beispiele für derartige DMS stellen ConSearch und SERprivateBrain dar.

Bei ConSearch werden dem Anwender verschiedene Zugriffsmöglichkeiten geboten: Neben eigenen verbalen Suchformulierungen stehen weitere Möglichkeiten für Modifikationen wie die Einteilung in Sachgebiete, Themen und Topics sowie Hilfsfragen zur Verfügung.

SERprivateBrain besitzt eine sehr schlichte Suchmaske. Der Anwender kann auswählen, welche Dokumen-

tenmenge die Grundlage für eine Suche darstellen soll, und dann einzelne Suchbegriffe oder auch einen Beispieltext eingeben. Anschließend wird eine Trefferliste zusammengestellt und angezeigt.



Suchmaske von SERprivateBrain

Welches Verfahren ist das Richtige?

Bereits an diesen wenigen Softwarebeispielen lässt sich zeigen, dass in Dokumentenmanagement-Systemen unterschiedliche Kombinationen von Ordnungsprinzipien eingesetzt werden:

- Dr. Doc basiert auf der Beschreibung durch Kategorien und bietet ergänzend eine Volltextindexierung. Zwar stellt das DMS keinen Thesaurus zur Verfügung; es besteht jedoch die Möglichkeit, einen solchen selbst zu erstellen und mit Hilfe der Software zu nutzen.

- Office Manager Pro arbeitet gleichfalls mit Kategorien und bietet die Möglichkeit, einen Thesaurus selbst zu erstellen.
- Das an den Dokumenteninhalten orientierte DMS ConSearch basiert hingegen auf dem Ordnungsprinzip der Klassifikation. Zusätzlich bietet es eine Volltextsuche. Einen Thesaurus nutzt das System nur intern.
- SERprivateBrain arbeitet ebenfalls auf Basis von Klassifikationen. Mit dem so genannten Associative Access stellt die Software zudem eine der Volltextindexierung vergleichbare Funktionalität bereit.

Der Wert kombinierter Ordnungsprinzipien in Dokumentenmanagement-Systemen zeigt sich daran, dass die Suchmöglichkeiten nach vorhandenen Dokumenten entscheidend erweitert und so die Eindimensionalität gegenständlicher Verfahren aufgehoben werden. Beide Gruppen von DMS, sowohl die Systeme mit eher konventionellen Ablageverfahren als auch diejenigen mit einer Orientierung an Dokumenteninhalten, besit-

zen je nach Anwendungszusammenhang und Bedarf der Anwender bestimmte Vorteile:

- Möglicherweise ist immer nur der Zugriff über ein bestimmtes Kriterium, z.B. über den Ersteller eines Dokumentes, erforderlich. Dann sind die von der gegenständlichen Ablage übernommenen Verfahren völlig ausreichend.
- Falls jedoch häufig unterschiedliche Zusammenstellungen einander ähnlicher Dokumente benötigt werden, empfiehlt sich der Einsatz von Systemen, die mit den Dokumenteninhalten arbeiten.

- Für den Fall, dass bei der Recherche ausreichende und möglichst eindeutige formale oder inhaltliche Merkmale der Dokumente, z.B. eine Kundennummer, bekannt sind, ist die erstgenannte Gruppe von DMS sinnvoll.
- Stehen derartige Informationen für Suchprozesse nie oder nur selten zur Verfügung, ist hingegen der Einsatz der zweiten Gruppe von DMS ratsam.
- Die Volltextindexierung ist immer dann eine sinnvolle Ergänzung der anderen Ordnungsprinzipien, wenn die einzelnen Dokumente hinsichtlich ihres enthaltenen Vokabulars sehr verschieden sind.

Weiterhin beeinflussen die Eigenschaften einer Dokumentenmenge maßgeblich, welches Suchverfahren für einen bestimmten Anwendungszusammenhang geeignet ist. Je nachdem, ob die Merkmale der Dokumente erschlossen sind oder nicht, lassen sich für die Suche folgende Rückschlüsse ziehen:

- Die Suche über Klassen sowie die Suche über die Deskriptoren eines Thesaurus setzen eine inhaltliche Erschließung der Dokumentenmerkmale voraus.
- Die Suche über festgelegte Kategorien ist sowohl bei formaler als auch bei inhaltlicher Dokumentenerschließung einsetzbar.
- Lediglich die Volltextsuche kann auch bei einer nicht erschlossenen Dokumentenmenge zum Einsatz kommen.

Fazit:

Optimales Dokumentenmanagement heißt also einerseits, Ordnungsprinzipien so einzusetzen, dass sowohl die Gegebenheiten der Dokumentenmenge als auch die daraus resultierenden Möglichkeiten voll ausgeschöpft werden. Ebenso wichtig ist es, die Bedürfnisse der Anwender, die sich in deren Suchanforderungen äußern, angemessen zu berücksichtigen. Wenn dies alles beim Aufbau einer Ordnung für Dokumente umgesetzt wird, bleiben ergebnislose Rechercheprozesse wie der zu Anfang beschriebene die Ausnahme.

Die Autoren:



Prof. Dr. Frank Linde ist Dozent für Wirtschaftswissenschaften am Institut für Informationswissenschaft der Fachhochschule Köln. Nach mehrjähriger Tätigkeit in verschiedenen Funktionen im Konzern Deutsche Telekom AG lehrt und forscht er an der Fachhochschule Köln unter anderem zum Thema Wissensmanagement.

linde@wissensmanagement.net



Ina van Kann studierte Bibliothekswesen an der Fachhochschule Köln mit dem Spezialisierungsgebiet Auskunftsdienst und Informationsvermittlung. Derzeit ist sie freiberuflich tätig und befasst sich mit dem Aufbau und der Pflege von Kundendatenbanken.

kann@wissensmanagement.net

Information Retrieval der nächsten Generation - made by empolis

Explizite Ontologien und intelligentes Retrieval bieten eine Vielzahl von Funktionen, die man bei anderen Systemen vergeblich sucht. Enrichment und Erläuterung von Inhalten sowie automatisierte Nutzerdialoge ergeben ein effizientes und bedienerfreundliches Wissensmanagement-System.

empolis knowledge manager - erfolgreich im Einsatz u.a. bei Fujitsu Siemens, Otto und Merck.

Vereinbaren Sie einen Termin für Ihr kostenloses Beratungsgespräch und Technologie-Demo vor Ort - wir sind für Sie da!

Gerhard.Koehn@empolis.com
Telefon: +49(9365)8062 201

empolis - Transforming Information into Value
www.empolis.de

