

Dokumentenmanagement, Archivierungs- und Verschlüsselungsverfahren^[1]

Dozent: Prof. Dr. Michael Eichberg
Kontakt: michael.eichberg@dhbw.de, Raum 149B
Version: 1.0.1

Folien: [HTML] <https://delors.github.io/dm-einfuehrung/folien.de.rst.html>
[PDF] <https://delors.github.io/dm-einfuehrung/folien.de.rst.html.pdf>
Fehler melden: <https://github.com/Delors/delors.github.io/issues>

[1] Dieser Foliensatz basiert auf Folien von: Klaus Götzer.
Dokumenten-Management von *Klaus Götzer, Patrick Maué, und Ulrich Emmert*,
dpunkt.verlag, 2023.
Alle Fehler sind meine eigenen.

Dokumentmanagement, Archivierungs- und Verschlüsselungsverfahren - W3WI_EH304.1

Verortung: Drittes Studienjahr (5. oder 6. Semester)

Umfang: 25 Vorlesungsstunden und 50 Stunden Selbststudium (2,5 ECTS)

Übergeordnetes Modul:

Fortgeschrittene Konzepte des Informationsmanagements
im Gesundheitswesen (W3WI_EH304)

Prüfungsleistung: Portfoliogespräch

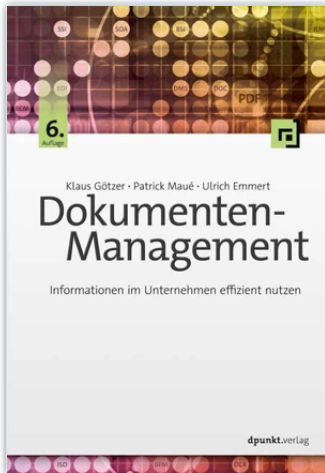
(d. h. mündl. Prüfungsgespräche - voraussichtlich 10min pro Person)

Kerninhalte gem. MHB

- Definitionen, Grundsätze, Aufgaben, Funktionen und Prozesse von Dokumentenmanagement
- Archivierungssysteme
- Struktur elektronischer Verwaltungsunterlagen, Archivierungsvarianten der Software- und Hardware-Lösungen
- Rechtliche, technische, organisatorische und wirtschaftliche Anforderungen
- Schnittstellen
- Aktuelle Standards
- Verfügbarkeit elektronischer Verwaltungsakten
- IT-Sicherheit von digital erzeugten und gescannten Dokumenten
- Elektronische Signaturen
- Archivierungsverfahren
- Verschlüsselungsverfahren

Literatur

■ Bzgl. Dokumentenmanagement:



■ weitere bzgl. der IT sicherheitsrelevanten Aspekte: Verschlüsselungsverfahren und Signaturen

1. Warum Dokumentenmanagement?

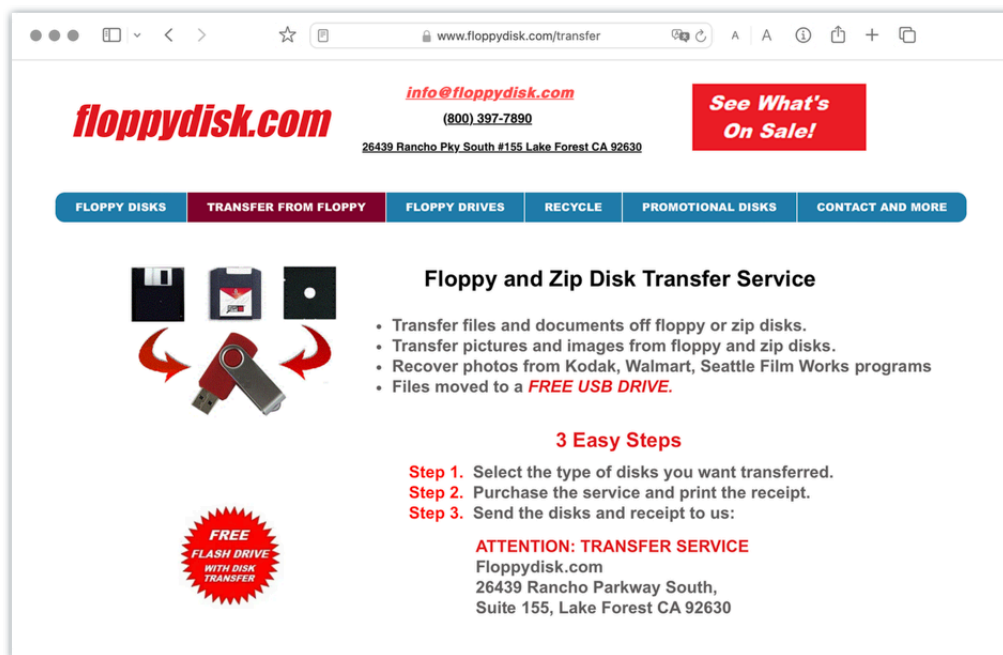
Motivation

- Die meisten strukturierten Daten liegen elektronisch vor (ERP, CRM, etc.).
- Dokumente liegen aber (noch immer) in Papierform vor (insbesondere in Deutschland) und es gilt dieses Rationalisierungspotential zu heben.
- Alle reden von elektronischen Geschäftsprozessen und Digitalisierung - Dokumente sind ein (der?) wesentlicher Bestandteil.
- Wichtige gesetzliche und technische _Grundlagen_ sind geschaffen.
- Experten schätzen, dass 95% der Papierdokumente nach Ablage nicht mehr genutzt werden, da sie zu schwer zu finden sind.
Ähnlich hohe Werte werden für digitale Dokumente genannt, die ohne weitere Strukturierungshilfen in Dateisystemen oder Datenbanken liegen.
- Die Menge an Dokumenten, die heute entstehen kann durch eine einzelne Person nicht mehr bewältigt werden.

Beispiele für die schleichende Digitalisierung

- an der DHBW wurden erst zum Jahrgang 2024/25 die Schriftform bei (einigen) der Dokumenten zu Anmeldung abgeschafft.
- An der DHBW war es bis Ende 2024 noch erforderlich Projektarbeiten und Bachelorarbeiten in Papierform einzureichen, obwohl dies von vielen Dozenten schon lange nicht mehr verlangt wurde. Für PA I arbeiten gilt dies auch im Jahr 2025 noch.
- Für ein einzelnes konventionelles Kraftwerk fallen bereits ca. 500 000 Dokumente.

Digitalisierung - ein langwieriger Prozess...



Jan. 2024

Verwendung von Disketten im Jahr 2024

Bundeswehrschiffe mit Diskettenlaufwerken: Marine sucht Ersatz für Floppy Disks

In kleinen Schritten schreitet auch die Bundeswehr bei der Digitalisierung voran: Jetzt will sie ihre Fregatten mit moderner Technik ausstatten – und verabschiedet sich dafür vom technischen Standard der Neunzigerjahre: der Diskette.

—19.07.2024 **RND**

Digitalisierung - ein langwieriger
nicht-aufzuhaltender Prozess...

Japanese government finally bids sayonara to the 3.5" floppy disk

Businesses can at long last submit digital docs to government agencies

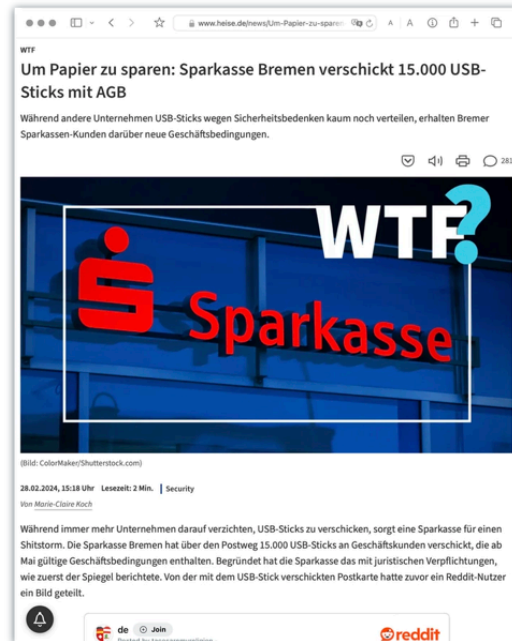
Japan is saying sayonara to the floppy disk, which until now was a required medium for submitting some 1,900 official documents to the government.

The announcement (Japanese, machine translated) last week from the Ministry of Economy, Trade and Industry brings decades of physical media submission requirements in Japan to an end. [...]

*Despite being a world leader in cutting-edge technology, Japan has an odd relationship with legacy tech. It's still a land of **cash-only payments and fax machines** that has moved slowly to embrace the modern digital economy.*

—The Register, Mon 29 Jan 2024 // 19:00 UTC

Digitalisierung - ein Prozess mit Stilblüten...



Zumindest in der Übergangszeit - aber nicht nur dann - kann es vorkommen, dass die Digitalisierung zu einem deutlichen Mehraufwand führt, da beide Systeme (Papier und Digital) gepflegt werden, weil die Nutzer dem neuen System noch nicht vertrauen oder es noch nicht vollständig implementiert ist und die echten Anforderungen (noch) nicht erfüllt.

Wie die zahlreichste Bibliothek, wenn ungeordnet, nicht so viel Nutzen schafft, als eine sehr mäßige, sorgfältig geordnete; eben so ist die größte Menge von Kenntnissen, die nicht gehörig durch eigenes Denken bearbeitet werden, viel weniger Wert als eine viel geringere Menge von Kenntnissen, die gehörig durchdacht werden.

—Arthur Schopenhauer (mutmaßlich)

Was kann ein Dokumentenmanagement System (DMS) leisten?

■ Beschleunigung der Prozesse

- Unabhängigkeit des Zugriffs von Ort und Zeit
- Schnelle Verfügbarkeit der Dokumente
- Gleichzeitiger Zugriff auf die Dokumente durch mehrere Mitarbeiter

■ Revisionssicherheit der Ablage

- Strukturierte Ablage und Suche von Dokumenten
- Sicherheit der Dokumente vor Verfälschung und Verlust
- Transparenz der Prozesse

■ Redundanzfreie Archivierung

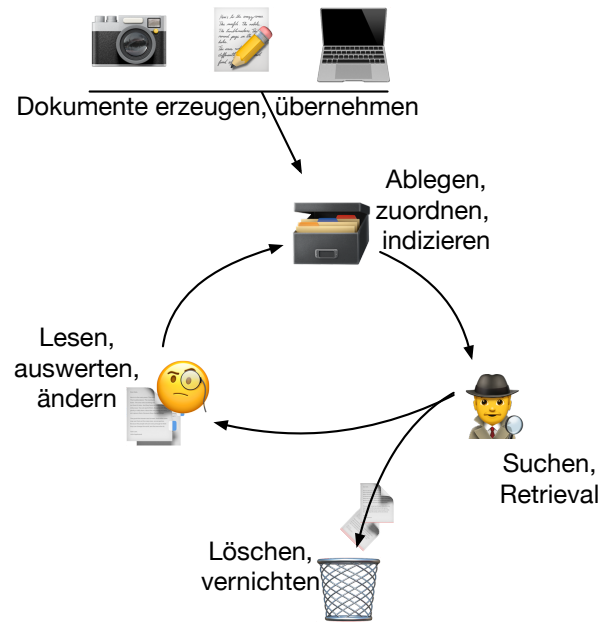
■ Kostenreduktion (Bearbeitungszeiten, Archivkosten....)

Revisionssichere Archivsysteme stellen sicher, dass Informationen wieder auffindbar, nachvollziehbar, unveränderbar und verfälschungssicher archiviert sind.

Gegenüberstellung konventionelles Archiv zu DMS

	Konventionell	DMS
Ablagestruktur	hierarchisch, meist nach Dokumententypen getrennt	datenbankgestützt, Suchbeginn nach jedem Suchkriterium möglich, Suche wahlweise z. B. je Kunde, Vorgang oder Beleg
Auskunftsbereitschaft	zeitverzögert, nach Sichtung bzw. Entnahme aller Dokumente	unmittelbar durch vorgangsbezogene bzw. dokumentenübergreifende Suche
Auskunftsaufwand	für Belegsuche, Entnahme und wieder einordnen, evtl. kopieren und Versand	nur für Recherche am Bildschirm, gegebenenfalls direkter Versand per E-Mail
Redundanz	Abteilungsablagen = mehrfacher Aufwand	Einmalablage
Vollständigkeit	leidet unter jeder Entnahme	bleibt immer gewahrt
Sicherheit	Ordner u. Dokumente lassen sich leicht entfernen oder kopieren	klare Regelung der Zugriffsrechte, keine ungewollte Entnahme möglich

Zyklus von Dokumenten



Grundlegende Voraussetzungen für Dokumentenmanagementorganisation

In einer Organisation ist zu regeln:

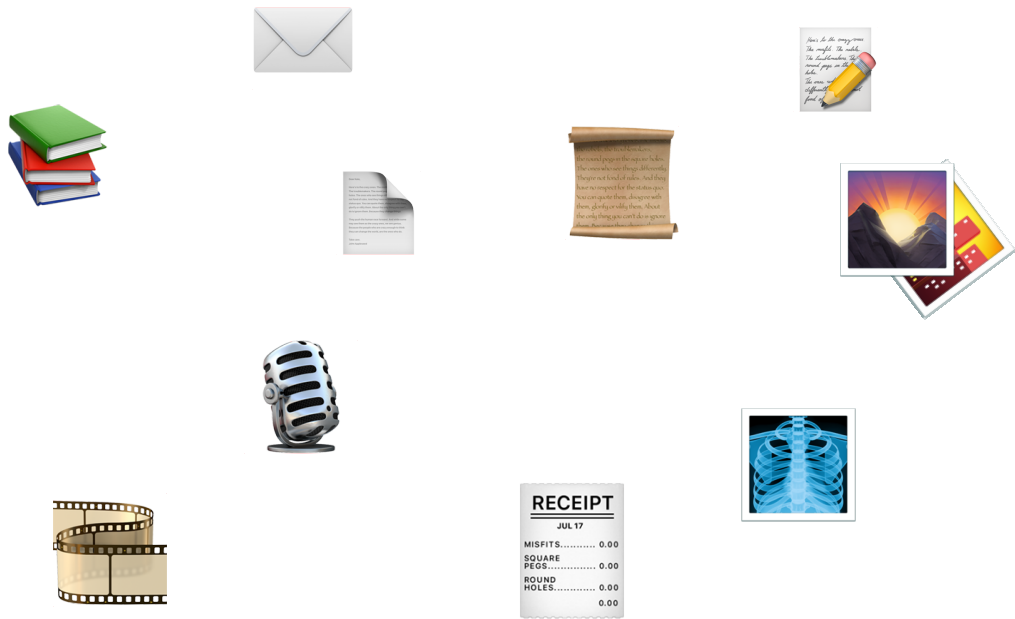
- Kennzeichnung und Beschreibung von Dokumenten
- Fortschreibung und Historienverwaltung von Dokumenten
- Ablage und Archivierung von Dokumenten
- Verteilung und Umlauf von Dokumenten
- Suche nach Dokumenten bzw. Dokumenteninhalten
- Vernichtung von Dokumenten
- Regelung von Verantwortlichkeiten für Inhalt und Verwaltung von Dokumenten

Bemerkung

Ein Dokumentenmanagementsystem ist ein IT System zur Unterstützung der Dokumentenmanagementorganisation.

2. Dokument und Dokumentenmanagement

Dokumente - Beispiele



Was ist ein Dokument aus logischer Sicht?

Definition

Ein Dokument fasst inhaltlich zusammengehörende Informationen strukturiert zusammen, die nicht ohne erheblichen Bedeutungsverlust weiter unterteilt werden könnten.

- Die Information ist für einen gewissen Zeitraum zu erhalten.
- Dokumente dienen dem Nachweis von Tatsachen.

Bemerkung

Das Dokument ist somit eigentlich der Träger, der die Informationen speichert, unabhängig davon ob das Dokument ein Stück Papier, eine Datei auf einem Rechner, ein Videoband oder eine Tontafel ist etc.

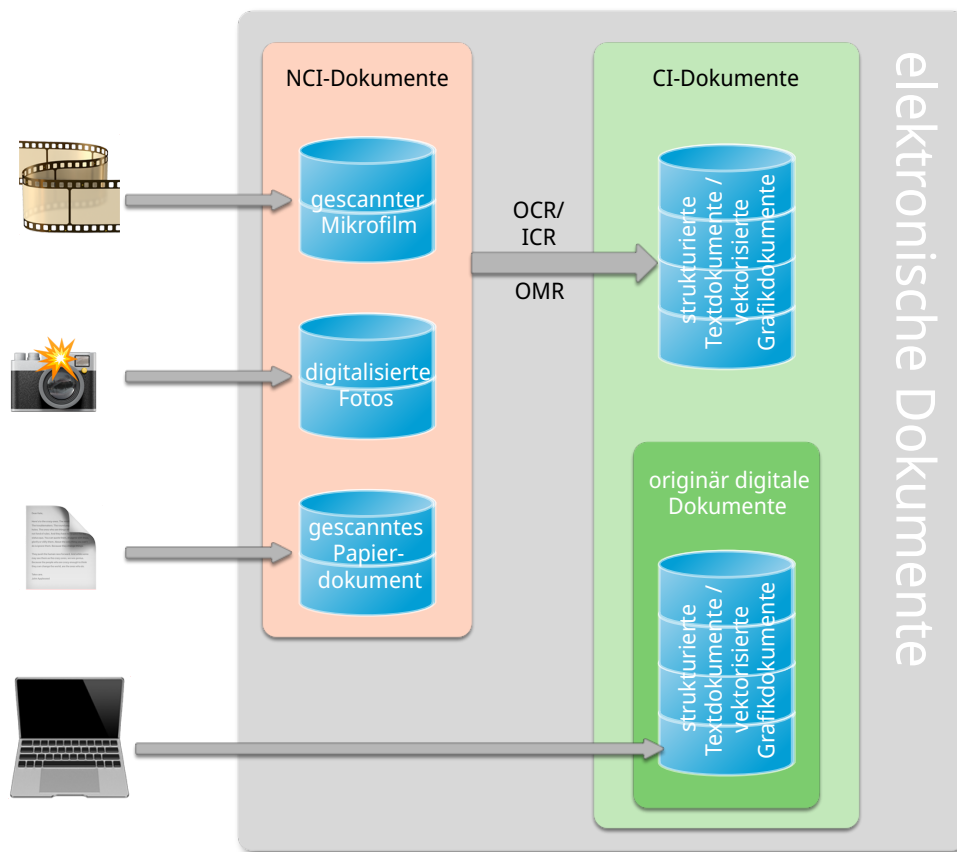
Was ist ein Dokument aus technischer Sicht?

Definition

Ein Dokument ist ein Objekt, das in einer Datenbank beschrieben wird.

Das beschriebene Objekt kann selbst elektronisch gespeichert werden.

Es ist somit eine reine Deklarationsfrage, was ein Dokument ist!



NCI: *Non-Coded-Information (NCI)-Dokumente* sind eingescannte Unterlagen, die als Bild vorliegen, also keine direkte Bearbeitung/Verarbeitung ermöglichen.

CI: Ein *CI-Dokument* ist ein digital erstelltes Dokument, das durch Zeichensätze kodiert ist und von Programmen direkt ausgewertet werden kann.

OCR (Optical Character Recognition): Text einer gedruckten Vorlage wird durch einfachen Mustervergleich automatisch in maschinenlesbare Zeichen transformiert.

ICR (Intelligent Character Recognition): Die Qualität der Texterkennung wird durch Kontextanalyse verbessert. Typische Fehler von OCR-Systemen, wie zum Beispiel Fehlerkennungen von optisch nahe beieinanderliegenden Zeichen (z. B. „8 und B“ oder „0 und O“), werden vermieden. (Wie nahe Zeichen beieinander liegen, ist stark vom verwendeten Schrifttyp abhängig.)

OMR (Optical Mark Recognition): Liest mit großer Sicherheit spezielle Markierungen in vordefinierten Feldern aus - zum Beispiel in Multiple-Choice-Tests und Vordrucken.

Bestandteile eines Dokumentes

Das "Dokument" →

Eindeutige Dokumenten-ID →

Metadaten des Dokuments:
- Erstellungsdatum
- Ersteller/Unterzeichner

Stempel →

Eheurkunde
Standesamt Otzberg
Registernummer Niederschrift über die Eheschließung
E 80/2010

Eheschließung
Ort, Tag Otzberg, 08.09.2010

Ehemann
Familienname Eichberg
Geburtsname
Vorname(n) Michael
Geburtsdag, Geburtsort
Religion
Familienname nach Eheschließung
Geburtsname nach Eheschließung


Ehefrau
Familienname
Geburtsname
Vorname(n)
Geburtsdag, Geburtsort
Religion
Familienname nach Eheschließung
Geburtsname nach Eheschließung

Ort, Tag Otzberg, 08.09.2010

Urkundsperson
(Ohlemüller, Standesbeamter)

12/1110 Verlag für Standesamtswesen GmbH, Frankfurt am Main · Berlin 2009

Struktur eines Dokumentes



DHBW
Duale Hochschule
Baden-Württemberg
Präsidium

Duale Hochschule Baden-Württemberg - Präsidium
Präsidentenstr. 14, 70174 Stuttgart

Herrn
Dr. Michael Eichberg

Prof. Dr. Martina Kärle
Präsidentin
DHBW
Präsidentenstr. 14
70174 Stuttgart
Berufsstelle:
Frau Kärle
Telefon + 49 711 32 06 60-19
Telefax + 49 711 32 06 60-65
www.karle@dhbw.de
www.dhbw.de
Altkanzler
DHBW 13.2.2017

Berufung auf eine W2-Professur Wirtschaftsinformatik an der Dualen Hochschule Baden-Württemberg Mannheim,
Verfahrensnummer: 2021-MA-W-4

11.07.2023

Sehr geehrter Herr Dr. Eichberg,

Im November 2022 haben Sie sich auf die genannte Professur beworben. Wir dürfen Ihnen die erfreuliche Mitteilung machen, dass Sie das Präsidium der Dualen Hochschule Baden-Württemberg auf den ersten Listenplatz gesetzt hat.


Wir berufen Sie deshalb auf diese Professur.

Die Einstellung soll im Rahmen einer Versetzung und Ernennung zum Professor an der Dualen Hochschule Baden-Württemberg im Beamtenverhältnis auf Lebenszeit erfolgen. Ihre Bezüge werden sich nach der Besoldungsgruppe W2 richten. Bitte beachten Sie, dass sich die DHBW für vier Wochen an diesen Ruf gebunden hält. Sie sollten daher innerhalb dieser Frist entscheiden, ob Sie den Ruf annehmen wollen. Aus Gründen der Personalplanung würden wir es jedoch begrüßen, wenn Sie sich möglichst frühzeitig bei uns melden.


Wegen des Einstellungsverfahrens wird sich die DHBW Mannheim mit Ihnen in Verbindung setzen. Für weitere Fragen im Zusammenhang mit dieser Berufung steht Ihnen Herr Prof. Dr. Georg Nagler als Rektor der DHBW Mannheim gerne zur Verfügung (Tel.: 0621 4105 1500; E-Mail: georg.nagler@dhbw-mannheim.de).

Wir würden uns freuen, Sie in Kürze als Professor an unserer Hochschule begrüßen zu dürfen und wünschen Ihnen hierfür bereits viel Erfolg und gutes Gelingen.

Mit freundlichen Grüßen



Prof. Dr. Martina Kärle
Präsidentin



Prof. Dr. Georg Nagler
Rektor der DHBW Mannheim

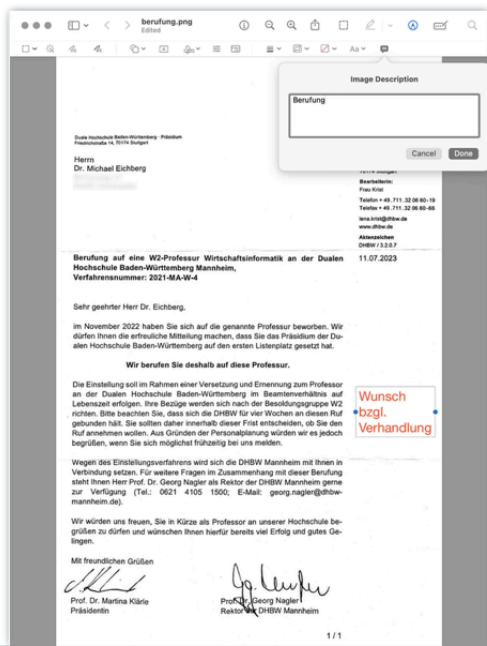
1 / 1

Kopf

Rumpf

Fuß

Dokumente annotieren



Renditions eines Dokumentes

- Man unterscheidet zwischen dem Originalformat des Dokuments (z. B. von MS-Word oder LibreOffice) und Renditions (wie PDF/a und TIFF).
- Formate wie PDF/a und TIFF sind in der Regel langlebiger. Für beide Formate gibt es weitverbreitete Viewer und sie erhalten besser den ursprünglichen optischen Zustand.
- Preview Images sind oft zusätzlich eingebettet.

? Frage

Was beeinflusst den ursprüngliche optische Zustand?

Wo sind die Probleme beim Erhalt des ursprünglichen optischen Zustandes?

Hashwerte und Signaturen [2]

- Ein Hashwert ist wie ein (eindeutige) mathematischer Fingerabdruck des Dokumentes.
 - Ein (kryptografischer) Hashwert erlaubt keinen Rückschluss auf die Inhalte des Dokuments.
 - Die kleinste Änderung am Dokument führt zu einem (völlig) anderen Hashwert.
 - Dieser Hashwert – verknüpft mit einer persönlichen Signatur – zeigt, dass dieses Dokument von dem Absender der Signatur stammt und das Dokument nicht verändert wurde.
-

[2] Hashwerte und Signaturen sind Gegenstand der Vorlesungen im Bereich IT Security.

Volltext(-indizierung)

- Für Volltextindizierung wird oft der Volltext - insbesondere wenn er durch OCR/ICR gewonnen wurde - mit dem Dokument abgespeichert.
- Die indizierten Begriffe werden oft mit der Angabe der Fundstelle im Dokument abgespeichert, um innerhalb des Dokuments das Suchergebnis anzeigen zu können.

Versionierung von Dokumenten

- Was ist die aktuelle gültige Version?
- Was hat sich gegenüber den Vorgängern geändert?
- Was ist für die nächste in Bearbeitung?
 - Vorgängerversion(en)
 - Freigegebene Version
 - Bearbeitungsversionen
- Versionen des Dokumentes
- Versionen der Metadaten des Dokumentes

Meta-Daten

- Strukturierte Daten, die das Dokument klassifizieren und beschreiben

Beispiele:

- Eindeutige Schlüssel wie Personalnr., Produktnr., ...
- Stichwörter zum Klassifizieren des Textes
- Datum der Erstellung, Änderung, ...
- Autor
- Kategorien wie Mahnung, Anfrage, Branche, Land, ...
- Quelle des Dokuments (Zeitschrift...)

Dies ist insbesondere ein Thema der Datenmodellierung, d. h. welche Daten möchte man wie erfassen.

Arten von Meta-Daten

- Eindeutiger Schlüssel im DMS
- Fremdschlüssel (z. B. Buchungsnummern)
- Statische Metadaten (unveränderlich)
- Dynamische Metadaten (wie Status oder Version der Dokumente)

Beispiele von Meta-Daten einer Verwaltungssoftware für grafische Zeichnungen

Nr.	Attribut	Muss	Funktion	Quelle	Bemerkung
1	Zeichnungsnummer	M	Eindeutiger Schlüssel	Manuelle Vergabe durch Benutzer	Identifiziert Zeichnung
2	Zeichnungsmappennummer	M	Fremdschlüssel		
3a	Version	M	Version der Zeichnung verwalten	Automatische Vergabe durch DMS bei Check-in	Benutzer entscheidet, ob <i>minor</i> oder <i>major</i>
3b	Check-In-Datum	M	Datum des Check-in der Version	Automatische Vergabe durch DMS	Check-in Datum
3c	Dokumenten-Owner	M	Gruppe aus letzten Bearbeitern	Aus USER-ID abgeleitet	
3d	Letzter Bearbeiter	M	Identifikation	USER-ID	Beim Check-in

Zusammenfassung: Dokumente in einem DMS

Ein Dokument in einem DMS ist ein komplexes Objekt, das aus verschiedenen Komponenten bestehen kann:

- Das Dokument im Originalformat (z. B. odt, docx, xlsx, txt, ...)
- Verschiedene Renditions (pdf, tiff, xml,)
- Vorschaubild
- Volltext
- Annotationen (Layer für Anmerkungen, Stempel, ...)
- Hashwert, um elektronische Signaturen zu erzeugen und/oder zu prüfen
- Elektronische Signaturen
- Versionen des Dokumentes
- Metadaten des Dokumentes bzw. der Komponenten des Dokumentes

Dokumentenstrukturen

- Welche Dokumente bilden eine logische Einheit („Mappen“, „Ordner“, „Vorgang“)?
- Metadaten zu diesen Mappen definieren.
- Ein Dokument kann in mehreren Mappen sein.
- Der Inhalt einer Mappe unterteilt sich in:
 1. Dokumente, die immer da sein müssen,
 2. solche, die optional da sind und
 3. in nicht vorhersehbare Exoten.

Ein Beispiel einer Mappe wäre eine Vorgangsakte mit:

- einem Antrag,
- ggf. E-Mails
- ggf. Unteraufträgen
- ggf. handschriftlichen Notizen
- (am Ende) einem Gutachten

Zusammengesetzte Dokumente

Komplexes Objekt aus mehreren Dokumenten mit eigener Verwaltungsstruktur:

- Metadaten
- Versionen
- Rechte

Zusammengesetzte Dokumente \cong  *Compound Documents*

3. Dokumentenlebenszyklus

Dokumentenlebenszyklus - Überblick



Dokumentenlebenszyklus

Dokumente ...

- entstehen,
- verändern sich,
- werden festgeschrieben,
- dienen als Nachweis / Infoquelle,
- müssen bestimmte Zeit aufbewahrt werden,
- können bzw. müssen gelöscht werden.

Erstellen von Dokumenten

- Scannen analoger Dokumente (Papier, Mikrofilm, ..)
- Neuerstellung von Dokumenten (Vorlagen im DMS,..)
- Vorhandene Dokumente einstellen (*drag and drop*)
- Dokumente aus Applikationen übernehmen (SAP-Archive-Link, Mail-Archivierung, ...)
- Spezielle Verfahren bei Migration und Massenimporten
- Indizieren der Dokumente entweder automatisch oder manuell

Nutzen und Bearbeiten von Dokumenten

- Suchen und Retrieval:

- Volltext

- Indizes

- Verknüpfungen (z. B. in Applikationen)

- Ausgabe der Dokumente auf Bildschirm, Drucker, Mail

- Check-out / bearbeiten / Check-in

Rahmenbedingungen für die Lebensdauer von Dokumenten

- Betriebliche Notwendigkeiten
- Gesetzliche Aufbewahrungs- bzw. Löschungspflichten
- Datenschutzbestimmungen

-
- Konzept zur intelligenten Verwaltung, Bewertung und Nutzung von Daten bei möglichst geringen Kosten (Geschäftsregeln, Servicelevel, ..)
 - Betrachtung der Daten und Dokumente über ihren gesamten Lebenslauf, aber nicht in Abhängigkeit vom Alter sondern von der Wichtigkeit
 - Entwicklung optimaler Verwaltungsstrategien in Abhängigkeit von der aktuellen Wichtigkeit und Nutzung
 - Enge Verzahnung von Speicherhardware, Archivierung und Daten-, Dokumenten- und Content-Management

Löschen von Dokumenten

- Falsche Dokumente (z. B. Fehler beim Indizieren)
- Nicht mehr benötigte Dokumente
- Logisches Löschen
- Physikalisches Löschen

Archivierung (von Dokumenten)

Definition

Archiv: Ein realer oder elektronischer Ort, in dem Dokumente, die zur laufenden Aufgabenerfüllung nicht mehr benötigt werden, erfasst, ausgewertet und zugänglich gemacht werden.

Kerneigenschaften elektronischer Archive

Archivsysteme müssen die folgenden Eigenschaften bei langjähriger Aufbewahrung unterstützen:

Integrität:	Nachweis, dass die Informationen während der Aufbewahrung nicht verändert oder gelöscht wurden.
Authenzität:	Herkunft lässt sich zweifelsfrei nachweisen.
Lesbarkeit:	Informationen müssen auch in Zukunft lesbar sein.



Beobachtung

Eigenschaften wie die Unterstützung von Volltextindizierung sind nicht in allen Fällen notwendig.

Wikipedia: **Authentizität** bedeutet Echtheit im Sinne von Ursprünglichkeit.

Z. B. müssen Privatpersonen Handwerkerrechnungen für (mind.) 2 Jahre aufbewahren.

Z. B. muss ein Insolvenzverwalter nach dem Abschluss des Verfahrens die Dokumente 10 Jahre aufbewahren obwohl er in dieser Zeit die Dokumente höchstwahrscheinlich nicht mehr benötigt.

Informationen über die Ergebnisse von arbeitsmedizinischen Untersuchungen (dies sind ggf. sehr vielfältige und umfangreiche Daten (z. B. Röntgenbilder, EKGs)) müssen 30 Jahre bis nach der letzten Untersuchung aufbewahrt werden, da sich daraus ggf. Rentenansprüche ableiten. Daraus leiten sich ggf. Fristen für Dokumente von 50 Jahren und mehr ab.

Revisionssichere elektronische Archivierung

Merkmale revisionssicherer Archivsysteme sind:

- Informationen sind wieder auffindbar,
- Änderungen sind nachvollziehbar,
- Informationen sind unveränderbar und
- die Informationen sind verfälschungssicher archiviert

Aufgrund von HGB-Vorschriften gilt für die Revisionssicherheit bei der Archivierung:

- Richtigkeit
- Vollständigkeit
- Sicherheit des Gesamtverfahrens
- Schutz vor Veränderung und Verfälschung
- Sicherung vor Verlust
- Nutzung nur durch Berechtigte
- Einhaltung der Aufbewahrungsfristen
- Dokumentation des Verfahrens
- Nachvollziehbarkeit
- Prüfbarkeit

Strategien für die langfristige Aufbewahrung von Dokumenten

- *Technikmuseum* (Variante: Alte Umgebung emulieren)
- Dauerhafte Formate nutzen
- Migration der Dokumente auf neue Umgebung

Bewertung von Dateiformaten in Hinblick auf die Dauerhaftigkeit

- vollständige und offene Dokumentation (am besten mit Standardisierung)
- Plattformunabhängigkeit
- nicht-proprietär (herstellerunabhängig)
- keine „verlustbehaftete“ oder proprietäre Komprimierung
- keine eingebetteten Dateien, Programme oder Skripte
- keine vollständige oder teilweise Verschlüsselung
- kein Passwortschutz
- relevante Nutzerbasis

Langfristige Aufbewahrung von Dokumenten[3]

Content Type	High probability for long-term preservation	Medium probability for long-term preservation	Low probability for long-term preservation
Text	<ul style="list-style-type: none"> ■ Plain text (encoding: USASCII, UTF-8, UTF-16 with BOM) ■ XML (includes XSD/XSL/XHTML, etc.; with included or accessible schema) ■ PDF/A-1 (ISO 19005-1) (*.pdf) 	<ul style="list-style-type: none"> ■ Cascading Style Sheets (*.css) ■ DTD (*.dtd) ■ Plain text (ISO 8859-1 encoding) ■ PDF (*.pdf) (embedded fonts) ■ Rich Text Format (*.rtf) ■ HTML (include a DOCTYPE declaration) ■ SGML (*.sgml) ■ Open Office (*.sxw/*.odt) ■ OOXML (ISO/IEC DIS 29500) (*.docx) 	<ul style="list-style-type: none"> ■ PDF (*.pdf) (encrypted) ■ Microsoft Word (*.doc) ■ WordPerfect (*.wpd) ■ All other text formats not listed here
Content Type	High probability for long-term preservation	Medium probability for long-term preservation	Low probability for long-term preservation
Raster Image	<ul style="list-style-type: none"> ■ TIFF (uncompressed) ■ JPEG2000 (lossless) (*.jp2) ■ PNG (*.png) 	<ul style="list-style-type: none"> ■ BMP (*.bmp) ■ JPEG/JFIF (*.jpg) ■ JPEG2000 (lossy) (*.jp2) ■ TIFF (compressed) ■ GIF (*.gif) ■ Digital Negative DNG (*.dng) 	<ul style="list-style-type: none"> ■ MrSID (*.sid) ■ TIFF (in Planar format) ■ FlashPix (*.fpx) ■ PhotoShop (*.psd) ■ RAW ■ JPEG 2000 Part 2 (*.jpf, *.jpx) ■ All other raster image formats not listed here
Content Type	High probability for long-term preservation	Medium probability for long-term preservation	Low probability for long-term preservation
	<ul style="list-style-type: none"> ■ SVG (no Java script 	<ul style="list-style-type: none"> ■ Computer Graphic 	<ul style="list-style-type: none"> ■ Encapsulated Postscript (EPS) ■ Macromedia

Vector Graphics	binding) (*.svg)	Metafile (CGM, WebCGM) (*.cgm)	Flash (*.swf) ■ All other vector image formats not listed here
-----------------	------------------	--------------------------------	---

Content Type	High probability for long-term preservation	Medium probability for long-term preservation	Low probability for long-term preservation
Audio	<ul style="list-style-type: none"> ■ AIFF (PCM) (*.aif, *.aiff) ■ WAV (PCM) (*.wav) 	<ul style="list-style-type: none"> ■ SUN Audio (uncompressed) (*.au) ■ Standard MIDI (*.mid, *.midi) ■ Ogg Vorbis (*.ogg) ■ Free Lossless Audio Codec (*.flac) ■ Advance Audio Coding (*.mp4, *.m4a, *.aac) ■ MP3 (MPEG-1/2, Layer 3) (*.mp3) 	<ul style="list-style-type: none"> ■ AIFC (compressed) (*.aifc) ■ NeXT SND (*.snd) ■ RealNetworks 'Real Audio' (*.ra, *.rm, *.ram) ■ Windows Media Audio (*.wma) ■ Protected AAC (*.m4p) ■ WAV (compressed) (*.wav) ■ All other audio formats not listed here

Content Type	High probability for long-term preservation	Medium probability for long-term preservation	Low probability for long-term preservation
Video	<ul style="list-style-type: none"> ■ Motion JPEG 2000 (ISO/IEC 15444-4) (*.mj2) ■ AVI (uncompressed, motion JPEG) (*.avi) ■ QuickTime Movie (uncompressed, motion JPEG) (*.mov) 	<ul style="list-style-type: none"> ■ Ogg Theora (*.ogg) ■ MPEG-1, MPEG-2 (*.mpg, *.mpeg, wrapped in AVI, MOV) ■ MPEG-4 (H.263, H.264) (*.mp4, wrapped in AVI, MOV) 	<ul style="list-style-type: none"> ■ AVI (others) (*.avi) ■ QuickTime Movie (others) (*.mov) ■ RealNetworks 'Real Video' (*.rv) ■ Windows Media Video (*.wmv) ■ All other video formats not listed here

Content Type	High probability for long-term preservation	Medium probability for long-term preservation	Low probability for long-term preservation
Spreadsheet / Database	<ul style="list-style-type: none"> ■ Comma Separated Values (*.csv) ■ Delimited Text (*.txt) 	<ul style="list-style-type: none"> ■ DBF (*.dbf) ■ OpenOffice (*.sxc/*.ods) ■ OOXML (ISO/IEC 	<ul style="list-style-type: none"> ■ Excel (*.xls) ■ All other spreadsheet/ database formats

	■ SQL DDL	DIS 29500) (*.xlsx)	not listed here
Content Type	High probability for long-term preservation	Medium probability for long-term preservation	Low probability for long-term preservation
Virtual Reality	■ X3D (*.x3d)	■ VRML (*.wrl, *.vrmf) ■ U3D (Universal 3D file format)	■ All other virtual reality formats not listed here
Content Type	High probability for long-term preservation	Medium probability for long-term preservation	Low probability for long-term preservation
Computer Programs	■ Computer program source code, uncompiled (*.c, *.c++, *.java, *.js, *.jsp, *.php, *.pl, etc.)		■ Compiled / Executable files (EXE, *.class, COM, DLL, BIN, DRV, OVL, SYS, PIF)
Content Type	High probability for long-term preservation	Medium probability for long-term preservation	Low probability for long-term preservation
Presentation		■ OpenOffice (*.sxi/*.odp) ■ OOXML (ISO/IEC DIS 29500) (*.pptx)	■ PowerPoint (*.ppt) ■ All other presentation formats not listed here

[3] Recommended File Formats for Long-Term Data Curation - Georgia Southern University | University Libraries

Dateiformate und ihre Eignung für die Langzeitarchivierung

SPASSPROJEKT: Entwicklerin erstellt PDF-Dokument in der Größe der Welt

Sind PDFs in ihren Ausmaßen in der Größe limitiert? Eine Frau wollte es genau wissen und erstellte ein Dokument, das größer ist als Deutschland.

PDF-Enthusiastin Alex Chan hat ein Experiment durchgeführt, um ein extrem großes PDF-Dokument zu erstellen – lediglich, um zu sehen, ob es möglich ist. Mit ihrem Wissen über das PDF-Dateiformat machte sich Chan daran, ein PDF-Dokument zu erstellen, das größer ist als die Bundesrepublik Deutschland. [...]

Sie lädt andere Dateiformat-Enthusiasten ein, mit ihr die Möglichkeiten jenseits der dokumentierten Spezifikationen zu erforschen.

—Andreas Donath 3. Februar 2024, 14:21 Uhr

E-Mail-Archivierung - eine besondere Herausforderung

Geschäftliche E-Mails sind Dokumente und müssen mit der gleichen Sorgfalt (revisionssicher) archiviert werden wie andere Dokumente.

Herausforderungen:

- nicht alle E-Mails sind (geschäftlich) relevant (z. B. private E-Mails oder Spam)
(Wer entscheidet, was relevant ist? Dezentral die Mitarbeiter oder zentral die IT?)
- manche E-Mails sind sehr komplex (z. B. Anhänge)
- E-Mails enthalten links auf flüchtige Dokumente (z. B. Webseiten)
- E-Mails enthalten Schadsoftware
(Diesbezüglich gilt das Fernmeldegeheimnis seit dem 1.12.2021 nicht mehr.)

Aufgrund rechtlicher Rahmenbedingungen wird im Allgemeinen die zentrale Archivierung von E-Mails bevorzugt.

Pflege des Systems

Regelmäßiges Umkopieren der Bestände ist häufig notwendig:

- Datensätze müssen gelöscht werden (Datenschutz)
- Dokumente und Mappen sollen zusammengefasst werden (Performance)
- Datenträger altern (Sicherheit)

Hinweis

Kann mit einer Migration auf andere Formate bzw. Formatversionen verbunden sein!

Sichere Speicherung

- Dokumente und Metadaten können grundsätzlich folgendermaßen gespeichert werden:
 - zentrale Speicherung
 - dezentrale, verteilte Speicherung
 - dezentrale Speicherung mit Replikation

Daraus ergeben sich Unterschiede in Hinblick auf:

- Administration
 - Transferzeiten (Latenz)
 - Speicherplatz
- Dokumente und die Metadaten können, müssen aber nicht gemeinsam gespeichert werden.

Die Frage ob die Metadaten oder Dokumente zentral oder dezentral gespeichert werden, stellt sich insbesondere bei Behörden und großen Unternehmen, die über mehrere Standorte verfügen.

Warnung

Das Speichern auf lokalen Netzlaufwerken oder gar auf dem Arbeitsplatzrechner ist sowohl aus betrieblicher (und ggf. auch rechtlicher Sicht) nicht ausreichend.

Drucken von Dokumenten

DM-Systeme müssen das Drucken von Dokumenten unterstützen, da Nutzer oft Ausdrücke benötigen (z. B. auf Baustellen)

Relevante Anforderungen an den Ausdruck:

- eindeutige Dokumentenkennungen müssen mit ausgedruckt werden
- relevante Dokumentenmerkmale (z. B. Autoren, Version, Status) sollten auf den Ausdrucken erscheinen
- ggf. automatisches Paginieren bei mehrseitigen Dokumenten; insbesondere wenn Signaturnachweise gefordert und mitgedruckt werden sollen

4. Integration

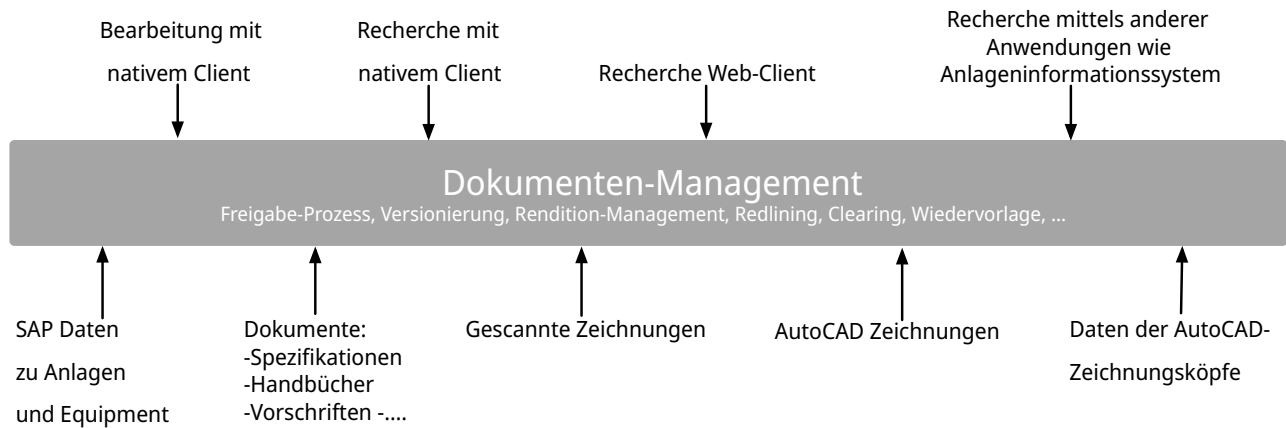
DMS als Infrastruktur

- DMS kann nie sinnvoll für sich allein stehen
- DMS bietet Dienste für andere Applikationen an, um Dokumente zu verwalten.
- Integration in:
 - Standardsysteme (wie MS-Office)
 - Mail-Systeme
 - ERP- / CRM- / PLM-Systeme und andere
 - Suchmaschinen
 - Workflow
 - Scanning
 - ...
- Referenzen auf Dokumente außerhalb des DMS

Typische (Web-) Dienste eines DMS

- Create document
- Move document
- Copy document
- Delete document
- Search document
- Search full text
- Retrieve full document info
- ...
- Create folder
- Move folder
- Copy folder
- Delete folder
- Search folder
- ...

Beispiel: Anlagendokumentation



Redlining bzw. die Redlining-Funktion, oder auch Rotstiftfunktion, bezeichnet das Markieren und Ändern von elektronischen Dokumenten oder Zeichnungen zu Feedback-Zwecken. Anmerkungen und Änderungen sind ersichtlich, ohne dass die Originaldatei verändert wird.

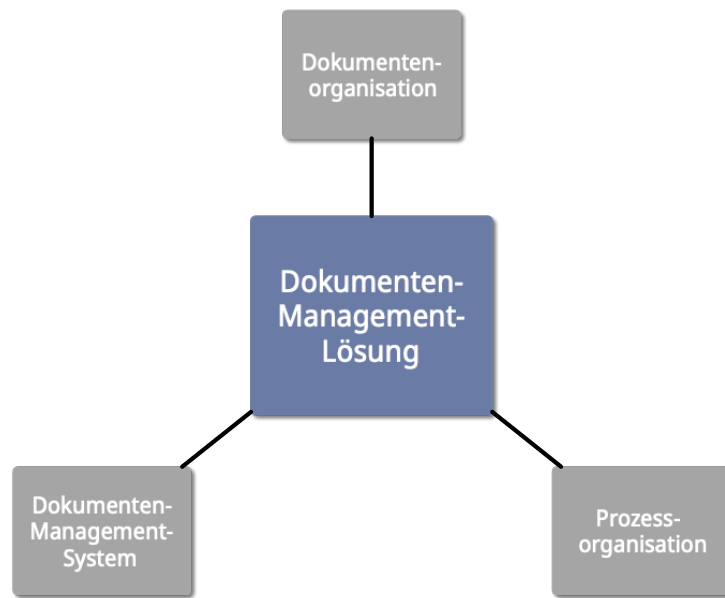
(Vgl. <https://www.fme.de/blog/redlining-spezielle-anwendungsfaelle-auf-basis-von-opentext-documentum-for-life-sciences-teil-5/>)

Beispiel: Anlagendokumentation

- Integrierte und konsistente Gesamtsicht auf alle Anlagen und ihre Bestandteile (Pläne und sonstige Dokumente)
- SAP als führendes System für Schlüssel (Datenqualität)
- Aktuelle Version der Dokumente überall und sofort verfügbar (Web)
- Revisionssichere Prozesse für Bereitstellung und Freigabe der Prozesse
- Eine verbindliche Quelle für alle Dokumente und alle Nutzer
- DMS wird für weitere Bereiche genutzt (z. B. SAP-Eingangs- und Ausgangsrechnungen, Magazin, E-Mail-Archivierung).

Typische Fragestellungen bei Integration

- Welches System ist bezüglich der Schlüssel das führende System?
- Einheitliche Nutzer- und Zugriffsrechte
- Wie werden die Systeme synchronisiert (permanent-online oder zyklisch im Batch)?
- Schnittstellenrealisierung
- ...



5. Produkte

Kategorisierung von DMS Produkten

- nach Funktionsbreite

- nach Funktionstiefe

Zum Beispiel in Hinblick auf unterstützte Dateiformate, unterstützte Barcode-Typen, ...

- nach Skalierung

In Hinblick auf die Anzahl der Dokumente, die Anzahl der gleichzeitigen Nutzer, die Anzahl der Standorte, ...

- nach Unterstützung spezifischer Kontextabhängiger Anforderungen

Bei internationalen Unternehmen ist ggf. die Unterstützung von mehreren Sprachen notwendig bzw. die Einhaltung verschiedener rechtlicher Rahmenbedingungen. Weiterhin können besondere Schnittstellen zu anderen Systemen erforderlich sein.

Enterprise Content Management (ECM)

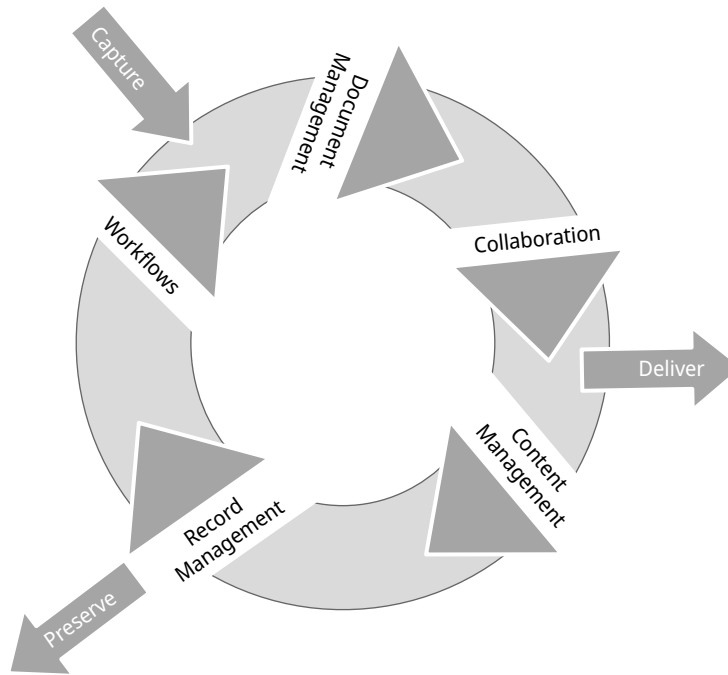
- Lösungen, die über das reine Dokumentenmanagement hinausgehen und auch Content-Management oder Records-Management (🇩🇪 *Schriftgutverwaltung*) unterstützen.
- Häufig aus Content-Management-Systemen (CMS) hervorgegangen.
- unterstützt auch schwach- oder unstrukturierte Daten (z. B. E-Mails, Webseiten, ...)

Beim Records-Management (Schriftgutverwaltung) ist die Betrachtung des gesamten Lebenszyklus von Akten in Hinblick auf einen sachlichen Kontext bzw. Geschäftsvorfall (Transaktion) und die Dokumentation desselbigen im Fokus.

Records-Management erstreckt sich ggf. auch auf Papierdokumente.

Records-Management bedingen den Einsatz von DMS und Workflow-Systemen.

Enterprise Content Management (ECM) im AIIM Modell^[4]



[4] Association for Intelligent Information Management (AIIM)

Open-Source Lösungen (Stand 2024)

- Agorum Core Open
- LogicalDOC Open Source DMS
- Bitfarm Archiv

Open-Source Lösungen bieten meist grundlegende Funktionalität

*Für bitfarm-Archiv Enterprise sind zumindest die Serverinstallation und eine Schulung des / der IT-Betreuer durch uns oder einen unserer Dokumentenmanagementsystem Partner notwendig. **So können wir dafür garantieren, dass das System funktioniert und den technischen Teil der zum Nachweis der Revisionssicherheit erforderlichen Verfahrensdokumentation übernehmen.***

Gerade in kleineren Unternehmen mit übersichtlichen Anforderungen und knapper Kostenkalkulation kann das DMS anschließend in Eigenleistung konfiguriert werden. Tauchen dabei Fragen oder Probleme auf, stehen Ihnen unsere Supportmitarbeiter zur Seite. [...]Inklusive professionellem Support für ein Jahr sowie Installation, aller Module & Plugins, ist die Enterprise Version schon ab etwa 2500,- € erhältlich.