

Konzeption eines Information-Lifecycle-Management-Frameworks im Dokumenten-Management-Kontext

Michael G. Kaiser, Stefan Smolnik, Gerold Riempp

Lehrstuhl für Wirtschaftsinformatik
European Business School (EBS)
International University Schloss Reichartshausen
Rheingastr. 1
65375 Oestrich-Winkel
[michael.kaiser|stefan.smolnik|gerold.riempp]@ebs.edu

Abstract: Information Lifecycle Management (ILM) wurde in den letzten Jahren ein beliebter Terminus für ein recht unscharf definiertes Bündel von Produkten, Technologien und Prozessen. Ziel dieses Beitrags ist es, eine schärfere Abgrenzung von ILM gegenüber anderen Technologien wie insbesondere dem Enterprise Content Management zu ermöglichen und weiterhin eine Methodik zur Entwicklung der notwendigen Konfigurationsparameter zur Verfügung zu stellen, mit deren Hilfe Organisationen in der Lage sind, ILM effizient zu implementieren.

1 Einführung

Information Lifecycle Management (ILM) erfährt derzeit große Aufmerksamkeit von Seiten der Speicher- und Informationsmanagement-Industrie. Hersteller wie EMC oder Hewlett-Packard richten ihre gesamten Produktlinien im Storage- und Content-Management-Bereich nach dem Paradigma des aktiv verwalteten Informationslebenszyklus aus. Auch das Interesse der Benutzer bzw. Kunden scheint geweckt, wobei umfangreiche und vollständige Implementierungen bisher in der Praxis nur selten anzutreffen sind.

Das Konzept stammt in erster Linie aus dem nicht-akademischen Umfeld, so ergeben sich einerseits definitorische Probleme und andererseits herrscht eine gewisse Unsicherheit bezüglich der genauen Abgrenzung von Vorgänger- oder Paralleltechnologien. Ziel dieses Beitrags ist es daher, ILM mit Hilfe grafischer Artefakte logisch zu strukturieren, einen Vorschlag für den Umfang und die Inhalte dieses Konzepts zu liefern und Prozesse zu konzipieren, die dieses Verständnis von ILM unterstützen. Kern dieser Strukturierung ist der ILM-Navigator, der alle essentiellen Komponenten von ILM darstellt und verknüpft.

Das hier vorgestellte Framework bezieht sich explizit auf den Dokumenten-Management-Kontext. Auch wenn ILM grundsätzlich für jegliche Art von Daten geeignet ist, bietet sich die Nutzung insbesondere für kohärente Informationsobjekte wie Dokumente an, da deren Wert für die Organisation vergleichsweise einfach auch isoliert bestimmbar ist, während dies etwa bei Datensätzen in einer Datenbank auf Grund der starken Interdependenzen ungleich schwieriger ist.

Im Folgenden wird erst der Ursprung und eine gängige Definition von ILM vorgestellt. Anschließend wird das ILM-Framework mit besonderem Fokus auf die Bewertungsprozesse und den typischen Lebenszyklus eines Dokuments durch die Organisation erläutert. Es wird im Detail der Prozess von der Bestimmung der Anforderungen über die Definition von Dokumentenklassen und Service Level Objectives bis zur Ableitung der relevanten Policies erläutert.

2 Information Lifecycle Management: Ursprung und Definition

Information Lifecycle Management (ILM) hat seine Ursprünge im Datenmanagement. Als technologische Vorgänger werden in der Literatur meist System-Managed Storage [Ge89], Hierarchical Storage Management (HSM) [Ch73] und Storage Resource Management (SRM) [CO05] genannt. ILM verfolgt jedoch einen breiteren Ansatz als die genannten Technologien, da nicht nur die effiziente Nutzung von Speicherkapazitäten, sondern auch eine allgemeine Zentralisierung querschnittlicher Aspekte des Informationsmanagement verfolgt wird. Der zusätzliche Nutzen liegt also vor allem in dieser Verallgemeinerung und auch in einer Einbeziehung zusätzlicher Entscheidungskriterien in das Management von Informationen, insbesondere aus dem Bereich des Enterprise Content Management (siehe Abb. 1).

Wurde der Begriff des ILM in seinen Anfängen hauptsächlich zur Vermarktung existierender Produkte im Bereich des Speichermanagements verwandt, haben sich mittlerweile mehrere Industrievereinigungen des Themas angenommen, um eine klarere Definierung des Begriffs zu erreichen. Eine weit verbreitete Definition, die aus den bisherigen Standardisierungsinitiativen hervorging, ist die der ILM-Initiative der Storage Networking Industry Association (SNIA):

“Information Lifecycle Management is comprised of the policies, processes, practices, and tools used to align the business value of information with the most appropriate and cost effective IT infrastructure from the time information is conceived through its final disposition. Information is aligned with business requirements through management policies and service levels associated with applications, metadata, and data.” [PS04]

Diese Definition dient als Grundlage für die folgenden Ausführungen und das ILM-Framework; verwandte Terminologien werden, soweit möglich, übernommen.

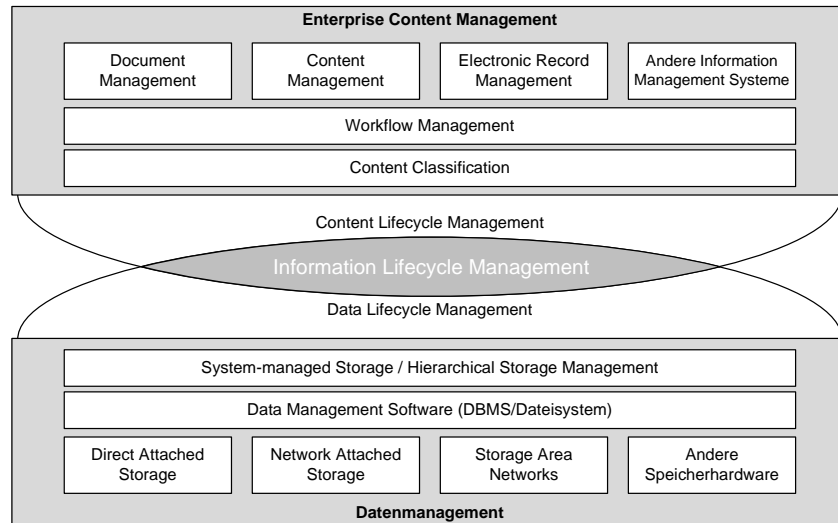


Abbildung 1: Positionierung von Information Lifecycle Management

3 ILM-Framework

Im Folgenden werden die einzelnen Komponenten des ILM-Frameworks beschrieben und in Zusammenhang gebracht. Die Systematik des Frameworks ist im ILM-Navigator verdeutlicht (siehe Abb. 2). Es werden hier allerdings nur die Informationsbewertung und der Lebenszyklus im Detail beschrieben, die Infrastruktur soll nur kurz im Ausblick erwähnt werden, da sie stark von den jeweils verwendeten Technologien bzw. Produktfamilien abhängt.

3.1 Geschäftswertermittlung

Vorraussetzung von effektivem Lebenszyklusmanagement von Dokumenten ist die Ermittlung des Geschäftswerts eines Dokuments. Von den Methoden, die dem Zweck der Geschäftswertermittlung von Informationen dienen, zielen die meisten (etwa [G193]) aber auf eine Bewertung im monetären Sinne, etwa im Zuge einer Due Diligence, ab. Der hohe Abstraktionsgrad dieser Methoden ermöglicht eine Ermittlung des Anteils von Informationen am Firmenwert, es macht sie jedoch auf Grund der zu geringen Detailtiefe für ILM ungeeignet.

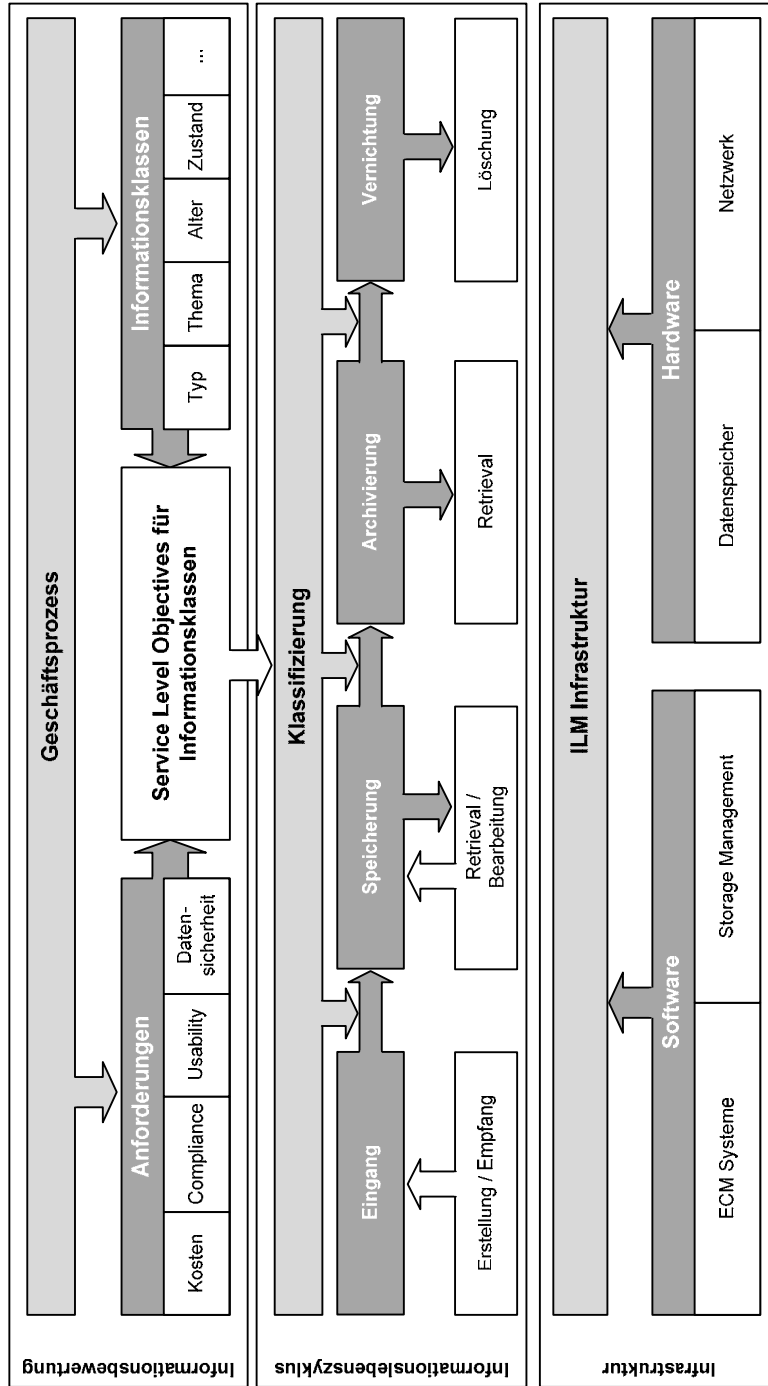


Abbildung 2: ILM-Navigator

Um eine Ausrichtung der IT-Infrastruktur einer Organisation sinnvoll durchzuführen, muss der Wert von exakt bestimmten Klassen von Dokumenten eruiert werden. Die Aspekte, die den Geschäftswert eines Dokuments beeinflussen, können in folgende Gruppen unterteilt werden:

- **Kosten:** Eine effizientere Nutzung und Speicherung von Dokumenten ermöglicht, die Kosten von Geschäftsprozessen und der IT-Infrastruktur zu reduzieren. Die für ein Dokument anfallenden Kosten sollten immer im Verhältnis zu seinem Wert für die Unternehmung stehen.
- **Usability:** Eine höhere Usability von Dokumenten erhöht ebenfalls ihren Wert. Im Bezug auf ILM bedeutet dies vor allem, die Performanz und Verfügbarkeit zu steigern. Je höher der Wert eines Dokuments ist, desto wichtiger ist eine hohe Zugriffperformanz und Verfügbarkeit der Daten.
- **Compliance:** Unter dem Begriff Compliance wird die Übereinstimmung der Datenhaltung und -verarbeitung mit allen relevanten Gesetzen und internen bzw. externen Bestimmungen zusammengefasst. Dies sind einerseits Aufbewahrungsfristen und Datenschutzvorschriften, andererseits die Bewahrung von Betriebsgeheimnissen.

Eine direkte, monetäre Bewertung der Dokumentenklassen ist im Kontext von Information Lifecycle Management nicht zwingend notwendig, da letztlich limitierte Ressourcen auf die Speicherung und Verwaltung der einzelnen Informationseinheiten rationiert werden müssen. Vielmehr muss also ein relativer Wert der Informationen zueinander und Mindestanforderungen an die IT der Organisation etabliert werden, um diese Zuteilung so adäquat wie möglich durchzuführen. Im Folgenden soll eine Methodik beschrieben werden, die eine strukturierte Bewertung der Dokumentklassen erlaubt.

3.1.1 Bewertungsmethodik

Es handelt sich bei der Bewertung von Informationen prinzipiell nicht um ein exaktes Verfahren. Vielmehr wird immer in einem Spannungsfeld verschiedener Interessen agiert, die Kompromisse unumgänglich machen. Die vorgestellte Methode versucht deshalb, Dokumenten einen relativen Wert zuzuordnen, der es ermöglicht, bei gegebenen Infrastrukturbudgets eine Priorisierung der Ressourcenallokation vorzunehmen.

Grundsätzlich lässt sich sagen, dass eine stark überproportionale Gewichtung des Kosten-, Compliance- oder Usability-Aspekts zu ungünstigen Ergebnissen führen kann, da sie tendenziell zu einer Vernachlässigung der anderen Aspekte führt. Ein zu starker Fokus auf die Kostenreduktion kann so etwa zu einer unannehmbar schlechten Usability und Vernachlässigung von Compliance und Datensicherheit führen. Es ist somit darauf zu achten, dass bestimmte Grenzen in der Planung, etwa der Kostenplanung, nicht unter- bzw. überschritten werden sollten. Diese Minima bzw. Maxima werden häufig durch externe Faktoren bestimmt, etwa festgesetzte Budgets, herstellerseitige Mindestanforderungen an die Performance oder rechtliche Bestimmungen. In den nächsten Abschnitten soll nun die grundlegende Terminologie der Informationsbewertung vorgestellt und anschließend der Bewertungsprozess genauer erläutert werden.

3.1.2 Dokumentenklassen

Als Grundlage für die spätere Bewertung müssen Dokumentenklassen definiert werden, für die dann Infrastrukturanforderungen ermittelt werden. Grundlage für diese Klassenbestimmung sind in der Regel die Geschäftsprozesse, in denen die Dokumente verwandt werden, aber auch einige der oben erwähnten externen Faktoren wie etwa rechtliche Vorgaben können eine Rolle spielen.

Eine Dokumentenklasse ist im Kontext der hier vorgestellten Methodik eine Sammlung von Charakteristiken, so genannten Attributen, die von einem Dokument erfüllt werden müssen, um zu der entsprechenden Klasse zu gehören. Besitzt ein Dokument alle Attribute einer Klasse, so ist es dieser Klassen zugehörig. Attribute sind vielgestaltig, es kann sich um eindeutige Bedingungen („Das Dokument ist ein Kreditantrag.“), Spannen von Werten („Das Dokument ist zwischen 3 und 10 Jahren alt.“) aber auch um statistische Eigenschaften, die durch eine Inhaltsanalyse ermittelt wurden, handeln. Welche Attributtypen zur Verfügung stehen, entscheidet sich letztlich durch die Leistungsfähigkeit des Klassifizierungssystems, das innerhalb der ILM-Lösung eingesetzt wird.

Während einfache Attribute, wie etwa das Alter des Dokuments (in einem Informationssystem), ohne großen Aufwand zugänglich sind und von den meisten ILM-Systemen unterstützt werden, sind statistische Klassifizierungen, die den Inhalt eines Dokuments einer bestimmten Thematik zuordnen, technisch wesentlich anspruchsvoller. Tabelle 1 zeigt exemplarisch eine solche Dokumentenklasse. Dabei ist zu beachten, dass ILM-Systeme bei der Eruiierung von Attributwerten nicht auf das das Dokument speichernde Informationssystem beschränkt sein müssen, sondern durchaus auch über entsprechende Querverweise, etwa ein Vorgangsnummer, auf andere Informationsquellen zurückgreifen können. Im vorliegenden Fall wäre dies das Datenbanksystem des Kreditverwaltungssystems einer Bank, in dem der Tilgungsstatus des betreffenden Kredits hinterlegt ist.

Attribut	Wert	Quelle
Typ	Kreditantrag	Dokumenten-Management-System (DMS)
Alter	3-5 Jahre	DMS oder Dateisystem
Prozessstatus 1	Antrag genehmigt	Workflow Management System
Prozessstatus 2	Kredit nicht getilgt	Datenbanksystem

Tabelle 1: Beispiel einer Dokumentenklasse

3.1.3 Anforderungen

Sind die Dokumentenklassen definiert, müssen die Anforderungen spezifiziert werden. Anforderungen spiegeln einerseits den Wert wider, der Dokumenten in einer Organisation zugeordnet wird, und sie sind andererseits die Grundlage für die Regeln, die nach der Bewertung in den ILM- und Informationsmanagement-Systeme implementiert werden.

Zur Ermittlung der Anforderungen können die bewährten Methoden des Requirements-Management übernommen werden, wie es beispielsweise in der Softwareentwicklung weit verbreitet ist. Nuseibeh und Easterbrook [NE00] schlagen die folgenden, generischen Schritte für die Anforderungsermittlung vor:

- Sammeln der Anforderungen
- Modellieren und analysieren der Anforderungen
- Kommunizieren der Anforderungen
- Einigung über die Anforderungen erzielen
- Anforderungen weiterentwickeln

Im Kontext von ILM bedeutet dies, dass die Anforderungen der verschiedenen involvierten Interessengruppen (Risikomanagement, Nutzer, etc.) an die Infrastruktur bezogen auf die einzelnen Dokumentenklassen harmonisiert werden müssen. Dies entspricht dem Vorgang der Modellierung von Anforderungen für ein Service Level Agreement (SLA), wodurch Prozess-Know-How des IT-Managements in der Formulierung von SLAs auf den ILM-Prozess übertragen werden kann. In dieser Prozessphase sollten auch die relevanten Minimal- bzw. Maximalanforderungen an die Infrastruktur ermittelt werden, um die Spielräume, die bei der Gestaltung und Gewichtung der Infrastruktur genutzt werden können, zu verstehen.

3.1.4 Service Level Objectives

Aus den Anforderungen können nachfolgend die technischen Ziele, die Service Level Objectives (SLO), abgeleitet werden. Die Definition der SLOs fällt weitgehend in den Bereich der IT-Abteilung, da hierfür bereits ein hohes Maß technischen Sachverstands bezüglich des Verhaltens der eingesetzten Systeme Voraussetzung ist.

SLOs unterteilen sich in verschiedene Felder, in denen entsprechende Parameter bestimmt werden (vgl. Abb. 3). Typische Parameter sind beispielsweise die Verfügbarkeit von Daten, also der Anteil einer bestimmten Zeiteinheit (typischerweise Tag, Monat oder Jahr), in dem auf Daten zuverlässig zugegriffen werden kann, die Anzahl der redundanten Kopien oder auch die Zeit, die ein Datensatz maximal unerreichbar sein darf, wenn ein signifikanter Zwischenfall (Feuer, Überschwemmung, etc.) eintritt.

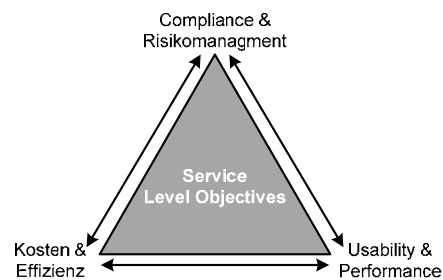


Abbildung 3: SLO-Spannungsfeld

Die Bestimmung von SLOs für jede Dokumentenklasse kann sehr aufwendig sein, besonders wenn eine große Anzahl von Klassen vorliegt. Unter diesen Umständen sollte in Erwägung gezogen werden, entweder verschiedene Klassen zusammenzufassen oder eine hierarchische Ordnung zu etablieren. Diese Hierarchie kann nicht nur den Aufwand der Zielbestimmung verringern sondern auch verhindern, dass einige Dokumente von keiner Klasse erfasst werden. Bedingung für solch eine Klassenhierarchie ist ein Klassifizierungssystem, das intelligent genug ist, Dokumente immer in die Klasse mit der höchsten Spezifität einzuordnen. Dies ermöglicht zunächst, eine grobe Aufgliederung der Dokumente vorzunehmen und diesen sicheren SLOs zuzuordnen, die je nach Bedarf weiter heruntergebrochen und präzisiert werden können.

3.1.5 Policies

Bei Policies handelt es sich schließlich um die eigentlichen Konfigurationsanweisungen in den ILM-Systemen. Da diese Konfigurationen implementierungsspezifisch sind und es sich dabei letztlich nur um die technische Reformulierung der SLOs handelt, soll hier nicht näher auf sie eingegangen werden.

3.2 Klassifizierung

Abhängig von dem eingesetzten ILM-System stehen unterschiedliche Methoden zur Dokumentenklassifizierung bereit. Die Leistungsfähigkeit und Integration mit den vorhandenen Informationssystemen beeinflusst sowohl die Möglichkeiten, die bei der Formulierung der Dokumentenklassen bereitstehen, als auch, inwiefern ILM umfassend und organisationsweit eingesetzt werden kann oder lediglich auf bestimmte Systeme beschränkt bleibt.

Klassifizierungsmethoden lassen sich in manuelle, Meta-Daten-, Inhalts- und Nutzungs-basierte Ansätze einteilen. Manuelle Klassifizierung erfordert eine direkte Interaktion des Nutzers mit dem Klassifizierungssystem. Er ist dafür verantwortlich, individuelle Dokumente einer Klasse zuzuordnen. Dieses Verfahren ist aus mehreren Gründen problematisch. Zum einen handelt es sich dabei um einen sehr zeitintensiven Prozess, der gerade bei einer größeren Zahl von Dokumenten und möglichen Klassen hohen Personalaufwand und dadurch Kosten verursacht. Da ILM meist auf größere Organisationen abzielt, die folglich eine große Zahl diverser Dokumente nutzen, ist der Einsatz manueller Klassifizierung deshalb meist aus Kostengründen nicht möglich. Dazu kommt, dass eine solche Klassifizierung potenziell fehlerbehaftet ist und je nach Domäne des Dokuments für eine Einordnung beträchtliches Fachwissen des Klassifizierenden voraussetzt.

Die Klassifizierung auf Basis von Meta-Daten ist dagegen praktikabler und weiter verbreitet. Meta-Daten werden in der Regel bei der Erstellung eines Dokuments (manuell oder automatisch) miterzeugt und stehen somit dem Klassifizierungssystem ohne zusätzlichen Aufwand zur Verfügung, solange passende Schnittstellen existieren. Ihre hohe semantische Eindeutigkeit und strukturierte Speicherung ermöglicht eine relativ einfache Spezifizierung zuverlässiger Klassifizierungsattribute. Die beiden Hauptquellen für Meta-Daten sind einerseits Dokumenten-Management-Systeme, die sie hauptsächlich für das Retrieval durch den Nutzer aber auch zur internen Verwaltung und Flusststeuerung verwenden. Andererseits können primitive Meta-Daten, etwa der Erstellungszeitpunkt oder die Größe eines Dokuments, auch direkt aus dem Dateisystem ausgelesen werden.

Inhaltsbezogene Klassifizierung ist technisch anspruchsvoller als die vorgenannten Methoden. Zuerst müssen die Inhalte eines Dokuments für die Klassifizierung zugänglich gemacht werden. Dies bedeutet im Falle von gescannten Dokumenten meist eine Verarbeitung mit einem OCR-System. Liegt der Text in kodierter Form vor, kann dieser dann einer Klasse zugeordnet werden. Die Klassifizierung kann an Hand bestimmter Schlüsselwörter, statistischer Analysen und verschiedener anderer Methoden vorgenommen werden, auf die an dieser Stelle nicht detailliert eingegangen werden soll (vgl. für eine Übersicht [YL99]).

Schließlich besteht die Möglichkeit einer Klassifizierung nach Kriterien, die sich über die Zeit ermitteln lassen. Beispiele sind die Nutzungshäufigkeit oder die verstrichene Zeit seit der letzten Änderung eines Dokuments. Prinzipbedingt sind diese Kriterien bei einer initialen Klassifizierung nicht verfügbar. Sie können erst über einen bestimmten Zeitraum ermittelt werden und dienen meist zur Bestimmung des Zeitpunkts, an dem eine Archivierung beziehungsweise Vernichtung der Dokumente stattfindet [LM06].

Der Klassifizierungsprozess ist nebenläufig zu dem eigentlichen Informationslebenszyklus. Er wird in bestimmten Intervallen oder ausgelöst durch bestimmte Ereignisse (etwa das Ende eines Geschäftsjahres) wiederholt und der Übergang eines Dokuments von einer Klasse in die nächste hat häufig auch den Übergang in die nächste Phase des Lebenszyklus zur Folge.

3.3 Lebenszyklus

Der Lebenszyklus von Informationen ist der Kern von ILM; er stellt die Prozesse dar, die ein Dokument während seiner Zeit in einer Organisation durchläuft. Im hier beschriebenen Ansatz wird der Lebenszyklus implizit über Reklassifizierungen realisiert. Dokumentenklassen werden bestimmten SLOs zugewiesen, die sich in der Regel nur langfristig ändern. Allerdings besteht bei jeder Klassifizierung eines Dokuments die Möglichkeit, dass dieses in eine neue Klasse mit eingeordnet wird und damit auch deren neue SLOs übernimmt. Dieses Vorgehen ermöglicht eine hohe Flexibilität, bedarf aber auch einer entsprechenden Werkzeugunterstützung, um die Komplexität der hohen Klassenanzahl beherrschbar zu machen.

Ein typischer Lebenslauf eines Dokuments reicht vom Eintritt in die Organisation über die Benutzung und Archivierung bis hin zur Vernichtung. In den nächsten Abschnitten soll kurz auf die Rolle von ILM in den einzelnen Lebenszyklusphasen eingegangen werden.

3.3.1 Eintritt

Es können grundsätzlich zwei Alternativen unterschieden werden, wie Dokumente in eine Organisation gelangen und im Rahmen des ILM betrachtet werden. Entweder stammen sie von einer externen Quelle und werden durch elektronischen Versand oder durch Einscannen der Papierform in ein ILM-System eingebracht oder sie werden von der Organisation selbst erstellt und direkt in dem internen Informationsmanagement-System abgespeichert.

Aus der Perspektive von ILM besteht ein wesentlicher Unterschied zwischen diesen beiden Quellen in der Kontrolle, die auf die interne Struktur der Dokumente genommen werden kann. Während bei einem elektronischen Versand und vor allem bei gescannten Dokumenten keine oder wenige Meta-Daten vorliegen und für die Klassifizierung auf inhaltliche Aspekte zurückgegriffen werden muss, um ein Dokument in eine Klasse einzuordnen, werden intern erstellte Dokumente in der Regel in einem strukturierten Format, wie etwa XML, abgelegt und schon bei ihrer Erstellung mit diversen Meta-Daten versehen, die den Aufwand bei der Klassifizierung erheblich vereinfachen.

3.3.2 Speicherung

Nach Eintritt des Dokuments in die Organisation und nach einer ersten Klassifizierung wird es in der IT-Infrastruktur gespeichert. Zum einen bedeutet dies in der Regel, dass es in ein Dokumenten-Management-System überführt wird, das den Nutzern die Möglichkeit des Zugriffs bietet. Zum anderen muss hier über die Art der physikalischen Speicherung entschieden werden. Abhängig von der zuvor bestimmten Klasse wird das Dokument automatisch auf ein Medium verschoben, das die mit der Klasse verbundenen Service Level Objectives erfüllt. Dabei wird auch die nötige Redundanz hergestellt und überwacht.

3.3.3 Archivierung

Sobald das Dokument nicht mehr operativ benötigt wird oder keine Änderungen mehr erwartet werden, kann es in ein digitales Archiv verschoben werden. Wann dieser Zeitpunkt erreicht wird, wird wiederum im Zuge einer Klassifizierung bestimmt. Ändert sich die Klasse eines Dokuments in einer Weise, die es für das Archiv qualifiziert, wird die Migration eingeleitet.

Im Unterschied zu der beschriebenen Speicherung sind Objekte im Archiv in der Regel nicht mehr editierbar. Damit werden nicht nur die Compliance-Anforderungen hinsichtlich einer Unterbindung von nachträglicher Datenmanipulation erfüllt, es erlaubt auch eine Optimierung der Infrastruktur, da beispielsweise Dokumente nur einmalig für eine Volltextsuche indiziert werden müssen und dieser Index bis zur Vernichtung des Dokuments gültig bleibt sowie generell Medien verwandt werden können, deren Leistung überproportional auf das Lesen von Daten optimiert ist. Abhängig von den Anforderungen können auch preisgünstigere Medien für das Archiv verwandt werden, da Wiederbeschreibbarkeit nicht erforderlich ist.

3.3.4 Vernichtung

Am Ende des Lebenszyklus steht schließlich die Vernichtung des Dokuments. Auch diese Phase wird durch eine entsprechende Klassifizierung eingeleitet, die ein Dokument als entbehrlich einstuft. Je nach der Bedeutung und rechtlichen Situation müssen verschiedene Maßnahmen zur Vernichtung getroffen werden. Bei weniger wichtigen Dokumenten reicht in der Regel eine Markierung zur Löschung, die dann in Intervallen ausgeführt wird. Andere Dokumente, insbesondere solche, die persönliche Daten enthalten, müssen unter Umständen sofort und unwiderruflich gelöscht werden.

Welcher Aufwand zur Löschung betrieben werden muss, hängt auch von den eingesetzten Medien ab. Nicht wiederbeschreibbare Medien, beispielsweise DVD-Rs, müssen manuell vernichtet werden und magnetische Medien müssen mehrfach überschrieben werden, um eine Wiederherstellung der Daten zu verhindern. All diese Konfigurationsdaten sind im Idealfall in den SLOs hinterlegt, so dass auch die Vernichtung weitgehend automatisiert durchgeführt wird.

4 Zusammenfassung und Ausblick

In diesem Beitrag wurde ein ILM-Framework vorgestellt, das sowohl die Abgrenzung von ILM gegenüber anderen Technologien als auch die Implementierung von Informationslebenszyklus-Management in einer Organisation erleichtern soll. Zu diesem Zweck wurde ein Prozess beschrieben, der die organisationalen Vorbedingungen einer ILM-Implementierung von der Bestimmung der Anforderung in den relevanten Gebieten Kosten, Usability und Compliance bis hin zu den System-Policies umfasst. Dies wurde zusammengeführt mit einer Methodik zur Definition und Zuweisung von Dokumentenklassen, um einen Lebenszyklus umfassend und flexibel abbilden zu können.

Derzeit ist der Stand der ILM-Implementierungen noch unbefriedigend. Hauptgrund dafür ist insbesondere der Mangel ausgereifter Standards für die Kommunikation der involvierten Komponenten. Dadurch sind Organisationen weitgehend auf Single-Vendor-Lösungen angewiesen. Architektonisch handelt es sich bei ILM derzeit häufig um Zusatzmodule zu bestehenden Backup- und Datenmanagement-Systeme. Dies führt dazu, dass das hier beschriebene übergreifende Informationsmanagement, das nicht nur Datenmigration und Backup sondern auch viele weitergehende Aspekte umfasst, derzeit nur mit großem Aufwand umzusetzen ist.

Um diese Situation zu verbessern, müssten die derzeitig vorherrschenden Architekturen für ILM-Systeme modifiziert und damit Klassifizierung, Policy-Verwaltung und das Management von Dokumentidentitäten zentralisiert werden. Ob dies in naher Zukunft realisiert werden kann, ist offen, da es eine Neuentwicklung großer Teile der heutigen Informations- und Datenmanagementsysteme nach sich ziehen würde.

Trotz dieser Problematik beschreibt ILM einen zukunftsweisenden Ansatz für den Umgang mit digitalen Dokumenten in großen Organisationen, der eine weitere Verbreitung finden wird sowie die Produktivität und Compliance nachhaltig positiv beeinflussen kann.

Literaturverzeichnis

- [Ch73] Chen, P.: Optimal file allocation in multi-level storage hierarchies. In: Proceedings of National Computer Conference and Exposition. AFIPS Conference Proceedings, Band 42. AFIPS Press, Montvale, NJ, USA, 1973, S. 277-282.
- [CO05] Chen, Y.; Ong, S.: Holistic Information Management Solutions. IBM Research Division, San Jose, CA, USA, 2005.
- [Ge98] Gelb, J.P.: System-Managed Storage. In: IBM Systems Journal, Band 28. 1998, S. 77-103.
- [GI93] Glazer, R.: Measuring the value of information: The information-intensive organisation, IBM Systems Journal, Band 32. 1993, S. 99-110.
- [LM06] Lotlikar, R.; Mohania, M. K.: Adaptive policies in information lifecycle management. In: Proceedings of the 17th international conference on database and expert systems applications Band 4080. Springer, Heidelberg etc, S. 612–621.
- [NE00] Nuseibeh, B.; Easterbrook, S.: Requirements engineering: A roadmap. In: Proceedings of the conference on the future of software engineering (ICSE '00). ACM Press, New York, NY, USA, 2000. S. 35-46.
- [PS04] Peterson, M.; St. Pierre, E.: Information Lifecycle Management Roadmap. SNIA Data Management Forum, 2004.
- [YL99] Yang Y.; Liu, X.: A re-examination of text categorization methods. In: Proceedings of the 22nd annual international ACM SIGIR conference on research and development in information retrieval (SIGIR '99). ACM Press, New York, NY, USA, 1999, S. 42-49.