

Bestandteile und Funktionen von Software-Lösungen zur E-Mail-Archivierung

Noch 2004 meinten bei einer Umfrage von SofTrust Consulting die meisten Unternehmen, dass eine organisierte E-Mail-Archivierung in naher Zukunft keine Priorität für sie habe. Inzwischen sieht das anders aus.

Als Ursache für das große Interesse wird meist auf die rechtliche Situation hingewiesen. Diese verlange eine Archivierung des E-Mail-Verkehrs. Das klingt einleuchtend, ist aber nur ein Teil der Wahrheit. Schließlich gelten die rechtlichen Vorschriften der AO und GPDdU nicht erst jetzt. Geschäftsrelevante E-Mails müssen nämlich schon seit spätestens 2002 vollständig, revisionssicher und in elektronischer Form für sechs, beziehungsweise für zehn Jahre aufbewahrt werden.

In Wirklichkeit kommt bei vielen Unternehmen der Hauptdruck für die E-Mail-Archivierung nicht von der Rechtsabteilung, sondern von den Anwendern, die nach größeren E-Mail-Boxen rufen und vom IT-Management, das in der E-Mail-Auslagerung eine Möglichkeit sieht, die E-Mail-Systeme wieder schlanker, effizienter und sicherer zu machen.

Diese beiden Hauptinteressenlager (Rechtsabteilung vs. Anwender und IT) muss man unterscheiden, denn sie führen zu vollkommen anderen Anforderungen an die E-Mail-Archivierung. Das spiegelt sich auch in den angebotenen Systemen wider: Jedem Archivierungssystem kann man ansehen, ob es ursprünglich primär für die Erfüllung rechtlicher Anforderungen (Compliance) oder für den Komfort der Anwender (unendliche Mailbox) entwickelt wurde.

Eine Sonderrolle haben jene E-Mailarchivierungslösungen, die in Dokumentenmanagement-Systeme integriert sind. Bei diesen steht die Integration der E-Mail-Dokumente in das übergeordnete Dokumentenmanagement und in den Workflow im Mittelpunkt. In einer von SofTrust Consulting 2008 durchgeführten Markterhebung bezeichneten sich 30 von 54 E-Mail-Archivierungslösungen als Dokumentenmanagementsystem (Details der Untersuchung sind unter <http://www.softrust.com/> kostenlos verfügbar). Wer ein solches Dokumentenmanagementsystem bereits einsetzt oder anschaffen will, sollte prüfen, ob dessen E-Mail-Integration nicht bereits seine Anforderungen erfüllt. Alle anderen Unternehmen kommen mit Spezialsystemen schneller zur Lösung. In Folge wird auf diese Spezialsysteme fokussiert, wobei das Gesagte auch für Dokumentenmanagementsysteme gilt.

Wer E-Mails archivieren will, muss in folgenden Bereichen Entscheidungen treffen:

1. Welche E-Mails sollen archiviert werden?
2. Wo und wie sollen sie erfasst werden?
3. Welche Bearbeitungsschritte sollten erfolgen?
4. Wie sollen sie gespeichert werden?
5. Wie ist der Zugriff organisiert?
6. Wie wird mit ihnen nach Ablauf der Aufbewahrungszeit verfahren?
7. Wie wird die Plattform verwaltet und betrieben?

Umfang der Archivierung

Aus gesetzlicher Sicht sind in Deutschland lediglich jene „Schriftstücke“ (und damit auch E-Mails) aufzubewahren, die entweder Handelsgeschäfte betreffen oder die für die Besteuerung von Bedeutung sind. Ein E-Mail betrifft dann ein Handelsgeschäft, wenn es die Vorbereitung, den Abschluss, die Durchführung oder die Rückgängigmachung eines Geschäfts zum Gegenstand hat. Von steuerlicher Relevanz sind jene E-Mails, die die Besteuerung oder bestimmte Bewertungen zum Gegenstand haben. Damit muss das Unternehmen von Gesetzes wegen lediglich für einen Teil seiner Mitarbeiter eine Archivierungslösung anbieten, beispielsweise für Mitarbeiter in den Finanzabteilungen und im Vertrieb. Und auch bei diesen muss laut Gesetz nur ein Teil der E-Mails archiviert werden. Beispielsweise braucht der Vertrieb Angebote, die zu keinem Abschluss führten nicht aufzubewahren - noch weniger natürlich E-Mails mit dem aktuellen Speiseplan der Kantine, die private E-Mail oder gar Spam-Mail.

Mit dem gesetzlich geforderten Archivierungsumfang geben sich viele Anwender allerdings nicht zufrieden. Sie wollen, dass die Archivierung auch Abteilungen erfasst, die keine Handels- oder bewertungsrelevante E-Mails schreiben und auch E-Mails erfassen, die von Gesetzes wegen nicht aufbewahrt werden müssten.

Unternehmen stehen damit vor der Herausforderung zu entscheiden, welche E-Mails in welchen Bereichen von einer E-Mail-Archivierung erfasst werden sollten. Im Vorfeld der Auswahl einer Lösung sollte deshalb unbedingt eine genaue Analyse und Definition stattfinden. Die Archivierung von privaten E-Mails stellt dabei eine besondere Problematik dar: Vertreter oder Nachfolger eines Mitarbeiters dürften eigentlich nicht mehr auf dessen Archiv zugreifen, da sie auch private Post lesen könnten. Das Herausfiltern privater E-Mails ist wiederum ein Akt, der gegen Fernmeldegesetze verstößt. Deshalb sollte – sofern noch nicht vorhanden – spätestens vor Einführung einer Archivierungslösung eine Betriebsvereinbarung über die Handhabung privater E-Mails getroffen werden.

Erfassung der zu archivierenden E-Mails

E-Mails können von Archivierungslösungen auf unterschiedlichste Weise erfasst werden. Am vollständigsten sind jene Systeme, die sich mit spezieller Hardware ins Netz einklinken und von jeder E-Mail automatisch eine Kopie ziehen –und zwar bevor die E-Mail das Unternehmen verlässt oder bevor es in die Eingangsbox gestellt wird. Diese Systeme sind garantiert vollständig und praktisch nicht manipulierbar. Der Nachteil besteht darin, dass viele E-Mails unnötig archiviert werden – was zu höherer Speicherbelastung führt, vor allem aber zur erschwerten Suche. Das Archiv kann auch viele E-Mails mit Texten enthalten, die besser nie geschrieben worden wären.

Andere Archivierungslösungen verlassen sich bei der Archivierung einzig auf die manuelle Direktive des Anwenders. Der Vorteil bei dieser Vorgehensweise liegt zweifellos darin, dass viele überflüssige E-Mails nicht archiviert werden und damit dem Archiv viel Schrott – und gegebenenfalls (rechtlich) belastendes Material – erspart bleibt. Nachteilig ist, dass die vom Gesetzgeber verlangte Vollständigkeit aller relevanten E-Mails schlecht nachzuweisen ist. Deshalb mag die manuelle Archivierung für bestimmte Abteilungen die richtige Lösung sein – für die Finanz- und Vertriebsabteilung sollten aber andere Erfassungsverfahren eingesetzt werden.

Die meisten Systeme mit manueller Archivierung bieten meist weitere Modi, die für bestimmte Mail-Boxen oder bestimmte Mail-Server automatisch Kopien für das Archiv zieht. Bei Exchange wird dazu in der Regel die Journalfunktion genutzt. Exchange schreibt für jede E-Mail eine Kopie in das Journal, die vom Archivierungsserver dann abgeholt wird. Im Gegensatz zu den ins Netzwerk eingeklinkten Hardwarelösungen kann hier nach Benutzergruppen unterschieden werden.

Neben der selektiven, benutzergesteuerten Archivierung und der automatischen, vollständigen Archivierung gibt es zusätzliche Archivierungsmodi, die eher an das hierarchische Speichermanagement denn an eine traditionelle Archivierung erinnern. In diesem Fall untersuchen die Archivierungslösungen bestimmte E-Mail-Boxen in regelmäßigen Abständen auf zu archivierende E-Mails. Seit dem letzten Archivierungslauf neu hinzugekommene E-Mails werden archiviert. In diesem Fall kann der Anwender zwischen den Archivierungsläufen noch alle jenen E-Mails löschen, die nicht archivierungswert sind.

Eine weitere Variante prüft die E-Mailboxen zunächst auf bestimmte Voraussetzungen - beispielsweise ob die Mail-Box eine bestimmte kritische Größe erreicht oder ob sie besonders große E-Mails oder ältere E-Mails enthält. Erst in diesem Fall werden E-Mails archiviert.

Bearbeitung

E-Mails, die von einer Archivierungslösung erfasst sind, können unterschiedliche Bearbeitungsschritte durchlaufen, bevor sie endgültig abgelegt werden. Nicht jedes System unterstützt jede Bearbeitung und auch innerhalb der Bearbeitungsart gibt es zwischen den Systemen teilweise erhebliche Unterschiede. Während manche Lösungen die Bearbeitung selbst vornehmen, binden andere Systeme Zusatzprodukte ein oder verlassen sich einfach darauf, dass das Unternehmen diese Werkzeuge ohnehin einsetzt. Als Bearbeitungsschritte sind vor allem Filtern, Volltextindizierung, Kategorisierung und Virenschutz zu erwähnen. Keiner dieser Bearbeitungsschritte ist gesetzlich vorgeschrieben, jedoch helfen sie, das Archiv aussagekräftig, sicherer und einfacher nutz- und verwaltbar zu machen.

Beim Filtern werden E-Mails aufgrund bestimmter Informationen im Header oder Text ausgefiltert. So können beispielsweise alle E-Mails, die von der Kantine geschickt werden oder die einen bestimmten Kürzel in der Betreffzeile haben von der Archivierung ausgeschlossen werden. Auch Spam soll natürlich nicht im Archiv landen und sollte deshalb vorher aussortiert werden.

Eine Volltextindizierung der E-Mails und deren Anhänge ist wichtig, wenn die E-Mails später wieder einfach aufgefunden werden sollen. Manche Systeme indizieren lediglich den E-Mail-Text. Andere Lösungen indizieren auch die Attachments, wobei die Anzahl der indizierbaren Dateitypen teilweise stark variiert. Es gibt Systeme, die die Dateien von hunderten von Applikationen analysieren können. Andere sind auf die populärsten Office-Anwendungen beschränkt. Unternehmen, die komprimierte Attachments indizieren wollen, sollten ferner darauf achten, dass das gewählte System dies auch ermöglicht.

Bei der Kategorisierung können die E-Mails aufgrund von Informationen im Header, im E-Mail-Text oder im Attachment bestimmten Kategorien zugewiesen werden. So können die E-Mails bestimmten Kundenakten oder bestimmten Dokumenttypen zugeordnet werden. Auch bestimmte Gefahrenklassen können festgelegt werden. Sofern die Kategorisierung noch vor dem Versand ausgehender E-Mail erfolgt, können E-Mails aufgrund der Gefahrenklassen-

Kategorisierung sogar noch aufgehoben werden – und brauchen deshalb auch nicht archiviert zu werden. Als „gefährlich“ oder „unerwünscht“ könnten beispielsweise E-Mails eingestuft werden, die Geheimnisverrat vermuten lassen oder rassistischen oder sexuellen Inhalt haben. Hierzu gibt es einen wachsenden Markt an „electronic discovery“-Modulen.

Im einfachsten Fall erfolgt die Kategorisierung manuell. Dabei muss der Anwender jedem zu archivierenden E-Mail Kategorieinformationen zuweisen. Das kann durch Ausfüllen einer Bildschirmmaske, durch Auswahl vorgegebener Optionen aus Drop-Down-Menues oder durch die Verwendung der Outlook-Kategorien geschehen. Zuverlässiger und für große E-Mail-Volumen geeigneter sind automatische Verfahren, die sich hinsichtlich ihrer Mächtigkeit allerdings sehr voneinander unterscheiden.

Sofern vorher noch nicht geschehen, sollten alle E-Mails vor der Ablage im E-Mail-Archiv noch auf Viren und Würmer untersucht werden. Viele Anbieter verweisen einfach auf kommerzielle Antiviren-Systeme. Andere Anbieter binden die Systeme direkt ein, was besonders beim Retrieval von E-Mails mit ursprünglich nicht erkannten Viren ein Vorteil sein kann.

Speicherung

Bezüglich der Speicherung der E-Mails verlangt der deutsche Gesetzgeber lediglich, dass die E-Mails originalgetreu und unveränderbar auf elektronisch lesbaren Speichern gehalten werden. Papierausdrucke oder Mikrofilme sind unzulässig da sie nicht elektronisch sind. Dagegen akzeptiert das Gesetz veränderbare Medien wie Festplatten, solange organisatorisch und softwaretechnisch gewährleistet ist, dass die gespeicherten E-Mail-Daten weder gelöscht noch verändert werden können. Solange nicht außerdeutsche Regelungen (z.B. SEC) erfüllt werden müssen, sind aus rechtlicher Sicht unveränderbaren Medien wie WORM also nicht notwendig. Jedes E-Mail muss mit einem eindeutigen, unveränderlichen Index versehen sein und über den gesetzlich vorgeschriebenen Aufbewahrungszeitraum (je nach Mail-Inhalt sechs bzw. zehn Jahre, in bestimmten Branchen bis 30 Jahre) für Prüfungen elektronisch verfügbar sein.

Archivierungslösungen, die primär auf die Erfüllung gesetzlicher Vorgaben ausgerichtet sind, schreiben einfach eine Kopie der zu archivierenden E-Mail ins Archiv und kümmern sich dann nicht weiter um das Originalmail. Von dieser Vorgehensweise unterscheiden sich jene Lösungen, die sich primär als „unendliche Mail-Box“ verstehen: diese Systeme migrieren beim Erreichen bestimmter Schwellenwerte ältere E-Mails und Attachments automatisch aus der Mail-Box in das Archiv. Zurück bleiben meistens nur kleine Platzhalter, die die verlagerten E-Mails weiterhin in der Mail-Box sichtbar halten. Jedes E-Mail ist mit seinem Attachment bei diesen Lösungen also idealerweise nur einmal vorhanden: entweder im E-Mail-System oder im Archiv. Mainframe- und Unix-Profis kennen diese Vorgehensweise vom Hierarchischen Storage-Management. Aus Anwendersicht haben diese Systeme den Vorteil, dass das Archiv im Prinzip für sie transparent ist. Es gibt Systeme, die in Abhängigkeit der E-Mail-Attribute das beste Speichermedium wählen, die Qualität der Datenträger überwachen, die E-Mails und Attachments verschlüsseln und vieles mehr.

Um Speicherplatz zu sparen komprimieren einige Systeme E-Mails und/oder Attachments. Weitere Systeme legen von E-Mails, die an mehrere interne Empfänger gingen nur eine einzige Kopie ab. 72 Prozent aller von SofTrust untersuchten E-Mail-Archivierungssysteme

bieten diese „Single-Instance“-Speicherung an. Auch hier gibt es große Unterschiede. Manche Archivierungsanbieter bieten Single-Instance nur für Attachments an. Andere brechen E-Mails und Attachments sogar auf und bieten Single-Instance auf Blockebene an.

Nicht jeder Anbieter löst die Storageaufgaben selbst. Manche setzen auf vorhandene Hierarchische-Speichermanagement-Software auf oder nutzen funktionsreiche Speichersubsysteme die bestimmte Funktionen in der Firmware realisiert haben (EMC, etc.).

Retrieval

Staat und Finanzamt verlangen, dass archivierte E-Mails für Prüfungszwecke schnell auffindbar sein müssen. Sie müssen für Prüfer im elektronischen Originalerscheinungsbild bereitgestellt werden. Je leistungsfähiger die Retrievalmethoden eines Systems sind, desto schneller und genauer wird man dieser Pflicht nachkommen können. Stringverarbeitung, Boolesche Operatoren, Wildcards und Unschärfetools werden nicht von allen Systemen in gleichem Maße angeboten. Ferner ist es natürlich wichtig, sich bei der Suche auf möglichst viele Elemente des E-Mails beziehen zu können. Je besser die archivierten E-Mail-Daten dabei im Vorfeld indiziert und kategorisiert wurden, desto besser werden naturgemäß die Suchresultate sein. Da eine Prüfungssituation vorliegt, ist es aus Sicht des Unternehmens wichtig, das Prüfgebiet möglichst weit einschränken zu können – also beispielsweise auf bestimmte Bereiche, Dokumentarten, Kunden, etc. Diese Begrenzungsmöglichkeiten sind sowohl für die Online-Prüfung durch einen Prüfer vor Ort, als auch für die Bereitstellung von Export-Dateien wichtig. Nur so kann verhindert werden, dass bei einer Prüfung plötzlich belastende Sachverhalte auf einem vollkommen anderen Gebiet auftauchen.

Wo Systeme lediglich der gesetzlichen Archivierungspflicht genügen müssen, reichen Retrievalzugänge für Verwalter und Prüfer aus. Anwenderorientierte Archivierungssysteme kommen damit natürlich nicht aus. Bei ihnen soll jeder Anwender Zugang zu seinen Archivdaten haben.

Archivierungssysteme, die Kopien aller E-Mails archivieren, haben normalerweise ein eigenes Log-In, ein eigenes Benutzerinterface und ein eigenes Suchsystem. Dagegen nutzen jene Systeme, die E-Mails transparent aus den Mail-Boxen in das Archiv migrieren für gewöhnlich das E-Mail-System als Front-End. Der Anwender sieht dabei über seinen E-Mail-Client (z.B. Outlook) in das Archiv. Er braucht keinen eigenen Log-In und auch keine Ausbildung bezüglich des Interfaces oder der Suchfunktionen. Das ist für den Anwender sehr bequem – die Suchoptionen sind aber in diesem Fall auf die vom E-Mail-System angebotenen Möglichkeiten beschränkt. Die Qualität der Suche hängt auch davon ab, wie groß der Platzhalter ist, in dem die Standardsuchfunktion suchen kann. Sofern die Suche des Clients nicht ausreicht, muss die Suche des Archivierers verwendet werden. Diese ist meistens in den E-Mail-Client (z.B. Outlook) integriert.

Die Systeme unterscheiden sich teilweise sehr darin, was mit den archivierten E-Mails getan werden kann. Anwenderorientierte Archivierungslösungen ermöglichen für gewöhnlich das einfache Zurückladen in die aktuelle Mailbox. Diese E-Mails können dann über den Client wie jedes andere E-Mail genutzt werden. Für Lösungen, die auf Compliance fokussieren ist dies nicht so wichtig. Diese Archivierungslösungen ermöglichen deshalb häufig kein Zurückladen der E-Mails ins E-Mail-System.

Entsorgung

Über die gesetzlichen Aufbewahrungsfristen hinaus braucht kein Schriftstück aufgehoben zu werden. Länger aufbewahrte E-Mails können bei Prüfungen aber durchaus mit geprüft werden. Deshalb sollten E-Mails, die ausschließlich zur Erfüllung gesetzlicher Auflagen gehalten werden, nach Ablauf der Frist gelöscht werden. Wie gut der Verwalter dabei unterstützt wird, differiert zwischen den Systemen sehr stark. Obwohl das Thema „Endbearbeitung“ erst mit dem erstmaligen Löschen nach sechs Jahren wirklich akut wird, sollte bereits bei der Auswahl auf die Erfüllung der eigenen Anforderungen geachtet werden.

Betrieb der Plattform

Mit der Entscheidung für eine Archivierungslösung legt sich eine IT-Abteilung für einen längeren Zeitraum fest. Deshalb sollte die Lösung einfach zu administrieren und zukunftssicher sein. Besonders ist abzuklären wie geschlossen oder offen das System sein sollte (Schnittstellen, unterstützte E-Mail-Systeme und E-Mail-Clients, Interfaces, etc.) und wie ausfallsicher die Archivierung und das Retrieval sein müssen.

Alle am Markt erhältlichen Lösungen sind relativ schnell zu implementieren. Abhängig von der Anzahl der aktiven Anwender kann die Anwenderschulung schnell zum größten Aufwand werden. 47 Prozent der Systeme sehen allerdings keine Anwenderschulung vor – sei es, weil der Anwender selbst keinen direkten Zugang erhält oder weil sich das System transparent in den E-Mail-Client einklinkt.

Alternativen

Vor allem kleinere Unternehmen die primär die rechtlichen Vorschriften erfüllen wollen sollten prüfen, ob sie nicht eine Appliance oder einen externen Archivierungsservice nutzen wollen. Es gibt eine zunehmende Anzahl von Anbietern, die schlüsselfertige Archivierungssysteme oder die „Archivierung in die Steckdose“ anbieten.

Fazit

Da unterschiedliche Interessengruppen nach E-Mail-Archivierung rufen, ist es nicht einfach, das richtige System zu finden, das allen gerecht wird. Zwischen den vielen angebotenen Softwarelösungen gibt es große Unterschiede hinsichtlich Positionierung, Leistungsumfang und Kosten. Jedes Unternehmen muss für sich das Produkt finden, das die eigenen Anforderungen am Besten erfüllt. Hierzu muss ein Anforderungsprofil erstellt werden, an dem sich die verschiedenen Systeme messen lassen müssen.

Zu SofTrust Consulting:

Die Unternehmensberatung SofTrust Consulting (Pullach bei München & Wien) beschäftigt sich mit dem Betrieb und der Nutzung von E-Mail-Systemen sowie mit der Gestaltung von E-Mail-Kultur. SofTrust Consulting veröffentlicht in regelmäßigen Abständen Marktübersichten über E-Mail-Archivierungslösungen. Kostenlose Kurzstudien sind bei www.softrust.com erhältlich.

SofTrust Consulting GmbH, Deutschland: Tel: +49-700-SOFTRUST, Österreich Tel: +43-1-33 33 634