

AUF PAPIER ODER ELEKTRONISCH? Grundsätzliche Überlegungen zur Archivierung im Labor

Dr. Stefan Schmitz*), Roland Heider), Dr. Iris Retzko*)**

*) CMC Pharma GmbH, Richard-Wagner-Straße 23, 68165 Mannheim, www.cmc-pharma.de

**) rh it consulting & services, Lessingstr. 12, 66186 St. Ingbert, rheider@t-online.de

Publikation in „LABO – Magazin für Labortechnik + Life Sciences“, Heft Januar 01/2007, Seiten 22-26

Viele Labors stehen heute vor der Frage wie sich die in großer Menge anfallenden Daten in einer behördlich akzeptierten Form elektronisch archivieren lassen. Es ist zurzeit üblich Papierlösungen neben elektronischen Lösungen zu betreiben, wobei die elektronisch basierten Systeme zum Teil erhebliche Mängel bezüglich Authentifizierung, Datensicherheit, Datenintegrität und Archivierung der Dateien und Datenbanken aufweisen. Ein häufiges Problem ist das herstellerabhängige Format welches die Frage nach der möglichen Wiederherstellung der Daten über die gesetzlich geregelte Aufbewahrungsfrist aufwirft.

Eine weitere Fragestellung ergibt sich nach der rechtlichen Freigabe (Authentifizierung) solcher Dokumente (Dateien), da die bisher üblichen Handunterschriften ohne weiteres nicht auf elektronischen Dokumenten anzuwenden sind.

Die meisten pharmazeutischen Unternehmen und deren Zulieferer arbeiten noch mit der Aussage: „Unsere Rohdaten sind Papier, unterschrieben, datiert und archiviert“. In der Praxis kommen zu den Rohdaten weitere Daten, denn längst wird durch die immer besser integrierten elektronischen Werkzeuge eine große Menge von elektronischen Daten erzeugt, die im Unternehmen, zwischen Unternehmen und Behörden ausgetauscht wird. Diesen Daten widmet man spätestens seit der Einführung des amerikanischen Gesetzes 21 CFR Part 11 im August 1997 [1, 2, 3], welches den Umgang mit elektronischen Aufzeichnungen und elektronischen Signaturen beschreibt, mehr Beachtung.

In der Praxis allerdings finden sich häufig nur halbherzig umgesetzte Strategien zum korrekten Umgang mit diesen elektronischen Aufzeichnungen. Man akzeptiert dabei bewusst Medienbrüche, teure paral-

lele Haltung von Papier neben elektronischen Aufzeichnungen (Hybridsituation) und nicht zuletzt die mangelnde Übereinstimmung mit behördlichen Forderungen (Compliance – Probleme), was Validierung, Datensicherheit, Wiederherstellbarkeit und Datenintegrität betrifft.

Die Gründe für die zögerliche Haltung der Industrie liegen zum Teil in den grundlegend unterschiedlichen Eigenschaften von Papierdokumenten im Vergleich zu elektronischen Aufzeichnungen (s. Tabelle 1).

Zur Veranschaulichung dient folgendes Gedankenexperiment. Wie können elektronische Daten in einem behördlichen Audit den äquivalenten rechtlichen Anspruch wie Papierdaten erfüllen?

Dazu müssen grundsätzliche Überlegungen über die Unterschiede von Papier und elektronischen Aufzeichnungen angestellt werden.

CMC Pharma GmbH

www.cmc-pharma.de
info@cmc-pharma.de

Tabelle 1: Unterschiede zwischen Papier- und elektronischen Aufzeichnungen

Anforderung	Papierdokumente	Elektronische Aufzeichnungen
Authentifizierung	Unterschrift und Datum per Hand auf dem Dokument ausgeführt sind weltweit gültig und anerkannt	Elektronische Signaturen mit Name, Datum/Uhrzeit (Zeitstempel) und Grund der Unterschrift unterliegen verschiedensten internationalen gesetzlichen Regelungen
Datensicherheit	Gewährleistet durch physikalische Aufbewahrung der Dokumente im Archiv	Gegeben durch physikalische und logische Zugangskontrollen (Benutzerzugangsregelung, z.B. Nur-Lesezugriff auf Archivdaten)
Datenintegrität	Die Nichtveränderung von archivierten Papierdokumenten wird postuliert	Prüfsummen erlauben die Erkennung der Nichtveränderung elektronischer Aufzeichnungen
Wiederauffindbarkeit	Indexierte Dokumente im Archiv sind manuell durchsuchbar	Automatisierbare Suchmechanismen
Wiederherstellbarkeit	In der Regel einfach, da es keine zusätzlichen Mittel benötigt	Abhängigkeit von verwendeter Hard- und Software
Langzeitarchivierbarkeit	Belegte Archivierbarkeit von Papier über mehrere Jahrhunderte	Abhängigkeit von verwendeter Hard- und Software und zusätzlich keine reale Langzeiterfahrung mit Datenträgern

Extrahiert man die Gründe, warum sich die meisten Unternehmen mit elektronischen Aufzeichnungen schwer tun, findet man drei Kernpunkte:

- **Authentifizierung** der Daten mit elektronischen Signaturen wird oft nicht umgesetzt wegen Abhängigkeiten von Hardware- und Softwaresystemen, sowie unterschiedlicher internationaler Signaturgesetze
- **Wiederherstellbarkeit von Daten** über den gesetzlichen Aufbewahrungszeitraum in Abhängigkeit von Hardware und Software
- **Datenträgeralterung**, die nach einer gewissen Zeit das Umkopieren auf andere Medien erfordert

Authentifizierung: Auf Papier ist es einfach eine weltweit rechtsgültige Unterschrift zu leisten. Für eine elektronische Unterschrift sind verschiedene rechtliche, aber auch technische Aspekte zu bedenken. Elektronische Unterschriften lassen sich durch verschiedene Systeme realisieren, z.B. durch passwortbasierte Lösungen (Benutzerkennung und Passwort), mit Smartcard und PIN oder durch Anwendung biometrischer Merkmale. Nach Meinung von Juristen gewährleisten nur biometrische Signaturen die eindeutige rechtssichere Zuordenbarkeit von Person zur Unterschrift und verhindern wirksam die Möglichkeit des Ableugnens einer Unterschriftsleistung.

In der Europäischen Gemeinschaft ist seit 1999 die EG Signaturrechtlinie (EGSIGRL) gültig [4], während in Deutschland seit 2001 das Signaturgesetz (SigG), die Signaturverordnung (SigV), sowie das Signaturänderungsgesetz (SigÄndG) [5, 6, 7] gelten. Es unterscheidet drei Stufen elektronischer Signaturen (Definitionen siehe Glossar):

einfache elektronische Signatur:

Anwendung von Benutzerkennung und Passwort zur Unterschrift

fortgeschrittene elektronische Signatur:

Anwendung von Signaturschlüsseln (private key) und Signaturprüfchlüsseln (public key) zusätzlich zu Passwort und Kennung; alternativ Anwendung biometrischer Verfahren

qualifizierte elektronische Signatur:

benötigt zusätzliche so genannte sichere Signaturerstellungseinheit und qualifizierte Zertifikate (Trust Center) für jeden Unterzeichner

Wiederherstellbarkeit von Daten: Als häufigster Grund für die mangelnde Akzeptanz elektronischer Archivlösungen wird die mangelnde Rückwärtskompatibilität von Softwarelösungen angeführt. Es ist durchaus richtig, dass viele Softwarehersteller nicht gewährleisten, dass mit der aktuellen Version ältere Daten wieder gelesen werden können.

Bei der Aktualisierung des Softwaresystems, in der Regel nach zwei bis drei Jahren stellt sich die Frage nach dem Umgang mit den Altdaten. Es können nun drei verschiedene Möglichkeiten eintreten:

- a) Die Altdaten können problemlos gelesen werden
- b) Die Altdaten können nur nach vorheriger Konvertierung gelesen werden
- c) Die Altdaten können nicht mehr gelesen werden

Im Fall a) kann ohne weiteren Aufwand mit den Altdaten weiter gearbeitet werden. Es ist aber durch Audit Trail und andere Mechanismen sicherzustellen, dass diese Altdaten nicht unkontrolliert verändert werden können (z.B. manuelle Nachintegration o.ä.).

Im Fall b) muss sichergestellt werden, dass die Daten mit einer validierten Pro-

zedur unter Erhaltung der Datenintegrität in das neue lesbare Format umkopiert werden.

Im Fall c) muss sichergestellt sein dass die Daten zumindest auf Papier oder elektronisch in einem Hersteller-unabhängigen Format archiviert wurden. Dies gilt auch für alle typischen Standard-Office-Anwendungen wie Word- oder Excel-Formate.

Durch die Verwendung von Dateien im PDF/A-1 – Format, welches durch die ISO – Norm 19005 [8] seit dem 11.10.2005 als langzeitarchivierbares Format gilt, werden die beiden Probleme Herstellerabhängigkeit und Wiederherstellbarkeit der Daten lösbar.

Datenträgeralterung: Datenträger, z.B. Magnetbänder und optische Speichermedien wie CD-ROM oder DVD unterliegen

im Gegensatz zu Papier einer schnelleren Alterung, die bei ungeeigneter Handhabung und Lagerung zu Ausfällen führen kann. Obwohl einige Hersteller Haltbarkeiten von bis zu 100 Jahren prognostizieren, sind Prozeduren einzurichten, die eine regelmäßige Integritätsprüfung der Archivdatenbestände durchführen und Daten in einem festgelegten Intervall auf neue Datenträger umkopieren. Hierzu gelten die gleichen Anforderungen wie an die Erstar Archivierung. Wegen der guten Automatisierbarkeit solcher Vorgänge können diese Verfahren schnell und kostengünstig durchgeführt werden.

Bei genauerer Betrachtung der Vor- und Nachteile beider Lösungen zeigt sich, dass sie sich antagonistisch verhalten, d.h. die Vorteile des Papiersystems stellen gleichzeitig Nachteile, bzw. Aufwand für das elektronische System dar und umgekehrt.

Tabelle 2: Vor- und Nachteile eines Papierarchivs gegenüber einem elektronischen Archiv

	Papierdokumente	Elektronische Aufzeichnungen
Vorteile	Kein aktiver Aufwand nach Archivierung zur Erhaltung der Gültigkeit, Datensicherheit, Datenintegrität	Leichte Integrierbarkeit, Verteilung, Prozessierbarkeit, Wiederauffindbarkeit (Retrieval), geringer Platzbedarf, niedrige laufende Kosten
Nachteile	Verteilung (Unterschriftenumläufe), Einzug, Aktualisierung, Dearchivierung (Wiederfinden), Vernichtung, da es immer nur ein Original gibt, Platzbedarf, Kosten	Erhöhter Aufwand zur Umsetzung einer rechtssicheren Umgebung, Migrationsaufwand auf neue Datenträger
Kosten	Raumkosten, Personalkosten, Druckkosten, Vervielfältigungskosten, Verwaltungskosten Archiv	Implementierung (Hardware/Software), Validierung, Signaturerstellung, Personalkosten, Schlüsselerstellung und -verwaltung, Migrationskosten, Datenträgerkosten

Tabelle 3: Kostenbeispiel Papierarchiv (ohne Berücksichtigung von Personal-, Raum-, Verwaltungskosten)

Papierarchiv	Anzahl	Kosten/Stück	Kosten
Aktenschranke (F90-Qualität)	3	333 €	1.000 €
Aktenordner	180	2,50 €	450 €
500 gedruckte Papierseiten/Ordner	90.000	0,05 €	4.500 €
Kosten für 90.000 Seiten			Ca. 6.000 €

Tabelle 4: Kostenbeispiel elektronisches Archiv (ohne Berücksichtigung von Personal-, Raum-, Verwaltungskosten)

Elektronisches Archiv	Anzahl	Kosten/Stück	Kosten
Datensafe (Kapazität für ca. 1.000 DVDs)	1	850 €	850 €
DVD (double layer)	1000	2 €	2.000 €
DVD Brenner	1	150 €	150 €
Zwischensumme für Archiv mit 1.000 DVDs			3.000 €
Kosten für Äquivalent zu 90.000 Seiten			Ca. 3 €

Bei einer angenommenen Datenmenge von durchschnittlich 100 Kb je Seite entsteht ein Datenvolumen von 9 GB (100 Kb x 90.000 Seiten), welches auf einer double layer DVD gespeichert werden kann.

Der Kostenvergleich basierend auf reinen Materialkosten ergibt einen Faktor von ca. 2.000 zu Gunsten der elektronischen Archivlösung. Vergleicht man umgekehrt das volle elektronische Archiv mit 1.000 DVDs mit dem äquivalenten Papierarchiv (180.000 Ordner) ergeben sich Kosten von

3.000 € zu 6.000.000 €. Selbst wenn man für das Papierarchiv die Aufbewahrungskosten (Schränke) auf Null setzt, belaufen sich die Kosten auf ca. 5.000.000 € (Faktor ca. 1.600 zu Gunsten der elektronischen Archivierung). Obwohl dieses Zahlenbeispiel konstruiert ist, veranschaulicht es doch die Größenordnung der Einsparungspotentiale einer elektronischen Archivlösung.

Eine einfache und kostengünstige Archivlösung in rein elektronischer Form könnte wie folgt aussehen:

CMC Pharma GmbH

www.cmc-pharma.de
info@cmc-pharma.de

Die zu archivierenden Dateien werden nicht ausgedruckt, sondern mit Standard-PDF-Tools in das PDF/A-1 Format umgewandelt und in einer Verzeichnisstruktur abgelegt.

Auf diese Verzeichnisstruktur wird nun ein Prüfsummenprogramm angewendet, wie z.B. md5 [9].

Dieses erzeugt für jede Datei eine Prüfsumme, die zu einem beliebigen späteren Zeitpunkt verifiziert werden kann und damit die Integrität dieser Datei belegt.

Im nächsten Schritt wird diese Verzeichnisstruktur mitsamt ihrer Inhalte auf ein geeignetes Archivmedium kopiert. Um Übertragungsfehler auszuschließen wird das Archivmedium ebenfalls mit dem Prüfsummenprogramm auf seine Integrität geprüft.

Ergibt sich für die zu archivierenden Daten die identische Prüfsumme wie für die Quelldaten, werden die Daten als zu archivierende Aufzeichnungen akzeptiert.

Zur Absicherung des nun erstellten Archivs kann das Protokoll der Prüfsummenberechnung als pdf-Datei vom Archivar elektronisch signiert und ebenfalls archiviert werden. Dieser Archivbeleg kann ohne großen Aufwand auch in das Quellverzeichnis zurückgespielt werden, so dass jederzeit auch ohne die Zuhilfenahme eines Archivars vom Anwender gezeigt werden kann, dass die elektronischen Aufzeichnungen auf einem bestimmten Medium archiviert wurden. Das Archiv wird nach diesem Vorgang geschlossen und

mit einem Nur-Lesezugriff ausgestattet. Nach erfolgreicher Archivierung der elektronischen Aufzeichnungen werden die Quelldaten gelöscht.

Archivierung in 10 Schritten

1. Selektion aller Dateien im Quellverzeichnis
2. Generierung und Speicherung der MD5 Prüfsummen der Dateien im Quellverzeichnis
3. Kopieren aller Dateien ins Archiv
4. Selektion aller Dateien im Archivverzeichnis
5. Generierung und Speicherung der MD5 Prüfsummen der Dateien im Archivverzeichnis
6. Inhaltlicher Vergleich der beiden gespeicherten MD5 Prüfsummendateien
7. Generierung einer PDF-Datei aus den gespeicherten MD5 Prüfsummen mit Informationen über das Archivmedium (Bezeichnung, Lagerort etc.) und mit elektronischer Signatur versehen
8. Die signierte PDF-Datei wird in Archiv- und Quellverzeichnis übertragen
9. Das Archiv wird geschlossen und mit Nur-Lesezugriff freigegeben
10. Die nun archivierten Dateien im Quellverzeichnis werden gelöscht.

Die oben beschriebene Archivelösung stellt eine einfache Vorgehensweise dar, die in beliebiger Weise skalierbar ist und damit jeder Unternehmensgröße angepasst werden kann.

Glossar

Begriff	Erklärung
21 CFR Part 11	Amerikanisches Bundesgesetz seit dem 20.08.1997, regelt den Umgang mit elektronischen Aufzeichnungen und elektronischen Unterschriften
Authentifizierung	Echtheitserklärung, d.h. im klassischen Fall eine Handunterschrift auf einem Papierdokument; im elektronischen Fall bedarf es einer elektronischen Signatur unterschiedlicher Ausprägung je nach Rechtslage (Weltweit verschiedene Signaturgesetze und Anerkennung)
Compliance	Übereinstimmung mit behördlich und selbst gestellten Anforderungen an ein Qualitätssystem, z.B. Archiv
Datenintegrität	Die Unversehrtheit der Daten nach Abspeicherung, d.h. die Daten sind unverändert. Manipulation und Datenträgeralterung können der Datenintegrität entgegenwirken
Datensicherheit	Hierunter ist zu verstehen der Schutz der Daten gegenüber Zugriffen von nicht berechtigten Bedienern. Der Schutz kann physikalisch erfolgen (Zugangsregelung) oder durch logischen Schutz in Form einer Benutzerzugriffsregelung. Darüber hinaus der Schutz der Daten vor Hard- und Softwarefehlern
Digitales Signaturgesetz SigG	Gesetz über die Rahmenbedingungen für elektronische Signaturen vom 16. Mai 2001
Digitale Unterschrift Definition nach 21 CFR Part 11, §11.3 (5)	Ist eine elektronische Unterschrift, basierend auf kryptographischen Methoden einer Urheber-Authentifizierung, berechnet durch einen Satz von Regeln und Parametern derart, dass die Identität des Unterzeichners und die Integrität der Daten verifiziert werden können
„Einfache elektronische Signatur“ nach SigG	Daten in elektronischer Form, die anderen elektronischen Daten beigefügt oder logisch mit ihnen verknüpft sind und die zur eindeutigen Identifizierung dienen
Elektronische Aufzeichnung Definition nach 21 CFR Part 11, §11.3 (6)	Als elektronische Aufzeichnungen definiert die FDA alle beliebigen Kombinationen von Text, Grafik, Daten, Ton, Bildern oder anderer Information in digitaler Form, welche generiert, verändert, gespeichert, archiviert, wiederhergestellt oder über ein Computersystem verteilt werden.
Elektronische Unterschrift Definition nach 21 CFR Part 11, §11.3 (7)	Eine elektronische Unterschrift bedeutet eine computerlesbare Übersetzung eines beliebigen Symbols oder einer Serie von Symbolen, die ausgeführt, angenommen oder durch eine Einzelperson autorisiert wird und das rechtlich gültige Äquivalent der individuellen handgeschriebenen Unterschrift darstellt.
„Fortgeschrittene elektronische Signatur“ nach SigG	Ist ausschließlich dem Signaturschlüssel-Inhaber zugeordnet, die Identifizierung des Signaturschlüssel-Inhabers ist möglich und ist mit Mitteln erzeugt, die der Signaturschlüssel-Inhaber unter seiner alleinigen Kontrolle halten kann, und mit den Daten auf die sie sich beziehen so verknüpft, dass eine nachträgliche Veränderung der Daten erkannt werden kann
Hash-Wert	Stellt eine nahezu eindeutige Kennzeichnung einer Datenmenge dar, ähnlich einem Fingerprint bei einem Menschen diesen nahezu eindeutig identifiziert
MD5	Message digest 5 ist ein Algorithmus zur Berechnung einer eindeutigen 128 Bit – Prüfsumme aus über einen Datensatz und

	erzeugt diese als 32-stellige Hexadezimalzahl, als so genannten Hash-Wert
PDF	Portable Document Format, Datenformat ursprünglich von der Firma Adobe eingeführt und gilt als offen gelegter Standard
PDF/A-1	Datenformat, welches in der ISO 19005 vom 01.10.2005 als langzeitarchivierbares Format gilt; anwendbar für Zeichen, Vektor- und Rastergrafiken
„Qualifizierte elektronische Signatur“ nach SigG	Zusätzlich zu den Anforderungen der qualifizierten elektronischen Signatur werden hier ein zum Zeitpunkt der Unterschrift gültiges qualifiziertes Zertifikat und eine sichere Signaturerstellungseinheit benötigt.
Qualifiziertes Zertifikat	elektronische Bescheinigungen für natürliche Personen, die von Zertifizierungsdiensteanbietern im Sinne des SigG ausgestellt werden
SigG	Deutsches Signaturgesetz in der Fassung vom 16. Mai 2001
SigÄndG	Änderung des Signaturgesetzes (SigÄndG) vom 04. Jan. 2005
SigV	Deutsche Signaturverordnung (SigV) vom 21. Nov. 2001 beschreibt die technische Umsetzung des SigG
Signaturschlüssel	einmalige elektronische Daten wie private kryptographische Schlüssel, die zur Erstellung einer elektronischen Signatur verwendet werden
Signaturschlüssel-Inhaber	natürliche Personen, die Signaturschlüssel besitzen und denen die zugehörigen Signaturprüfchlüssel durch qualifizierte Zertifikate zugeordnet sind
Signaturprüfchlüssel	elektronische Daten wie z.B. öffentliche kryptographische Schlüssel, die zur Überprüfung einer elektronischen Signatur verwendet werden
Zertifizierungsdiensteanbieter Certificate Authority (CA) Trust Center	natürliche oder juristische Personen, die qualifizierte Zertifikate oder qualifizierte Zeitstempel ausstellen, z.B. Bundesnetzagentur (BNetzA), Globalsign, Signtrust, Verisign u.a.

Literatur

- [1.] 21 CFR Part 11 Final rule, datiert 20. Aug. 1997
- [2.] 21 CFR Part 11 Guideline „Scope and Application“ Final Guidance Aug 2003
- [3.] Deutsche Übersetzung des 21 CFR Part 11 Final Rule, S. Schmitz, 2001
- [4.] Richtlinie 1999/39/EG EGSIGRL vom 13. Dez. 1999
- [5.] Deutsches Signaturgesetz (SigG) vom 16. Mai 2001
- [6.] Deutsche Signaturverordnung (SigV) vom 21. Nov. 2001
- [7.] Änderung des Signaturgesetzes (SigÄndG) vom 04. Jan. 2005
- [8.] ISO 19005:2005 vom 01.10.2005
- [9.] Wikipedia, md5 Algorithmus

Weblinks

<http://www.cmc-pharma.de>
<http://www.bsi.de>
<http://www.bundesnetzagentur.de>
<http://www.iso.ch>

<http://www.adobe.de>
<http://www.pdfa.org>
<http://estri.org/eCTD/>
<http://www.wikipedia.org>

CMC Pharma GmbH

www.cmc-pharma.de
info@cmc-pharma.de