

Langzeitarchivierung

Oliver Abraham, Abraham Beratung, November 2020

1. Einleitung

Langzeitarchivierung unterstützt bei der Archivierung und Lagerung von digitalen Daten auf externen Speichermedien.

Wenn man Daten auf Medien sichert, tut man dies vermutlich, um sie für sich selbst oder für die Nachkommen sicher aufzubewahren. Hierbei sind verschiedene Dinge zu beachten:

2. Funktionsübersicht

- Analysieren
- Prüfinformationen
- Auf Medium sichern
- Differenzen finden
- Medium prüfen
- Medium umkopieren
- Medienbestand

3. Lizenz

Die aktuelle Programmversion ist Freeware. Sie kann kostenlos und für alle Zwecke verwendet werden. Der Hersteller haftet nicht.

4. Auswahl von Medien

Die Medienvielfalt hat sich zum Glück auf wenige Typen reduziert. In den Anfängen der Computertechnik wurde viel ausprobiert. Dinge wie Disketten oder ZIP-Laufwerke sind in der Versenkung verschwunden. Bandlaufwerke sind nur noch in Firmen und Forschungsabteilungen vorhanden, wo hohe Datenvolumen gespeichert werden müssen.

Zur erschwinglichen Lagerung kommen nur drei Formate in Betracht:

- **Festplatten**
Magnetisches Medium, günstig und hohe Kapazität. Haltbarkeit 10-15 Jahre
Kein separates Laufwerk erforderlich
- **DVDs und BluRays**
Optisches Medium, günstig und geringere Kapazität, günstige Laufwerke
Haltbarkeit 10-20 Jahre
Speicherung etwas aufwendiger, weil ein Medium nur 50 GB fasst.
- **Cloudspeicher**
Der Anbieter gewährleistet die Datensicherheit, aber teurer.
Und es ist nicht gewiß, wie lange es den Anbieter gibt.
Man muss sich bewußt sein, dass Regierungen oder dessen Erfüllungsgehilfen immer Zugriff auf alle Daten haben. Daten besser verschlüsseln. Zudem ist unklar, ob die Erben überhaupt Zugriff bekommen.

Nicht zu empfehlen sind:

- Flash-Medien (USB-Sticks, SD-Karten, SSD-Festplatten)
Die Haltbarkeit beträgt nur wenige Jahre bis 10 Jahre.
Die Daten müßten ständig umkopiert werden.
- Bandlaufwerke
Hohe Kapazität, sehr lange Haltbarkeit, aber teure Laufwerke.
Es nicht gewährleistet, dass Leselaufwerke in 50 Jahren noch erhältlich sind.

Medienpreise

Hier exemplarisch einige Medienpreise (Stand 2020):

- LTO 7-Band: 12 € / Terabyte
- USB-Stick: 96 € / Terabyte
- BluRay Disc: 29 € / Terabyte
- Festplatte: 20 € / Terabyte

Empfehlung

Die Empfehlung ist klar: Externe Festplatte mit einer Kapazität etwas größer als der eigene Computer. Daten nehmen zu und irgendwann ist der Platz im Computer erschöpft und man wird aufrüsten. Es wäre ärgerlich, wenn die externen Medien ebenfalls getauscht werden müßten. Der Preis für externe Festplatten ist gering und es sollte gleich in zwei Laufwerke investiert werden.

5. Lagerung

Der wichtigste Rat: Immer mindestens eine Kopie außer Haus lagern! Die Archivierung soll ja auch dem Schutz vor Brand, Löschwasser, Einbruch, Vandalismus und Naturkatastrophen dienen! Dies bedeutet konkret: Wenn man eine externe Festplatte zur Neubespielung aus dem Schließfach holt, muss es noch mindestens ein weiteres Medium außer Haus geben!

Auf gar keinen Fall darf eine externe Festplatte ständig mit dem Computer verbunden sein! Sie ist leichte Beute für Verschlüsselungstrojaner.

Medien sollten bei Zimmertemperatur und normaler Luftfeuchtigkeit gelagert werden. Schließfächer bei Banken sind ebenfalls gut geeignet. Bei Self-Storage-Anbietern sollte nachgefragt werden, ob im Winter geheizt wird.

Beispiele für Lagerorte:

- Dachboden: Im Sommer zu warm, im Winter zu kalt. Kein Schutz vor Brand.
- Keller: Im Winter evtl. zu kalt. Kein Schutz vor Brand.
- Auto: Zu warm/zu kalt. Ständige Temperaturschwankungen senken die Lebensdauer. Diebstahlgefahr.
- Schließfächer: gut geeignet

6. Verschlüsselung

Für die Langzeitarchivierung für die Nachkommen muss von einer Verschlüsselung dringend abgeraten werden, aus den gleichen Gründen wie für die Leseprogramme selbst.

Für die Datensicherung oder Lagerung für 10-20 Jahre macht eine Verschlüsselung Sinn, da man dann die Medien auch im Büro oder bei Freunden zur Lagerung unterbringen kann. Das Verschlüsselungsprogramm und evtl. eine Anleitung muss unbedingt mit auf das Medium kopiert werden.

Von externen Festplatten mit dem Versprechen, selber die Daten zu verschlüsseln, sollte Abstand genommen werden. Zu viele Sicherheitslücken sind in der Vergangenheit bekannt geworden.

Zur Verschlüsselung kommt nur ein Programm in Betracht: VeraCrypt. Es ist der Nachfolger des bekannten TrueCrypt und gilt heute, 2020, als Stand der Technik. Es ist nachgewiesen sicher. Experten haben die Sicherheit des Programms attestiert. VeraCrypt erfordert Einarbeitung.

7. Die Zukunft von Dateitypen

Die Frage, ob ein bestimmter Dateityp / ein Dateiformat in 50 oder 100 Jahren noch maschinenlesbar ist, ist ein Blick in die Kristallkugel. Für die Vorhersage gibt es aber Indizien:

- Wie viele Hersteller gibt es, die das Format unterstützen?
- Wie alt ist das Format bereits?
- Kann das Dateiformat notfalls direkt mit einem Textprogramm wie dem einfachen „Windows Notepad“ gelesen werden?

Nach diesen Kriterien gibt die App Handlungsempfehlungen für einzelne Formate. Für Formate, die die App nicht kennt, empfiehlt sich eine Suche im Internet. Hier ein paar Beispiele:

- **TXT**
Eines der ältesten und absolut robusten Dateiformate.
Es handelt sich meistens um Textdateien, die im ASCII-Code kodiert sind. Dieses Format gibt es seit mindestens 70 Jahren und es wird wegen seiner Einfachheit problemlos noch in 100 Jahren lesbar sein.
- **HTML**
Das am weitesten verbreitete Textformat im Internet.
Selbst mit einem einfachen Texteditor können HTML-Dateien gelesen werden, auch wenn eingebettete Bilder und Formate (CSS) nicht mehr vorhanden sind.
Sonderzeichen wie deutsche Umlaute sind problemlos lesbar (ä = das deutsche kleine ä)
Wird ebenfalls in 100 Jahren noch problemlos lesbar sein.
- **PDF**
Das am weitesten verbreitete Format für Text und Bilder mit Layout. Es gibt eine für Langzeitspeicherung vorgesehene Variante namens PDF/A. PDF-Dateien lassen sich nicht mehr mit einem einfachen Texteditor lesen.
Wird ebenfalls in 100 Jahren noch problemlos lesbar sein, weil es garantiert Anzeigeprogramme geben wird.
- **JPG**
Das am weitesten verbreitete Bildformat. Alle heutigen Digitalkameras beherrschen bzw. erzeugen dieses Format.
Bilddateien lassen sich nicht mit einem einfachen Texteditor lesen.
Aufgrund der weltweiten Verbreitung wird es auch in 100 Jahren noch Anzeigeprogramme geben.
- **PNG, BMP, PCX, TIF, AIFF**
Eine Auswahl weiterer heute gebräuchlicher Bildformate. Aufgrund der Variantenvielfalt ist davon abzuraten, auf diese Formate zu setzen. Teilweise sind es Nischenformate, teilweise herstellereigene Formate.
- **DOC, DOCX, XLS, XLSX, PPT, PPTX**
Heute gebräuchliche Office-Formate. Diese und noch viele mehr sind proprietäre Formate von Microsoft. Ob diese in 20 Jahren noch lesbar sind, ist fraglich. Am ehesten werden wohl DOCX-Dateien lesbar sein, weil es ein ZIP-komprimiertes Archiv mit reinem Text ist. Den Inhalt kann man zur Not in 200 Jahren mit einem Notepad noch lesen. Dennoch ist zu empfehlen, in TXT oder PDF zu wandeln. Problematisch sind eingebettete Bilder.

- ODT
Open Office-Dateiformate. Sie sind zwar ISO-standardisiert, das schützt sie aber nicht davor, dass es in 50 Jahren kein Anzeigeprogramm mehr gibt. Die Empfehlung: ebenfalls in TXT oder PDF wandeln. Problematisch sind ebenfalls eingebettete Bilder.
- RAW-Dateien von Digitalkameras
RAW-Formate sind hersteller-eigene Formate. Sie haben eine sehr kurze Haltbarkeit und sind auch nicht für die Langzeitspeicherung gedacht. Auf jeden Fall in JPG umwandeln.
- Bildformate von Bildbearbeitungsprogrammen
Hier gilt das gleiche wie für RAW-Dateien. Sie werden bald mangels Leseprogrammen nicht mehr anzeigbar sein.
- MP3, WAV
Audioformate. Das robusteste Format ist WAV, weil es unkomprimiert und eine einfache Abfolge von Samples ist. Das am weitesten verbreitete Format ist MP3. Beide Formate werden auch in 100 Jahren noch hörbar sein.
- MP2, OGG, FLAC, AU
Herstellereigene und Nischenformate für Musik. Es ist davon abzuraten, diese Formate zu archivieren. Die Empfehlung: in WAV (verlustfrei) oder MP3 mit höchster Qualität wandeln.

8. Archivierung von Leseprogrammen

Die Idee, für die Dateiformate die benötigten Leseprogramme mit auf Archivmedien zu sichern, ist eine gute Idee für die nächsten 20 Jahre. Darüber hinaus bietet es keinen Wert, aus mehreren Gründen:

- Zum einen beinhalten viele Programme Lizenzprüfungen, die in 20 Jahren entweder aufgrund des Alters eine abgelaufene Lizenz feststellen oder aufgrund der Tatsache, dass der Hersteller den Lizenzserver nicht mehr betreibt.
- Der andere Grund ist die ständige Weiterentwicklung der Betriebssysteme. Ein Beispiel: In der Anfangszeit des Betriebssystems Windows waren 16-Bit-Programme der Stand der Technik („Windows 95“). Diese Programme lassen sich heute auf Windows 10 gar nicht mehr ausführen, und dies ist nur 25 Jahre her.

9. Funktionen

Einstellungen

Hier müssen zunächst die Archivordner festgelegt werden, die das eigene Archiv bilden.

Analysieren

Diese Funktion steht am Anfang und macht eine Bestandsaufnahme des Archivs. Es durchsucht das Archiv und gibt eine Übersicht über die vorkommenden Dateitypen und die Menge. Zu jedem Dateityp zeigt die App eine Handlungsempfehlung. Diese ist wie schon erwähnt als grobe Richtung und kein Dogma zu sehen.

Prüfinformationen

Erstellt für jede Datei im Archiv eine Prüfinformation, einen sogenannten Hash. Diese können später zur Prüfung verwendet werden, ob die Datei verändert wurde. (durch einen Datenträgerdefekt). Die Funktion erstellt eine Datei, in der alle Archivdateien mit ihren Prüfinfos aufgelistet werden. Die Datei wird mit im Archiv abgelegt.

Auf Medium sichern

Kopiert alle Archivdateien auf eine externe Festplatte. Für den Einsatz von DVDs oder BluRays muss ein Herstellerprogramm oder frei erhältliches Programm wie Nero Burning Rom verwendet werden.

Differenzen finden

Vergleicht die Dateien im Archiv mit denen auf einem externen Datenträgermedium. Hier wird es Unterschiede geben, da man in der Regel im Archiv ja Daten ändert und hinzufügt.

Medium prüfen

Prüft einerseits die Lesbarkeit der Dateien auf einem externen Datenträger und andererseits die Prüfinformationen, sofern vorhanden. Der Lesetest liest schlicht alle Dateien auf dem Medium einmal. Dies hat für Festplatten (nicht für DVDs oder BluRays) auch einen Mehrwert: Date werden unter Umständen repariert.

Um zu verstehen, wie das funktioniert, muss man wissen, wie eine Festplatte arbeitet. Im Innern einer Festplatte ist ein kleiner Computer, wir nennen ihn zur Unterscheidung mit dem „großen“ Computer Controller. Dieser verarbeitet alle Lese- und Schreibaufträge des Computers. Die Speicherschicht der Platte ist in kleine Blöcke aufgeteilt. Wenn ein solcher Block nicht mehr gelesen werden kann, gibt der Controller das Problem an den Computer weiter. Bevor der Controller aufgibt, versucht er es diverse Male. Sollte es doch noch gelingen, den Block zu lesen, kopiert der Controller die Daten in einen frischen Block, der noch intakt ist. Jede Festplatte hat hierfür reservierte Blöcke, die beim Bespielen nicht vergeben werden. Ein solche „Reparatur“ eines Leseproblems behält die Festplatte für sich und macht dies im Stillen (Nach aktuellen Wissen gibt es für uns Kunden keine Möglichkeit, hiervon zu erfahren, dies kann nur der Hersteller).

Zwingt man nun die Festplatte, alle Daten einmal zu lesen, werden sozusagen „schlecht gewordene“ Blöcke ausgesondert. Dies tut der Lagerfähigkeit für die nächsten Jahre natürlich gut.

Medium umkopieren

Hier kann ein externes Medium auf ein Neues umkopiert werden. Bitte denken sie daran, das alte ggf. zu vernichten. Oder, falls sie vorhaben, es zu verkaufen oder zu verschenken, die Inhalte sicher zu löschen! Verwenden Sie keine Schnellformatierung, da diese die Daten nicht löscht!

Medienbestand

Zeigt die verwendeten externen Sicherungsmedien.

App selbst sichern

Diese Funktion kopiert sich selbst selbst auf ein externes Medium. Dies kann der Archivdatenträger oder auch ein USB-Stick sein, da es nur darauf ankommt, die Programmdateien gegen den Defekt des eigenen Computers zu schützen (also ein Backup).

10. Umwandlung der Dateien

Zur Umwandlung gibt es nicht das eine Hilfsmittel. Dieses Kapitel gibt Tipps, wie die Dateitypen gewandelt werden können:

- Office-Dokumente (DOC, XLS, PPT)
Der Doxillion Document Konverter bietet für wenig Geld (ca. 15 Euro) die Umwandlung von vielen Dateien am Stück. Er kann mehrere Office-Formate in PDF und auch zurück wandeln.
<https://www.nchsoftware.com/>
- Office-Dokumente können mit Microsoft Office (Microsoft) und Open Office (kostenlos) ebenfalls kostenlos umgewandelt werden, aber immer nur eine Datei zur Zeit.
- Bilddateien (TIF, PCX, BMP, JPG)
Bilder können mit dem kostenlosen Irfan View mit der eingebauten Stapelverarbeitung in Massen umgewandelt werden. Irfan View beherrscht auch mehr, z.B. drehen oder Größenreduzierung.
Irfan View hat seine Stärken als Bildbetrachter und zur Sortierung. Man kann es dazu nutzen, Bilder zu sichten und bequem in bis zu 10 Ordner einzusortieren.
- Hersteller-eigene Bildformate (PSPIMAGE u.a.)
Hier ist das jeweilige Herstellerprogramm nötig
- Musikdateien (WAV, OFF, FLAC, MP3)
Hier gibt es auch viele Programme. Empfehlenswert ist der Spesoft Audio Converter. Er beherrscht Stapelverarbeitung.
<https://spesoft-audio-converter.de.uptodown.com/windows>
- Videos
Hier ist das Programm „Any Video Converter“ zu empfehlen. Er beherrscht ebenfalls

Stapelverarbeitung

<https://any-video-converter.de.uptodown.com/windows>