



GEVER
Geschäftsverwaltung für Luzern

Richtlinie Dateiformate für Langzeitarchivierung

Version 1.0 vom 23. Februar 2010
Status: Verabschiedet

1. Einleitung	3
2. Grundlagen	3
3. Dateiformate	5

Diese Vorgaben entsprechen der Weisung des Staatsarchivs über Mindestanforderungen an Geschäftsverwaltungssysteme vom 1. September 2009, gestützt auf das Luzerner Archivgesetz (SRL Nr. 585).

Inkraftsetzung

Die vorliegende Richtlinie ist für alle GEVER-Installationen der kantonalen Verwaltung verbindlich.



.....
Dr. Markus Hodel
Staatsschreiber



.....
Irene Saieva
Projektleiterin GEVER

Versionen

Version	Datum	Status	Ersteller
0.1	12.11.2009	Entwurfsvorlage	EgGr
0.2	Januar 2009	Vernehmlassung	EgGr
1.0	23.2.2010	Verabschiedung	AGR GEVER

1. Einleitung

Basierend auf seinen gesetzlichen Aufgaben¹ und aufgrund von Anfragen aus Dienststellen definiert das Staatsarchiv Luzern einen Standard für archivtaugliche Dateiformate.

In GEVER-Installationen werden weitere Formate verwendet (besonders MS Office), die nicht für die Langzeitarchivierung geeignet sind.

Der vorliegende Standard soll

- die langfristige Verstehbarkeit und Nutzung der digitalen Unterlagen gewährleisten.
- den Aufwand für die Archivierung in Grenzen halten.
- zukünftige Aufwände für Erhaltungsmaßnahmen minimal halten.

Um dies zu erreichen, favorisiert das Staatsarchiv die Einschränkung der Anzahl Formate. Für die Dienststellen hat dies zur Konsequenz, dass archivtaugliche Formate so früh wie möglich im Data Lifecycle erzeugt werden sollen.

Dieser Standard dient in zweierlei Hinsicht als Handreichung:

- Datenproduktion: Im Kontakt mit den Dienststellen legt er dar, welche Formate aus archivischer Sicht im aktiven Lifecycle verwendet werden sollen. Für die Planung und Einführung neuer Systeme ist er bindend. Die dort verwendeten Dateiformate müssen archivtauglich sein bzw. müssen von der anbietepflichtigen Dienststelle gemäss dem vorliegenden Standard derart aufbereitet werden können, dass sie im Staatsarchiv ohne zusätzlichen Aufwand archivierbar sind.
- Datenarchivierung: Bei Angeboten digitaler Unterlagen zur Archivierung legt er dar, welche Formate archivtauglich sind und in welche Formate migriert werden soll.

2. Grundlagen

Der vorliegende Standard wurde im Staatsarchiv erarbeitet. Er basiert weitgehend auf Arbeiten von:

- KOST (Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen). Deren Katalog archivischer Dateiformate ist zu finden unter <http://www.kost-ceco.ch/wiki/whelp/KaD/> (weiterführend betr. Hintergründe, Auswahlkriterien etc.)
- BAR (schweizerisches Bundesarchiv). Deren Archivtaugliche Dateiformate – Standards für die Archivierung digitaler Unterlagen sind zu finden unter <http://bar.admin.ch>, Themen > Beratung > Digitale Archivierung (weiterführend betr. Anforderungen etc.)
- Best practice anderer Archive und internationale / ISO-Standards

¹ [aus Archivgesetz, SRL 585] §5, Abschnitt 2 [Die öffentlichen Organe] berücksichtigen bei der Beschaffung von Hilfsmitteln für die Unterlagenverwaltung, namentlich von elektronischen Datenverarbeitungssystemen und Geschäftskontrollen, die Bedürfnisse der Archivierung. [aus Archivverordnung, aus SRL 586] §2, Abschnitt 3 Das Staatsarchiv ist [...] bei der Planung elektronischer Datenverarbeitungssysteme und Geschäftskontrollen frühzeitig beizuziehen.

- Weitere, im Web frei verfügbare Ideen und Formulierungen, die wir nicht einzeln nachgewiesen haben.

Für eine vertiefende Lektüre empfehlen wir: nestor Handbuch – Eine kleine Enzyklopädie der digitalen Langzeitarchivierung, 2. Auflage 2009, zu finden unter <http://nestor.sub.uni-goettingen.de/handbuch>.

3. Dateiformate

Katalog archiverischer Dateiformate, die für die Langzeitarchivierung benutzt werden können.

Anwendungsbereich	Archivtaugliche Formate
Text	PDF/A TXT
Bild	TIFF JPG / JPEG2000 (nach Absprache)
Tabellen	PDF/A CSV XLS/XLSX (nach Absprache)
Datenbanken	CSV PDF/A SIARD (nach Absprache)
Audio	WAVE
Video	nach Absprache

Text

PDF/A:

PDF/A ist als Format für die Langzeitarchivierung sehr geeignet. Textdateien sollen bereits bei der Dossierschließung in das PDF/A-Format umgewandelt werden. Für Dokumente mit archivwürdigem Layout muss darauf geachtet werden, dass alle Inhalte, insbesondere Bilder, Schriften und Farbräume, eingebettet werden.

TXT:

Nur-Text erfüllt die meisten archivischen Anforderungen maximal (mit der wesentlichen Ausnahme der Layout-Erhaltung) und ist deshalb als Archivformat unbestritten.

Bild

Für die Archivierung von Rastergrafiken wird allgemein die Verwendung von TIFF als Archivformat empfohlen. Für bestehende Bildsammlungen kommt JPG / JPEG2000 ebenfalls in Frage. Für Vektorgrafiken soll abgeklärt werden, ob eine Umwandlung in PDF/A oder TIFF sinnvoller ist.

Tabellen

Wenn bei der Archivierung von Tabellenkalkulationsdaten die Bewahrung der Funktionalität als essentiell betrachtet wird, muss im Originalformat archiviert werden, also in der Regel in XLS/XLSX. Dies kann auch allgemein als provisorische Lösung betrachtet werden, bis ein zukünftiges, besser geeignetes Archivformat vorliegt.

Wenn nur die optische Erscheinung, nicht aber die Funktionalität, als essentielle Eigenschaft gilt, können Arbeitsblätter in tabellenform (Spreadsheets) in PDF/A konvertiert werden. Wenn nur die Werte der Tabelle erhalten werden sollen, können Spreadsheets in CSV konvertiert werden.

CSV

Grosse Datenbestände sind bereits in diesem Format archiviert worden, deshalb wird das CSV-Format für Tabellen seine Bedeutung behalten. Da jedoch keine Formeln in diesem Format festgehalten werden können, werden nur gut dokumentierte Daten in CSV-Format ihren Wert behalten.

PDF/A:

PDF/A ist ein geeignetes Archivformat; allerdings nur, wenn auf die Bewahrung der Funktionalität verzichtet werden kann.

XLS/XLSX

Trotz der archivischen Bedenken (proprietäres Format bzw. übermässig komplexe Spezifikation) kann XLS/XLSX als vorläufiges Archivformat empfohlen werden. Wenn die Funktionalität eines Spreadsheets als essentiell angesehen wird, ist eine vorläufige Archivierung in einem Originalformat wie XLS/XLSX oder OOXML sogar unabdingbar. Dokumente im Excel-Format sollen in diesem Fall so übernommen oder allenfalls in das aktuelle Excel-Format konvertiert werden. Für Dokumente in obsoleten Tabellenkalkulations-Formaten empfiehlt sich eine Konvertierung ins Excel-Format.

Datenbanken

Die Archive werden im Augenblick nicht darum herum kommen, Datenbanktabellen in CSV-Format zu übernehmen. Bei geeigneter Übernahme von Metainformationen zur archivierten Fachanwendung können diese dennoch von einigem Wert sein. In Zukunft werden sich wahrscheinlich umfassendere Lösungen wie SIARD vom Schweizerischen Bundesarchiv etablieren können.

Wenn nur die Werte der Tabelle oder einzelne Reports erhalten werden sollen, können Datenbanken in CSV oder PDF/A exportiert / konvertiert werden.

CSV

Grosse Datenbestände sind bereits in diesem Format archiviert worden, deshalb wird das CSV-Format für Tabellen seine Bedeutung behalten. Da jedoch keine Beziehungen, Metadaten und Strukturinformationen in diesem Format festgehalten werden können, werden nur gut dokumentierte Datenbestände in CSV-Format ihren Wert behalten.

PDF/A:

PDF/A ist ein geeignetes Archivformat; allerdings nur, wenn auf die Bewahrung der Funktionalität verzichtet werden kann.

SIARD

SIARD bietet eine neue Möglichkeit für die Archivierung von relationalen Datenbanken. Es implementiert eine einfache und international akzeptierte SQL99 / XML-Kodierung für die langfristige Aufbewahrung von Archivdaten.

Audio

Für die Archive empfiehlt es sich, vorläufig die anfallenden Audiodateien unkomprimiert abzulegen. Dafür kann das verbreitete WAV-Format (RIFF WAVE) verwendet werden. Wenn sich ein verlustfrei komprimiertes Audioformat durchgesetzt oder wenigstens eine gewisse Verbreitung gefunden hat, wird sich dieses als Archivformat aufdrängen. Als Kandidat kommt der ISO-Standard MPEG-4 Audio Lossless Coding ALS (ISO/IEC 14496-3:2005/Amd 2:2006) in Frage.

WAV:

Aus archivarischer Sicht kann das WAV-Format zur Verwendung empfohlen werden. Die Audiodateien weisen eine hohe Qualität auf, sind seit langem in Gebrauch, weit verbreitet und weitgehend von spezifischer Hardware und Betriebssystemen unabhängig. Die Formatspezifikation ist bekannt und nicht

durch Lizenz- oder Patentansprüche der Urheberfirmen eingeschränkt. Der Nachteil des Formates liegt eindeutig in der Grösse der Dateien.

Video

Für die Archivierung von Videodaten muss erwogen werden, ob es der Umfang der zu archivierenden Videodaten erlaubt, diese in unkomprimierter oder verlustfrei komprimierter Form zu speichern. Besonderes Augenmerk muss möglichst langen Migrationszyklen gelten. Deshalb ist für Videodaten die Migrationsvermeidung höher zu gewichten als die Reduktion der Formatanzahl im Archiv. Mindestens provisorisch können mehr Formate als archivtauglich akzeptiert werden als für andere Formatkategorien.

MJPEG2000:

Wegen seiner Vorteile gegenüber anderen Videoformaten, insbesondere der verlustfreien, spatialen Kompression, kommt MJPEG2000 auf jeden Fall als Archivformat für Videodaten in Frage.

MPEG2

Bestehende DVD werden provisorisch in diesem Datei-Format (als Image, nicht auf diesem Datenträger) aufbewahrt.

Richtlinie: Erlaubte Dateiformate

Allgemein

Gesetzliche Grundlagen	<ul style="list-style-type: none"> • Archivgesetz und Archivverordnung (siehe oben) • Weisung "Standards für archivtaugliche Dateiformate im Kanton Luzern" vom 1.7.2009
Bezugsobjekt	<ul style="list-style-type: none"> • Dateiojekt
Zeitlich Bestimmung	<ul style="list-style-type: none"> • Dateiobjekte werden bei der Dossierschliessung automatisch ins archivische Dateiformat migriert.
Bemerkungen	<ul style="list-style-type: none"> • Aus praktischen Gründen wird zwischen Dateiformaten in den Organen (Aktenführung) und Dateiformaten im Archiv unterschieden. Die Migration in die archivischen Dateiformate muss sichergestellt sein. • Die Richtlinie muss periodisch auf ihre technische Aktualität überprüft werden.