

# Handreichung für die Archivierung von digitalen Daten aus Kommunalverwaltungen

Peter Worm, LWL-Archivamt für Westfalen  
Stand: 07.09.2011, Version: 0.4

---

1. VORAB ZU KLÄRENDE FRAGEN .....	2
2. ALTVERFAHREN BEWERTEN .....	2
3. VORGABEN FÜR DIE DATENHALTUNG .....	3
3.1. Die Ablage in sog. CSV-Dateien .....	3
3.2. Die Speicherung als XML-Datei .....	4
3.3. Die Speicherung im Langzeitformat PDF/A.....	4
3.4. Anforderungen an die Datenträger .....	4
4. MITWIRKUNG BEI DER ANSCHAFFUNG NEUER SOFTWARE .....	5
4.1. Einflussmöglichkeiten bei der Softwareanschaffung.....	5
4.2. Dokumenten-Management-Systeme .....	6
5. WELCHE HILFESTELLUNGEN GIBT DAS LWL-ARCHIVAMT? .....	6

---



Für die Menschen.  
Für Westfalen-Lippe.

**Diese Handreichung richtet sich an Kommunalarchive, die den Einstieg in das Thema „Archivierung elektronischer Unterlagen“ suchen und dabei mit den sogenannten Fachverfahren beginnen wollen.**

## **1. Vorab zu klärende Fragen**

Gemeinden, Städte und Kreise erledigen in zunehmender Weise ihre Aufgaben unter Zuhilfenahme der elektronischen Datenverarbeitung. Schon seit den 60er und 70er Jahren des letzten Jahrhunderts haben sich organisatorische Strukturen herausgebildet, die die Fachabteilungen bei der Auswahl und dem Betrieb von Hard- und Software unterstützen.

Wichtig ist es, die in Ihrer Kommune herrschenden Organisationsstrukturen gut zu kennen – Gibt es eine TUIV- oder IT-Abteilung in Ihrer Stadt oder Ihrem Kreis? Gibt es eine Anbindung an ein Kommunales Rechenzentrum? Wie eng ist die Verbindung zwischen diesen Strukturen, arbeitet die Kommune u.U. sogar mit mehreren Rechenzentren für unterschiedliche Verwaltungszweige oder Aufgabenfelder zusammen?

Suchen Sie das Gespräch, finden Sie feste Gesprächspartner für die Fragen der elektronischen Archivierung. Beachten Sie, dass IT-Fachleute unter "Archivierung" oft etwas anderes verstehen als Archivare: Für erstere ist jeder Datenbestand, der sich nicht mehr in der täglichen Nutzung befindet, ein "Archivbestand". Auch komprimierte Daten werden oft als "Archiv" bezeichnet (ZIP, RAR ...). Zur Unterscheidung davon hat sich der Begriff "Langzeitarchivierung" oder "Endarchivierung" etabliert. Auch wenn er für uns Archivare "doppelt-gemoppelt" scheint, ermöglicht er doch eine Differenzierung.

## **2. Altverfahren bewerten**

Da archivische Belange bei der Beschaffung von Programmen naturgemäß nicht im Vordergrund stehen – schließlich geht es der Verwaltung um eine effizientere Aufgabenerledigung – wurden bisher keine Vorkehrungen für die Aussonderung und Langzeitsicherung von Datenbeständen getroffen. Archivare haben sich oft nur theoretisch mit dem Thema beschäftigt, auf die entsprechenden Paragraphen des Archivgesetzes verwiesen und gehofft (oder gefürchtet), dass die IT-Verantwortlichen auf sie zukommen würden.

Die so entstandenen "Altlasten" müssen in einem ersten Schritt angepackt werden. Dafür ist es notwendig, eine möglichst vollständige Liste der Fachanwendungen zu erstellen. Hierbei können die IT-Abteilungen oder die Rechenzentren behilflich sein. Auch der Datenschutzbeauftragte hat i.d.R. eine Verzeichnisliste für die Software, in der personenbezogene Daten vorgehalten werden. Dann muss ein Bewertungsprozess stattfinden, der die archivrelevanten von den nicht relevanten Anwendungen trennt. Anders als bei der Bewertung von traditionellem Registraturgut, das meist erst nach dem Ablauf der gesetzlichen Aufbewahrungsfristen auf seine Archivwürdigkeit geprüft wird, muss der Archivar oder die Archivarin bei EDV-gestützten Verfahren schon im Vorgriff eine Bewertung vornehmen. Im Einzelfall ist dazu eine enge Zusammenarbeit mit den betroffenen Fachabteilungen nötig.

Generell gelten zusätzlich zu den sonst angewendeten Bewertungsgrundsätzen folgende Leitlinien:

- Archivwürdig sind Daten aus Anwendungen, die archivwürdige Papierakten komplett ersetzen. Man spricht hier vom "papierlosen Büro" oder sog. Dokumenten-Management-Systemen – kurz DMS – oder Enterprise Content Management – kurz ECM.
- Ähnliches gilt für die elektronischen Verfahren, die ein Papierregister / Karteikartensystem ersetzen (z.B. im Einwohnermeldewesen oder beim Gewerberegister).
- Archivwürdig sind Daten aus Anwendungen, die wichtige Teile archivwürdiger Papierakten komplett ersetzen. Es entstehen sog. "Hybridakten", deren Verzahnung auch im Archiv erhalten werden muss.

- Archivwürdig sind Hilfsdatenbanken, die bei der Verwaltung entstanden sind und die auch für die Archivarbeit hilfreich erscheinen. Das kann beispielsweise eine zentral gepflegte Übersicht der Prozesse sein, die die Stadt oder der Kreis geführt haben, oder eine digital erstellte Zeitungsausschnittssammlung.
- Schließlich kann es sinnvoll sein, eine Grundgesamtheit in digitaler Form abzubilden und recherchierbar zu halten und gleichzeitig das Bild der exemplarisch übernommenen Papierakten zu komplettieren. Im Stadtarchiv Stuttgart werden beispielsweise elektronische Personalisten aus der Abrechnungsstelle mit den Personal-Kerndaten aller Arbeiter, Angestellten und Beamten übernommen und durch die exemplarische Übernahme einzelner (papierener) Personalakten ergänzt.
- Nicht archivwürdig sind auch in Papierform nicht aufbewahrenswerte Daten (z.B. Verfahren für die Vergabe von Kfz-Kennzeichen, Hundesteuer- oder Personalressourcenverwaltung).
- Nicht archivwürdig sind Programme, die die Bescheiderstellung unterstützen. Kriterium ist hier: Gibt es weiterhin eine Papierakte?
- Nicht archivwürdig sind Einzelschreiben in elektronischer Form, sprich E-Mails!
- Nicht archivfähig sind Unterlagen, die in die Kompetenz der Staatsarchive fallen (z.B. statistische oder Wahldaten, die an IT.NRW gemeldet werden).

Hat man die archivrelevanten Programme identifiziert, geht es an die Einzelbewertung der Datenfelder. Nicht jeder Eintrag bzw. jede Datentabelle ist erhaltenswert. Eine Verringerung bei den Datenmengen und eine Vereinfachung der Datenstruktur erleichtern die Erhaltung und ggf. eine spätere Wiederzugänglichmachung der archivierten Daten. Hier gilt es in Zusammenarbeit mit den Fachabteilungen und ggf. mit den IT-Fachleuten eine geeignete Aussonderungsschnittstelle zu definieren. Es empfiehlt sich, auf die Vorarbeiten von anderen Kolleginnen und Kollegen zurückzugreifen. In manchen Fällen wird die Programmierung hausintern möglich sein, vielfach wird aber auf die Herstellerfirma der Software zugegangen werden müssen. Es ist klar, dass hier Kosten entstehen können!

Die vollständige Umkopierung in neuere Programmversionen (sog. Datenmigration) kann eine kurzfristige Lösung sein, um Daten lesbar und recherchierbar zu halten. Langfristig ist es m.E. unumgänglich die Altdaten auszusondern, um sie zu sichern und die laufende Programmversion performant zu halten.

Insgesamt gilt: Nicht Programme, sondern die in ihnen enthaltenen Daten sind archivwürdig. Für die Langzeitsicherung ergeben sich die folgenden Vorgaben für die Datenhaltung:

### **3. Vorgaben für die Datenhaltung**

Eine Speicherung der Archivdaten in einem proprietären – d.h. nur von einer Spezial-Software lesbar zu machendem – Format ist nicht empfehlenswert. Für Datenbankinhalte existieren zwei Speichervarianten:

#### **3.1. Die Ablage in sog. CSV-Dateien**

(Character Separated Value = Zeichen getrennte Werte):

Die Datensätze werden in ASCII-Kodierung oder bei Vorkommen von Sonderzeichen in Unicode abgespeichert. Die so entstehende Textdatei speichert jeden Datensatz in einer neuen Zeile; die Werte untereinander werden durch Semikola oder ein anderes Feldtrennzeichen voneinander unterschieden. Da der so entstehende Datenfluss sich nur erklärt, wenn man die aufeinander folgenden

den Felder und ihre Inhalte erklären kann, braucht man ein dazu gehöriges Codebuch. Hier werden die Verkettungen der Datentabellen erläutert, die Tabellenspalten betitelt und evtl. vorhandene Zahlencodes aufgelöst (z.B. "0" für männlich, "1" für weiblich o.ä.).

### 3.2. Die Speicherung als XML-Datei

(eXtensible Markup Language = erweiterbare Auszeichnungssprache):

Die Daten werden i.d.R. in einer Unicode-Kodierung (UTF-8) abgelegt, wobei die Daten hierarchisch strukturiert abgelegt werden. Dieses Format versucht in großen Teilen selbsterklärend zu sein und benötigt deshalb nur eine geringere Dokumentation.

Ein Beispieldatensatz könnte wie folgt aussehen:

```
<Datensatz>
  <Datensatz-Nummer>1</Datensatz-Nummer>
  <Name>Müller</Name>
  <Vorname>Gerhard</Vorname>
  <Straße>Talweg 27</Straße>
  <Ort>Münster</Ort>
  ....
</Datensatz>
```

Durch die Vorteile, die diese Speicherung bietet, hat sich XML im zwischenbehördlichen Datenaustausch auf breiter Front durchgesetzt. Bundesweit einheitliche XML-Schnittstellen werden XÖV-Standard genannt und auf einer für jedermann zugänglichen Internetplattform („XRepository“<sup>1</sup>) veröffentlicht. Es gibt sie inzwischen für zahlreiche Verwaltungsfelder; oft eignen sie sich als Grundlage für eine Archivschnittstelle.

### 3.3. Die Speicherung im Langzeitformat PDF/A

Die Firma Adobe hat in Zusammenarbeit mit amerikanischen und europäischen Nationalarchiven und Fachleuten eine PDF-Variante entwickelt, dessen Kodierung als ISO 19005-1:2005 jedem zugänglich ist. Selbst wenn es also in ferner Zukunft keine entsprechenden Ansichtsprogramme ("Viewer" wie der Acrobat Reader) mehr gibt, wird es Programmierern möglich sein, einen solchen Viewer nachzuprogrammieren und die Informationen wieder sichtbar zu machen. Das Format eignet sich besonders für die Speicherung von Plänen (da es Vektorgrafiken bis 6x6m abbilden kann) und für die Sicherung von Dokumenten aus DMS. Anders als bei dem früher gängigen TIFF – einem reinen Bildformat – bleibt im PDF Text als solcher erhalten und durchsuchbar. Die 2011 verabschiedete Version 2 des PDF/A erlaubt auch sogenannte Layer (eine Art „elektronische Klarsichtfolie“, die über eine Hintergrundgrafik gelegt werden kann), was besonders für Kartenmaterial eine wichtige Funktion darstellt.

### 3.4. Anforderungen an die Datenträger

Die Speicherung im sichersten Format nutzt nichts, wenn der Datenträger, auf dem die Dateien liegen, nicht mehr lesbar ist. Da es noch keinen unbegrenzt haltbaren Speicher gibt, arbeitet man bei der Datensicherung heutzutage i.d.R. mit SAN-Systemen (Storage Area Networks), die die zu sichernden Daten auf unterschiedlichen Speichern (Festplatten, WORM-Medien, Bandsystemen ...) an räumlich getrennten Orten zur Verfügung halten. Wichtig ist eine redundante Datenhaltung, d.h. dass die gleiche Datei an mehreren Stellen vorgehalten wird. Falls eine Datei nicht mehr lesbar ist, können die anderen Kopien an ihre Stelle treten und genutzt werden. Ein Mittel zur raschen Überprüfung der Datenunversehrtheit stellen Prüfsummen (z.B. sog. MD5-Summen) dar, die eine Art Quersumme aus einer Datei bilden. Wird nur ein Byte willentlich oder unwillentlich an der Datei ver-

---

<sup>1</sup> [www.xrepository.deutschland-online.de/](http://www.xrepository.deutschland-online.de/).

ändert, passt die so erzeugte Prüfsumme bei einem Testlauf nicht mehr und die veränderte Datei kann gegen eine unversehrte Fassung ausgetauscht werden. Sicherung auf optischen Datenträgern (CD, DVD) ist nicht empfehlenswert<sup>2</sup>. Wenn es keine sicherere Alternative gibt, sollte man die Daten dann zumindest auf drei Datenträger kopieren, die an unterschiedlichen Orten bei möglichst geringer Luftfeuchte und kühlen Temperaturen aufbewahrt werden. Stehende Lagerung der Medien verhindert, dass sie sich verbiegen und dadurch unbrauchbar werden.

Hohe Datensicherheit und einen Schutz vor ungewollter Datenveränderung bieten sogenannte CAS-Systeme (**C**ontent **A**dressed **S**torage), die auch die Anforderung einer revisionssicheren Speicherung erfüllen können (z.B. Centera, NetApp u.ä.). Dieser Speicher ist noch relativ teuer, aber für die Langzeitarchivierung am besten geeignet.

Speichersysteme, die im Bereich Langzeitarchivierung verwendet werden, sollten nach dem internationalen Standard OAIS (**O**pen **A**rchival **I**nformation **S**ystem, ISO 14721:2003<sup>3</sup>) betrieben werden und dem nestor-Kriterienkatalog vertrauenswürdige digitale Langzeitarchive Version II / hrsg. von der nestor-Arbeitsgruppe Vertrauenswürdige Archive – Zertifizierung<sup>4</sup> entsprechen.

## 4. Mitwirkung bei der Anschaffung neuer Software

### 4.1. Einflussmöglichkeiten bei der Softwareanschaffung

Die nachträgliche Programmierung von Aussonderungsschnittstellen benötigt Zeit und Geld. Es ist deshalb sinnvoller, dass im Vorfeld einer Softwareanschaffung das Archiv konsultiert wird um die folgenden Fragen zu klären:

- Betrifft die Softwareanschaffung einen "archivkritischen" Bereich?
- Wenn ein Softwarewechsel in einem solchen Bereich stattfindet, was passiert mit den Altdaten?
- Verfügt das neue Programm über eine archivgerechte Schnittstelle zur Datenaussonderung?
- Falls nicht, wie müsste diese Schnittstelle beschaffen sein?

Hat man diese Vorgaben gemacht, können schon bei der Ausschreibung und den folgenden Verhandlungen mit dem Softwarehersteller die richtigen Weichenstellungen vorgenommen werden.

Besondere Aufmerksamkeit ist geboten bei Systemen, die jetzt schon eine Vorgangsbearbeitung unterstützen, bei denen aber noch Papierakten geführt werden. Besonders starke Trends zur elektronischen Akte gibt es in allen Massenverfahren (Sozial- und Jugendsachen) sowie in den Bauverwaltungen (fast alle heutzutage erstellten Pläne werden schon elektronisch mit sog. CAD-Systemen erzeugt; auch bei Baugenehmigungen läuft vieles über sog. E-Government-Lösungen auf rein elektronischem Weg). Eine Archivierungspflicht besteht für die Kreise bei der Sicherung der Katasterunterlagen, die nach 1948 entstanden sind – hier ist die Umstellung von ALK (Automatisierte Liegenschaftskarte) und ALB (Automatisiertes Liegenschaftsbuch) auf das sog. ALKIS (Amtliche Liegenschaftskataster-Informationssystem) nahezu abgeschlossen.

---

<sup>2</sup> The Relative Stabilities of Optical Disc Formats by JOE IRACI, in: Restaurator 26 (2005).

<sup>3</sup> <http://public.ccsds.org/publications/archive/650x0b1.pdf>; neuere Fassung, derzeit noch „Draft“-Status: <http://public.ccsds.org/sites/cwe/rids/Lists/CCSDS%206500P11/CCSDSAgency.aspx>.

<sup>4</sup> nestor-Kriterien - Kriterienkatalog vertrauenswürdige digitale Langzeitarchive Version II / hrsg. von der nestor-Arbeitsgruppe Vertrauenswürdige Archive – Zertifizierung (nestor-materialien 8) Frankfurt am Main : nestor c/o Deutsche Nationalbibliothek, 2008. - 40 S. URN: urn:nbn:de:0008-2008021802

## 4.2. Dokumenten-Management-Systeme

Wird auf einen elektronischen Posteingang (Einscannen der Post) und eine elektronische Postverteilung umgestellt, so spricht man von einem **DMS** (s.o.). Diese Systeme unterstützen die Arbeitsabläufe im Haus ("Workflow"), indem sie die Laufmappen und Botendienste ersetzen. Besitzt man eine "intelligente Scanstelle", so werden formalisierte Anträge und Schreiben automatisch erkannt und dem richtigen Sachbearbeiter zugewiesen. Wurde das DMS gut in die bestehende Softwarearchitektur integriert, werden bestimmte Angaben aus den gescannten Formularen sofort in die entsprechenden Felder der Fachanwendungen übergeben und so die Bearbeitung vereinfacht. Die zweite Aufgabe eines solchen DMS ist die Datenspeicherung und –archivierung im Rahmen der gesetzlichen Aufbewahrungsfristen. Es muss deshalb logische Strukturen der Schriftgutverwaltung nachbilden (Akte – Vorgang – Dokument) und den Sachbearbeiter zu einer Ablage nach Aktenplan anhalten. Da erst detaillierte Aktenpläne ein solches System handhabbar und langfristig lauffähig und performant halten, ist gute Planung und Vorbereitung das A und O<sup>5</sup>. Fachabteilungen und Archive als "Schriftgut-Sachverständige" sollten die Produktauswahl und –einführung deshalb eng begleiten und für Akzeptanz werben. Gute Hinweise und einen ausführlichen Anforderungskatalog bietet das Bundesministerium des Innern herausgegebene DOMEA-Konzept 2.1<sup>6</sup>. Im Laufe des Jahres 2011 soll der Nachfolger des DOMEA-Konzepts unter der Bezeichnung „Organisationskonzept Elektronische Verwaltungsarbeit“ vorgestellt werden.

## 5. Welche Hilfestellungen gibt das LWL-Archivamt?

Auf unserer Homepage [www.lwl-archivamt.de](http://www.lwl-archivamt.de) finden sich in der Rubrik „Archiv und IT“ weiterführende Hinweise, Links und Handreichungen u.a. zum Thema „Einwohnermeldedaten-Archivierung“, Anforderungen an Verzeichnungsprogramme und Langzeitarchiv-Software sowie zur elektronischen Registerführung im Personenstandswesen. Es lohnt also ab und zu ein Blick auf neue Inhalte!

Darüber hinaus bieten wir die Mitarbeit in regionalen Arbeitskreisen zum Themenkreis „Archivierung elektronischer Unterlagen“ an, hier macht es Sinn auch die zuständigen Rechenzentren in die Arbeit einzubinden. Und wir bieten Fortbildungsveranstaltungen z.B. zur Rolle von Archiven bei der Einführung von DMS an.

Schließlich gibt es natürlich auch die Möglichkeit der Einzelberatung unserer Mitgliedskommunen.

---

<sup>5</sup> Orientierung bieten Musteraktenpläne der Kommunale Gemeinschaftsstelle für Verwaltungsmanagement (KGSt) oder des Städte- und Gemeindebundes (zus. mit der Fa. REGIS), die zunehmend auch die Anforderungen elektronischer Systeme berücksichtigen.

<sup>6</sup> Bundesministerium des Innern – Informationsseiten „Verwaltung innovativ“ unter <http://www.verwaltung-innovativ.de>. Eine Einführung in die Vorteile und Ziele von DOMEA finden Sie in Barbara HOEN, Das neue DOMEA-Konzept – ein Standard und die Praxis. In: Texte und Untersuchungen zur Archivpflege 19 (2006), S. 57-63.