

Amtliche Daten – für die Ewigkeit?

Von Thomas Mehner und Peter Schliwa, Magdeburg

Zusammenfassung

Seit dem Ende der 1990er Jahre wurde in Sachsen-Anhalt damit begonnen, die bisher in „handgefertigter“ analoger Form vorliegenden Dokumente des amtlichen Vermessungswesens zu digitalisieren. Mittlerweile steht der Umstieg auf die zweite technische Plattform an. Bedenkt man die aus der wesentlich kürzeren Haltbarkeit digitaler Medien heraus resultierenden Gefahren und stellt diese wesentlich günstigeren Eigenschaften hinsichtlich Zugriffsmöglichkeiten und „Datenlagerung“ gegenüber, befindet man sich mitten in einer Technologiediskussion. Der Artikel stellt die aktuellen Entwicklungen bei der Erneuerung des Fachdokumentenmanagementsystems vor.

I Die Dokumente des amtlichen Vermessungswesens

Wie alle Liegenschaftskataster auf dem Gebiet der heutigen Bundesrepublik Deutschland hat auch das preußische Grund- und Gebäudesteuerkataster seine Wurzeln in einer Reform der Grundsteuer. Dieser Anlass der Katasterentstehung – die Forderung nach einer gerechteren Verteilung von Steuern – wurde später durch weitere Zweckaspekte ergänzt: Die inzwischen als klassische Hauptfunktion eingeordnete Sicherung des Grundeigentums, des Grundstücksverkehrs und der Ordnung von Grund und Boden sowie die Basisfunktion, vor allem für Bedürfnisse der Landesplanung, Bauleitplanung, Grundstückswertermittlung, Umwelt- und Naturschutz. Alle diese Aufgaben haben eines gemeinsam: Sie sind Bestandteil der staatlichen Daseinsvorsorge. Dies zu Beginn eines eher technisch geprägten Beitrages zu betonen, erscheint wichtig, um herauszustellen, welche Bedeutung und Wertigkeit die heute „zu Bits und Bytes verkommenen“ Unterlagen haben.

Kein Selbstzweck – staatliche Daseinsvorsorge.

Die Digitalisierung der Dokumente war dabei anfänglich keine Selbstverständlichkeit. Vielmehr wurden im Jahr 1998 in einer Projektvorstudie drei Lösungsalternativen untersucht [Elsner 1999]. Neben der Möglichkeit der Nulllösung – alles bleibt analog, wie es ist – wurde dabei auch die Mikroverfilmung mit analogen und digitalen Rechercheinstrumenten erwogen. Schließlich setzte sich die vollständig digitale Führung auf Basis eines Dokumentenmanagementsystems (DMS) als wirtschaftlichste Lösung im Verlauf der Projektbearbeitung durch. Es wurde eine Amortisation der neuen Verfahrenslösung nach vier bis fünf Jahren erwartet und dies trotz der erheblichen Kosten für die Ersterfassung der Dokumentenbestände [Elsner, Frommhagen 2000].

Die Ersterfassung für die primären 12 Dokumententypen wurde bis 2006 abgeschlossen. Das Scannen der Dokumente der sogenannten Stufe I ist mittlerweile also erledigt, ebenso die Ersterfassung der Stufe 2. Die Dokumente der sogenannten Stufe 3 werden seit Ende 2006 erfasst, wobei mit einem Gesamtzeitaufwand von rund 4,5 Jahren zu rechnen ist. Die Ersterfassung findet dabei zentral in einem „Scan-Pool“ statt. Dieser hat zur Zielsetzung, die datentechnischen Voraussetzungen dafür zu schaffen, dass mit Abschluss der Ersterfassung der dritten Stufe alle für

das Fachdokumentenmanagementsystem vorgesehenen Dokumente in die digitale Form überführt sind [Fehse, Krimmling 2008]. Die ständig aktuell hinzukommenden Dokumente werden seit Jahren lokal direkt nach ihrer Entstehung digital erfasst.

Erfassungsstufe	erfasste Dokumente	Zeitablauf	
		Scannen	Indizierung
Stufe 1	<ul style="list-style-type: none"> ◆ Dokumente über die Veränderung (z.B. Fortführungsriss, Niederschrift über den Grenztermin) ◆ bedeutende historische Katasterdokumente (z.B. Separationskarte) ◆ Dokumente der Grundlagenvermessung 	erfasst bis Ende II. Quartal 2006	bis Ende 2011
Stufe 2	<ul style="list-style-type: none"> ◆ mit der Erneuerung außer Kraft gesetzte Blätter der analog geführten Liegenschaftskarte ◆ Rahmenkartenblätter, die Liegenschaftsvermessungen enthalten, die nicht nach VV LiegVerm durchgeführt worden sind 	erfasst bis Ende I. Quartal 2006	Ende 2006 abgeschlossen
Stufe 3	<ul style="list-style-type: none"> ◆ weitere historische Katasterdokumente und Katasterbücher 	seit III. Quartal 2006	seit 2006

Tab. 1: DMS-Erfassungsstufen
[LVermGeo 2008]

Nachdem die Dokumente gescannt sind, werden sie in einer Technologiekette mit Attributen versehen, also in eine fachliche Zuordnungslogik gebracht, und schließlich im Dokumentenmanagementsystem gespeichert. Diese Indizierung der Dokumente ist die Voraussetzung dafür, die gescannten „Rohdaten“ später im System wieder auffindbar zu machen. Dabei werden die eigentlichen Dokumente systemlogisch getrennt von den Attributdaten gespeichert (Indexdatenbank). Die Indizierung der oben genannten Bestandsdaten macht dabei den wesentlichen Zeitanteil an der vollständigen Inbetriebnahme des Fachdokumentenmanagementsystems aus. Die vollständige Indizierung über alle Erfassungsstufen für die gesamte Landesfläche wird somit noch geraume Zeit in Anspruch nehmen.

Trotz des insgesamt langen Zeitrahmens vom Beginn der Erfassungstätigkeiten bis zur landesweiten Betriebsfähigkeit werden für die digitale Führung von Beginn an solide Mengenabschätzungen benötigt, um die Rechen- und insbesondere Datenhaltungskapazitäten sachgerecht dimensionieren zu können. Erste Abschätzungen gingen für die Stufe I von rund 4,1 Millionen Dokumenten (Blättern) aus [Elsner, Frommhagen 2000]. Aktuell stellt sich das prognostizierte Mengengerüst wie in Tab. 2 zu sehen dar. Die erste technische Realisierung (Abschnitt 2) ging dabei im Ansatz von den Werten der Stufe I aus, während die aktuelle Konsolidierung (Abschnitt 3) bereits die weiteren Dokumentenklassen berücksichtigt. Diese Lösung muss dabei auch so flexibel sein, die ständig wachsenden Datenmengen durch aktuell hinzukommende Dokumente zu beherrschen.

DMS	Anzahl Dokumente (Stufe I)	Speicherbedarf in GB (Stufe I)	Anzahl Dokumente (vollständig)	Speicherbedarf in GB (vollständig)
Salzwedel, Stendal	563 504	763	1 860 000	1 900
Magdeburg, Haldensleben, Staßfurt, Wernigerode	1 142 060	2 210	2 410 000	3 300
Halle	1 211 761	1 913	3 260 000	3 700
Dessau-Roßlau	1 068 153	1 565	1 880 000	2 300
LVerGeo gesamt	3 985 479	6 452	9 410 000	11 200

Tab. 2: Mengengerüst und Speicherbedarf

Bei dem aktuell durchgeführten Konsolidierungsvorhaben ist also von der hier prognostizierten Datenmenge von rund 12 TeraByte ausgegangen worden. Anforderung war außerdem, ein skalierbares System zu erwerben, um einen zukünftig möglicherweise größeren Speicherbedarf durch Hinzufügen einer skalierten Speichereinheit abbilden zu können.

2 Der Beginn der digitalen Führung der Fachdokumente

Bei der Auflegung des ersten digitalen Systems zur Führung der Fachdokumente standen im Wesentlichen drei fachliche Anforderungen im Fokus, die es abzubilden galt:

- ◆ dauerhafte Bestandssicherung der Originale vor Verfall und Verschleiß,
- ◆ vereinfachte und beschleunigte Benutzung der Fachdokumente und
- ◆ Verminderung von Medienbrüchen.

Dies sind Anforderungen, die aus Sicht der Informationstechnik dem Begriff der Langzeitdokumentation zugeordnet werden können. Unter **Langzeitdokumentation** versteht man allgemein die Erfassung, die langfristige Aufbewahrung und die Erhaltung der dauerhaften Verfügbarkeit von Informationen. Bei der Langzeitdokumentation digitaler Informationen ist hierbei nicht nur die sichere Speicherung der Dokumente von Bedeutung, sondern gleichgewichtig auch Fragen der Erstellung der Dokumente und Attribute sowie deren Verwaltung.

Geht man also den im Projekt eingeschlagenen Weg der Langzeitdokumentation digital vorliegender Informationen, so muss man sich mit völlig anderen Themenkomplexen auseinandersetzen, als im analogen Zeitalter. So reicht es für die geforderte dauerhafte Bestandssicherung nicht aus, die Garantieerklärung der Medienhaltbarkeit eines Herstellers über 10 oder auch 100 Jahre zu erhalten. Vielmehr müssen verantwortliche und vorausschauende Strategien entwickelt werden, um den beständigen, durch permanente Innovation des informationsverarbeitenden Marktes verursachten Wandel bewältigen zu können.

Für die Bewältigung der Aufgaben einer digitalen Langzeitdokumentation werden in der Regel spezielle Dokumentenmanagementsysteme eingesetzt. Diese bestehen in der Regel aus einer Software, Datenbanken und Speichersystemen. Der fachliche Ansatz für diese Systeme basiert meistens auf dem Verweis über eine Referenzdatenbank mit den Verwaltungs- und Indexinformationen, die auf einen externen Speicher zeigen, in dem die Informationsobjekte gehalten werden.

Diese sogenannte **Referenz-Datenbank-Architektur** ist notwendig, um große Mengen von Informationen von den relativ schnellen, aber kostenintensiven externen Speichern zu extrahieren. Die Datenbank erlaubt, über die Indexinformationen jederzeit das Dokument wieder zu finden und mit einem entsprechenden fachbezogenen Programm dem Anwender bereitzustellen.

Anforderungen an die Langzeitdokumentation digitaler Informationen

Systeme für digitale Langzeitdokumentation müssen u.a. folgende Anforderungen erfüllen:

- ◆ programmgestützter, direkter Zugriff auf einzelne Informationsobjekte (Dokumente) oder Informationskollektionen, z.B. Listen,
- ◆ datenbankgestützte Verwaltung der Informationsobjekte auf Basis von Metadaten und gegebenenfalls Volltexterschließung der Inhalte der Informationsobjekte,
- ◆ Verwaltung von Speichersystemen mit nur einmal beschreibbaren Medien einschließlich des Zugriffs auf Medien die sich nicht mehr im Speichersystem direkt befinden,
- ◆ Sicherstellung der Verfügbarkeit der gespeicherten Informationen über einen langen Zeitraum, der Jahrzehnte betragen kann,
- ◆ Bereitstellung von Informationsobjekten unabhängig von der sie ursprünglich erzeugenden Anwendung auf verschiedenen Clients und mit Übergabe an andere Programme,
- ◆ Absicherung der gespeicherten Informationen gegen unberechtigten Zugriff und gegen Veränderbarkeit der gespeicherten Information,
- ◆ eigenständige Wiederherstellungsfunktionalität (Recovery), um inkonsistent gewordene oder gestörte Systeme aus sich heraus verlustfrei wieder aufbauen zu können,
- ◆ sichere Protokollierung von allen Veränderungen an Strukturen und Informationsobjekten, die die Konsistenz und Wiederauffindbarkeit gefährden können und dokumentieren, wie die Informationen im System verarbeitet wurden,
- ◆ Unterstützung von Standards für die spezielle Aufzeichnung von Informationen auf Speichern mit WORM-Verfahren, für gespeicherte Dokumente und für die Informationsobjekte beschreibende Metadaten um eine langfristige Verfügbarkeit und die Migrationssicherheit zu gewährleisten sowie
- ◆ Unterstützung von automatisierten, nachvollziehbaren und verlustfreien Migrationsverfahren.

In Anlehnung an diese Anforderungen und im Ergebnis eines europaweiten Ausschreibungsverfahrens setzte sich schließlich folgende Lösung durch:

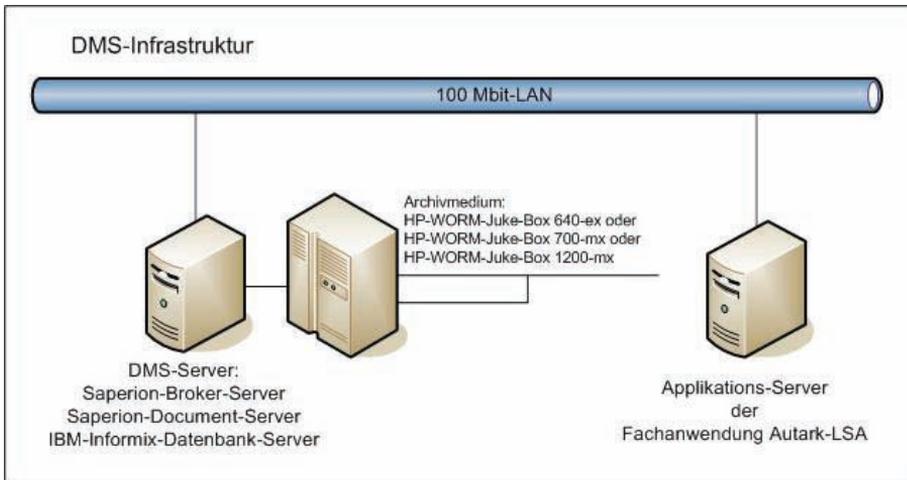


Abb. 1: DMS-Infrastruktur je Kataster (ab 2001)

- ◆ Dokumentenmanagementsoftware: Saperion der Firma Saperion,
- ◆ Referenz- und Indexdatenbank: IBM-Informix der Firma IBM,
- ◆ Erstellung der Fachapplikation nach Anforderungskatalog des LVermGeo: Autark-LSA der Firma HL GIS.

Unter Hinzuziehung des Mengengerüsts für das zu haltende Datenvolumen war dabei festzulegen, welche Medien und Speichertechnologie unter Beachtung der vom Hersteller der Dokumentenmanagementsoftware vorgegebenen Kompatibilitäten, angewendet werden sollte. Dabei war zu berücksichtigen:

- ◆ Welche Zugriffszeiten sind im Mittel zu erbringen?
- ◆ Wie hoch ist die Zahl gleichzeitiger Zugriffe im Mittel?
- ◆ Welche Aufbewahrungsfristen sollen durch das Medium abgedeckt werden?

Zum Zeitpunkt der Einführung waren folgende Medien verfügbar, die für eine Speicherung genutzt werden konnten:

Medium	Kapazität	Verwendung	Revisionssicherheit
CD-R	5,25 Zoll - 700 MB	kleine Datenbestände	ja
DVD-R	5,25 Zoll - 4 GB	mittelgroße Datenbestände	ja
ISO 9171-WORM	5,25 Zoll - 9,1 GB	mittelgroße bis große Datenbestände	ja

Tab. 3: verfügbare Medien im Jahr 2001

WORM-Medien – „Datengarantie“ für 30 Jahre!(!?)

Die wesentlichen Unterschiede zwischen den Medien bestehen in der Verarbeitung zuverlässigerer Materialien und erweiterten Garantieerklärungen der Hersteller. Diese garantieren für einmal beschreibbare Medien eine Datenstabilität zwischen 10 und 100 Jahren und für WORM-Medien zwischen 30 und 100 Jahren, je nach Hersteller und unter jeweiliger Vorgabe optimaler Einsatzbedingungen. Unter weiterer Beachtung des zum Einführungszeitpunkt vorliegenden Mengengerüsts und einer Hochrechnung der Entwicklung des Datenbestandes über mindestens fünf Jahre im Voraus wurde die Anwendung von WORM-Medien fixiert (HP Surestore OpticalJukebox von Hewlett-Packard).

Der ursprüngliche Projektansatz, je katasterführender Stelle (ehemaliges Katasteramt) ein DMS einzuführen, konnte infolge haushaltswirtschaftlicher Rahmenbedingungen und innerer struktureller Änderungen nicht durchgehend umgesetzt werden. Aktuell sind die in Tab. 2 aufgeführten acht „Einzel-DMS“ an den vier Standorten des LVerGeo in Betrieb. Sie sind per WAN über das Informationstechnische Netz des Landes (ITN-LSA) vernetzt. Die Vernetzung gewährleistet für alle genannten DMS eine zentrale Systemadministration, ein zentrales Verfahrensmanagement und ermöglicht jedem autorisierten Benutzer den Zugang zu den Daten eines jeden DMS, sofern dies erforderlich ist.

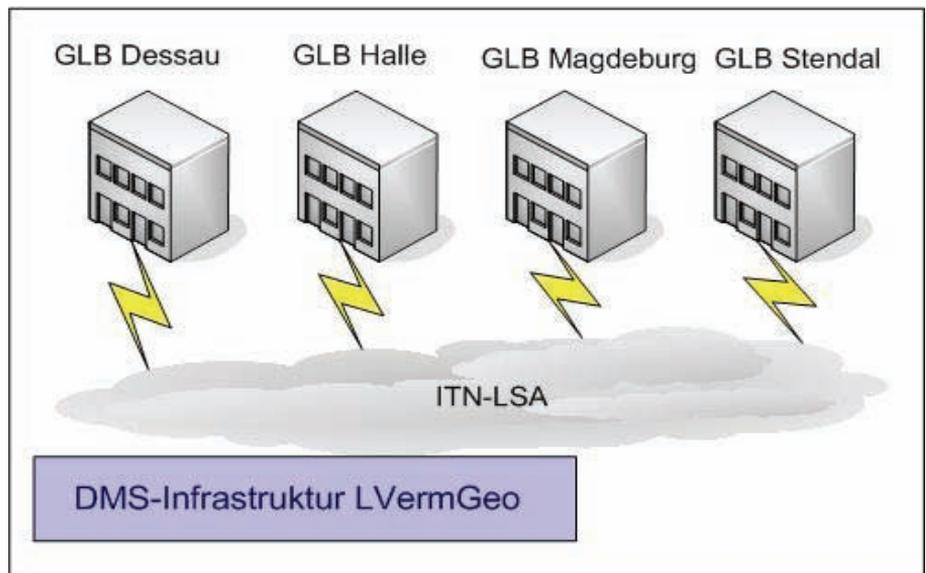


Abb. 2: dezentrale
Infrastruktur (2001 bis 2009)

Mittlerweile haben die DMS des LVerGeo eine Betriebsdauer von sechs bis acht Jahren erreicht. Dabei müssen stetig steigende Aufwendungen, sowohl für die Systemadministration, das Verfahrensmanagement und steigende finanzielle Aufwendungen, festgestellt werden, die auf unterschiedlichen Ursachen basieren:

- ◆ Teilweise keine Ersatzteilversorgung im Serverbereich mehr, da der Hersteller nicht mehr am Markt ist.
- ◆ Steigende Kosten für Reparaturen der Medienspeicher HP Surestore OpticalJukebox.
- ◆ Meldungen, die darauf hinweisen, dass Dokumente nicht mehr von den Medien extrahiert werden können, mit zunehmender Tendenz.

Während die ersten beiden Punkte mit entsprechenden finanziellen Aufwendungen behoben werden können, so bedroht der letztgenannte Fakt die Verfügbarkeit der Daten. Die Analysen und Tiefenprüfungen der Systemadministration und des Verfahrensmanagements bestätigen in der Regel diese Meldungen und lassen nur die eine Schlussfolgerung zu, dass das Dokument nicht mehr vom Medium gelesen werden kann.

Bedrohung der Verfügbarkeit

Weist also das Medium WORM eine „Gedächtnisschwäche“ auf, trotz einer definierten Datenstabilität für WORM-Medien von 30 Jahren? Leider bestätigen die aufgetretenen Fälle diese Frage und es ist jeweils eine zeitaufwendige Restauration von den zugehörigen Kopiemedien notwendig. Warum abgelegte Dateien nicht mehr lesbar sind, kann nur vermutet werden. Die physikalische Beschädigung des Mediums nach wiederholtem Zugriff des Speichers HP Surestore OpticalJukebox scheint die wahrscheinlichste Ursache zu sein.

Stellt man die Zunahme der Schwierigkeiten mit den Medien und die erreichte Betriebsdauer der gesamten DMS-Infrastruktur sachlich der für die amtlichen Daten bestehenden „Ewigkeitsforderung“ gegenüber, so gelangt man zu dem Schluss, dass es notwendig ist, eine Konsolidierungsmaßnahme umzusetzen. Diese wurde im Jahr 2008 begonnen.

Es ist Zeit für eine Konsolidierung der digitalen DMS-Bestände des LVerMGeo.

3 Der technologische Wandel

Jedes Konsolidierungsvorhaben setzt geeignete Zieldefinitionen voraus. Selbstverständlich gelten hierbei an erster Stelle die unter Abschnitt 2 genannten allgemeinen Anforderungen an die Langzeitdokumentation fort. Daneben sind Rahmenbedingungen zu beachten, die weniger dem technischen Umfeld zuzuordnen sind:

- ◆ Reduzierung des Aufwandes für die Systemadministration und für das Verfahrensmanagement (Personalrückgang),
- ◆ Einheitlicher Zugriff auf alle Dokumente und deren Reproduktion von allen Standorten des LVerMGeo (Zusammenführung),
- ◆ Prüfung, Anwendung und Migration von Verfahrenskomponenten auf strategische Softwareausrichtungen der Landesverwaltung Sachsen-Anhalt (Standardisierungsansatz) und
- ◆ Kostenreduzierung durch Integration in die vorhandene IT-Infrastruktur (Sparsamkeit).

Aus den genannten Fakten ließen sich nun konkrete Vorgaben für die informationstechnische Ausprägung des erneuerten DMS ableiten:

- ein System** ♦ Betrachtet man die rasante Technologieentwicklung der letzten Jahre und berücksichtigt die äußeren und inneren Strukturänderungen im LVerGeo, so ist die Führung der bisher separaten digitalen DMS-Bestände in einem LVerGeo-weiten DMS-Bestand anzustreben. Dies wird gestützt durch die Erfahrung der Erstinvestition, dass die Realisierung des Speichers (Backend-Storage-System), einschließlich der Medien, mit ca. 50 % der Gesamtinvestitionskosten der kostenintensivste Bestandteil der Ausstattung einer digitalen Langzeitdokumentation war. Somit wird eine maximierende Kosteneinsparung durch eine Führung der Liegenschaftskatasterakte in einem zusammengeführten digitalen Datenbestand erreicht, da die Kosten für vier Systeme bei einer geoleistungsbereichsweisen Führung wesentlich höher liegen würden, als die Kosten für ein gemeinsames System mit großer Speicherkapazität.
- Virtualisierung und Hochverfügbarkeit** ♦ Die Erneuerung der Servertechnik erfolgt unter der Maßgabe der Erhöhung von Flexibilität, Skalierbarkeit und Verfügbarkeit durch die Anwendung von Virtualisierungstechnologien. Damit war begleitend auch die Einführung von Hochverfügbarkeitstechnologien zur Gewährleistung der Verfügbarkeit, des nach der Konsolidierung zusammengeführten Verfahrens notwendig. Die Nichtverfügbarkeit des Fachverfahrens beträfe andernfalls dann nicht mehr nur einen Standort, sondern alle Standorte.
- neue Technologie** ♦ Aufgrund der Negativerfahrung soll die Ablösung der WORM-Technologie für die Langzeitdokumentation der Daten durch zeitgemäße, innovative Speicher bzw. Technologien erfolgen.
- neues DBMS** ♦ Die beschriebene Referenz-Datenbank-Architektur wird im Wesentlichen beibehalten, die DMS-Software Saperion bleibt ebenso in Betrieb, wie die aufgesetzte Fachapplikation Autark-LSA. Allerdings wird das zugrunde liegende Datenbankmanagementsystem (DBMS) gewechselt. Die Migration der Referenz- und Indexdatenbank vom relationalen Datenbankmanagementsystem IBM-Informix zu Oracle entspricht einer Vorgabe der Landesleitstelle IT-Strategie zur Einführung und Durchsetzung einheitlicher Softwarestandards und deren Anwendung bei der Konsolidierung von Fachverfahren.

Nach der Sicherstellung der erforderlichen Mittel wurde im Jahr 2008 ein europaweites Ausschreibungsverfahren durchgeführt. Die beschriebenen Kriterien waren dabei von potenziellen Bietern zu erfüllen und mit konkreten technischen Realisierungsvorschlägen zu hinterlegen. Das Ausschreibungsergebnis hat schließlich zum Zuschlag für folgende Lösung geführt (Abb. 3).

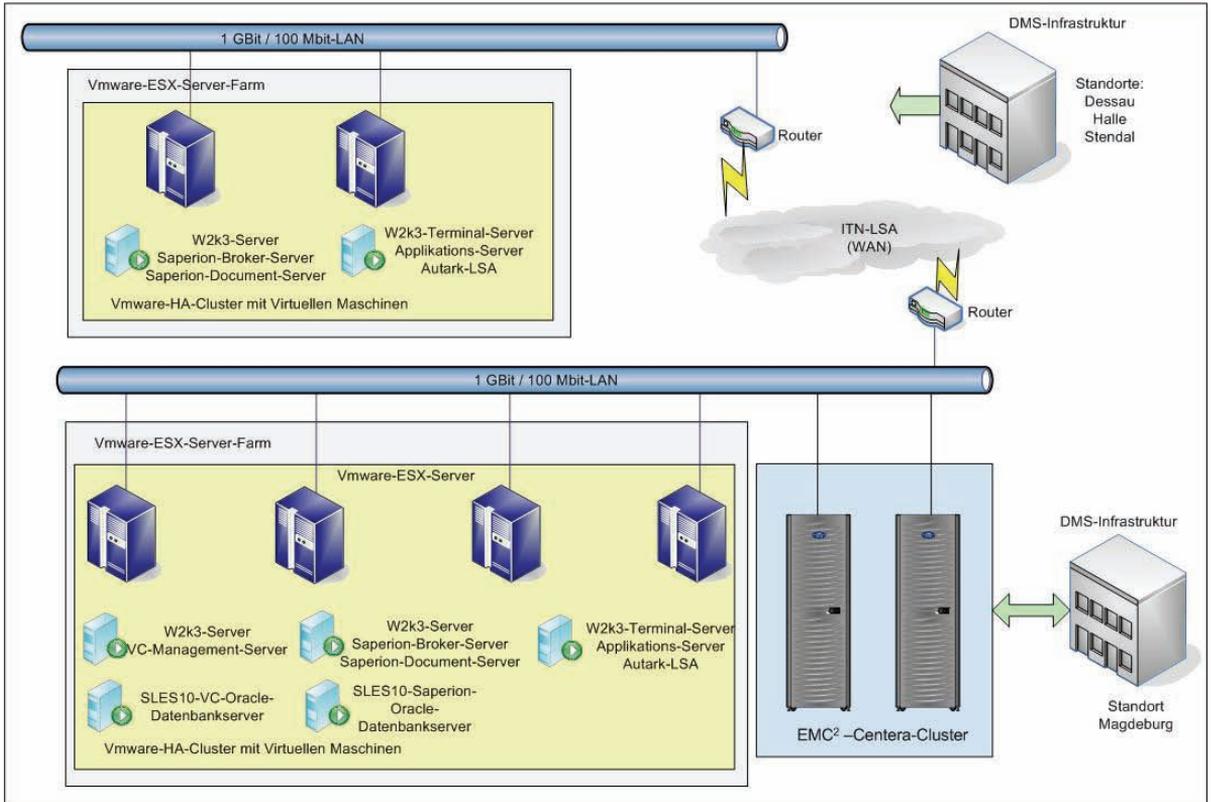


Abb. 3: Konsolidierte DMS-Infrastruktur (2008)

Wesentliche Neuerung der Gesamtarchitektur ist die Realisierung einer aktuellen Technologie und damit verbunden der Einsatz neuer Medien. Die eingesetzte Technologie wird als CAS (Content Addressed Storage) bezeichnet. Sie ist besonders geeignet zur Speicherung von Fixed-Content, also von unveränderlichen Daten. Diese Form des Content liegt beim DMS vor, da es sich um fertig gestellte Dokumente handelt, die bei Änderungen in Form neuer Revisionen abgelegt werden, wobei die ursprünglichen Dokumente erhalten bleiben müssen.

CAS arbeitet dabei mit Objekten mit festem Inhalt und nicht mit Datenblöcken oder Dateien. Soll ein Dokument gespeichert werden, überträgt die Anwendung das Datenobjekt an das CAS-System. Dieses berechnet dann eine sogenannte Content-Adresse (CA), also eine Art Schlüssel, die um weitere CA für die Metadaten ergänzt wird. Die CA werden an die Anwendung zurück gegeben und schließlich in der Datenbank für die Zukunft gespeichert. Der Zugriff auf die Objekte erfolgt dann ausnahmslos über die generierten CA.

CAS-Speichersysteme basieren hinsichtlich der Speichermedien auf Festplatten. Durch spezielle Software wird erreicht, dass die (eigentlich überschreibbaren und somit in der Vergangenheit für Langzeitaufbewahrung nicht relevanten) Medien die gleichen Eigenschaften wie herkömmliche WORM-Medien erreichen. Durch die Codierung bei der Speicherung und die spezielle Adressierung (CA) wird ein Überschreiben oder Ändern der Information auf dem Speichersystem verhindert. Bei

den CAS-Speichersystemen handelt es sich um abgeschlossene Subsysteme, die allerdings nahezu wie herkömmliche Festplattensysteme direkt in die IT-Umgebung integriert werden können. Sie bieten Speicherkapazitäten mit hoher Performance bis in den TeraByte-Bereich.

Für die Anwendung von CAS-Technologie wurde im Ausschreibungsverfahren abgesichert, dass CAS durch die Dokumentenmanagementsoftware „Saperion“ unterstützt wird. Die Saperion-Kompatibilitätsliste lässt neben der bekannten WORM-Technologie auch die Nutzung einer CAS-Technologie durch eine Integration zum CAS-Speichersystem Centera der Firma EMC² (Saperion-EMC²-Centera-Integration) zu. EMC²-Centera ist eine kombinierte Hard- und Software-Lösung für Content Adressed Storage. Die Prüfung der Parameter und die gute Skalierbarkeit der Speicherkapazität der EMC²-Centera führten zur Anwendungsentscheidung für dieses CAS-System als Speichermedium in der Konsolidierung der digitalen Langzeitdokumentation im LVerGeo.

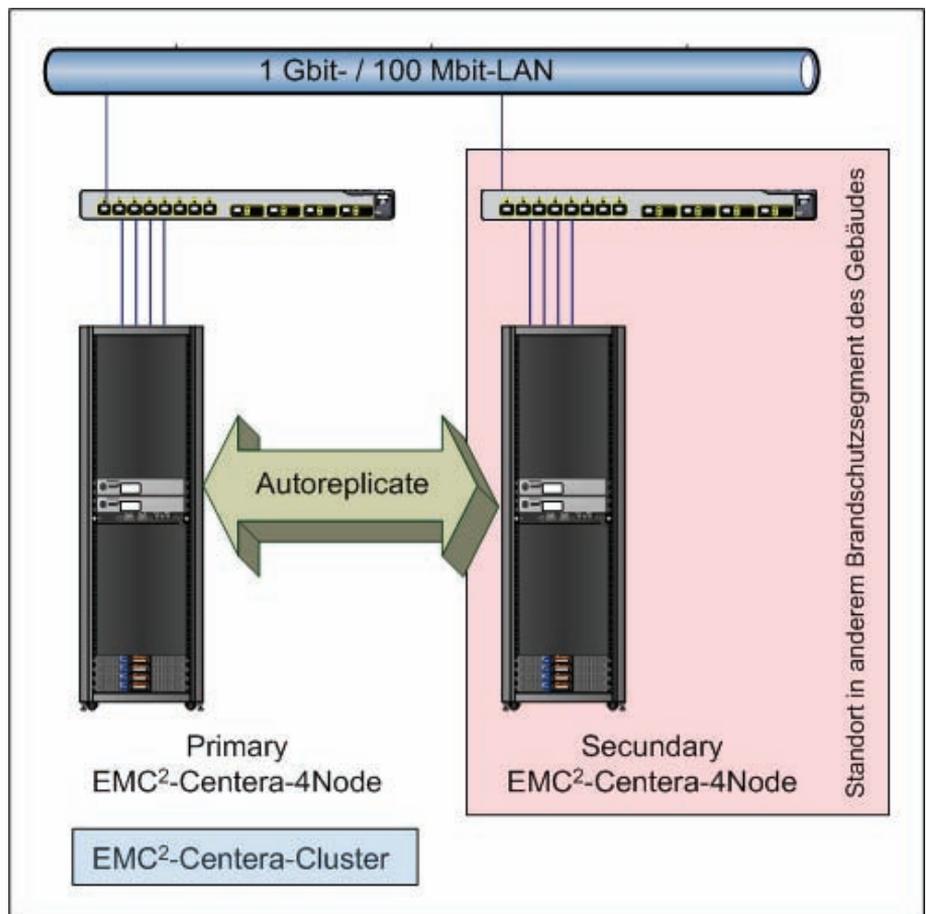


Abb. 4: Realisierung des EMC²-Centera-Clusters als konsolidiertes Zielsystem

Realisiert wurde letztendlich eine 4-Node-EMC²-Centera mit einer Bruttospeicherkapazität von 12 TeraByte. Dies entspricht gemäß Tab. 2 dem Vollausbau für die aktuell geplanten Erfassungsstufen. Ein Nachbestücken zusätzlicher Nodes für später zuwachsende Datenmengen ist möglich. Die Backup- und Disaster-Recovery-Absicherung wird über eine zweite EMC²-Centera realisiert, die die Daten durch Anwendung der asynchronen Replikation spiegelt.

Das neue zusammengeführte System hat dabei die folgenden Funktionen, die gegenüber bisherigen Ansätzen verschiedene Vorteile zeigen:

Funktionen	Vorteile
automatische Konfiguration, automatische Fehlerbehebung und automatisches Management	Kostensenkung, da weniger Managementaufwand und Ressourcen erforderlich als bei Band-, optischen oder herkömmlichen Festplattenlösungen
Single-Instance-Speicher	erfordert weniger Speicherplatz als Alternativlösungen, da jedes eindeutige Content-Element nur einmal gespeichert wird
Content-Authentifizierung	Informationen bleiben verfügbar und unveränderlich
Online-Zugriff	schneller Zugriff auf Informationen, unabhängig von der Größe der Datenbestände
Management auf Datensatzebene	Erfüllung von Unternehmens- und Anwendungs-Policies für die Aufbewahrung und Löschung von Informationen
skalierbar auf PetaByte	nahtlose Kapazitätserweiterung ohne Unterbrechung möglich
objekt-basierter Speicher	zukunftssichere Investition, denn die Technologie veraltet nicht

Tab. 4: Systemvorteile
[EMC² 2009]

Doch wie sieht es nun mit dem Ewigkeitsanspruch bezüglich der Datenhaltbarkeit aus? Voran zu stellen ist hier, dass der Transfer der Daten von der Konfiguration aus dem Jahr 2001 auf die Konfiguration aus dem Jahr 2008 nicht trivial ist. Die Datenbestände in erheblichem Umfang müssen zuerst ausgelesen werden. Aufgrund zu geringer Bandbreiten ist der Transport zum neuen System nur auf Wechseldatenträgern und nicht direkt über das ITN-LSA möglich. Anschließend werden die Daten auf dem neuen System eingespielt, unter gleichzeitigem Vollzug des Wechsels des Datenbanksystems. Das Zuspielden der bisher lokalen Datenbestände in einen einzigen Datenbestand erfordert darüber hinaus eine fachliche Bereinigung, insbesondere an den „Grenzen“ der einzelnen Altsysteme. Die auf Tests basierende Migrationsplanung sieht dabei vor, die Daten von drei Altsystemen noch im Jahr 2009 auf die neue Plattform zu übertragen und die verbleibenden Daten im Jahr 2010. Somit wird aus heutiger Sicht eingeschätzt, dass Ende 2010 die CAS-Technologie die bisherige WORM-Technologie abgelöst haben wird. Der wesentliche Anspruch dabei ist es, das die Daten vollkommen verlustfrei im neuen System

ankommen. Dies müsste im Negativfall zu Lasten der Migrationszeiträume sicher gestellt werden.

Sind alle Daten schließlich überführt, wird die Datensicherheit in der konsolidierten Lösung primär durch die beschriebene Redundanz gewährleistet (Abb. 4), da sämtliche Datenobjekte mehrfach gespiegelt werden. Bei einem lokalisierten Fehler ist das System in der Lage, den Fehler selbst zu beheben:

- ◆ Es werden neue, (rück-)gespiegelte Kopien von den bereits gespiegelten Datenobjekten erstellt.
- ◆ Ausgefallene Daten-Objekte werden unterbrechungsfrei ersetzt.
- ◆ Zusätzlich wird die redundante Architektur durch den Anschluss von zwei separaten Stromanschlüssen unterstützt, sodass ein durchgehender, ausfallfreier Betrieb möglich ist.

Mit dieser Systemausprägung eines CAS soll erreicht werden, das Phänomen der „digitalen Gedächtnisschwäche“, die unter anderem (nach noch nicht einmal einem Jahrzehnt) zur Ablösung der WORM-Technologie führte, zu umgehen. Das durch Innovationszyklen bestehende Migrationsrisiko muss jedem Betreiber einer digitalen Langzeitdokumentation bewusst sein und veranlasst die Autoren des Artikels, abschließend einen kritischen Blick in die Zukunft zu wagen.

4 ... und später

Die Frankfurter Allgemeine beschreibt in einem Artikel [Wirwalski 2007] das Szenario, dass nach dem Untergang der Menschheit der dann leere Planet von Außerirdischen besucht würde. Nach dem Auffinden unserer medialen Artefakte hätten sie wohl keine größeren Schwierigkeiten, sich gebaute, gedruckte oder gemalte Überbleibsel vor Augen zu führen. Mit heutigen Formen von digitalen Datenträgern, so schätzt es der Autor ein, könnten sie allerdings höchstens noch Frisbee spielen.

Dieses Szenario ist aus Sicht der digitalen Langzeitdokumentation einerseits zwar eine Horrorvision, andererseits aber völlige Normalität, wenn man sich in Erinnerung ruft, dass kaum ein Computernutzer heute wohl noch in der Lage wäre, eine 15 Jahre alte 5,25-Zoll-Diskette einzulesen. Hat man sich also einmal der digitalen Langzeitdokumentation „verschrieben“, so kann man sich diesem „Frisbee-Effekt“ nicht mehr entziehen. Die Frage ist nun, wie man sich auf den „Frisbee-Effekt“ einstellen und ihm möglichst vorausschauend begegnen kann.

Ein einfacher Weg, den das obige Zitat nahezu schlussfolgernd aufdrängt, ist die Rückkehr zu den Medien, die nicht der digitalen Kurzlebigkeit unterliegen. Und interessant ist, dass von „klassischen Archivkreisen“ aktuell eine Diskussion über den Einsatz der Mikroverfilmung geführt wird. So heißt es zum Beispiel in einem Positionspapier der ARK (Archivreferentenkonferenz) NRW, dass der Mikrofilm „abgesehen von seiner bis heute unerreichten Funktion als alterungsbeständiges Sicherungsmedium – und damit auch als Digitalisierungsvorlage – außerdem als analoger Übergangsspeicher für digitale Daten dienen und somit Migrationsprobleme und -kosten reduzieren kann“ [Archivar 2008]. Zu beachten ist bei dieser Argumentation jedoch, dass beim klassischen Archivgut häufig eine der Hauptforderungen bei der Einführung des DMS, die vereinfachte und beschleunigte Benutzung der Fachdokumente, nur eine untergeordnete Rolle spielt.

Dieser „Ausweg“ ist für die „Benutzungsdokumente“ des amtlichen Vermessungswesens aus Sicht der Autoren also nicht gegeben. Vielmehr sind zwei Faktoren möglichst vorausschauend zu würdigen und ständig weiter zu beobachten:

- ◆ die Konstanz von Basistechnologien auch über den Wechsel von Technikgenerationen hinaus und
- ◆ die Marktgängigkeit von Produkten und insbesondere auch von Herstellern.

Dem erstgenannten Punkt wurde mit der Referenz-Datenbank-Architektur bereits über zwei Technikgenerationen Rechnung getragen. Der objektbasierte Speicheransatz mit Content-Authentifizierung des neuen Speichersystems ist gewählt worden, da er einen weiteren festen Bestandteil zukünftiger Lösungen darstellen soll. Dagegen ist das physikalisch hinterlegte Medium vermutlich in Kürze bereits wieder veraltet, da es sich abzeichnet, dass Festplatten von Flash-Speichern abgelöst werden. Hier wird also kontinuierlich im fünf- bis zehnjährigen Rhythmus Investitionsbedarf entstehen.

Zum zweiten Punkt lässt sich ausführen, dass aus der Sicht des Betreibers der digitalen Langzeitdokumentation alle Komponenten von am Markt platzierten Unternehmen entwickelt, hergestellt und vertrieben werden. Die Auswahl wird zuerst durch das Produktportfolio des Herstellers der Dokumentenmanagementsoftware bestimmt. Die Realisierung der Fachapplikation wiederum setzt genauso auf eine veröffentlichte Schnittstelle der Dokumentenmanagementsoftware auf, wie die Dokumentenmanagementsoftware auf die veröffentlichte Schnittstelle des Herstellers des Speichersystems aufsetzt. Die Auswahl und Festlegung für die Referenz- und Indexdatenbank ist aufgrund der führenden Marktposition des Datenbankherstellers erfolgt, ist aber Privileg des Herstellers der Dokumentenmanagementsoftware, an dem sich der Kunde zu orientieren hat.

Voraussetzung, um die erwartete permanente zyklische Erneuerung durchführen zu können, ist eine erfolgreiche, langfristige Platzierung der Unternehmen am Markt. Kritisch für den Bestand der digitalen Langzeitdokumentation könnte die Situation dann werden, wenn eines der beteiligten Unternehmen (Softwarehersteller) sich gar nicht mehr am Markt behaupten kann. Diesen Effekt kann man nur kritisch, aufmerksam und kontinuierlich beobachten...

Anschrift der Autoren **Thomas Mehner**

Landesamt für Vermessung und Geoinformation Sachsen-Anhalt
Otto-von-Guericke-Straße 15
39104 Magdeburg
E-Mail: Thomas.Mehner@LVermGeo.Sachsen-Anhalt.de

Peter Schliwa

Landesamt für Vermessung und Geoinformation Sachsen-Anhalt
Otto-von-Guericke-Straße 15
39104 Magdeburg
E-Mail: Peter.Schliwa@LVermGeo.Sachsen-Anhalt.de

Literaturverzeichnis

Elsner, J. 1999: Bearbeitungsmethodik von Entwicklungsprojekten am Beispiel „Automatisierte Dokumentenführung“, in: Zeitschrift für das Öffentliche Vermessungswesen des Landes Sachsen-Anhalt (LSA VERM), 5. Jahrgang, Heft 1, S. 29-36, Magdeburg 1999.

Elsner, J., Frommhagen, R. 2000: Automatisierte Dokumentenführung in der VuKV LSA, in: Zeitschrift für das Öffentliche Vermessungswesen des Landes Sachsen-Anhalt (LSA VERM), 6. Jahrgang, Heft 2, S. 167-184, Magdeburg 2000.

Fehse, S., Krimmling, U. 2008: Der Scan-Pool – Ein Bericht aus der Produktion, in: Zeitschrift für das Öffentliche Vermessungswesen des Landes Sachsen-Anhalt (LSA VERM), 14. Jahrgang, Heft 1, S. 69-76, Magdeburg 2008.

EMC², 2009: Funktionen und Vorteile EMC-Centera, <http://germany.emc.com/products/detail/hardware/centera.htm>, 24.04.2009.

Wirwalski, A. 2007: Zurück auf den Mikrofilm, in: F.A.Z. Nr. 65, S. 7, vom 17.03.2007.

Archivar 2008: Positionspapier der ARK „Digitalisierung von Archivgut im Kontext der Bestandserhaltung“ in: Archivar – Zeitschrift für Archivwesen, 61. Jahrgang, Heft 04 November 2008, S. 395-398, http://www.archive.nrw.de/archivar/hefte/2008/ausgabe4/ARCHIVAR-04-2008_Internet.pdf, 24.04.2009.

LVermGeo 2008: Festlegung zur Ordnung der Dokumente des Liegenschaftskatasters, der Landesvermessung und der Bodenordnung. In: Verfügung des Landesamtes für Vermessung und Geoinformation Sachsen-Anhalt vom 11.02.2008 – 3.22-23080, n.v., Magdeburg 2008.