

Rolf Lang, Kai Naumann

## **Bei Umzug Übernahme – Bewertung und Ablieferung elektronischer Unterlagen im Rahmen von Systemmigrationen**

Archivarinnen und Archivare wissen, dass die Beziehungen zur abgebenden Stelle bei einer Standortschließung, einer Zusammenlegung oder dem Bezug eines Neubaus aufblühen. Schon immer ist ein Umzug von Registraturen willkommener Anlass zur Bewertung und Übernahme. Im Projekt des Landesarchivs Baden-Württemberg zum Aufbau eines Digitalen Archivs zeigte sich, dass auch der Umzug von Altdaten auf ein neues Softwaresystem für Archivare ein guter Grund zum Handeln sein kann.

In diesem Zusammenhang werden zunächst die Übernahmeprojekte dargestellt, die beim Landesarchiv im Zusammenhang mit Systemmigrationen standen. Hieraus und aus der Beobachtung der Entwicklungsperspektiven in der IT werden wir versuchen, den Wert der genutzten Methode bestimmen. Abschließend versuchen die Referenten, die zukünftige Rolle der Methode in der elektronischen Archivierung zu ermesen.

### **I. Vorstellung der übernommenen Unterlagen**

Die im Rahmen der Systemmigration übernommenen Altdaten stammen aus datenbankbasierten Fachanwendungen. Die ersten drei Bestände wurden aus Systemen migriert, die zwischen 1968 und 1989 entstanden und deren Architektur zur Zeit seltener wird. Die am Schluss dieses Kapitels zu behandelnde Polizeiliche Kriminalstatistik stellt dagegen einen Fall dar, der in Zukunft häufiger auftreten dürfte.

Die Fachanwendung PLUS (Personalverwaltung der Lehrpersonen und Stellenbewirtschaftung, HStAS EA 3/156), die Personalangaben von allen Lehrern des Landes enthält, war im Jahr ihrer Ablösung seit über zwanzig Jahren im Betrieb. Die von den Oberschulämtern und dem Kultusministerium genutzte Software wurde 2006 stillgelegt und in das Personalverwaltungssystem DIPSY<sup>1</sup> des Landesamts für Besoldung und Versorgung (LBV) integriert. Nachdem das Landesarchiv auf die Notwendigkeit einer Archivierung elektronischer Unterlagen hingewiesen hatte<sup>2</sup>, wurde eine Beschreibung der Inhalte von PLUS abgeliefert. Die Daten wurden als vollständig archivwürdig eingestuft und anschließend abgeliefert.

Die Daten aus PLUS stellen aber keinen abgeschlossenen Informationsbestand dar. Der aktive Datenbestand zu den Lehrern des Landes wird seit 2006 regelmäßig von DIPSY auf ein Datenbanksystem der Kultusverwaltung übertragen. Man ändert die Daten dort nicht, sondern führt bestimmte Auswertungen durch, die das LBV nicht leisten kann. Zu diesem Zweck haben die Daten aus PLUS in der DIPSY-Umgebung ein Kennzeichen erhalten, das ihren Ursprung aus dem Bereich der Kultusverwaltung markiert. Diese Kennzeichnung der schulischen Landesbediensteten dürfte aus organisatorischen Gründen auf Dauer erhalten bleiben. Insofern wäre es unter den gegebenen Umständen möglich, für eine Längsschnittstudie im Jahr 2020 die Daten zur Lehrerschaft von PLUS aus 2006 mit den Daten zur Lehrerschaft von DIPSY aus 2016 zu vergleichen.

POVO (Praxisorientierte, ganzheitliche Vorgangsorganisation, HStAS EA 7/105) wurde 1989 als erstes elektronisches Registratursystem der Landesverwaltung beim Landwirtschaftsministerium Baden-Württemberg installiert. Es enthält Angaben zu allen als relevant eingestuften Postein- und -ausgängen des Ministeriums (34.000 bis 44.000 pro Jahr). Der Bestand wurde

<sup>1</sup> DIPSY steht für „Dialogisiertes Integriertes Personalverwaltungssystem“.

<sup>2</sup> Vgl. LANG/KEITEL/NAUMANN: Handlungsfähige Archive, in: K. ERNST (Hg.): *Erfahrungen mit der Übernahme digitaler Daten. Elfte Tagung des AK „Archivierung von Unterlagen aus digitalen Systemen“ vom 20./21. März 2007*, Stuttgart/Leipzig 2007, S. 10-14.

vollständig übernommen. Für 2008 ist die Einführung des Scannens eingehender Post in dem DMS-System DoRIS geplant. Damit ist POVO ein exemplarisches Beispiel für die Frühzeit des digitalen Dokumentenmanagement und hat bereits einen Überlieferungsstrang konstituiert, der sich mit zukünftigen Übernahmen elektronischer Akten aus DoRIS fortsetzen lassen wird.

Die Straßendatenbank (StAL EL 75 V) ist nach ihrer Systemarchitektur das älteste übernommene Fachverfahren des Landesarchivs. Alle Daten sind nach einem Lochkartenschema organisiert, und Lochkarten waren bei der Errichtung des Systems im Jahre 1968 das tatsächliche Speichermedium. Später kamen Magnetbänder und Festplatten zum Einsatz. Das System wurde aber nach diesem Prinzip bis ins Jahr 2000 betrieben, um dann einem Nachfolger zu weichen, der den Straßenbaubehörden endlich die Einbindung von digitalen Karten und Plänen erlaubte. Es enthielt auf dem Stand von 2000 lückenlose Angaben über alle Autobahnen, Bundes-, Landes- und Kreisstraßen in Baden-Württemberg. Unter anderem sind Gemeindezugehörigkeit des Abschnitts, Fahrspuren, Baulast, Hindernisse, Einmündungen, Bauwerke, Durchlässe, Querschnittsaufbau in Segmenten, Schichtaufbau der Segmente, Baujahr, Höhe über NN und Zählstellen der Straßenverkehrszählung vermerkt. Die Übernahme des Fachverfahrens, die sechs Jahre nach der Migration erfolgte, war nur möglich, weil die Landesstelle für Straßentechnik es sicherheitshalber nicht hatte abschalten lassen. Es ruht nach wie vor in einem Rechenzentrum und wird irgendwann mit der Ausmusterung der Hardware erlöschen, die in ihm enthaltenen Informationen aber werden im Staatsarchiv Ludwigsburg weiterleben.

Die polizeiliche Kriminalstatistik (PKS, StAL EL 48/6) wird in Baden-Württemberg seit den frühen siebziger Jahren mit Hilfe von Computern erstellt. Ab 1984 wurden Einzelfalldatensätze als Basis der Auswertungen über Tatumstände, Täter, später auch über Tatmittel und Opfer erzeugt und bis heute aufbewahrt. Der Jahrgang 1985 ist inzwischen nicht mehr lesbar, obwohl die Daten auf professionellen Bandlaufwerken gesichert wurden – wir Archivare sollten demnach nicht müde werden, auf die Notwendigkeit verlässlicher Sicherungsmechanismen gegenüber den Behörden hinzuweisen.

Alle Mikrodaten sind durch das LKA aus dem größeren Informationsbestand der polizeilichen Fachverfahren ausgewählt worden. Insbesondere sind die Daten pseudonymisiert, die Namen also durch Codes ersetzt, die jedes Jahr wechseln. Diese Entscheidungen waren nicht beeinflussbar und mussten hingenommen werden (zur Bewertung dieser Faktenlage weiter unten mehr). Als Mikrodatenformat diente von 1984 bis 2002 eine Ableitung des Formats PAD/MOD (Personen Auskunft- und Fall-Datei/Modus-Operandi-Datei), das nicht nur für die Statistik, sondern auch für andere Fachanwendungen eingesetzt wurde. In den folgenden Jahren wurden die Mikrodaten im Zuge der bundesweiten Umstellung der Systeme der Polizei in zwei weiteren Formaten abgelegt. Wieder waren diese aus den umfangreicheren Informationsbeständen der Vorgangsbearbeitungssysteme abgeleitet.

Die Erhebung von Zahlen aus diesem inhomogenen Material erfordert derzeit technisches Spezialwissen, weshalb das Landeskriminalamt (LKA) eine Migration aller Mikrodaten in ein einheitliches Datenbankformat anstrebt. Sie sollen mittels einer browserbasierten Anwendung auch von technischen Laien nach beliebigen Kriterien ausgewertet werden können. Das Landesarchiv wird die Daten aus diesem Migrationsprozess sukzessive erhalten. Bisher sind die Jahrgänge 2003 bis 2006 abgeliefert worden, und bis September 2008 sollen die ersten Jahrgänge aus der Reihe 1996 bis 2002 vorliegen.

### III. Übernahme und Aufbereitung

Übernahme und Aufbereitung der vier übernommenen Systeme folgten einem recht einheitlichen Muster. Für die Datenmigration erzeugten die Behörden einige (4-15) Tabellen im Festbreitenformat, die zusammen mit einer Dokumentation an die mit der Systemmigration beauftragten IT-Dienstleister und in Kopie an das Landesarchiv übergeben wurden. Diese Daten bildeten den Zugang (Repräsentation 1). Anschließend übertrug das Landesarchiv die Tabellen für die dauerhafte Erhaltung in ein CSV-Format und versah sie mit einer definierten Tabellenbeschreibung in XML. Lediglich beim LKA war das Zugangsformat MS Access 2003, die Repräsentation 2 wurde jedoch nach dem gleichen Muster wie in den übrigen Fällen gebildet. Auch hier blieb die Anzahl der Datentabellen unter zehn.

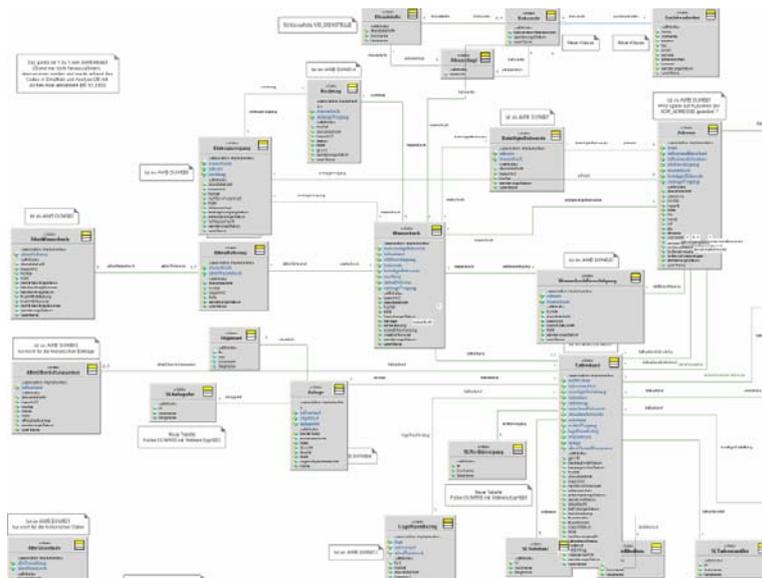


Abbildung 1: Beziehungsdiagramm einer relativ einfachen Fachanwendung mit normalisierten Daten (Ausschnitt).

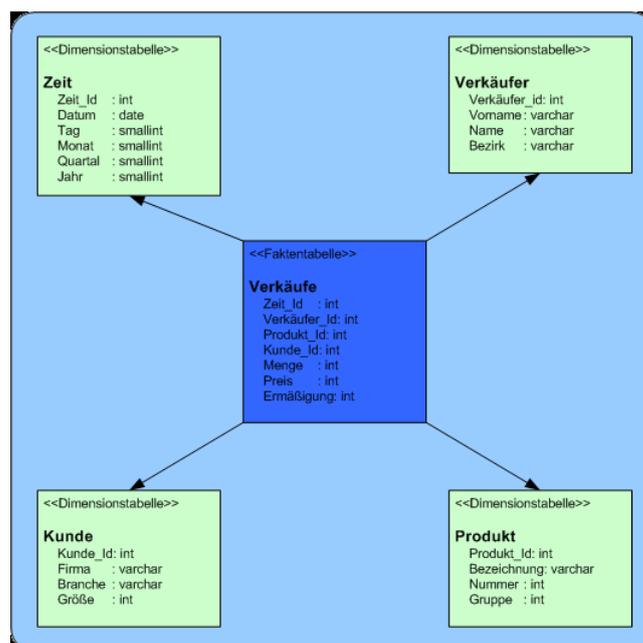


Abbildung 2: Beziehungsdiagramm eines Sternschemas für eine OLAP-Anwendung mit denormalisierten Daten. Der Zugriff auf die Fakten erfolgt über die Schlüsselattribute (mit der Endung \_Id). (Abbildung: Jörn Schimmelpfeng für Wikipedia)

## II. Die PKS als Beispiel für „Data Warehousing“

Begehungsweise, Tatmittel, Ort, erstrebte und erlangte Güter, Opfertyp, Angaben zu Verletzungen, Schulbildung, Staatsangehörigkeiten und Geburtsland des Verdächtigen: Die filigranen Daten der PKS laden dazu ein, nach verborgenen Zusammenhängen zu suchen. In der Wirtschaft liegt die systematische Fahndung nach solchen Verbindungen sehr im Trend. Eine Supermarktkette hat beispielsweise entdeckt, dass junge Väter, die abends noch schnell Windeln kaufen müssen, gern einen Sixpack Bier mitnehmen. Im Ergebnis wurde die Getränkeabteilung näher an das Babyregal gerückt. Auch in der Verwaltung findet die Methode Interesse. Bei der PKS sind zur Zeit fast alle Länder und der Bund in Entwicklungsgemeinschaften engagiert, die mit unterschiedlichen Mitteln das ähnliche Ziel verfolgen, beliebige Auswertungen aus den Mikrodaten erstellen zu können. Die PKS ist nur ein Beispiel – in vielen Verwaltungszweigen Baden-Württembergs sind entsprechende Projekte in der Planung.

Aus den verwaltenden Datenbanksystemen werden Auszüge in ein „Data Warehouse“ exportiert, in dem ein „Online Analytical Processing“ (OLAP) nach Zusammenhängen suchen und beliebige statistische Berichte erstellen kann. Um der OLAP-Anwendung die Suche zu erleichtern, werden die Datenbankinhalte umgruppiert. Während sie in der Anwendung, um Speicher- und Rechnerkapazität zu optimieren, in hunderten Tabellen abgelegt und nur durch Verweise aufeinander bezogen sind (Normalisierung, vgl. Abb. 1), führt man sie für den OLAP-Export in einigen wenigen Tabellen zusammen und nimmt auch Doppelungen in Kauf (Denormalisierung, Abb. 2).

Das so entstehende „Sternschema“ erfüllt viele Anforderungen, die zukünftige Nutzer an eine archivische Darstellung von Daten stellen könnten:

- Es ist übersichtlich (eine Zentraltabelle, wenige Kindtabellen).
- Die Tabellen sind nicht nur innerhalb der OLAP-Anwendung, sondern nach einem Export in ein dauerhaftes Format auch außerhalb ohne den Einsatz eines Datenbanksystems nutzbar.
- Ein recht großes Fallvolumen (Monate, Jahre) ist in einer Datei versammelt.
- In der Nutzung sind sowohl Angaben über Einzelfälle als auch statistische Auswertungen möglich.

Nicht ohne Grund ähnelten die im vergangenen Jahr vom Landesarchiv – ohne Kenntnis dieser Entwicklungen in der IT – zur Archivierung angeforderten Auszüge aus Datenbanken recht stark dem beschriebenen Modell.

## IV. Bewertung der Methode

Genau genommen, gibt es zwei Arten von Übernahmen im Rahmen von Migrationsprozessen. Im einen Fall wird einmalig im Moment der Migration auf ein Nachfolgesystem übernommen. Im anderen Fall werden Informationen, die ein bestehendes System an ein Data Warehouse weitergibt, regelmäßig übernommen. Beide Arten haben positive und negative Seiten, deren Relevanz im konkreten Fall zu bewerten ist. Überwiegend negativ schlagen zu Buche:

- Archivare können sich nicht darauf verlassen, dass bei einer Systemmigration alle Unterlagen des Ausgangssystems in das Folgesystem umziehen: wenn etwa Daten zu ausgedehntem Personal nicht ins Folgesystem übertragen werden sollen, muss das Archiv dies erkennen und eine Übernahme veranlassen können. Es dürfte möglich sein, diesen Export „nebenbei“ auf dem gleichen Weg wie die eigentliche Systemmigration ablaufen zu lassen.
- Der beschriebene Weg kann zu Redundanzen im Archiv führen. Wenn etwa das DoRIS-System des Landwirtschaftsministeriums in Zukunft elektronische Akten aussondert, wird

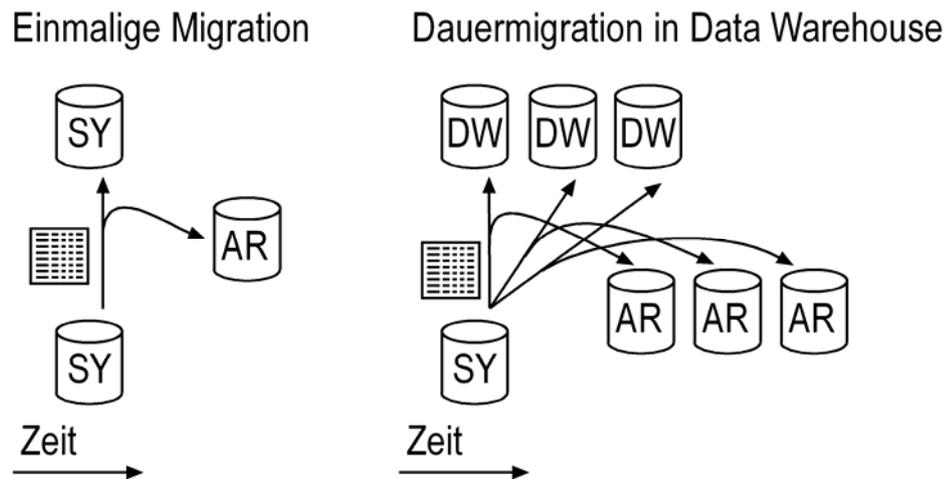


Abbildung 3: Mögliche Arten der Übernahme von archivwürdigen Informationen im Rahmen von Migrationsvorgängen. SY = Fachanwendung, AR = öffentliches Archiv, DW = Data Warehouse.

unter Umständen der vollständige POVO-Datenbestand mit angeboten werden. Wie nachteilig eine Redundanz ist, wenn man sie nicht vermeiden kann, wäre im Einzelfall zu erwägen. Im Beispiel wäre sie vielleicht erträglich, da sie Nutzern eine Recherche über einen größeren Datenbestand ermöglichen würde, und da Speicherkosten irrelevant wären (POVO als Archivbestand belegt nur rund 200 MB). Außerdem stünde POVO früher als der DoRIS-Auszug zur Nutzung bereit, da seine Schutzfristen eher ablaufen.

- Das Archiv kann, wenn es sich auf den Systemwechsel verlässt, den Zeitpunkt der Übernahme nicht bestimmen. Dieser Makel ist auch in der Papierwelt zu Hause, aber deshalb nicht weniger schwer. Insbesondere die Vergleichbarkeit von Datenbeständen leidet, wenn diese nicht in regelmäßigen Abständen vorliegen.
- In den vier geschilderten Beispielen stellte das Übertragungsformat eine gute Basis für die Übernahme ins Archiv dar. Nicht jede Systemmigration läuft aber nach dem beschriebenen Muster durch Übergabe von einzelnen Textdateien ab. Falls kompliziertere Wege eingeschlagen werden müssen, nähern sich die Kosten denen einer Übernahme aus laufenden Systemen an.
- Bei Übernahmen aus OLAP-Systemen nimmt die Behörde vorab eine Bewertung vor, beeinträchtigt also die archivische Bewertungshoheit. Eine Übernahme aus den ungleich komplizierteren Fachverfahren könnte aber an den Kosten scheitern. OLAP-Systeme liefern immerhin, wie im Falle der PKS erläutert, alle für die Aufgabenerfüllung einer Organisation wichtigen Informationen. Damit dürfte es dem Archivar unter Umständen leichtfallen, sich die an sich unberechtigte Bewertungsentscheidung zu eigen zu machen.

Andere Aspekte sind eher als Vorteile einzustufen.

- Die Finanzierung eines Aussonderungsmoduls ist bei VBS oder DMS, die über mehrere Jahre an vielen hunderten Arbeitsplätzen im Einsatz sind, wohl zu rechtfertigen. Bei kleineren Systemen wird sich eine solche Lösung aber nicht immer gegen fiskalische Zwänge durchsetzen lassen. Die Möglichkeit, bei Systemmigrationen Kopien der Daten zu erhalten, schafft also Bewegungsfreiheit, da sie die Archive von Geldfragen unabhängiger macht.
- Ein Umstieg auf eine neue Systemkomponente erfolgt aus Kostengründen. Eine abliefernde Stelle ist also im Normalfall gut motiviert, die Daten vollständig, lesbar und logisch konsistent zu übertragen. Diese Tatsache hebt sich angenehm von der Motivationslage ab, die sich einem Archivar bei einer üblichen Aussonderung bietet. So sehr auch gesetzliche

Regelungen die Behörde binden, ist doch die von außerhalb kommende Motivation oft zu gering, um eine reibungslosen Ablieferungsvorgang zu gewährleisten.

- Für die Systemmigration wird in der Regel eine Beschreibung der zu exportierenden Datenelemente und ihrer Menge angefertigt. Die Daten in den exportierten Textdateien können am Arbeitsplatz-PC in Augenschein genommen werden. Diese Umstände ermöglichen eine Bewertung der Inhalte, die über die Möglichkeiten der Sachbearbeiter-Sicht auf dem Ausgangssystem hinausgehen. In allen vier Beispielen war die Entscheidung, alle Daten als archivwürdig zu übernehmen, damit sauber begründbar.

Insgesamt stellt sich die Übernahme im Zuge der Systemmigration derzeit als Methode dar, die sich vor allem für kleinere Systeme mit einem relativ geringen Informationsanfall eignet. Sie ist auch anwendbar, wenn sich die Behörde gegen aufwändigere Methoden sträubt. Die Inhalte solcher Systeme können archivisch durchaus wertvoller als die Inhalte ressortübergreifender Massenverfahren sein. Das wichtigste Argument für ihren Einsatz sind die geringen Kosten und die gute Motivation der Behörde. Die Nachteile sind weniger gravierend, können aber, wenn sie geballt auftreten, gegen die Anwendung sprechen.

## **V. Ausblick**

Wie wird sich die geschilderte Methode weiterentwickeln? Folgende Entwicklungen halten die Verfasser aufgrund des bisher Gesagten für wahrscheinlich, ohne eine Garantie für ihr Eintreten übernehmen zu wollen:

### ***Grundsicherung aus der Gesamtheit gewinnt an Bedeutung***

Immer mehr Privatleute entdecken die Archive als Quelle der eigenen Familiengeschichte. Selbst Joschka Fischer als prominentes Mitglied der 68er-Generation hat sich 2004 ein Familienwappen mit Abstammungsnachweis zugelegt. Gleichzeitig wächst der Anteil älterer Mitbürger, die im Ruhestand nach Beschäftigung suchen. Schon jetzt profitieren Archive, die Kirchenbücher, Volkszählungsunterlagen oder Spruchkammerakten zu bieten haben, von wachsenden Benutzungszahlen, denn sie halten einen Unterlagentyp vor, der potenziell jedem Bürger ein individuelles Vergangenheitserlebnis bieten kann. Demnach könnte ein Archiv, das heute eine Grundsicherung digitaler Informationen über die Gesamtheit aller Fälle vornimmt, in einigen Jahrzehnten mit guten Nutzerzahlen rechnen.

### ***Systemeinführung und Systemwechsel stehen im Terminkalender***

Sowohl Systemeinführung als auch Systemmigration werden für Archivare wichtige Momente im Lebenszyklus von Unterlagen sein. Der Moment der Systemeinführung ist geeignet, um auf die Belange der Archive aufmerksam zu machen und bei der Systemarchitektur gewisse Weichen zu stellen. Selten jedoch kann die Aussonderung schon hinreichend geplant werden. Sie ist zu diesem Zeitpunkt oft technisch nicht konkret formulierbar. Die Migration ist ein Moment, in dem die gleichen Belange erneut geltend gemacht werden können. Die Aussonderung ist inzwischen konkretisierbar, kann kurz- und mittelfristig eingeplant und relativ kostengünstig dargestellt werden. Eine Intervention bei Systemeinführung dient also gleichsam als „Eisbrecher“ für eine spätere Übernahme.

### ***OLAP-Daten werden ein wichtiger archivischer Unterlagentyp***

Die durch Data Warehousing in den nächsten Jahren entstehenden OLAP-Systeme der öffentlichen Verwaltung werden in verdichteter Form relevante Angaben über alle von einer Behörde bearbeiteten Einzelfälle enthalten. In bestimmten Fällen werden Archivare diese Inhalte als ideale Form für die Überlieferung von massenhaft gleichförmigen Geschäftsvorgängen einstufen. Die Archive werden beim Transfer von der Fachanwendung in das OLAP-System eine Kopie der denormalisierten Daten in ihre Bestände übernehmen.

### **Speziallösungen zur Datenbankarchivierung entstehen**

Bislang galt der Export in ein Textformat mit einer rudimentären Beschreibung der Verknüpfungen als einzige Methode zur langfristigen Erhaltung von Datenbanken. Seit einiger Zeit entstehen auch verschiedene Softwarelösungen, die sich mit einem Datenbankserver verbinden, alle Datenbankinhalte exportieren und in ein einheitliches XML-Format überführen<sup>3</sup>. Dieses Vorgehen kann alle von dem Ausgangssystem geführten Inhalte und relationalen Logiken abbilden.

Gleichwohl lösen solche Werkzeuge nicht das Problem der Bewertung: Was kann ein Nutzer mit einem Datenmodell aus hunderten Tabellen anfangen, deren Zusammenhänge er nicht ohne ein dreiwöchiges Studium der Programmcodes versteht? Hier ist eine Verdichtung der Einzelinformationen notwendig, sind die Kenntnisse und Fähigkeiten der Software-Entwickler und der Archivare zugleich gefragt, die eben bei der Planung einer Systemmigration besonders günstig aufeinander treffen.

In der Welt der Informationstechnologie halten nur wenige Lösungen länger als einige Jahre. Wenn in Zukunft Anwendungen auftauchen, die sowohl eine herstellerneutrale Darstellung als auch eine archivgerechte Bewertung von Datenbankinformationen ermöglichen, könnte sich die Sachlage ändern. Aber erst wenn diese Anwendungen sich in der Praxis durchsetzen, wird die Methode der Übernahme bei Systemmigration, nachdem sie einige Jahre im Gebrauch war, allmählich wieder an Relevanz verlieren.

---

<sup>3</sup> Stefan BRANDL, Peter KELLER-MARXER: *Long-term Archiving of Relational Databases with Chronos*, International Workshop on Database Preservation, Edinburgh, 23.3.2007, [http://homepages.inf.ed.ac.uk/hmueller/presdb07/papers/Germany\\_CHRONOS\\_PresDB07.pdf](http://homepages.inf.ed.ac.uk/hmueller/presdb07/papers/Germany_CHRONOS_PresDB07.pdf);  
*HP StorageWorks Reference Information Manager for Databases Open Archive* (Produktbeschreibung), [http://h18006.www1.hp.com/products/storageworks/rimdb/modules/rim\\_db\\_openarchive.html](http://h18006.www1.hp.com/products/storageworks/rimdb/modules/rim_db_openarchive.html);  
Gabriel DAVID: *Data Warehouses in the Path from Databases to Archives*, International Workshop on Database Preservation, Edinburgh, 23.3.2007, [http://homepages.inf.ed.ac.uk/hmueller/presdb07/papers/PRESDB\\_GTD.pdf](http://homepages.inf.ed.ac.uk/hmueller/presdb07/papers/PRESDB_GTD.pdf)