



nestor

Verfahren zur Langzeitarchivierung
von Datenbankinhalten aus Fachanwendungen
und die Dokumentation dazugehöriger Prozessvorgänge

Donato Raselli

Universität Bern, Philosophisch-historische Fakultät
Archiv-, Bibliotheks- und Informationswissenschaft (MAS ALIS)

nestor-edition 7





Verfahren zur
Langzeitarchivierung
von Datenbankinhalten
aus Fachanwendungen und die
Dokumentation dazugehöriger
Prozessvorgänge

Donato Raselli

Universität Bern, Philosophisch-historische Fakultät
Archiv-, Bibliotheks- und Informationswissenschaft (MAS ALIS)

nestor edition 7

Herausgegeben von

nestor - Kompetenznetzwerk Langzeitarchivierung und
Langzeitverfügbarkeit Digitaler Ressourcen für Deutschland

nestor - Network of Expertise in Long-Term Storage
of Digital Resources

www.langzeitarchivierung.de

nestor Kooperationspartner:

- Bayerische Staatsbibliothek
- Bibliothekservice-Zentrum Baden-Württemberg
- Bundesarchiv
- Computerspiele Museum Berlin
- Deutsche Kinemathek – Museum für Film und Fernsehen
- Deutsche Nationalbibliothek
- FernUniversität Hagen
- GESIS - Leibniz-Institut für Sozialwissenschaften
- Hochschulbibliothekszentrum des Landes Nordrhein-Westfalen
- Humboldt-Universität zu Berlin
- Institut für Deutsche Sprache
- Institut für Museumsforschung (Stiftung Preußischer Kulturbesitz)
- Konrad-Zuse-Zentrum für Informationstechnik Berlin
- Landesarchiv Baden-Württemberg
- Landesarchiv Nordrhein-Westfalen
- Niedersächsische Staats- und Universitätsbibliothek Göttingen
- PDF Association
- Rechenzentrum der Universität Freiburg
- Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden
- Verbund der Zentralen Fachbibliotheken

© 2017

nestor - Kompetenznetzwerk Langzeitarchivierung und Langzeitverfügbarkeit
Digitaler Ressourcen für Deutschland

Der Inhalt dieser Veröffentlichung darf vervielfältigt und verbreitet werden, sofern der Name des Rechteinhabers "nestor - Kompetenznetzwerk Langzeitarchivierung" genannt wird. Eine kommerzielle Nutzung ist nur mit Zustimmung des Rechteinhabers zulässig.

URN: [urn:nbn:de:0008-2017080108](http://nbn-resolving.org/urn:nbn:de:0008-2017080108)

[<http://nbn-resolving.org/urn:nbn:de:0008-2017080108>]

Die Schriftenreihe **nestor edition** präsentiert ausgewählte wissenschaftliche Arbeiten mit dem Schwerpunkt Langzeitarchivierung. Sie wird in loser Folge von **nestor – Kompetenznetzwerk Langzeitarchivierung** herausgegeben. Damit entsteht ein Forum, in dem Beiträge zu verschiedenen Aspekten der digitalen Langzeitarchivierung einer breiten Öffentlichkeit zugänglich gemacht werden.

Die Arbeiten werden von ausgewiesenen Experten aus den jeweiligen Fachgebieten für die **nestor edition** gezielt ausgewählt, wenn sie einen besonderen Beitrag zu wichtigen Themenfeldern oder zu neuen wissenschaftlichen Forschungen auf dem Gebiet leisten.

Bemerkungen zu dieser Publikation, aber auch Vorschläge für die Aufnahme weiterer Beiträge in der Edition gerne an: VL-nestor@dnb.de

Für die Partner von nestor – Kompetenznetzwerk Langzeitarchivierung
Sabrina Kistner
Deutsche Nationalbibliothek

Inhaltsverzeichnis:

1. Einleitung	3
1.1 Ausgangslage und Fragestellung	3
1.2 Methodische Vorgehensweise und inhaltliche Gliederung	4
1.3 Literatur- und Quellenlage	6
2. Aspekte der digitalen Langzeitarchivierung	10
3. Archivierung von Datenbankinhalten aus Fachanwendungen	13
3.1 Fachanwendungen und Grundlagen relationaler Datenbanken	13
3.2 Aspekte der dauerhaften Archivierung von Datenbankinhalten	15
4. Verfahren zur Datenbankarchivierung	18
4.1 Analyse der Fallbeispiele und Archivierungsziele	18
4.1.1 Erstes Fallbeispiel: Die Fachanwendung „socialWEB“ des „Kompetenzzentrums Jugend und Familie Schlossmatt“.....	18
4.1.2 Zweites Fallbeispiel: Die Fachanwendung „ESCADA ² “ des „Erziehungsdepartements Kanton Basel-Stadt“.....	22
4.2 Die Spezifikation und Funktionsweise der Archivierungsverfahren	27
4.2.1 IngestList (Landesarchiv Baden-Württemberg).....	27
4.2.2 SIARD (Schweizerisches Bundesarchiv).....	29
4.2.3 CHRONOS (Softwarehersteller CSP GmbH).....	31
4.3 Evaluation der Archivierungsverfahren in Bezug auf die Fallbeispiele	35
5. Archivierung datenbankgestützter Geschäftsprozesse	42
5.1 Empfehlungskatalog zur Dokumentation von Prozessvorgängen	42
5.2 Anwendung des Empfehlungskatalogs auf die Fallbeispiele	45
6. Schlussfolgerungen	49
7. Bibliografie	53
Anhang 1: Elemente des relationalen Datenmodells	60
Anhang 2: Migration von Daten aus einer Datenbank und Aufnahme deskriptiver Anmerkungen mit IngestList	61
Anhang 3: Migration von Daten aus einer Datenbank und Aufnahme deskriptiver Anmerkungen mit SIARD	62
Anhang 4: Migration von Daten aus einer Datenbank und Aufnahme deskriptiver Anmerkungen mit CHRONOS	63
Anhang 5: Empfehlungskatalog: Zusatzdokumentation für Archivdaten	64

1. Einleitung

1.1 Ausgangslage und Fragestellung

Informationen werden zunehmend digital erzeugt und verbreitet. Die dadurch entstehenden kulturellen und wissenschaftlichen Ressourcen sind aber von Kurzlebigkeit bedroht, da technische Systeme zur Interpretation digitaler Informationen infolge rascher Entwicklungen in der Computertechnologie stets zu überaltern drohen. Angesichts dessen ist die digitale Langzeitarchivierung bemüht, Strategien zu entwickeln, um digitale Informationen dauerhaft verfügbar zu halten.¹ Konzentrierten sich entsprechende Bemühungen anfänglich auf Objekte wie digitale Bilder und Textdokumente, rückte in den letzten Jahren die Archivierung von Datenbankinhalten in den Vordergrund.² Zentral sind in diesem Zusammenhang datenbankgestützte Fachanwendungen, da diese in Privatunternehmen sowie in der öffentlichen Verwaltung maßgeblich die Bearbeitung von Geschäftsprozessen unterstützen.³

Von Datenbankbetreibern abgesehen ist die Archivierung von Datenbankinhalten aus Fachanwendungen insbesondere im Archivbereich bedeutend, weil Fachanwendungen gegenüber öffentlich-rechtlichen Archiven einer Anbietungspflicht unterliegen und Datenbankanhalte oftmals aufgrund ihres Evidenz- und Informationswertes in Archivbestände aufgenommen werden.⁴ Dabei ist erforderlich, dass archivwürdige Geschäftsprozesse und dazugehörige Prozessinstanzen (Prozessdurchläufe), die zur Datenerzeugung in einer Fachanwendung geführt haben, im Archivgut nachvollzogen werden können. Zudem ist es unerlässlich, dass die Integrität und Authentizität der Archivdaten gewährleistet wird.⁵

Entsprechend der Relevanz der Archivierung von Datenbankinhalten aus Fachanwendungen für den Archivbereich wurde die Entwicklung spezifischer Lösungen zur Datenbankarchivierung primär aus dem Archivbereich angestoßen. Konkret kam es auf nationaler Ebene – im Zuge digitaler Langzeitarchivierungsprojekte – zur Entwicklung unterschiedlicher Verfahren bzw. Applikationen, die Inhalte aus Fachanwendungen in geeigneter Form in digitale Archive übernehmen. Mit Verzögerung entwickelten vermehrt auch private Softwareentwickler kommerzielle Lösungen für die digitale Langzeitarchivierung von Datenbankinhalten.⁶

¹ Altenhöner/Schrimpf, Langzeitarchivierung, S. 529-531; Scheffel, Einführung, Kap. 2, S. 1; Borghoff/Rödig, Technik der digitalen Langzeitarchivierung, S. 6.

² Heuscher/Keller-Marxer, Archival Access to Complex Relational Data, S. 1; Christophides/Buneman, Database Preservation, S. 55.

³ Dässler/Schwarz, Fachverfahren, S. 6.

⁴ Ebd., S. 8; Bernstein, Database Preservation, S. 1; Hoen, Auslandberichterstattung, S. 347; ERPANET, Long-term Preservation of Databases, S. 5.

⁵ ERPANET, Long-term Preservation of Databases, S. 11; Giaretta, Digital Preservation, S. 203; KOST, Bewertung von Datenbanken, o.S..

⁶ Dässler/Schwarz, Fachverfahren, S. 15; ERPANET, Long-term Preservation of Databases, S. 4f.

Vor dem Hintergrund der Entwicklung von Verfahren zur Datenbankarchivierung und der archivischen Anforderung, dass im künftigen Archivgut Prozesse und Prozessinstanzen nachvollziehbar sein sollten und dass Integrität und Authentizität der Daten zu gewährleisten sind, soll in dieser Untersuchung die Frage beantwortet werden, *wie sich die unterschiedlichen Verfahren dazu eignen, diese inhaltlichen und qualitativen Anforderungen zu erfüllen.* In diesem Zusammenhang wird zudem danach gefragt, *welche Informationen bei der Anwendung der jeweiligen Verfahren unter Umständen zusätzlich zu den Datenbankdaten archiviert werden müssen, um das Archivierungsziel zu erreichen.*

Die Untersuchung konzentriert sich auf die Auseinandersetzung mit drei Verfahren. Es handelt es sich um die Anwendungen „SIARD“, ein verbreitete Archivierungsanwendung des Schweizerischen Bundesarchivs, und „CHRONOS“, ein führendes kommerzielles Produkt des privaten Softwareentwicklers CSP. Schließlich wird ein Verfahren untersucht, das durch die Software „IngestList“, ein in Deutschland häufig verwendetes Programm des Landesarchivs Baden-Württemberg, unterstützt wird.⁷ Um die Untersuchung anhand von Fallbeispielen durchzuführen, werden zwei datenbankgestützte Fachanwendungen aus der Praxis herangezogen. Dabei handelt es sich einerseits um die Fachanwendung „socialWEB“ der Firma „socialweb GmbH“ und andererseits um die Fachanwendung „ESCADA²“ der Firma „JCS Software AG“. Die Anwendung „socialWEB“ wird vom Kompetenzzentrum Jugend und Familie „Schlossmatt“ Bern (KJFS) und die Anwendung „ESCADA²“ vom Erziehungsdepartement Basel-Stadt (ED) eingesetzt. Die Fallbeispiele konnten durch die Unterstützung des Stadtarchivs Bern und dem Staatsarchiv Basel-Stadt ermittelt werden.

1.2 Methodische Vorgehensweise und inhaltliche Gliederung

A) Methodische Vorgehensweise

Zur Beantwortung der Fragestellung wird folgende methodische Vorgehensweise gewählt: In einem ersten Schritt werden die Fallbeispiele analysiert und Archivierungsziele formuliert. In diesem Zusammenhang werden die Aufgaben der ausgewählten Institutionen und die Funktionsweise der darin angewendeten Fachanwendungen sowie die durch diese unterstützten Geschäftsprozesse analysiert. Dabei soll aufgezeigt werden, welche Rolle den Fachanwendungen konkret zukommt und welche analogen Unterlagen zusätzlich zu den Datenbankinhalten erzeugt werden. Auf der Grundlage dieser Analyse wird in Form von Archivierungszielen festgelegt, welche Prozesse und damit verknüpften Datenbankinhalte archiviert werden sollen. Für die Analyse werden u.a. verfügbare Quellen (z.B. Detailspezifikationen und Benutzerhandbücher) sowie Interviews mit den Fachanwendungsbetreibern verwendet.

⁷ Vgl. Hollmann, Archivierung genuin elektronischer Unterlagen, S. 3.

In einem zweiten Schritt werden die Archivierungsverfahren untersucht. Hierbei interessieren insbesondere Spezifikation und Funktionsweise bzw. die Archivierungsmöglichkeiten der Verfahren. Zu diesem Zweck werden einerseits Fachliteratur berücksichtigt und andererseits die Archivierungsverfahren, soweit wie möglich, direkt untersucht. In einem dritten Schritt werden die Archivierungsmöglichkeiten, welche die Archivierungsverfahren zulassen, sowie die durch die Verfahren gewonnen Archivdaten in Bezug auf die Archivierungsziele evaluiert. Dabei wird auf theoretischer Ebene beurteilt, inwiefern sich die untersuchten Verfahren dazu eignen, die als archivwürdig erachteten Datenbankinhalte der Fallbeispiele so zu archivieren, dass im Archivgut zum einen die mit den Datenbankinhalten in Verbindung stehenden Prozesse und Prozessinstanzen nachvollzogen und zum anderen Integrität und Authentizität der Archivdaten gewährleistet werden können.

Als letzter Schritt werden Möglichkeiten aufgezeigt, wie Archivdaten allenfalls durch Zusatzinformationen ergänzt werden können, damit durch Fachanwendungen unterstützte Prozesse (Prozessinstanzen) im Archivgut nachvollzogen werden können. In der Folge werden die aufgezeigten Möglichkeiten in einem Empfehlungskatalog zusammengefasst. Geeignete Zusatzinformationen werden auf der Grundlage der Fachliteratur ermittelt. Abschließend wird dargelegt, mit welchen Zusatzinformationen die Fallbeispiele vor dem Hintergrund der individuellen Möglichkeiten der Archivierungsverfahren ergänzt werden müssen, damit die Archivierungsziele erreicht werden können.

B) Inhaltliche Gliederung

Im Anschluss an die Einleitung (Ausgangslage / Fragestellung, methodische Vorgehensweise, inhaltliche Gliederung sowie Literatur- und Quellenlage) werden in einem ersten Teil Aspekte der digitalen Langzeitarchivierung dargestellt. Hierbei interessieren u.a. die Ziele, Strategien (Migration / Emulation) und Anforderungen (Integritäts- und Authentizitätswahrung) der digitalen Archivierung sowie das „Open Archival Information System“, das die konzeptionelle Grundlage digitaler Archive bildet (Kapitel 2). Vor diesem Hintergrund liegt der Fokus in einem zweiten Teil auf der Archivierung datenbankgestützter Fachanwendungen (Kapitel 3). Hierbei interessieren zunächst Gestalt und Funktion von Fachanwendungen sowie die theoretischen Grundlagen relationaler Datenbanksysteme. Als Nächstes werden grundlegende Aspekte der Datenbankarchivierung diskutiert.

Nach der Erarbeitung der theoretischen Basis werden in einem dritten Teil die Fallbeispiele analysiert und die Archivierungsziele festgelegt (Kapitel 4.1). Im Weiteren werden die Spezifikation und die Funktionsweisen bzw. Archivierungsmöglichkeiten der ausgewählten Archivierungsverfahren untersucht (Kapitel 4.2) und in Bezug auf die Fallbeispiele bzw. die festgelegten Archivierungsziele evaluiert (Kapitel 4.3). In einem vierten Teil (Kapitel 5) werden

Möglichkeiten aufgezeigt, wie archivierte Datenbankinhalte im Hinblick auf die Überlieferung der damit verbundenen Prozesse und Prozessinstanzen mit Zusatzinformationen ergänzt werden können. Die Ergebnisse werden in einem Empfehlungskatalog zusammengefasst (Kapitel 5.1). Der Empfehlungskatalog wird sodann im erforderlichen Umfang auf die Fallbeispiele angewendet (Kapitel 5.2). Im fünften und letzten Teil werden die Ergebnisse zusammengefasst (Kapitel 6). Die Ergebnisse können als Entscheidungsgrundlage für die künftige Wahl eines Datenbankarchivierungsverfahrens dienen.

1.3 Literatur- und Quellenlage

A) Literaturlage

Die Literaturlage zur digitalen Langzeitarchivierung kann als sehr gut bezeichnet werden. Exemplarisch sind für den deutschen Sprachraum das Werk von Borghoff (2003) sowie für den französischen und angelsächsischen Sprachraum die Publikationen von Banat-Berger (2009), Giaretta (2011) und Brown (2013) hervorzuheben.⁸ Diese Veröffentlichungen beschäftigen sich umfassend mit Schwerpunkten der digitalen Archivierung und begründen einen zentralen Bestandteil des aktuellen Forschungsstandes. Nebst diesen Publikationen gibt es eine Vielzahl von Artikeln, die einerseits in die Thematik einführen und andererseits vertiefend auf spezifische Themen der digitalen Archivierung eingehen.⁹ Relevant sind zudem die Arbeiten der schweizerischen Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen (KOST) und des deutschen Netzwerks für die Langzeitarchivierung digitaler Ressourcen (nestor). Sowohl KOST als auch nestor veröffentlichen einschlägige Beiträge in lexikalischer Form und anhand von Abstracts über ihren Webauftritt. Im Weiteren publizierte nestor ein Handbuch zur digitalen Langzeitarchivierung (2009), das im deutschen Sprachraum ebenfalls als Standardwerk zu betrachten ist.¹⁰

Literatur zur Datenbankarchivierung bzw. zur Archivierung von Datenbankinhalten aus Fachanwendungen ist verhältnismäßig wenig vorhanden. Wichtige Beiträge wurden im Rahmen internationaler Workshops bzw. Konferenzen zur Datenbankarchivierung publiziert. Sie beschäftigen sich mit der Präsentation unterschiedlicher Projekte, in welchen ausschließlich die Migration von Datenbankinhalten im Vordergrund steht (2003-2007).¹¹ Weitere Veröffentlichungen entstanden direkt im Zuge von Archivierungsprojekten einzelner Archivinstitutionen

⁸ Vgl. Borghoff/Rödiger et al., Langzeitarchivierung; Brown, Practical digital preservation; Banat-Berger, l'archivage numérique à long terme; Giaretta, Digital Preservation.

⁹ Exemplarisch: Altenhöner/Schrimpf, Langzeitarchivierung; Ohnesorge/D'Incau et al., digitale Archivierung.

¹⁰ Vgl. <<http://www.kost-ceco.ch>> (20.6.2014); <<http://www.langzeitarchivierung.de>> (22.6.2014); Neuroth/Osswald et al., nestor Handbuch.

¹¹ Vgl. ERPANET, Long-term Preservation of Databases; Hoen, Auslandberichterstattung; Christophides/Buneman, Database Preservation.

(2003-2009).¹² Ferner erschienen Publikationen, die sich allgemein mit der Thematik der Datenbankarchivierung auseinandersetzen. Hervorzuheben sind an dieser Stelle die Monografie von Olson (2009), der sich aus einer allgemeinen Perspektive mit Vorgehensweisen bezüglich der Umsetzung eines Datenbankarchivierungsprojekts beschäftigt, und die Zeitschriftenartikel von Ribeiro (2009), Naumann (2010) sowie von Dässler/Schwarz (2010).¹³ Letzterer Arbeit kommt eine zentrale Bedeutung zu, da sie sich eingehend mit der Archivierung von Datenbankinhalten aus Fachanwendungen beschäftigt und dabei explizit eine archivistische Perspektive einnimmt. Literatur, die sich allgemein mit der Theorie von (relationalen) Datenbanksystemen auseinandersetzt – ohne Bezug auf die Archivierung von Datenbankinhalten –, ist in großer Menge verfügbar; exemplarisch sind die Publikationen von Schubert (2007), Steiner (2009) und Kemper (2013) zu erwähnen.¹⁴

Literatur zu den ausgewählten Archivierungsverfahren ist ausreichend vorhanden. Für die Anwendung „SIARD“ ist insbesondere die SIARD-Formatspezifikation (2011) zu erwähnen, die auf der Internetseite des schweizerischen Bundesarchivs frei zugänglich ist.¹⁵ Ferner ist SIARD Gegenstand zahlreicher Veröffentlichungen, die sich allgemein mit der Datenbankarchivierung beschäftigen (vgl. oben). Ähnlich verhält es sich mit „CHRONOS“. Namentlich zu erwähnen sind hier der Aufsatz von Brandl/Keller-Marxer (2007), das „Whitepaper“ von Keller-Marxer (2012) sowie das Benutzerhandbuch der Firma CSP.¹⁶ Was die Anwendung „IngestList“ anbelangt, steht verhältnismäßig wenig Literatur zur Verfügung. Hervorzuheben ist an dieser Stelle ein Zeitschriftenartikel von Keitel (2010).¹⁷

Einschlägige Studien, die sich aus einer kritischen Perspektive mit den Anwendungen „SIARD“ und „CHRONOS“ auseinandersetzen, wurden von Lindley (2012) vom „Austrian Institut of Technology“ und von Fitzgerald (2013) von den „Queensland State Archives“ veröffentlicht. Lindley untersucht technische Aspekte der Anwendungen, während sich Fitzgerald mit der Frage beschäftigt, wie die Anwendungen Bestandteil von Verfahren sein können, die Datenbankinhalte in digitale Archive überführen. Die Studie von Fitzgerald ist nicht abgeschlossen, weshalb sich der Artikel auf einen Zwischenbericht beschränkt.¹⁸

Literatur, die sich explizit mit der Frage beschäftigt, mit welchen zusätzlichen Informationen Datenbankinhalte bei der Archivierung u.U. zu ergänzen sind, damit Prozesse (Prozessin-

¹² Vgl. Bernstein, Database Preservation; Heuscher/Keller-Marxer, Archival Access to Complex Relational Data; Heuscher/Keller-Marxer, XML – ein strategisches Instrument für Archive, S. 16-18; Keitel/Lang, Ingest von Fachverfahren.

¹³ Vgl. Olson, Database Archiving; Ribeiro, Database Preservation; Naumann, Übernahme von Daten aus Fachanwendungen; Dässler/Schwarz, Fachverfahren.

¹⁴ Vgl. Schubert, Datenbanken; Steiner, Relationale Datenbanken; Kemper/Eickler, Datenbanksysteme.

¹⁵ Vgl. BAR, SIARD. Formatbeschreibung.

¹⁶ Vgl. Brandl/Keller-Marxer, Long-term Archiving; Keller-Marxer, CHRONOS (=Whitepaper); CSP GmbH, Benutzerhandbuch.

¹⁷ Vgl. Keitel/Lang, DIMAG und IngestList.

¹⁸ Vgl. Lindley, Database Preservation; Fitzgerald, data archiving tools.

stanzen), die von Fachanwendungen unterstützt werden, im Archivgut nachvollzogen werden können, ist nicht vorhanden. Die Thematik wird aber in zahlreichen Artikeln als Bestandteil der Datenbankarchivierung erwähnt und in knapper Form besprochen. Vor diesem Hintergrund ist insbesondere auch die Auseinandersetzung mit den Fallbeispielen von Bedeutung.

B) Quellenlage

Die Quellenlage kann insgesamt als befriedigend bezeichnet werden. Für die Fachanwendung „socialWEB“ des KJFS stehen ausführliche Benutzerhandbücher zur Verfügung, in welchen die Anwendung mit zahlreichen Screenshots illustriert wird.¹⁹ Zudem enthält die Homepage des Anwendungsherstellers weiterführende Informationen zur Fachanwendung.²⁰ Das KJFS hat für diese Untersuchung zudem eine „Matrix“ zugänglich gemacht, die aufzeigt, in welcher Art und Weise bestimmte Informationen in die Fachanwendung aufgenommen werden.²¹ Was das Datenmodell der proprietären Fachanwendung „socialWEB“ anbelangt, hat der Hersteller darauf hingewiesen, dieses, sollte zwischen dem Autor, dem Hersteller, der Universität Bern und dem Stadtarchiv Bern eine Vereinbarung getroffen werden, partiell zur Verfügung zu stellen. Dabei müssten aber in Bezug auf eine allfällige Publikation der Untersuchung bestimmte Rahmenbedingungen eingehalten werden.²² Vor diesem Hintergrund wurde auf die Ausarbeitung einer Vereinbarung verzichtet. Hinsichtlich der Aufgaben des KJFS sind ein ausführliches Leitbild und einschlägige gesetzliche Erlasse vorhanden. Für die Fachanwendung „ESCADA²“ des ED Basel-Stadt stehen ausführliche Benutzerhandbücher und zusätzlich technische Spezifikationen zur Verfügung.²³ Auch in diesem Fallbeispiel sind der Homepage des Anwendungsherstellers weiterführende Informationen zu entnehmen.²⁴ Das Datenmodell der ebenfalls proprietären Anwendung „ESCADA²“ wird durch den Hersteller „JCS Software AG“ nicht zugänglich gemacht.²⁵ Über die Tätigkeit bzw. die Aufgaben des ED Basel-Stadt geben die Homepage des ED sowie gesetzliche Erlasse

¹⁹ Vgl. socialweb Software GmbH, Benutzerhandbuch. Arbeitsplanung; socialweb Software GmbH, Benutzerhandbuch. Arbeitszeiterfassung; socialweb Software GmbH, Benutzerhandbuch. Buchhaltung; socialweb Software GmbH, Benutzerhandbuch. Einsatz/Freiwünsche und Kontingente; socialweb Software GmbH, Benutzerhandbuch. KlientInnenmodul; socialweb Software GmbH, Benutzerhandbuch. KlientInnenmodul – Zielprozesse und Ein-/Austritte von KlientInnen; socialweb Software GmbH, Benutzerhandbuch. Kontaktmodul; socialweb Software GmbH, Benutzerhandbuch. MitarbeiterInnenmodul; socialweb Software GmbH, Benutzerhandbuch. Protokolle; socialweb Software GmbH, Produktdokumentation.

²⁰ Vgl. URL: <www.socialweb.ch> (30.4.2014).

²¹ Vgl. KJFS, Matrix.

²² Interview von Donato Raselli mit Matthias Andenmatten, 2.5.2014.

²³ Vgl. JCS Software AG, Benutzerhandbuch. FakturaGST; JCS Software AG, Benutzerhandbuch. PLSGST; JCS Software AG, Benutzerhandbuch. SchuleGST; JCS Software AG, ESCADA². Grundfunktionen; JCS Software AG, ESCADA². Grundschulen; JCS Software AG, ESCADA². Stammdaten; JCS Software AG, ESCADA².

²⁴ Vgl. URL: <<http://www.jcssoftware.ch>> (16.5.2014).

²⁵ Korrespondenz zwischen Donato Raselli und Viviane Cantaluppi, 14.4.2014.

Auskunft. Für die zwei Fallbeispiele wurden keine Testdaten zur Verfügung gestellt, da in beiden Fachanwendungen sensible Personendaten verwaltet werden.²⁶

Um weiterführende Informationen über die Fachanwendungen sowie die Institutionen zu erlangen, in welchen die Anwendungen zum Einsatz kommen, wurden vier Interviews durchgeführt. Für das erste Fallbeispiel wurden Martina Suter, Leiterin des KJFS, und Matthias Andenmatten, Geschäftsführer der „socialweb Software GmbH“, interviewt.²⁷ Für das zweite Fallbeispiel wurden Christoph Döbeli, Leiter Information und Dokumentation des ED Basel-Stadt, und Claude Stettler sowie Ali Aydin der Fachstelle Applikationen des ED Basel-Stadt (Leiter/IC-Berater II) befragt.²⁸

²⁶ Interview von Donato Raselli mit Matthias Andenmatten, 2.5.2014; Korrespondenz zwischen Donato Raselli und Viviane Cantaluppi, 13.3.2014.

²⁷ Interview von Donato Raselli mit Martina Suter, 25.4.2014; Interview von Donato Raselli mit Matthias Andenmatten, 2.5.2014.

²⁸ Interview von Donato Raselli mit Christoph Döbeli, 14.5.2014; Interview von Donato Raselli mit Claude Stettler und Ali Aydin, 14.5.2014.

2. Aspekte der digitalen Langzeitarchivierung

Ziel der digitalen Langzeitarchivierung ist die dauerhafte Erhaltung digitaler Informationen. Eine dauerhafte Erhaltung bedeutet, dass digitale Informationen auch in einer veränderten technischen Umgebung zur Anzeige gebracht werden können.²⁹ Die zentrale Herausforderung an die digitale Langzeitarchivierung besteht diesbezüglich darin, Daten in einem digitalen Archiv unversehrt und unverfälscht zu erhalten. Zu diesem Zweck sind ohne zeitliche Einschränkung sowohl die Integrität als auch die Authentizität digitaler Informationen zu gewährleisten.³⁰ Das bedeutet, dass archivierte Daten weder absichtlich noch unabsichtlich verändert oder zerstört werden (Integrität) und weiterhin darstellen, was sie vorgeben darzustellen (Authentizität). Dabei besteht zwischen der Integrität und der Authentizität eine Abhängigkeit. So umfasst die Beurteilung der Authentizität immer auch die Feststellung der Integrität. Weitere Anforderungen an die digitale Langzeitarchivierung sind die Sicherung der technischen Interpretierbarkeit, der Benutzbarkeit sowie der Nachvollziehbarkeit digitaler Informationen, damit diese von künftigen Nutzergruppen verwendet und in ihrem historischen Kontext verstanden werden können.³¹

Um auf die beschränkte Lebensdauer technischer Systeme zur Interpretation digitaler Informationen zu reagieren, stehen der digitalen Langzeitarchivierung im Wesentlichen die Migrations- und Emulationsstrategie als Handlungsmöglichkeiten zur Verfügung. Bei der Migration werden digitale Objekte periodisch auf neue Datenträger migriert und/oder in neue Formate konvertiert. Sollen bei der Datenträgermigration die Daten physisch erhalten bleiben, werden bei der Konvertierung die Daten verändert, damit sie künftig von einer neuen technischen Umgebung interpretiert werden können. Dabei besteht aber die Gefahr, dass Informationen teilweise oder sogar vollständig verloren gehen können.³²

Das Problem eines allfälligen Datenverlustes bei der Datenmigration kann durch die Strategie der Emulation umgangen werden. Zu diesem Zweck werden die von einem digitalen Objekt ursprünglich benötigten Hardware- und Softwarekomponenten in einer neuen Umgebung nachgebildet. Durch diese Strategie müssen digitale Objekte nicht verändert werden und die Daten verbleiben in ihrem originalen Zustand. Der Nachteil der Emulation besteht darin, dass

²⁹ Altenhöner/Schrimpf, Langzeitarchivierung, S. 529-531; Borghoff/Rödiger, Technik der digitalen Langzeitarchivierung, S. 6.

³⁰ Ebd., S. 531; Ohnesorge/D'Incau et al., digitale Archivierung, S. 12; Giaretta, Digital Preservation, S. 203.

³¹ Ebd. Vgl. Authenticity Task Force, Assessing and Maintaining the Authenticity, S. 1-11; Internationale Organisation für Normung, Schriftgutverwaltung (ISO 15489-1), S. 12f; Giaretta, Digital Preservation, S. 203-205. Vgl. Raselli, Verfahren zur Wahrung der Integrität und Authentizität.

³² Borghoff/Rödiger, Technik der digitalen Langzeitarchivierung, S. 7f; Ullrich, Bitstream Preservation, Kap. 8, S. 3-9; Funk, Migration, Kap. 8, S. 10-15; Giaretta, Digital Preservation, S. 200-202; Ohnesorge/D'Incau et al., digitale Archivierung, S. 13.

die Emulationssoftware für jede weiterentwickelte Hardwarearchitektur aktualisiert oder aufwändig neu entwickelt werden muss.³³

Infolge des technischen und organisatorischen Aufwandes der Emulation wird in der Praxis primär die Datenmigration für die digitale Langzeitarchivierung verwendet.³⁴ Für eine erfolgreiche Umsetzung dieser Strategie müssen Archivdaten aber durch umfassende Metadaten ergänzt werden. Konkret besteht die Notwendigkeit, technische, beschreibende, strukturierende und administrative Metadaten zu erheben. Einerseits können mit diesen Daten die Anforderungen an die digitale Archivierung – insbesondere die Gewährleistung der Integrität und Authentizität – sichergestellt werden; andererseits ermöglichen sie die Recherche und Verwendung der digitalen Informationen.³⁵ Vor diesem Hintergrund grenzt sich die digitale Langzeitarchivierung – zumindest im Archiv- und Bibliotheksbereich – von der Datensicherung ab, die nur die Speicherung von Daten auf einem Datenträger umfasst.³⁶

Was die grundsätzliche Organisation digitaler Archive anbelangt, stößt in der digitalen Langzeitarchivierung das „Open Archival Information System“ (OAIS) auf große Akzeptanz und wird entsprechend breit rezipiert. Als Referenzmodell definiert das OAIS Schlüsselbegriffe und gibt für Archivierungsabläufe einen formal-strukturellen Rahmen vor – von der Datenübernahme bis zum Datenzugriff durch Nutzergruppen –, ohne dabei aber konkrete technische und betriebsorganisatorische Lösungsmöglichkeiten vorauszusetzen.³⁷ Ein zentrales Element des OAIS ist das so genannte Informationspaket. In diesem bilden die zu archivierenden Inhaltsdaten zusammen mit Repräsentationsinformationen, die der Interpretation der Inhaltsdaten dienen, und den übrigen Metadaten (z.B. beschreibende Metadaten) eine konzeptionelle Einheit. Je nach dem Bearbeitungszeitpunkt wird zwischen Submission (SIP), Archival (AIP) und Dissemination Information Package (DIP) unterschieden (Übersetzt: Übernahme-, Archivierungs- und Nutzungspaket).³⁸

Die Informationspakete dienen im OAIS als Eingabedaten für sechs definierte Arbeitsprozesse: Bei der *Übernahme* (Ingest) werden digitale Objekte als SIP in das Archiv übernommen, in AIPs umgewandelt (z.B. Datenmigration) und im *Archivspeicher* (Archival Storage) gespeichert. Die Metadaten werden durch die *Datenverwaltung* (Data Management) verwaltet

³³ Ebd., S. 8; Funk, Emulation, Kap. 8, S. 16-23; Giaretta, Digital Preservation, S. 123f.

³⁴ Altenhöner/Schrimpf, Langzeitarchivierung, S. 530; Ohnesorge/D'Incau et al., digitale Archivierung, S. 13.

³⁵ Funk, Migration, Kap. 8, S. 10; Ohnesorge/D'Incau et al., digitale Archivierung, S. 16; Giaretta, Digital Preservation, S. 177-190.

³⁶ Ebd., S. 530f; Liegmann/Neuroth, Einführung, Kap. 1, S. 3.

³⁷ Brübach, Das Referenzmodell OAIS, Kap. 4, S. 1-7; Borghoff/Rödiger et al., Langzeitarchivierung, S. 25f; Ohnesorge/D'Incau et al., digitale Archivierung, S. 13f; Altenhöner/Schrimpf, Langzeitarchivierung, S. 534. Vgl. Nestor-Arbeitsgruppe, Referenzmodell für ein Offenes Archiv-Informationssystem.

³⁸ Ebd., Kap. 4, S. 8f; Banat-Berger, l'archivage numérique à long terme, S. 51; Brown, Practical digital preservation, S. 129-153; Giaretta, Digital Preservation, S. 53-62.

und dienen der Auffindbarkeit digitaler Objekte. Beim *Zugriff* (Access) werden die Metadaten und die Inhaltsdaten den Nutzern als DIP zugänglich gemacht. Parallel zu diesen Kernprozessen stellt die *Erhaltungsplanung* (Preservation Planning) die dauerhafte Nutzbarkeit der AIPs sicher und die *Administration* sorgt für die allgemeine Funktion des digitalen Archivs.³⁹

³⁹ Ebd., Kap. 4, S. 9-14; Borghoff/Rödiger et al., Langzeitarchivierung, S. 29-31; Banat-Berger, l'archivage numérique, S. 52-55; Giaretta, Digital Preservation, S. 63-65.

3. Archivierung von Datenbankinhalten aus Fachanwendungen

3.1 Fachanwendungen und Grundlagen relationaler Datenbanken

In der digitalen Langzeitarchivierung stellt die Archivierung strukturierter Daten eine Herausforderung dar.⁴⁰ Dabei stehen infolge ihrer Verbreitung insbesondere Fachanwendungen im Vordergrund. Fachanwendungen sind „Spezialapplikationen“⁴¹, die im administrativen Umfeld von Behörden und Unternehmen für die Bearbeitung bestimmter Aufgaben bzw. Geschäftsprozesse zur Anwendung gelangen. Sie speichern Informationen in formalisierter Form in Datenbanksystemen ab, damit diese über definierte Abfragen recherchiert, verknüpft und zu Berichten zusammengefügt werden können.⁴² In welcher Struktur die Informationen dabei konkret gespeichert werden, wird durch die Spezifikation des Datenbanksystems, das die informationstechnische Grundlage einer Fachanwendung bildet, bestimmt.⁴³

Ein Datenbanksystem dient dazu, große Mengen an Informationen persistent zu speichern, zu verwalten und Anwendungen ohne Redundanz und Inkonsistenz zur Verfügung zu stellen.⁴⁴ Datenbanksysteme setzen sich grundsätzlich aus einer Datenbank und einem Datenbankmanagementsystem (DBMS) zusammen. Das DBMS ist eine Software bzw. eine Steuerungseinheit, um Daten, die in der Datenbank gespeichert sind, unabhängig von Anwendungsprogrammen zu verwalten bzw. Systemroutinen wie z.B. die Speicherung, Überschreibung oder Löschung von Daten durchzuführen. Anwendungsprogramme können nur über ein DBMS und mittels einer Anwendungsprogramm-Schnittstelle auf Datenbankinhalte zugreifen. Vor diesem Hintergrund kann in Bezug auf Datenbanksysteme eine physische, konzeptionelle und externe Abstraktionsebene unterschieden werden.⁴⁵

Auf der physischen Ebene wird definiert, wie und wo Daten in einem Datenbanksystem gespeichert werden (z.B. auf einem Plattenspeicher). Demgegenüber wird auf der konzeptionellen Ebene mittels eines Datenmodells die logische Struktur der Datenverwaltung durch ein DBMS und damit die theoretische Grundlage eines Datenbanksystems festgelegt. Im Wesentlichen können hierbei relationale, hierarchische und objektorientierte Datenbankmodelle unterschieden werden. In der Praxis kommen aber überwiegend nur relationale Datenbanken bzw. relationale Datenbankmanagementsysteme (RDBMS) zur Anwendung.⁴⁶

⁴⁰ Vgl. KOST, Katalog archivischer Dateiformate (strukturierte Daten: Datenbanken), o.S..

⁴¹ Toebak, Records Management, S. 588.

⁴² Ebd.; VdA, Bewertung elektronischer Fachverfahren, S. 1; Dässler/Schwarz, Fachverfahren, S. 6.

⁴³ Steiner, Relationale Datenbanken, S. 8-12. Vgl. Giaretta, Digital Preservation, S. 38.

⁴⁴ Schubert, Datenbanken, S. 22-28; Kemper/Eickler, Datenbanksysteme, S. 19f.

⁴⁵ Kemper/Eickler, Datenbanksysteme, S. 21-23; Steiner, Relationale Datenbanken, S. 6f; Unterstein/Matthiessen, Relationale Datenbanken, S. 7-17.

⁴⁶ Ebd.; Schubert, Datenbanken, S. 50-53; Dässler/Schwarz, Fachverfahren, S. 9-11.

In einer relationalen Datenbank werden Daten für bestimmte Zwecke zusammengeführt und in Form von Tabellen (Relationen) abgelegt. Eine Tabelle setzt sich aus Zeilen (Tupeln, Datensätzen) zusammen (z.B. Nutzer A), deren Inhalte über Spalten (Attribute) festgelegt werden (z.B. Name und Vorname des Nutzers A). Die einzelnen Bestandteile einer Zeile befinden sich folglich in Tabellenzellen und werden als Attributwerte⁴⁷ bezeichnet (z.B. Name des Nutzers A). In einer relationalen Datenbank werden die Menge der Tabellen sowie deren Beziehung untereinander durch ein Datenbankschema definiert (Vgl. Anhang 1).⁴⁸

Es ist unerlässlich, dass Tabellenzeilen identifiziert werden. Dies wird durch die Verwendung eines Primärschlüssels bzw. durch ein eindeutiges Schlüsselattribut (z.B. Nutzer-ID) oder eine Attributkombination gewährleistet. In einer Tabelle kann aber auch ein so genannter Fremdschlüssel enthalten sein. Dabei handelt es sich um ein Attribut (bzw. Attributfolge), dem in einer anderen Tabelle eine Primärschlüsselfunktion zukommt. In der Funktion eines Fremdschlüssels können Primärschlüssel daher Beziehungen zwischen zwei Tabellen aufzeigen. In der Praxis setzen sich relationale Datenbanken aus einer Vielzahl untereinander in Beziehung stehender Tabellen zusammen.⁴⁹ Zusätzlich kann eine Datenbank weitere Datenbankobjekte wie u.a. Prozeduren⁵⁰ oder Triggers⁵¹ enthalten.⁵²

Während auf der konzeptionellen Ebene das Datenbankschema ein Modell der ganzen Informationsmenge in einer Datenbank darstellt, werden auf der externen Ebene nur Teilmengen der Informationen zur Anzeige gebracht. Dabei besteht die Möglichkeit, mittels einer Datenbankabfrage die Daten unterschiedlicher Tabellen miteinander zu verknüpfen. Auf diese Weise entstehen „virtuelle Ergebnistabelle[n]“⁵³ (Sichten), die in dieser Form im Datenbankschema nicht vorhanden sind. Vor diesem Hintergrund zeigt sich ein Charakteristikum relationaler Datenbanken. Die Daten können in beliebiger Form abgefragt werden, wobei sich das Abfrageresultat „dynamisch aus der Nutzerinteraktion“⁵⁴ ergibt. Fachanwendungen erzeugen auf ihrer Benutzeroberfläche nutzungsspezifische Sichten auf die Datenbankinhal-

⁴⁷ Für relationale Datenbanken können primär drei Datentypen als Attributwert unterschieden werden: Zahlen, Zeichenketten und Zeitangaben. Durch ihre zunehmende Effizienz können relationale Datenbanken auch so genannte „Large Object“-Datentypen (LOB) speichern. Hierbei handelt es sich z.B. um „Character Large Objects“ (CLOB), in welchen umfangreiche Texte bzw. Zeichenketten gespeichert werden können, oder „Binary Large Objects“ (BLOB), wobei es sich um grosse binäre Objekte wie z.B. Audiodaten oder Videosequenzen handelt. Vgl. Kemper/Eickler, Datenbanksysteme, S. 112, 420f.

⁴⁸ Steiner, Relationale Datenbanken, S. 16f; Unterstein/Matthiessen, Relationale Datenbanken, S. 19-29; Schubert, Datenbanken, S. 68-73; Dässler/Schwarz, Fachverfahren, S. 9-11; Glöde, Archivierung relationaler Datenbanken, S. 12-14.

⁴⁹ Ebd., S. 18-23; Unterstein/Matthiessen, Relationale Datenbanken, S. 30-32; Schubert, Datenbanken, S. 62-68; Glöde, Archivierung relationaler Datenbanken, S. 14f.

⁵⁰ Ein gespeichertes und ausführbares Programm in einem DBMS zur Erfüllung bestimmter Aufgaben. Vgl. Glöde, Archivierung relationaler Datenbanken, S. 135.

⁵¹ Eine benutzerdefinierte Prozedur, die zur Erfüllung einer bestimmten Bedingung automatisch durch das DBMS gestartet wird. Vgl. Kemper/Eickler, Datenbanksysteme, S.167.

⁵² Glöde, Archivierung relationaler Datenbanken, S. 16.

⁵³ Dässler/Schwarz, Fachverfahren, S. 12.

⁵⁴ Ebd., S. 9.

te. Eine solche Sicht kann einer Funktion entsprechen, die ein Anwender mit einer Fachanwendung ausführt. Sie stellt daher ein relevantes Informationsobjekt dar und kann sich aus Attributwerten mehrerer Tabellen⁵⁵ zusammensetzen.⁵⁶

Für Datenbankabfragen sowie allgemein für alle Datenbankoperationen (u.a. die Erzeugung und Veränderung von Tabellen sowie die Änderung und Hinzufügung von Daten) wird bei relationalen Datenbanken überwiegend die standardisierte Skript- bzw. Datenbanksprache SQL (Structured Query Language) verwendet. In der Praxis wird der SQL-Standard teilweise mit nicht-standardisierten Elementen erweitert. In der Konsequenz entwickeln sich SQL-Dialekte, die je nach Produkt eine unterschiedliche Ausprägung haben können.⁵⁷

3.2 Aspekte der dauerhaften Archivierung von Datenbankinhalten

Analog zu den allgemeinen Anforderungen an die digitale Langzeitarchivierung wird auch in Bezug auf die Datenbankarchivierung bzw. die Archivierung von Datenbankinhalten aus Fachanwendungen das Ziel verfolgt, Daten in einer lesbaren und systemneutralen Form zu erhalten und dabei die Integrität und Authentizität der Daten zu gewährleisten. Dabei kommt in der Datenbankarchivierung wie auch in Bezug auf andere digitale Objekte die Datenmigration zur Anwendung, während der emulative Archivierungsansatz infolge fehlender Lösungsansätze kaum verfolgt wird. Ein weiterer wesentlicher Aspekt bei der Archivierung von Datenbankinhalten ist aus archivischer Perspektive die Bewertung (Überlieferungsbildung). So ist grundsätzlich festzulegen, was der eigentliche Gegenstand bei der Archivierung von Datenbankinhalten ist bzw. in welchem Umfang Datenbankinhalte erhalten bleiben sollen. Zudem muss berücksichtigt werden, dass Datenbanken nach einem Archivierungsprojekt oftmals weiterbetrieben und nur selten vollständig „stillgelegt“ werden.⁵⁸

Im Rahmen der digitalen Langzeitarchivierung können in Bezug auf die Migration von Daten aus Datenbanken im Wesentlichen zwei Methoden unterschieden werden. Die erste Methode besteht darin, dass Inhalte einer Datenbank in Form statischer elektronischer Dokumente wie z.B. CSV-Dateien (Comma-Separated-Values) aus Datenbanken extrahiert werden (Datenextraktion). Durch diese Methode wird zwar eine Systemneutralität zur Quelldatenbank (Fachanwendung) geschaffen, doch werden dabei Informationsverluste und damit einhergehend der Verlust der Datenbankfunktionalität in Kauf genommen. So können beispielsweise

⁵⁵ Informationsobjekte können auch einzelne Datensätze, Tabellen oder auch ein ganzes Datenbanksystem umfassen. Vgl. Glöde, Archivierung relationaler Datenbanken, S. 19.

⁵⁶ Dässler/Schwarz, Fachverfahren, S. 9-12; Kemper/Eickler, Datenbanksysteme, S. 22. Vgl. Glöde, Archivierung relationaler Datenbanken, S. 18f; Keller-Marxer, Metadaten in CSP CHRONOS, S. 12.

⁵⁷ Ebd., S. 11; Glöde, Archivierung relationaler Datenbanken, S. 16f; Kemper/Eickler, Datenbanksysteme, S. 27.

⁵⁸ Ebd., S. 8; VdA, Bewertung elektronischer Fachverfahren, S. 1, 4; Kortyla, Übernahme aus Ratsinformationssystemen, S. 72; Keitel/Lang, Ingest von Fachverfahren, S. 35.

durch das CSV-Format alleine weder Metadaten noch Tabellenbeziehungen, die in einem relationalen Datenbankschema bestehen, festgehalten werden.⁵⁹

Mit der zweiten Methode wird das Ziel verfolgt, Inhalte so aus einer Datenbank (Fachanwendung) zu migrieren, dass möglichst geringe Informationsverluste auftreten und die Datenbankfunktionalität (z.B. Tabellenbeziehungen) erhalten bzw. rekonstruiert werden kann. Zu diesem Zweck werden mehrere offene und anerkannte Standards miteinander kombiniert. In diesem Zusammenhang sind insbesondere XML (Extensible Markup Language) und SQL (vgl. Kapitel 3.1) hervorzuheben. Während XML u.a. für Metadaten und Datenbankinhalte zur Anwendung gelangt, eignet sich SQL dazu, die Struktur einer Datenbank zu repräsentieren. In Anbetracht dessen, dass bei dieser Methode die Datenbankfunktionalität erhalten bleibt, kann nur hier „im eigentlichen Sinne von einer Datenbankarchivierung“⁶⁰ gesprochen werden. In der Archivpraxis haben Grundzüge beider Methoden zur Entwicklung kommerzieller und nicht kommerzieller Verfahren bzw. Anwendungsprogrammen geführt, um Inhalte für eine dauerhafte Archivierung aus Datenbanken zu migrieren (vgl. Kapitel 4.2).⁶¹

Auch bei der Migration von Informationen aus Datenbanken ist die Erhebung von Metadaten in Bezug auf die Gewährleistung der Integrität und Authentizität von Archivdaten zentral (vgl. Kapitel 2). Dabei muss berücksichtigt werden, dass Datenbanken in der Regel ein Bestandteil größerer Systeme wie Fachanwendungen sind, die ihrerseits wichtige Informationen in ihrer Programmlogik enthalten.⁶² Durch Metadaten bzw. Zusatzinformationen sollen daher Zusammenhänge hergestellt werden, die zuvor von den übergeordneten Systemen, in diesem Beispiel von Fachanwendungen, geleistet wurden.⁶³

Was die Identifikation des Archivierungsgegenstandes bei der Archivierung von Datenbankinhalten anbelangt, ist zunächst mittels einer archivischen Bewertung die grundsätzliche Archivwürdigkeit einer Datenbank bzw. einer Fachanwendung zu bestimmen. Wird eine Fachanwendung infolge einer inhaltlichen Analyse und unter Berücksichtigung übergreifender Bewertungsmodelle als archivwürdig beurteilt, ist festzulegen, ob Datenbankinhalte vollständig oder nur partiell archiviert werden sollen. Bei einer partiellen Archivierung muss analysiert werden, welche Prozessvorgänge, die durch eine Fachanwendung unterstützt werden, bzw. welche Datenbankinhalte oder Informationsbereiche einer Fachanwendung zu archivieren sind. Dabei ist zu beachten, dass sich Datenbankinhalte dem Nutzer einer Fachanwen-

⁵⁹ KOST, Katalog archivischer Dateiformate (strukturierte Daten: Datenbanken), o.S.; Dässler/Schwarz, Fachverfahren, S. 15f; Kortyla, Übernahme aus Ratsinformationssystemen, S. 72f.

⁶⁰ Dässler/Schwarz, Fachverfahren, S. 16.

⁶¹ Ebd., S. 15f; Kortyla, Übernahme aus Ratsinformationssystemen, S. 73f; ERPANET, Long-term Preservation of Databases, S. 15f; Heuscher/Keller-Marxer, Archival Access to Complex Relational Data, S. 4f.

⁶² KOST, Katalog archivischer Dateiformate (strukturierte Daten: Datenbanken), o.S..

⁶³ Keitel/Lang, Ingest von Fachverfahren, S. 35, 38.

dung nicht in ihrer konzeptionellen Struktur präsentieren, sondern als Informationsobjekte auf einer Benutzeroberfläche in Gestalt spezifischer Sichten (vgl. Kapitel 3.1). Vor diesem Hintergrund ist unter Berücksichtigung nutzungsspezifischer Sichten eine archivrelevante Datenauswahl zu treffen, um dadurch das Datenbankmodell einer Datenbank auf bestimmte relevante Tabellen zu reduzieren (Teildaten).⁶⁴ Die Schwierigkeit liegt dabei darin, dass Datenbankinhalte oft in einer komplexeren Struktur gespeichert sind, als diese in einer nutzungsspezifischen Sicht wahrgenommen werden können. Aus archivischer Perspektive ist daher oftmals eine Unterstützung durch die Anwender und insbesondere durch den Hersteller einer Datenbank bzw. Fachanwendung notwendig.⁶⁵

Eine weitere Schwierigkeit bei der Archivierung von Datenbankinhalten aus Fachanwendungen besteht darin, dass archivwürdige Datenbankinhalte noch vor ihrer Übernahme in ein Archiv infolge automatischer Löschroutinen und einer fehlenden Datenbankhistorisierung oder aufgrund von Datenbankmodifizierungen (Weiterentwicklung einer Datenbank) verloren gehen können. Eine mögliche Strategie, mit Löschungen und Veränderungen umzugehen, besteht darin, Datenbanken frühzeitig in periodischen Abständen, z.B. bevor eine Löschung ansteht, ganz oder partiell zu archivieren bzw. so genannte Snapshots zu erzeugen. Diese Strategie führt aber dazu, dass Datenbankinhalte in einem Archiv redundant gespeichert werden. Eine alternative Möglichkeit bietet die inkrementelle Archivierung. Hierbei werden in definierten Zeitabständen bestimmte Datenbankinhalte in ein digitales Archiv überführt, wodurch dieses kontinuierlich wächst und die Datenredundanz minimiert werden kann.⁶⁶ Um einen Datenverlust zu verhindern, ist es wichtig, dass Archive sich frühzeitig für eine Archivierung aus laufenden Datenbanksystemen bzw. Fachanwendungen einsetzen.⁶⁷

⁶⁴ VdA, Bewertung elektronischer Fachverfahren, S. 1f; Dässler/Schwarz, Fachverfahren, S. 12f; KOST, Bewertung von Datenbanken, o.S.; Naumann, Übernahme von Daten aus Fachanwendungen, S. 31f. Vgl. Keller-Marxer, Metadaten in CSP CHRONOS, S. 11-13; Glöde, Archivierung relationaler Datenbanken, S. 18f.

⁶⁵ Keitel, Digitale Archivierung beim Landesarchiv Baden-Württemberg, S. 23.

⁶⁶ Ebd., S. 4; ERPANET, Long-term Preservation of Databases, S. 12f; Ribeiro, Database Preservation, S. 1. Vgl. Dässler/Schwarz, Fachverfahren, S. 8.

⁶⁷ Dässler/Schwarz, Fachverfahren, S. 8.

4. Verfahren zur Datenbankarchivierung

4.1 Analyse der Fallbeispiele und Archivierungsziele

4.1.1 Erstes Fallbeispiel: Die Fachanwendung „socialWEB“ des „Kompetenzzentrums Jugend und Familie Schlossmatt“

Das „Kompetenzzentrum Jugend und Familie Schlossmatt“ (KJFS) ist dem Jugendamt bzw. der Direktion für Bildung, Soziales und Sport der Stadt Bern unterstellt (Art. 25 OV).⁶⁸ Die Finanzierung der Einrichtung wird durch den Kanton, namentlich über einen Leistungsvertrag zwischen der Gesundheits- und Fürsorgedirektion (GEF) und dem KJFS wahrgenommen.⁶⁹ Bei der Umsetzung seiner Aufgaben wird das KJFS durch die datenbankgestützte Fachanwendung „socialWEB“ der „socialweb Software GmbH“ unterstützt. Die Fachanwendung ist im KJFS seit dem Jahre 2009 in Betrieb und soll zurzeit von inaktiven Daten entlastet werden. Vor diesem Hintergrund wird sich das Stadtarchiv Bern in absehbarer Zeit der Archivierung von Datenbankinhalten aus der Fachanwendung „socialWEB“ annehmen.⁷⁰

Was die Aufgaben des KJFS anbelangt, wird im Leistungsvertrag zwischen dem GEF und dem KJFS festgehalten, dass „die Entwicklung und Förderung der betreuten Kinder und Jugendlichen“⁷¹ im Zentrum stehe. Im Leitbild des KJFS wird diese Aufgabe konkretisiert. So besteht die Aufgabe darin, Kinder, Jugendliche, junge Mütter und Familien bei der Bewältigung von Krisen und schwieriger Lebenssituationen zu unterstützen. Konkret sollen Jugendliche, die in ihrer Entwicklung „gefährdet“ sind, geschützt und in ihrer Entwicklung gefördert werden. Junge Mütter sollen lernen, während das Wohl des Kindes gesichert ist, ihre Mutterrolle wahrzunehmen. Zudem werden Familien unterstützt, innerfamiliäre Situationen zu klären. Zu diesem Zweck führt das KJFS stationäre Einrichtungen (Kinder- und Jugendheim, Wohngemeinschaften) und bietet ambulant eine Familienbegleitung an.⁷²

Die Basis der Software „socialWEB“, die das KJFS bei der Umsetzung seiner Aufgaben bzw. Prozessvorgänge unterstützt, bildet eine MySQL-Datenbank (Back-End). Das Datenbankschema umfasst ca. 250 Tabellen, wobei die Schlüsselbeziehungen applikatorisch definiert sind.⁷³ Die Software „socialWEB“ (Front-End) ist modular aufgebaut. In den Modulen „*Klient*“ und „*Mitarbeiter*“ werden Informationen über die Klientenunterstützung sowie zur Mitarbeiterorganisation festgehalten. Im Modul „*Kontakte*“ werden Stamm- und Personendaten verwaltet. Spezialfunktionen zur Erstellung von Protokollen und Arbeitsplänen sowie für die Buch-

⁶⁸ „Verordnung über die Organisation der Stadtverwaltung“ (SSSB), 27.2.2001, S. 13f.

⁶⁹ Vgl. GEF, Allgemeine Vertragsbestimmungen zum Leistungsvertrag 2014, 30.4.2014.

⁷⁰ Interview von Donato Raselli mit Martina Suter, 25.4.2014.

⁷¹ GEF, Allgemeine Vertragsbestimmungen zum Leistungsvertrag 2014, 30.4.2014, S. 3.

⁷² KJFS, Leitbild des Kompetenzzentrums Jugend und Familie Schlossmatt, S. 1-3.

⁷³ Interview von Donato Raselli mit Matthias Andenmatten, 2.5.2014.

haltung und das Rechnungswesen finden sich im Modul „*Diverses*“. Die Klienten können ferner nach Gruppen geordnet werden, um Zugriffsrechte zu definieren.⁷⁴

Im Modul „*Klient*“ wird die Betreuung eines Klienten durch so genannte Journaleinträge dokumentiert (vom Erstkontakt bis zum Austritt). Konkret werden in das Journal Angaben zur Lebenssituation, Befindlichkeit und Familienstruktur eines Klienten sowie zu Abklärungen und Maßnahmen, die während eines Aufenthaltes getroffen wurden, aufgenommen. Zudem können im Journal Klienten bezogene Ereignisse, Auffälligkeiten und Vereinbarungen aufgeführt werden. Ferner können im Modul die Termine und die Aufgaben der Klienten dokumentiert werden. Ein weiteres Modulelement bilden die Prozessvariablen. Einerseits zeigen Sie auf, in welcher Betreuungsphase sich ein Klient befindet (Klärungs-, Aufenthalts- oder Austrittsphase), andererseits dienen sie dazu, anhand von Kategorien phasenspezifische Informationen aufzunehmen (z.B. Schulsituation).⁷⁵

Im Modul „*Mitarbeiter*“ können die Mitarbeiter individuell ihre Termine und Aufgaben verwalten, „Freiwünsche“ (z.B. Ferien) angeben oder die Arbeitszeit erfassen.⁷⁶ Das Modul „*Kontakt*“ dient zur Erfassung von Personalangaben sowie von Klienten bezogenen Stamm- und Personendaten, die in der Anwendung auch statistisch ausgewertet werden können. Insbesondere die Stammdaten enthalten umfassende Informationen zu den Klienten. So werden mittels standardisierter Kategorien beispielsweise Angaben über die Familienstruktur, den Inhaber der Obhut und des Sorgerechts, den kulturell-religiösen Hintergrund, die Sprachkenntnisse, die Nationalität und den Wohnsitz der Klienten erfasst.⁷⁷

Im Modul „*Diverses*“ ist die Protokollfunktion zentral, da hiermit Besprechungen mit den Klienten protokolliert und anschließend evaluiert werden können. Konkret sind dies Klärungsgespräche, die verfasst werden, wenn ein Klient ein Angebot des KJFS in Anspruch nimmt, Zielgespräche, die beschreiben, welche Ziele ein Klient während seines Aufenthaltes erreichen möchte, sowie Abschlussgespräche. Ferner kann das Personal mit dieser Funktion auch interne Teamsitzungen protokollieren. Die Funktion „Arbeitsplanung“ ermöglicht dem Personal, detaillierte Einsatzpläne für die Betreuung der Klienten zu generieren, und mittels

⁷⁴ socialweb Software GmbH, Produktdokumentation, S. 4-36; Interview von Donato Raselli mit Martina Suter, 25.4.2014.

⁷⁵ socialweb Software GmbH, Benutzerhandbuch. KlientInnenmodul, S. 3-39; socialweb Software GmbH, Benutzerhandbuch. KlientInnenmodul – Zielprozesse und Ein-/Austritt von KlientInnen, S. 2-20; socialweb Software GmbH, Produktdokumentation, S. 5-7.

⁷⁶ socialweb Software GmbH, Benutzerhandbuch. MitarbeiterInnenmodul, S. 2-22; socialweb Software GmbH, Benutzerhandbuch. Arbeitszeiterfassung, S. 2-19; socialweb Software GmbH, Benutzerhandbuch. Einsatz/Freiwünsche und Kontingente, S. 2-12; socialweb Software GmbH, Produktdokumentation, S. 8-10, 20f; Interview von Donato Raselli mit Martina Suter, 25.4.2014.

⁷⁷ socialweb Software GmbH, Benutzerhandbuch. Kontaktmodul, S. 2f; socialweb Software GmbH, Produktdokumentation, S. 11-13, 28-29; Interview von Donato Raselli mit Martina Suter, 25.4.2014. Vgl. KJFS, Matrix.

der Funktionen „Buchhaltung“ und „Rechnungswesen“ können die Ausgaben und die Einnahmen verwaltet und die Ausgaben mit den Aufenthaltstagen verrechnet werden.⁷⁸

Vor dem Hintergrund der Aufgaben des KJFS und der Funktionsweise der Fachanwendung „socialWEB“ lassen sich mehrere Prozesse identifizieren, die durch die Fachanwendung unterstützt werden. So können die Kinder- und Jugendbetreuung, die Betreuung junger Mütter und die Familienbegleitung als Kernprozesse identifiziert werden. Ausgehend von den Prozessvariablen im Klientenmodul bestehen diese Prozesse, mit Ausnahme der Familienbegleitung, aus den Teilprozessen „Klärungs-, Aufenthalts- und Austrittsphase“. Bei der Familienbegleitung wird der stationäre Teilprozess durch einen ambulanten ersetzt. In Bezug auf die Kernprozesse besteht die Rolle der Fachanwendung darin, Hintergrundinformationen zu den Klienten zu verwalten und den Verlauf der Betreuung zu organisieren und zu dokumentieren (z.B. Protokolle, Journaleinträge). Dabei kommen bei allen Kernprozessen die gleichen Funktionen zur Anwendung. Für die Familienbegleitung und die Betreuung junger Mütter werden leicht modifizierte Stammdaten erfasst (Angaben zum Kleinkind). Im Weiteren unterstützt die Fachanwendung Supportprozesse. Ein erster Supportprozess umfasst die Mitarbeiterorganisation (z.B. Zeiterfassung) und ein zweiter Supportprozess die Buchhaltung und das Rechnungswesen. In Bezug auf diese Prozesse besteht die Rolle der Fachanwendung darin, ein Instrumentarium für die Prozessumsetzung zur Verfügung zu stellen.⁷⁹

Werden die durch die Fachanwendung unterstützten Prozessvorgänge im Hinblick auf analoge Unterlagen untersucht, wird sichtbar, dass Papierakten entstehen, die mit der Anwendung nicht unmittelbar in Verbindung stehen. Wiederum fließen Daten aus der Fachanwendung in Papierakten ein. Konkret wird bei den Kernprozessen die Mehrheit der in der Anwendung erfassten Klientendaten am Ende eines Betreuungsverhältnisses ausgedruckt und in analoge Klientendossiers aufgenommen (Stammdaten, Prozessvariablen, Journaleinträge, Protokolle). Tritt jedoch ein Klient nach der Klärungsphase nicht in ein Betreuungsangebot ein, entsteht kein analoges Dossier und die erfassten Daten bleiben ausschließlich in der Anwendung. Die Termine und Aufgaben der Klienten sowie die Evaluationen zu den protokollierten Klientengesprächen werden nicht ausgedruckt. Unabhängig von der Fachanwendung enthalten die Klientendossiers Unterlagen wie z.B. Protokolle von Elterngesprächen, Zielsetzungsentwürfe der Klienten, Datenschutzpapiere, Suchtmittelvereinbarungen oder Berichte und Verfügungen von Behörden. Was das Personal anbelangt, werden die Team-sitzungsprotokolle ausgedruckt und archiviert; die Personen- und Stammdaten sowie die

⁷⁸ socialweb Software GmbH, Benutzerhandbuch. Protokolle, S. 2-13; socialweb Software GmbH, Benutzerhandbuch. Arbeitsplanung, S. 2-49; socialweb Software GmbH, Benutzerhandbuch. Buchhaltung, S. 2-23; socialweb Software GmbH, Produktdokumentation, S. 22-27; Interview von Donato Raselli mit Martina Suter, 25.4.2014.

⁷⁹ Interview von Donato Raselli mit Martina Suter, 25.4.2014.

Daten zu Terminen und Aufgaben bleiben in der Fachanwendung. In Bezug auf den Supportprozess der Mitarbeiterorganisation werden die Arbeitspläne und Arbeitszeitabrechnungen ausgedruckt und abgelegt. Auch beim Supportprozess Buchhaltung und Rechnungswesen werden zentrale Daten in Papierform aufbewahrt (z.B. Jahresabschluss, Rechnungsstellung). Zusätzlich werden diese Daten laufend an die Stadtbuchhaltung übertragen.⁸⁰

Die Hauptaufgaben des KJFS werden durch die Kernprozesse umgesetzt, weshalb die dabei entstehenden Datenbankinhalte grundsätzlich als archivwürdig eingestuft werden können. Eine wichtige Rolle spielt dabei der Umstand, dass bestimmte Daten nicht konsequent in analogen Klientendossiers übernommen werden und in der Folge nur in der Anwendung „socialWEB“ vollständig erhalten sind. Nicht in die Dossiers fließen die Protokollevaleuierungen ein; ferner alle Informationen zu Klienten, die nach einem Erstkontakt nicht in ein Angebot eingetreten sind. Die Personen- und Stammdaten (Klientel/Personal) sind, obwohl auch analog verfügbar, in Bezug auf eine Archivierung von besonderem Interesse, da sie als digitale Datensammlung vielfältige Auswertungsmöglichkeiten zulassen.⁸¹ Für die digitale Archivierung weniger interessant sind die Teamsitzungsprotokolle, da diese konsequent ausgedruckt werden, sowie aus inhaltlicher Perspektive die Termin- und Aufgabenverwaltung der Klienten und des Personals. Die Datenbankinhalte des Supportprozesses „Mitarbeiterorganisation“ beziehen sich nur indirekt auf die Kernaufgaben des KJFS und sind daher inhaltlich, von den Personen- und Stammdaten abgesehen, ebenfalls von geringerem Interesse (z.B. Zeiterfassung).⁸² Ähnlich verhält es sich mit den Daten des Supportprozesses Buchhaltung und Rechnungswesen, da diese zum Kerngeschäft der Stadtbuchhaltung gehören. Vor diesem Hintergrund kann die Archivierung der Datenbankinhalte, die während der Umsetzung der Kernprozesse (die Betreuung der Kinder/Jugendlichen und der jungen Mütter sowie die Familienbegleitung) entstanden sind, als Archivierungsziel festgelegt werden. Dabei sollen im Archivgut die Prozesse (Prozessinstanzen) nachvollzogen werden können und die Integrität und Authentizität der Archivdaten muss gewährleistet sein. Konkret handelt es sich bei der Datenbank um die Informationsobjekte „Personen- und Stammdaten“ (Klientel/Personal), „Prozessvariablen“, „Journaleinträge“ und klientenbezogene „Protokolle“ (inkl. Evaluationen). Diese Informationsobjekte umfassen ca. 70-80 der insgesamt ca. 250 Tabellen.⁸³

⁸⁰ Ebd.

⁸¹ Vgl. Dässler/Schwarz, Fachverfahren, S. 7f.

⁸² Auch das Stadtarchiv Bern erachtet diese Datenbankinhalte als nicht archivwürdig. Vgl. Interview von Donato Raselli mit Martina Suter, 25.4.2014.

⁸³ Interview von Donato Raselli mit Matthias Andenmatten, 2.5.2014.

4.1.2 Zweites Fallbeispiel: Die Fachanwendung „ESCADA²“ des „Erziehungsdepartements Kanton Basel-Stadt“

Das Erziehungsdepartement (ED) ist eines von sieben Departementen der Staatsverwaltung des Kantons Basel-Stadt und ist zuständig für die Wahrnehmung staatlicher Aufgaben in den Bereichen Erziehung, Bildung und Sport (§ 17/18 KV, § 26 OG).⁸⁴ Für die Umsetzung schuladministrativer Aufgaben setzt das ED kantonal die datenbankgestützte Fachanwendung „ESCADA²“ der Firma „JCS Software AG“ ein, die im Zuge des Projekts „Einführung einer neuen Schulverwaltungssoftware Basel-Stadt“ (NeSuBa) im Schuljahre 2004/2005 eingeführt wurde. Mit der Einführung der Fachanwendung „ESCADA²“ wurde das Ziel verfolgt, die Administration öffentlicher Bildungsinstitutionen in den Bereichen der Volks- und Mittelschule sowie der Berufsbildung zu vereinheitlichen. Als Schulverwaltungssoftware dient die Fachanwendung primär als Arbeitsinstrument für Schulleitungen (Volksschule) und Rektorate (Mittelschule, Berufsbildung). Die Datenbankinhalte von „ESCADA²“ kommen aber auch in weiteren Bereichen und auf anderen Ebenen des Bildungssystems zum Einsatz (z.B. ED, Volksschulleitung, Zentrum Brückenangebote, Institut für Unterrichtsfragen und Lehrerfortbildung). Nach neunjährigem Betrieb soll die Anwendung – wie bereits im ersten Fallbeispiel – im Zuge eines Archivierungsprojekts (Staatsarchiv Basel-Stadt) von inaktiven Daten entlastet werden. Infolge des großen Einsatzgebietes der Fachanwendung „ESCADA²“ konzentriert sich diese Untersuchung auf den Einsatz der Fachanwendung auf der Ebene der einzelnen Schuladministrationen im Bereich der Volksschule (GST).⁸⁵

Aufgabe der Volksschule des Kantons Basel-Stadt ist es, die geistige und körperliche Entwicklung der Schüler zu fördern, damit diese „sowohl den allgemein menschlichen als auch den beruflichen Anforderungen des Lebens gewachsen sind“⁸⁶ (§ 3a SG). Zu diesem Zweck vermittelt die Volksschule Kenntnisse und Fähigkeiten, die „in der Gesellschaft und in der Berufswelt“⁸⁷ (§ 3b SG) benötigt werden. Die Volksschule des Kantons Basel-Stadt besteht aus einer Primar- und Sekundarstufe I. Die Führung der einzelnen Volksschulstandorte obliegt jeweils einer Schulleitung (§ 87c SG).⁸⁸ Die Aufgaben der Schulleitung (inkl. Schulsekretariate) umfassen die Erarbeitung des Schulprogramms, die Festlegung pädagogischer Schwerpunkte, die Klärung organisatorischer Belange und die Verwaltung betrieblicher Mittel. Im Weiteren trägt sie die Personalverantwortung (u.a. Personalauswahl, Klassenzutei-

⁸⁴ „Verfassung des Kantons Basel-Stadt“, 23.3.2005, S. 5; „Gesetz betreffend die Organisation des Regierungsrates und der Verwaltung des Kantons Basel-Stadt“, 22.4.1976, S. 7.

⁸⁵ Agnolazza, Projekt „NeSuBa“, S. 22; Interview von Donato Raselli mit Christoph Döbeli, 14.5.2014. Vgl. JCS Software AG, ESCADA², S. 2.

⁸⁶ „Schulgesetz“, 4.4.1929, S. 3.

⁸⁷ Ebd., S. 3.

⁸⁸ Diese Struktur entstand im Zuge eines Schulharmonisierungsprojektes und betrifft Kinder, die ab dem Frühjahr 2002 geboren wurden. Ältere Schulkinder besuchen die Primar-, Orientierungs- und Weiterbildungsschule. Vgl. URL: <<http://www.schulharmonisierung-bs.ch/grundlagen/schulsystem-alt-neu>> (17.5.2014); „Schulgesetz“, 4.4.1929, S. 34.

lung, Pensenplanung) und leitet den Schulbetrieb (u.a. Disziplin, Dispensation, Schülerbeurteilung, Klassenbildung, Stundenplanfestlegung) (§ 6-16 Schulleitungsverordnung).⁸⁹

Die Anwendung „ESCADA²“ unterstützt die Schulleitungen bei der Umsetzung ihrer Aufgaben. Die Basis von „ESCADA²“ besteht aus verteilten Datenbanken (Microsoft SQL-Server), die zusammen einen „einzigsten Datenbestand[...]“⁹⁰ bilden (Back-End). Konkret setzt sich „ESCADA²“ aus vier Datenbanken zusammen. Eine Datenbank (JCS-Gemeinsam) dient der Verwaltung von Stammdaten, die für alle Schultypen relevant sind (z.B. Adressen, zentrale Steuerdaten, Schülerdaten, Schülerchroniken und Personaldaten). Die übrigen Datenbanken enthalten schultypische Daten für die Volksschule (JCS-GST), die Mittelschule (JCS-GYM) sowie die Berufsbildung (JCS-AGS). Innerhalb dieser Datenbanken sind die Daten „nach Mandant (Schule) gekapselt“⁹¹. Das bedeutet, dass jede Schule als Mandant auf ihre Daten zugreifen kann. Zu diesem Zweck werden die Datensätze einer Schule über einen Mandantenschlüssel identifiziert. Die Schlüsselbeziehungen sind in „ESCADA²“ auf der Datenbankebene implementiert. Für die Volksschuladministration sind die Datenbanken JCS-GST und JCS-Gemeinsam, deren Datenbankschema je ca. 200 Tabellen umfassen, relevant.⁹²

Auf der Ebene der Applikation (Front-End) besteht die Fachanwendung „ESCADA²“ aus einem Konglomerat von Modulen, wobei jedes Modul einer eigenständigen Applikation entspricht. In jeder Applikation können über eine Symbolleiste bestimmte „Views“ (hier „Fenster“) geöffnet werden. Die „Views“ ermöglichen einerseits den Zugriff auf bestimmte Datentypen (z.B. Schülerdaten) und andererseits die Bearbeitung der Daten über ein Kontextmenü. Jedem Schultyp stehen mindestens drei Module zur Verfügung. In der Volksschulverwaltung kommen die Module „SchuleGST“ (z.B. Schülerverwaltung), „FakturaGST“ (z.B. Rechnungsstellungen) und „PLSGST“ (Personal- und Lohnverwaltungssystem) zur Anwendung. Zusätzlich gibt es Applikationen, die sich auf alle Module von „ESCADA²“ beziehen. Hierzu gehören die „Finanzadministration“ und die „Administration“ (z.B. User-Verwaltung). Ferner verwendet „ESCADA²“ das Reporting-Tool „Crystal Reports“, um spezifische Inhalte abzurufen und zu definierten Berichten zu verarbeiten (z.B. Zeugnisse, Stammbblätter).⁹³

Mit dem Modul „SchuleGST“ lassen sich primär Schüler- und Lehrerdaten sowie Angaben zu den Klassen (inkl. Kurse) verwalten. Konkret können beim Schuleintritt und während der Schulzeit Kontakt- und Personendaten der Schüler (z.B. Sprache, Nationalität, Konfession),

⁸⁹ „Verordnung für die Schulleitungen der Volksschulen“, 26.6.2012, hier S. 2-8.

⁹⁰ JCS Software AG, ESCADA², S. 1.

⁹¹ Ebd.

⁹² Ebd., S. 1-3; Interview von Donato Raselli mit Claude Stettler und Ali Aydin, 14.5.2014. Vgl. JCS Software AG, ESCADA². Stammdaten; JCS Software AG, Benutzerhandbuch. SchuleGST, S. 11-18.

⁹³ Ebd., S. 3; JCS Software AG, ESCADA². Grundfunktionen; Interview von Donato Raselli mit Claude Stettler und Ali Aydin, 14.5.2014.

Angaben zum Schulbesuch (z.B. Absenzen, Ein- und Austritt, Schülerlaufbahn) und disziplinarische Maßnahmen aufgenommen werden. Zudem werden im Modul die Klassen-, Niveau- und Schulhauszuteilung sowie die Notenverwaltung vorgenommen. Was die Klassenorganisation anbelangt, bietet das Modul die Möglichkeit, Schulklassen zu bilden (Schülerzuweisung), Studentafeln zu definieren und Fächerspiegel zu erfassen. Im Weiteren werden den Lehrpersonen in diesem Modul die entsprechenden Unterrichtsstunden zugeteilt (inkl. Pflichtstundenzahlfestlegung). Ferner können allgemeine Adressen (z.B. von Drittpersonen) verwaltet und bei Bedarf bestimmte Gremien (z.B. Elternrat) gebildet werden.⁹⁴

Das Modul „*PLSGST*“ unterstützt die Personal- und Lohnverwaltung. Dabei sind insbesondere die Aufnahme und Verwaltung von Personendaten sowie von administrativen Angaben zu den Lehrpersonen von Bedeutung (u.a. Adresse, Personalnummer, Ein- und Austrittszeitpunkt, Lohnklasse, allfällige Nebenbeschäftigung). Im Weiteren wird festgehalten, welche Ausbildung eine Lehrperson besitzt und welche Fächer diese unterrichten darf. Zentrale Modulfunktionen sind zudem die Erfassung und Bearbeitung der Arbeitsverträge und der Stundenzuteilung (inkl. Einzelstundenbuchhaltung). Hinsichtlich der Verträge werden u.a. Daten bezüglich der Funktion, der Anstellungsart, der Anstellungsdauer und des Beschäftigungsgrades aufgenommen. Weitere Modulfunktionen betreffen die Verwaltung zusätzlicher Ämter von Lehrpersonen (z.B. Amt der Pensenplanung), die Dokumentation von Absenzen (u.a. Dauer, Unfall/Krankheit) sowie die Regelung von Stellvertretungen (u.a. Fächer, Anzahl Lektionen, Lohnansätze). Schließlich werden in diesem Modul die Lohnzahlungen vorbereitet und in eine SAP-Anwendung exportiert.⁹⁵

Für die Verwaltung und Verrechnung zusätzlicher Angebote (z.B. Tagesstrukturangebot „Mittagstisch“) steht den Schulleitungen das Modul „*FakturaGST*“ zur Verfügung. Hier können die von den Schülern beanspruchten Zusatzangebote erfasst und in Rechnung gestellt werden. In einer Debitorenübersicht werden die Rechnungen verwaltet und es können z.B. Mahnungen ausgelöst sowie Zahlungsvereinbarungen (z.B. Teilrechnungen) getroffen werden.⁹⁶ Was die modulübergreifenden Zusatzapplikationen „*Administration*“ und „*Finanzadministration*“ anbelangt, können über die Anwendung „*Administration*“ die Datenbank-User definiert, die Zugriffsberechtigungen vergeben und schultypisch Steuerdaten aufgenommen werden (exemplarisch „*SchuleGST*“: u.a. Absenzgründe, diszipl. Maßnahmen, Kriterien So-

⁹⁴ JCS Software AG, Benutzerhandbuch. SchuleGST, S. 11-84, 186f; JCS Software AG, ESCADA². Grundschulen; Interview von Donato Raselli mit Claude Stettler und Ali Aydin, 14.5.2014.

⁹⁵ JCS Software AG, Benutzerhandbuch. PLSGST, S. 13-43; Interview von Donato Raselli mit Claude Stettler und Ali Aydin, 14.5.2014.

⁹⁶ JCS Software AG, Benutzerhandbuch. FakturaGST, S. 12-33; Interview von Donato Raselli mit Claude Stettler und Ali Aydin, 14.5.2014.

zialverhalten, Fächer, Stundentafeln). Mit der Applikation „*Finanzadministration*“ werden u.a. Schulbudget, Konten, Kostenstellen und Rechnungsnummern verwaltet.⁹⁷

Durch die Untersuchung der Schulleitungsaufgaben (Volksschule) und der Funktionsweise der Fachanwendung „*ESCADA*²“ lassen sich mehrere Prozessvorgänge identifizieren, die über die Fachanwendung umgesetzt werden. Konkret können dabei drei Kernprozesse unterschieden werden: Die „Schulbetriebsorganisation“ (Teilprozesse: Schüler- und Klassenverwaltung), die „Personalorganisation“ (Teilprozesse: Verwaltung der Lehrpersonen und Vorbereitung der Lohnzahlungen) und die „Organisation der Zusatzangebote“ (Teilprozesse: Verwaltung und Verrechnung des Tagesstrukturangebots). Bei der Umsetzung dieser Kernprozesse besteht die Rolle der Fachanwendung jeweils darin, Informationen zu verwalten (z.B. Noten und Personendaten der Schüler) und eine effiziente sowie strukturierte Umsetzung organisatorischer Aufgaben zu erleichtern (z.B. Verrechnungen, Klasseneinteilungen). In Bezug auf Supportprozesse unterstützt „*ESCADA*²“ die Budgetverwaltung (Finanzadministration). Die Rolle der Anwendung besteht auch hier darin, durch ein spezifisches Instrumentarium eine effiziente und organisierte Umsetzung des Prozesses zu ermöglichen.⁹⁸

Werden diese Prozessvorgänge hinsichtlich der Entstehung analoger Unterlagen untersucht, wird sichtbar, dass wie im ersten Fallbeispiel (vgl. Kapitel 4.1.1) bestimmte Daten aus der Fachanwendung direkt in Papierakten fließen, während andere Unterlagen unabhängig von der Anwendung entstehen. Beim Kernprozess der Schulbetriebsorganisation werden für die Schülerverwaltung diverse analoge Unterlagen wie z.B. Notizen, Unfallmeldungen und Korrespondenzen (z.B. bezüglich Absenzen oder disziplinarischer Maßnahmen) in Schülerdossiers aufgenommen. Entsprechend der jeweiligen Alltagspraxis einer Schulleitung können die Dossiers auch Blätter mit Stammdaten enthalten, die aus der Anwendung ausgedruckt werden. Weitere Anwendungsprodukte sind in diesem Kontext zudem Zeugnisse und Notenlisten. Ferner verfassen die Lehrpersonen individuelle Lernberichte zu den Schülern. Bei der Klassenverwaltung sind es insbesondere Statistiken, Stundentafeln, Absenz-, Klassen- und Schulbesuchslisten, die aus der Anwendung exportiert werden. Zusätzlich entstehen bei der Verwaltung der Klassen diverse Konferenzprotokolle und allgemeine Verwaltungsunterlagen (z.B. Projektunterlagen). Grundsätzlich werden Informationen im Rahmen der Schulbetriebsorganisation nur im Bedarfsfall explizit aus der Fachanwendung exportiert.⁹⁹

In Bezug auf den Kernprozess der Personalorganisation werden in den Schuladministratio-
nen für die Verwaltung des Lehrpersonals Personaldossiers geführt. Diese enthalten u.a.

⁹⁷ Ebd., S. 8-10; JCS Software AG, Benutzerhandbuch. SchuleGST, S. 11-18; JCS Software AG, Benutzerhandbuch. PLS-GST, S. 10f; Interview von Donato Raselli mit Claude Stettler und Ali Aydin, 14.5.2014.

⁹⁸ Interview von Donato Raselli mit Christoph Döbeli, 14.5.2014.

⁹⁹ Ebd. Vgl. JCS Software AG, *ESCADA*². Grundschulen.

Korrespondenzen und Unterlagen aus Mitarbeitergesprächen. Die Dossiers können teilweise auch Arbeitsverträge und Stammdatenblätter enthalten, die aus der Fachanwendung exportiert werden. Ferner können in der Anwendung z.B. Pensen sowie DAG-Formulare (Dienstaltersgeschenk) erstellt und ausgedruckt werden. Im Hinblick auf die Vorbereitung der Lohnzahlungen müssen in der Anwendung in analoger Form Auszahlungsbelege erzeugt und an das ED weitergeleitet werden. Was die Prozesse der Budgetverwaltung (Supportprozess) und der Zusatzangebotsorganisation anbelangt, sind bezüglich analoger Unterlagen primär Rechnungen und Korrespondenzen hervorzuheben. Ferner werden ausgehende Rechnungen sowie Listen zu den Tagesstrukturen aus der Anwendung exportiert.¹⁰⁰

Ausgehend von den Aufgaben der Schulleitungen (inkl. Schulsekretariaten) im Bereich der Volksschule und der Analyse der damit einhergehenden Prozessvorgänge, können diejenigen Datenbankinhalte von „ESCADA²“, die bei der Umsetzung der Kernprozesse entstehen, weitgehend als archivwürdig eingestuft werden. Diese Einschätzung scheint auch mit der bisherigen Übernahmepraxis des Staatsarchivs übereinzustimmen. Gemäß dem ED erachtet das Staatsarchiv Akten als archivwürdig, die sich direkt auf die Schüler und Lehrpersonen beziehen.¹⁰¹ In Bezug auf die Akten der Schulbetriebsorganisation ist zu berücksichtigen, dass die schülerbezogenen Personendaten (inkl. Noten) teilweise vom „Statistischen Amt“ übernommen werden. Zu diesem Zweck werden die Daten aus der Datenbank exportiert, wodurch künftige Auswertungen erschwert oder verhindert werden könnten. Zudem fließen diese Daten nicht konsequent in die Schülerdossiers ein. Bezüglich der Klassenverwaltung können Informationen wie z.B. Klassenzuteilungen oder Fächerspiegeldefinitionen einen Beitrag leisten, um Informationszusammenhänge bezüglich des behördlichen Handelns im Schulbereich zu erhalten.¹⁰² Bezüglich der lehrerbezogenen Personendaten (Personalorganisation) können ähnliche Argumente wie in Bezug auf die Schüler-Personendaten hervorgebracht werden. Die Lohnzahlungsvorbereitung ist von geringer Bedeutung, da sie weitgehend zu den Aufgaben des Finanzdepartements gehört. Hinsichtlich der schulischen Zusatzangebote steht in erster Linie das Rechnungswesen im Vordergrund, das inhaltlich vernachlässigt werden kann. Weniger relevant sind zudem die Daten, die in Verbindung mit der Budgetverwaltung (Supportprozess) entstehen, da diese ebenfalls über das Finanzdepartement weiterbearbeitet werden. Vor diesem Hintergrund kann auch bei diesem Fallbeispiel die Archivierung eines Großteils der Kernprozesse (Schulbetriebsorganisation: Schüler- und Klassenverwaltung; Personalorganisation: Verwaltung des Lehrpersonals) und der Mehrheit der in diesem Zusammenhang entstehenden Datenbankinhalte sowie die Gewährleistung der

¹⁰⁰ Ebd.

¹⁰¹ Das Staatsarchiv Basel-Stadt stellt für diese Untersuchung keine Informationen zur (Vor-)Bewertung der Datenbankinhalte (ESCADA²) zur Verfügung. Vgl. Interview von Donato Raselli mit Christoph Döbeli, 14.5.2014.

¹⁰² Vgl. VdA, Bewertung elektronischer Fachverfahren, S. 2.

Integrität und Authentizität der Archivdaten als Archivierungsziel festgelegt werden. Dabei stehen die folgenden Informationsobjekte im Vordergrund: „schülerbezogene Personendaten“ (inkl. Absenzen), „Schülerlaufbahnen“ (Chroniken), „disziplinarische Maßnahmen“, „Notenverwaltung“, „Zuteilungen“ (Klasse, Niveau, Schulhaus), „Fächerspiegel“ und „lehrerbezogene Personendaten“ (inkl. Absenzen, Ausbildung, Pensen, Verträge, Zusatzämter).

4.2 Die Spezifikation und Funktionsweise der Archivierungsverfahren

4.2.1 IngestList (Landesarchiv Baden-Württemberg)

Die Applikation „IngestList“¹⁰³ wurde im Rahmen des Projekts „Konzeption für ein digitales Landesarchiv“ (2005-2009) durch das Landesarchiv Baden-Württemberg entwickelt. Mit der Entwicklung der javabasierten Anwendung wurde das Ziel verfolgt, die Übernahme beliebiger Dateitypen und -Formate in ein digitales Archiv zu erleichtern. Konkret soll die Anwendung „elektronische Unterlagen in einem sicheren und glaubwürdigen Verfahren von einer abliefernden Stelle an das Landesarchiv [...] überführen“¹⁰⁴. Ein zentrales Element der Applikation bildet, von anderen Funktionen an dieser Stelle abgesehen, die Möglichkeit der Übernahme von Datenbankinhalten aus Fachverfahren über eine JDBC-Verbindung (Java Database Connectivity). Namentlich kann IngestList durch seine Datenbank-Anbindung Datenbankinhalte aus MySQL-, MS Access- und Oracle-Datenbanken migrieren.¹⁰⁵ Künftig soll zudem die Anbindung weiterer Datenbanksysteme ermöglicht werden.¹⁰⁶

In Bezug auf die Migration von Datenbankinhalten aus Fachanwendungen beruht IngestList grundsätzlich auf der Methode der Extraktion von Daten in Form statischer elektronischer Dokumente (vgl. Kapitel 3.2). So kann die Anwendung alle oder auch nur spezifische Tabellen oder Sichten aus einer relationalen Datenbank herauslesen und in CSV-Dateien migrieren. Zu diesem Zweck wird in einem ersten Schritt über die Anwendung eine Verbindung zum Datenbankserver hergestellt (Schnittstelle) und in der Folge werden die Tabellen und Sichten einer Datenbank mit ihren Attributen auf die Benutzeroberfläche der Anwendung ausgelesen (vgl. Anhang 2). In einem weiteren Schritt können mit IngestList Tabellen, Sichten und Attribute selektiert bzw. für die Migration ausgewählt werden. Dabei kann die Auswahl durch die Festlegung von Bedingungen in Bezug auf bestimmte Attributwerte verfeinert werden (z.B. nur Attributwerte, die „kleiner“ sind als X, migrieren). In einem letzten Schritt werden die ausgewählten Tabellen, Sichten und Attribute mittels eines „Exportbefehls“ aus

¹⁰³ „IngestList6“ (=6. Version der Applikation).

¹⁰⁴ Vgl. „IngestList6“ (=6. Version der Applikation), Hilfefunktion.

¹⁰⁵ Landesarchiv Baden-Württemberg, kostenlose Software zur digitalen Archivierung, o.S.; Keitel/Lang, Ingest von Fachverfahren, S. 39f; Keitel/Lang, DIMAG und IngestList, S. 53, 62; Naumann, Übernahme von Daten aus Fachanwendungen, S. 33.

¹⁰⁶ Keitel/Lang, DIMAG und IngestList, S. 62.

dem Datenbankserver migriert.¹⁰⁷ Da hierbei die Datenbankstruktur (z.B. Schlüsselbeziehungen) nicht erhalten wird, geht die ursprüngliche Datenbankfunktionalität verloren.¹⁰⁸

Nach Beendigung des Exports werden die Datenbankinhalte in Form von CSV-Dateien in einem manuell erstellten Exportverzeichnis gespeichert, wobei für jede Sicht oder Tabelle eine separate Datei entsteht, in der die (selektierten) Attribute und Attributwerte enthalten sind. Im Weiteren erzeugt IngestList für jede migrierte Sicht oder Tabelle ergänzende CSV-Dateien. Diese enthalten u.a. für jede Sicht oder Tabelle Attributsbeschreibungen (Data Typ, Size, Field Name) und Angaben über die Anzahl migrierter Zeilen (Tupel) und Spalten (Attribute) sowie die für den Export verwendeten Export-SQL-Statements.¹⁰⁹

Um die Integrität und Authentizität der aus ihrer „ursprünglichen Präsentationsumgebung [...] entnommen[en]“¹¹⁰ Datenbankinhalte zu gewährleisten, erhebt IngestList während der Migration automatisiert technische Metadaten und berechnet für jede Sicht oder Tabelle eine Prüfsumme (md5-Algorithmus). Im Weiteren bietet IngestList die Möglichkeit, deskriptive Metadaten zur Datenbank (Fachanwendung) sowie zu den migrierten Datenbankinhalten aufzunehmen (vgl. Anhang 2). Um zusätzlich auch den Exportprozess und die Datenübernahme zu dokumentieren, führt die Applikation über alle Migrationsprozesse und durchgeführten Kontrollen ein Protokoll. Sämtliche Metadaten, namentlich die technischen und deskriptiven Metadaten, Protokolle und Prüfsummen, werden in einer XML-Datei festgehalten, die ihrerseits mit einem Hashwert versehen und im Exportverzeichnis abgelegt wird.¹¹¹

Konkret verlangt IngestList vor dem Datenexport – was die Beschreibung der Datenbankinhalte anbelangt –, dass sowohl die Provenienz als auch der Inhalt der zu migrierenden Daten in einem Textfenster beschrieben werden. Nach dem Datenexport können optional weitere deskriptive Anmerkungen, Auffälligkeiten oder Unzulänglichkeiten in Bezug auf die Exportdaten im Allgemeinen oder im Hinblick auf ausgewählte Sichten oder Tabellen manuell aufgenommen werden. Als technische Metadaten werden die bereits als CSV-Dateien ausgegebenen Zusatzinformationen (Sichten-/Tabellennamen, Zeilen- und Spaltenzahlen, SQL-Statements und Attributbeschreibungen) sowie Angaben zur Datenbank, Migrationssoftware und zum Exportformat (CSV) in die XML-Datei aufgenommen. Für die Erhebung der Metadaten für das Exportformat sind bei der Software IngestList die Anwendungen „JHOVE“

¹⁰⁷ Ebd.; Naumann, Übernahme von Daten aus Fachanwendungen, S. 33; Hollmann, Archivierung genuin elektronischer Unterlagen, S. 4. Vgl. Software „IngestList6“, Allgemein (9.4.2014).

¹⁰⁸ Hollmann, Archivierung genuin elektronischer Unterlagen, S. 4.

¹⁰⁹ „IngestList6“ (=6. Version der Applikation), Hilfefunktion; Keitel/Lang, Ingest von Fachverfahren, S. 40f; Keitel/Lang, DIMAG und IngestList, S. 62.

¹¹⁰ Keitel/Lang, Ingest von Fachverfahren, S. 38.

¹¹¹ Ebd., S. 40f.

(JSTOR/ Harvard Object Validation Environment) der Harvard University Library (Formatvalidierung) und „DROID“ (Digital Record Object Identification) bzw. die Dateiformat-Datenbank „PRONOM“ der National Archives/GB (Formatidentifikation) implementiert. Die Verwendung von „JHOVE“ und „DROID/PRONOM“ dient aber primär der kontrollierten Übernahme unterschiedlicher digitaler Daten. Kommt IngestList für die Extraktion von Datenbankinhalten zur Anwendung, beschränkt sich diese Metadatenpezifikation auf die CSV-Exportdatei.¹¹²

4.2.2 SIARD (Schweizerisches Bundesarchiv)

Im Zuge des ARELDA-Projekts (Archivierung elektronischer Daten) entwickelte das Schweizerische Bundesarchiv (BAR) einerseits das offene Archivierungsformat „SIARD“ (Software-Independent Archiving of Relational Databases) für relationale Datenbanken sowie andererseits die Software „SIARD Suite“, die relationale Datenbanken in das SIARD-Format migrieren kann.¹¹³ Mit der Entwicklung des SIARD-Formats und der dazugehörigen Applikation verfolgt das BAR das Ziel, nicht nur die Inhalte, sondern auch die Struktur einer Datenbank in ein archivfähiges Format zu migrieren, damit eine Datenbank und ihre Funktionalität dauerhaft sowie unabhängig von einer proprietären Software erhalten werden können. Zu diesem Zweck analysiert die Software SIARD alle Datenbankelemente, migriert in der Folge die Datenbankinhalte und -struktur in reine Textdateien und speichert diese gemäss der SIARD-Formatspezifikation ab. Das SIARD-Format basiert auf den offenen ISO-Standards XML, SQL:1999¹¹⁴ und Unicode-Text. Durch das SIARD-Verfahren können Datenbankinhalte aus MySQL-, MS Access-, Oracle- und MS SQL Server-Datenbanken migriert werden.¹¹⁵

Das SIARD-Format sieht vor, dass Datenbankinhalte in einer Sammlung unterschiedlicher XML-Dateien gespeichert werden. Da aber die Formatspezifikation eine Datenbank wie „eine einzige zu archivierende Unterlage behandelt“¹¹⁶, werden alle XML-Dateien bei der Speicherung in einem unkomprimierten ZIP-Archiv vereint (Dateierweiterung „.siard“). Vor diesem Hintergrund enthält das ZIP-Archiv einerseits ein „content-Verzeichnis“ für die Primärdaten (Tabelleninhalte) und andererseits ein „header-Verzeichnis“ mit den Metadaten.¹¹⁷

Konkret befindet sich im Metadatenverzeichnis („header“) eine XML-Datei („metadata.xml“), die entsprechend der Spezifikation einer relationalen Datenbank hierarchisch strukturiert ist.

¹¹² Ebd., S. 39-45; Neubauer, Extraktion, Kap. 7, S. 16; Funk, Das JSTOR, Kap. 13, S. 21; „IngestList6“ (=6. Version der Applikation), Allgemein.

¹¹³ SIARD Suite 1.0 (SiardEdit 1.50) / SIARD-Format 1.0.

¹¹⁴ Für die Erhaltung der Datenstruktur verwendet SIARD ein Standard-SQL als „Archivformat“. Bei der Migration werden die unterschiedlichen SQL-Dialekte entsprechend standardisiert. Vgl. Lindley, Database Preservation, S. 33, 35f.

¹¹⁵ BAR, SIARD. Formatbeschreibung, S. 1-3; BAR, Datenbanken für die Zukunft sichern (=Factsheet), S. 2f; Heuscher/Keller-Marxer, XML – ein strategisches Instrument für Archive?, S. 17; Kortyla, Übernahme aus Ratsinformationssystemen, S. 73f.

¹¹⁶ Ebd., S. 5.

¹¹⁷ Ebd., S. 5-8; Lindley, Database Preservation, S. 36.

Sie enthält Metadaten über die Datenbankstruktur und den Speicherort der Primärdaten sowie deskriptive und technische Metadaten. Auf der obersten Ebene werden allgemeine Datenbankangaben festgehalten (z.B. Angaben über das Archiv, den Inhalt oder den Datenbankbesitzer). Im Weiteren folgen Informationen zu den Ebenen „Schema“, „Tabelle“ und „Spalte“. Hierbei werden u.a. Namen und Beschreibungen festgehalten sowie Listen zu den Tabellen (Schema-Ebene), Attributen und Primär- und Fremdschlüsseln (Tabellen-Ebene) aufgeführt. Auf der untersten Ebene enthält die Metadatendatei u.a. spezifische Informationen zu den Primär- und Fremdschlüsseln und zu vorhandenen Sichten.¹¹⁸

Die Primärdatendatei („content“) enthält Unterverzeichnisse. Für jedes Datenbankschema besteht ein Ordner (z.B. schema1), der wiederum für alle vorhandenen Tabellen in einem Schema einen Unterordner enthält (*table1*, etc.). Alle Daten einer Tabelle (Primärdaten) werden separat auf dieser letzten Ebene in einer XML-Datei (z.B. *table1.xml*) gespeichert. Im Weiteren ist auf dieser Ebene für jede Tabelle eine XML-Schemadefinition (z.B. *table1.xsd*) aufgeführt, um auf das Speicherformat der Primärdaten zu verweisen. Sollten in einer Datenbank zusätzlich so genannte BLOB- und CLOB-Daten (vgl. Fussnote 47) enthalten sein, werden diese in Form binärer Dateien, separiert von den XML-Dateien, in das Primärdatenverzeichnis („content“) aufgenommen und in den XML-Dateien referenziert.¹¹⁹

Die Software SIARD Suite, die Datenbanken in das SIARD-Format konvertiert, besteht aus den Anwendungen „SiardFromDb“, „SiardEdit“ und „SiardToDb“. Bei der Software SiardFromDb handelt es sich um eine Befehlszeilenanwendung, die über die graphische Benutzeroberfläche der Anwendung SiardEdit gestartet werden kann. Mit SiardToDb wird eine Datenbank in das Siard-Format konvertiert. Der Umfang der Datenbankarchivierung wird dabei durch die Datenbank-User-ID (spezifizierte „Nutzerrechte“), mit der die Datenbankverbindung hergestellt wird, bestimmt. So werden nur diejenigen Datenbankobjekte migriert, die in einer Datenbank für eine User-ID freigegeben sind. Im Anschluss an den Migrationsprozess können die konvertierten Datenbankinhalte nicht weiter selektiert werden.¹²⁰

Nach der Migration wird die Datenbank auf der Benutzeroberfläche von SiardEdit in Gestalt einer Verzeichnisstruktur angezeigt. SiardEdit bietet die Möglichkeit, die während der Migration erstellte XML-Metadatendatei zu ergänzen. Konkret können die unterschiedlichen Ebenen einer relationalen Datenbank (z.B. Schema, Tabellen), entsprechend der im SIARD-Format definierten Metadatenstrukturierung, über die Verzeichnisstruktur angewählt und

¹¹⁸ Ebd., S. 8-15; Bernstein, Database Preservation, S. 2.

¹¹⁹ Ebd., S. 15f; Bernstein, Database Preservation, S. 2; Fitzgerald, data archiving tools, S. 270.

¹²⁰ Hartwig, SIARD Suite. Handbuch, S. 24-29; BAR, Datenbanken für die Zukunft sichern (=Factsheet), S. 3; Lindley, Andrew, Database Preservation, S. 33, 37.

über Textfenster dokumentiert werden (vgl. Anhang 3). Wurden bei der Migration u.a. auch Views oder Routinen übernommen, können auch diese mit deskriptiven Metadaten versehen werden. Demgegenüber besteht keine Möglichkeit, die Primärdaten zu verändern. Im Weiteren kann mit SiardEdit in der XML-Metadatendatei nach Primärdaten gesucht werden.¹²¹ Für die Durchführung komplexer Abfragen muss die SIARD-Datei mit der Befehlszeilenanwendung SiardToDb, die ebenfalls über SiardEdit ausgeführt wird, wieder in ein Datenbankformat geladen werden. Bei der Datenbankreproduktion werden aber nur die Tabellen mit ihren Beziehungen wieder hergestellt. Übrige Datenbankelemente, die bei der Migration erfasst wurden (z.B. Sichten oder Prozeduren), können nicht reproduziert werden.¹²²

Was die Gewährleistung der Integrität anbelangt, berechnet SiardFromDb während der Migration für die Primärdaten eine Prüfsumme (md5-Algorithmus) und speichert diese in der XML-Metadatendatei auf der Datenbankebene ab. In der Folge kann in der Anwendung SiardEdit die Datenintegrität kontrolliert werden. Die Ergebnisse der Integritätsprüfungen werden in den Metadaten aber nicht protokolliert.¹²³ Im Weiteren ist an dieser Stelle hinsichtlich der Authentizitätswahrung auf die deskriptiven Metadaten zu verweisen, welche die automatisch generierten technischen und strukturellen Metadaten ergänzen. Ferner erzeugt SiardFromDb ein Migrationsprotokoll, das als separate Datei gespeichert und allenfalls in eine zusätzliche Dokumentation zu den Archivdaten aufgenommen werden kann. Für allfällige Dokumentationsunterlagen ist in der SIARD-Formatspezifikation im Rahmen der Beschreibung der Dateistruktur eines entsprechenden Informationspaketes (nach dem OAIS-Referenzmodell) ein separates Verzeichnis definiert.¹²⁴

4.2.3 CHRONOS (Softwarehersteller CSP GmbH)

Die Anwendung „CHRONOS“¹²⁵ ist ein kommerzielles Produkt zur Archivierung relationaler Datenbanken, das auf der Basis eines gemeinsamen Forschungs- und Entwicklungsprojekts des Softwarehersteller CSP GmbH und der Fachhochschule für angewandte Wissenschaften in Landshut entwickelt wurde. Mit der Entwicklung der Software CHRONOS wurde das Ziel verfolgt, strukturierte Daten aus Datenbanken zu migrieren, so dass diese systemneutral archiviert und bei Bedarf, unter Beibehaltung der ursprünglichen Datenbankfunktionalität, wieder in eine Datenbank importiert werden können. Datenbankinhalte sollten dabei nicht nur aus „stillzulegenden“ Datenbanken bzw. Fachanwendungen übernommen werden. Es soll auch die Möglichkeit bestehen, Daten aus laufenden Systemen, deren Produktivumgebungen einem stetigem Wandel unterliegen, inkrementell zu migrieren und einer Nutzerschaft

¹²¹ Ebd., S. 10-23; BAR, Datenbanken für die Zukunft sichern (=Factsheet), S. 3.

¹²² Ebd., S. 30-34; BAR, Datenbanken für die Zukunft sichern (=Factsheet), S. 3; Lindley, Database Preservation, S. 32, 34.

¹²³ Ebd., S. 21. Vgl. BAR, SIARD. Formatbeschreibung, S. 9.

¹²⁴ Ebd., S. 18, 36; BAR, SIARD. Formatbeschreibung, S. 4; Fitzgerald, data archiving tools, S. 270.

¹²⁵ CHRONOS (=Version 4.5.1), Softwarehersteller CSP GmbH (Deutschland) / ikeep AG (Schweiz).

einheitlich zugänglich zu machen.¹²⁶ Um Datenbankinhalte systemneutral archivieren zu können, migriert CHRONOS die Datenbankinhalte und -struktur in die reinen Textdateien CSV (ASCII/UTF) und XML.¹²⁷ Mit CHRONOS können Datenbankinhalte aus Oracle-, DB2-, Informix- und MS SQL Server-Datenbanken migriert werden.¹²⁸

Was die Systemarchitektur anbelangt, setzt sich CHRONOS aus sechs javabasierten Server-Komponenten (Softwareanwendungen) und aus zwei Clients zusammen. Namentlich werden über den „Configuration-Server“ neue oder bereits definierte Archivkonfigurationen – spezifische Parameter und Einstellungen bezüglich der Archivdaten – geladen oder gespeichert. Mittels des „Export-Servers“ werden Daten (Primärdaten und Datenbankschema) aus einer Datenbank migriert (JDBC-Verbindung) und dem „Storage-Server“ zur Validierung sowie zur Ablage in einem Speichersystem übergeben. Der „Search-Server“ führt Suchanfragen aus, die von den Benutzer-Clients gesendet werden, und wird dabei durch den „Index-Server“ unterstützt, der Archivdaten indiziert. Mit dem „Import-Server“ können archivierte Datenbankinhalte wieder in eine Datenbank importiert werden. Im Weiteren benötigt CHRONOS ein DBMS (Repository), um Log-Daten und Konfigurationen zu speichern.¹²⁹

Die CHRONOS-Clients ermöglichen einerseits die Durchführung einer Datenbankarchivierung und machen andererseits die Archivdaten einer Nutzerschaft zugänglich. Der javabasierte Client „CHRONOS Admin Suite“ dient primär als Arbeitsinstrument dazu, Archivkonfigurationen zu erstellen (z.B. Bestimmung des Archivumfangs), Datenbankmigrationen zu initialisieren und in den Archivdaten SQL- oder Volltextabfragen durchzuführen. Zudem bietet „CHRONOS Admin Suite“ die Möglichkeit, zusätzliche Verfahrensdokumentationen in die Archivdatei aufzunehmen. Der zweite CHRONOS-Client, die Webapplikation „CHRONOS Archive Explorer“, dient dazu, den Endnutzern die Archivdaten zugänglich zu machen. So kann ein Nutzer mit dem Explorer über benutzerdefinierte Archiv-Views bzw. Such- und Report-Vorlagen (JasperReports) sowie Archiv-Deskriptoren (inkl. Web Designer) auf Archivdaten (Informationsobjekte) zugreifen und das allenfalls vorhandene Dokumentationsmaterial einsehen. Dadurch wird „über den technischen Archiven eine geschäftsorientierte Sicht“¹³⁰

¹²⁶ Brandl/Keller-Marxer, Long-term Archiving, S. 1f; Täuber, Herstellererklärung, S. 5; Lindley, Database Preservation, S. 31.

¹²⁷ Im Gegensatz zu SIARD werden die SQL-Dialekte der verschiedenen DBMS erhalten bzw. nicht standardisiert. Bei einer allfälligen Reproduktion einer Datenbank, werden die Archivdaten in den SQL-Dialekt des Ziel-DBMS transformiert. Vgl. Keller-Marxer, CHRONOS (=Whitepaper), S. 5.

¹²⁸ Täuber, Herstellererklärung, S. 12; Brandl/Keller-Marxer, Long-term Archiving, S. 2f; Fitzgerald, data archiving tools, S. 270.

¹²⁹ Keller-Marxer, CHRONOS (=Whitepaper), S. 25-27.

¹³⁰ Ebd., S. 19.

auf die Archivdaten ermöglicht. Ferner können über den „CHRONOS JDBC-Treiber“ auch Drittsysteme mittels SQL-Abfragen auf Archivdaten zugreifen.¹³¹

Aufgrund der CHRONOS-Systemarchitektur können verschiedene „geografische[...] Skalierung[en]“¹³² realisiert werden. So können die CHRONOS-Instanzen (Server, Client, DBMS) zentralisiert an einem Ort installiert werden (zentralisierte Archivierung). Ausgehend von dieser Installation kann CHRONOS über Netzwerke auf Datenbanken an unterschiedlichen Standorten zugreifen, um diese zu archivieren. Die CHRONOS-Instanzen können aber auch aufgeteilt werden (verteilte Archivierung). Dazu werden einzelne Instanzen an unterschiedlichen Standorten installiert, die vor Ort z.B. eine Archivierung vornehmen und in der Folge die Archivdaten z.B. über eine FTP-Verbindung (File Transfer Protocol) oder auf mobilen Datenträgern an den Hauptstandort übertragen. Über den Client „CHRONOS Archive Explorer“ können Nutzer standortunabhängig auf Archivdaten zugreifen.¹³³

Die Anwendung CHRONOS kann „stillzulegende“ Datenbanken/Fachanwendungen archivieren oder Datenbankinhalte inkrementell aus laufenden Systemen in ein Archiv exportieren. Während bei einer Datenbankstillegung die Datenbankinhalte in einem Archivierungslauf in eine Archivdatei migriert werden, erzeugt CHRONOS bei der inkrementellen Archivierung periodisch so genannte Archivpakete, welche die Datenbankinhalte sowie das Datenbankschema (Metadaten) enthalten. Dabei vergleicht CHRONOS vor jedem Archivierungsdurchlauf das jüngste Datenbankschema mit den Metadaten der bereits archivierten Daten, hält Schemaänderungen fest (z.B. Tabellenlöschungen) und speichert das aktuelle Datenbankschema als „neue Archiv-Revision“¹³⁴ ab. Bei einer Suchabfrage berücksichtigt CHRONOS die Änderungen, damit ein Nutzer virtuell immer über das aktuellste Datenbankschema auf die Archivdaten zugreifen kann. Dabei werden die Daten älterer Archivpakete nicht verändert und behalten ihre ursprüngliche Struktur. Werden die Archivdaten aber wieder in ein DBS importiert, verwendet CHRONOS das jüngste Datenbankschema und passt die älteren Daten entsprechend an. Wird eine Datenbank „stillgelegt“, aus der zuvor inkrementell archiviert wurde, kann CHRONOS, sollten sich Archivierungsziele verändert haben, sich selbst als Datenquelle nutzen und aus mehreren Archivpaketen eine abschließende Archivdatei erzeugen. Ferner unterstützt CHRONOS einen inkrementellen Import der Archivdaten in ein Aktivsystem. Um hierbei eine erhöhte Flexibilität zu erreichen (z.B. kann CHRONOS Resultate einer Datenbanksuche gezielt in ein System importieren), werden beim Import generell keine Schlüsselbeziehungen reproduziert. Sowohl bei der Datenbankstillegung als auch bei der

¹³¹ Ebd., S. 6, 19; Täuber, Herstellererklärung, S. 8-10; Brandl/Keller-Marxer, Long-term Archiving, S. 5; Keller-Marxer, Metadaten in CSP CHRONOS, S. 13-17.

¹³² Ebd., S. 28.

¹³³ Ebd., S. 28-31; Brandl/Keller-Marxer, Long-term Archiving, S. 2.

¹³⁴ Ebd., S. 14.

inkrementellen Archivierung können Datenbanken partiell oder vollständig archiviert werden. Die entsprechenden Einstellungen erfolgen mittels einer Archivkonfiguration.¹³⁵

Um einen Archivierungslauf zu spezifizieren, wird in der CHRONOS-Anwendung „CHRONOS Admin Suite“ eine Archivkonfiguration erstellt. Das zentrale Element einer Konfiguration ist der so genannte Strukturbaum. Wird eine Datenbank vollständig archiviert, können alle Tabellen einer Datenbank flach auf einem Niveau des Strukturbaums angezeigt und migriert werden (vgl. Anhang 4). Im Hinblick auf eine partielle Datenbankarchivierung können an dieser Stelle beliebig viele Tabellen aus dem Strukturbau gelöscht werden. Um aber gezielt vollständige Informationsobjekte zu migrieren, wird eine Tabelle oder Sicht (View) einer zu archivierenden Datenbank, in der Regel diejenige Tabelle oder Sicht, welche die wesentlichen Daten eines Informationsobjekts enthält, als „Mastertabelle“ definiert. Ausgehend von einer Mastertabelle baut CHRONOS in der Folge automatisiert einen hierarchischen Strukturbaum auf, wobei sämtliche Tabellen systematisch aneinander gereiht werden, die über Schlüsselbeziehungen in Verbindung stehen (vgl. Anhang 4). Damit CHRONOS automatisiert vorgehen kann, ist es wichtig, dass die Schlüsselbeziehungen in der Quelldatenbank enthalten sind. Ansonsten können die Schlüsselbeziehungen, aber auch deskriptive Anmerkungen über das Konfigurationspanel manuell beschrieben werden. Diese Zusatzinformationen werden getrennt von den Primärdaten in den Metadaten gespeichert. Im Weiteren kann die Mastertabelle mit einem SQL-Statement versehen werden, um spezifische Datenbankinhalte zu migrieren. Dabei besteht bei der inkrementellen Archivierung die Möglichkeit, von weiteren Konfigurationseinstellungen abgesehen, Löschungen zu definieren. So können Datenbankinhalte nach ihrer Archivierung automatisiert aus der Quelldatenbank gelöscht werden. Ist die Archivkonfiguration für eine partielle Archivierung abgeschlossen, werden die definierten Datenbankinhalte entsprechend aus der Quelldatenbank migriert.¹³⁶

Die durch CHRONOS exportierten Datenbankinhalte werden in ein „CHRONOS-Archiv“ gespeichert, das beliebig viele Archivdateien (bzw. Pakete) enthalten kann (z.B. infolge der inkrementellen Archivierung). In einer Archivdatei befinden sich eine XML-Metadatendatei (*db_metaArchiv_Y_M_D.xml*), eine XML-Schemadefinition (*db_metainfo.xsd*) und CSV-Zip-Archive mit den Primärdaten. Dabei besteht für jede Tabelle ein separates CSV-Zip-Archiv. Ferner kann eine Archivdatei – sofern in der Quelldatenbank vorhanden – BLOB- und CLOB-Elemente (separate Text- oder Binärdateien) sowie zusätzliches Dokumentationsmaterial enthalten. Die XML-Metadatendatei enthält primär technische Metadaten: z.B. Angaben zur

¹³⁵ Brandl/Keller-Marxer, Long-term Archiving, S. 3-5; Keller-Marxer, CHRONOS (=Whitepaper), S. 13-19; CSP GmbH, Benutzerhandbuch, S. 131-137; Korrespondenz zwischen Donato Raselli und Peter Keller, 5.6.2014; Lindley, Database Preservation, S. 31, 34. Vgl. Täuber, Herstellererklärung, S. 22f.

¹³⁶ Ebd., S. 3f; Keller-Marxer, CHRONOS (=Whitepaper), S. 35-52.

Quelldatenbank, zu Tabellen, Attributen und Beziehungen, zu Prozeduren sowie Angaben zu Views. Unabhängig vom Archivpaket ist ein zusätzlicher Metadatenbestand im CHRONOS-Repository gespeichert, der als separate XML-Metadaten-datei exportiert und in Ergänzung zu den Archivpaketen archiviert werden kann. Hierzu zählen insbesondere umfangreiche Erhaltungsmetadaten und Paketbeschreibungen (u.a. strukturelle und deskriptive Metadaten zu Attributen und Tabellen, Angaben zu optional definierten Archiv-Views und -Reports, Zugriffsrechten oder zur Datenbankreproduktion). Grundsätzlich enthalten die „Repository-Metadaten“ die im „CHRONOS Admin Suite“ festgelegte Archivkonfiguration.¹³⁷

Für die Gewährleistung der Integrität der Archivdaten berechnet der „Export-Server“ für jede Tabellenzeile (Datensatz) eine Prüfsumme, die als letzter Wert einer Tabellenzeile in den CSV-Zip-Archiven gespeichert wird. Zusätzlich erzeugt CHRONOS eine Prüfsumme über eine ganze Archivdatei. Die Prüfsummen werden sodann durch den „Storage-Server“ vor der Ablage der Daten in ein Speichersystem und auch später kontinuierlich validiert.¹³⁸ Im Weiteren können die Archivdaten über CHRONOS in externe spezialisierte Speichersysteme überführt mit einer Verschlüsselung (Advanced Encryption Standard) oder digitalen Signaturen (inkl. Zeitstempel) versehen werden.¹³⁹ Was die Wahrung der Authentizität anbelangt, erhebt CHRONOS umfangreiche strukturelle und technische Metadaten, lässt die Erhebung deskriptiver Metadaten zu und führt in seinem Repository einen Audit-Log, der die Aktivitäten der Archivierung und Benutzung sowie des Systembetriebs protokolliert. Der Audit-Log kann als CSV-Datei aus dem Repository exportiert und zusätzlich archiviert werden.¹⁴⁰

4.3 Evaluation der Archivierungsverfahren in Bezug auf die Fallbeispiele

Die Untersuchung der Fallbeispiele und die Darstellung der Archivierungsverfahren zeigen auf, dass die Fachanwendungen infolge ihrer Konzeption unterschiedliche Herausforderungen an die Archivierung stellen (z.B. im Datenbankschema hinterlegte oder applikatorisch definierte Schlüsselbeziehungen) und zur Umsetzung einer Datenbankarchivierung grundsätzlich verschiedene Verfahren zur Verfügung stehen. In der Folge wird dargestellt, welche Möglichkeiten diese Verfahren in Bezug auf die als archivwürdig identifizierten Datenbankinhalte der Fallbeispiele bieten. Dabei wird evaluiert, inwiefern im erzeugten Archivgut Prozesse bzw. Prozessinstanzen nachvollzogen werden können und ob die Archivierungsverfahren die Integrität und Authentizität der Archivdaten zu gewährleisten vermögen. Bei dieser Evaluation wird die partielle und die vollständige Datenbankarchivierung berücksichtigt. Denn

¹³⁷ Keller-Marxer, Metadaten in CSP CHRONOS, S. 2-15; Keller-Marxer, CHRONOS (=Whitepaper), S. 43, 52-56; Lindley, Database Preservation, S. 32, 36; Fitzgerald, data archiving tools, S. 270; Brandl/Keller-Marxer, Long-term Archiving, S. 6.

¹³⁸ Brandl/Keller-Marxer, Long-term Archiving, S. 4; Keller-Marxer, Metadaten in CSP CHRONOS, S. 1; Lindley, Database Preservation, S. 36.

¹³⁹ Keller-Marxer, CHRONOS (=Whitepaper), S. 21f; Fitzgerald, data archiving tools, S. 270.

¹⁴⁰ Ebd., S. 21; Keller-Marxer, Metadaten in CSP CHRONOS, 18.

eine vollständige Datenbankarchivierung verhindert, dass Informationen unbeabsichtigt – entweder bei der Bewertung oder bei der Datenmigration – verloren gehen.¹⁴¹

A) IngestList

Bei der Archivierung der Fallbeispiele durch die Applikation „IngestList“ (vgl. Kapitel 4.2.1) besteht grundsätzlich die Schwierigkeit, dass die Applikation die Datenbank MS SQL Server der Fachanwendung „ESCADA²“ nicht unterstützt. Um dieses Fallbeispiel mit IngestList archivieren zu können, muss die Datenbank erst in ein kompatibles Datenbanksystem migriert werden. Hierzu stellt z.B. „Oracle“ das Datenbank-Verwaltungswerkzeug „Oracle SQL Developer“ zur Verfügung, wodurch die Datenbank von „ESCADA²“ theoretisch auf eine mit IngestList kompatible Oracle-Datenbank migriert werden kann.¹⁴² Zwischen der MySQL-Datenbank der Fachanwendung „socialWEB“ und IngestList ist Kompatibilität gegeben. Weitere Voraussetzung ist, dass „IngestList“ über eine Schnittstelle (JDBC-Verbindung) auf die Datenbanken/Fachanwendungen zugreifen kann.

Um die Datenbankinhalte beider Fallbeispiele vollständig zu archivieren, muss IngestList mit einer Schnittstelle auf die Daten zugreifen, die über umfassende Leserechte auf alle Datenbanktabellen verfügt (Superuser-Leserechte). Ist eine Schnittstelle eingerichtet – im Falle von „ESCADA²“ über mehrere Datenbanken hinweg –, kann IngestList sämtliche Tabellen der Datenbanken in einzelne CSV-Dateien migrieren. Um den Fokus der Archivierung – im Sinne einer partiellen Datenbankarchivierung – auf diejenigen Tabellen zu richten, die zu den als archivwürdig identifizierten Informationsobjekten der Fallbeispiele gehören, müssen über IngestList manuell zu bestimmende Tabellen (und evtl. Attribute) von der Archivierung ausgeschlossen werden. Eine weitere Möglichkeit besteht darin, der Anwendung IngestList eingeschränkte Leserechte zu erteilen (spezifische „Archivschnittstelle“), damit nur archivwürdige Tabellen aus der Datenbank gelesen werden können. Bei diesen Vorgehensweisen muss bezüglich der Schnittstellenkonfiguration explizit sichergestellt werden, dass die Vollständigkeit der Daten gesichert ist (referentielle Integrität). Das heißt, dass zu einem archivwürdigen Informationsobjekt alle dazugehörigen Tabellen berücksichtigt werden, so dass die Fremdschlüssel in den betroffenen Tabellen aufgelöst werden können.¹⁴³ Ferner besteht die Möglichkeit, dass ein Archivar in den Quelldatenbanken der Fallbeispiele z.B. archivwürdige Attribute zu spezifischen Sichten synthetisiert und über IngestList migriert. Dadurch werden die Sichten zu eigenständigen Tabellen im CSV-Format materialisiert.¹⁴⁴

¹⁴¹ Vgl. Fitzgerald, data archiving tools, S. 271.

¹⁴² Eine Migrationsanleitung sowie ein Download-Link für den „Oracle SQL Developer“ sind unter folgender URL zu finden: <<http://www.oracle.com/technetwork/database/migration/sqlserver-095136.html>> (31.5.2014).

¹⁴³ Vgl. Kemper/Eickler, Datenbanksysteme, S. 160f.

¹⁴⁴ Vgl. Keller-Marxer, CHRONOS (=Whitepaper), S. 43.

Unabhängig von der Methode ist bei der Datenbankarchivierung mit IngestList zu beachten, dass Schlüsselbeziehungen bei der Migration nicht erhalten werden. Diese müssen in IngestList, sollen sie überliefert werden, über die Bemerkungsfelder zu den einzelnen Tabellen manuell dokumentiert und in die Metadaten aufgenommen werden. Im Weiteren handelt es sich bei den Fallbeispielen um Systeme, die weiterhin in Betrieb sein werden. Hierbei müssen, ausgehend von der gewählten Archivierungsmethode, mit IngestList periodisch Datenbank-Snapshots erzeugt werden, wodurch mehrere und teilweise redundante Archivdateien entstehen. Zusätzlich müssen die konfigurierten „Archivschnittstellen“ immer auf ihre Aktualität hin überprüft werden, da sich die Datenbanken möglicherweise verändern.¹⁴⁵

Werden die durch IngestList archivierten Fallbeispiele bzw. Informationsobjekte (z.B. die „Journaleinträge“ von „socialWEB“) hinsichtlich der Nachvollziehbarkeit der Prozesse bzw. Prozessinstanzen im Archivgut untersucht, ist zu bedenken, dass die Archivdaten bzw. sämtliche Tabellen in separaten CSV-Dateien vorliegen (für die „Journaleinträge“ ca. fünf von ca. 250 Dateien/Tabellen) und nicht mehr über die Benutzeroberfläche einer Fachanwendung synthetisiert werden (externe nutzungsspezifische Sicht). Für Nutzer wird es daher, obwohl CSV-Dateien gut lesbar sind, eine Herausforderung sein, in der Menge und der potentiellen Komplexität der Archivdaten (im Falle von „socialWEB“ ca. 70-80 Dateien/Tabellen) die zu einem Informationsobjekte gehörenden Dateien zu identifizieren und daraus Prozessinstanzen zu rekonstruieren.¹⁴⁶ Wird eine Datenbank vollständig archiviert, gestaltet sich diese Aufgabe noch anspruchsvoller. Sind die Schlüsselbeziehungen auf den Tabellen definiert (z.B. „ESCADA²“) oder in den Metadaten dokumentiert (z.B. „socialWEB“), können Informationsobjekte möglicherweise „syntaktisch ungefähr erahn[...][t]“¹⁴⁷ werden. Die „in [der] Applikationssoftware verwendete Semantik lässt sich daraus aber nicht erschließen“¹⁴⁸. Hat der Datenbankhersteller die Tabellen verständlich bezeichnet und werden durch ein Archiv die Dokumentationsmöglichkeiten, die IngestList bietet (z.B. zusätzliche Bemerkungen zu den einzelnen Tabellen) und verlangt (z.B. Inhalts- und Provenienzangaben zu den Archivdaten), ausreichend wahrgenommen, können die Identifikation von Informationsobjekten und der Rückschluss auf Prozesse bzw. Prozessinstanzen gefördert werden.¹⁴⁹

Was die Sicherung der Integrität und Authentizität angeht, berechnet IngestList für alle Archivdateien Prüfsummen und protokolliert den Migrationsprozess. Nebst der manuellen Erhebung deskriptiver Metadaten, erhebt die Anwendung zudem automatisiert umfassende strukturelle und technische Metadaten, die mit den Archivdaten „ein Geflecht gegenseitiger

¹⁴⁵ Vgl. Naumann, Übernahme von Daten aus Fachanwendungen, S. 34.

¹⁴⁶ Heuscher/Keller-Marxer, Archival Access to Complex Relational Data, S. 4.

¹⁴⁷ Keller-Marxer, Metadaten in CSP CHRONOS, S. 12.

¹⁴⁸ Ebd.

¹⁴⁹ Vgl. Ribeiro, Database Preservation, S. 1.

Verweise [bilden], die kaum umfassend gefälscht werden können“¹⁵⁰. Aufgrund der XML-Metdatendatei können die archivierten Primärdaten über IngestList beliebig oft validiert werden, während zugleich das zur Metdatendatei gehörende Protokoll aktualisiert wird. Vor diesem Hintergrund ist festzuhalten, dass durch IngestList die Integrität und Authentizität der Daten grundsätzlich gewährleistet werden kann.¹⁵¹

B) SIARD

Werden die Fallbeispiele über die Applikation „SIARD“ (vgl. Kapitel 4.2.2) archiviert, ist die Kompatibilität mit den Quelldatenbanken gegeben. Wie für die Applikation IngestList gilt auch für SIARD, dass die Anwendung über eine JDBC-Verbindung auf die Datenbanken der Fachanwendungen „socialWEB“ und „ESCADA²“ zugreifen muss, damit Datenbankinhalte migriert werden können. Was die Möglichkeiten bezüglich der Archivierungsmethode von SIARD anbelangt, sind diese mit denjenigen von IngestList vergleichbar. Der Unterschied besteht aber darin, dass SIARD, vor dem Hintergrund der Archivierungspraxis des Schweizerischen Bundesarchivs, prinzipiell für die vollständige Datenbankarchivierung entwickelt wurde. Dementsprechend bietet die Applikation SIARD keine Möglichkeit, die nicht archivwürdigen Datenbankinhalte der Fallbeispiele von der Archivierung auszuschließen. Verfügt SIARD über einen Superuser-Zugang auf die Quelldatenbanken, werden die Datenbankinhalte vollumfänglich in das Zip-Archiv bzw. die Content-XML-Dateien migriert.

Sollen bei einer partiellen Datenbankarchivierung dennoch nur ausgewählte Tabellen bzw. archivwürdige Informationsobjekte der Fallbeispiele archiviert werden, geschieht dies, wie dies auch bei IngestList eine Möglichkeit darstellt, indirekt über die Einschränkung der Lese- bzw. Zugriffsrechte (User-ID). Für den Fall, dass zusätzlich bestimmte Attribute von der Archivierung ausgeschlossen (z.B. allfällige Telefonnummern der Schüler im Informationsobjekt „schülerbezogene Personendaten“ der Fachanwendung „ESCADA²“) oder nur synthetisierte Sichten migriert werden sollen, müssen entsprechende Einschränkungen in der Quelldatenbank mittels eines SQL-Statements definiert und die Ergebnisse „als Tabellen [...] realisiert“¹⁵² werden. Diese „künstlich“ erstellten Tabellen können anschließend über SIARD migriert werden.¹⁵³ Bei der Archivierung mit SIARD muss die Quelldatenbank ebenfalls gründlich analysiert werden, damit bei der manuellen Ausschließung bestimmter Datenbankinhalte die referentielle Integrität der zu archivierenden Daten nicht gefährdet wird.

¹⁵⁰ Keitel/Lang, Ingest von Fachverfahren, S. 40.

¹⁵¹ Ebd.; Keitel, Ways to deal with complexity, S. 3f; Naumann, Übernahme von Daten aus Fachanwendungen, S. 33.

¹⁵² BAR, SIARD Suite: FAQ (=Frequently Asked Questions), S. 1.

¹⁵³ Ebd.

Die im Schema der Anwendung „ESCADA²“ hinterlegten Schlüsselbeziehungen werden bei der Migration durch SIARD in die Metadaten übernommen und beim Datenimport in einem neuen Datenbanksystem wieder hergestellt; dies trifft aber nicht auf Sichten zu, die z.B. als Tabelle „realisiert“ werden. Die Archivierung der Fachanwendung „socialWEB“, bei der die Schlüsselbeziehungen applikatorisch bestimmt sind, müssen die Verknüpfungen über SiardEdit auf der Tabellen-Ebene dokumentiert und manuell in die Metadaten integriert werden. Bei der Reproduktion der Datenbank werden die Beziehungen nicht automatisch wiederhergestellt. Bezüglich der Archivierung aus aktiven Systemen ist die Vorgehensweise gleich wie bei IngestList: In periodischen Abständen werden Snapshots erzeugt, währenddessen die Schnittstellenaktualität kontinuierlich zu überprüfen ist. Die Archiv-Snapshots können sich, sollten die Quelldatenbanken weiterentwickelt werden, strukturell unterscheiden.

Werden die Archivdaten bzw. die Content-XML-Dateien sowie die Metadaten, die durch SIARD erzeugt werden, im Hinblick auf die Nachvollziehbarkeit von Prozessen bzw. Prozessinstanzen untersucht, ist festzuhalten, dass mit entsprechenden Kenntnissen auch die XML-Dateien von den Nutzern gelesen und nachvollzogen werden können. Wie bei IngestList ist aber auch hier davon auszugehen, dass Informationsobjekte und die dazugehörenden Prozessinstanzen sich nur erahnen lassen. Dennoch besteht grundsätzlich die Möglichkeit, ausgehend von den Schlüsselbeziehungen und den aufgenommenen deskriptiven Metadaten, auf Prozesse bzw. Prozessinstanzen zu schließen. Dabei kann insbesondere der Umstand, dass im Falle von „ESCADA²“ die Archivdaten wieder mit den Schlüsselbeziehungen in ein Datenbanksystem importieren werden können, eine wichtige Unterstützung bieten.

In Bezug auf die Integritäts- und Authentizitätswahrung erstellt SIARD eine Prüfsumme, erhebt strukturelle sowie technische Metadaten und bietet die Möglichkeit, ergänzende deskriptive Metadaten aufzunehmen. Im Gegensatz zu IngestList erzeugt SIARD nur eine Prüfsumme über alle Primärdaten, wodurch sich Integritätsverletzungen nicht auf Tabellenniveau nachverfolgen lassen; zudem werden die Integritätsprüfungen nicht in den Metadaten protokolliert. Im Weiteren ist an dieser Stelle bezüglich der Authentizitätswahrung darauf hinzuweisen, dass allfällige Herstellerinformationen zu den Tabellen, die in einer Datenbank hinterlegt sein können, automatisch in die Textfelder zur Aufnahme deskriptiver Metadaten aufgenommen werden. Dabei besteht die Gefahr, dass solche Primärdaten mit Anmerkungen des Archivs vermengt werden. Archivische Beschreibungen müssen daher explizit als solche gekennzeichnet werden. Ansonsten kann SIARD grundsätzlich die Integrität und infolge der Metadaten die Authentizität der Archivdaten gewährleisten.

C) CHRONOS

Die Archivierung der Fallbeispiele durch „CHRONOS“ (vgl. Kapitel 4.2.3) erfordert, dass die Kompatibilität zwischen CHRONOS und der Anwendung „socialWEB“ hergestellt wird. CHRONOS unterstützt die Archivierung aus MySQL-Datenbanken nicht, weshalb die Datenbank der „socialWEB-Anwendung“ erst in ein kompatibles Datenbanksystem migriert werden muss. Zu diesem Zweck kann erneut, wie im Beispiel von IngestList und „ESCADA²⁴“ erwähnt wurde, das Datenbank-Werkzeug „Oracle SQL Developer“ verwendet werden, das die MySQL-Datenbank von „socialWEB“ theoretisch in eine Oracle-Datenbank migrieren kann. In Bezug auf „ESCADA²⁴“ ist die Kompatibilität mit CHRONOS gegeben.¹⁵⁴

Würden zwischen CHRONOS und den Fallbeispielen eine JDBC-Verbindung mit Superuser-Rechten erstellt, können die Datenbanken ohne Zusatzeinstellungen vollständig migriert werden. Für die partielle Archivierung der Datenbanken bzw. die Migration der in Bezug auf die Fallbeispiele als relevant erachteten Informationsobjekte werden entsprechende Datenbankreduktionen direkt in der Archivkonfiguration von CHRONOS vorgenommen. Zu diesem Zweck wird für jedes zu archivierende Informationsobjekt eine Tabelle lokalisiert, welche die relevanten Informationen enthält und dadurch als Mastertabelle verwendet werden kann.¹⁵⁵ In der Folge können ausgehend von der Mastertabelle manuell oder automatisiert alle Tabellen identifiziert werden, die zur Auflösung der Fremdschlüsselbeziehungen benötigt werden (Child-Parent-Beziehung). Dadurch ergeben sich – im Falle einer automatisierten Durchführung – vollständig referentiell integre Informationsobjekte. Diese Archivdaten können im Weiteren automatisiert durch Tabellen ergänzt werden, die nicht zu einem Informationsobjekt gehören, aber dennoch auf die Mastertabelle referenzieren (Parent-Child-Beziehung).¹⁵⁶ Für das Fallbeispiel „socialWEB“, in welchem die Schlüsselbeziehungen applikatorisch definiert sind, müssen die Verknüpfungen in der CHRONOS-Archivkonfiguration manuell nachgeführt werden. Um vorhandene oder durch Archive erzeugte Sichten zu migrieren, können diese als Mastertabellen festgelegt und über die Migration als eigenständige (CSV-)Tabellen materialisiert werden. Mittels SQL-Statements kann die Datenmigration zudem gezielt auf spezifische Tabelleninhalte fokussiert werden. Da Informationsobjekte gleiche Datenbankinhalte verwenden können, kann es in den Archivdaten, sofern mehrere Informationsobjekte in die selbe Archivdatei migriert werden, zu Redundanzen kommen. So werden z.B. die Informationsobjekte „Protokolle“ und „Journaleinträge“ der Anwendung „socialWEB“ die gleichen Daten verwenden, um die Einträge und Protokolle nach Kindern zu differenzieren.

¹⁵⁴ Eine Migrationsanleitung sowie ein Download-Link für den „Oracle SQL Developer“ sind unter folgender URL zu finden: <<http://www.oracle.com/technetwork/database/migration/mysql-093223.html>> (5.6.2014).

¹⁵⁵ Vgl. Keller-Marxer, CHRONOS (=Whitepaper), S. 35.

¹⁵⁶ Vgl. Ebd., S. 44.

Die Funktionalität von CHRONOS ist besonders darauf ausgerichtet, Datenbankinhalte aus laufenden Systemen zu archivieren. In Bezug auf die Fallbeispiele können dabei durch CHRONOS inkrementell spezifische Inhalte in ein Archiv übernommen werden, während die Anwendung allfällige Schemaänderungen automatisch analysiert und im Hinblick auf die Nutzung berücksichtigt. In diesem Zusammenhang könnten zudem Löschungen konfiguriert werden. Dies würde z.B. im Falle von „ESCADA²“ die Möglichkeit bieten, dass etwa die „schülerbezogenen Personendaten“ nach Ablauf der Aufbewahrungsfristen automatisiert einem digitalen Archiv übertragen und aus der Quelldatenbank gelöscht werden.

Die durch CHRONOS erzeugten Archivdaten (CSV- und XML-Dateien) bieten, was die Nachvollziehbarkeit von Prozessen bzw. Prozessinstanzen anbelangt, ähnliche Voraussetzungen wie die Archivdaten von IngestList oder SIARD. Die Archivdaten sind grundsätzlich gut lesbar und die in die Metadaten aufgenommenen Schlüsselbeziehungen sowie allfällige deskriptive Bemerkungen ermöglichen Rückschlüsse auf Informationsobjekte und dazugehörige Prozessinstanzen. CHRONOS bietet jedoch zusätzliche Möglichkeiten, die eine Nutzerschaft dabei unterstützen können, Prozesse bzw. Prozessinstanzen in den Archivdaten besser nachzuvollziehen. Eine Möglichkeit besteht darin, archiwürdige Informationsobjekte mit verhältnismäßig geringem Aufwand referentiell integer in separate Archivdateien zu migrieren, wodurch diese durch die Nutzer leichter nachvollzogen werden könnten. Im Weiteren stellt CHRONOS einen Client zur Verfügung („Archive Explorer“), der über frei definierbare Such- und Report-Vorlagen sowie Archiv-Deskriptoren die Archivdaten einer Nutzerschaft zugänglich machen kann. Konkret können relevante Inhalte von Informationsobjekten unmittelbar nach der Archivierung wieder synthetisiert werden, was eine Annäherung an die ursprüngliche nutzungsspezifische Sicht auf die Datenbankinhalte ermöglicht. So können bei der Archivierung der Fallbeispiele in Zusammenarbeit mit den Herstellern und Institutionen bestimmte Sichten erhalten oder teilweise direkt übertragen werden (z.B. von „Crystal Reports“ bei der Anwendung „ESCADA²“). Über den „Archive Explorer“ können die Archivnutzer zudem auf Dokumentationsmaterial zugreifen – sofern solches in das Archivgut aufgenommen wird –, um z.B. Zusatzinformationen über die Quelldatenbank zu erhalten. CHRONOS stellt Unterstützungsmöglichkeiten zur Verfügung, welche die Nachvollziehbarkeit von Prozessen bzw. Prozessinstanzen wesentlich fördern können.

Auch CHRONOS erzeugt zur Sicherung der Integrität Prüfsummen. Im Gegensatz zu den anderen Archivierungsanwendungen werden diese aber für jeden einzelnen Datensatz in der Archivdatei erstellt, wodurch Integritätsverletzungen genau lokalisiert werden können. Der Umstand, dass CHRONOS auch eine Prüfsumme für die Archivdatei berechnet, ermöglicht es, auch die Dokumentationsunterlagen, die sich getrennt von den Primärdaten in der Archivdatei befinden, zu sichern. Im Weiteren bietet CHRONOS weitere Optionen, um die Da-

tenintegrität zu erhöhen (Verschlüsselung, digitale Signatur, externes Speichersystem). Die Authentizität der Archivdaten kann durch umfangreiche deskriptive, strukturelle und technische Metadaten sowie einen Audit-Log gesichert werden. Mit CHRONOS besteht die Möglichkeit, die Integrität und Authentizität der Archivdaten im hohen Masse zu sichern.

D) Zwischenfazit

Die hier untersuchten Archivierungsverfahren ermöglichen, dass Datenbanken partiell oder vollständig aus laufenden oder „stillzulegenden“ Fachanwendungen archiviert werden können und dabei die Integrität und Authentizität gewährleistet wird. An dieser Stelle ist aber darauf hinzuweisen, dass ein Archiv durch die Erzeugung und Migration „künstlicher“ Sichten möglicherweise Datenkompositionen schafft, die in dieser Form in der Verwaltungspraxis nicht verwendet werden. Dabei stellt sich bezüglich der Authentizität die Frage, inwiefern diese Methode „Überlieferung bildet oder Historie erschafft“¹⁵⁷. Unabhängig von der gewählten Vorgehensweise migrieren alle Archivierungsanwendungen die Datenbankinhalte (inkl. Metadaten) in lesbare Formate (CSV und XML) und bieten die Möglichkeit, ergänzende deskriptive Metadaten aufzunehmen. Ferner können über die Metadaten auch Schlüsselbeziehungen erhalten werden. Vor diesem Hintergrund können Prozessinstanzen in den Archivdaten theoretisch erahnt werden. Bei größeren und komplexeren Datenbanken muss aber die Möglichkeit in Betracht gezogen werden, dass entsprechende Rückschlüsse nicht mehr möglich sind. Damit die Prozessinstanzen und insbesondere die Geschäftsprozesse im Archivgut nicht nur erahnt, sondern nachvollzogen und verstanden werden können, ist es notwendig, dass die Archivdaten mit weiteren Informationen angereichert werden. Insbesondere im Falle einer partiellen Archivierung könnten Informationen über die Beschaffenheit der Quelldatenbank bzw. die Fachanwendung zudem die Authentizität der Daten zusätzlich erhöhen, da die archivische Bewertung bzw. die Auswahl von Datenbankinhalten immer auch von einer gewissen Willkür geprägt ist.¹⁵⁸

5. Archivierung datenbankgestützter Geschäftsprozesse

5.1 Empfehlungskatalog zur Dokumentation von Prozessvorgängen

Für die Langzeitarchivierung von Datenbankinhalten aus Fachanwendungen ist es unerlässlich, dass nicht nur Integrität und Authentizität der migrierten Datenbankinhalte gewährleistet werden, sondern dass die Daten auch über eine lange Zeitspanne hinweg verstanden werden können. Konkret sollen sich die Prozesse und Prozessinstanzen, die zur Erzeugung der Daten geführt haben, aus dem Archivgut erschließen lassen. Zu diesem Zweck müssen die

¹⁵⁷ Dässler/Schwarz, Fachverfahren, S. 13.

¹⁵⁸ Vgl. KOST, Bewertung von Datenbanken, o.S.; Heuscher/Keller-Marxer, Archival Access to Complex Relational Data, S. 4.

z.B. in Form von CSV- und XML-Dateien vorliegenden Primär- und Metadaten mit Dokumentationsmaterial angereichert werden. In diesem Sinne wurde auch im Abschlussbericht des internationalen ERPANET-Workshops (Electronic Resource Preservation And Access Network) zur Datenbankarchivierung im Jahre 2003 festgehalten:

„The functionality of the system where the data have been created can not be preserved with data conversion and migration. The only way to document how the business processes created data and how the database application supported these processes is to collect system documentation material covering the whole life cycle of the system.“¹⁵⁹

In der Folge soll aufgezeigt werden, mit welchen Zusatzinformationen das Archivgut ergänzt werden kann, damit Nutzer die Prozesse und Prozessinstanzen nachvollziehen können. Dabei werden die unterschiedlichen Ebenen einer Fachanwendung – namentlich die Datenbank- und Applikationsebene – berücksichtigt.¹⁶⁰ Zudem ist der Kontext, in welchem eine Anwendung zum Einsatz kommt, zu beachten. Das Resultat ist ein Katalog (siehe Anhang 5), der die Durchführung von Archivierungsprojekten in der Praxis unterstützt.

Ein zentrales Element, um ein Datenbanksystem einer Fachanwendung zu verstehen, ist das Datenbankmodell, da es die wesentlichen Strukturinformationen enthält (vgl. Kapitel 3.1). Es gibt u.a. Auskunft über die in einem Datenbanksystem verwendeten Tabellen und deren Attribute sowie über die Tabellenverknüpfungen (Primär- und Fremdschlüssel).¹⁶¹ In diesem Zusammenhang sind auch beschreibende Informationen zu den Daten wichtig. Insbesondere können Angaben zu den Datenfeldern einer Tabelle weiterführende Hinweise über die Bedeutung und Nutzung eines Attributes enthalten (Datenbeschreibung).¹⁶² Werden als Attributwerte besondere Codierungen verwendet, müssen entsprechende Codelisten beschafft werden, damit diese Daten aufgelöst und verstanden werden können. Durch Zusatzinformationen über das einer Fachanwendung zugrundeliegende Datenbanksystem kann die Möglichkeit erhöht werden, in den Archivdaten die Prozessinstanzen zu identifizieren.¹⁶³

Im Hinblick auf das Anwendungsprogramm, das über seine Benutzeroberfläche die Datenbankinhalte zu nutzungsspezifischen Sichten synthetisiert, bilden Benutzerhandbücher, Benutzungsreglemente und Schulungsunterlagen zentrale Zusatzinformationen. Insbesondere Handbücher und Schulungsunterlagen können detaillierte Angaben darüber enthalten, welche Daten zu welchem Zweck in einer Fachanwendung verwendet werden. Dadurch wird ein

¹⁵⁹ ERPANET, Long-term Preservation of Databases, S. 11.

¹⁶⁰ VdA, Bewertung elektronischer Fachverfahren, S. 1.

¹⁶¹ BAR, Dokumentation Datenbanken; Fitzgerald, data archiving tools, S. 271; Glöde, Archivierung relationaler Datenbanken, S. 27; Ribeiro, Database Preservation, S. 1; Dässler/Schwarz, Fachverfahren, S. 17.

¹⁶² Ebd.; Heuscher/Keller-Marxer, Archival Access to Complex Relational Data, S. 4; Dässler/Schwarz, Fachverfahren, S. 17; Glöde, Archivierung relationaler Datenbanken, S. 27.

¹⁶³ Zürcher, Bewertung von Datenbanken S. 20; BAR, Dokumentation Datenbanken, o.S.; Dässler/Schwarz, Fachverfahren, S. 17.

Einblick in die Systemanwendung ermöglicht.¹⁶⁴ Zudem können auch allfällige Erzeugnisse einer Fachanwendung (z.B. Reports) gesammelt sowie z.B. ausgewählte Screenshots zentraler Datenerfassungsmasken erzeugt werden. Durch diese Zusätze werden einer Nutzerschaft Informationen zugänglich gemacht, die nicht nur die Identifikation von Prozessinstanzen, sondern möglicherweise auch die Rekonstruktion von Prozessen fördern.¹⁶⁵

Weitere Dokumentationsunterlagen, die sowohl Angaben zum Datenbanksystem als auch zum Anwendungsprogramm enthalten, sind allgemeine Systemspezifikationen (z.B. Broschüren, Homepage-Informationen) und Projektunterlagen zur Genese bzw. Konzeption einer Fachanwendung.¹⁶⁶ Ferner kann sogar daran gedacht werden, mit Anwendern einer Fachanwendung Interviews durchzuführen und diese in schriftlicher Form oder als Tondatei in das Archivgut aufzunehmen.¹⁶⁷ Eine andere Möglichkeit wären Filmaufnahmen, die beispielsweise Anwender bei der täglichen Arbeit mit der Fachanwendung zeigen.¹⁶⁸

Werden der Kontext und die Prozesse identifiziert und analysiert, in welchen Fachanwendungen zum Einsatz kommen, wird sichtbar, dass nicht alle Informationen, die bei Prozessdurchläufen entstehen, in eine Datenbank aufgenommen werden (vgl. Kapitel 4.1). Konkret entstehen analoge Dokumente, die mit den Datenbankinhalten einer Fachanwendung in Verbindung stehen können. Vor diesem Hintergrund sollte untersucht werden, welche archivwürdigen analogen Akten bei der Umsetzung eines Prozesses entstehen, die Datenbankinhalte ergänzen könnten. Diese analogen Unterlagen tragen unter Umständen wesentlich dazu bei, dass sich Prozesse, in welchen Fachanwendungen eine spezifische Rolle spielen, vertieft nachvollziehen lassen. In diesem Zusammenhang könnten auch – sofern vorhanden – Prozessdokumentationen in das Archivgut aufgenommen werden, um die Einbindung einer Fachanwendung in Geschäftsprozesse detailliert aufzuzeigen.

Es muss davon ausgegangen werden, dass in der archivischen Praxis Zusatzinformationen nur seltenen verfügbar sind. So ist dem erwähnten Abschlussbericht des internationalen ERPANET-Workshops zu entnehmen: „Participants noted that there is always a danger of insufficient documentation being available, and therefore a tendency to safeguard whatever documentation exists.“¹⁶⁹ Eine weitere Herausforderung ist der Umstand, dass Dokumentationsunterlagen wahrscheinlich sowohl in digitaler als auch in analoger Form vorgefunden

¹⁶⁴ BAR, Dokumentation Datenbanken, o.S.; VdA, Bewertung elektronischer Fachverfahren, S. 4; Fitzgerald, data archiving tools, S. 271; Naumann, Übernahme von Daten aus Fachanwendungen, S. 32; Heuscher/Keller-Marxer, Archival Access to Complex Relational Data, S. 4.

¹⁶⁵ Ebd.; VdA, Bewertung elektronischer Fachverfahren, S. 4; Fitzgerald, data archiving tools, S. 271.

¹⁶⁶ Ebd.; Naumann, Übernahme von Daten aus Fachanwendungen, S. 32; VdA, Bewertung elektronischer Fachverfahren, S. 4.

¹⁶⁷ Fitzgerald, data archiving tools, S. 271.

¹⁶⁸ Heuscher/Keller-Marxer, Archival Access to Complex Relational Data, S. 4.

¹⁶⁹ ERPANET, Long-term Preservation of Databases, S. 11.

werden. Analoge Zusatzinformationen könnten daher digitalisiert werden, um die Dokumentation einheitlich mit den Datenbankinhalten zu archivieren. In diesem Falle müsste gewährleistet sein, dass die Dokumentation ihrerseits die Anforderungen der digitalen Langzeitarchivierung erfüllt.¹⁷⁰ Unabhängig davon ist es zentral, dass auch eine Verfahrensdokumentation zur Archivierung einer Nutzerschaft zur Verfügung gestellt wird (z.B. dokumentierter Übernahmeprozess, Aufbereitungsbericht zu vollzogenen Entscheidungen).¹⁷¹

5.2 Anwendung des Empfehlungskatalogs auf die Fallbeispiele

Wie in der Evaluation der Archivierungsverfahren in Bezug auf die Fallbeispiele aufgezeigt wurde (Kapitel 4.3), besteht im Hinblick auf die Archivierungsziele die Notwendigkeit, migrierte Datenbankinhalte mit Zusatzinformationen anzureichern (Kapitel 5.1). Dabei ist die Ausgangslage bei den Archivdaten, die durch die Verfahren IngestList, SIARD und CHRONOS erzeugt wurden, weitgehend gleich. So sollten für alle Archivdaten in einem ähnlichen Umfang Zusatzinformationen in das Archivgut aufgenommen werden; zumal Dokumentationsmaterial u.U. generell nur mangelhaft vorhanden ist. Unterschiede, die sich auf die Auswahl zusätzlicher Informationen auswirken können, bestehen darin, dass SIARD und CHRONOS die Archivdaten wieder in eine Datenbank importieren können und CHRONOS aufgrund seiner Clients Vorteile hinsichtlich der Nutzung der Archivdaten bietet. In der Folge soll, ausgehend vom erstellten Empfehlungskatalog (vgl. Anhang 5), in Bezug auf die Fallbeispiele und unter Berücksichtigung der durch die Archivierungsverfahren erzeugten Archivdaten dargestellt werden, welche Zusatzinformationen in das Archivgut aufgenommen werden müssen.

A) Fallbeispiel „socialWEB“ (KJFS)

Was das Fallbeispiel „socialWEB“ (KJFS) angeht, besteht eine Schwierigkeit darin, dass es sich um eine proprietäre Fachanwendung handelt. Dies kann zur Folge haben, dass Informationen (z.B. Datenbankmodell) nicht zur Verfügung stehen, da diese durch den privaten Hersteller infolge geschäftlicher Interessen zurückgehalten werden. In diesem Zusammenhang hat der Hersteller der Anwendung „socialWEB“ aber darauf hingewiesen, dass er im Zuge einer Schutzvereinbarung (vgl. Kapitel 1.3) entsprechende Unterlagen, konkret das Datenbankmodell, für ein Archivierungsprojekt zur Verfügung stellen würde.¹⁷² Dieser Umstand ist im Falle der Anwendung „socialWEB“ relevant, da hier die Schlüsselbeziehungen applikatorisch definiert sind. Anhand des Datenbankmodells kann die Ermittlung der Tabellenverknüpfungen vereinfacht werden. Gemäß Hersteller gibt es zur Datenbank der Fachan-

¹⁷⁰ Ebd.

¹⁷¹ Keller-Marxer, CHRONOS (=Whitepaper), S. 21; KOST, Bewertung von Datenbanken, o.S.; Naumann, Übernahme von Daten aus Fachanwendungen, S. 33.

¹⁷² Interview von Donato Raselli mit Matthias Andenmatten, 2.5.2014.

wendung „socialWEB“ keine deskriptiven Zusatzinformationen zu den Tabellen und Attributen. Zudem sind in den Datenbanktabellen keine Codierungen enthalten.¹⁷³

Zur Anwendung „socialWEB“ steht ein umfangreiches Benutzerhandbuch zur Verfügung, das mit seiner Struktur, seinen Inhalten sowie den zahlreichen Screenshots ein wichtiges Hilfsmittel bildet, um die Anwendung zu verstehen sowie Prozesse / Prozessinstanzen nachzuvollziehen. Die Schwierigkeit besteht darin, dass das Handbuch nur online (HTML) abrufbar ist. Für die Archivierung müsste dieses in strukturierter Form in ein archivisches Dateiformat migriert werden (z.B. PDF/A).¹⁷⁴ Was die Ergänzung des Archivguts durch Anwendungserzeugnisse anbelangt, könnte für die archiwürdigen Informationsobjekte eine Auswahl entsprechender Dokumente, die sich über die Anwendung ausdrucken lassen, archiviert werden (z.B. Stammdatenblätter, Journaleinträge, Protokolle). Diese würden die Differenzierung von Prozessinstanzen fördern. Im Falle einer Archivierung durch die Archivierungsanwendung CHRONOS könnte auf die Archivierung von Anwendungserzeugnissen verzichtet werden, sofern der Nutzer-Client „Archive-Explorer“ entsprechend eingesetzt würde. Ferner kann auf die Aufnahme separater Screenshots an dieser Stelle verzichtet werden, da das Benutzerhandbuch zahlreiche und repräsentative Abbildungen enthält.¹⁷⁵

Technische Systemspezifikationen zur Anwendung „socialWEB“ wurden nicht erstellt.¹⁷⁶ Zudem erzeugen die Informationen, welche der Homepage des Herstellers zu entnehmen sind, gegenüber dem Benutzerhandbuch keinen Mehrwert.¹⁷⁷ Projektunterlagen sind vorhanden, da die Anwendung „socialWEB“ in Zusammenarbeit mit dem KJFS entwickelt wurde. Diese Unterlagen müssten aber im Hinblick auf eine Archivierung bewertet und erschlossen werden.¹⁷⁸ Um fehlende Angaben zu technischen Systemspezifikationen zu erhalten und die Umsetzung der Prozesse, die durch die Fachanwendung unterstützt werden, vertieft verstehen zu können, wäre die Ergänzung des Archivguts durch Interviews und Filmaufnahmen denkbar; dies ist aber abhängig von den zeitlichen und finanziellen Ressourcen, die ein Archiv für ein Datenbankarchivierungsprojekt einsetzen kann.

Analoge Akten, die bei der Umsetzung der Prozesse zusätzliche entstehen, können einen wichtigen Bestandteil des Archivguts bilden. Im Beispiel des KJFS handelt es sich dabei um Klientendossiers und Teamsitzungsprotokolle. Die Dossiers enthalten Berichte, die aus der Anwendung exportiert wurden (z.B. Stammdaten, Journaleinträge), während Dokumente in die Dossiers einfließen, die nicht aus der Anwendung stammen (z.B. Verfügungen, Korres-

¹⁷³ Ebd.

¹⁷⁴ Ebd. Vgl. KOST, Katalog archivischer Dateiformate (Textdaten).

¹⁷⁵ Vgl. socialweb Software GmbH, Benutzerhandbuch (diverse).

¹⁷⁶ Interview von Donato Raselli mit Matthias Andenmatten, 2.5.2014.

¹⁷⁷ Vgl. URL: <www.socialweb.ch> (30.4.2014).

¹⁷⁸ Interview von Donato Raselli mit Martina Suter, 25.4.2014.

pondenz, Elterngesprächsprotokolle). Die Dossiers können wesentlich dazu beitragen – in der Regel entspricht ein Dossier einem Prozessdurchlauf –, Prozessinstanzen zu identifizieren und Gesamtprozesse nachzuvollziehen.¹⁷⁹ Da die Dossiers Anwendungserzeugnisse enthalten, könnte darauf verzichtet werden, solche Erzeugnisse separat zu archivieren. Eine explizite Prozessdokumentation ist nicht vorhanden. Es ist jedoch eine „Matrix“ verfügbar, welche die einzelnen Prozessschritte der archivwürdigen Kernprozesse chronologisch darstellt und aufzeigt, welche Inhalte mit welchen Funktionen der Fachanwendung bearbeitet werden. Hierbei handelt es sich um ein Schlüsselement zur Nachvollziehbarkeit der Prozesse, die durch die Anwendung „socialWEB“ unterstützt werden. Abschließend kann das Archivgut mit einer Verfahrensdokumentation ergänzt werden, damit eine Nutzerschaft archivistische Entscheidungen nachvollziehen kann (vgl. Kapitel 5.1).¹⁸⁰

B) Fallbeispiel „ESCADA²“ (ED Basel-Stadt)

Auch die Fachanwendung „ESCADA²“ ist ein proprietäres Produkt. Die „Fachstelle Applikationen“ des ED Basel-Stadt geht davon aus – obwohl keine rechtliche Absicherung getroffen wurde –, dass der Hersteller für ein Archivierungsprojekt, das z.B. durch das Staatsarchiv Basel-Stadt durchgeführt würde, Zusatzinformationen wie das Datenbankmodell zur Verfügung stellt.¹⁸¹ Vermutlich müsste aber auch in diesem Falle die Informationsnutzung über eine Vereinbarung geregelt werden. Da die Schlüsselbeziehungen in der Datenbank der Anwendung „ESCADA²“ nicht applikatorisch definiert sind, könnten Informationen über das Datenbankmodell durch SIARD und CHRONOS auch indirekt hergeleitet werden. Deskriptive Zusatzinformationen zu den Tabellen und Attributen der Quelldatenbank sind nicht vorhanden. Möglicherweise enthält die Datenbank aber Codierungen. Entsprechende Codelisten – sofern vorhanden – müssten über den Hersteller besorgt werden.¹⁸²

Das Benutzerhandbuch ist für die Fachanwendung „ESCADA²“ ebenfalls als Internetversion verfügbar und muss für die Archivierung migriert werden. Das Handbuch ist detailliert und bildet ein zentrales Hilfsmittel, um die Funktionsweise der Anwendung zu verstehen und Prozesse / Prozessinstanzen zu identifizieren und nachzuvollziehen. Die Bearbeitung und Verwaltung der Daten wird beschrieben und anhand von Screenshots illustriert. Diese sind aber mit der Benutzeroberfläche der Anwendung nicht deckungsgleich, da sie teilweise veraltet sind. Daher müssten Screenshots neu erstellt werden, damit eine repräsentative nutzungsspezifische Sicht auf die Benutzeroberfläche bzw. Datenerfassungsmasken der An-

¹⁷⁹ Vgl. Gollino, Dossierbildung und Fachanwendungen, S. 365.

¹⁸⁰ Ebd. Vgl. KJFS, Matrix.

¹⁸¹ Interview von Donato Raselli mit Claude Stettler und Ali Aydin, 14.5.2014.

¹⁸² Ebd.

wendung ermöglicht wird.¹⁸³ Im Weiteren können über die Software „Crystal Reports“ zahlreiche Anwendungserzeugnisse erstellt werden, die als Bestandteil des Archivguts die Nachvollziehbarkeit der Prozesse / Prozessinstanzen erleichtern (z.B. Zeugnisse, Klassenlisten, Studentafeln). Auch hier könnte beim Einsatz der Archivierungsanwendung CHRONOS auf eine separate Übernahme von Anwendungserzeugnissen verzichtet werden.

Systemspezifikationen zur Fachanwendung „ESCADA²“ sind vorhanden und enthalten Angaben über die technische Beschaffenheit der Datenbank.¹⁸⁴ Ferner sind der Hersteller-Homepage zusätzliche Angaben zur Systemspezifikation zu entnehmen.¹⁸⁵ Im Weiteren sind aufgrund der öffentlichen Ausschreibung des Projekts „NeSuBa“ (vgl. Kapitel 4.1.2) Projektunterlagen vorhanden, die das Archivgut ergänzen und einer Nutzerschaft vertiefte Informationen über die Anwendung sowie die Prozesse, die durch sie unterstützt werden, liefern könnten. In diesem Zusammenhang wären insbesondere auch neuere Unterlagen von Bedeutung, da die Fachanwendung regelmäßig weiterentwickelt wird. Fehlende oder weiterführende Informationen könnten über Interviews kompensiert werden.¹⁸⁶

Bei der Umsetzung der Prozesse, die durch die Anwendung „ESCADA²“ unterstützt werden, entstehen vergleichbare analoge Unterlagen wie bei der Anwendung „socialWEB“. So werden im Rahmen des Prozesses „Schulbetriebsorganisation“ Schülerdossiers geführt und z.B. Zeugnisse und Lernberichte erstellt (Schülerverwaltung). Im Zuge der Klassenverwaltung entstehen u.a. Klassenlisten und Studentafeln sowie Protokolle und allgemeine Verwaltungsakten. In Bezug auf den Prozess „Personalorganisation“ sind an dieser Stelle die Personaldossiers zu erwähnen (Kapitel 4.1.2). Es handelt sich um Aktenbestände, die gemäß der bisherigen Praxis des Staatsarchivs Basel-Stadt bereits kontinuierlich in das Archiv übernommen werden.¹⁸⁷ Vor diesem Hintergrund könnte auf die Aufnahme von Fachanwendungserzeugnissen in eine Zusatzdokumentation verzichtet werden. Spezifische Prozessdokumentationen zu den Geschäftsprozessen sind nicht vorhanden. Auch in diesem Fallbeispiel sollte das Archivgut mit einer Verfahrensdokumentation ergänzt werden.

C) Zwischenfazit

Die Untersuchung der Zusatzdokumentation zeigt auf, dass für beide Fallbeispiele vergleichbare Materialien vorhanden sind. Materialien wie z.B. Handbücher, Projektunterlagen oder zusätzliche analoge Aktenunterlagen sind für jedes Fallbeispiel verfügbar. Jeweils nur für ein

¹⁸³ Vgl. Ebd.

¹⁸⁴ JCS Software AG, ESCADA².

¹⁸⁵ Vgl. JCS Software AG, ESCADA². Grundfunktionen.

¹⁸⁶ Interview von Donato Raselli mit Claude Stettler und Ali Aydin, 14.5.2014.

¹⁸⁷ Interview von Donato Raselli mit Christoph Döbeli, 14.5.2014.

Beispiel vorhanden sind z.B. Systemspezifikationen und Prozessdokumentationen. Es bietet sich an, fehlende Informationen durch Interviews (evtl. Filmaufnahmen) zu kompensieren. Sofern proprietäre Systeme eingesetzt werden, ist die Beschaffung des Datenbankmodells und allenfalls von Systemspezifikationen problematisch. Es ist daher ratsam, dass sich Datenbankbetreiber bereits bei der Anschaffung diesbezügliche Rechte zusichern lassen. Im Falle von SIARD und CHRONOS kann die Zusatzdokumentation an dem dafür vordefinierten Ort in das Archivgut aufgenommen werden. Für IngestList bleibt noch zu bestimmen, wie die Archivdaten mit Dokumentationsunterlagen ergänzt werden können. Was die zusätzlichen analogen Aktenunterlagen anbelangt, muss bei der Übernahme der migrierten Archivdaten in ein OAIS, im Zuge einer weiterführenden archivischen Erschließung eine Verbindung zu den entsprechenden analogen Archivbeständen aufgezeigt werden.

6. Schlussfolgerungen

Die digitale Langzeitarchivierung verfolgt das Ziel, digitale Informationen dauerhaft zu erhalten. Eine Herausforderung kommt hierbei der Archivierung von Datenbankinhalten aus Fachanwendungen zu. In der Praxis unterstützen datenbankgestützte Fachanwendungen maßgeblich die Bearbeitung von Geschäftsprozessen. Daher müssen bei der Archivierung von Daten aus Fachanwendungen Prozesse bzw. Prozessinstanzen, die zur Erzeugung der Daten geführt haben, im Archivgut nachvollzogen werden können. Zudem müssen – ausgehend von den allgemeinen Anforderungen an die digitale Archivierung – die Integrität und Authentizität der Archivdaten gewährleistet sein. Für die Umsetzung Archivierungsprojekte wurden im Archivbereich kommerzielle und nichtkommerzielle Verfahren bzw. Applikationen entwickelt, um Daten mittels der Migrationsstrategie aus Fachanwendungen, die durch relationale Datenbanken gestützt werden, in digitale Archive zu übernehmen.

Vor diesem Hintergrund wurde einleitend die Frage ins Zentrum gestellt, wie sich die unterschiedlichen Verfahren eignen, diesen Anforderungen gerecht zu werden und welche Informationen bei der Anwendung der jeweiligen Verfahren möglicherweise zusätzlich zu den Datenbankdaten archiviert werden müssen, um das Archivierungsziel zu erfüllen. Zur Beantwortung dieser Fragen wurden exemplarisch die Applikationen bzw. Archivierungsverfahren „IngestList“, „SIARD“ und „CHRONOS“ analysiert und theoretisch auf die als Fallbeispiele dienenden Fachanwendungen „socialWeb“ (KJFS) und „ESCADA²“ (ED Basel-Stadt) angewendet. In diesem Zusammenhang galt es auch, die Fallbeispiele zu analysieren und konkrete Archivierungsziele festzulegen. Der anschließende Empfehlungskatalog zeigt auf, mit welchen Informationen archivierte Datenbankinhalte ergänzt werden können, um die festgelegten Archivierungsziele zu erreichen.

Grundsätzlich eröffnen die untersuchten Archivierungsverfahren unterschiedliche Möglichkeiten – wie anhand der Fallbeispiele aufgezeigt –, um Datenbankinhalte partiell oder vollständig aus laufenden oder „stillzulegenden“ Fachanwendungen mittels Datenmigration in digitale Archive zu übernehmen. Dabei migrieren die Applikationen, unabhängig von der gewählten Vorgehensweise, die Datenbankinhalte aus Fachanwendungen in CSV- oder XML-Dateien, erheben automatisiert Metadaten und bieten die Möglichkeit, die Archivdaten manuell mit zusätzlichen Informationen zu ergänzen. Teilweise können die Archivdaten wieder in eine Datenbank importiert werden (SIARD, CHRONOS).

Im Hinblick auf die qualitativen Anforderungen, namentlich die Gewährleistung der Integrität und Authentizität der Archivdaten, bieten die untersuchten Archivierungsverfahren ähnliche Voraussetzungen. Alle Verfahren erzeugen zur Sicherung der Datenintegrität so genannte Prüfsummen. Unterschiede bestehen darin, dass die Archivierungsverfahren die Prüfsummen in einer unterschiedlichen „Dichte“ erzeugen. So erzeugt SIARD eine Prüfsumme für alle Primärdaten, IngestList für jede Archivdatei (Tabelle oder Sicht) und CHRONOS für jede einzelne Tabellenzeile (Datensatz) in einer Archivdatei sowie für die Archivdatei selbst. Insbesondere im Falle von CHRONOS lassen sich Integritätsverletzungen genau lokalisieren; zudem besteht hier auch die Möglichkeit, die Integrität allfälliger zusätzlicher Dokumentationsunterlagen zu gewährleisten. Im Weiteren bietet CHRONOS optional ergänzende Maßnahmen, um die Datenintegrität zu sichern. So können die Archivdaten verschlüsselt und mit einer digitalen Signatur versehen werden. Ferner können die Archivdaten einem spezialisierten externen Compliance-Speichersystem übergeben werden.

In Bezug auf die Authentizitätswahrung erheben alle Archivierungsverfahren automatisiert strukturelle sowie technische Metadaten und lassen deskriptive Anmerkungen zu. Im Falle von IngestList ist es zudem obligatorisch, manuell bestimmte deskriptive Metadaten aufzunehmen (z.B. Provenienzangabe). Ferner nutzt IngestList für die Erhebung technischer Metadaten die Anwendungen „JHOVE“ und „DROID“ („PRONOM“). Im Weiteren protokollieren die Archivierungsverfahren den Migrationsprozess. Bei den Archivierungsverfahren IngestList und CHRONOS können diese Protokolle auch nach dem Migrationsprozess weitergeführt werden (z.B. Protokollierung nachträglicher Integritätsprüfungen). Vor diesem Hintergrund kann hinsichtlich der qualitativen Anforderungen an die Archivierung von Daten aus datenbankgestützten Fachanwendungen festgehalten werden, dass die untersuchten Verfahren die Voraussetzungen bieten, um die Integrität und Authentizität von Archivdaten zu gewährleisten. CHRONOS bietet zudem optional ergänzende Sicherheiten.

Was die Erfüllung der inhaltlichen Anforderungen bzw. die Gewährleistung der Nachvollziehbarkeit der Prozesse bzw. Prozessinstanzen im Archivgut anbelangt, ist festzuhalten, dass die Datenbankinhalte nach der Migration nicht mehr durch eine Fachanwendung synthetisiert

werden und sich einer Nutzerschaft in Gestalt nutzungsspezifischer Sichten (Informationsobjekte) darbieten. Obwohl die Datenbankinhalte in lesbare Dateiformate konvertiert werden, ist es für Archivnutzer – in Anbetracht der potentiellen Menge und Komplexität der Archivdaten – eine Herausforderung, Informationsobjekte zu identifizieren und auf ursprüngliche Prozesse bzw. Prozessinstanzen zu schließen. Wie anhand der Fallbeispiele angedeutet wurde, wird die Herausforderung im Falle einer vollständigen Datenbankarchivierung erhöht. Demgegenüber kann der Fokus bei einer partiellen Archivierung bereits zu Beginn auf archivwürdige Informationsobjekte gerichtet werden. Die in dieser Untersuchung analysierten Verfahren bieten in unterschiedlichem Umfang Unterstützung dafür, dass Prozesse bzw. Prozessinstanzen in den Archivdaten möglicherweise nachvollzogen werden können.

Eine wichtige, von allen Archivierungsverfahren geleistete Unterstützung besteht in der Möglichkeit, die Archivdaten manuell mit deskriptiven Anmerkungen zu ergänzen, damit eine Nutzerschaft z.B. Hintergrundinformationen zu den einzelnen Tabellen und ihrer Funktion erhalten kann. Im Weiteren können über die deskriptiven Anmerkungen die Schlüsselbeziehungen einer Quelldatenbank erhalten werden, sofern diese wie im Fallbeispiel „socialWEB“ lediglich applikatorisch definiert sind oder durch ein Archivierungsverfahren – namentlich IngestList – bei der Migration nicht automatisiert in die Metadaten aufgenommen werden können. Werden die Schlüsselbeziehungen über die Metadaten erhalten und werden bei einem Archivierungsprojekt die Archivdaten mit deskriptiven Anmerkungen ergänzt, erhöht sich die Wahrscheinlichkeit, dass Prozesse bzw. Prozessinstanzen, die zur Erzeugung der Daten geführt haben, erahnt oder sogar rekonstruiert werden können. Datenbankinhalte, die durch SIARD und CHRONOS migrierte wurden, können zudem wieder in ein Datenbanksystem importiert werden; im Falle von SIARD auch Schlüsselbeziehungen (sofern automatisiert migriert). Dieser Umstand kann eine zusätzliche Unterstützung bei der Rekonstruktion der Prozesse bzw. Prozessinstanzen im Archivgut bieten.

Auch im Hinblick auf die Gewährleistung der Nachvollziehbarkeit der Prozesse bzw. Prozessinstanzen im Archivgut bietet das Archivierungsverfahren CHRONOS zusätzliche Unterstützung. So kann im Falle einer partiellen Datenbankarchivierung der Datenbankinhalt mit relativ geringem Aufwand direkt durch CHRONOS reduziert werden. Konkret lassen sich dadurch, sofern Schlüsselbeziehungen nicht applikatorisch definiert sind, z.B. automatisiert referentiell integrale Informationsobjekte in separate Archivdateien migrieren. Dies bedeutet einen wesentlichen Vorteil, wenn es darum geht, zusammengehörende Tabellen zu identifizieren und ihr Zusammenwirken nachzuvollziehen. Zusätzlich stellt CHRONOS mit dem „Archive Explorer“ einen Client zur Verfügung, der explizit dazu dient, einer Nutzerschaft die Archivdaten zugänglich zu machen. Dabei können mittels definierter Archiv-Views bzw. Such- und Reportvorlagen sowie Archiv-Deskriptoren nach der Migration wieder bestimmte

Tabelleninhalte zu nutzungsspezifischen Sichten synthetisiert werden. Hierbei besteht die Möglichkeit auch bei einer inkrementellen Archivierung einen einheitlichen Zugriff auf die Archivdaten zu erhalten. Ferner kann über den „Archiv-Explorer“ direkt auf allfälliges Dokumentationsmaterial zugegriffen werden. Diese Zusatzoptionen können die Nachvollziehbarkeit der Prozesse bzw. Prozessinstanzen im Archivgut wesentlich begünstigen.

Obwohl die Archivierungsverfahren unterstützende Funktionen bieten, um Prozesse bzw. Prozessinstanzen in den Archivdaten nachzuvollziehen, ist es notwendig, das Archivgut mit zusätzlichen Informationen zu ergänzen. Dadurch ergibt sich die Möglichkeit, dass Prozesse bzw. Prozessinstanzen im Archivgut nicht nur „erahnt“, sondern auch umfassend „nachvollzogen“ werden können; insbesondere bei Archivierung großer und komplexer Datenbanken. In diesem Zusammenhang wurde im Rahmen dieser Untersuchung ein „Empfehlungskatalog“ (vgl. Anhang 5) erarbeitet, der aufzeigt, mit welchen Zusatzinformationen das Archivgut ergänzt werden kann. Der Katalog berücksichtigt dabei Informationen, die sich auf die Ebenen der Datenbank und der Applikation (Fachanwendung) beziehen. Zudem umfasst er Informationen aus dem allgemeinen Kontext, in welchem eine Fachanwendung eingesetzt wird, sowie Informationen, die sich aus dem jeweiligen Archivierungsprojekt ergeben.

In Bezug auf die Datenbankebene kommt insbesondere dem Datenbankmodell eine zentrale Bedeutung zu. Ergänzende Informationen liefern zudem allfällig vorhandene Datenbeschreibungen (z.B. Herstellerinformationen) und Codelisten, um Codierungen aufzuschlüsseln. Auf der Ebene der Applikation können z.B. Benutzerhandbücher, Benutzungsreglemente oder Schulungsunterlagen wichtige Informationen enthalten, um die Konzeption und den Verwendungszweck einer Fachanwendung zu verstehen. Um nutzungsspezifische Sichten auf die Datenbankinhalte zu erhalten, können ferner Anwendungserzeugnisse (z.B. Reports) und Screenshots (z.B. von Datenerfassungsmasken) archiviert werden. Dokumentationsunterlagen, die sowohl die Ebene der Datenbank als auch der Applikation betreffen, sind Systemspezifikationen und Projektunterlagen (z.B. die Fachanwendung betreffend) oder Filmaufnahmen und gezielt durchgeführte Interviews mit Anwendern.

Bei Geschäftsprozessen, die durch Fachanwendungen unterstützt werden, können bei den Prozessdurchläufen Informationen entstehen, die nicht in eine Datenbank aufgenommen, sondern in analoger Form festgehalten werden. Solchen analogen Aktenunterlagen kann u.a. ein hoher Evidenzwert zukommen. Sie stellen möglicherweise wichtige Kontextinformationen bereit, um Prozesse bzw. Prozessinstanzen im Archivgut detailliert nachvollziehen zu können. Werden Prozessdokumentationen überliefert, kann die Einbindung einer Fachanwendung in Geschäftsprozesse konkret aufgezeigt werden. Abschließend sollte das Archivgut mit einer Verfahrensdokumentation ergänzt werden, in welcher die Vorgehensweise bei der Durchführung eines Datenbankarchivierungsprojekts festgehalten wird.

Die Untersuchung zeigt auf, dass die analysierten Verfahren IngestList, SIARD und CHRONOS die qualitativen Anforderungen an ein Datenbankarchivierungsprojekt (Integritäts- und Authentizitätsgewährleistung) zu erfüllen vermögen. In Bezug auf die inhaltlichen Anforderungen (Nachvollziehbarkeit der Prozesse bzw. Prozessinstanzen) vermögen sie Unterstützung zu leisten. Die kommerzielle Anwendung CHRONOS bietet sowohl in Bezug auf die inhaltlichen als auch die qualitativen Anforderungen ergänzende Möglichkeiten. Um die inhaltlichen Anforderungen umfassend zu erfüllen, ist es notwendig, das Archivgut mit Dokumentationsmaterial anzureichern. Die Auseinandersetzung mit konkreten Fallbeispielen hat aufgezeigt, dass entsprechendes Dokumentationsmaterial in einem bestimmten Umfang verfügbar ist. Doch kann die Verfügbarkeit von Dokumentationsmaterial insbesondere bei proprietären Fachanwendungen eingeschränkt sein (z.B. Datenbankmodell). Es ist daher zu empfehlen, bei der Einsetzung proprietärer Fachanwendungen sich entsprechende Rechte zusichern zu lassen. Zudem sollte für ein Archivierungsprojekt eine enge Zusammenarbeit mit dem Betreiber und dem Hersteller einer Fachanwendung angestrebt werden.

7. Bibliografie

A. Literatur

Agnolazza, Daniele, Projekt „NeSuBa“ planmäßig abgeschlossen, in: Basler Schulblatt 2005 (8), S. 22.

Altenhöner, Reinhard/**Schrimpf**, Sabine, Langzeitarchivierung, in: Kuhlen, Rainer/Semar, Wolfgang et al. (Hg.), Handbuch zur Einführung in die Informationswissenschaft und -praxis, Berlin 2013, S. 529-540.

Authenticity Task Force (Hg.), Requirements for Assessing and Maintaining the Authenticity of Electronic Records. Appendix 2, o.O. 2002, <http://www.interpares.org/book/interpares_book_k_ap02.pdf> (12.03.2014).

Banat-Berger, Françoise, L'archivage numérique à long terme. Les débuts de la maturité?, Paris 2009.

BAR/Schweizerisches Bundesarchiv, Datenbanken für die Zukunft sichern. Die SIARD Archivierungslösung für relationale Datenbanken (=Factsheet), Bern o.D., Online: URL: <<http://www.bar.admin.ch/dienstleistungen/00823/00825/index.html?lang=de>> (16.2.2014).

BAR/Schweizerisches Bundesarchiv, SIARD. Formatbeschreibung (=Format Version 1.0, Dokumentversion 1.2), Bern 2011, Online: URL: <http://www.bar.admin.ch/dienstleistungen/00823/00825/index.html?lang=de&download=NHZLpZeg7t,Inp6l0NTU042l2Z6ln1acy4Zn4Z2qZpnO2YUq2Z6gpJCDdIR8fmym162epYbg2c_JjKbNoKSn6A-->> (16.2.2014).

BAR/Schweizerisches Bundesarchiv, SIARD Suite: FAQ (=Frequently Asked Questions), Online: URL: <http://www.bar.admin.ch/dienstleistungen/00823/00825/index.html?lang=de> (2.6.2014).

Bernstein, Amir, Database Preservation: The international Challenge and the Swiss Solution, in: Digital preservation Europe (briefing paper) 2008, Online: URL: <<http://www.digitalpreservationeurope.eu/publications/briefs/>> (16. 2.2014).

- Borghoff, Uwe/Rödig, Peter et al.**, Langzeitarchivierung. Methoden zur Erhaltung digitaler Dokumente, Heidelberg 2003.
- Borghoff, Uwe/Rödig, Peter**, Technik der digitalen Langzeitarchivierung. Zuverlässig von Nullen und Einsen zur Information – auch noch nach zehn, fünfzig oder mehr Jahren?, in: Zeitschrift für Datenrecht und Informationssicherheit. Digma 2010 (1), S. 6-19.
- Brandl, Stefan/Keller-Marxer, Peter**, Long-term Archiving of Relational Databases with CHRONOS, o.O. 2007, Online: URL: <http://omepages.inf.ed.ac.uk/hmueller/presdb07/papers/Germany_CHRONOS_PresDB07.pdf> (16.2.2014).
- Brown, Adrian**, Practical digital preservation. A how-to guide for organizations of any size, London 2013.
- Brübach, Nils**, Das Referenzmodell OAIS, in: Neuroth, Heike/Osswald, Achim et al. (Hg.), nestor Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.0, Boizenburg 2009, Kap. 4, S. 3-14.
- Christophides, Vassilis/Buneman, Peter**, Report on the First International Workshop on Database Preservation, in: SIGMOD Record 2007 (3), S. 55-58.
- CSP GmbH**, Benutzerhandbuch. CHRONOS Hilfe, o.O. 2010.
- Dässler, Rolf/Schwarz, Karin**, Archivierung und dauerhafte Nutzung von Datenbankinhalten aus Fachverfahren. Eine neue Herausforderung für die digitale Archivierung, in: Der Archivar 2010 (1), S. 6-18.
- ERPANET/Electronic Resource Preservation an Access Network (Hg.)**, The Long-term Preservation of Databases (=Workshop Report), Bern 2003, Online: URL: <http://www.erpanet.org/events/2003/bern/Bern_Report_final.pdf> (16.2.2014).
- Fitzgerald, Neal**, Using data archiving tools to preserve archival records in business systems – a case study, in: Borbinha, José/Nelson, Michael/Knight, Steven (Editors), iPres2013. Proceedings of the 10th International Conference on Preservation of Digital Objects (3-5 September, Lisbon/Portugal), Lisbon 2013, S. 268-271.
- Funk, Stefan E.**, Migration, in: Neuroth, Heike/Osswald, Achim et al. (Hg.), nestor Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.0, Boizenburg 2009, Kap. 8, S. 10-15.
- Funk, Stefan E.**, Emulation, in: Neuroth, Heike/Osswald, Achim et al. (Hg.), nestor Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.0, Boizenburg 2009, Kap. 8, S. 16-23.
- Funk, Stefan E.**, Das JSTOR/Harvard Object Validation Environment (JHOVE), in: Neuroth, Heike/Osswald, Achim et al. (Hg.), nestor Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.0, Boizenburg 2009, Kap. 13, S. 21-28.
- Giaretta, David**, Advanced digital preservation, Berlin 2011.
- Glöde, Julia**, Archivierung relationaler Datenbanken auf der Grundlage von XML. Konzeption eines OAIS-konformen Archivierungsmodells und Entwicklung eines neuen Ansatzes zur Archivierung von Datenbankabfragen (=Masterarbeit FH Potsdam), Berlin 2009, Online: URL: <<http://opus4.kobv.de/opus4-fhpotsdam/frontdoor/index/index/docId/110>> (16.2.2014).
- Gollino, Erich**, „Wie Feuer und Wasser?“ – Dossierbildung und Fachanwendungen, in: Coutaz, Gilbert/Knoch-Mund, Gaby/Reimer, Ulrich (Hg.), Informationswissenschaft: Theorie, Methode und Praxis (=MAS ALIS 2010-2012), Baden 2014, S. 363-385.

- Hartwig**, Thomas, SIARD Suite. Handbuch (Siard Suite 1.0, SiardEdit 1.32) (=herausgegeben vom Schweizerischen Bundesarchiv), Bern 2013.
- Heuscher**, Stephan/**Keller-Marxer**, Peter, XML – ein strategisches Instrument für Archive?, in: Arbido 2003 (3), S. 16-18.
- Heuscher**, Stephan/**Keller-Marxer**, Peter et al., Providing Authentic Long-term Archival Access to Complex Relational Data, Bern 2004, Online: URL: <<http://arxiv.org/pdf/cs/0408054v1.pdf>> (15.2. 2014).
- Hoer**, Barbara, Auslandberichterstattung. Internationales. Workshop „Long-term Preservation of Databases“ im Schweizerischen Bundesarchiv, in: Der Archivar 2003 (4), S. 347f.
- Hollmann**, Michael, Das „Digitale Archiv“ des Bundesarchivs. Die Archivierung genuin elektronischer Unterlagen im Bundesarchiv (=Vortrag, 8th European Conference on Digital Archiving, Genf, 28.-30. April), o.O. 2009, Online: URL: <<http://www.bar.admin.ch/aktuell/00568/00702/00861/01572/index.html?lang=de>> (16.2.2014).
- Internationale Organisation für Normung** (Hg.), Information und Dokumentation. Schriftgutverwaltung (ISO 15489-1), o.O. 2001.
- Keitel**, Christian, Ways to deal with complexity (=Paper, iPRES conference, British Library, 30.9.2008), Ludwigsburg 2008, Online: URL: <http://www.bl.uk/ipres2008/presentations_day2/45_Keitel.pdf> (16.2.2014).
- Keitel**, Christian, Digitale Archivierung beim Landesarchiv Baden-Württemberg, in: Der Archivar 2010 (1), S. 19-26.
- Keitel**, Christian/**Lang**, Rolf, Ingest von Fachverfahren im Landesarchiv Baden-Württemberg. Anmerkungen zu Authentizität, Prozessen und Softwareentwicklung, in: Staatsarchiv St. Gallen (Hg.), Entwicklung in den Bereichen Records Management / Vorarchiv – Übernahme – Langzeitarchivierung. 13. Tagung des Arbeitskreises „Archivierung von Unterlagen aus digitalen Systemen“ vom 27./28. April 2009 ausgerichtet vom Staatsarchiv St. Gallen, St. Gallen 2009, S. 35-45.
- Keitel**, Christian/**Lang**, Rolf, DIMAG und IngestList. Übernahme, Archivierung und Nutzung von digitalen Unterlagen im Landesarchiv Baden-Württemberg, in: Maier, Gerald/Fritz, Thomas (Hg.), Archivische Informationssysteme in der digitalen Welt. Aktuelle Entwicklungen und Perspektiven, Stuttgart 2010, S. 53-63.
- Keller-Marxer**, Peter, CHRONOS. Unternehmenslösung für die professionelle Archivierung von Datenbanken (=Whitepaper), Bern 2012.
- Keller-Marxer**, Peter, Umfang und Zweck von Metadaten in CSP CHRONOS Database Archiving, Bern 2012.
- Kemper**, Alfons/**Eickler**, André, Datenbanksysteme. Eine Einführung, München 2013.
- Kortyla**, Stephanie, Übernahme aus Ratsinformationssystemen (=Masterarbeit FH Potsdam), Potsdam 2013, Online: URL: <<http://opus4.kobv.de/opus4-fhpotsdam/frontdoor/index/index/docId/518>> (16.2.2014).
- KOST**/Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen (Hg.), Bewertung von Datenbanken, Online: URL: <http://kost-ceco.ch/cms/index.php?database_de> (16.2.2014).

- KOST**/Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen (Hg.), Katalog archivischer Dateiformate (V. 3.0). Datenbanken, Online: URL: <<http://www.kost-ceco.ch/wiki/whelp/KaD/>> (16.4.2014).
- KOST**/Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen (Hg.), Katalog archivischer Dateiformate (V. 3.0). Textdaten, Online: URL: <<http://www.kost-ceco.ch/wiki/whelp/KaD/>> (17.6.2014).
- Landesarchiv Baden-Württemberg**, Landesarchiv Baden-Württemberg veröffentlicht kostenlose Software zur digitalen Archivierung (=Pressebericht vom 5.5.2009), Stuttgart 2009, Online: URL: <<http://www.landearchiv-bw.de/web/49289>> (9.4.2014).
- Liegmann**, Hans/Neuroth, Heike, Einführung, in: Neuroth, Heike/Osswald, Achim et al. (Hg.), nestor Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.0, Boizenburg 2009, Kap. 1, S. 1-10.
- Lindley**, Andrew, Database Preservation Evaluation Report – SIARD vs. CHRONOS. Preserving complex structures as databases through a record centric approach?, in: Borbinha, José/Nelson, Michael/Knight, Steven (Editors), iPres2013. Proceedings of the 10th International Conference on Preservation of Digital Objects (3-5 September, Lisbon/Portugal), Lisbon 2013, S. 29-38.
- Naumann**, Kai, Übernahme von Daten aus Fachanwendungen – Schnittstellen, Erhaltungsformen, Nutzung, in: Generaldirektion der Staatlichen Archive Bayerns (Hg.), Neue Entwicklungen und Erfahrungen im Bereich der digitalen Archivierung: von der Behördenberatung zum Digitalen Archiv. 14. Tagung des Arbeitskreises „Archivierung von Unterlagen aus digitalen Systemen“ vom 1./2. März 2010 in München (=Sonderveröffentlichungen der Staatlichen Archive Bayerns, Nr. 7), München 2010, S. 26-36.
- Nestor-Arbeitsgruppe** (Hg.), Referenzmodell für ein Offenes Archiv-Informationssystem (=Deutsche Übersetzung), Frankfurt am Main 2012.
- Neubaur**, Mathias, Extraktion, technische Metadaten und Tools, in: Neuroth, Heike/Osswald, Achim et al. (Hg.), nestor Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.0, Boizenburg 2009, Kap. 7, S. 15-18.
- Neuroth**, Heike/Osswald, Achim et al. (Hg.), nestor Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.0, Boizenburg 2009.
- Ohnesorge**, Krystyna/D’Incau, Loris et al., Digitale Archivierung im Bundesarchiv. Wie sich das Schweizerische Bundesarchiv für die Übernahme von immer mehr elektronischen Dokumenten rüstet, in: Zeitschrift für Datenrecht und Informationssicherheit 2010 (1), S. 12-19.
- Olson**, Jack E., Database Archiving. How to Keep Lots of Data for a Very Long Time, Amsterdam 2009.
- Raselli**, Donato, Verfahren zur Wahrung der Integrität und Authentizität in digitalen Archiven. Die „Berner Archiv AG“ (=Hausarbeit, MAS ALIS Universität Bern), Bern 2013.
- Ribeiro**, Christina, Database Preservation, in: Digital preservation Europe (briefing paper) 2009, Online: URL: <<http://www.digitalpreservationeurope.eu/publications/briefs/>> (Stand: 16.2.2014).
- Scheffel**, Regine, Einführung, in: Neuroth, Heike/Osswald, Achim et al. (Hg.), nestor Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.0, Boizenburg 2009, Kap. 2, S. 1.

Schubert, Matthias, Datenbanken. Theorie, Entwurf und Programmierung relationaler Datenbanken, Wiesbaden 2007.

Steiner, René, Grundkurs Relationale Datenbanken. Einführung in die Praxis der Datenbankentwicklung für Ausbildung, Studium und IT-Berufe, Wiesbaden 2009.

Täuber, Mario, Herstellererklärung (=CHRONOS 4.2), Grossköllnbach 2013.

Toebak, Peter, Records Management. Ein Handbuch, Baden 2007.

Ullrich, Dagmar, Bitstream Preservation, in: Neuroth, Heike/Osswald, Achim et al. (Hg.), nestor Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Version 2.0, Boizenburg 2009, Kap. 8, S. 3-9.

Unterstein, Michael/**Matthiessen**, Günter, Relationale Datenbanken und SQL in Theorie und Praxis, Heidelberg 2012.

VdA/Verband deutscher Archivarinnen und Archivare (Hg.), Bewertung elektronischer Fachverfahren. Diskussionspapier des VdA-Arbeitskreises „Archivische Bewertung“, Online: URL: <<http://www.vda.archiv.net/aktuelles/meldung/257.html>> (15.2.2014).

Zürcher, Thomas, Sicherung und Bewertung von Datenbanken im Schweizerischen Bundesarchiv, in: Arbido 2003 (3), S. 19-21.

B. Quellen

B.a. Unpublizierte Quellen

BAR/Schweizerisches Bundesarchiv, Dokumentation Datenbanken / Fachapplikationen (=Informationsblatt, v1.0), o.O. o.D..

JCS Software AG, ESCADA² Benutzerhandbuch. FakturaGST (=in ein PDF konvertierte CHM-Datei), o.O. o.D..

JCS Software AG, ESCADA² Benutzerhandbuch. PLSGST (=in ein PDF konvertierte CHM-Datei), o.O. o.D..

JCS Software AG, ESCADA² Benutzerhandbuch. SchuleGST (=in ein PDF konvertierte CHM-Datei), o.O. o.D..

JCS Software AG, ESCADA². Software für die Schuladministration, Aesch 2014.

Kompetenzzentrum Jugend und Familie Schlossmatt (KJFS), Matrix. Prozessvariablen/Stammdaten, o.O. o.D..

socialweb Software GmbH, socialWEB Benutzerhandbuch. Anleitung Arbeitsplanung, o.O. o.D..

socialweb Software GmbH, socialWEB Benutzerhandbuch. Anleitung Arbeitszeiterfassung, o.O. o.D..

socialweb Software GmbH, socialWEB Benutzerhandbuch. Anleitung Buchhaltung, o.O. o.D..

socialweb Software GmbH, socialWEB Benutzerhandbuch. Anleitung Einsatz/Freiwünsche und Kontingente, o.O. o.D..

socialweb Software GmbH, socialWEB Benutzerhandbuch. Anleitung KlientInnenmodul, o.O. o.D..

socialweb Software GmbH, socialWEB Benutzerhandbuch. Anleitung KlientInnenmodul – Zielprozesse und Ein-/Austritt von KlientInnen, o.O. o.D..

socialweb Software GmbH, socialWEB Benutzerhandbuch. Anleitung Kontaktmodul, o.O. o.D..

socialweb Software GmbH, socialWEB Benutzerhandbuch. Anleitung MitarbeiterInnenmodul, o.O. o.D..

socialweb Software GmbH, socialWEB Benutzerhandbuch. Anleitung Protokolle, o.O. o.D..

socialweb Software GmbH, socialWEB Produktdokumentation. Branchensoftware für soziale Einrichtungen (=Version 2.0), o.O. o.D..

Interviews:

Interview von Donato Raselli und Brigitte Sacker (Beisitz, Stadtarchiv Bern) mit Martina Suter (Leiterin KJFS), 25.4.2014.

Interview von Donato Raselli mit Matthias Andenmatten (Geschäftsführer der „socialweb Software GmbH“), 2.5.2014.

Interview von Donato Raselli mit Christoph Döbeli (Leiter Information und Dokumentation des ED Basel-Stadt), 14.5.2014.

Interview von Donato Raselli mit Claude Stettler (Leiter Fachstelle Applikationen des ED Basel-Stadt) und Ali Aydin (IC-Berater II des ED Basel-Stadt), 14.5.2014.

Korrespondenz:

Korrespondenz zwischen Donato Raselli und Viviane Cantaluppi (=JCS Software AG), 14.4.2014.

Korrespondenz zwischen Donato Raselli und Peter Keller (=ikeep AG, CHRONOS-Distributor Schweiz), 5.6.2014.

B.b. Publierte Quellen

Gesundheits- und Fürsorgedirektion BE (GEF), Allgemeine Vertragsbestimmungen zum Leistungsvertrag 2014 zwischen dem Kanton Bern und den Institutionen für Kinder und Jugendliche, Online: URL: <http://www.gef.be.ch/gef/de/index/direktion/organisation/alba/formulare/institutionen_fuerkinderundjugendliche.assetref/content/dam/documents/GEF/ALBA/de/Formulare_Bewilligungen_Gesuche/Institutionen_Kinder_Jugendliche/LV_AVB_2014_KiJu_de.pdf> (30.4.2014).

JCS Software AG, ESCADA². Grundfunktionen, Online: URL: <<http://www.jcssoftware.ch/escada2/index.html>> (18.5.2014).

JCS Software AG, ESCADA². Grundschulen, Online: URL: <<http://www.jcssoftware.ch/escada2/index.html>> (18.5.2014).

JCS Software AG, ESCADA². Stammdaten, Online: URL: <<http://www.jcssoftware.ch/escada2/index.html>> (18.5.2014).

Kompetenzzentrum Jugend und Familie Schlossmatt (KJFS), Leitbild des Kompetenzzentrums Jugend und Familie Schlossmatt, Bern 2010.

Erlasse:

„Gesetz betreffend die Organisation des Regierungsrates und der Verwaltung des Kantons Basel-Stadt“, 22.4.1976, in: Systematische Gesetzessammlung Basel-Stadt (SG, 153.100).

„Schulgesetz“, 4.4.1929, in: Systematische Gesetzessammlung Basel-Stadt (SG, 410.100).

„Verfassung des Kantons Basel-Stadt“, 23.3.2005, in: Systematische Gesetzessammlung Basel-Stadt (SG, 111.100).

„Verordnung für die Schulleitungen der Volksschulen“, 26.6.2012, in: Systematische Gesetzessammlung Basel-Stadt (SG, 411.350).

„Verordnung über die Organisation der Stadtverwaltung“, 27.2.2001, in: Systematische Sammlung des Stadtrechts Bern (SSSB, 152.01).

C. Applikationen

CHRONOS (=Version 4.5.1), Softwarehersteller CSP GmbH (Deutschland) / ikeep AG (Schweiz).

IngestList6 (=6. Version der Applikation), Online: URL: <<http://sourceforge.net/projects/ingestlist/>> (9.4.2014).

SIARD Suite 1.0 (SiardEdit 1.50) / SIARD-Format 1.0, Online: URL: <<http://www.bar.admin.ch/dienstleistungen/00823/00825/index.html?lang=de>> (16.4.2014).

D. Internet

<<http://www.arxiv.org>> (15.2.2014).

<<http://www.bar.admin.ch>> (20.6.2014).

<<http://www.bl.uk>> (16.2.2014).

<<http://www.digitalpreservationeurope.eu>> (16.2.2014).

<<http://www.erpanet.org>> (16.2.2014).

<<http://www.interpares.org>> (12.3.2014).

<<http://www.jcssoftware.ch>> (16.5.2014).

<<http://www.kost-ceco.ch>> (20.6.2014).

<<http://www.landesarchiv-bw.de>> (9.4.2014).

<<http://www.omepages.inf.ed.ac.uk>> (16.2.2014).

<<http://www.opus4.kobv.de>> (16.2.2014).

<<http://www.oracle.com>> (31.5.2014).

<<http://www.schulharmonisierung-bs.ch>> (17.5.2014).

<<http://www.socialweb.ch>> (30.4.2014).

<<http://www.sourceforge.net>> (9.4.2014).

<<http://www.vda.archiv.net>> (15.2.2014).

Anhang 1: Elemente des relationalen Datenmodells¹⁸⁸

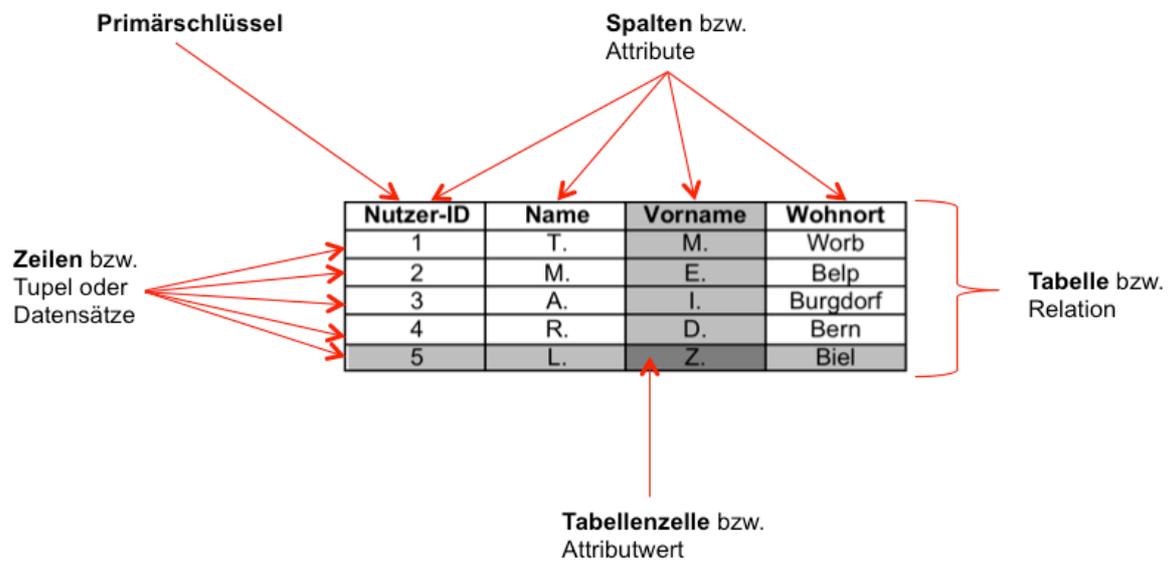


Abbildung 1: Elemente des relationalen Datenmodells

¹⁸⁸ Vgl. Glöde, Archivierung relationaler Datenbanken, S. 14.

Anhang 2: Migration von Daten aus einer Datenbank und Aufnahme deskriptiver Anmerkungen mit IngestList¹⁸⁹

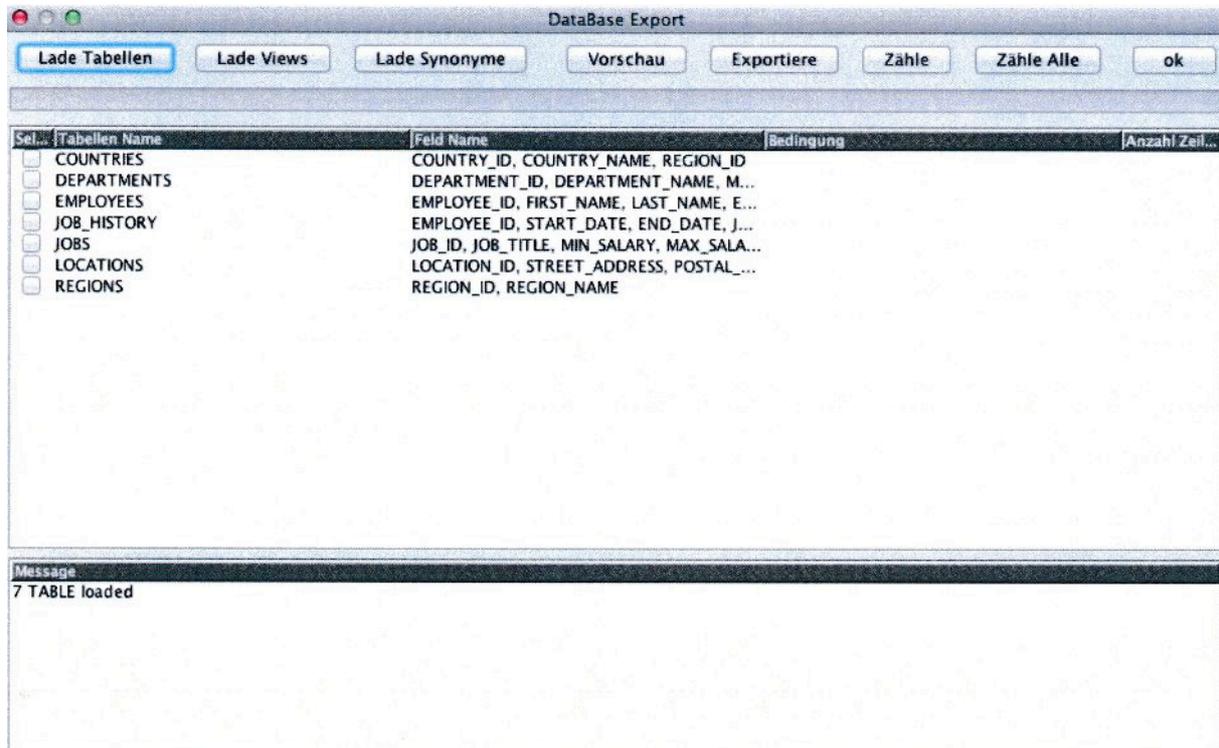


Abbildung 2: Die Tabellen einer Datenbank (hier „Human Resources-Testdatenbank“ von Oracle) werden auf die Benutzeroberfläche von IngestList ausgelesen und können selektiert werden.

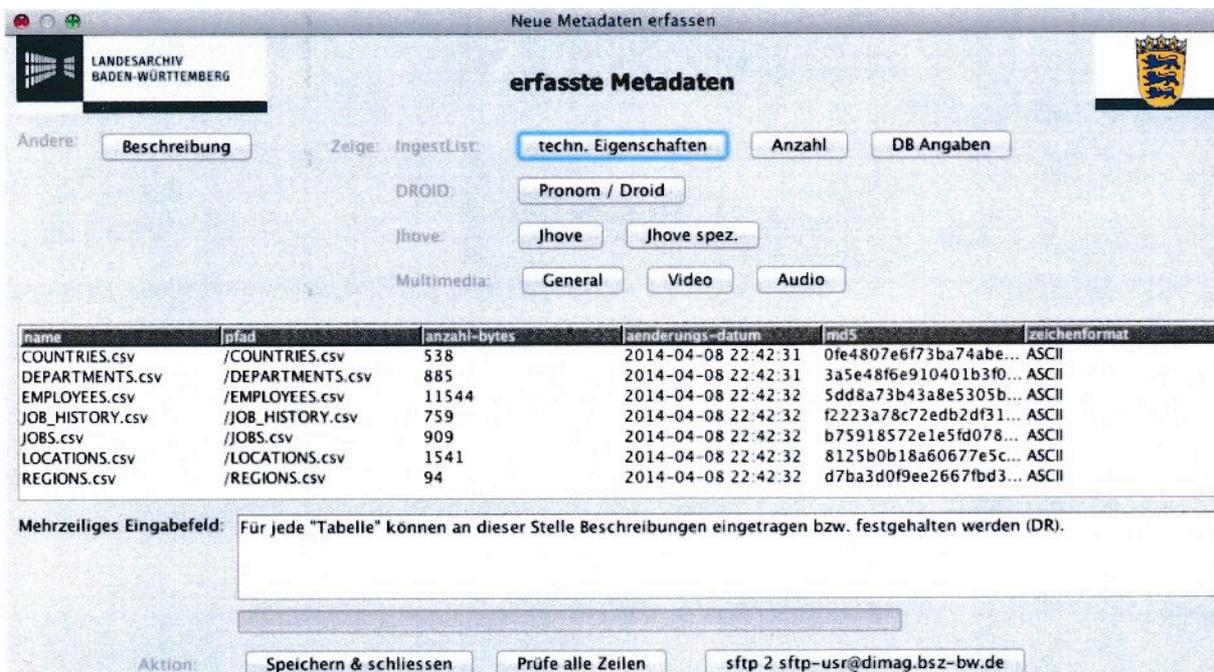


Abbildung 3: Für jede migrierte Tabelle können über IngestList deskriptive Metadaten aufgenommen werden.

¹⁸⁹ „IngestList6“ (=6. Version der Applikation).

Anhang 4: Migration von Daten aus einer Datenbank und Aufnahme deskriptiver Anmerkungen mit CHRONOS¹⁹¹

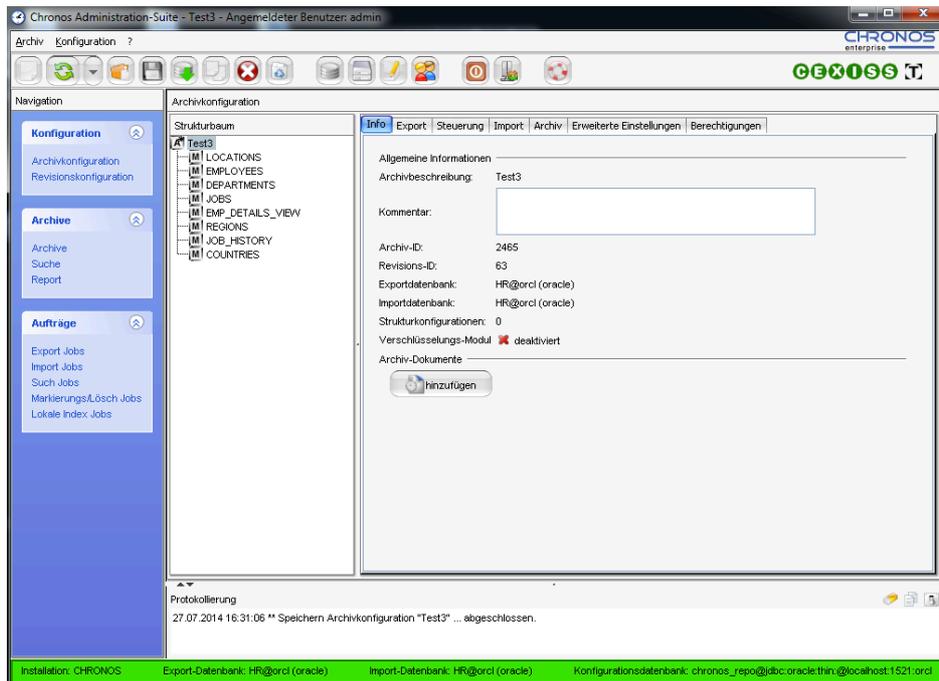


Abbildung 5: Bei der vollständigen Archivierung einer Datenbank werden alle Tabellen flach auf einem Niveau des Strukturbaums angezeigt (hier „Human Resources-Testdatenbank“ von Oracle). Auf der rechten Seite der Benutzeroberfläche können deskriptive Metadaten aufgenommen werden.

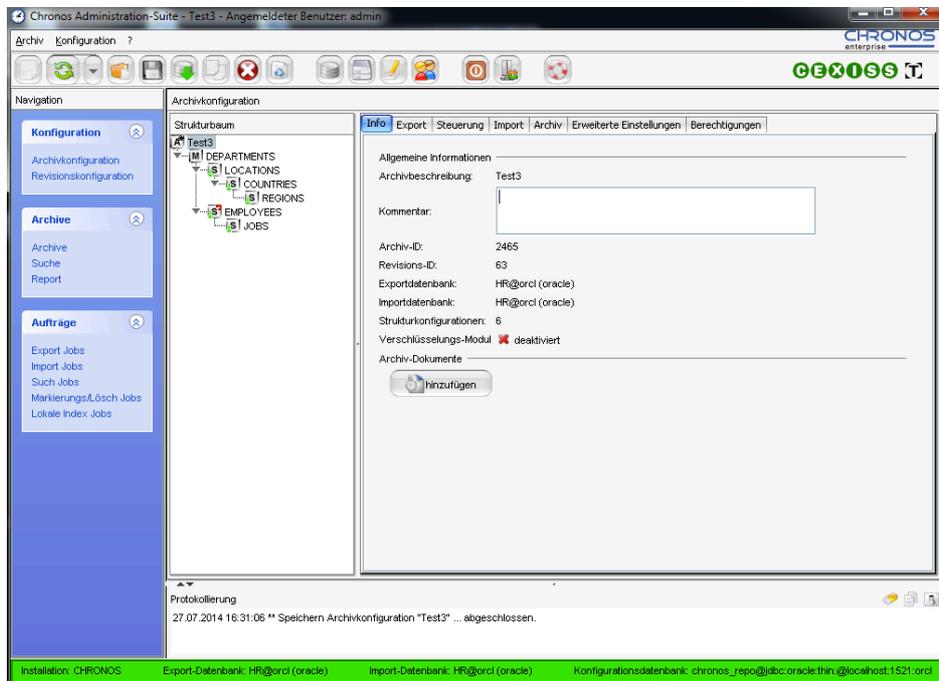


Abbildung 6: Bei einer partiellen Archivierung kann z.B. eine Tabelle als „Mastertabelle“ definiert werden (hier die Tabelle „Departments“). CHRONOS kann in der Folge automatisiert alle Tabellen herauslesen, die über Schlüsselbeziehungen mit der Mastertabelle in Verbindung stehen (Child-Parent-Beziehung). Dadurch können referentiell integrale Informationsobjekte migriert bzw. archiviert werden (hier „Human Resources-Testdatenbank“ von Oracle).

¹⁹¹ Chronos (=Version 4.5.1).

Anhang 5: Empfehlungskatalog: Zusatzdokumentation für Archivdaten¹⁹²

	Bestandteile einer Dokumentation	Erläuterungen
Datenbank-ebene	Datenbankmodell	Ein Datenbankmodell gibt Auskunft über alle Tabellen (teilweise inklusive Attribute) und Schlüsselbeziehungen in einer relationalen Datenbank (grafische Darstellung).
	Datenbeschreibung / Codeliste	Datenbeschreibungen sind deskriptive Zusatzinformationen zu den einzelnen Tabellen und Attributen (z.B. Herstellerinformationen). Werden in einer Tabelle Codierungen verwendet, müssen Codelisten zur Verfügung gestellt werden, damit die Daten übersetzt bzw. verstanden werden können.
Applikations-ebene	Benutzerhandbuch / Benutzungsreglement / Schulungsunterlagen	Diese Dokumentationsunterlagen enthalten in der Regel wichtige Informationen über die Konzeption und den Verwendungszweck einer Fachanwendung. Sie können zudem auch technische Angaben enthalten.
	Anwendungserzeugnisse	Produkte einer Fachanwendung (z.B. Reports) können eine nutzungsspezifische Sicht auf Datenbankinhalte ermöglichen. Sie zeigen, welche Teildaten für einen bestimmten Verwendungszweck synthetisiert wurden.
	Screenshots	Screenshots zentraler Datenerfassungsmasken unterstützen die Nachvollziehbarkeit der Funktionsweise einer Fachanwendung. Zudem ermöglichen sie nutzungsspezifische Sichten auf die Daten zu erhalten.
Datenbank- und Applikations-ebene	Systemspezifikation	Systemspezifikationen enthalten wichtige Information über die Architektur des Datenbanksystems und/oder die Beschaffenheit der Applikation.
	Projektunterlagen	Projektunterlagen können Informationen zur Konzeption einer Fachanwendung (vgl. Systemspezifikationen) enthalten. Sie zeigen auf, wie ein System entwickelt und verändert wurde. Möglicherweise enthalten sie auch Informationen über den Verwendungszweck einer Fachanwendung (vgl. Benutzerhandbücher).
	Interview	Interviews können – in Abhängigkeit von den Fragestellungen – auf einer breiten Ebene Informationen zur Verwendung und Beschaffenheit einer Fachanwendung enthalten. Es ist denkbar, dass ein Mangel an Dokumentationsunterlagen durch Interviews kompensiert wird.
	Filmaufnahme	Kommentierte Bildschirmaufzeichnungen zeigen beispielsweise, wie Anwender mit einer Fachanwendung arbeiten. Dadurch kann das Verständnis, wie Fachanwendungen in Prozesse eingebunden sind, erhöht werden.
Kontextebene	Analoge Aktenunterlagen	Analoge Aktenunterlagen enthalten Informationen, die möglicherweise migrierte Datenbankinhalte ergänzen und wesentlich dazu beitragen, dass im Archivgut Prozesse und Prozessinstanzen nachvollzogen werden können.
	Prozessdokumentation	Prozessdokumentationen zeigen die Struktur bzw. die Konzeption von Geschäftsprozessen auf. Anhand einer Prozessdokumentation könnte die Einbindung einer Fachanwendung in einen Prozess präzise nachvollzogen werden. Dabei wird ersichtlich, welche Rolle einer Fachanwendung bei der Umsetzung bestimmter Aufgaben zukommt.
Archivierung	Verfahrensdokumentation	Verfahrensdokumentationen zeigen auf, welche Vorgehensweise bei der Durchführung eines Datenbankarchivierungsprojekts gewählt und welche Entscheidungen getroffen wurden: z.B. muss festgehalten werden, welche Datenbankinhalte bei der Reduktion des Datenbankmodells (partielle Datenbankarchivierung) nicht berücksichtigt wurden.

¹⁹² Vgl. Fussnoten 164 bis 172.