

MediaFabric: ein Framework für standortübergreifende Archiv-Lösungen

MediaFabric: a framework for multi-site archiving solutions

Michael Zimmermann [VCS Aktiengesellschaft]

Klaus M. Heidrich [VCS Aktiengesellschaft]

- Ein Beitrag im Rahmen des Vortragsprogramms der Tonmeistertagung 2004 -

Einführung

Wenn man die Entwicklung der Hörfunkdigitalisierung der vergangenen Jahre betrachtet, so stand die Lösungsentwicklung unabhängiger Funktionsbereiche im Vordergrund. In den Bereichen Planung, Produktion und Playout sind verschiedenste Produkte entstanden, die in Ihrer Funktion, Ihren Formaten und Ihrer Datenhaltung sehr unterschiedlich sein können. Auf Grundlage dieser Gegebenheiten verfolgt der Framework-Ansatz von MediaFabric das Ziel, produktübergreifende Problemstellungen zu lösen, wie zum Beispiel zentralisierte Hörfunkarchive. Wie sich die Anforderungen und das Design im Vergleich zu herkömmlichen Anwendungen unterscheiden und welche zusätzlichen Risiken auftreten, ist Thema dieses Vortrags.

Grundlagen

Anforderungen an eine universelle, digitale Archivlösung

Die Anforderungen an ein *Archiv im allgemeinen* sind relativ offensichtlich, es sorgt für die sichere Verwahrung und Katalogisierung von Dingen. Es muß so beschaffen sein, dass dort abgelegte Dinge schnell und sicher wiederzufinden sind.

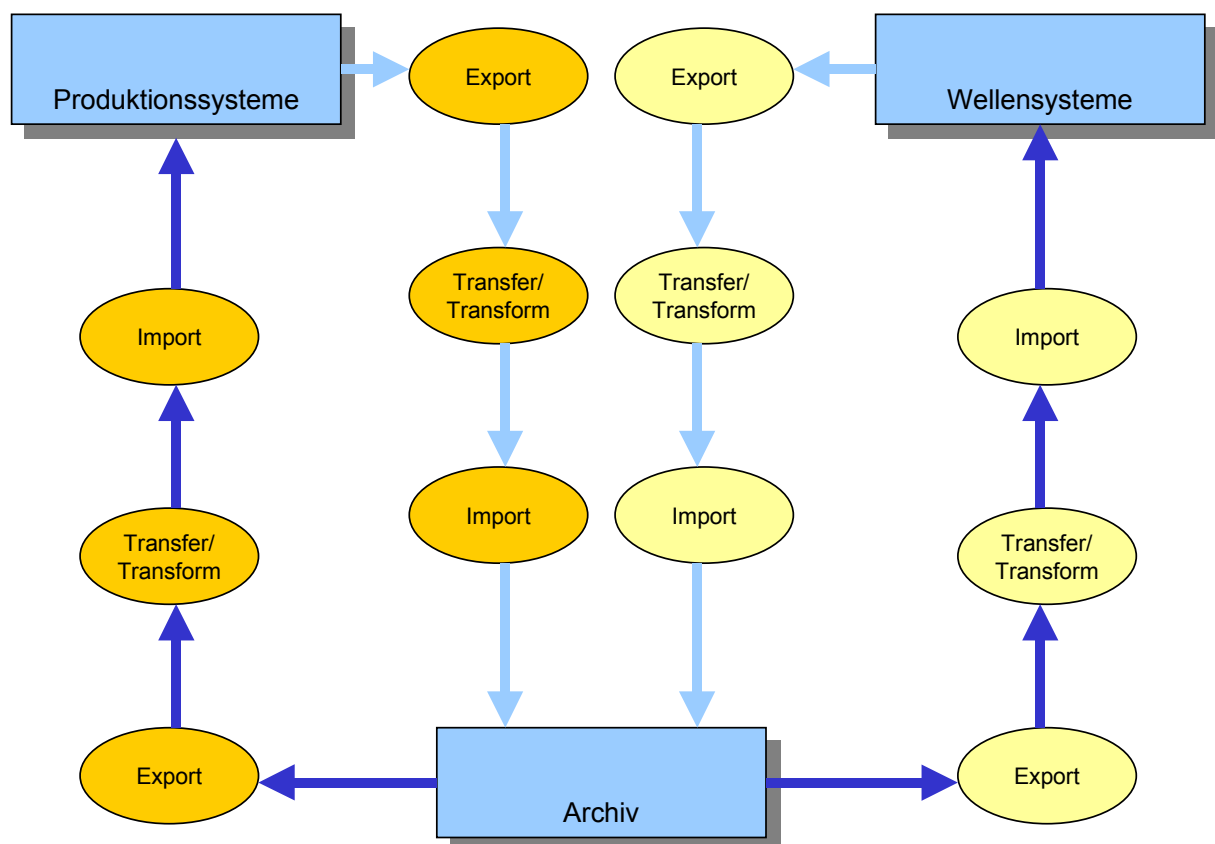
Die Anforderungen an ein *digitales Archiv im Broadcast-Umfeld* lassen sich schon konkreter formulieren. Die Dinge, um die es sich handelt, sind Dateien im digitalen Format, also Audio, Video, Bilder, Dokumente usw. oder allgemeiner Medienobjekte. Diese müssen, um bei der

Medien, wichtig ist jedoch, dass die Datensicherheit im Vordergrund steht. Das Thema Katalogisierung ist im Umfeld eines digitalen Archivs ein weites Gebiet. Es umfaßt die Erfassung, Indexierung und Aufbereitung von Metadaten, so dass über ein komfortable Suchoberfläche recherchiert werden kann. Über die Oberfläche ist auch der Zugriff auf die archivierten Medienobjekte möglich. Über die Standardfunktionalität hinaus gibt es in digitalen Archiven auch noch Funktionalität für den Umgang mit Speichersystemen, mit Medienobjekten und mit Metadaten.

Kommt als letzte Anforderung noch "universell" hinzu, so muß die Archivlösung so aufgebaut sein, dass sie nicht nur für einen konkreten Zweck einsetzbar ist, z.B. ein digitales Hörfunkarchiv mit einem festen Metadatensatz, sondern auch auf Fernsehen und alternative Metadaten anpaßbar sein, ohne dass sich die Kernfunktionalitäten des Archivs ändern.

Archiv: Standalone Lösung oder integriertes System?

Bei der Betrachtung, welchen Wert das Archiv für den Nutzer hat spielt neben der Basisarchivfunktionalität auch noch die Qualität der Anbindung eine Rolle, also die Integration mit der bestehenden Infrastruktur und den Prozessen. Als Standalone Lösung sehen die zentralen Import- und Exportvorgänge folgendermaßen aus:

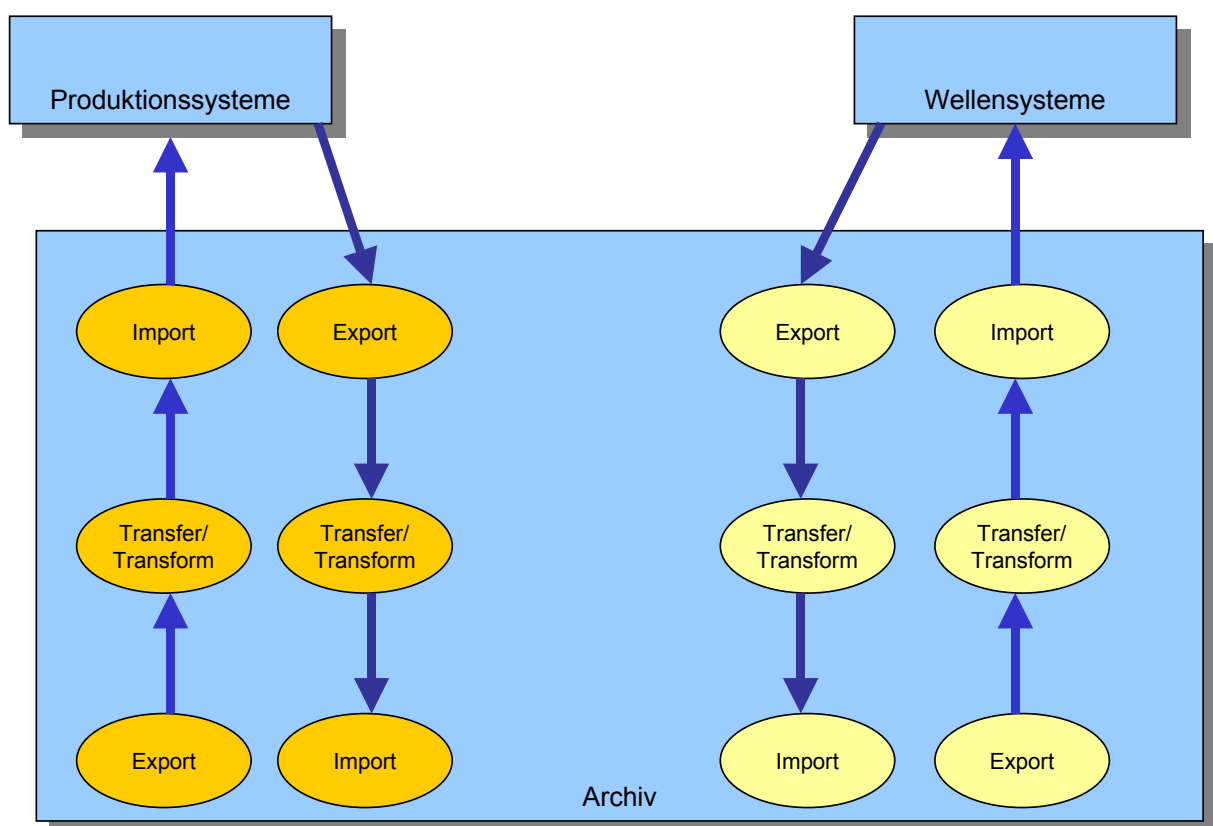


Das Archiv stellt Import- und Exportdienste für Medienobjekte und Metadaten zur Verfügung. Ebenso tun dies Produktions- und Wellensysteme. In den meisten Fällen passen allerdings

die Formate nicht zusammen und müssen transferiert und transformiert werden. Das alles findet als Archiv-Standalone-Lösung entweder manuell oder semiautomatisch statt. Wenn man bedenkt, dass diese Schritte bei jedem Import- und Exportvorgang durchlaufen werden müssen, so ist das für den Nutzer mit hohen Aufwänden verbunden.

Auf Basis dieser Betrachtung ist das Archiv anders als z.B. Produktions- oder Wellensystem zu charakterisieren, weil der Austausch von Daten mit anderen System integraler Bestandteil des Archivs ist. Dadurch ergeben sich andere Designanforderungen, die über die eigentliche Archivfunktionalität hinausgehen und einen großen Anteil am Gesamtsystem ausmachen.

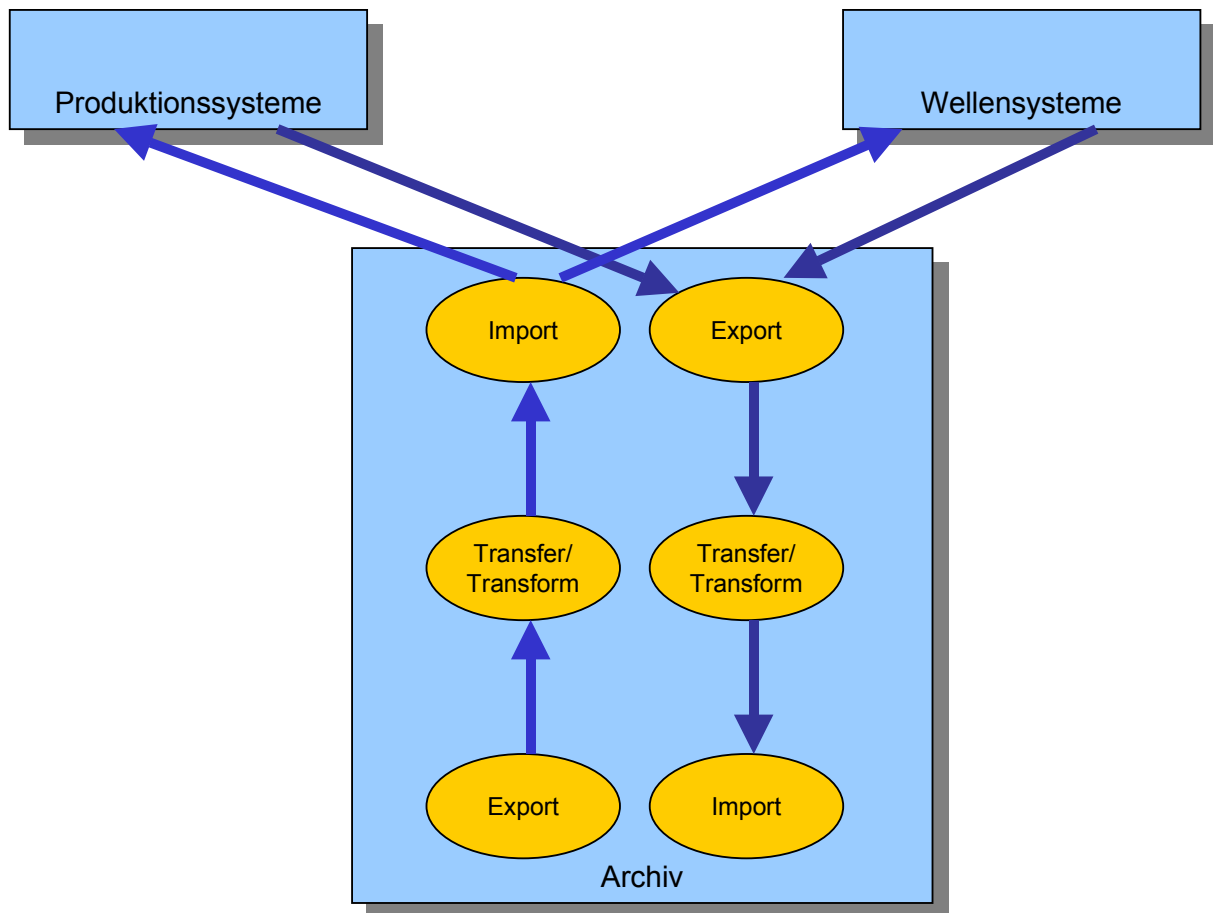
Daher könnte eine mögliche Verbesserung im Sinne der Integration folgendermaßen aussehen:



In dieser Architektur umfaßt das Archiv neben den Kernfunktionen auch die Integrationsfunktionen für die angeschlossenen Systeme. Für den Benutzer findet eine nahtlose Integration zwischen Produktions- und Wellensystem und Archiv statt, ohne das zusätzliche Arbeiten erforderlich sind. Das bedeutet für den Import vom Wellensystem in das Archiv, dass diese Aktion im Wellensystem angestoßen wird und einige Zeit später im Archiv liegt. Anders herum kann aus der Archivsuchoberfläche ein Beitrag für den Export aus dem Archiv ausgewählt werden und landet später in einem Speicher im Wellensystem.

Der so dargestellt Ansatz ist schon deutlich komfortabler, hat allerdings noch die Schwäche, dass bei vielen verschiedenen Systemen die Anzahl der Archivprozesse sehr groß werden

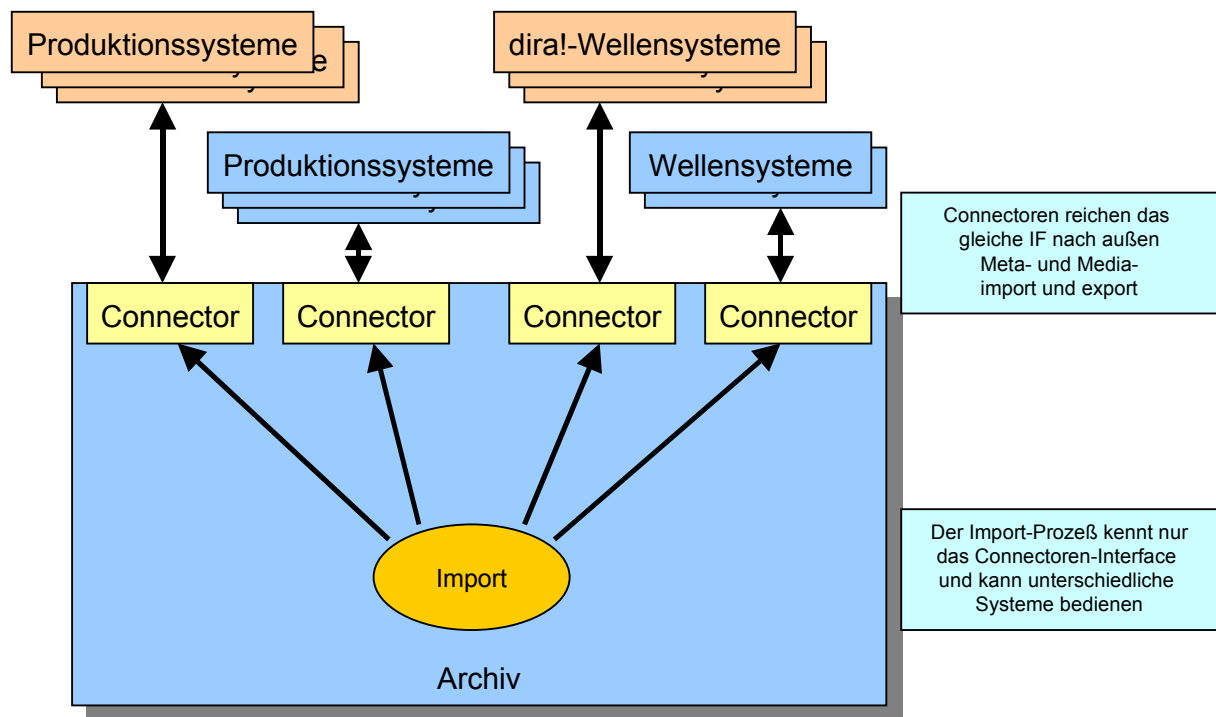
kann. Daher wäre eine weitere Vereinheitlichung aus Sicht der Wartbarkeit und der Komplexität hilfreich, der etwa folgender Idee entspricht:



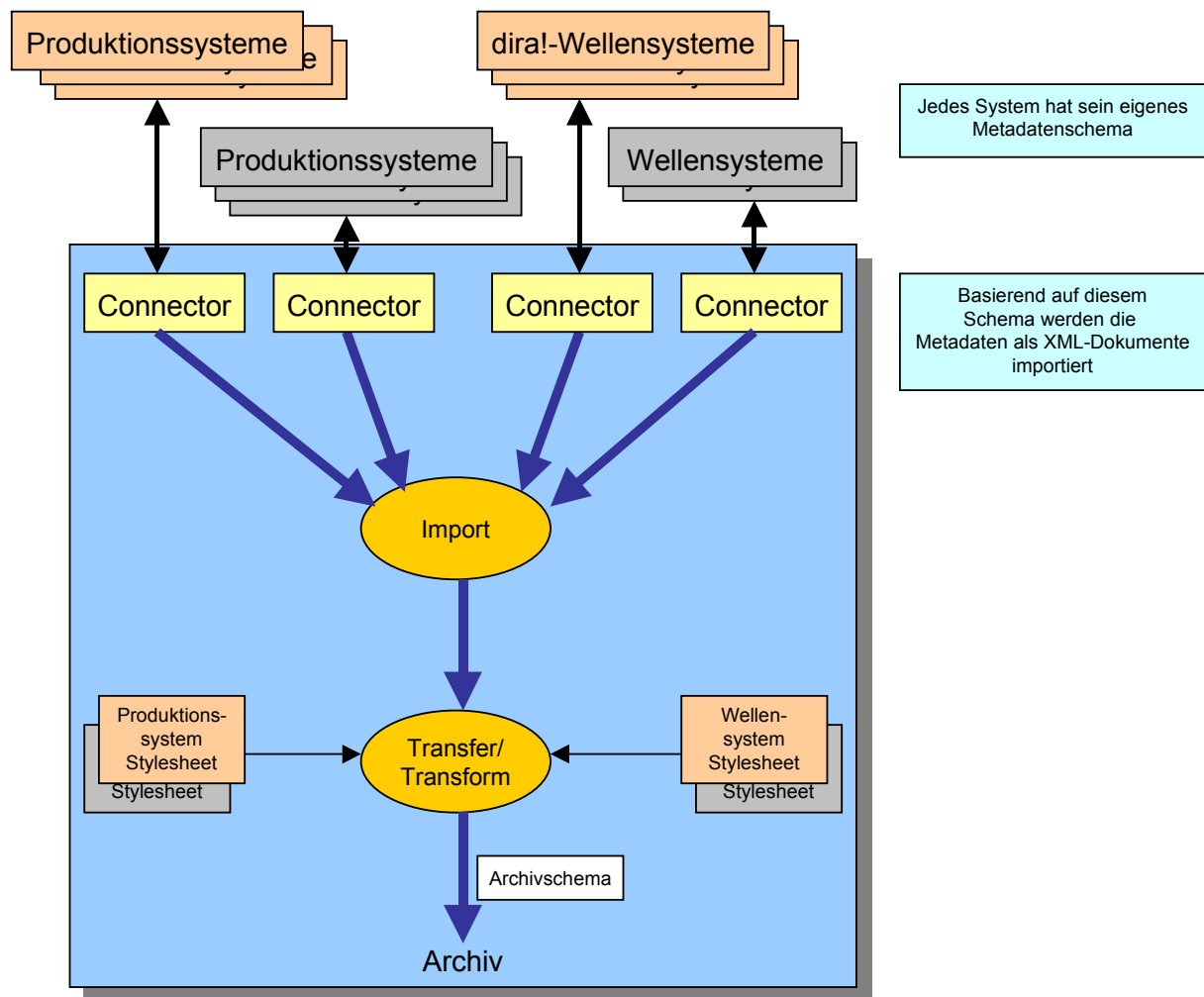
Das geht so natürlich nicht, weil die Systeme letztendlich anders aussehen bzw. funktionieren. Allerdings lassen sich die Komponenten in systemabhängige und systemunabhängige Elemente aufteilen und miteinander kombinieren.

Komponenten und Technologien des MediaFabric Frameworks

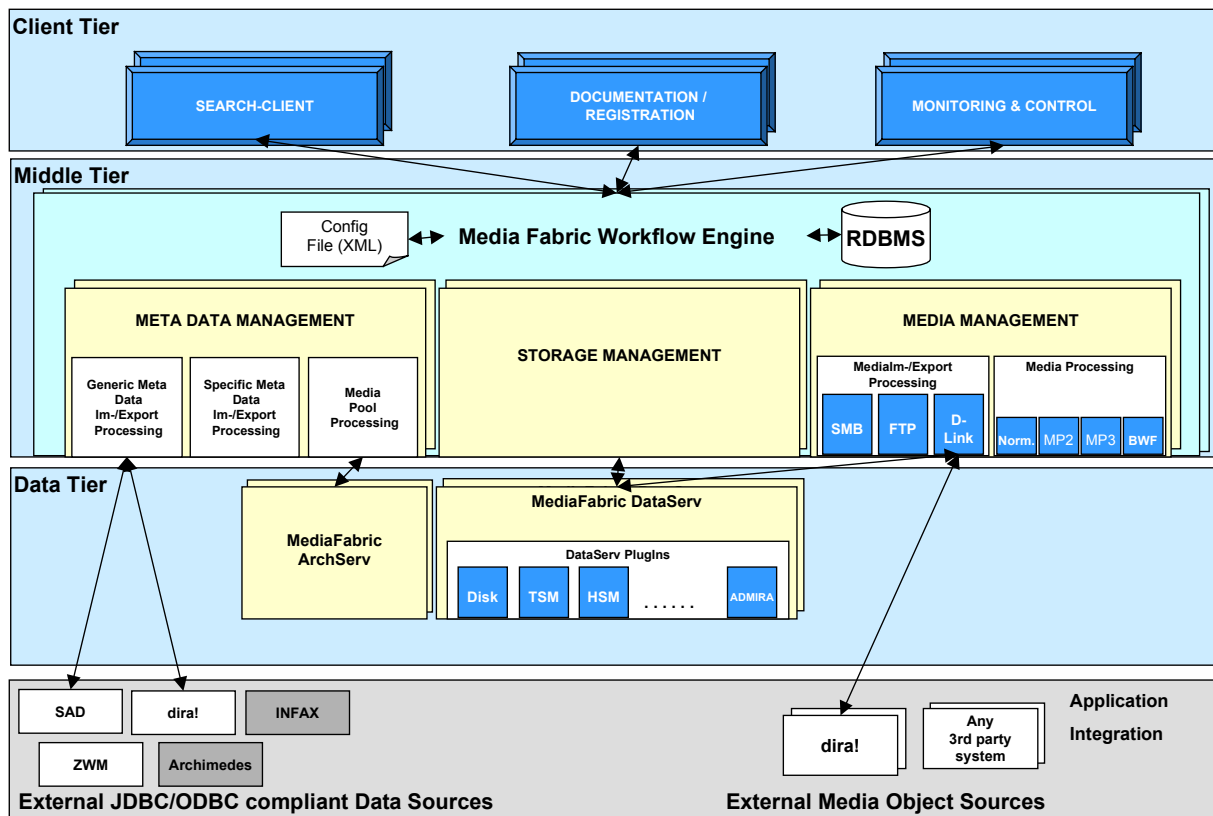
Aus den vorherigen Darstellungen ergibt sich das Design-Prinzip, dass es eine "dünne" Connectoren Schicht gibt, welche die proprietären Funktionen des externen Systems kapselt und einen systemunabhängigen Teil, welcher die tatsächliche Integration in das Archivsystem bewerkstelligt. Zur Veranschaulichung dieses Prinzips folgen zwei Beispiele. Das erste Beispiel zeigt den funktionalen Integrationsansatz:



Die Connectoren für die angeschlossenen Systeme reichen die gleiche Schnittstelle nach außen, z.B. für den Import und Export von Medienobjekten und Metadaten. Unabhängig, mit welchem Connector der Import Prozeß verbunden ist, ist er in der Lage den Datenaustausch zu steuern. Diese Art, Systeme zu integrieren, wird unter dem Begriff EAI (Enterprise Application Integration) zusammengefaßt. Selbst wenn die Schnittstellen nun technologisch in der Lage sind, Daten auszutauschen, gibt es noch ein weiteres Problemfeld, die Metadaten: Alle Systeme, sei es nun Produktionssystem, Wellensystem oder Archiv, haben ein eigenes Metadatenmodell, auf dem sie arbeiten. Damit die Schnittstellen wie beschrieben funktionieren, darf es keine Abhängigkeit vom Metadatenmodell geben. Dieses wird über XML-basierte Technologien umgesetzt.



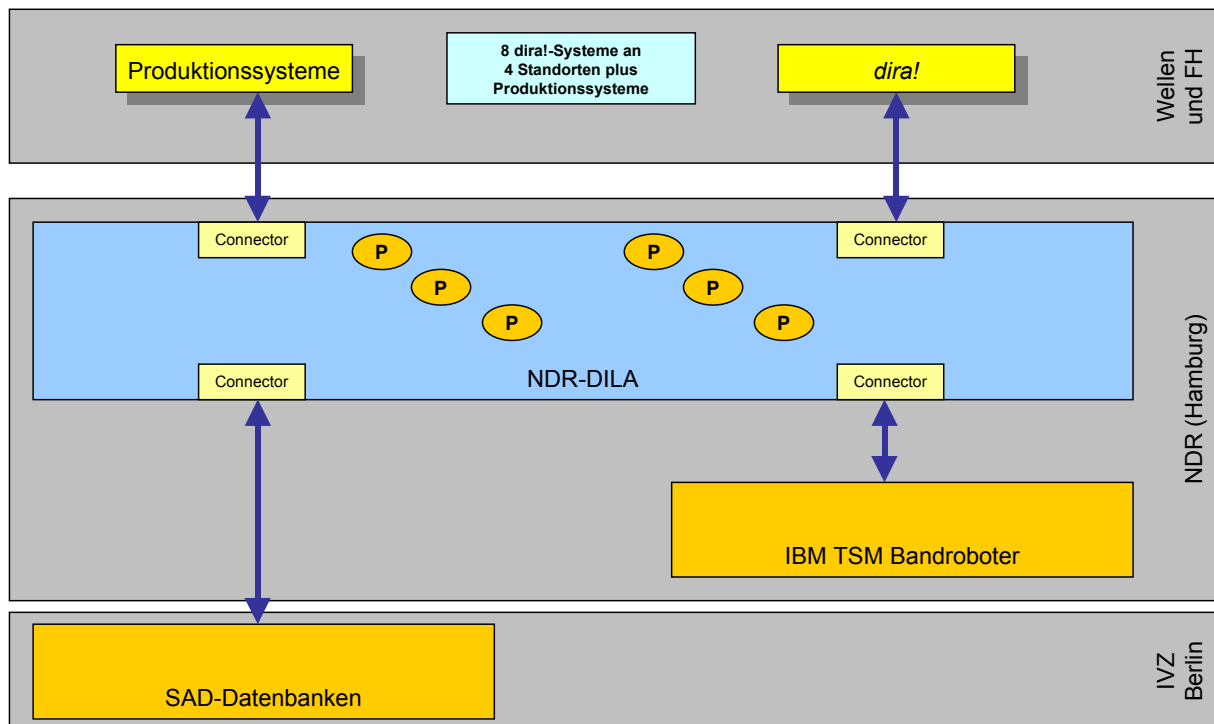
Mittels XML Schema und XML Stylesheet Transformationen läßt sich der Umgang mit Metadaten dahingehend vereinheitlichen, dass die zentralen Prozesse für den Export und den Import oder andere archivelevanten Prozesse gleich bleiben. Unter Einsatz dieser und anderer Technologien ist das Framework MediaFabric entstanden, welches technologisch den Rahmen für die Archivilösung setzt.



Die MediaFabric Archivlösung umfaßt drei Ebenen. Auf der Datenebene geht es um die Speicherung von Medienobjekten und Metadaten. Um unabhängig von dem zugrundeliegenden Medium (Tape, Jukebox, Disk, ..) zu sein, gibt es für die unterschiedlichen Systeme Plug-ins, so dass für die Integration das eingesetzte System transparent ist. Der größte Block des Archivs ist der MiddleTier. Dort findet sich ein Großteil der Komponenten wieder. Die Einsatzbereiche lassen sich sehr grob in drei Teilbereiche gliedern, in das Metadatenmanagement, das Mediamanagement und das Storamangement. Im Kontext dieser Bereiche findet auch die Integration der externen Systeme statt. Die Koordination der Komponenten wird durch eine Workflow-Engine realisiert, die ebenfalls der Bestandteil des Middle-Tiers ist. Auf dem Client-Tier finden sich dann die User-Interfaces für die Suche, die dokumentarische Erschließung und die Administration wieder. Auf Basis dieser Architektur wurden die Archive beim NDR und beim DeutschlandRadio realisiert, welche nun vorgestellt werden.

Beispiel 1 : NDR DILA

Beim NDR wurde ein Ansatz gewählt, in dem es ein zentrales Archivsystem für alle Wellen und Funkhäuser gibt, welches in Hamburg steht. Die Langzeitspeicherung findet auf einem Bandrobotersystem (IBM TSM/ADMIRA) statt. Als führende Metadatenbank ist SAD and das Archiv angebunden.



Die zentralisierte Architektur spiegelt sich auch in der Suchoberfläche wieder. Das Archiv hat eine zentrale Suchdatenbank für alle Bestände und Standorte.

DILA-Search | Suchmaske | Vollinfo | Drucken | Neue Pickliste | Beitragsübernahme aus DILA | ?

Sie haben im **Wortarchiv** (Alle Standorte) gesucht nach: Überall suchen="Schröder"

Seite 1 von 100 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | >> Vorwärts | 1000 Beiträge | Status: | Reimport:

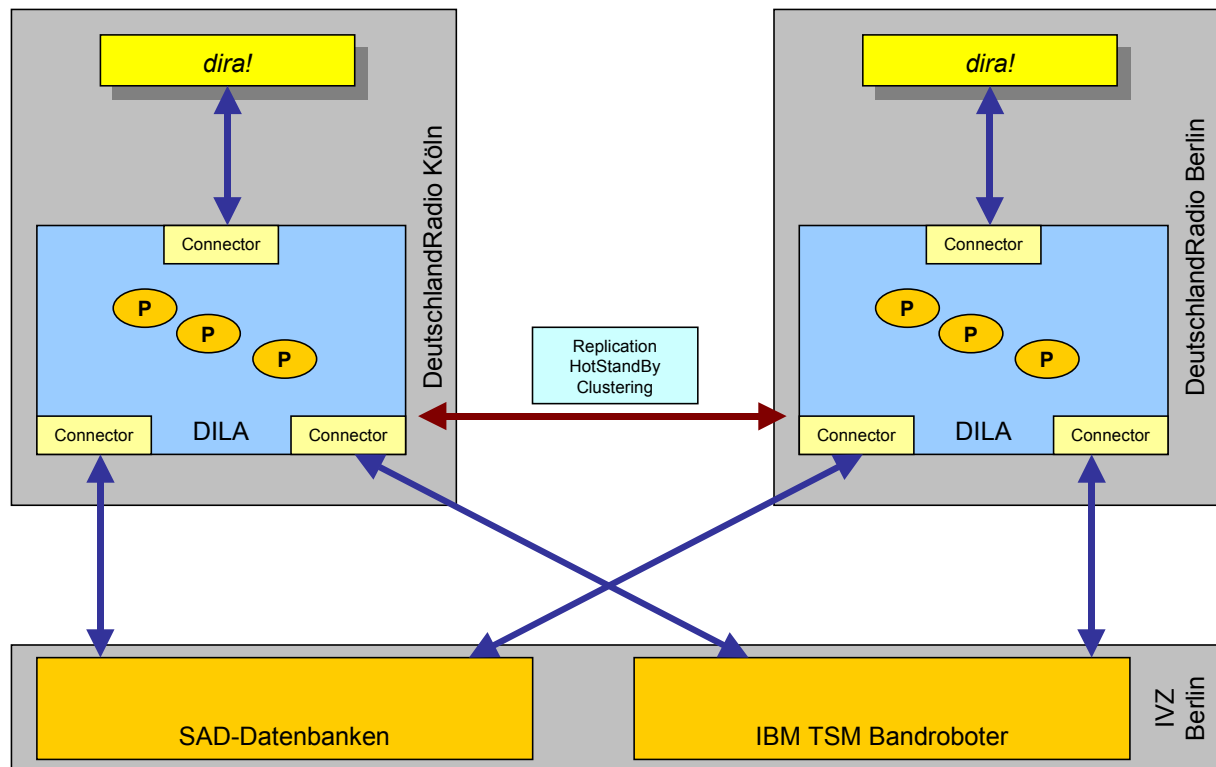
Keine Pickliste aktiv ⚠ Bitte evtl. Verwendungsbeschränkungen beachten!

Nr.	<input type="checkbox"/>	Titel	Dauer	Erstsendedat.	Tontr.-Art	Archivnr.	Standort
1	<input type="checkbox"/>	Museen im Wendland	8'28	20.02.2001	Audiofile	6000004 003	Hannover
2	<input type="checkbox"/>	Kommunalwahl: Bundespolitiker im niedersächsischen Wahlkampf	5'41	06.09.2001	Audiofile	6000063 002	Hannover
3	<input type="checkbox"/>	Bundeskanzler Gerhard Schröder besucht die CeBIT 2001	2'00	22.03.2001	Audiofile	6000101 000	Hannover
4	<input type="checkbox"/>	Eröffnung der Computermesse CeBIT 2001	2'35	22.03.2001	Audiofile	6000126 000	Hannover
5	<input type="checkbox"/>	Eröffnung der Hannover-Messe 2001	2'38	23.04.2001	Audiofile	6000128 000	Hannover
6	<input type="checkbox"/>	Einweihung des neuen Bundeskanzleramtes in Berlin	3'44	02.05.2001	Audiofile	6000142 000	Hannover
7	<input type="checkbox"/>	Zehn Jahre Deutsche Bundesstiftung Umwelt (DBU)	2'59	29.05.2001	Audiofile	6000143 000	Hannover
8	<input type="checkbox"/>	Einigung für das VW-Tarifmodell "5000 mal 5000" nahe	3'22	10.08.2001	Audiofile	6000194 001	Hannover
9	<input type="checkbox"/>	Bundeskanzler Gerhard Schröder zu Besuch in Oldenburg	1'07	10.08.2001	Audiofile	6000194 002	Hannover
10	<input type="checkbox"/>	VW-Modell 5000 x 5000 - Reaktion der Bundespolitiker	2'42	28.06.2001	Audiofile	6000199 004	Hannover

Beispiel 2 : DeutschlandRadio DILA

DeutschlandRadio hat in seinen Anforderungen einige Unterschiede zum NDR System. Zum einen wurde unter Havariegesichtspunkten das Archiv redundant ausgelegt. Beide Standorte

haben ein volles Archivsystem und zusätzliche Prozesse, welche den Abgleich zwischen den Standorten regeln.



DeutschlandRadio hat sich dazu entschieden, Speicherkapazität auf den Bandrobotern des IVZ zu reservieren anstatt sich ein eigenes Robotiksystem zuzulegen. Zwischen den Standorten und dem IVZ wird das Hybnet als Übertragungsmedium genutzt.

MediaPool

Suchtext: Schröder Erweiterte Suche Jobliste Pickliste Optionen Administration Hilfe

Suchergebnis Markierte Beiträge in Pickliste Pickliste Zimmermann transferieren Suchergebnisse drucken

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
		Sendedat.	Aufnahmed.	Länge	Titel	Person	ANR, Prod.Nr.	Quelle	Speicher			
<input type="checkbox"/>	<input type="checkbox"/>	25.07.2004	22.07.2004	50:43	Variationen mit Thema (Teil 1) - 25.07.2004-	Felber, Gerald	Z004553 000, DZ004553 000	DLRB	Wortarch			
<input type="checkbox"/>	<input type="checkbox"/>	05.07.2004	30.06.2004	4:49	Politisches Feuilleton: Deutsche Reformfähigkeit	Gauland, Alexander	Z007857 000, DZ007857 000	DLRB	Wortarch			
<input type="checkbox"/>	<input type="checkbox"/>	03.07.2004	29.06.2004	5:10	Politisches Feuilleton: Deutschland und Frankreich: Sündigen mit großer Geste	Stürmer, Michael	Z007725 000, DZ007725 000	DLRB	Wortarch			
<input type="checkbox"/>	<input type="checkbox"/>	26.06.2004	26.06.2004	3:08	Schlaglicht: Rücktritt Völler, Haushaltsplanung 2005, Finanzlage des Staates	Wassmund, Norbert	Z015045 000, DZ015045 000	DLRB	Wortarch			
<input type="checkbox"/>	<input type="checkbox"/>	28.06.2004	25.06.2004	15:03	LänderReport-L: Havel-Poker	Schibilsky, Maren	Z009402 000, DZ009402 000	DLRB	Wortarch			
<input type="checkbox"/>	<input type="checkbox"/>	15.07.2004	21.06.2004	8:21	Kostprobe: "Flattergeist"	Tschechow, Anton	Z005149 000, DZ005149 000	DLRB	Wortarch			
<input type="checkbox"/>	<input type="checkbox"/>	19.06.2004	18.06.2004	0:15	Schlaglicht 2: Gerhard Schröder zur Agenda 2010 - nicht immer der richtige Weg	Gerhard Schröder	Z014841 000, DZ014841 000	DLRB	Wortarch			
<input type="checkbox"/>	<input type="checkbox"/>	18.06.2004	17.06.2004	4:55	Politisches Feuilleton: Küsst die Hand, Kanzler! Anmerkungen zum Protokoll	Grawert-May, Erik von	Z007792 000, DZ007792 000	DLRB	Wortarch			
<input type="checkbox"/>	<input type="checkbox"/>	17.06.2004	16.06.2004	4:06	LänderReport-K: ein Dorf wählt braun (O-Ton)	Griess, Hanno	Z010416 000, DZ010416 000	DLRB	Wortarch			
<input type="checkbox"/>	<input type="checkbox"/>	16.06.2004	16.06.2004	2:58	Abschluß G-8-Gipfel	Schmiester, Carsten AUT	Z000210 000, DZ000210 000	DLFK	Wortarch			
<input type="checkbox"/>	<input type="checkbox"/>	16.06.2004	15.06.2004	4:51	Politisches Feuilleton: Liebe ich mein Land? (1)		Z007839 000, DZ007839 000	DLRB	Wortarch			
<input type="checkbox"/>	<input type="checkbox"/>	14.06.2004	11.06.2004	5:56	Kalenderblatt 14.06. (1954 - Kuratorium unteilbares Deutschland)	Weil, Marianne	Z002311 000, DZ002311 000	DLRB	Wortarch			

Im Gegensatz zum NDR System unterstützt die DeutschlandRadio Suchoberfläche auch die verteilte Suche über Wellensysteme, SAD-Bestände und Archivinhalte.

Erweiterungen, Ausblicke

Integriertes Systemmanagement

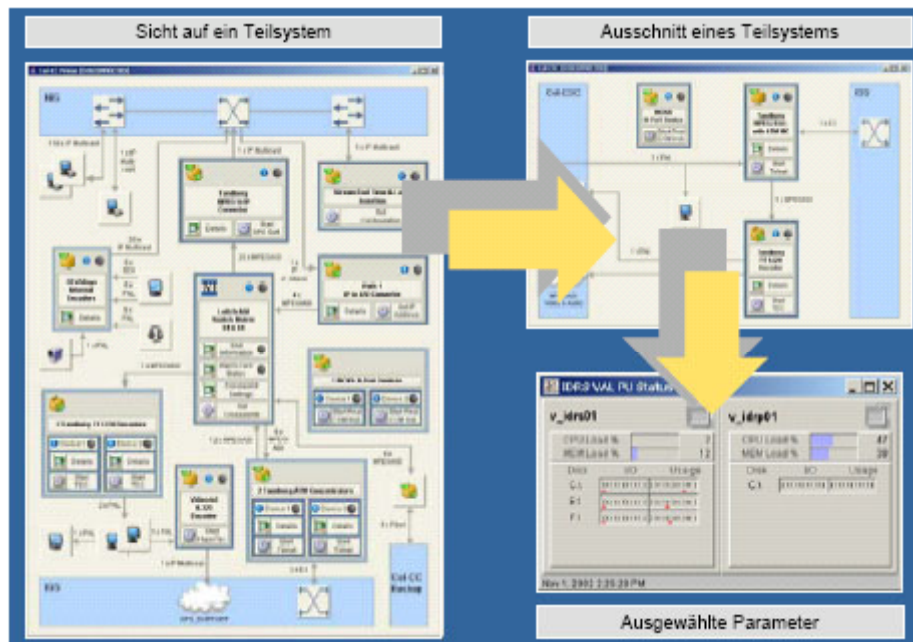
Die dargestellten Lösungen haben, wie wir schon festgestellt haben, folgende Beschaffenheit:

- Bestehen aus vielen Komponenten, die sich auf verschiedene Systeme, auch standortübergreifend, verteilen
- Hoher Integrationsgrad mit existierenden Systemen (Wellen-, Produktionssystemen)
- Ein Gutteil der Funktionalität liegt auf dem Server bzw. in der Middleware

Daraus ergeben sich im Betrieb folgende Problemstellungen.

- Den Überblick in dem komplexen System zu bewahren ist schwierig, insbesondere das Auffinden von möglichen Fehlerquellen (Platten, Netzwerk, Rechner) ist eine langwierige Angelegenheit.
- Die Abläufe im Archiv bedingen durch den hohen Integrationsgrad auch, dass angeschlossene Systeme verfügbar sein müssen, z.B. Dokumentationsdatenbanken, Bandrotbotik. Probleme im Betrieb müssen nicht zwingend aus Archivkomponenten selbst resultieren, sondern aus externen Faktoren oder dem Zusammenspiel aus beiden.

Daher ist gerade für ein Archivsystem ein ausgereiftes, integriertes Systemmanagement eine sinnvolle und notwendige Ergänzung, um schnell und effizient Probleme zu erkennen und zu beheben.



Hier ist ein Beispiel für ein integriertes Systemmanagement, wie es z.B. im Raumfahrtumfeld zum Einsatz kommt (VCS Framework egmc²). Das Systemmanagement umfaßt die Überwachung von Geräten, Hardware und Software und stellt sie in übersichtlicher Form dar. Meldungen aus den Systemen werden erfaßt und dargestellt. Die für die Administration des Systems (Konfiguration, Starten und Stoppen von Komponenten) notwendigen Funktionen sind über die Oberfläche erreichbar. Ein detaillierterer Vortrag zu diesem Thema findet am Samstag morgen unter dem Titel **"Von anderen lernen: Wirtschaftliche Betriebskonzepte durch integriertes Systemmanagement"** (25062) statt

Fazit

Durch den hohen Anteil an Integration in Verbindung mit der archiveigenen Komplexität ist eigentlich jedes Archivprojekt im Medioumfeld eine ehrgeizige Herausforderung. Moderne Integrationstechnologien wie J2EE, .NET, XML und EAI (nicht vollständige Aufzählung) erlauben es heute, Systeme zu kreieren, die vor ein paar Jahren nicht oder nur mit überproportionalem Aufwand zu realisieren waren. Vollständig integrierte Lösungen, die flexibel auf sich ändernde Rahmenbedingungen angepaßt werden können, sind die Anforderungen an die heute zu realisierenden Systeme.

Das soll aber nicht darüber hinweg täuschen, dass nach wie vor noch komplexe Aufgabenstellungen zu lösen sind und dieses lediglich durch die aktuellen Technologien in den Bereich des wirtschaftlich Machbaren rückt.

Kontakt: VCS Aktiengesellschaft
Michael Zimmermann
Solutions Manager
Borgmannstr. 2
44894 Bochum
Tel.: +49 234 9258-213
Fax: +49 234 9258-190
E-Mail: Michael.Zimmermann@vcs.de