

Inhaltsverzeichnis

1	Vorwort	9
2	Einleitung.....	11
3	Langzeitarchivierung	19
3.1	Herausforderungen.....	19
3.1.1	Hardware- & Datei-Korruption	21
3.1.2	Soft- und Hardwareobsoleszenz	22
3.2	Status Quo der Langzeitarchivierungs-Forschung.....	26
3.2.1	Bitstream Preservation	26
3.2.2	Integrität von digitalen Daten	27
3.2.3	Emulation	29
3.2.4	Migration.....	30
3.2.5	Significant Properties	32
3.2.6	Metadaten	36
3.2.7	Identifikation, Charakterisierung und Validierung	40
3.2.8	File Format Registries.....	42
3.2.9	Technology Watch	44
3.2.10	Eindeutige und persistente Identifier.....	46
3.2.11	OAIS	48
4	Das DA-NRW Projekt und seine Zielstellung	55
4.1	Ausgangssituation: Herausforderungen aus der Praxis	56
4.1.1	Born Digitals	58
4.1.2	(Retro-)Digitalisate	60
4.1.3	Objektdefinition	61
4.1.4	Fazit	63
4.1.5	Wünsche der Einlieferer	63
4.2	Architektonische Überlegungen im DA-NRW	66
4.2.1	Verteilte Lokationen & redundante Speicherung.....	67
4.2.2	Das DA-NRW-Objektmodell	68
4.2.3	Integritätsprüfung der Langzeitspeicher – Bitstream Preservation.....	73
4.2.4	Formatidentifikation & -migration	74
4.2.5	Contracts	76
4.3	Logische Architektur des DA-NRW	77

4.3.1	Datenmanagement im DA-NRW	80
4.3.2	ContentBroker	82
4.3.3	Presentation Repository	83
5	Das DA-NRW in der Praxis.....	85
5.1	Datenaufbereitung – (Pre-)Ingest.....	85
5.1.1	Der DA-NRW SIP-Builder.....	86
5.2	Benutzerschnittstellen im DA-NRW	91
5.2.1	Die Webanwendung „DA-Web“ im Überblick.....	94
5.2.2	DA-Web und WebDAV als Ausgangspunkt der Use Cases Ingest und Retrieval	95
5.2.3	Potentiale der Webanwendung und Ausblick	97
5.3	Paketverarbeitung.....	98
5.3.1	Grundlegende Architektur	99
5.3.2	Formatidentifikation und -konvertierung	106
5.3.3	Formatidentifikation	106
5.3.4	Konvertierungsroutinen	109
5.4	Datenhaltung im DA-NRW in der Praxis.....	121
5.4.1	Anbindung der heterogenen LZA-Speichermedien	121
5.4.2	Interaktion von ContentBroker zu iRODS	125
5.4.3	Logische Verwaltung der Objekte auf den Ressourcen	126
5.4.4	Technische Metadaten im Langzeitspeicher	127
5.4.5	Bitstream Preservation	128
5.4.6	Aufbau eigener „Zonen“ und „Federation“	133
5.4.7	Sicherheitsaspekte der Verwendung von iRODS.....	138
5.4.8	Ausblick	141
5.5	Präsentation	143
5.5.1	Publikation von Archivmaterial.....	143
5.5.2	Aufbereitung von Metadaten für die Publikation.....	147
5.5.3	Repositorysoftware & Präsentationsschnittstellen	154
5.6	Spezifikationen.....	158
5.6.1	Pakettypen.....	158
5.6.2	Original Name, Containerformate und Inhalte.....	159
5.6.3	SIP-Spezifikation im DA-NRW	159
5.6.4	AIP-Spezifikation im DA-NRW	162
5.6.5	Deltas.....	165
5.6.6	DIP-Spezifikation im DA-NRW	168
5.6.7	Identifier im DA-NRW	171
5.6.8	Verstandene und unterstützte Primärdatenformate	175
5.6.9	Unterstützte Metadatenformate	177

6	Ausblick	183
6.1	PreIngest.....	186
6.2	Paketverarbeitung.....	187
6.3	Ausblick iRODS.....	188
6.4	Ausblick „DA-Web“.....	189
6.5	Ausblick SIP-Builder.....	190
6.6	Ausblick Präsentation.....	190
7	Anhänge	191
7.1	Lizenzen	191
7.2	Use Cases.....	195
7.3	Externe Dokumentationen	195
7.4	Eine vom SIP-Builder generierte PREMIS-Datei	196
7.5	Die JSON-LD-Repräsentation eines Objekts in der Suchschnittstelle	200
8	Glossar	203
9	Literaturverzeichnis	211