



**Startpunt
Migratieonderzoek**

Uitgave: 1.1
Auteur: Caroline van Wijk
Datum: 5 juli 2006

Koninklijke Bibliotheek
Afdeling Digitale Duurzaamheid

Startpunt Migratieonderzoek

Ontstaansgeschiedenis Startpunt Migratieonderzoek

Versie Document	Datum wijziging	Auteur	Samenvatting wijzigingen
0.1 Concept	5 april	Caroline van Wijk	
1.0	20 april	Caroline van Wijk	Opmerkingen van H. van Wijngaarden en B. Sierman verwerkt.
1.1	5 juli	Caroline van Wijk	Opmerkingen uit PLANETS discussie verwerkt: Onderscheid gemaakt tussen werkwijzen migratie en projecten die zijdelings met migratie te maken hebben.

Aanverwante documenten

Project Initiatie Document – Migratieonderzoek, C. van Wijk (2006)

Startpunt Migratieonderzoek

Inhoudsopgave

1	Inleiding	4
1.1	Doel van dit document	4
1.2	Indeling van dit document	4
	<i>Overzicht instellingen en migratieonderzoek.....</i>	<i>4</i>
	<i>Overzicht werkwijzen migratie.....</i>	<i>4</i>
	<i>Voorlopige evaluatie</i>	<i>4</i>
2	Overzicht instellingen en migratieonderzoek	5
2.1	Inleiding	5
2.2	Overzichten	6
	<i>Praktijkgericht</i>	<i>6</i>
	<i>Theoriegericht.....</i>	<i>10</i>
3	Overzicht werkwijzen migratie.....	14
3.1	Inleiding	14
3.2	Overzicht	15
3.2	Overview projects related to migration	22
4	Voorlopige evaluatie.....	23
	Evaluatie	23
	Vervolg	23
5	Literatuurlijst	24
6	Termenlijst (voorlopig).....	26

Startpunt Migratieonderzoek

1 Inleiding

1.1 Doel van dit document

In dit document wordt het startpunt voor het migratieonderzoek bij de Koninklijke Bibliotheek beschreven. Dit startpunt bestaat uit een overzicht van internationale projecten, waarin migratieonderzoek centraal staat, en een overzicht van de verschillende typen migratie die in internationale literatuur beschreven zijn.

De overzichten in dit document zijn gebaseerd op informatie die via Internet beschikbaar was. Het is mogelijk dat de overzichten onvolledig zijn. Opmerkingen en toevoegingen zijn welkom en kunnen naar de auteur verstuurd worden (caroline.vanwijk@kb.nl).

1.2 Indeling van dit document

Overzicht instellingen en migratieonderzoek

In het overzicht van uitgevoerde projecten wordt een onderscheid gemaakt tussen projecten met een praktische inslag en projecten die een theoretische inslag hebben. De projecten worden beschreven met een korte beschrijving van de werkzaamheden, de resultaten, welke instelling het project heeft uitgevoerd en wanneer het project gedaan is.

Overzicht werkwijzen migratie

Het overzicht van 'werkwijzen migratie' geeft weer welke typen migratie (conversie) op het gebied van digitale duurzaamheid zijn beschreven. In het overzicht komen werkwijze, beschrijving en een opsomming van voor- en nadelen voor.

Voorlopige evaluatie

In het kort wordt een voorlopige evaluatie gegeven van de in dit document behandelde informatie.

2 Overzicht instellingen en migratieonderzoek

2.1 Inleiding

In dit hoofdstuk worden theoretische en/of praktische projecten wat betreft migratie van bestandsformaten in het kader van digitale duurzaamheid ingedeeld en beschreven in twee groepen. De eerste groep wordt gevormd door instellingen die een migratieproject met een praktische invulling hebben uitgevoerd. De tweede groep bestaat uit instellingen die een theoretisch onderzoek hebben uitgevoerd wat betreft migratie, maar (nog) geen praktische tests hieraan hebben verbonden.

Startpunt Migratieonderzoek

2.2 Overzichten

Praktijkgericht

In het onderstaande overzicht worden de projecten en uitvoerende instellingen genoemd waarvan een concrete test met (verschillende soorten) migratie onderdeel van het project uitmaakt.

Naam instelling en contactpersoon	Project	Datum	Activiteiten	Onderzoeksresultaat
Public Record Office Victoria (Australia)	VERS (Victorian Electronic Records Strategy)	1998 1e versie, 2002 2e versie	Ontwikkelen van tool om een wrapper van XML te maken om het archiefbestand (PDF) heen.	Wordt toegepast.
Nationaal Archief - Remco Verdegem	Testbed	2000-2003	Testen van migratie en conversie (normalisatie) van tekstdocumenten	<ul style="list-style-type: none">• Achterwaartse compatibiliteit niet geschikt voor lange termijn opslag en toegankelijkheid• Conversie van tekstdocument naar XML beste strategie als documentopmaak expliciet is. (gebruik gemaakt van templates voor de opmaak van het document)• Conversie van tekstdocument naar PDF beste strategie als documentopmaak

Startpunt Migratieonderzoek

Naam instelling en contactpersoon	Project	Datum	Activiteiten	Onderzoeksresultaat
				impliciet is. (geen gebruik gemaakt van templates)
			Test migratie databases	Conversie van database naar XML beste strategie voor databases
			Test migratie spreadsheets	Conversie van spreadsheet naar XML beste strategie voor spreadsheets
			Onderverdelen van 'migratie' in: <ul style="list-style-type: none"> • achterwaartse compatibiliteit • interoperabiliteit • conversie naar standaardformaten 	
University of Leeds- Phil Mellor, Paul Wheatly, Derek Sergeant	CAMiLEON	2002 (6 th ECDL proceedings)	Testen van 'Migration on request' (pas een conversie bij aanvraag van het bestand) en 'reversible migration' (converteren van het geconverteerde bestand naar oorspronkelijk formaat). Er is getest met vector graphics en een zelf ontwikkelde tool.	De modulaire migration on request tool heeft succesvol vectorbestanden geconverteerd. Reversible migration was niet zo makkelijk als het in theorie leek te zijn.

Startpunt Migratieonderzoek

Naam instelling en contactpersoon	Project	Datum	Activiteiten	Onderzoeksresultaat
Swiss Federal Archives – Stephan Heuscher, Stephan Järmann, Peter Keller-Marxer, Frank Möhle	SIARD: Software-invariant Archiving of Relational Databases	2002 (eerste publicaties)	<p>Uitvoeren van een experiment met het omzetten van relationele databases naar:</p> <ul style="list-style-type: none">• SQL 3 (datalogic)• Platte tekstbestanden (de inhoud van de tabellen)• XML voor de beschrijving van de context <p>Vervolgens deze gegevens weer inlezen in een relationele database via een zelf ontwikkelde reload tool</p>	De databases zijn succesvol omgezet naar SQL 3, platte tekst en XML. Het inlezen van deze gegevens in een relationele database was succesvol.
National Archives Australia – Andrew Wilson (is contactpersoon van de KB, niet specifiek voor dit project)		2002	<p>Er is een prototype gebouwd voor normalisatie (de normaliser en de viewer). Normalisatie (naar XML) en XML encapsulation van het origineel met behulp van de tool Xena.</p> <p>http://xena.sourceforge.net</p>	De tool wordt gebruikt vóór het ingest gedeelte van het archief.

Startpunt Migratieonderzoek

Naam instelling en contactpersoon	Project	Datum	Activiteiten	Onderzoeksresultaat
			t/download.html	
Florida Center for Library Automation / DAITSS – Priscilla Caplan		2003	Normalisatie van de objecten wordt uitgevoerd voor opname in DAITSS. PDF 1.x wordt XML en TIFF-bestanden. DTD's en XML-bestanden die verwijzen naar externe bestanden worden naar DTD_NORM_1 en XML_NORM_1 geconverteerd.	<p>Wordt toegepast.</p> <p>Per 16 maart 2006 zijn er 73,383 bestanden opgeslagen (1.3 TB).</p> <ul style="list-style-type: none"> • Gearchiveerd materiaal bestaat uit: Florida Heritage and West Florida Photohistory materiaal van UWF • Electronische dissertaties en proefschriften (ETDs) van UCF
Lister Hill National Center for Biomedical Communications / National Library of Medicine Bethesda, Maryland – Frank L. Walker, George R. Thoma		2004	Tool die verschillende type bestanden (BMP, TIFF, JPEG etc.) omzet naar PDF/A (MyMorph)	Bewaren van beeldmateriaal (en ook tekstverwerkerdocumenten) in de PDF/A standaard
Stanford University Libraries – David S. H. Rosenthal, Thomas lipkis, Seth Morabito, Thomas Robertson	LOCKSS	2004	Proof of concept implementeren van 'migration on access': conversie van gif bestanden in een website	Het project is geslaagd in het on-the-fly converteren van GIF-bestanden in een website. De conversie werd zodanig uitgevoerd, dat de aanvrager van de website deze niet heeft opgemerkt.

Startpunt Migratieonderzoek

Naam instelling en contactpersoon	Project	Datum	Activiteiten	Onderzoeksresultaat
			naar PNG-formaat bij aanvraag van de website. Voor de proof of concept is het conversiemechanisme ingebouwd in het LOCKSS systeem.	
Koninklijke Bibliotheek / IBM – Hilde van Wijngaarden	Universal Virtual Computer voor afbeeldingen	2004	Implementatie van UVC concept (R. Lorie). Ontwikkelen van een: <ul style="list-style-type: none">• UVC• JPEG decoder, LDS• GIF decoder, LDS• Viewer	De UVC voor afbeeldingen is ontwikkeld als een “vangnet tool” voor de publicaties in het e-Depot. De publicaties zijn voor het grootste gedeelte PDFs. Dit bestandsformaat is zeer complex. Het ontwikkelen van een LDS en PDF decoder kost veel tijd. Om het bereik van het UVC project haalbaar te houden is gekozen om een UVC voor afbeeldingen te maken. De PDF uit het e-Depot zal omgezet worden in JPEG formaat op het moment dat de UVC gebruikt zal worden.

Theoriegericht

In het onderstaande overzicht worden de projecten en uitvoerende instellingen genoemd die een theoretische grondslag voor migratie hebben ontwikkeld welke voor het migratieonderzoek van de KB van belang zou kunnen zijn.

Startpunt Migratieonderzoek

Naam instelling en contactpersoon	Project	Datum	Onderwerp	Opmerking
CEDARS: Consortium of University Research Libraries University of Leeds- Paul Wheatly, Derek Sergeant	CEDARS	1998 - 2001	Migration on request: conversie van een bestand wordt pas toegepast op het moment dat het bestand wordt aangevraagd.	In het CAMiLEON project is een praktische test uitgevoerd met "migration on request".
Council on Library and Information Resources – Gregory W. Lawrence, William R. Kehoe, Oya Y Rieger, William H. Walters, Anne R. Kenney		2000	Risk management: checklist voor conversies	Bruikbaar bij opzet test?
University of Pennsylvania - John Ockerbloom	Typed Object Model	2000-2001	Formalisering van conversies, de eigenschappen van een bestandsformaat. Respectful Type Converters Links niet meer bereikbaar!	Te formeel om direct toe te passen
University of Pennsylvania – John Ockerbloom		2001	Archiveren en conserveren van PDF in onderdelen: <ul style="list-style-type: none"> • Statische afbeeldingen van een 	Tools die genoemd worden zijn: Ghostscript voor conversie naar afbeelding, Pstotext of Prescript voor een conversie naar tekst.

Startpunt Migratieonderzoek

Naam instelling en contactpersoon	Project	Datum	Onderwerp	Opmerking
			bladzijde in een PDF <ul style="list-style-type: none">• Tekst in de PDF• Structuur (table of content) in de PDF• Dynamische onderdelen	
University of Leeds - Paul Wheatley	CAMiLEON	2001	Typering in niveaus van migratie	Typering bruikbaar bij opzet test eventueel?
Smithsonian institution archives - Dollar Consulting		2001	Aanbeveling voor digitale duurzaamheidstrategie: HTML naar XHTML, dus normalisatie	
Department of Software Technology and Interactive Systems, Vienna University of Technology - Carl Rauch, Andreas Rauber	DELOS- Testbed en Utility Analysis	2004	<ul style="list-style-type: none">• Ontwerpen van een raamwerk voor het kiezen en testen van digitale duurzaamheid strategieën.• Ontwikkelen van een testbed applicatie waarin verschillende digitale bestanden, met beschrijving van	Het project heeft een aantal case studies gedaan waarin het raamwerk is toegepast op audio- en videobestanden en tekstdocumenten en richt zich (voorlopig) alleen op de strategie migratie. Er zijn geen migratietrajecten uitgevoerd.

Startpunt Migratieonderzoek

Naam instelling en contactpersoon	Project	Datum	Onderwerp	Opmerking
			de eigenschappen van het bestand, worden opgeslagen.	
DAVID – Filip Boudrez		2005	Richtlijnen migratie naar archiveringsformaten (normalisatie)	
University of Minho, Portugal - Miguel Ferreira (promovendus)		2005 (ECDL)	Voorstel voor een pilot project/experiment. Onderwerp is Automatic Evaluation of Migration Quality in Distributed Networks of Converters	

3 Overzicht werkwijzen migratie

3.1 Inleiding

In de literatuur over “migratie” en “conversie” worden deze termen vaak door elkaar gebruikt voor hetzelfde proces. In het migratieonderzoek van de Koninklijke Bibliotheek worden “migratie” en “conversie” ook gebruikt om hetzelfde proces van het veranderen van een digitaal object aan te duiden.

De werkwijzen van migratie die in dit hoofdstuk beschreven worden, vallen onder wat in het OAIS Reference Model ¹ ‘transformatie’ genoemd wordt. Transformatie is in dit model gedefinieerd als een migratie van een digitaal object waarbij de Content Informatie (PDI bits) wordt veranderd om de gehele informatie content van het object te bewaren. Deze omschrijving past in de definitie voor de digitale duurzaamheid strategie migratie: het omzetten van een digitaal object in een bepaalde software en hardwareomgeving naar een andere software en hardwareomgeving om het digitale object toegankelijk te maken. In het OAIS Reference Model worden vier typen van migratie beschreven: verversen (refreshing), dupliceren (replication), opnieuw verpakken (repackaging) en dus transformatie (transformation). Alleen bij transformatie is sprake van het wijzigen van de Content Informatie. Het migratieonderzoek van de KB zal zich niet richten op de overige typen migratie (zoals bijvoorbeeld een verandering van drager). In het migratieonderzoek zullen de termen “migratie” en “conversie” worden gebruikt en niet de term “transformatie”.

Zoals in de definitie van transformatie naar voren komt wordt bij migratie het originele digitale object aangepast om de toegankelijkheid van het object te vergroten. In principe houdt dit een kans in op fouten die tijdens een migratie ontstaan, of het vervallen van eigenschappen van een oorspronkelijk object. In welke mate fouten optreden, is verschillend per werkwijze van migratie. Het feit dat bij wijziging van een digitaal object corruptie van het bestand of het verlies van informatie en/of functionaliteit kan optreden, is een nadeel van migratie als strategie in het geheel en vormt geen onderdeel in de beschrijving van

¹ Reference Model for an Open Archival Information System (OAIS), Blue Book, Consultative Committee for Space Data Systems, januari 2002
http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Startpunt Migratieonderzoek

de werkwijzen van migratie in dit hoofdstuk. In het overzicht worden per migratie werkwijze voor-en nadelen genoemd, hierin zal de mate waarin fouten kunnen optreden worden genoemd als voor- of nadeel.

3.2 Overzicht werkwijzen migratie

Startpunt Migratieonderzoek

Werkwijze	Beschrijving	Voordeel	Nadeel
Migratie naar nieuwere versie	<p>Met de werkwijze migratie wordt bedoeld het omzetten van een digitaal bestand in bestandsformaat A naar een nieuwere versie van hetzelfde bestandsformaat A.</p> <p>Dit proces zal elke keer herhaald moeten worden op het moment dat de nieuwere versie van het bestandsformaat verouderd en ontoegankelijk dreigt te worden.</p> <p>Het proces kan stapsgewijs of sprongsgewijs uitgevoerd worden. Onder stapsgewijze uitvoering wordt verstaan dat naar iedere nieuwere versie van het bestandsformaat wordt gemigreerd. Bij een sprongsgewijze uitvoering worden minder migraties uitgevoerd.</p>	<ul style="list-style-type: none">• Bekend proces (bewezen in de praktijk)• Eigenschappen van een document blijven goed bewaard in de gemigreerde versie van het originele bestand.• Mogelijk nieuwe functionaliteit aanwezig in de nieuwe applicatie.	<ul style="list-style-type: none">• Het migratieproces zal vele malen herhaald moeten worden.• Het aantal digitale bestanden die in aanmerking komen voor migratie wordt groter naarmate de inhoud het depot uitgebreid wordt. Migratietrajecten worden langduriger.• Waarschijnlijk niet geschikt voor het bewaren van de toegankelijkheid voor de lange termijn door de cumulatieve kans op fouten bij elk migratieproces. Kans op cumulatie van fouten is groter bij stapsgewijs uitvoeren van het migratieproces.²• Nieuwe inzichten kunnen ertoe leiden dat reeds uitgevoerde migraties tot een slechter resultaat hebben geleid, dan als de migratie door een verbeterde strategie had worden uitgevoerd (bijv. als A -> C een beter resultaat geeft dan A -> B -> C).

² Zie Testbed (Den Haag, 2003), uitgevoerd door het Nationaal Archief

Startpunt Migratieonderzoek

Werkwijze	Beschrijving	Voordeel	Nadeel
Migratie/conversie naar ander bestandsformaat	Met de werkwijze conversie wordt bedoeld het omzetten van een bestandsformaat A naar een bestandsformaat B (geen standaard formaat ³ , maar wel een gangbaarder of “nieuwer” formaat dan formaat A).	<ul style="list-style-type: none">Bekend proces (bewezen in de praktijk)	<ul style="list-style-type: none">Bestaande functionaliteit kan verdwijnen bij het gebruik van een nieuwe applicatie.Het conversieproces zal vele malen herhaald moeten worden.Het aantal digitale bestanden die in aanmerking komen voor migratie wordt groter naarmate de inhoud het depot uitgebreid wordt. Migratietrajecten moeten dan steeds eerder gepland worden.Waarschijnlijk niet geschikt voor het bewaren van de toegankelijkheid voor de lange termijn door cumulatieve kans op verlies van eigenschappen van het originele bestand.Conversie naar een ander bestandsformaat maakt de kans op informatie- of functionaliteitsverlies groter.

³ Definitie standaardformaat: Een de jure-standaard is een gangbaar bestandsformaat, gebaseerd op open standaard (is geen eigendom van een leverancier). Een de facto-standaard is een gangbaar bestandsformaat, gebaseerd op een gesloten standaard en dus eigendom van een leverancier. Als voor een bewaarstrategie gekozen wordt om bestanden op te slaan in een standaardformaat, heeft een de jure-standaard de voorkeur.

Startpunt Migratieonderzoek

Werkwijze	Beschrijving	Voordeel	Nadeel
Normalisatie	Met de werkwijze normalisatie wordt bedoeld de conversie van een bestandsformaat A naar een standaard bestandsformaat C.	<ul style="list-style-type: none"> • Bekend proces (bewezen in de praktijk) • Bij gebruik van enkele standaarden is er een kleiner aantal bestandsformaten die toegankelijk moeten worden gehouden voor de lange termijn. • Een standaard heeft vaak een groter draagvlak onder gebruikers dan een niet-standaard. Het grote draagvlak kan ervoor zorgen dat het behoud van toegankelijkheid ook commercieel aantrekkelijk is. • Verwachting is dat een standaard minder gevoelig is voor digitale veroudering en minder vaak geconverteerd hoeft te worden dan een bestandsformaat dat geen standaard is. 	<ul style="list-style-type: none"> • Het conversieproces zal herhaald moeten worden, een standaard bestandsformaat is niet uitgezonderd van digitale veroudering. • Het aantal digitale bestanden die in aanmerking komen voor migratie wordt groter naarmate de inhoud het depot uitgebreid wordt. Migratietrajecten moeten dan steeds eerder gepland worden. • Keuze maken welke bestandsformaten standaarden zijn. Dit brengt een risico met zich mee • Alleen conversie naar een standaard formaat is niet voldoende om de toegankelijkheid van digitale bestanden voor de lange termijn te garanderen. • Conversie naar een ander bestandsformaat maakt de kans op informatie- of functionaliteitsverlies groter.
“Migration on request”	Met “migration on request” kan het conversieproces van het omzetten van een bestandsformaat A naar een ander formaat B of C	<ul style="list-style-type: none"> • Het conversieproces hoeft pas ingezet te worden op het moment dat 	<ul style="list-style-type: none"> • De noodzaak van een periodiek proces, dat controleert of de

Startpunt Migratieonderzoek

Werkwijze	Beschrijving	Voordeel	Nadeel
	<p>uitgesteld worden tot het betreffende bestand wordt opgevraagd. Dit wil zeggen: het originele bestand wordt geconverteerd naar een hogere versie of naar een ander bestandsformaat op het moment dat een gebruiker het bestand aanvraagt. Bij deze werkwijze is het niet nodig om periodieke migratie van opgeslagen bestanden plaats te laten vinden.</p>	<p>een bestand ook daadwerkelijk aangevraagd wordt. Periodieke, grootschalige conversieprocessen zijn niet nodig.</p> <ul style="list-style-type: none"> • Verlies van eigenschappen van het originele bestand wordt geminimaliseerd door altijd het originele bestand te gebruiken voor conversie. 	<p>conversie van bestandsformaat A naar formaat B of C nog werkzaam is, is een risico. Wanneer formaat B en C niet meer toegankelijk zijn, moet op tijd een conversie naar bestandsformaat D in plaats van B of C mogelijk zijn.</p> <ul style="list-style-type: none"> • Conversietools moeten worden herschreven. • Proces wordt nog niet veel toegepast in de praktijk. • Wachtijd bij aanvraag en rendering van het bestand vanwege de conversie.
<p>“Migration on request” (modular)</p>	<p>Modulaire “migration on request” maakt gebruik van het principe van “migration on request”, dus is een conversieproces pas nodig op het moment dat een bestand wordt aangevraagd. Met modulaire migratie wordt nagestreeft dat de conversietools niet geheel opnieuw ontwikkeld moeten worden op het moment dat het doelbestandformaat wijzigt van B naar C. Door de modulaire opzet is het “input” deel van de tool statisch (gebaseerd op inlezen van het originele bestandsformaat A) en het “output” deel (gebaseerd op het doelbestandformaat) is flexibel te ontwikkelen</p>	<ul style="list-style-type: none"> • Het conversieproces hoeft pas ingezet te worden op het moment dat een bestand ook daadwerkelijk aangevraagd wordt. Periodieke, grootschalige conversieprocessen zijn niet nodig. • Verlies van eigenschappen van het originele bestand wordt geminimaliseerd door altijd het originele bestand te gebruiken voor conversie. 	<ul style="list-style-type: none"> • De noodzaak van een periodieke controle of de conversie van bestandsformaat A naar formaat B of C nog werkzaam is een risico. Wanneer formaat B en C niet meer toegankelijk zijn, moet op tijd een conversie naar bestandsformaat D i.p.v. B of C mogelijk zijn. • Conversietools moeten gedeeltelijk worden herschreven. • Er moet een keuze voor een

Startpunt Migratieonderzoek

Werkwijze	Beschrijving	Voordeel	Nadeel
	voor bestandsformaat B of C of een bestandsformaat dat nog niet bestaat. In de theorie wordt gesproken over een soort “tussenformaat” waarnaar het originele bestandsformaat A als tussenvorm wordt omgezet door het “input” deel voordat het “output” deel er bestandsformaat B of C van maakt.	<ul style="list-style-type: none">• Het is niet nodig om conversietools geheel te herschrijven door de modulaire opzet van de tools.	<ul style="list-style-type: none">• tussenformaat worden gemaakt. De keuze voor een tussenformaat zal de mogelijkheden van het doelbestandsformaat beperken.• Het concept modulaire migratie betreft een onderzoek. Het concept is (nog) niet in de praktijk toegepast.
UVC	<p>De UVC is een virtuele computer naar analogie van de bestaande computerarchitectuur. De UVC is zeer eenvoudig en hierdoor toepasbaar op iedere computer, ook in de toekomst.</p> <p>De implementatie voor JPEG en GIF, die gemaakt is door de KB in samenwerking met IBM, bestaat uit vier componenten:</p> <ol style="list-style-type: none">1. Universal Virtual Computer (UVC)⁴2. Format decoder3. Logical Data Schema (LDS)4. Viewer <p>Het originele object (bij de implementatie een JPEG- of GIF-bestand) wordt met behulp van de Format decoder omgezet naar een Logical Data View (LDV, deze stap is een conversie naar een “standaard” formaat dat platformonafhankelijk is). De LDV kan vervolgens geïnterpreteerd worden</p>	<ul style="list-style-type: none">• Bekend proces (bewezen in de praktijk) voor JPEG- en GIF-formaat.• In de huidige tijd kunnen Format decoder en LDS ontwikkeld worden. Implementatie specificaties van de “bovenste” laag van de UVC, waarop de decoder gaat draaien wordt vastgelegd. Als specificaties voor de Viewer wordt vastgelegd hoe de Viewer met behulp van de LDS het digitale object in de toekomst zal tonen, nog niet hoe de Viewer op een toekomstig platform wordt geïmplementeerd. Hoe de UVC op een toekomstig platform	<ul style="list-style-type: none">• Vooralsnog niet gerealiseerd voor “complexere” bestandsformaten zoals PDF.• Het blijft de vraag of de specificaties voor UVC en Viewer voldoende zijn voor een toekomstige programmeur om een werkende UVC en Viewer te ontwikkelen.

⁴ Zie voor informatie http://www.kb.nl/hrd/dd/dd_onderzoek/uvc_voor_images.html

Startpunt Migratieonderzoek

Werkwijze	Beschrijving	Voordeel	Nadeel
	met behulp van het Logical Data Schema en het object kan bekeken worden met de Viewer, die de LDV vertaald. ⁴	wordt geïmplementeerd speelt pas op het moment in de toekomst als de UVC gebruikt gaat worden. <ul style="list-style-type: none">• Het conversieproces hoeft pas ingezet te worden op het moment dat een bestand ook daadwerkelijk aangevraagd wordt. Periodieke, grootschalige conversieprocessen zijn niet nodig.	

Startpunt Migratieonderzoek

3.2 Overview projects related to migration

Variations of migration	Description	Status
Automatische evaluatie van migratiekwaliteit in een gedistribueerd netwerk van conversieprogramma's	Een opzet voor een architectuur van gedistribueerde conversieprogramma's (op basis van bijvoorbeeld webservices). Het ontwerp maakt deels gebruik van het raamwerk van Rauch en Rauber waarmee een keuze voor een strategie kan worden gemaakt die gebaseerd is op de verwachte Utility. Het ontwerp is geschikt voor meerdere vormen van migratie (normalisatie, migration on request etc.)	<ul style="list-style-type: none">Het is nog een concept, geen praktijkervaring.
PANIC	Het PANIC systeem heeft een geïntegreerde en uitbreidbare architectuur, gebaseerd op preservation metadata, diensten met automatische notificatie, software en bestandsformaat registers en semantische grid diensten. PANIC biedt een duurzaam, dynamische aanpak voor de opslag voor lange termijn van grote collecties met heterogene wetenschappelijke data. Migratiediensten worden beschikbaar gesteld via dit PANIC systeem.	<ul style="list-style-type: none">Het is nog vooral een theoretisch concept.
GRACE	Grace is een http proxy server waarmee op transparante wijze voor een browser ongeschikte en verouderde bestanden geconverteerd kunnen worden. De browser kan de geconverteerde bestanden tonen zonder het gebruik van extra plug-ins. Grace kan per gebruiker geconfigureerd worden en kan worden uitgebreid met vele conversiediensten. Het prototype van Grace transformeert diverse beeldformaten (XBM, PNG, JPEG2000) zodat deze kunnen worden getoond in Internet Explorer.	<ul style="list-style-type: none">Het betreft een prototype.

4 Voorlopige evaluatie

Evaluatie

Veel instellingen geven aan migratie als voornaamste strategie op het gebied van digitale duurzaamheid te hebben. Het is echter niet makkelijk om via Internet te achterhalen of er concrete tests zijn uitgevoerd. Het aantal instellingen dat heeft gepubliceerd over een migratieproject is kleiner dan verwacht.

Het PLANETS project is een goede gelegenheid om internationaal onderzoek te doen naar digitale duurzaamheidsstrategieën en om concrete tests (gezamenlijk) uit te voeren.

Vervolg

Op basis van de verzamelde informatie over migratieonderzoeken zal worden bekeken welke werkwijzen van belang kunnen zijn voor de KB.

Vervolgens zullen op kleine schaal enkele werkwijzen getest worden.

5 Literatuurlijst

Bennet, J. C. A: Framework of data types and formats, and issues affecting the long term preservation of digital material, British Library Research and Innovation Report 50, 1997.

Boudrez, F.: Migratie naar Archiveringsformaten, Richtlijn DAVID.

Caplan, P.: Building a digital preservation archive: tales from the front, VINE: The Journal of Information and Knowledge Management Systems, vol. 34 (1), (2004), pp 38-42.

Clausen, L. R.: Handling file formats, State and University Library and Royal Library, Denmark, May 2004.

Ferreira, M.: Automatic Evaluation of Migration Quality in Distributed Networks of Converters, ECDL Doctoral Consortium Poster, 2005.

Hedstrom, M., Lampe, C.: Emulation vs. Migration: Do Users Care?, RLG DigiNews, vol. 5, (6), 2001.

Heslop, H., Davis, S., Wilson, A.: An Approach to the Preservation of Digital Records Australia, 2002.

Hoeven, J.R. van der, Diessen, R.J. van en Meer, K. van der, Development of a Universal Virtual Computer (UVC) for long-term preservation of digital objects, Journal of Information Science, vol. 31(3), p. 196-208, 2005.

Lawrence, G. W., Kehoe, W. R., Rieger, O. Y., Walters, W. H., Kenney, A. R.: Risk Management of Digital Information: A File Format Investigation, Council on Library and Information Resources, Washington, D.C, June 2000.

Lorie, R. The UVC: method for preserving digital documents – proof of concept. LTP report series, number 4, 2002.

Ockerbloom, J.M.: Archiving and Preserving PDF files, RLG DigiNews, vol. 5, (1), 2001.

Rauch, C., Rauber, A., Z. Chen et al. (Eds.): Preserving Digital Media: Towards a Preservation Solution Evaluation Metric, ICADL vol. 3334, pp. 203-212, 13-17 Proceedings, 2004.

Rosenthal, D. S. H., Lipkis, T., Morabito, S., Robertson, T.: Transparent Format Migration of Preserved Web Content, November 22, 2004.

Testbed Digitale Bewaring: Bewaren van tekstdocumenten: Van digitale vluchtigheid naar digitaal houvast, Den Haag, 2003.

Testbed Digitale Bewaring: Migratie: Context en huidige stand van zaken, Den Haag, 2001.

Startpunt Migratieonderzoek

Verheul, I.: Networking for Digital Preservation: Current Practice in 15 National Libraries, IFLA Publications 119, 2006.

Walker, F. L., Thoma, G. R.: Image Preservation Through PDF/A, Lister Hill National Center for Biomedical Communications, National Library of Medicine, Bethesda, 2004.

Wijngaarden, H. van en Oltmans, E., Digital Preservation and Permanent Access: the UVC for images, 2004.

Wing, J.M., Ockerbloom, J.: Respectful Type Converters for Mutable Types, IEEE Transactions on Software Engineering, vol. 26, (7), July 2000.

Startpunt Migratieonderzoek

6 Termenlijst (voorlopig)

Term	Beschrijving
Migratie Conversie	Een proces dat het digitale object / digitale objecten wijzigt om het object bruikbaar en/of toegankelijk te houden.
Migration on access Migration on request	Een proces dat het digitale object / digitale objecten wijzigt op het moment van aanvraag om het object bruikbaar en/of toegankelijk te maken.
Batch migratie	Een proces dat een batch digitale objecten wijzigt om deze bruikbaar en/of toegankelijk te houden.
Online migratie	Een proces dat het digitale object / digitale objecten wijzigt om het object / de objecten bruikbaar en/of toegankelijk te maken via een conversie tool die online wordt aangeboden. Het is niet nodig om de conversie tool te downloaden of te installeren.
Automatische migratie	Het proces dat het digitale object / digitale objecten wijzigt om het object / de objecten bruikbaar en/of toegankelijk te houden wordt via een tijdschakelaar aangezet.
Handmatige migratie	Het proces dat het digitale object / digitale objecten wijzigt om het object / de objecten bruikbaar en/of toegankelijk te houden wordt met behulp van een handmatige actie aangezet.