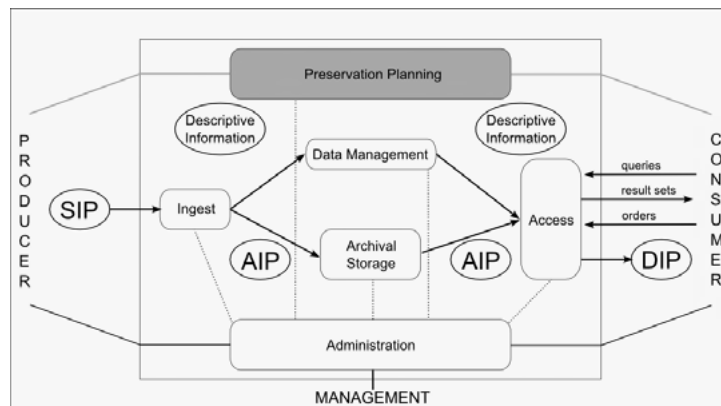


HANDLEIDING ONTWIKKELING PRESERVERINGSSTRATEGIEËN EN PRESERVERINGSACTIEPLAN

V1.1

BEELD EN GELUID





ENGLISH SUMMARY

One of the organisational goals of the Netherlands institute of Sound and Vision is to become an OAIS-compliant trustworthy digital archive. To achieve this a project was started in 2013 to define a normative reference model to use as the basis for a gap analysis between the OAIS-goals and the current practices within the organisation. This document addresses the Preservation planning functional entity from the OAIS functional model. This entity is made up of four functions that can be loosely summed up as: community monitoring, technology watch, defining preservation strategies and defining preservation plans. At the time of writing none of these functions were addressed within the organisation even though some of these functions were occasionally carried out in a non-standardized and somewhat informal way. This document functions as the groundwork for setting up the preservation planning functional entity in a standardized and structured way.

According to the model preservation strategies are written as recommendations by the preservation planning functional entity based on the reports of the community monitoring and technology watch-functions. These recommendations are filed to the Administration functional entity that decides on migration goals based on the preservation policy and organisational and financial considerations. These migration goals are worked out into preservation plans that describe practical solutions and actions to counter detected preservation risks as well as all possible choices and considerations that might apply to these solutions. These plans also need to detail what role preservation metadata plays in the solutions, how the actions affect the digital objects and how the preservation of significant properties of the files are safeguarded throughout the process of migration. To make informed decisions, preservation strategies and preservation plans need to take into account information from a number of different domains:

- Technical information: reports from Risk assessment, disaster and recovery policies, technology watch
- Object information: preservation metadata, preservation levels, provenance information, significant properties
- Organisational policies: Mission, collection policy, retention policy, preservation policy, financial information, organisational standards
- Information related to producers: Submission agreements, Order agreements, reports from monitoring designated communities

The process of working out a preservation plan can be structured as follows:

1. Identifying the preservation risk
2. Identifying the collection concerned
3. Determine constraints that apply when formulating a solution
4. Formulating possible solutions
5. Reviewing and testing solutions
6. Work out solution into a complete preservation plan
7. Evaluate and adjust preservation plan

The need for creating a preservation plan is two-fold: on the one hand it offers a complete view on all risks and actions that make up the solution to counter a preservation risk, on the other hand it offers a historical view on all actions that have been performed on the digital objects in order to prove to the designated communities the objects are authentic and have been handled according to specification.

When talking about migration it is important to consider the different types of migration that are specified by OAIS. For instance the type refreshment when only underlying carriers



are changed but not the objects themselves, replication when creating a copy of the object, repackaging when changing the form of the object but not the essence and transformation when changing the form and the essence. In the last case it is important that preservation planning takes into account preservation of the transformational migration properties. These properties are aspects of the archival object that need to be maintained as much as possible in the new format. How much of this information can be maintained in the new format depends on technical aspects of the input and output format, type of use by the designated community and any possible limitations that might apply on the side of the producer or the archival infrastructure as well as organisational policies and financial considerations.

Considering the above a typical preservation plan needs to contain the following:

- Identification of the plan
- State of the plan and detected preservational risk
- Determination of constraints and limitations
- Collection concerned
- Details on the process of selection of the best solution
- Roles, responsibilities and costs
- Preservation action plan

The action plan needs to contain:

- Actions that need to be performed
- Planning of which actions are performed when
- What software/hardware is needed
- Which mechanisms are used to check whether actions have been succesful
- Who is responsible for which action

INHOUD

<i>Inleiding</i>	6
Het OAIS Proces Preservation Planning	6
Rol van Preserveringsmetadata	9
Doel van dit document	9
<i>Benodigde informatie voor het formuleren van preserveringsstrategieën</i>	11
<i>Het formuleren van een preserveringsplan</i>	12
Identificeren aanleiding	13
Identificeren collectie	13
Bepalen randvoorwaarden voor oplossing	13
Speciale randvoorwaarden bij de preserveringsstrategie migratie/transformatie	15
Opstellen mogelijke strategieën	17
Kwantificeren en/of testen strategieën	17
Strategie uitwerken tot een preserveringsplan	17
Plan evalueren en bijstellen	17
<i>Wat is de opbouw van een preserveringsplan?</i>	18
Identificatie	18
Status en aanleiding	18
Bijschrijving van randvoorwaarden	19
Beschrijving collectie	19
Vastlegging proces van selectie	19
Rollen, verantwoordelijkheden, kosten	20
Preserveringsactieplan	20
<i>Use Case: Fixity-Informatie voor DPX-Bestanden: Een praktijkvoorbeeld van Preservation Planning</i>	21

INLEIDING

Een van de doestellingen in het meerjarenplan 2012-2015 van Beeld en Geluid is het bereiken van overeenstemming met het OAIS-model met als doel het verkrijgen van de status van 'Trusted Digital Repository'. Het TDR-project is opgezet ten einde de randvoorwaarden om dit te bereiken in kaart te brengen. Binnen het OAIS-referentiemodel is een basisopzet gedefinieerd die aangeeft hoe een 'trusted' digitaal archief ingericht moet zijn om te kunnen zorgen dat een digitaal object van ingest tot aan access op een betrouwbare manier verwerkt, opgeslagen en uitgeleverd kan worden. Het is hierbij belangrijk dat de gebruikers van het archief, de designated communities, het object beschouwen als authentiek en bruikbaar. Het is de taak van het archief om deze authenticiteit aan te tonen aan de hand van een 'audit trail' van acties die op het object zijn uitgevoerd. Het archief moet zorgen dat de objecten bruikbaar blijven door acties op de objecten uit te voeren die ervoor zorgen dat deze bruikbaarheid in de loop der tijd in stand blijft.

HET OAIS PROCES PRESERVATION PLANNING

Binnen het functionele model van OAIS is de functionele entiteit 'Preservation planning' van groot belang aangezien deze verantwoordelijk is voor het betrouwbaar en begrijpelijk houden van de digitale objecten in het archief voor de lange termijn. Het onderdeel 'preservation planning' is verantwoordelijk voor het opstellen van beleidsbeslissingen en richtlijnen die betrekking hebben op het behoud van digitale objecten voor de lange termijn. Dit betekent onder andere het opstellen en uitvoeren van preserveringsplannen ten einde de digitale objecten die gepreserveerd moeten worden beschikbaar en begrijpelijk te houden voor de gebruikers.¹

Binnen het functionele OAIS-model zijn een viertal functies van preserveringsplanning expliciet benoemd.

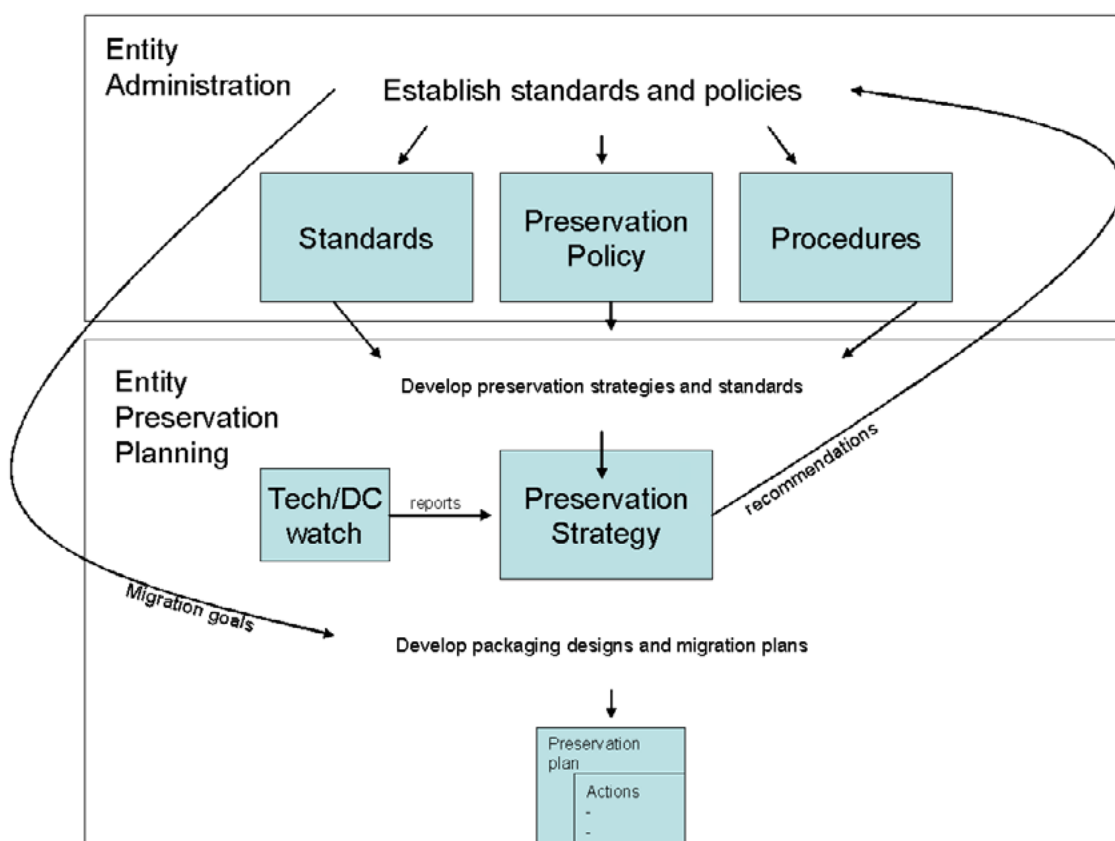
- **Communitymonitoring:** deze functie heeft als doel de begrijpelijkheid van de objecten blijvend te kunnen toetsen en eventuele nieuwe eisen op basis van technische veranderingen tijdig te kunnen opmerken.
- **Technologiemonitoring:** deze functie heeft als doel om te waken voor de toegankelijkheid van in het archief gebruikte hardware en software en tijdig alternatieven aan te dragen wanneer dit noodzakelijk is
- **Opstellen preserveringsstrategie:** deze functie gebruikt de input vanuit de communitymonitoring en de technologiemonitoring om adviezen te kunnen uitbrengen over de te volgen preserveringsstrategie

¹ Reference model for an Open Archival Information System, 4-2, <http://public.ccsds.org/publications/archive/650x0m2.pdf>

- **Opstellen preserveringsplan:** deze functie heeft als doel op basis van beleid zoals geformuleerd binnen de preserveringsstrategie concrete plannen op te stellen en deze te testen.²

Voor al deze functies binnen 'Preservation planning' moeten beleidsstukken ontwikkeld worden. Dit betekent dat er nog documenten opgeleverd moeten worden aangaande 'monitoring Designated communities', 'Technologiemonitoring', 'Risicoanalyse van hardware, software en mediaformaten' en een document dat de preserveringsstrategie bepaald zoals deze gevolgd wordt binnen het archief.

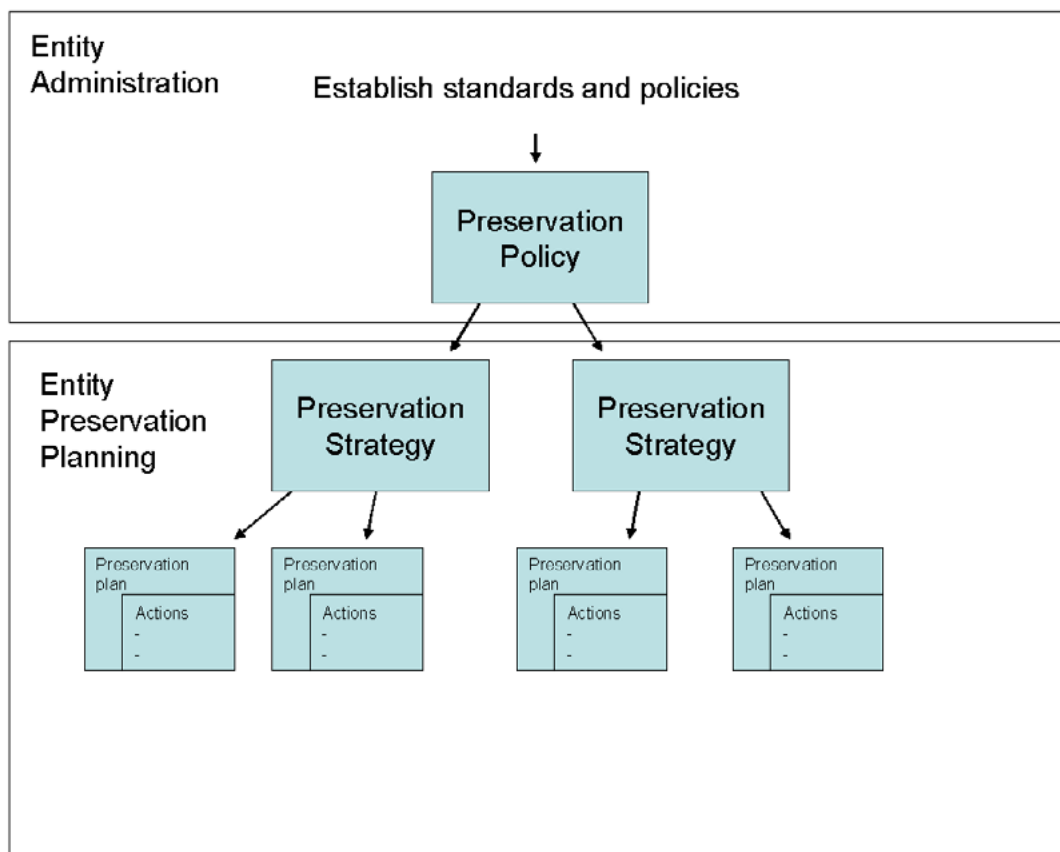
Monitoring van zowel technologie als de designated communities moet ervoor zorgen bijvoorbeeld dat risico's rondom mediaformaten tijdig gesignaleerd en verholpen kunnen worden. Het maken en bewaren van preserveringsplannen moet tenslotte het bewijs vormen dat voorziene risico's adequaat zijn aangepakt.



Figuur 1. Hoe het preserveringsplan tot stand komt op basis van verschillende informatiestromen

² Reference model for an Open Archival Information System, 4-14-4-16, <http://public.ccsds.org/publications/archive/650x0m2.pdf>

In figuur 1 zien we hoe dit proces verloopt en de wisselwerking tussen de entiteit Administration en de entiteit Preservation planning: de entiteit Administratie is verantwoordelijk voor het opstellen van standaarden en beleidstukken op een overkoepelend niveau. Hiertoe behoort ook het preserveringsbeleid. Op basis hiervan en op basis van input vanuit Technology monitoring en monitoring Designated communities worden binnen de entiteit Preservation Planning strategieën, standaarden en risicorapporten ontwikkeld die als aanbeveling terug gestuurd worden naar Administratie ter goedkeuring. Administratie bepaalt in hoeverre deze adviezen omgezet worden in migratiedoelen. Deze worden vervolgens binnen de functie ‘Develop packaging designs and migration plans’ uitgewerkt tot concrete preserveringsplannen met actiepunten die uitgevoerd kunnen worden. Deze plannen bevatten dus zowel de aanleiding, het risico dat vermeden moet worden, evenals de strategie die gevolgd wordt om te komen tot een aantal concrete stappen die gezet moeten worden om het risico weg te nemen.



Figuur 2. Het preserveringsbeleid binnen het archief leidt tot het opstellen van een aantal strategieën die weer uitgewerkt worden tot een aantal concrete plannen met daarin beschreven de acties die uitgevoerd moeten worden.



ROL VAN CONSERVERINGSMETADATA

Conserveringsplanning behelst echter meer dan alleen plannen opstellen om te voorkomen dat mediaformaten verouderd raken. Het is ook de taak van deze functie om beleid te formuleren rondom implementatie van voldoende conserveringsmetadata om conserveringsstrategieën te kunnen uitvoeren. Metadata draagt tenslotte ook bij aan de begrijpelijkheid en de toegankelijkheid van een digitaal object.

Om de digitale objecten in een bepaalde collectie te kunnen conserveren is het noodzakelijk de kenmerken van het object vast te stellen en deze vast te leggen in conserveringsmetadata evenals de verschillende conserveringsevents die hebben plaatsgevonden. Dit om aan te kunnen tonen dat de eigenschappen van het object en de acties die zijn gedaan in de levenscyclus van het object overeenkomen met de strategie zoals vastgesteld in het conserveringsbeleid.

Aangezien de ontwikkeling van conserveringsstrategieën en actieplannen afhankelijk is van monitoring van de hele omgeving is het van belang dat de gehele digitale workflow en alles er om heen in kaart is gebracht om de planning op een geïnformeerde manier te kunnen uitvoeren. Het is dan ook niet mogelijk om conserveringsplanning te reduceren tot alleen een aantal technische eisen die aan digitale objecten worden gesteld. Het hele proces van ingest tot aan uitlevering moet in beschouwing genomen worden inclusief alle factoren en beslissingen die daarbij een rol spelen.

Het opstellen, volgen en bewaren van conserveringsplannen dient een tweeledig doel binnen het OAIS-model. Enerzijds zorgt het voor een concrete handleiding aan de hand waarvan conserveringsacties eenduidig en gestandaardiseerd uitgevoerd kunnen worden, anderzijds zorgt het voor een geschiedenis van uitgevoerde conserveringsacties op basis waarvan het archief aantoonbaar kan maken hoe informatie vanuit Technology monitoring, Monitoring designated communities en Risk analysis wordt ingezet om de digitale objecten blijvend te conserveren. Op basis van de conserveringsplannen moet het archief kunnen aantonen dat er oplossingen zijn bedacht voor gedetecteerde risico's, dat er maatregelen zijn genomen om te zorgen dat de objecten overeenkomen met de wensen van de designated communities en dat de doorgevoerde maatregelen voldoen. Het is dan ook belangrijk om aan te kunnen tonen dat conserveringsplannen regelmatig worden geëvalueerd en bijgewerkt.

DOEL VAN DIT DOCUMENT

Dit document bevat een beschrijving voor de functies 'Opstellen conserveringsstrategieën' en 'Opstellen migratieplannen'. Het beschrijft hoe het proces van het opstellen van conserveringsstrategieën ingericht kan worden en wat er in een conserveringsplan moet staan. De functie 'Opstellen migratieplannen' zal vooral zorgen voor het beschrijven van de praktische acties zoals deze in het conserveringsactieplan moeten komen. De functie

‘Opstellen preserveringsstrategieën’ zal vooral een bijdrage leveren door op een hoger niveau risico’s te signaleren en oplossingen te formuleren. Allebei deze bijdrages komen tenslotte samen in een preserveringsplan, dat zowel de aanleiding en achtergrond van de preserveringsstrategieën beschrijft alsmede de concrete acties die uitgevoerd moeten worden.

Tenslotte is een praktijkvoorbeeld van het preserveringsplanningproces uitgewerkt aan de hand van een actueel scenario.

Dit document bevat verder geen preserveringsplannen. Deze zullen in de loop der tijd ontwikkeld moeten worden in de lijn door personen die hier vanuit hun functie verantwoordelijk voor zijn. Aangezien de besluitvorming rond deze materie afhankelijk is van een veelheid aan factoren zoals kosten, doorlooptijd, beleid en afspraken in contracten, is het aannemelijk dat personen uit verschillende afdelingen betrokken worden bij dit proces.

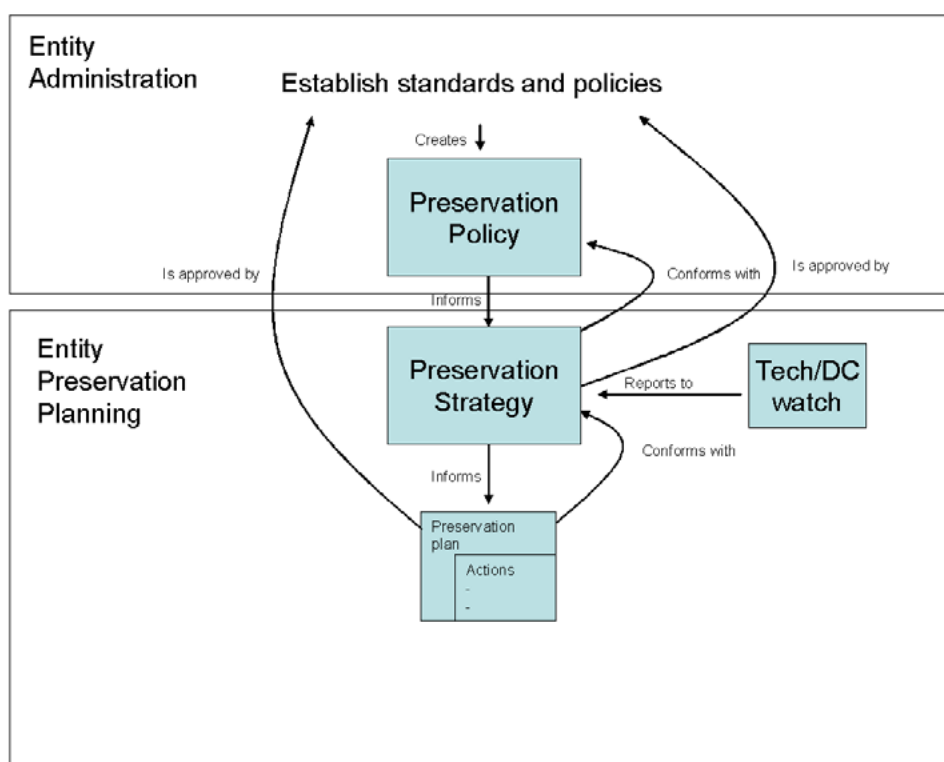
Veel archieven zullen in de praktijk op een impliciete manier al bezig zijn met het ontwikkelen van preserveringsstrategieën doordat ze bijvoorbeeld keuzes maken voor het updaten van software of het vervangen van hardware maar deze keuzes komen misschien niet altijd via een gedefinieerd proces tot stand of worden niet gedocumenteerd op een manier die voor iedereen inzichtelijk is.

Dit document heeft vooral tot doel een gestandaardiseerd model te bieden voor het maken, testen en documenteren van keuzes die invloed hebben op het bewaren en beschikbaarstellen van archiefmateriaal met het oog op de lange termijn.

BENODIGDE INFORMATIE VOOR HET FORMULEREN VAN PRESERVERINGSSTRATEGIEËN

Om verantwoorde keuzes te kunnen maken in het opstellen van een preserveringsstrategie en een preserveringsplan, is het belangrijk voor het archief om op een aantal gebieden benodigde informatie beschikbaar te hebben:

1. Techniek: resultaten van Risk Assessment, Disaster recovery planning, Technology monitoring
2. Digitale objecten: Preserveringsmetadata, representatieinformatie, geldende preserveringsniveau's, provenance-informatie, Transformational information properties
3. Beleid: Missie, Collectiebeleid, Retentiebeleid, Preserveringsbeleid, budgettering, geldende standaarden
4. Depotgevers/Gebruikers: voorwaarden uit Submission Agreements, Order Agreements, resultaten uit Community monitoring, beschrijving Designated communities



Figuur 3. De relaties tussen de verschillende onderdelen die een rol spelen bij preserveringsplanning.



HET FORMULEREN VAN EEN CONSERVERINGSPLAN

Het conserveringsplan dient een aantal zaken te bevatten, zoals aanleiding en randvoorwaarden, die duidelijk maken wat de achtergrond is van een bepaalde conserveringsstrategie. Dit om een verantwoording te kunnen vormen voor acties die invloed hebben op het digitale archief object, de AIP. In het conserveringsplan moeten tevens zaken als kosten en doorlooptijd ingeschat zijn om de acties in te kunnen plannen en budget hiervoor vrij te maken.

Een conserveringsplan moet uiteindelijk ook resulteren in een lijst van zeer concrete en uitvoerbare acties, het conserveringsactieplan.

Deze twee componenten, de aanleiding, randvoorwaarden en verantwoording enerzijds en de concrete acties in het actieplan anderzijds vormen tezamen het conserveringsplan.

Om tot een niveau van detail te kunnen komen zoals dit nodig is voor het actieplan, moeten een hoeveelheid variabelen in kaart gebracht worden en een aantal verschillende scenario's voorgesteld en geëvalueerd worden.

Het proces kan er in het kort als volgt uitzien:

1. identificeren aanleiding
2. identificeren collectie
3. bepalen randvoorwaarden voor oplossing:
 - a. techniek (hardware, software, formaten)
 - b. gebruikers/producers (monitoring designated communities, SA, OA)
 - c. procedureel (personeel, kosten, doorlooptijd, foutafhandeling)
 - d. beleid (missie, strategisch conserveringsplan)
4. opstellen mogelijke scenario's
5. kwantificeren en/of testen scenario's
6. scenario uitwerken tot een conserveringsplan
7. plan evalueren en bijstellen (op basis van verandering in techniek, wensen, organisatie, beleid)³

Het is hieruit al duidelijk dat het plan niet het einde vormt van het proces maar altijd bijgesteld kan worden en regelmatig geëvalueerd moet worden om te kunnen anticiperen op nieuwe ontwikkelingen.

Het plan dient een tweeledig doel, enerzijds biedt het een concreet overzicht van risico's en acties die leiden tot het bepalen van een plan om deze risico's te voorkomen, anderzijds biedt het een historisch overzicht van de besluitvorming rondom digitale objecten dat bijdraagt aan de geloofwaardigheid van de digitale objecten en tevens de gemaakte keuzes begrijpelijk maakt voor de gebruikers van het archief.

³ Deze indeling is met enkele aanpassingen ontleend aan Becker et al., Systematic Planning for Digital Preservation, 4.2, <http://www.ifs.tuwien.ac.at/~becker/pubs/becker-ijdl2009.pdf>

IDENTIFICEREN AANLEIDING

Een bepaalde aanleiding vormt altijd de start van het proces. Er kunnen diverse zaken aanleiding vormen om een preserveringsplan op te stellen of te herzien. Organisatorische zaken kunnen een rol spelen, bijvoorbeeld veranderd beleid rondom preservering. Ook veranderingen in de infrastructuur kunnen een rol spelen, bijvoorbeeld wanneer software verandert of hardware vervangen moet worden. Tenslotte kunnen ook veranderingen van buitenaf een rol spelen zoals wanneer uit monitoring designated communities blijkt dat gebruikers de huidige dienstverlening niet meer voldoende vinden.

IDENTIFICEREN COLLECTIE

Na het identificeren van de aanleiding is het van belang te bepalen welk deel van de collectie getroffen wordt door een mogelijke verandering. Bij collectie moet gedacht worden aan een willekeurige indeling van objecten in een categorie. Dit kan een genre zijn maar kan evengoed bestaan uit objecten die inhoudelijk geen verband hebben maar alleen technisch, bijvoorbeeld doordat ze van hetzelfde bestandstype zijn of omdat ze op dezelfde server staan.

BEPALEN RANDVOORWAARDEN VOOR OPLOSSING

Na deze twee identificatiestappen is het vervolgens van belang om een beeld te krijgen van alle factoren die een rol spelen bij een mogelijke verandering.

Deze kunnen opgedeeld worden in een aantal categorieën:

Technische voorwaarden: de oplossing moet passen binnen een bestaande infrastructuur, bepaalde gewenste functionaliteit moet blijven werken na een wijziging, de objecten mogen niet in kwaliteit achteruit gaan, bepaalde metadata moet behouden blijven.

Gebruikersvoorwaarden: gebruikers moeten baat hebben bij de verandering, ze moeten de veranderingen kunnen begrijpen, wijzigingen moeten passen binnen afspraken die zijn gemaakt met depotgevers.

Procedurele voorwaarden: de oplossing moet binnen een bepaalde tijd opgelost kunnen worden, de oplossing mag meer kosten dan een bepaald bedrag, er moet voldoende personeel zijn om de wijzigingen te kunnen begeleiden, het moet duidelijk zijn hoe fouten worden afgehandeld.

Beleidsvoorwaarden: De oplossing moet in lijn zijn met de missie statement, de oplossing moet passen in de lijn van het preserveringsbeleid,, het collectiebeleidsplan en geldende preserveringsniveau's.

Om deze voorwaarden goed in beeld te krijgen is het belangrijk om relevante documentatie op een centrale plek raadpleegbaar te hebben. Om bijvoorbeeld te controleren of een oplossing in overeenstemming is met gebruikerswensen is het goed om documenten zoals opgeleverd door monitoring designated communities bij de hand te hebben.

Het preservingsbeleid binnen het archief zal ook een grote rol spelen. Binnen het preservingsbeleid van het archief moet vastgelegd zijn welke strategie gekozen is voor preservering. Voorbeelden van verschillende preservingsstrategieën zijn emulatie, migratie of de keuze om formaten te houden zoals ze zijn ook wel 'bit preservation' genoemd.

De Preserveringsstrategie Migratie

De preservingsstrategie 'migratie' kan soms leiden tot verwarring aangezien onder deze noemer doorgaans een viertal verschillende technieken geschaard worden die het beste onderscheiden kunnen worden door middel van aparte termen.

- **Verversing:** het vervangen van het onderliggende medium waarop de objecten staan, zonder dat dit effect heeft op de objecten zelf, bijvoorbeeld het migreren van MXF-files van LTO-4 naar LTO-5.
- **Replicatie:** het bit voor bit kopiëren van de data van een digitaal object naar een nieuw bestand, zonder dat er aan het bestand zelf iets verandert, bijvoorbeeld bij het maken van een backup. De context kan hierdoor overigens wel veranderen doordat er bijvoorbeeld een nieuwe locatie moet worden opgenomen in de metadata.
- **Repackaging:** het vervangen van de verpakking zonder dat de data in de verpakking wijzigt maar hooguit de structuur van de data, bijvoorbeeld wanneer een DPX-bestand in een directorystructuur wordt gemigreerd naar een AXF-bestand.
- **Transformatie:** het daadwerkelijk veranderingen aanbrengen in de onderliggende data van een digitaal object, bijvoorbeeld wanneer een MOV-bestand met MP4-video wordt omgezet naar een MXF met MPG2-video. Deze vorm van migratie valt bovendien uiteen in twee soorten namelijk omkeerbare transformatie en onomkeerbare transformatie, waarbij bovengenoemd voorbeeld behoort tot de laatstgenoemde categorie vanwege de lossy codering die in dit geval een rol speelt bij de migratie.



SPECIALE RANDVOORWAARDEN BIJ DE PRESERVERINGSSTRATEGIE MIGRATIE/TRANSFORMATIE

Wanneer een conserveringsplan wordt opgesteld als uitwerking van de conserveringsstrategie 'migratie' is het van belang dat er ook rekening gehouden wordt met de zogenaamde '*transformational information properties*'. Dit zijn kenmerken van een digitaal object die helpen te bepalen of een transformatie als geslaagd beschouwd kan worden. Het vaststellen van deze kenmerken heeft als doel om bij een transformatie zoveel mogelijk eigenschappen van het bronformaat mee te nemen naar het nieuwe formaat.

Wat hierbij mogelijk is wordt bepaald door:

- a. de technische eigenschappen van het bronformaat
- b. de technische eigenschappen van het doelformaat
- c. het type gebruik vanuit de designated communities
- d. beperkingen vanuit de producer
- e. beperkingen vanuit de environment

Om te bepalen welke eigenschappen te behouden is het allereerst van belang om de technische eigenschappen van het bronformaat in beeld te hebben, bijvoorbeeld doordat deze als technische metadata zijn opgeslagen als onderdeel van de conserveringsmetadata. De eigenschappen van het doelformaat zijn ook van belang. Bij een transformatie van DV naar D-10 MXF is het bijvoorbeeld belangrijk om te beseffen dat D-10 MXF alleen 4:2:2-chroma subsampling ondersteunt. Hierdoor is het niet mogelijk om de oorspronkelijke chroma-waarde van DV (4:2:0) te handhaven binnen het nieuwe formaat.⁴

De manier waarop het digitale object wordt gebruikt door de designated community is ook van belang bij het vaststellen van de Transformational information properties. Het kan bijvoorbeeld zijn dat de gebruiker bij bepaalde data heeft bepaald dat de inhoud van de data belangrijker is dan het uiterlijk. Het archief kan dan bepalen dat uiterlijk bij transformatie geen rol speelt wanneer dit de transformatie haalbaarder maakt.⁵

Ook kunnen bepaalde afspraken met de producer/rechthebbende beperkingen opleggen omtrent de eigenschappen die meegenomen kunnen worden bij een transformatie.

Tenslotte speelt de environment ook een rol binnen de besluitvorming omtrent transformatie. De softwareomgeving waarbinnen een bepaald formaat afgespeeld kon worden hoeft na een transformatie niet meer dezelfde te zijn. Wanneer het van belang is dat het nieuwe formaat ook in een bepaalde softwareomgeving blijft werken moet dit als een transformational information property opgenomen worden.

⁴ De transformatie zelf hoeft hierdoor overigens niet als mislukt beschouwd te worden zolang vooraf is bepaald dat het upscalen van de chroma een minder groot risico is als het behouden van het oorspronkelijke formaat, bijvoorbeeld doordat het oorspronkelijke bronformaat niet benoemd is als een preserveerbaar formaat binnen het conserveringsbeleid omdat het geen open standaard is.

⁵ DANS heeft dit bijvoorbeeld zo bepaald in hun beleid, zie guideline 12, https://assessment.datasealofapproval.org/assessment_47/seal/html/

De transformational information properties zijn dus niet inherent aan het object maar worden vastgesteld in het kader van een specifieke transformatieactie. Het is dus bijvoorbeeld niet mogelijk om te stellen dat bij video de aspect ratio altijd een transformational information property vormt. Bij het maken van een DIP ligt het voor de hand dat een 4:3-video niet wordt uitgerekt tot 16:9 omdat dit het beeld vervormt. Bij het scannen van 4:3-film naar HD is het echter noodzakelijk om de oorspronkelijke aspect ratio aan te passen naar 16:9 omdat HD geen 4:3 ondersteunt.⁶ De transformational information properties kunnen dus alleen vastgelegd worden per specifieke transformatie zoals deze in een preserveringsplan kan zijn opgenomen. Ook is het belangrijk om te beseffen dat deze eigenschappen niet alleen een enkelvoudige waarde hoeven te hebben maar dat een reeks van toegestane waarden ook van toepassing kan zijn.

Een transformatie kan als geslaagd beschouwd worden wanneer de vastgestelde transformational information properties in een nieuw formaat behouden zijn en de designated community het object in deze nieuwe vorm beschouwt als een betrouwbare en bruikbare representatie van het origineel.⁷ Bij de andere vormen van migratie verandert de inhoud van het data-object niet en kan het succes van de migratieactie simpelweg gecontroleerd worden aan de hand van een berekening van de checksum van de essence.

⁶ Door het toevoegen van zwarte balken aan de zijkant blijft het actieve beeld echter in de praktijk wel 4:3.

⁷ Reference model for an Open Archival Information System, 5-4-5-9,
<http://public.ccsds.org/publications/archive/650x0m2.pdf>

OPSTELLEN MOGELIJKE STRATEGIEËN

Nadat alle randvoorwaarden in kaart zijn gebracht kan gekeken worden naar mogelijke oplossingen. Sommige strategieën zullen daarbij al meteen afvallen omdat ze niet kunnen voldoen aan sommige van de gestelde randvoorwaarden. Van de overgebleven strategieën kan bekeken worden of een keuze te maken is op basis van een vergelijking van strategieën of dat een test noodzakelijk is om de beste strategie te kiezen.

KWANTIFICEREN EN/OF TESTEN STRATEGIEËN

Bij het evalueren van verschillende strategieën is het aan te raden om producten te vergelijken op basis van zo objectief mogelijk te kwantificeren variabelen zoals kosten, verwachte doorlooptijd en concrete technische uitkomsten. Om verschillende producten met uitleenlopende resultaten toch zo objectief mogelijk te kunnen vergelijken kan een waarde toegekend worden aan bepaalde eisen. Tools die ontwikkeld zijn om dit proces te ondersteunen, zoals PLATO⁸, bieden een interface om dergelijke gegevens op een uniforme manier vast te leggen en het resultaat inzichtelijk te maken.

STRATEGIE UITWERKEN TOT EEN PRESERVERINGSPLAN

Na het kiezen van de beste strategie kan deze vervolgens uitgewerkt worden tot een preserveringsplan. Een preserveringsplan bevat een vastlegging van bovengenoemde proces inclusief een prerveringsstappenplan waarin concrete acties worden opgesomd in de volgorde waarin deze dienen te verlopen inclusief rollen, verantwoordelijkheden en eventuele condities. Het is de bedoeling dat betrokkenen met dit deel van het plan direct aan de slag kunnen.

PLAN EVALUEREN EN BIJSTELLEN

Aangezien een preserveringsplan, afhankelijk van de inhoud en strekking ervan, een aanzienlijke tijd kan bestrijken is het noodzakelijk om het plan te evalueren wanneer hier aanleiding voor is, bijvoorbeeld door veranderingen in de organisatie. Zie Identificeren aanleiding.

⁸ PLATO is een online tool die een structuur aanbiedt om preserveringsplanning vorm te geven. Deze bevat tevens de mogelijkheid om bepaalde services te integreren, zoals extractie van technische metadata, zie <http://www.ifs.tuwien.ac.at/dp/plato/intro.html>

WAT IS DE OPBOUW VAN EEN CONSERVERINGSPLAN?

Een conserveringsplan dient vooral een praktisch uitvoerbaar stappenplan met acties te bevatten, echter is ook van belang om te beschrijven hoe men tot deze lijst van acties is gekomen. Het is van belang om in het kader van de navolgbaarheid en de transparantie, zoals deze binnen TRAC⁹ vereist worden, kort de randvoorwaarden en de redenen achter de gemaakte keuzes te beschrijven zoals deze zijn bepaald tijdens het opstellen van de conserveringsstrategie. De essentiële onderdelen van een conserveringsplan kunnen als volgt weergegeven worden:

- Identificatie
- Status en aanleiding
- Beschrijving van randvoorwaarden
- Beschrijving collectie
- Vastlegging proces van selectie
- Rollen, verantwoordelijkheden, kosten
- Conserveringsstappenplan¹⁰

IDENTIFICATIE

Het is belangrijk om het document allereerst te voorzien van een unieke ID om het conserveringsplan te kunnen koppelen aan een verzameling digitale objecten en om het document uniek te kunnen identificeren binnen de groeiende verzameling van conserveringsplannen die tezamen het conserveringsbeleid vormen.

STATUS EN AANLEIDING

De status van het document is belangrijk om aan te geven of het document al dan niet in behandeling is genomen, af is geweest, goedgekeurd of uitgevoerd inclusief een datering van deze status. Tot de status behoren ook attributen die duidelijk moeten maken hoe het document zich verhoudt tot andere documenten binnen het conserveringsbeleid.

Bijvoorbeeld wanneer het een aanpassing vormt van een ouder document of wanneer het vervangen is door een nieuwer document.

De aanleiding kan, zoals hierboven beschreven, gevormd worden door meerdere factoren, resultaten van community monitoring, technische ontwikkelingen, signalen vanuit Risk assessment of veranderd beleid.

⁹ Het document CCSDS Audit and Certification of Trustworthy Digital Repositories 2011 dat de richtlijnen bevat voor het behalen van de status 'Trusted digital repository'.

¹⁰ Deze indeling is met enkele aanpassingen ontleend aan Becker et al., Systematic Planning for Digital Preservation, 3.2, <http://www.ifs.tuwien.ac.at/~becker/pubs/becker-ijdl2009.pdf>



BESCHRIJVING VAN RANDVOORWAARDEN

Zoals eerder genoemd bestaan de randvoorwaarden van een conserveringsplan uit een veelheid aan factoren die de besluitvorming beïnvloeden.

In het plan moet tenminste beschreven worden hoe de acties zich verhouden tot ondergenoemde onderdelen.

- Beleidsvoorwaarden
 - Missiestatement
 - Conserveringsbeleid
- Technische voorwaarden
 - Transformational information properties
 - Representatie informatie
 - Preservation Description Information
 - Gebruikte hardware
 - Gebruikte software
 - Software licenties/SLA's
- Gebruikersvoorwaarden
 - Beschrijving/monitoring Designated communities
 - Juridische voorwaarden verbonden aan SA's/OA's
- Procedurele voorwaarden
 - Doorlooptijd
 - Informatie over budgettering

BESCHRIJVING COLLECTIE

In het plan moet verder komen te staan welke objecten binnen het archief het plan betreft. Dit kan door middel van opname van een lijst met id's of een technische beschrijving van de objecten voor zover deze voldoende is om het onderscheid te kunnen maken. Eventueel kan ook verwezen worden naar een voorbeeld-object zodat duidelijk is welk type object bedoeld wordt.

VASTLEGGING PROCES VAN SELECTIE

Dit onderdeel van het plan vormt de vastlegging van de besluitvorming rondom de verschillende scenario's die denkbaar zijn bij de praktische uitwerking rondom het wegnemen van een bepaald risico of het anticiperen op een bepaalde verandering binnen de instelling. Dit om de vereiste inzichtelijkheid te kunnen bieden rondom de geschiedenis van het conserveringsbeleid binnen de instelling.

De onderdelen die hier tenminste genoemd moeten zijn:

- Een korte beschrijving van de verschillende scenario's

- De resultaten van evaluering van de scenario's
- De beslissing die is gemaakt voor een bepaald scenario inclusief redenen
- Het effect dat de beslissing zal hebben op de digitale objecten die het betreft

ROLLEN, VERANTWOORDELIJKHEDEN, KOSTEN

Dit onderdeel moet inzichtelijk maken wie er nodig zijn om het plan uit te voeren, welke rollen er te onderscheiden zijn en wat doorvoering van dit plan kost. Hierbij moet ook in beschouwing genomen worden dat preserveringsplannen doorgaans voor de langere duur zijn geschreven waarbij het dus van belang is dat kosten worden doorberekend op een manier die hiermee rekening houdt zoals in het geval van het model van Total Cost of Ownership.¹¹

PRESERVERINGSACTIEPLAN

Tenslotte volgt het belangrijkste onderdeel van het preserveringsplan; de concrete acties die gevolgd moeten worden om het gestelde doel te bereiken. De beschrijving van wat hier nodig is kan variëren al naar gelang de gekozen oplossing. Het kan simpelweg de implementatie van een tool zijn maar kan ook bestaan uit een workflow met meerdere stappen en verschillende condities. Er moet in dit deel vooral gespecificeerd zijn wat er moet gebeuren en wanneer.

Er moet tenminste beschreven staan:

- welke acties uitgevoerd moeten worden
- wanneer deze acties plaatsvinden
- welke software en/of hardware hiervoor nodig is
- welke mechanismes worden ingezet om te verifiëren of de acties geslaagd zijn
- welke rollen verantwoordelijk zijn voor uitvoer van de acties.

¹¹ http://en.wikipedia.org/wiki/Total_cost_of_ownership

USE CASE: FIXITY-INFORMATIE VOOR DPX-BESTANDEN: EEN PRAKTIJKVOORBEELD VAN PRESERVATION PLANNING

1. IDENTIFICEREN AANLEIDING

Aanleiding

De aanleiding voor dit conserveringsplan wordt gevormd door het feit dat er bij DPX-bestanden binnen de huidige infrastructuur geen fixity-informatie¹² wordt opgeslagen. De noodzaak om fixityinformatie op te slaan bij de objecten is binnen het conserveringsbeleid geformuleerd als een belangrijke voorwaarde om een garantie van integriteit te kunnen bieden aan de Designated communities.

Het garanderen van de fixity van objecten door het aanmaken van checksums en het opslaan hiervan als conserveringsmetadata is de gekozen strategie binnen de instelling om aan het conserveringsbeleid te kunnen voldoen.

Het conserveringsplan moet tenslotte duidelijk maken hoe fixity-informatie praktisch gezien uiterekend en opgeslagen kan worden bij objecten waarbij dit nu nog niet gebeurt.

Risico

Het risico is dat als de wijziging niet plaatsvindt DPX-bestanden corrupt kunnen raken zonder dat dit gedetecteerd wordt. Een ander risico is dat de Designated communities de objecten niet zullen beschouwen als integer wanneer hier geen bewijs van geleverd kan worden. Een risico op hoog niveau is tenslotte dat het archief het gestelde conserveringsbeleid niet waar kan maken.

Doel

Het doel is om extra fixityinformatie op te slaan, gekoppeld aan de DPX-bestanden en opgeslagen als conserveringsmetadata zodat fixityinformatie van het gehele object bijgehouden kan worden bij kopieer- en migratieacties en deze ook uitgeleverd kan worden aan gebruikers.

Impact op digitale objecten

Dit plan heeft vooral invloed op fixity-garanties met betrekking tot DPX-bestanden. De actie betreft een aanvulling van de metadata. De data van de essence zelf blijft daarmee ongewijzigd.

¹² Onder fixity-informatie wordt binnen deze use case een Message digest verstaan, oftewel het resultaat van een cryptografische functie die als invoer de data van een digitaal object heeft en als resultaat een nagenoeg unieke reeks karakters oplevert. Wanneer de invoer van een volgende controleactie ook maar 1 bit verschilt met het origineel levert dit doorgaans een volstrekt afwijkende reeks karakters op. Met dit mechanisme kan dus gecontroleerd worden of bestanden tijdens opslag of kopieeracties niet onbedoeld veranderd zijn.



2. IDENTIFICEREN COLLECTIE

De objecten waarop dit plan van toepassing is wordt gevormd door de collectie van gedigitaliseerd filmmateriaal dat beschikbaar is in DPX-formaat. Het betreft zo'n vijftienduizend titels.

3. BEPALEN RANDVOORWAARDEN VOOR OPLOSSING:

Techniek

- Hardware: de informatie moet binnen bestaande storagemanagementomgevingen opgeslagen kunnen worden, de checksums moeten uitgerekend kunnen worden zonder aanschaf van extra hardware.
- Software: de checksums moeten uitgerekend en opgeslagen kunnen worden aan de hand van bestaande software dus zonder het schrijven van custom code. De oplossing mag uitlevering/transcodering van de DPX-bestanden niet bemoeilijken. De checksums moeten geëxporteerd kunnen worden om opgeslagen te kunnen worden als conserveringsmetadata.

Gebruikers

- Fixityinformatie moet geëxporteerd kunnen worden zodat gebruikers de integriteit van de digitale objecten kunnen evalueren.

Procedureel

- Het uitrekenen van checksums moet geautomatiseerd kunnen worden, het moet uitgevoerd worden zonder onbeschikbaarheid van de collectie, er mogen geen kosten per file aan verbonden zitten alleen kosten voor de server/software, het mag geen zware belasting vormen voor de huidige infrastructuur. Het moet op korte termijn doorgevoerd kunnen worden.

Beleid

- Het archief moet kunnen aantonen dat bestanden niet onbedoeld veranderd zijn gedurende de tijd dat ze opgeslagen zijn of bij kopieer- of migratieacties.



4. OPSTELLEN MOGELIJKE SCENARIO'S

Bij het bedenken van mogelijke scenario's gelden in ieder geval bovenstaande overkoepelende eisen aan de te kiezen oplossing. Het uitgangspunt voor mogelijke oplossingen is dat er fixity-informatie gegenereerd, opgeslagen en uitgeleverd moet kunnen worden binnen de bestaande infrastructuur en dat deze acties uitlevering en transcoding niet mogen vertragen of bemoeilijken.

De uitwerking van de oplossing en welke infrastructuur daarbij gebruikt moet worden verschilt echter per scenario.

Scenario 1: Fixityinformatie opslaan in de TAR op Stornext:

Binnen dit scenario wordt de huidige locatie van de DPX-bestanden en het huidige TAR-packaging format gebruikt om fixity-informatie toe te voegen. Op basis van de huidige informatie lijkt het dat de software dit niet 'natively' kan en dat het genereren en opslaan van de checksums een proces is dat zelf gebouwd moet worden. Dit kan op korte termijn in gang gezet worden omdat de bestaande locatie gebruikt wordt.

Scenario 2: DPX-bestanden opslaan als AXF binnen DivArchive 7

Binnen dit scenario worden de DPX-bestanden in een nieuw packaging formaat gestopt, namelijk AXF, zoals ondersteund binnen DivArchive 7. De bestanden worden daarmee binnen een andere technische oplossing beschikbaar gesteld. Checksums maken een integraal deel uit van het ingestproces en worden binnen het AXF-formaat opgeslagen door DivArchive. Dit kan niet op korte termijn gebeuren aangezien een versieupdate van DivArchive vereist is om dit mogelijk te maken.

5. KWANTIFICEREN SCENARIO'S

Beide scenario's hebben voor- en nadelen en geen van beide vervullen alle gestelde randvoorwaarden precies. Bij het kwantificeren van de scenario's moet dus ook een weging van de randvoorwaarden meegenomen worden.

Scenario 1 heeft als voordeel dat deze sneller te realiseren lijkt. Strategie 2 kan nog een paar jaar duren zodat een oplossing in de tussentijd niet mogelijk is.

Scenario 2 is daarentegen haalbaarder doordat het genereren van checksum native ondersteund wordt binnen DivArchive 7 en er geen custom aanpassingen nodig zijn om het genereren van checksuminformatie en het controleren hiervan bij kopieeracties mogelijk te maken.

Scenario 1 heeft als nadeel dat deze gebruik maakt van de huidige infrastructuur waardoor een risico is dat uitvoer van dit scenario een te grote last legt op de huidige infrastructuur.

6. HET UITWERKEN TOT EEN PRESERVERINGSPLAN

Op basis van de kwantificering in de vorige stap wordt een keuze gemaakt voor het beste scenario. In dit geval lijkt scenario 2 de meest haalbare omdat het de meeste voordelen biedt. Na weging van de randvoorwaarden kan bijvoorbeeld besloten zijn dat het niet hoeven doen van softwareaanpassingen in een bestaande oplossing belangrijker is dan het feit dat de oplossing niet op korte termijn mogelijk is. Doordat de keuze van scenario's en de weging van de randvoorwaarden op papier is gezet vormt deze daarmee ook meteen een verantwoording van de besluitvorming rondom preservering op basis waarvan altijd achteraf aangetoond kan worden dat er gehandeld is in overeenstemming met het geldende preserveringsbeleid en geldende strategieën.

Het scenario kan nu als aanbeveling gestuurd worden naar Administratie ter goedkeuring. Als de aanbeveling wordt omgezet in een migratiedoel kan deze vervolgens verder uitgewerkt worden tot een preserveringsactieplan.

Aangezien het hier slechts een voorbeeld betreft zal het actieplan hier verder niet uitgewerkt worden omdat de details die benodigd zouden zijn om een bruikbaar plan op te leveren in dit geval nog ontbreken.

Bij het opstellen van een preserveringsactieplan zou er bijvoorbeeld in dit stadium al meer duidelijkheid moeten zijn over precieze kosten en concrete acties die uitgevoerd moeten worden inclusief rollen en verantwoordelijkheden. Ook kan de keuze voor een bepaald scenario weer extra onderzoek met zich meebrengen. In dit voorbeeld is het bijvoorbeeld noodzakelijk om bij de uitwerking van scenario 2 te verifiëren dat het veranderen van packaging format of het updaten van de software van DivArchive 7 bijvoorbeeld geen nadelen met zich meebrengt. Het vooraf testen van het scenario zou in dit geval dus ook een noodzakelijke stap moeten vormen binnen het proces.

Ondanks de onvolledigheid van dit voorbeeld illustreren bovenstaande stappen hopelijk voldoende het denkproces en de besluitvorming rondom preserveringsplanning zoals deze uiteindelijk in praktijk gebracht moet worden en vastgelegd binnen een preserveringsplan.