



Pr s rv r n

Stappen zetten in
een nieuw vakgebied

Jaarboek 19

redactie

Margriet van Gorsel, Erika Hokke, Bart de Nil, Marcel Ras



GOED VOOR UW LOOPBAAN

BENEFICIAL FOR YOUR CAREER

De Hogeschool van Amsterdam leidt archiefprofessionals op binnen het hoger beroepsonderwijs. Daarnaast verzorgen we na- en bijscholing voor de archiefsector.

Kenmerken:

- ▶ onafhankelijk
- ▶ professioneel
- ▶ praktijkgericht

Onze expertise:

- ▶ digitaal archief- en informatiebeheer
- ▶ duurzame toegankelijkheid
- ▶ privacybescherming en openbaarheid
- ▶ archiefrecht en auteursrecht

Maatwerk

Wat kan scholing voor uw organisatie betekenen? Op hva.nl/archiefschool vindt u het *complete cursusaanbod*.
Neemt u gerust contact op voor een oriënterend gesprek.

Contact

☎ 06 - 211 58 210

✉ archiefschool@hva.nl

🌐 hva.nl/archiefschool

HVA.NL/ARCHIEFSCHOOL

CREATING TOMORROW

Preserveren
Stappen zetten in een nieuw vakgebied

Preserveren
Stappen zetten in een nieuw vakgebied

Redactie: Margriet van Gorsel, Erika Hokke, Bart de Nil, Marcel Ras

© 2019 de auteurs en redacteuren

ontwerp - www.absoluutdesigners.com

ISBN/EAN 978-90-71251-47-4
Stichting Archiefpublicaties, 's-Gravenhage, 2018

Deze uitgave is mogelijk gemaakt door:
Archiefschool/Hogeschool van Amsterdam (pagina 2 en 8)
Picturae (pagina 6 en 8)
Axiell ALM Netherlands (pagina 8 en 10)
De Ree archiefsystemen (pagina 8 en 219)
STROOM in, advies en opleidingen (pagina 8)

Preserveren

Stappen zetten in een nieuw vakgebied

redactie

Margriet van Gorsel

Erika Hokke

Bart de Nil

Marcel Ras

Jaarboek 19
Stichting Archiefpublicaties

s@p



PICTURAE



uw partner, in iedere fase

Ten geleide

Zo'n anderhalf jaar geleden stelde het bestuur van de Stichting Archiefpublicaties via twitter de vraag 'waar zou het volgende jaarboek over moeten gaan?' Per kerende tweet volgden de suggesties: 'e-depot', 'duurzaamheid', 'preserving', digitale toegankelijkheid', 'digital curation', vanzelfsprekend voorzien van uitroeptekens en soms getypt in kapitalen. Het bestuur zag een informatiebehoefte, dit jaarboek Preserveren is het resultaat.

De artikelen in dit jaarboek bewegen tussen 'vroeger en nu' en behandelen zowel het '(be)denken als het doen', de theorie en de praktijk. Beide weerbarstig, toch laten de artikelen vooruitgang zien: e-depots worden ingericht, kennis ontwikkeld en ervaringen gedeeld, nagedacht over (toekomstige) samenwerking en rol/taakverdeling, gepeinsd over andere benaderingen en nieuwe strategieën. Er lijkt een nieuw vakgebied 'digitale preserving' te ontstaan, waarin verschillende disciplines elkaar vinden en versterken. Een vakgebied waarin overigens nog veel discussie is en zelfs de term 'preserving' niet eenduidig gedefinieerd lijkt te zijn.

Ruimte geven aan nieuwe ontwikkelingen in het vakgebied, deze beschrijven en discussies hierover stimuleren is de primaire doelstelling van de Stichting Archiefpublicaties. Met dit jaarboek Preserveren hoopt de stichting de nodige inspiratie te geven aan iedereen die in de dagelijkse praktijk wil realiseren dat digitale informatie duurzaam toegankelijk blijft.

Dit jaarboek was er niet geweest, als niet een flink aantal mensen heel hard aan het werk waren gegaan. Namens het bestuur van de Stichting Archiefpublicaties dank ik Margriet van Gorsel, Marcel Ras en Bart de Nil voor hun tomeloze inzet en enthousiasme voor dit jaarboek. Als redactie waren zij innovatief en inspirerend en wisten zij een grote schare auteurs te bewegen een bijdrage te leveren. De auteurs verdienen minsten zoveel dank. Binnen strakke deadlines en met een kritische redactie leverden zij artikelen die gezamenlijk een mooi beeld geven van het 'denken en doen' in dit nieuwe vakgebied. Een dankwoord aan Marc Meeuwissen, die in hele korte tijd alle teksten en plaatjes omvormde tot een digitale publicatie. Tot slot natuurlijk een woord van dank aan de adverteerders, die de uitgave van dit jaarboek mede mogelijk hebben gemaakt.

Een jaarboek dat overigens alleen digitaal verschijnt, want waarom zou je een jaarboek over duurzame toegankelijkheid van digitale informatie op papier uitgeven?

Ik wens u veel leesplezier en een duurzame digitale toekomst,

Erika Hokke
Secretaris Stichting Archiefpublicaties

Deze uitgave is mogelijk gemaakt door:



Axiell ALM Netherlands

Safariweg 18-22, 3605 MA Maarssen
Postbus 1436, 3600 BK Maarssen
T +31 (0) 346-5 86 800
ALM.info@axiell.com
www.alm.axiell.com/nl/



Erg goed in Erfgoed

DE REE archiefsystemen

Lijnbaanstraat 2, 9711 RV Groningen,
The Netherlands
T +31 (0)50-3112600
info@de-ree.nl
www.de-ree.nl



Hogeschool van Amsterdam

Hogeschool van Amsterdam | Archiefschool Faculteit Digitale Media en Creatieve Industrie

Postbus 1025, 1000 BA Amsterdam,
The Netherlands
T +31 (0)6-21158210
archiefschool@hva.nl
www.hva.nl/archiefschool



Picturae

Droogmakerij 12, 1851 LX Heiloo,
The Netherlands
T +31 (0)72-5320444
contact@picturae.com
www.picturae.com

● STROOM IN ●

STROOM in, advies en opleidingen

T +31(0)6 4705 6871
info@stroomin.nl
www.stroomin.nl
twitter: @stroomin

Inhoud

~ Woord vooraf 11

HOOFDSTUK 1

Pioniers in e-land

- Inleiding [17](#)
Hans Hofman Een wereld van verschil? Van MLG's naar duurzame toegankelijkheid [19](#)
Barbara Sierman OAIS: Keurslijf of kans? [36](#)
Martin Wrigley, Becky McGuinness, Carl Wilson The Open Preservation Reference Toolset. Extracting meaning from stored bits and bytes [45](#)
Barbara Sierman, Margriet van Gorsel, Remco van Veenendaal Preservering: van Beleid naar Beheer [55](#)
Barbara Sierman, Marcel Ras Certificering van digitale archieven in Nederland [65](#)

HOOFDSTUK 2

Ontwerpen, bouwen en verbouwen

- Inleiding [75](#)
Annemieke de Jong Van proefondervindelijk naar gestuurd. Duurzaam preserveren bij het Nederlands Instituut voor Beeld en Geluid [77](#)
Afelonne Doek, Donald Weber Social History Portal en het project Heritage of People's Europe (HOPE). Het preserveren van een gezamenlijke inspanning [90](#)
Wim Lowet, Frans Neggers 10 jaar digitale architectuurarchieven: een stand van zaken [96](#)
Gaby Wijers Continue in beweging. Het beheer en behoud van digitale kunst [105](#)
Laurents Sesink Casus Universiteit Leiden: Duurzame toegang tot research data [113](#)
Matthias Priem Een audiovisueel archief voor Vlaanderen. Waar staat VIAA na vijf jaar? [123](#)

HOOFDSTUK 3

Voorbereiden op volgende stappen

- Inleiding [133](#)
Mette van Essen Digital Preservation – wat is het niet (of niet alleen). Tijd voor een andere benadering [135](#)
Annet Dekker Capturing online cultures and storytelling as method [142](#)
Joost van der Nat Generieke voorzieningen en samenwerking [155](#)
Herman Uffen, Tamar Kinkel De kosten van digitale duurzaamheid beheersbaar [163](#)
William Kilbride Minding the Gaps: Digital Preservation then and now [174](#)
Hans Hofman Een wereldbeeld vergruizeld: uitdagingen voor archiveringsdeskundigen [184](#)
Charles Jeurgens De overbrenging voorbij. Het institutionele geheugen van de overheid [201](#)

Bijlage

~ Over de auteurs [215](#)



Webbased collectiebeheer in de cloud

Online collectiebeheer waar en wanneer u maar wilt



Webbased

Hardware
onafhankelijk

Intuïtief

Moderne en
innovatieve interface

Support

inclusief helpdesk-
ondersteuning en updates

Samenwerking

Eenvoudig
data delen

Efficiënt

Op ieder moment
overal beschikbaar

Flexibiliteit

Toekomstbestendige
technologie

Vraag nu uw gratis demo aan op www.axiell.com/nl/collections



an Axiell product

Woord vooraf

Preservering, voer voor discussie

Er is maar weinig in het domein van digitale preservering dat niet tot discussie leidt, het is voer voor collectiebeheerders, IT-ers, bibliothecarissen, webmasters, informatiemanagers en archivariissen. Alleen daarom al is de redactie verheugd over de totstandkoming van dit Jaarboek. Dit Jaarboek is niet alleen een momentopname van dat wat er in de praktijk gebeurt bij collectiebeherende instellingen op het gebied van digitale preservering. Het is ook een overzicht van de ontwikkeling van dit kennisdomein. De gevraagde auteurs komen dan ook uit deze wereld.

Het doel van digitale preservering is om digitale informatie te beheren en te bewaren zodat het authentiek en bruikbaar blijft. Juist omdat discussies over digitale preservering over het algemeen een open einde hebben, de digitale techniek staat immers niet stil, bestaat het risico dat een moderne informatiebeheerder of erfgoed professional het gevoel heeft dat zij op drijfzand staat. De redactie herkent dit gevoel en heeft voor haar indeling van dit Jaarboek gekozen voor het model van Simon Wardley¹ waarin de ontwikkeling van een bedrijf wordt geschetst aan de hand van de ontwikkeling van een gebied en het type mens dat zich in dat gebied ophoudt.

Het model biedt houvast bij het organiseren van bedrijvigheid in een wereld waar sprake is van voortdurende innovatie. Het cyclische model spreekt van 'pioneers', 'settlers' en 'town planners'. Deze typen staan voor de ontwikkeling van het pionieren, de eerste bewoning van een gebied tot aan de planning in de stabiele leefomgeving. Al deze fasen vragen andere soorten gedrag, vaardigheden, werkwijze, en persoonlijkheid. In het eerste hoofdstuk van dit jaarboek komen de verkenner aan bod. De eerste schreden op het gebied van digitaal archiveren, digitale duurzaamheid als serieus onderzoeksgebied, de theorie, het ontwikkelen van standaarden, de internationale projecten en het certificeren van een digitale bewaaromgeving. Het tweede hoofdstuk sluit aan bij het 'Settelen', het geeft een gevarieerd beeld van de eerste bewoners van de digitale wereld. Hierin geven zes collectiebeherende instellingen de staat weer van hun digitale archief, de stand van zaken van de huidige preservering praktijk, de hulpmiddelen en de samenwerking. In het laatste hoofdstuk komen de planners aan het woord die, op basis van de huidige strategieën voor digitale preservering, de methodiek bieden om door te ontwikkelen en ook weer nieuwe ontwikkelingen verkennen.

Digitaal informatiebeheer is meer dan ooit het beheer van informatie en alle technische en logische facetten om informatie leesbaar en bruikbaar te houden. Dit vraagt van collectiebeherende instellingen een innovatieve opstelling en het organisatorische vermogen om planners, bewoners en verkenner naast elkaar te laten leven. Dit Jaarboek wil hen een overzicht geven en een steun in de rug zijn.

¹ Simon Wardley On Pioneers, Settlers, Town Planners and Theft (2015) via <https://blog.gardeviance.org/2015/03/on-pioneers-settlers-town-planners-and.html> geraadpleegd op 21 oktober 2018.

Overwegingen

Om te voorkomen dat de auteurs en redactie in een definitie discussie terecht zouden komen is in dit Jaarboek de definitie van SCAPE² voor digitale preservering als uitgangspunt genomen. Volgens die definitie bestaat digitale preservering uit de volgende onderdelen:

- 1) Bit preservation gedefinieerd als “de werkzaamheden die benodigd zijn om de bit streams intact en leesbaar te houden”.
- 2) Functionele preservering, ook wel omschreven als content preservering of logische preservering: “ (...) is bedoeld om – door de tijd heen – de duurzame toegankelijkheid van digitale bronnen te waarborgen, door via actieve interventie de verstoring te verkleinen die wordt veroorzaakt door technologische veranderingen. Het proces genereert nieuwe technische versies van de bronnen door bijvoorbeeld format migratie. Deze nieuwe versies worden geïncorporeerd in de preservation storage omgeving, voor blijvend behoud op het niveau van de bits. Functionele preservering omvat verschillende preservering strategieën die alle bedoeld zijn om verschillende soorten digitale materialen duurzaam te behouden. Het kiezen van een bepaalde strategie moet goed worden overwogen. De gekozen strategie moet immers de integriteit waarborgen van het materiaal, ook na verschillende preservering acties en er zorg voor dragen dat het ook voor de toekomst toegankelijk is, voor zowel de organisatie als voor de doelgroepen, de Designated Communities.”³

Dit onderscheid tussen passieve en actieve preservering is slechts het begin van de verschillen in reikwijdte en invalshoeken die op het onderwerp digitale preservering van toepassing kunnen zijn. De redactie heeft niet gekozen tussen de brede en de smalle definitie van digitale preservering. De smalle definitie lijkt veel op de bovenstaande definitie van functionele preservering, de brede definitie betreft tevens het geheel van organisatie, personeel, middelen, procesinrichting en kosten. In beide gevallen spelen de preserveringsactiviteiten zich af binnen een organisatie die een bepaalde collectie of informatieverzameling beheert. Dit is goed vergelijkbaar met het beheren van een papieren collectie of informatieverzameling. De controle is geheel in handen van die organisatie.

Tijdens de voorbereidingen van dit Jaarboek was al duidelijk dat deze vorm van controle en *custody* met betrekking tot digitale informatie aan het veranderen is. Er is overduidelijk een beweging zichtbaar van *custodial* naar *post custodial* beheer. De *postcustodial theory of archives* zegt dat “*The idea that archivists will no longer physically acquire and maintain records, but that they will provide management oversight for records that will remain in the custody of the record creators.*”⁴ Daarnaast is een digitaal informatie object geen enkelvoudig, plat document meer, als het dat al geweest is, maar een meervoudig, gelaagd, samengestelde verzameling van informatie componenten. En naast die complexiteit heeft de toename in omvang en hoeveelheid van informatie objecten grote invloed om de mate van beheer van digitale objecten. Hierdoor wordt ook de noodzaak van automatisering van

² Voor meer over SCAPE zie <http://scape-project.eu> geraadpleegd op 21 oktober 2018.

³ Voor beide definities zie http://wiki.ncdd.nl/index.php?title=Duurzaamheidsbeleid/Beleidsuitgangspunt-Bit_Preservering en http://wiki.ncdd.nl/index.php?title=Duurzaamheidsbeleid/Beleidsuitgangspunt-Functionele_Preservering geraadpleegd 21 oktober 2018.

⁴ *Postcustodial theory of archives*, zie <https://www2.archivists.org/glossary/terms/p/postcustodial-theory-of-archives> geraadpleegd 21 oktober 2018.

preservingsactiviteiten met de dag groter of anders gezegd, preservering moet in de productiefase van de ontwikkeling van preservingsystemen komen.

Er is bewust gekozen voor de term ‘preserveren’ om het over de functie te hebben en niet voor het begrip eDepot. De succesvolle lobby voor eDepot’s door archiefinstellingen heeft digitaal archiveren op de kaart gezet. Tegelijkertijd wordt duidelijk dat een eDepot vooral geassocieerd wordt met een digitale bewaaromgeving en niet met het preservingswerk dat in alle omgevingen moet worden uitgevoerd, dat voortdurend monitoring van digitale informatie vereist, dat specialistische kennis vereist en dat zich buigt over de authenticiteit en betrouwbaarheid van de informatie. De aanschaf van een eDepot is de start van de organisatieverandering, van eerste bewoning naar planning van de organisatie die preservering ondersteunt.

Een andere invalshoek die de redactie in overweging heeft genomen is de vraag waar preservering begint in de life cycle van een informatieobject. Zijn alle elementen voor digitale preservering bij het begin van de creatie van een digitaal informatie object beïnvloedbaar? De voorwaarden die van invloed zijn op digitale preservering zijn zeker van belang bij de creatie, opslag, beschrijving, waardering, bewaring voor lange termijn en beschikbaarstelling. De redactie heeft een strikte scheiding gehanteerd tussen het records management in de dynamische fase bij de overheid, in wetenschappelijk onderzoek, in de ontwerpfasen van architectenbureaus, bij het maken van programma’s door de publieke omroep en de creatieve fase van de maker van digitale kunst.

Nationaal en internationaal

Dit jaarboek richt zich op ontwikkelingen in Nederland (heden, verleden en toekomst), maar de ontwikkelingen in Nederland staan niet op zichzelf. Ook internationaal is er van alles gedacht, gedaan en geëxperimenteerd. Al in 1995 schreef Jeff Rothenberg⁵ over de uitdagingen van digitale duurzaamheid.⁶ In die tijd lag de focus nog sterk op de technische uitdagingen. Zoals uit de artikelen in dit Jaarboek zal blijken, is dit verschoven naar een meer organisatorische kwestie. Ook al in 1995 vestigde de Europese Commissie de aandacht op dit onderwerp. Dit was met name gericht op vragen rondom elektronische publicaties en rechtenkwesties. Dit mondde uiteindelijk uit in de Open Content beweging. Europese samenwerkingsprojecten gefinancierd door de EU hebben gezorgd voor een aantal forse stappen voorwaarts. ERPANET, DELOS, PRESTOSPACE en met name PLANETS zorgden voor een brede samenwerking waarin erfgoedinstellingen en onderzoekers aan elkaar gekoppeld werden. Deze projecten zijn van groot belang geweest voor de verdere ontwikkeling van een nieuw vakgebied. Zowel voor de ontwikkeling van software tools en systemen, als voor de bewustwording en kennisontwikkeling.

Er zijn in de loop der tijd kleine en grote samenwerkingsverbanden opgericht. Een belangrijke rol speelden daarbij de Digital Preservation Coalition, PrestoCenter,

⁵ Jeff Rothenberg ‘Ensuring the Longevity of Digital Documents’ in *Scientific American* Vol 272 no 1 (1995) p. 42 Bibcode:1995SciAm.272a..42R. doi:10.1038/scientificamerican0195-42.

⁶ In die tijd spraken we over digitale duurzaamheid, vanaf het begin van de 21e eeuw spreken we over duurzame toegankelijkheid of duurzame toegang.

P I O N E E R S

Pioniers in e-land

Inleiding

“*That’s one small step for [a] man, one giant leap for mankind*”, op 21 juli 2019 zal het een halve eeuw geleden zijn dat Neil Amstrong als eerste mens een voet op de maan zette. Het was een iconische gebeurtenis en het werd dé referentie voor latere ruimtereizen.

Ongetwijfeld zal er op dit moment wel ergens in wereld een *would-be* astronaut dromen over zijn of haar eerste woorden als eerste mens die een voet op Mars zal zetten. We hebben nu eenmaal iconische ankerpunten nodig om ons een perspectief op het verleden, heden en de toekomst te geven. En niet toevallig is dit ook het bruggetje naar de kleine geschiedenis van de pioniers van het digitaal preserveren.

De beginperiode van het digitaal preserveren van erfgoeddata heeft ook een Neil Amstrong-momentje gekend, met de ontwikkeling van de internationale standaard OAIS. En ook hier weer is er de link met de ruimte want deze standaard werd in het midden van de jaren 2000 ontwikkeld door een organisatie voor het duurzaam beheren van data over ruimtereizen. OAIS werd heel snel dé wereldwijde standaard die aanbevelingen geeft bij het inrichten van een archief dat gericht is op langetermijnbewaring van en toegang tot digitale informatie. Het biedt een samenhangend geheel van uitgangspunten en terminologie.

De pioniersfase van e-depotvoorzieningen wordt gekenmerkt door uitproberen, testen, uitproberen, testen, enzovoort. Een ankerpunt dat zeker zorgde voor een versnellingsmoment was het EU-project PLANETS. Gestart in 2006 bracht het PLANETS consortium de expertise van nationale bibliotheken en archieven én enkele toonaangevende onderzoeksinstituten en bedrijven in Europa samen. De resultaten van PLANETS leidde tot een internationale benadering en nieuwe initiatieven. Zoals de Open Preservation Foundation dat de leiding neemt in een gezamenlijke inspanning om een set van duurzame en open source digitale bewaarinstrumenten en ondersteunende bronnen te ontwikkelen.

In deze ontdekkingsperiode beginnen de pioniers zich ook te verenigen vanuit de gedachte dat de uitdagingen en gevaren die samengaan met het ontdekken van het ongerepte maar gevaarlijke *e-land* veiliger is als je dit samen doet. Er ontstaan samenwerkingsvormen zoals DPC, NCDD of nestor waarbij het ene acroniem wat geslaagder is dan het andere. Er worden op het einde van de pioniersperiode ook al stapjes gezet naar het gebruiken van duurzame digitale informatie. Wiki wordt niet enkel een buzz-woord onder techneuten maar het krijgt ook weerklank bij de meer

‘digitaal georiënteerde’ medewerkers in erfgoedinstellingen. Dit geeft aan dat het bewustzijn rond duurzame toegankelijkheid begint te groeien en dat het al wordt gezien als een belangrijk uitdaging die deel zou moeten uitmaken van één van de kerntaken van een collectiebeherende instelling.

Zoals dat gaat met pioniers die zich beginnen te settelen worden er ook enkele regels afgesproken om te vermijden dat er anarchie ontstaat. Dat leidt tot procedures en softwaretools die het mogelijk maken dat erfgoedinstellingen zichzelf kunnen evalueren met aan de horizon de beloning van een certificaat. Die certificering is het ankerpunt voor de settelaars bij het ontwikkelen van een duurzame digitale preservingsstrategie.

Een wereld van verschil?

Van MLG's naar duurzame toegankelijkheid

Inleiding

Tot de publicatie van het rapport van de Algemene Rekenkamer *Machineleesbare Gegevensbestanden. Archivering en beheer bij het Rijk* in 1991, had de archiefwereld zich nog weinig bekommerd om het beheren en bewaren van digitale informatie.¹ Eén van de conclusies van de Rekenkamer was dat er bij het Rijk geen beleid was op dit punt en dat hier dringend aandacht aan besteed moest worden. Reden voor de Algemene Rijksarchivaris om al snel een projectgroep in te stellen die daartoe een eerste voorstel moest doen. Het waren de eerste, voorzichtige stappen op onbekend terrein en het begin van een lange weg naar een nieuwe wereld, met grote consequenties voor erfgoedinstellingen en archiefdiensten in het bijzonder, zowel op gebied van de theorie en praktijk van archiveren en archiefbeheer als voor de wijze van communiceren door en werken bij de instellingen zelf.

In dit artikel geef ik een beknopt overzicht van de belangrijkste ontwikkelingen op het gebied van het denken over en het bewaren en beheren van digitale informatie, in het bijzonder digitale archiefbescheiden, sinds begin jaren '90 van de vorige eeuw. Daarbij ga ik ook in op de rol van de belangrijkste betrokken partijen. Mijn invalshoek is daarbij in eerste instantie die van goede archivering (*recordkeeping*). De situatie bij bibliotheken, musea, datacentra en andere organisaties waar digitale informatie langdurig bewaard moet worden, komt slechts zijdelings aan bod. Het zal duidelijk zijn dat het binnen het beperkte bestek van dit artikel onmogelijk is een volledig beeld te schetsen van wat er de laatste ruwweg dertig jaar is gebeurd op dit terrein. Ik zal mij daarom beperken tot een aantal hoofdlijnen waarbij ik een periodisering toepas, die mij op dit moment het meest geschikt lijkt.

Het onderwerp 'preserving' kent vele definities van heel breed tot nauw geformuleerd, als ook verschillende benamingen, zoals digitale bewaring, 'digital preservation' of 'digital curation'. Het gebruik ervan is niet consistent en kan leiden tot onduidelijkheid. Het geeft aan dat het vakgebied nog in ontwikkeling is. In dit artikel wordt voorlopig een ruime definitie gevolgd, vergelijkbaar met die door het Nationaal Archief wordt gehanteerd: "Preservation is: Het op zodanige wijze vastleggen, bewaren, beheren en beschikbaar stellen van digitale documenten (in de brede zin van het woord), dat deze ook na verloop van tijd raadpleegbaar,

¹ Gepubliceerd ter gelegenheid van het afscheid van de president van de Algemene Rekenkamer, F.G. Kordes op 31 oktober 1991

toegankelijk en authentiek zijn.”² Deze omschrijving omvat ook het creëren van digitale archiefbescheiden of informatie met waarschijnlijk als belangrijkste reden, dat goed onderhoud begint met een goed gemaakt product. Niettemin is deze definitie zo breed dat er vrijwel geen onderscheid meer is met de archief- of archiveringsfunctie (*recordkeeping*³). Dat roept de vraag op waarom een nieuwe term nodig is?

Achtergrond

Zoals hierboven al aangegeven, begon men pas in 1991 serieus na te denken over digitale overheidsinformatie. Een belangrijke reden was dat men nog niet of nauwelijks geconfronteerd was met zogenaamde *machineleesbare gegevensbestanden* (MLG's), zoals digitale archiefbescheiden toen nog genoemd werden. Het potentiële probleem was wel al gesignaleerd, maar had in Nederland nog niet tot actie geleid.⁴ Opvallend in deze context is dat statistische of onderzoeks-gegevensbestanden in Nederland vrijwel geen aandacht kregen. Blijkbaar werden datasets niet als archief gezien.⁵ In de samenvatting van het rapport *Documenten uit de tijd* (1993) wordt dat nog eens benadrukt: “Tot voor kort werd aangenomen dat elektronische gegevensbestanden niet voor permanente bewaring in aanmerking kwamen, of geen archief waren, zoals in het geval van statistische bestanden. In deze situatie is een verandering opgetreden.”⁶ In het buitenland waren vanaf de jaren 60 in landen als bijvoorbeeld de Verenigde Staten, Canada en Zweden, wel al ‘data libraries’ in de nationale archieven opgericht om dergelijke bestanden te beheren. Opvallend is wel dat deze ‘datasets’ ook daar niet als archiefbescheiden, maar wel als te beheren objecten gezien werden. Vandaar ook de term ‘libraries’. Deze datasets waren resultaten van de verwerking van geautomatiseerde gegevensbestanden, die zelf niet bewaard werden, net zomin als de documentatie die iets zei over wat de verwerking precies was geweest. Pas begin jaren 80 kwam daar verandering in. In Nederland was in 1964 wel het Steinmetzarchief opgericht voor het beheer van gegevens met betrekking tot sociaal onderzoek, maar dat was buiten archiefdiensten om. In 1989 kwam daar het Nederlands Historisch Data Archief (NHDA) bij voor gegevensbestanden voortkomend uit historisch onderzoek.⁷ Binnen de International Council on Archives (ICA) was al in 1972 een *Committee on Automation* ingesteld, die zich in het begin vooral bezig hield met het opstellen van “Guidelines for Administering Machine-readable Archives”.⁸ Ook hier ging het om wat genoemd werd ‘automatic data processing’ (ADP).

² Preservation policy, p. 6. Zie ook: <https://kia.pleio.nl/groups/view/41371832/kennisplatform-preservation/blog/view/52308822/een-kwestie-van-definitie-geraadpleegd-5-juli-2017>. Een ander voorbeeld van een definitie is: “Het geheel van activiteiten dat zorgt voor het technische en intellectuele behoud van digitale informatieobjecten.” Zie <https://labyrinth.riensjonker.nl/lexicon/terminologie-term-preservering-geraadpleegd-5-juli-2017>

³ De term wordt hier gebruikt in de Australische betekenis: “Making and maintaining complete, accurate and reliable evidence of business transactions in the form of recorded information.” (zie AS 4390 Part 1 Clause 4.19). Dit omvat zowel creatie als beheer van archiefbescheiden, en het ontwerp, de invoering en werking van archiveringsystemen

⁴ Zie bijvoorbeeld het rapport van de Rekenkamer *Archiefbeheer en -behoud bij het Rijk*, (Den Haag, 1988, Tweede Kamer stuk nr. 20665/2), p.24-25. Daar wordt het ontbreken van enig beleid op dit terrein geconstateerd, zowel bij het ministerie van Welzijn, Volksgezondheid en Cultuur (WVC) / Rijksarchiefdienst als bij BiZa.

⁵ In het Nederlands Archiefblad van die jaren, toch het vakblad bij uitstek, is hierover ook niets te vinden.

⁶ Rapport van de Projectgroep MLG (fase 2A) *Documenten uit de tijd* (Den Haag 1993), p.1.

⁷ Het NHDA met onderzoeksdata in de alfa- en gamma wetenschappen is in 1997 opgegaan in het Nederlands Instituut voor Wetenschappelijke Informatiediensten (NIWI) dat vervolgens in 2005 weer opging in Data >

Het denken over digitale bestanden in het laatste decennium van de 20ste eeuw kan niet los worden gezien van een aantal andere ontwikkelingen op gebied van archieven. De belangrijkste noem ik hieronder.

Automatisering

Allereerst is er de automatisering bij archiefdiensten vanaf begin jaren '80. In de jaren tot ca. 1980 waren het vooral grote computers (mainframes) waar mee gewerkt werd. Daarna kwamen de minicomputers en de eerste microcomputers, later meestal aangeduid als personal computers (PC's zoals Microsoft ze noemde) in zwang. Deze braken echt door in de jaren '90, toen ook het internet en email ter beschikking kwamen. Daarmee ontstond de kantoorautomatiseringsomgeving evenals de grafische user interface (vooral Windows). De toepassing van informatietechnologie verliep tot op zekere hoogte nog tamelijk geleidelijk. Eén en ander betekende dat eigenlijk pas vanaf midden jaren '90 het ontstaan van digitale documenten echt op gang kwam.⁹ Bij de Rijksarchiefdienst waren in de jaren '80 stap voor stap microcomputers ingevoerd en was de automatisering van werkprocessen opgepakt, zoals voor het toegankelijk maken en het beheer van archieven. De ontwikkeling van dergelijke softwareprogramma's stelde de manier van werken met en het denken over archieven nogal ter discussie.

Naar een nieuwe Archiefwet

In deze jaren speelde ook de wijziging van de Archiefwet 1962, waarbij de belangrijkste verandering de verkorting van de overbrengingstermijn van 50 naar 20 jaar was.¹⁰ Dat zou een grote toevloed van (papier) archief tot gevolg hebben. Om deze in goede banen te leiden werd in 1991 het Project Invoering Verkorting OverbrengingsTermijn (PIVOT) ingericht.¹¹ Het is goed te bedenken dat de bestaande overbrengingstermijn van 50 jaar inhield dat men nog vooral met papieren archief te maken had en dat archiefdiensten helemaal achter in de 'keten' zaten. Een andere belangrijke verandering in de nieuwe archiefwet was de toevoeging 'ongeacht de vorm' aan de definitie van archiefbescheiden. Digitale archiefbescheiden kwamen hierdoor nadrukkelijk onder de werking van de archiefwet. Eén en ander betekende dat archiefdiensten in deze jaren zich enerzijds met een inhaalslag bezig moesten houden en anderzijds toekomstgericht bezig moesten houden met een nieuw type archiefbescheiden, waarvan zich toen nog slechts een eerste begin aftekende.

⁹ en onderdeel van het KNAW. Zie voor Steinmetzarchief o.a.: https://www.edata.nl/0403_011209/pdf/0403_011209_7.pdf geraadpleegd 5 oktober 2018.

⁸ Gepubliceerd in 1980, plaats onbekend. Zie ook M.H. Fischbein (ed.), *An introduction to archival automation*, 1981, inclusief bijdragen van A. Arad en M.E. Olsen. Ook werd in 1972 besloten een tijdschrift te beginnen, genaamd *ADPA: Automation-Archives-Informatique*. ADPA= Automatic Data Processing in Archives.

⁹ Naast de proces-ondersteunende maatwerksoftware (informatiesystemen) kwam er geleidelijk aan ook kantoorautomatisering. Eerst beperkt tot tekstverwerking, zoals Wordstar dat eind jaren '80 werd vervangen door Wordperfect (v.4.1) en simpele databaseprogramma's (zoals Dbase3 en 4). In de tweede helft van de jaren '90 kwam daar email bij voor iedereen, al hadden eerder wel enkele mensen toegang tot een email programma als Eudora. Was tot circa 1995 alleen in een MS-DOS omgeving gewerkt, in 1995 werd overgegaan naar een complete Microsoft-omgeving met Windows en MS Office. Deze was gebaseerd op een grafische omgeving gekenmerkt door icoontjes op het scherm en niet langer tekstgestuurde commando's die je moest intypen. En tenslotte was er natuurlijk de komst van internet vanaf circa 1995.

¹⁰ De nieuwe Archiefwet en het bijbehorende Archiefbesluit werden in 1995 van kracht.

¹¹ Zie voor een evaluatie van PIVOT: Rini den Brabander, Hans Poppelaars, Wim de Haart, *De handelingen gewaardeerd. Evaluatie van het Project Invoering Verkorting Overbrengingstermijn (PIVOT)*, Den Haag, 2000, rapport in opdracht van ministerie van OCW.

Veranderende overheid

Intussen waren er ook de nodige ontwikkelingen bij de overheidsadministratie. Eén ervan was een toenemende decentralisatie, waarbij verantwoordelijkheden zo laag mogelijk in de organisatie werden gelegd. Dat had weer gevolgen voor de wijze waarop de DIV (documentaire informatievoorziening) was georganiseerd. Voorts waren er experimenten om de archivering bij de overheidsadministratie zelf te verbeteren. De ideeën die daarbij een belangrijke rol speelden zijn vastgelegd in de publicatie *Omslag in Opslag. Terugkeer naar informatievoorziening*.¹² De benadering stelde de rol van DIV en het gebruik van termen als archief of archiefbescheiden ter discussie en legde de nadruk op welke belangen bij het bewaren van informatie speelden.

Archivistische theorie op de schop

Nieuwe ontwikkelingen in het buitenland op het gebied van archieftheorie drongen steeds meer door. Een belangrijk rol speelde hierin bijvoorbeeld David Bearman die de archiefwereld opschudde met toen voor de gemiddelde archivaris nogal 'revolutionaire' ideeën en onder andere de bestaande archivistische methoden als ook de rol van archiefdiensten ter discussie stelde.¹³ Zo was er onder zijn leiding het zogenaamde Pittsburgh project, dat in 1996 met een set functionele eisen 'for evidence in recordkeeping' kwam.¹⁴ Ook aan de andere kant van de wereld, in Australië, waren er baanbrekende ontwikkelingen met betrekking tot de archivistische theorie. In 1996 verscheen de eerste publicatie over het 'records continuum', waarbij het traditionele denken met een centrale plaats voor het fysieke archief en de indeling in dynamisch, semi-statisch en statisch, in een ander perspectief werd geplaatst.¹⁵ Het zou het begin inluiden van wat wel een paradigmaverschuiving in de archiefwetenschap genoemd wordt. Deze nieuwe ideeën bleken bovendien bij uitstek geschikt in de nieuwe digitale wereld.

Al deze ontwikkelingen hadden tot gevolg dat er wereldwijd veel discussie ontstond over hoe de archiveringsfunctie het beste kon worden vormgegeven.

Beginjaren: oriëntatie en discussie (1990-2000)

In deze context kwam begin jaren '90 de uitdaging van machineleesbare gegevensbestanden (MLG's) om de hoek kijken. Uit het eerdergenoemde Rekenkamerrapport bleek dat er bij de verantwoordelijke ministeries, Binnenlandse Zaken (destijds

¹² Ministerie van Binnenlandse Zaken *Omslag in Opslag* (Den Haag, 1991)

¹³ Zie onder andere David Bearman and Margaret Hedstrom "Reinventing Archives for Electronic Records: Alternative Program Delivery Options" in: Margaret Hedstrom ed. *Electronic Records Management Program Strategies* (Pittsburg, Archives & Museum Informatics, 1993) p. 82-98 zie http://www.archimuse.com/publishing/archival_strategies/footnotes.html#ixzz59omNCN4W geraadpleegd 5 oktober 2018; en David Bearman, 'New models for management of electronic records by archives' in: *Cadernos BAD* (Journal of the Portuguese Association of Librarians, Archivists and Documentalists, 1992, vol. 2) p. 61-70 herdrukt in 1994, zie http://www.archimuse.com/publishing/electronic_evidence/ElectronicEvidence.Ch10.pdf geraadpleegd 5 oktober 2018; en David Bearman 'Archival Methods. Archives and Museums Informatics Technical Report #9' in: *Archives and Museum Informatics* (Pittsburg, 1989) zie http://www.archimuse.com/publishing/archival_methods/index.html#ixzz2yGcxM4JW geraadpleegd 5 oktober 2018.

¹⁴ Zie onder andere <http://www.archimuse.com/papers/nhprc/>. De oorspronkelijke website van de University of Pittsburgh is helaas verloren gegaan (!).

¹⁵ Een belangrijk denker in deze is Frank Upward, maar daarnaast speelden ook grootheden als Sue McKemmish, Barbara Reed, Livia Iacovino, Chris Hurley een grote rol in de verdere uitwerking en toepassing van deze ideeën. Zie voor meer informatie: https://en.wikipedia.org/wiki/Records_Continuum_Model. >

afgekort als BiZa)¹⁶ en OCW/Rijksarchiefdienst, geen enkel inzicht was van wat er bij de (rijks)overheidsadministratie gebeurde en hoe de stand van zaken was met betrekking tot automatisering en de productie van digitale bestanden. Dit leidde ertoe dat de Algemeen Rijksarchivaris een MLG-project startte, met als doel om op dit gebied tot een beleid te komen. Centraal in de eerste fase van het MLG-project stonden dan ook drie onderwerpen dan wel vragen:

- 1) wat is de stand van zaken met betrekking tot de informatievoorziening bij de overheid, de automatisering en het ontstaan van digitale bestanden (welke soorten zijn er)?
- 2) wat is nu precies het probleem, en dan vooral wat is nu eigenlijk het verschil tussen papieren en digitale archiefbescheiden?
- 3) wat betekent dat voor de rol van de Rijksarchiefdienst en archiefdiensten in het algemeen? Wat zou die moeten zijn en welk beleid met betrekking tot digitale bestanden was gewenst?

Vooral het tweede punt leidde tot veel discussie en analyses, waarbij het besef ontstond dat in een digitale omgeving het fysieke (de bits en bytes, hard- en software) en het logische (of intellectuele) niet langer gefixeerd waren, zoals bij papier, maar gefragmenteerd en tot op zekere hoogte ontkoppeld. De software interpreteert de bits en bytes en presenteert de inhoud (bijvoorbeeld het document) via een verwerkingsproces op het scherm. Dat leidde er uiteindelijk toe dat er een vijftal essentiële elementen of aspecten werden onderscheiden die onlosmakelijk deel uitmaken van een document in welke vorm dan ook, nl. inhoud, structuur, verschijningsvorm, context en in een aantal gevallen gedrag. Deze elementen of kenmerken komt men steeds weer tegen als het gaat om de vraag wat er nu eigenlijk bewaard worden?¹⁷ Het MLG-project produceerde drie rapporten, die achtereenvolgens de probleemstelling en een begrippenapparaat formuleerden, een schets van de 'gewenste situatie voor het beheer van MLG's' gaven en een beleidsplan bevatten.¹⁸ De belangrijkste aanbevelingen waren dat

- er beter inzicht verkregen moest worden in wat er bij de overheidsadministratie gebeurde;
- er aan de bewustwording zowel bij overheidsorganisaties als archiefdiensten gewerkt moest worden; en
- er een 'technische infrastructuur' voor het beheer zou moeten komen in eerste instantie gericht op tekstbestanden en databases.

> Het 'continuum'-denken bouwde onder andere voort op de ideeën van Peter Scott in de jaren 1960, waarbij het fysieke en het intellectuele aspect van archieven uit elkaar werden gehaald en het zogenaamde 'series-system' ontstond.

¹⁶ In 1998 werd dit het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties: BZK.

¹⁷ Wanneer deze kenmerken voor het eerst gebruikt werden, is niet duidelijk, maar drie ervan komen onder andere voor in een definitie van een *record* in de *Guide for managing electronic records from an archival perspective* (1997), van het ICA Committee on electronic records, p.22: "A record is recorded information produced or received in the initiation, conduct or completion of an institutional or individual activity and that comprises content, context and structure sufficient to provide evidence of the activity. Deze drie kenmerken zijn later aangevuld met twee andere: verschijningsvorm en gedrag." Zie bijvoorbeeld Jeff Rothenberg en Tora Bikson, *Digital Preservation. Carrying authentic, understandable and usable digital records through time*, (programma Digitale Duurzaamheid, Den Haag, 1999) p. 14.

¹⁸ Die rapporten waren achtereenvolgens: *Rapport Fase 1: Probleemstelling* (1991); *Documenten uit de Tijd (fase 2a)* (1993); en *Het papieren tijdperk voorbij. Beleid voor een digitaal geheugen (fase 2b)* (1995). Was de eerste fase nog een project van de Rijksarchiefdienst alleen, de volgende fasen werden gekenmerkt door samenwerking met ook de archiefdiensten van de vier grote steden, de KVAN en de KALO.

Ook het ministerie van Binnenlandse Zaken was aangesproken door het Rekenkamer rapport en daar wilde men eerst inzicht krijgen in wat er elders in de wereld gebeurde. Dat leidde tot een onderzoek in 1993 door EAC/RAND Europe in opdracht van BiZa, waarbij ook de Rijksarchiefdienst werd betrokken. Daarbij werden de nationale archiefdiensten en enkele overheidsinstellingen in Canada, de Verenigde Staten, Zweden en Duitsland bezocht, als ook ministeries en uitvoerende diensten in Nederland geïnterviewd. Het leidde tot het rapport *Preserving the present. Toward viable electronic records* waarin op basis van organisatorische, archivistische en technologische aspecten de situatie in Nederland en elders werd beschreven.¹⁹ De drie belangrijkste constateringingen waren niet veel anders dan wat elders ook werd gezegd:

- er is geen beleid met betrekking tot het beheer van digitale archiefbescheiden geformuleerd en daar zou snel werk van gemaakt moeten worden door organisaties zelf, maar ook via regelgeving
- er moet snel begonnen worden met ontwikkelen van ondersteunende software voor beheer van digitale archiefbescheiden, en
- de traditionele ('papieren') werkwijze voldoet niet meer, nieuwe methoden moeten worden ontwikkeld die meer recht doen aan 'interactive information environments'.²⁰

Op al deze punten werd de volgende jaren actie ondernomen.

Op weg: digitale duurzaamheid

Midden jaren '90 kwam het tot nauwere samenwerking tussen OCW/Rijksarchiefdienst en BiZa.²¹ Eind 1995 leidde dit tot de instelling van een gezamenlijk programma Digitale Duurzaamheid (DD). Ook de andere bestuurslagen participeerden in dit programma, door het beschikbaar stellen van gelden en mensen. Het programma maakte deel uit van de, door beide staatssecretarissen gezamenlijk uitgebrachte, derde Beleidsnota Informatiebeleid Openbare Sector (BIOS-3) getiteld *Terug naar de Toekomst*. Thema's daarin waren naast digitale duurzaamheid, o.a. toegankelijkheid van overheidsinformatie, overheidsloket 2000 en elektronische burgerparticipatie.

Het programmabureau Digitale Duurzaamheid richtte zich primair op overheidsorganisaties met als doel goede digitale archivering te stimuleren. Een kader hiervoor was het zgn. *Proefreglement* dat in 1996 was gepubliceerd. Deze voorloper van de latere Archiefregeling gaf in tien artikelen richtlijnen voor de duurzaamheid, toegan-

¹⁹ Het rapport, geschreven door T.K. Bikson en E.J. Frinking (1993) van het European-American centre for policy analysis (EAC/RAND) is ook in Nederlandse vertaling gepubliceerd onder de titel: *'Het heden onthouden'*.

²⁰ Zie T.K. Bikson en E.J. Frinking, *Preserving the present*, (Den Haag, 1993), p. 16-17.

²¹ Brief van Staatssecretarissen van BiZa en OCW d.d. 27-11-1995 met stand van zaken over het project 'Digitale Duurzaamheid', aan Tweede Kamer, nr. 20644/25.

²² Deze regeling is in 1998 op enkele punten herzien en vormde de basis voor de latere formele Regeling Geordende en toegankelijke staat archiefbescheiden (2001), ex art. 12 van de Archiefbesluit 1995. In dit reglement was voorlopig gekozen voor het PDF-bestandsformaat. Inmiddels is deze regeling vervangen door de Archiefregeling (2009), zie <http://wetten.overheid.nl/BWBR0027041/2014-01-01>.

²³ Zie voor een overzicht van activiteiten bijvoorbeeld de tweede voortgangsrapportage over de BIOS 3 nota aan de Tweede Kamer, november 1997.

²⁴ De grote archiefdiensten waren: Rijksarchiefdienst, en vier grote gemeentearchieven Amsterdam, Den Haag, Rotterdam en Utrecht.

kelijkheid en geordende staat van digitale archiefbescheiden.²² Vanuit het programmabureau werden pilot-projecten gevolgd en begeleid, voorlichting gegeven en workshops georganiseerd, gewerkt aan functionele eisen voor informatiesystemen, en onderzoek gedaan naar standaarden voor bestandsformaten.²³ Ook fungeerde het als kenniscentrum dat handzame boekjes publiceerde, zoals over email archivering, de nieuwe rol van de archieffunctie, digitale archivering, en juridische bewijskracht. Het bevorderde de bewustwording en het denken over digitale duurzaamheid bij de overheidsadministratie, al zou de verandering slechts traag op gang komen.

Uit de rapporten van het MLG-project volgde ook de ontwikkeling van een 'technische infrastructuur' (incl. een digitaal depot). De voortgang was ook hier echter traag. Sinds 1996 waren de grote archiefdiensten bezig een digitaal depot (aanvankelijk 'Digitale Kluis' genaamd) te ontwikkelen, in eerste instantie voor tekstbestanden om ervaring mee op te doen.²⁴ In 1998 schreef de Rijksarchiefdienst, samen met de grote gemeentearchieven, een projectplan voor de bouw van een digitaal depot. Het project is niet van de grond gekomen, vermoedelijk omdat er geen subsidie verkregen kon worden uit het Nationaal Actieprogramma Elektronische Snelweg (NAPES). Wel werd een eerste ontwerp voor een digitaal depot ('depot 2000') gemaakt en gepubliceerd, dat onder meer wegens gebrek aan financiën voorlopig zonder gevolg bleef.²⁵

De Rekenkamer constateerde in haar evaluatierapport van begin 1998 dat een en ander sinds 1991 slechts langzaam op gang gekomen was en dat pas in 1997 echt voortgang geboekt werd.²⁶ Het was allemaal nogal beleidsmatig en gericht op bewustwording, weinig concreet.

Per 1 juli 1998 werd het programmabureau Digitale Duurzaamheid opgeheven. Het programma werd voortgezet door zowel BiZa/ACIB, als OCW/ de Rijksarchiefdienst ieder onder hun eigen verantwoordelijkheid.²⁷ Het opheffen van het bureau leidde tot versplintering van activiteiten en verdwijnen van kennis. BiZa richtte zich op de vraag hoe de *archieffunctie* in een digitale wereld gepositioneerd en vorm gegeven zou moeten worden, om daarmee overheidsorganisaties te kunnen ondersteunen bij de inrichting van hun (digitale) archiefvorming, terwijl de Rijksarchiefdienst opdracht gaf tot een onderzoek naar langdurige bewaring van digitale informatie.²⁸ Het daaruit voortkomende advies leidde ertoe dat in oktober 2000 een nieuw project werd gestart waarin de nadruk zou liggen op het doen van experimenteren en ontwikkelen van praktische richtlijnen, het Testbed digitale bewaring. Dit was opnieuw een samenwerking tussen BiZa en Rijksarchiefdienst.²⁹

²⁵ Gepubliceerd door de Rijksarchiefdienst: Hans Hofman, Nico van Egmond, Jacqueline Slats, Tamara van Zwol, *Depot 2000. Functional design for a digital depot*, (Den Haag, 2000).

²⁶ Algemene Rekenkamer, *Beheer en archivering van digitale bestanden*, rapport aan de Tweede Kamer, nr. 25970/1, (Den Haag, 1998).

²⁷ Het programmabureau werd opgeheven na amper 2 jaar. Het besluit het programmabureau op te heffen was gebaseerd op een rapport van Het Expertise Centrum, waarin werd geconstateerd dat in deze constructie de eigen verantwoordelijkheden van de twee ministeries te weinig tot hun recht kwamen. BiZa kreeg de coördinatie van de activiteiten onder de vlag van Digitale Duurzaamheid.

²⁸ Het eerste leidde tot de publicatie van het boekje *Het geheugen als actieve kracht. De archieffunctie binnen de digitale overheid* (Den Haag, 1999), het tweede tot een rapport van Jeff Rothenberg getiteld: *Digital Preservation. Carrying authentic, understandable and usable digital records through time*, (Den Haag, 1999). Beide waren publicaties onder de vlag van digitale duurzaamheid. De Rijksarchiefdienst bleef ook bezig met regels voor digitale archivering en met het opzetten van een kenniscentrum.

²⁹ Het werd na ruim één jaar ondergebracht bij ICTU (ICT-uitvoeringsorganisatie ingesteld door BZK). Vanaf 1 juli 2003 werd het project weer ondergebracht bij het Nationaal Archief. Het eindigde begin 2004. Elders in dit boek wordt op dit Testbed nader ingegaan.

Al deze projecten probeerden met behulp van experimenten te komen tot praktische bewaarmethoden van digitale informatie. Men wilde bepalen wat de beste bewaarstrategie was, welke stappen daarvoor nodig waren en uitzoeken wat de uiteindelijke keuze van een bewaarstrategie bepaalde. Dat hing onder andere samen met het type digitaal object en natuurlijk de beschikbare technologie(ën). Eén van de uitkomsten was dat vaak niet één strategie, maar een mix van strategieën noodzakelijk is (migratie met emulatie bijvoorbeeld).⁴⁴ Bij de vraag wat de beste bewaarstrategie(en) in welke context was, speelde het aspect van essentiële kenmerken ('significant properties') van een digitaal object een belangrijke rol, al was dat ook al enkele jaren eerder aan de orde gekomen.⁴⁵

Het onderzoek (inclusief testbeds) en de samenwerking op het gebied van 'digital preservation' in Europa en dan vooral bij erfgoedinstellingen werd in deze jaren gestimuleerd door middel van onderzoeksprogramma's van de Europese Commissie, zoals IST (Information Society Technologies). In de tabel hiernaast zijn enige van de vele projecten opgesomd en kort beschreven.

Belangrijke thema's bij deze projecten en bij onderzoek op gebied van 'preserving' in het algemeen waren onder andere (niet uitputtend):

Processen en modellen op basis van OAIS

In veel van de projecten, zoals CASPAR, Planets, Protage en Pericles, werden modellen ontwikkeld en 'preservingsprocessen' uitgewerkt op basis van OAIS. Hoe kunnen de verschillende processen, zoals *ingest*, *preservation planning*, *preservation action*, beschikbaarstellen e.d. worden vormgegeven? In het Planets project is gewerkt aan *preservation planning* en *preservation action* en een functioneel model waarin de samenhang wordt aangegeven, terwijl in CASPAR de aandacht was gericht op informatiestromen in samenhang met processen en in Pericles zogenaamde digitale ecosystemen en een bijbehorende ontologie zijn ontwikkeld.⁴⁷

Ontwikkeling softwarehulpmiddelen

Daarnaast was er behoefte aan softwarehulpmiddelen die de feitelijke 'preserving' van digitale informatie mogelijk maken en ondersteunen. Dat was vooral ook nodig

⁴⁴ Zie o.a. de resultaten van het CAMiLEON project.

⁴⁵ Zie o.a. Jeff Rothenberg, *Digital Preservation*, (Den Haag, 1999), p. 14. In het Planets project werd hiervoor een methode ontwikkeld als onderdeel van het *preservation planning* proces: zie bijv. <https://www.dpconline.org/docs/miscellaneous/events/138-preservation-planning-part-1-workflow-and-the-plato-tool-becker/file> en <http://www.ifs.tuwien.ac.at/dp/plato/documentation/#ijdl2009>. In het Planets project is ook een poging gedaan om tot een formele beschrijvingstaal te komen waarmee de essentiële kenmerken beschreven konden worden, eXtensible Characterisation Definition/Extraction Language, XCDL/XCEL (ontwikkeld door Manfred Thaller, Universiteit van Keulen). Met behulp daarvan kunnen 'significant properties' worden geëvalueerd bijvoorbeeld na migratie. Zie bijvoorbeeld Gareth Knight en Maureen Pennock, 'Data without meaning: establishing the significant properties of digital research', in: *The International Journal of Digital Curation*, Issue 1, Volume 4, 2009 p. 159-174. Zie ook http://www.ifs.tuwien.ac.at/~becker/pubs/becker_sac08.pdf en <http://www.ifs.tuwien.ac.at/dp/ipres2010/papers/tarrant-65.pdf>. Ook in de VS was men er mee bezig bij California Digital Library en Harvard's University Library (zie bijv. <https://blogs.bodleian.ox.ac.uk/archivesandmanuscripts/tag/significant-properties/>). Voorzover bekend is het niet verder ontwikkeld.

⁴⁶ Voor PLATO zie <http://www.ifs.tuwien.ac.at/dp/plato/intro/>

⁴⁷ Voor CASPAR zie onder andere: https://www.researchgate.net/publication/250150721_The_CASPAR_approach_to_digital_preservation?_sg=FWz9fXX0l2BdfXkjkRgze6gB3jKSfZrj-nPpTVhkjObguS9nyvVIGF-FXITuHWPTTz7mT8sejcuKjXt7ssDv65G8gqFnFKahmw en https://www.researchgate.net/figure/CASPAR-Key-Components_fig1_250150721. Voor Protage: <https://ercim-news.ercim.eu/en80/special/protage> en <http://journals.uic.edu/ojs/index.php/fm/article/view/3069/2629>. Voor Pericles: <http://www.pericles-project.eu>.

Naam project	Doelstelling	Periode	Resultaten	Opmerkingen / URL
ERPANET	opbouw netwerk, kennis delen en verspreiden	2001-2004	seminars, workshops, publicaties, website, hulpmiddelen, gidsen	http://www.erpanet.org
DELOS	netwerk (digital libraries), multidisciplinair	2004-2007	kader/ model voor testbed, 'summerschool on digital preservation',	WP6 Preservation Cluster, http://delos-old.isti.cnr.it/WP6.html
DIGicult	scanning, bewaring van digitaal erfgoed in Europa	2001-2005	Vele rapporten met stand van zaken. Lund principles en actieplan	http://collections.internetmemory.org/haeu/20160920164245/ http://cordis.europa.eu/ist/ka3/digicult/home.html
DPE	bevorderen bewustwording bij organisaties en het brede publiek, delen van kennis, networking, samenwerking (WePreserve)	2005-2009	DRAMBORA, PLATTER, trainingen/seminars, animaties (Team Digital Preservation), Digital Preservation Challenge (wedstrijd)	voortzetting van ERPANET, http://www.digitalpreservation.gov/series/edge/dpe.html
Planets	ontwikkeling van hulpmiddelen, testbed, training	2006-2010	hulpmiddelen zoals PLATO (planning tool) ⁴⁶ , herkenningsoftware (DROID, FIDO)	uit dit project ontstond in 2010 de Open Planets Foundation (OPF), nu Open Preservation Foundation: http://openpreservation.org/about/organisation
Caspar	Cultureel en wetenschappelijk erfgoed	2006-2010	architectuurmodel, testbed	http://www.dcc.ac.uk/resources/briefing-papers/technology-watch-papers/caspar
KEEP	ontwikkelen emulatie-omgeving	2007-2012	Emulatiekader en Virtual Machine.	voortbordurend op DIOSCURI http://emuframework.sourceforge.net en https://www.kb.nl/nieuws/2011/emulatie-bruikbare-oplossing-voorduurzame-digitale-toegang
Protage	ontwikkelen van software tools (intelligent software agents) tbv. digitale bewaring	2007-2010	model, prototype, rapporten	https://ercim-news.ercim.eu/en80/special/protage
SCAPE	Realiseren van schaalbaarheid, automatisering en planning van 'preservation actions'	2011-2014	softwarehulpmiddelen, zoals Jpylyzer	http://scape-project.eu
Pericles	toegankelijkheid van digitale bronnen waarborgen	2013-2017	modellen (digitale ecosysteem), ontologie, software hulpmiddelen	http://www.pericles-project.eu
E-ARK	verbeteren/harmoniseren van methoden en technologieën voor digitale archivering in Europa	2014-2017	Richtlijnen, software-hulpmiddelen	http://www.eark-project.com/about

NB. Van veel van de bovenstaande projecten is de website die ten tijde van het project bestond verdwenen, behalve van ERPANET, SCAPE, Pericles en E-Ark (raadpleegdatum juli 2018). Het feit dat al deze projecten zich bezighielden met digitale bewaring is helaas geen waarborg gebleken voor het behoud van de informatie op die websites.

om in de testbeds experimenten te kunnen uitvoeren. Voorbeelden zijn: *characterisation tools*, zoals FIDO voor het identificeren en valideren van bestandsformaten, ondersteuning van het beslissingsproces om te komen tot een goede bewaarstrategie voor een bepaald type digitale objecten (*preservation planning*).

Identificatie en validatie van bestandsformaten

Een ander thema was 'hoe ga je om met de vele verschillende bestandsformaten?' Een belangrijke rol in dit alles speelt een register van bestandsformaten. The National Archives (TNA, VK) heeft zo'n register aangelegd, PRONOM genaamd. Pogingen om tot één globaal register te komen, zoals het Unified Digital Format Registry (UDFR, University of California Curation Centre) en Global Digital Format Registry (GDFR, Harvard University) zijn op niets uitgelopen.⁴⁸ Het gevolg is dat PRONOM nu die rol vervult.⁴⁹ In nauwe samenhang hiermee speelt het karakteriseren en valideren van de gebruikte bestandsformaten bij informatie die in e-depots wordt opgenomen. Hiervoor zijn door verschillende organisaties softwarehulpmiddelen ontwikkeld zoals DROID (TNA), of JHOVE (Harvard), Jpylyzer (SCAPE project) en FIDO (Planets project).⁵⁰

Metadata

Welke metadata⁵¹ zijn nodig en met welke beschrijvingsstructuren? Het gaat om informatie met betrekking tot het beheer, de context (ontstaan, betrokkenen, in welk proces enz.), de structuur en de inhoud van digitale archiefbescheiden, dan wel objecten. Vanuit verschillende invalshoeken, disciplines als ook landen werden metadata sets voor de digitale omgeving ontwikkeld. Zo werd bijvoorbeeld voor de web-omgeving Dublin Core (DCMI) en voor 'digital preservation' PREMIS (preservation metadata, gebaseerd op OAIS) ontwikkeld.⁵² Het ging in het laatste geval vooral om gegevens over aspecten als 'provenance, fixity, significant properties, technical specifications and current activity', begrippen in het OAIS-model.

In de wereld van *recordkeeping* (incl. archieven) werden onder de vlag van ISO, door TC46/SC11, eveneens standaarden op dit gebied ontwikkeld, maar dan meer gericht op de onderliggende principes en concepten dan op het vaststellen van metadata sets.⁵³ Naast informatie over structuur, verschijningsvorm en technische gegevens, gaat het dan om informatie over de inhoud en context (waarom, waar, wanneer, door wie). Tenslotte is de informatie over het beheer van digitale archiefbescheiden essentieel om vast te kunnen stellen wat er gebeurd is sinds de creatie. Zoals aangegeven dienen metadata onverbreekelijk met de informatie (zoals publicaties of archief-

⁴⁸ Hierbij was o.a. het PLANETS project betrokken.

⁴⁹ Zie <http://www.udfr.org>, als ook <http://udfr.org/project/UDFR-final-report.pdf> en https://library.harvard.edu/preservation/digital-preservation_gdfr.html.

⁵⁰ Zie <https://www.nationalarchives.gov.uk/PRONOM/Default.aspx>. Andere herkennings/validatiesoftwarehulpmiddelen zijn o.a. JHOVE (<https://github.com/openpreserve/jhove/releases/tag/v1.20.1>), FIDO (<http://openpreservation.org/technology/products/fido/>) en Jpylyzer (<http://openpreservation.org/technology/products/jpylyzer/>).

⁵¹ De term metadata komt op gelijk met de ontwikkeling van informatiesystemen. In de archivaliek werd dit gevangen onder de noemer van beschrijven of toegankelijk maken. Het omschrijft of duidt kort gezegd andere informatie. Vooral bij uitwisseling van informatie tussen systemen, maar ook bij het beschikbaar stellen ervan zijn deze metadata van belang, omdat die informatie dan betekenis krijgt. Het is vergelijkbaar met bijv. een formulier waarin de velden nader worden aangeduid met een term of omschrijving. Een getal op zich zegt niet zoveel, maar als er 'jaartal' voor staat krijgt het betekenis. Zonder die metadata zou het bijv. moeilijk zo niet onmogelijk zijn om de digitale informatie te begrijpen, te zoeken of te beheren. Zij maken onverbreekelijk deel uit van die digitale informatie.

bescheiden) verbonden te zijn. Bovendien zijn metadata niet statisch, maar zullen gedurende het bestaan van digitale objecten metadata toegevoegd blijven worden. Een manier om metadata met de informatie vast te leggen is inkapseling ('encapsulation'). Een voorbeeld is de zgn. VEO (VERS encapsulated object) die door het Public Records Office van Victoria (PROV, Australië) is ontwikkeld.⁵⁴

De grootste uitdaging zit hem in het feit dat deze metadata veelal los van de eigenlijke informatie/documenten staan, vaak systeemgebonden zijn en in talloze formaten en samenstellingen voorkomen. Elke (overheids)organisatie heeft zijn eigen invulling daarvan. Voor een goed beheer in een depot moeten deze echter tot een bepaalde voor het depot geldende (in de regel 'open') standaard kunnen worden getransformeerd en dat blijkt ingewikkeld.

Unieke identificatie

Hoe kun je in een virtuele, 'onbegrensde' omgeving 'informatie-eenheden' (documenten, datasets e.d.) uniek identificeren? Een onderdeel daarvan was ook de zgn. *persistent identifier* (PID), een uniek kenmerk, waarmee digitale objecten (blijvend) kunnen worden geïdentificeerd. Verschillende systemen zijn daarvoor ontwikkeld, zoals Digital Object Identifier (DOI) of OpenURL.⁵⁵

Migratie tussen systemen

Hoe migreer je informatie tussen verschillende systemen? Een project op dit terrein was het Australische CRKM project (2003-2006) dat informatie uitwisseling tussen informatiesystemen heeft onderzocht.⁵⁶ Doel was de ontwikkeling van een model op basis waarvan door slim gebruik van metadata digitale informatie toegankelijk, bruikbaar en authentiek kan worden gehouden zo lang als nodig is. Migratie tussen systemen, vaak met verschillende metadata-schema's is daarbij een belangrijk aandachtspunt. Om dat te ondervangen en te vergemakkelijken zou gebruik gemaakt moeten worden van metadata schema registers, en een 'metadata broker' waarin de verschillende schema's met elkaar 'gematcht' zijn of worden. Daarmee wordt uitwisseling mogelijk gemaakt. Het in standhouden van digitale informatie vraagt sowieso steeds migratie ervan tussen systemen met al dan niet nieuwere technologie. Veelal gaat dat ook gepaard met conversie naar een ander bestandsformaat. Migratie van informatie kan bijvoorbeeld plaatsvinden van uit het oorspronkelijke systeem naar een e-depot, maar ook vanuit Word-omgeving naar een DMS (documentmanagementsysteem) dan wel RMS (recordmanagementsysteem) en vandaar naar e-depot en vervolgens tussen e-depots. Belangrijk probleem hierbij is vaak het exporteren van de metadata uit systemen en het gebrek aan standaardisatie ervan.

⁵² Zie voor DCMI <http://dublincore.org> en voor PREMIS: <https://www.loc.gov/standards/premis/>. Voor enige geschiedenis, zie o.a.: <http://www.dcc.ac.uk/sites/default/files/documents/resource/curation-manual/chapters/preservation-metadata/preservation-metadata.pdf>. Voor PREMIS zie ook: <https://www.dpconline.org/handbook/organisational-activities/metadata-and-documentation>

⁵³ Dat is de serie onder nr. ISO 23081 Metadata for records, waarvan de eerste in 2004 werd gepubliceerd.

⁵⁴ Zie <https://www.prov.vic.gov.au/recordkeeping-government/a-z-topics/vers>. Een ander voorbeeld is METS: <https://www.loc.gov/standards/mets/> dat de structuur van een uitwisselingscontainer definieert. Weer een andere methode is die van zgn. 'self documenting objects', waarbij bij elke handeling met het digitale object (of record) de informatie erover 'automatisch' wordt vastgelegd.

⁵⁵ Zie bijv.: <http://www.ncdd.nl/pid/> voor een korte inleiding en <http://www.ncdd.nl/projecten/netwerk-digitaal-erfgoed/project-persistent-identifiers/>. Als ook: <https://www.oclc.org/research/publications/2017/oclc-research-convenience-compliance-rim-europe.html>

⁵⁶ Zie <https://www.monash.edu/it/our-research/research-centres-and-labs/cosi/projects/past-projects/crkm>

Beoordelen van niveau

Op welke wijze kun je het niveau en de organisatie van beheer verbeteren en de overgang naar goed beheer van digitale informatie sturen en beheersen? Daartoe werd gekeken naar zogenaamde 'capacity' en 'maturity' modellen die in andere disciplines, zoals de financiële wereld, bestonden. In Canada is daar veel werk aan besteed met name aan de Information Management Capacity Check (IMCC).⁵⁷ In navolging van dergelijke modellen is ook in Nederland zo iets voor digitale bewaring ontwikkeld, het SCORE-model⁵⁸, waarbij de stand van zaken en de risico's in beeld worden gebracht. Het levert een actieplan op voor de organisatie.

Het onderzoek en de experimenten leverden heel veel rapporten en studies op, en ook enkele nuttige software-hulpmiddelen. Opvolgende projecten bouwden daarop soms voort, maar veel is inmiddels onder het stof (en van het internet) verdwenen. Dat wil niet zeggen dat het niet belangrijk of nuttig was. Integendeel, het hielp de gedachtenvorming en leidde tot nieuwe initiatieven. Heel belangrijk was ook dat het deskundigen uit allerlei disciplines en nationaliteiten bij elkaar bracht. Men deelde 'het probleem', waardoor er een levendige en boeiende uitwisseling van gedachten, ideeën en initiatieven ontstond. Deelnemers kwamen van onder andere bibliotheken, archiefdiensten, universiteiten, bedrijven, onderzoeksinstituten, musea en data centra, als ook de IT-sector, zoals Microsoft Research, Tessella en IBM. De nalatenschap van de bovenstaande projecten is voor een deel terug te vinden bij samenwerkingsverbanden als OPF (Open Preservation Foundation), Digital Preservation Coalition (DPC), Digital Curation Centre (DCC), en voor een deel helaas in de vergetelheid geraakt.⁵⁹

Ontplooiing en toepassen in de praktijk (2010- heden)

Waren de experimenten nog veelal in een laboratorium-omgeving (= testbed) uitgevoerd, vanaf ongeveer 2007 begon ook de bouw van digitale depots serieuze vormen aan te nemen. De mogelijkheid en noodzaak de ervaringen in een 'echte' digitale omgeving te testen ontstond. Voorlopig beperkte dat zich tot het opnemen van digitale objecten in een 'digitaal depot'. Dat ging dan om zaken als identificeren en valideren van het bestandformaat, het eventueel converteren naar een acceptabel en onderhoudbaar formaat, en het controleren van de vereiste bijgevoegde metadata op aanwezigheid en volledigheid. Voorbeelden daarvan in Nederland zijn: het Nationaal Archief, Gemeentelijke archiefdiensten Amsterdam en Rotterdam, NHDA, KB, en Beeld en Geluid. De laatste jaren zijn ook kleinere gemeentelijke archiefdiensten gestart met het inrichten van digitale depots, zoals het Erfgoedcentrum Achterhoek en de Liemers.

⁵⁷ Zie o.a. <https://www.slideserve.com/finn-gardner/information-management-capacity-check-imcc>. Er bestaan echter meer modellen. Zie bijv. <https://benchmarkdigitalpreservation.files.wordpress.com/2013/09/caiw2013katuu.pdf>. In de VS zijn de NDSA (National Digital Stewardship Alliance) levels of preservation in 2012 ontwikkeld, zie <https://blogs.loc.gov/thesignal/2016/04/expanding-nds-a-levels-of-preservation/> en het *Digital Preservation Capability Maturity Model* (DPCMM) van Charles Dollar en Lori J. Ashley (zie: <http://www.securelyrooted.com/blog/2015/7/7/dpcmm-27-released>, versie 2.7, 2015).

⁵⁸ Zie <https://scoremodel.org/site/home>, ontwikkeld door PACKED (kennisplatform voor de archivering en conservering van audiovisuele kunsten in Vlaanderen) en DEN.

⁵⁹ De OPF, opgericht in 2010 door de samenwerkende partners van het Planets project, heette oorspronkelijk Open Planets Foundation. Zie verder: <http://planets-project.eu> en <http://openpreservation.org>. Voor DPC zie: <https://dpcnline.org> en voor DCC: <http://www.dcc.ac.uk>. Jaarlijks vinden er ook veel conferenties plaats, zoals iPRES, JDCL of ICDL, waar ervaringen, kennis en informatie op dit gebied worden uitgewisseld >

In Vlaanderen was vooral het DAVID project (nu expertisecentrum David) te Antwerpen dat aan de weg timmerde. Het Felix archief/Stadsarchief Antwerpen heeft sinds 2007 een digitaal depot en heeft veel gedaan om tot praktische oplossingen voor digitale bewaring te komen.

Dat wat bewaard wordt beperkt zich al lang niet meer tot documenten of boeken, maar omvat inmiddels het brede scala van digitale uitingen en sociale media, zoals email, blogs, vlogs, sms, whatsappjes, tweets/twitter, video's, enz. Ook websites worden al sinds de oprichting van The Internet Archive in 1996 in toenemende mate bewaard, al is dit wat Nederlandse websites betreft nog met grote hiaten.⁶⁰ De KB speelt hierin sinds 2007 een belangrijke rol met het binnenhalen van de belangrijkste Nederlandse websites.

Het ontwikkelen van een digitaal depot dat meer is dan een opslagsysteem blijkt in de praktijk niet zo 'eenvoudig' als het lijkt. Een belangrijke reden hiervoor is dat de aangeleverde digitale informatie in de regel zeer divers is en gebrekkig gedocumenteerd. Dat betekent dat het opnameproces (*ingest*) ingewikkeld is, allerlei controles dient te bevatten en in bepaalde gevallen tot conversie van bestanden leidt. Eerste versies van digitale depots moeten al snel vervangen worden door nieuwe verbeterde varianten. Belangrijkste redenen zijn bijvoorbeeld dat men het te groot aanpakt (zoals NARA⁶¹), of dat men zaken over het hoofd heeft gezien of onvoldoende geïmplementeerd (zoals met betrekking tot *event history*, waarbij leveranciers of organisaties denken dat een *audit trail* voldoende is⁶²). Ook is er het risico dat het te ingewikkeld wordt gemaakt door te veel functies in één systeem te willen inbouwen.

De bestaande digitale depots hebben meestal wel in meer of mindere mate de basisfuncties voor opnemen, bewaren en beschikbaar stellen. In de regel (nog) niet planning en het in standhouden: de functies *preservation planning*, *preservation action*. Dergelijke functionaliteit hoeft evenwel geen onderdeel te zijn van de e-depots, maar kan afzonderlijke dienstverlening zijn. Een ander punt is in hoeverre e-depots moeten zijn ontwikkeld als 'open source', waarbij duidelijk is hoe het precies intern werkt, of als commerciële software, waar dat niet altijd duidelijk is uit concurrentie-overwegingen?⁶³

Digitale depots zijn slechts één aspect dat van belang is. Minstens zo belangrijk is het hebben van beleid. Dat geldt zowel voor de beherende instelling, zoals een Nationaal Archief, maar ook voor de archiefvormende instellingen. Er zijn wat de laatste betreft wel kaders en richtlijnen, maar toepassing ervan in de praktijk blijkt weerbarstig.

> tussen projecten, deskundigen en belangstellenden. Voor een deel is informatie en kennis verloren gegaan, nadat projecten waren afgerond. Websites werden niet meer onderhouden of verdwenen gewoon, omdat niemand zich verantwoordelijk voelde en het ook niet was geregeld.

⁶⁰ Zie voor Internet Archive (of WayBackMachine) opgericht door Brewster Kahle: <https://archive.org/web/web.php>. Ook is er de International Internet Preservation Coalition (IIPC) uit 2003 die de samenwerking en standaardisatie op dit gebied heeft bevorderd. In Europa is er ook de Internet Memory Foundation (oorspronkelijk European Archive Foundation) sinds 2004.

⁶¹ Zie: <https://www.archives.gov/era/about>.

⁶² Dat was bijvoorbeeld het geval bij de eerste versies van Preservica, toen nog SDB (Safety Deposit Box) geheten.

⁶³ Er zijn inmiddels verschillende versies van digitale depots in gebruik, zoals het commerciële Preservica, en de open source versies RODA en Rosetta.

Het kost tijd voor organisaties om die richtlijnen in hun systemen toe te passen, ook omdat niet altijd duidelijk is hoe dat moet.⁶⁴

Naast het verder ontwikkelen van digitale depots is er ook een toenemend streven om een netwerk aan diensten te ontwikkelen op gebied van digitale bewaring. In Nederland is dat te zien in het Netwerk Digitaal Erfgoed (NDE).⁶⁵ Het doel is om een stelsel van landelijke voorzieningen voor het behoud van digitale erfgoed in erfgoedinstellingen te ontwikkelen. Niet iedereen hoeft in zo'n constructie alles zelf te doen, maar voorzieningen kunnen worden gedeeld.⁶⁶ Het is gebaseerd op de *Nationale Strategie Digitaal Erfgoed (2015)*, waarin een duurzaamheidsbeleid is geformuleerd.⁶⁷ Grote instellingen als KNAW, Nationaal Archief, Koninklijke Bibliotheek, Nederlands Instituut voor Beeld en Geluid, Rijksdienst voor het Cultureel Erfgoed zijn daarbij betrokken. Archiefdiensten zitten hier wat dubbel in. Zij staan immers ook met één been in de overheidsadministratie. Door zich te veel of zelfs uitsluitend met erfgoedinstellingen te verbinden, ontstaat het risico dat de verbinding met de overheidsadministratie in het gedrang komt en dat kan dan bijvoorbeeld leiden tot meer problemen bij overbrenging van digitale archiefbescheiden.⁶⁸

Samenvatting en enige voorlopige conclusies

In het voorgaande is in grote lijnen een beeld geschetst van wat er de afgelopen drie decennia gebeurd is op het gebied van bewaren en beheren van digitale informatie. Daarbij lag de aandacht vooral op de archiefwereld. Het is al gezegd: het betekende dat archiefdiensten in Nederland in deze jaren zich enerzijds met een inhaalslag ten aanzien van papieren archieven bezig moesten houden vanwege de verkorting van de overbrengingstermijn en anderzijds toekomstgericht bezig moesten zijn met een nieuw type archiefbescheiden en met de virtuele omgeving van het worldwideweb, die beide een grote impact zouden hebben op hun rol, structuur en werkwijze.

Het beeld dat eruit naar voren komt is dat de opkomst van digitale informatie dankzij de nieuwe informatietechnologie grote gevolgen heeft voor alle informatie- of data beherende instellingen, zowel technisch, organisatorisch als maatschappelijk. Zij moeten zich als het ware opnieuw uitvinden. De aard van de digitale informatie in samenhang met de komst van internet vraagt een andere technische infrastructuur, zowel voor beheer als voor beschikbaarstellen, terwijl diezelfde infrastructuur ook de rol van de instellingen verandert. De beherende instelling staat vanuit het oogpunt van de 'gebruiker' of onderzoeker op het internet niet meer centraal, maar de informatie, ongeacht waar die bewaard wordt.

⁶⁴ Voorbeelden van dit soort richtlijnen zijn standaarden als de ISO 15489:2016, of metadatrichtlijnen als Toepassingsprofiel Rijk of TML0, eisen sets als DUTO (en voorheen Baseline Informatiehuishouding Rijksoverheid dan wel gemeenten).

⁶⁵ Zie <http://www.den.nl/pagina/511/netwerk-digitaal-erfgoed/> en <http://www.netwerkdigitaal-erfgoed.nl>.

⁶⁶ Zo zijn op het digitale depot van het Nationaal Archief verschillende regionaal historische centra aangesloten.

⁶⁷ http://www.netwerkdigitaal-erfgoed.nl/wp-content/uploads/2018/02/Nationale_Strategie_Digitaal_Erfgoed_MinOCW.pdf

⁶⁸ In dit verband is ook de huidige trend om de overbrengingstermijn verder te verkorten naar 8-10 jaar interessant. Het brengt de archiefdienst nog dichter bij de overheidsadministratie (en omgekeerd).

'Preservering' is daarvan slechts één, zij het niet onbelangrijk aspect, dat evenwel in die bredere samenhang bekeken moet worden. De afgelopen decennia zijn de nodige stappen gezet: van bewustwording naar experimenteren en praktische ervaring opdoen en naar ontwikkeling en implementatie van een digitale infrastructuur voor beheer. We staan echter nog pas aan het begin. De geïmplementeerde 'e-depots' zijn nog tamelijk primitief en vooral gericht op gecontroleerd binnenhalen, bewaren van de digitale informatie en beschikbaar stellen. De eigenlijke 'preserverings'-activiteiten, *preservation planning* en uitvoering van (noodzakelijke) onderhoudswerkzaamheden, zijn nog slechts in zeer beperkte mate mogelijk, als ze al mogelijk zijn. Verder is duidelijk dat 'preservering' niet begint met het binnenhalen van digitale informatie, maar dat ook de 'producenten' een belangrijke rol spelen in het maken van 'bewaarbare' informatie-eenheden.

Tenslotte staan de ontwikkelingen in informatietechnologie niet stil, maar gaan die in razend tempo voort. Het is vergelijkbaar met 'shooting at a moving target' en dat vraagt voortdurende aanpassing, ontwikkeling en flexibiliteit.⁶⁹

⁶⁹ J. Hofman, 'Shooting at a moving target'. The development of a repository for the preservation of digital information, in: *Proceedings of the DLM-Forum EUROPEAN on electronic records. European citizens and electronic information: the memory of the Information Society*, INSAR (Information Summary on Archives), Supplement IV (Brussel, 2000), p. 80-89. (2000), p 80-89.

OAIS: Keurslijf of kans?

Inleiding

De ISO-standaard OAIS behoort al twee decennia tot de standaarduitrusting van iedereen die zich met digitale duurzaamheid bezighoudt en wordt gezien als de leidraad bij het in praktijk brengen van duurzame bewaring van digitaal materiaal. Omdat de standaard zowel een functioneel model biedt dat aangeeft welke acties bij digitale duurzaamheid horen, als een informatie model dat aangeeft welke metadata aan het digitale object meegegeven wordt, is de algemene notie dat OAIS beschrijft wat digitale duurzaamheid of 'digital preservation' is.

Dit artikel beschrijft de ontstaansgeschiedenis van de standaard (wat was er voordat OAIS bestond?) en gaat in op de invloed die de standaard wereldwijd heeft. Die invloed betreft zowel systemen als opleidingen, certificering, en praktijk in vele instellingen. De dominante positie van OAIS heeft positieve effecten als gedeelde terminologie en begrippen maar ontvangt ook kritiek. Hoewel het ISO bureau een 5-jaarlijkse review van de standaard organiseert, leidt deze review in de praktijk zelden tot fundamentele wijzigingen. Is de standaard robuust genoeg om organisaties 'in de geest van OAIS' een oplossing te kunnen bieden voor de ontwikkelingen in de digitale wereld, of loopt de OAIS-standaard de kans de aansluiting te verliezen en ontstaan er daarnaast 'de facto standaarden'? Wat is daar het risico van?

Het ontstaan van de OAIS-standaard

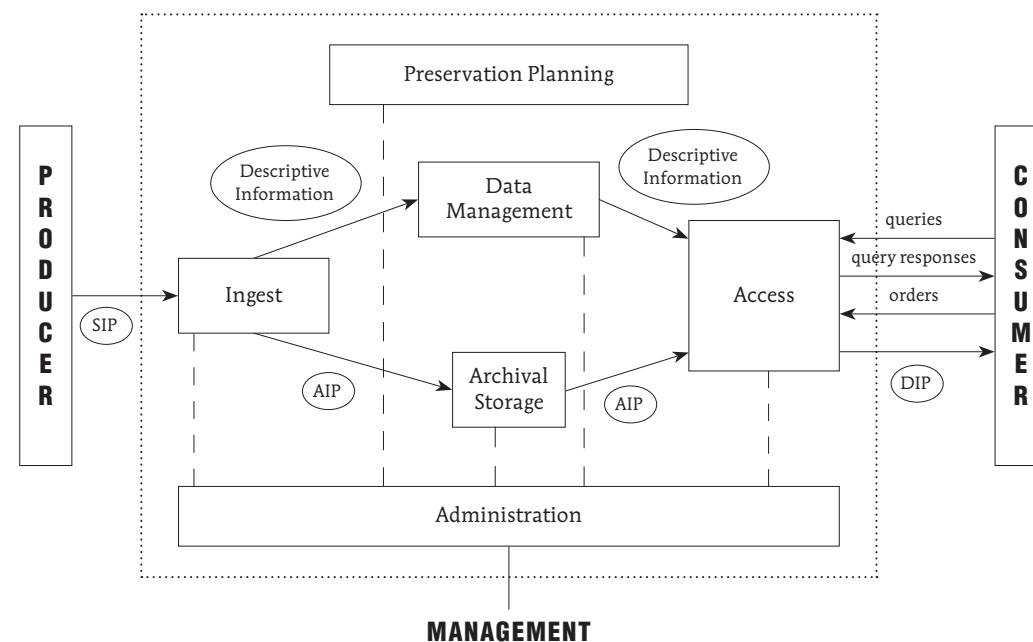
Ruim 20 jaar geleden publiceerde de Task Force on Archiving of Digital Information in de Verenigde Staten een rapport¹, dat enkele jaren later uitgebreider zou verschijnen als de eerste versie van de ISO standaard 14721 *The Open Archival Information System*² standaard. In de wandelgangen noemen we deze standaard 'het OAIS'. De opstellers van het rapport waren afkomstig uit bibliotheken, archieven en automatiseringsbedrijven en luidden de noodklok over de toenemende hoeveelheid digitaal materiaal dat voor de lange termijn bewaard zou moeten worden. De Consultative Committee for Space Data Systems (CCSDS³) is een internationale organisatie van ruimtevaartinstellingen, die standaarden formuleert voor deze instellingen. Daar werd het initiatief genomen om de OAIS-standaard te ontwikkelen, met bijdragen uit andere vakgebieden. Een deel van de opstellers van het Task Force rapport nam deel aan dit initiatief. De versie van CCSDS is daarna voorgedragen en geaccepteerd als ISO-standaard 14721 in 2002. Nog steeds is het CCSDS die de standaard onderhoudt.

¹ Waters, Donald; Garrett, John *Preserving Digital Information, Report of the Task Force on Archiving of Digital Information* (1996) <https://www.clir.org/pubs/reports/pub63/> Geraadpleegd 1 oktober 2018

² De finale versie is te vinden op de site van CCSDS <https://public.ccsds.org/pubs/650x0m2.pdf> of te koop bij het ISO bureau <https://www.iso.org/standard/57284.html> Geraadpleegd 1 oktober 2018

³ De officiële website van CCSDS is <https://public.ccsds.org/default.aspx> Geraadpleegd 1 oktober 2018

Ook in Europa ontstonden binnen de bibliotheek- en archiefwereld initiatieven. De Koninklijke Bibliotheek leverde input aan de samenstellers van OAIS-standaard. Een groep Europese bibliotheken onder de naam van Networked European Deposit Libraries⁴ schreef verscheidene rapporten over het inrichten van systemen voor 'electronic publications'. In de archiefwereld werd in 1999 de internationale InterPARES⁵ groep opgericht met als doel beleid, strategieën en standaarden te ontwikkelen voor duurzame bewaring. Na het verschijnen van de OAIS-standaard in 2002 gaven verschillende nationale bibliotheken leveranciers opdracht een systeem volgens deze richtlijn te bouwen, zoals bijvoorbeeld de Koninklijke Bibliotheek de opdracht gaf aan IBM (dit systeem werd in 2004 in gebruik genomen). Het verschijnen van de OAIS standaard was een eerste aanzet om grip te krijgen op het bewaren van digitaal materiaal.



Figuur 1 OAIS Functioneel model

Wat is OAIS?

De OAIS-standaard beschrijft verschillende componenten van een archief waarin digitale objecten duurzaam toegankelijk worden gehouden. Het concept Digitaal Archief omvat in de standaard niet alleen het systeem dat gebruikt wordt om de objecten op te slaan, maar ook de organisatie eromheen. Daarom begint de standaard met duidelijk te maken welke verantwoordelijkheden een organisatie op zich neemt als men start met een digitaal archief. Welk (juridisch) mandaat heeft de

⁴ De NEDLIB publicaties zijn te vinden op <https://www.kb.nl/en/organisation/research-expertise/research-on-digitalisation-and-digital-preservation/publications-on-digital-preservation-and-digitization/the-nedlib-publications> Geraadpleegd 1 oktober 2018

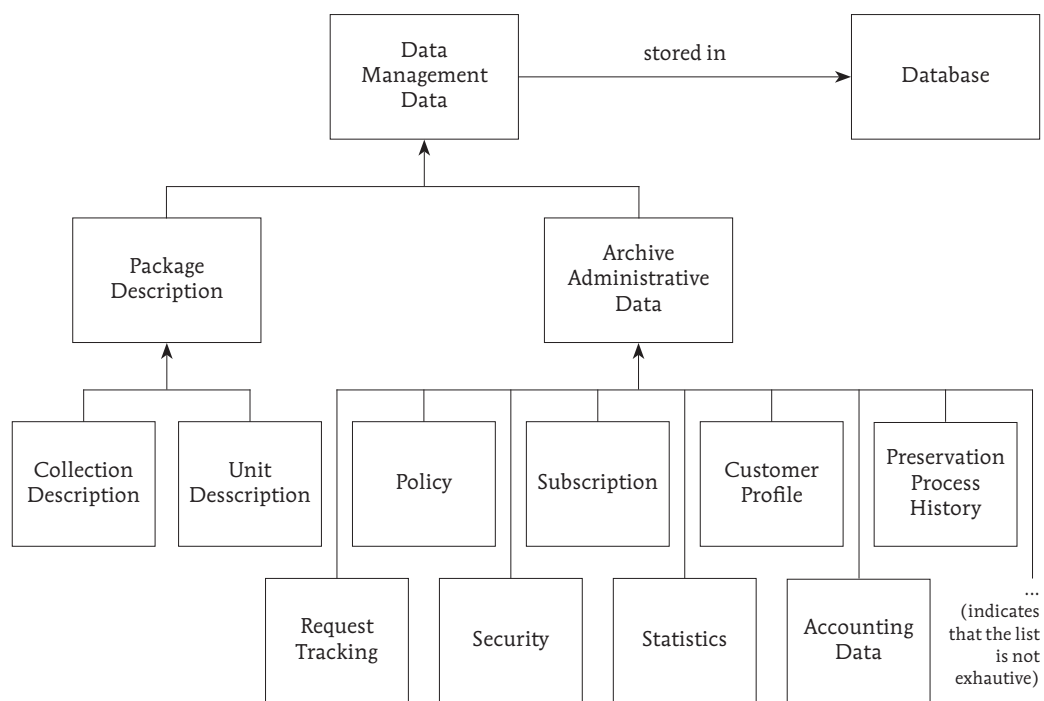
⁵ Meer over het InterPARES project is te vinden op de website http://www.interpares.org/ip1/ip1_index.cfm Geraadpleegd 1 oktober 2018

organisatie om materialen te bewaren? Wie is de doelgroep van deze activiteit, met andere woorden: wie is de 'Designated Community'? Hoe zal het materiaal ter beschikking worden gesteld en hoe denkt de organisatie de authenticiteit te kunnen bewaren?

Vervolgens beschrijft de standaard op hoofdlijnen uit welke zes onderdelen een digitaal archief bestaat, met daarbij de bijbehorende functies:

- Ingest: het proces van ontvangst van het digitale object tot en met het gereed maken voor opslag,
- Storage: de wijze van opslag,
- Data Management: het beheer van de verschillende soorten metadata die bij het object horen
- Administration: alle activiteiten die samenhangen met het onderhoud van het digitale archief,
- Preservation Planning: het risicomanagement en de acties die genomen moeten worden om risico's voor te zijn dan wel op te vangen,
- Access: de manier waarop de gebruiker toegang krijgt tot de digitale objecten.

Tevens beschrijft de standaard het informatiemodel van de objecten, met andere woorden welke extra informatie moet aan het digitale object worden toegevoegd om de begrijpelijkheid en toegankelijkheid op termijn te kunnen garanderen. Deze extra informatie zit in de metadata.



Figuur 2. OAIS Data model

De OAIS-standaard beschrijft drie fases van het digitale object. Het Submission Information Package (SIP) is het object zoals dit bij het archief binnenkomt. Deze SIP kan verschillende bewerkingen ondergaan zoals toevoegingen van metadata en slaat men vervolgens op als Archival Informatie Package (AIP). Uiteindelijk krijgt de gebruiker de beschikking over een Dissemination Information Package (DIP). In dit artikel is geen ruimte om hier uitgebreid op in te gaan, daarvoor verwijs ik naar mijn artikel over OAIS.⁶

Samenvattend beschrijft de OAIS-standaard alle aspecten van een digitaal archief. Maar een waarschuwing is op zijn plaats: de OAIS-standaard is een 'referentiemodel' en geen kant-en-klaar voorschrift.

De OAIS-standaard als referentiemodel

Een referentiemodel is een poging om een fenomeen te modelleren, ook al is er op het moment van opstellen nog niet veel praktijkervaring. Toen OAIS werd opgesteld, werd duurzame bewaring nog op beperkte schaal toegepast. In het OAIS zijn deze ervaringen gecombineerd met kennis en inzicht uit verschillende vakgebieden (bibliotheken, archieven, uitgevers, automatisering, wetenschap) en uiteraard 'gezond verstand'. Omdat de standaard breed toepasbaar moet zijn en op conceptueel niveau wordt vastgesteld, is het abstractieniveau hoog en zal men een koppeling met de praktijk zelf moeten zoeken.

OAIS geeft geen definitie van 'digital preservation' maar biedt, zoals in de standaard staat, het 'framework' voor 'Long Term Preservation', waarbij toegevoegd wordt "Long Term is long enough to be concerned with the impacts of changing technologies, including support for new media and data formats, or with a changing user community." Anderen hebben zich wel gewaagd aan definities. Zo beschrijft bijvoorbeeld het Digital Preservation Handbook van de Digital Preservation Coalition de termals "[Digital Preservation] Refers to the series of managed activities necessary to ensure continued access to digital materials for as long as necessary."⁷

Hoewel in de loop der tijd gebleken is dat de OAIS standaard vrij compleet is in het benoemen van relevante elementen voor duurzame bewaring, is het voor elke individuele instelling vaak een opgave om een eigen invulling aan de elementen te geven. Het niveau van abstractie moet immers vertaald worden naar de eigen instelling. Dat dwingt ons zowel de concepten uit OAIS te doorgronden maar ook om goed na te denken waar de raakvlakken liggen met onze eigen organisatie. Een nationale bibliotheek bijvoorbeeld is opgericht ten dienste van een algemeen publiek: hoe omschrijft men dan de 'designated community'? 'Iedereen' is te vaag en biedt onvoldoende houvast voor relevante acties als bijvoorbeeld het bepalen van essentiële kenmerken van een object (die kunnen voor 'iedereen' anders zijn) of het vaststellen hoe de Dissemination Information Package voor toegang eruit moet zien – dat kan een andere samenstelling hebben voor een wetenschapper of een scholier. Kortom, de OAIS-standaard is geen checklist die je kunt afvinken, maar geeft in grote lijnen de principes weer van duurzame bewaring, om die in de praktijk nader in te vullen – in de 'geest van OAIS'.

⁶ B. Sierman, 'Het OAIS model: leidraad voor duurzame toegankelijkheid' in: *Handboek Informatiewetenschap*, 2012, https://www.kb.nl/sites/default/files/docs/sierman_oaismodelned.pdf Geraadpleegd 1 oktober 2018

⁷ Het DPC Handbook en de glossary zijn te vinden op <https://www.dpconline.org/handbook/glossary> Geraadpleegd 1 oktober 2018

Keurslijf of kans?

De volgende gedachtenoefening is interessant: hoe zou digitale duurzaamheid eruit zien als er geen OAIS-standaard was? Over welke concepten waren we het dan eens en oneens, konden we het eigenlijk wel ergens eens over worden? Welke termen gebruikten we voor verschillende onderwerpen? Zou iedereen zelf uitvinden door 'trial and error' hoe het best digitaal materiaal in een systeem op te slaan? Hadden we een manier om te beoordelen of de ene instelling het beter deed dan de andere? Of het ene systeem beter was dan het andere? Waren er criteria voor 'slechte duurzaamheid'? Had iedereen wel overal aan gedacht? Zouden we iets over de domeinen heen kunnen bereiken, iets waar bibliotheken, archieven, datacenters, commerciële instellingen, wetenschap en industrie het over eens zouden zijn? Konden we überhaupt dit wel met elkaar bespreken en elkaar dan ook begrijpen?

Ik betwijfel het. Velen zijn het erover eens dat de OAIS-standaard als referentiemodel antwoord geeft op een aantal van bovenstaande vragen. Het geeft een gemeenschappelijk begrip van termen en concepten over domeinen heen (vandaar dat de termen zelden vertaald worden en iedereen de Engelse OAIS-terminologie gebruikt) en een kader voor alle aspecten die bij digitale duurzaamheid een rol spelen. De OAIS-standaard beschrijft geen specifieke technieken of implementaties en is daardoor dus niet verbonden met bepaalde technische oplossingen. Uiteraard zijn er leveranciers die hun systemen hebben gebaseerd op de standaard en de functionele entiteiten en het informatiemodel hebben overgenomen.

Doe ik het goed?

Er zijn instellingen die inmiddels ruim 50 jaar lang digitale collecties beheren. Ze begonnen met ponsbanden en verwerkten alle opvolgende dragers van digitale informatie zoals tapes, CD-cassettes, floppy's in soorten en maten, ZIP-drives en via FTP aangeleverde bestanden. Een van die instellingen in Nederland is bijvoorbeeld DANS, die net als hun collega's in het Verenigd Koninkrijk UKDA en de Verenigde Staten ICPSR al sinds de jaren '50 gegevens op het gebied van sociale wetenschappen duurzaam toegankelijk houdt. Die instellingen hebben migraties gedaan, schaduwkopieën van hun bestanden gemaakt, over metadata nagedacht, voorkeursformaten vastgesteld en systemen ontwikkeld om de gegevens op te slaan. Je zou kunnen zeggen dat hun gezond verstand en technische kennis ze op weg hebben geholpen. Maar uiteindelijk wilde ook DANS aantonen aan haar klanten dat ze de duurzame bewaring op een verantwoorde manier deden en volgens de internationale praktijk. En dan komt certificering om de hoek kijken.

Al in het rapport uit 1994 werd geadviseerd een certificering proces te ontwerpen, met name om de betrouwbaarheid van de digitale archieven te kunnen aantonen "to create a climate of trust in this field of endeavor".⁸ En toen de OAIS-standaard eenmaal was opgesteld, werd de vraag hoe 'kunnen we checken of de instellingen ook werken volgens deze standaard?' Er ontstond een werkgroep die een lijst criteria opstelde en in 2005 publiceerde de research Libraries Group het document Trusted digital repositories, roepnaam TRAC. Dit document was de basis voor een nieuwe

⁸ Zie noot 1

⁹ De standaard is te vinden op <https://public.ccsds.org/pubs/652x0m1.pdf> (geraadpleegd 1 oktober 2018) of te koop bij de ISO organisatie.

¹⁰ Zie figuur 1

¹¹ <https://www.netwerkdigitaalergoed.nl/kennis-en-voorzieningen/digitaal-erfgoed-zichtbaar/project-kosten-van-duurzame-toegang/> Geraadpleegd 1 oktober 2018

ISO standaard, die in 2012 verscheen als ISO 16363 *Audit and Certification of Trustworthy Digital Repositories*.⁹ Zie het artikel over certificering elders in deze bundel.

De invloed van OAIS

Een aantal jaren geleden kon je geen presentatie over digitale duurzaamheid zien, of het [bovenstaande¹⁰] plaatje van OAIS werd getoond. Inmiddels is men zo familiair met het OAIS-model dat het tonen van de diagrammen uit de standaard op conferenties niet meer nodig is. Naast de wereld van bibliotheken en archieven en de ruimtevaart, is het model ook geïntroduceerd bij overheidsinstellingen, commerciële instellingen, en verschillende wetenschapsgebieden als medici, bio medici en sociale wetenschappen etc. Er zijn vertalingen verschenen van de standaard onder meer in het Frans, Tsjechisch, Slowaaks. Verschillende cursussen leggen het model uit en Wikipedia geeft inleidende informatie. Zowel commerciële leveranciers als open source initiatieven hebben systemen ontwikkeld om objecten van een OAIS-archief op te slaan en baseren zich daarbij op de standaard, die inmiddels als de 'golden standard' wordt gezien. Diverse kostenmodellen gaan uit van de functionaliteiten die in OAIS genoemd worden, zoals onder meer het model dat in het project 4C werd ontwikkeld en nu binnen NDE verder is uitgewerkt¹¹. En uiteindelijk zijn er ook verschillende wijzen van certificering van een OAIS archief (de Core Trust Seal, nestor/DIN en de ISO 16363), aangevuld met de Nederlandse ED3 standaard. Er komt steeds meer transparantie over de implementatie van de standaard in de praktijk door publicatie van workflows en de architectuur van ontworpen systemen en men deelt gewijzigde inzichten.¹²

Kortom: rondom de OAIS-standaard is een ecosysteem van aanvullende diensten en systemen ontstaan in verschillende domeinen. De standaard biedt kennelijk voldoende aanknopingspunten om de verschillen te kunnen overbruggen. De brede toepassingsgebieden van de huidige systemen als bijvoorbeeld Preservica en Archivematica (beiden gebaseerd op OAIS) laten dit ook zien.

OAIS en de samenhang met andere standaarden

Het duurzaam bewaren van digitale objecten kent vele aspecten die in de OAIS-standaard niet uitgebreid worden behandeld maar wel relevant zijn. Enkele daarvan worden genoemd in de bijlage bij de OAIS-standaard (annex B). Het gaat hier bijvoorbeeld om de metadata standaard PREMIS, IT-standaarden en kwaliteitsstandaarden. Een interessant overzicht voor archieven is te vinden op de site van NEN¹³, waar relevante ISO standaarden voor archiefbeheer zijn weergegeven en temidden daarvan staat dan OAIS.

CCSDS werkt intussen doelbewust aan een reeks standaarden die bepaalde aspecten uit OAIS nader invullen. Inmiddels zijn er standaarden verschenen die de relatie tussen de leverancier en het archief nader definieert zoals de Producer-Archive Interface Specification (PAIS) en de Producer-Archive Interface Methodology

¹² Zie bijvoorbeeld de evaluatie op iPRES 2017 van B.Caron e.a *Life and Death of an Information Package: Implementing the Lifecycle in a Multi-Purpose Preservation System*. (iPRES 2017 Conference Kyoto) <https://ipres2017.jp/wp-content/uploads/25Bertrand-Caron.pdf> Geraadpleegd 1 oktober 2018

¹³ Normcommissie Informatie- en Archiefmanagement, Plaat overzicht NEN en ISO normen 2017-12-01 <https://www.nen.nl/Home/Informatie-en-archiefmanagement.htm> Geraadpleegd 1 oktober 2018

Abstract Standard (PAIMAS). Men werkt aan het Data Exchange Protocol for Interoperability and Preservation (DEPIP) en Data Archive Architecture Description Document waarin uitwisseling van gegevens en de architectuur van de digitaal archief systeem nader worden uitgewerkt. Op termijn zullen dit ook ISO-standaarden worden. Inmiddels zijn de standaarden PAIS en PAIMAS redelijk ingeburgerd in de wereld van digitale duurzaamheid en worden deze ook door verschillende collectie beherende instellingen gebruikt in hun praktijk.

Van wie is OAIS?

Hoewel de eerste versie van de OAIS standaard gemaakt werd door een groep vrijwilligers die dit namens CCSDS deden, kan in principe iedereen bijdragen aan de standaard via de nationale standaardcommissies, gerelateerd aan ISO. In de praktijk zijn weinig instellingen vertegenwoordigd in hun nationale ISO vertegenwoordiging (in Nederland is dit NEN¹⁴).

Maar net als elke ISO-standaard kent ook de OAIS-standaard een periodieke review om de 5 jaar. De review van OAIS vindt plaats in een werkgroep van CCSDS (die het resultaat vervolgens aan ISO 'doorgeeft'). Dit keer (de review van 2017) is gepoogd het proces transparant te maken, door de inrichting van een website review.oais.org, waarop wijzigingsvoorstellen ingediend konden worden. Ook binnen de preservation community ontstond een initiatief om hieraan bij te dragen en via iPRES¹⁵ en een speciale website¹⁶ werd men vanaf 2016 opgeroepen te helpen de standaard te actualiseren. Uiteindelijk zijn ruim 200 wijzigingen voorgesteld. Op de website zijn de voorstellen te lezen, als ook de discussie die binnen de CCSDS-werkgroep is gevoerd en tenslotte de uiteindelijke beslissing die hierover is genomen. De volgende versie van de OAIS-standaard zal een aantal nieuwe concepten introduceren die inmiddels in de wereld van digitale duurzaamheid al gemeengoed zijn geworden. Emulatie heeft sinds de vorige versie van OAIS een grote ontwikkeling doorgemaakt en dat zal in de standaard te zien moeten zijn. De introductie van het concept van 'Preservation Watch' (ontstaan in het Europese project Planets 2007-2011¹⁷) is geaccepteerd als voorstel, net als het voorstel om 'Distributed Preservation / Inner OAIS-Outer OAIS'¹⁸ op te nemen.

Door een tijdige voorbereiding en transparantie van het proces kan de standaard blijven aansluiten bij de actualiteit en de praktijk van de instellingen. Maar de huidige ontwikkelingen op digitaal gebied gaan zo snel, dat de digitale archieven ze al nauwelijks kunnen bijhouden. Laat staan dat OAIS op enig moment naadloos aan kan sluiten bij de werkelijkheid. Wijzigingen in de standaard zullen daarnaast altijd beperkt blijven omdat men de standaard 'backwards compatible' wil houden en zo voorkomen dat bestaande implementaties ineens niet meer aan de standaard voldoen.

Nieuwe ontwikkelingen

Van oudsher is de OAIS-standaard voornamelijk toegepast in de wereld van bibliotheken, archieven, archeologie en ruimtevaart. De toenemende aandacht voor onderzoek data en het werken met grote datasets in de wetenschap heeft ook daar de aandacht voor goed databeheer gestimuleerd. Niet alleen beheer van data in het hier en nu maar ook op de lange termijn, opdat onderzoek gereproduceerd kan worden en eerder gemaakte datasets herbruikt kunnen worden. Dit streven is samengevat in

de 'FAIR principles', een acroniem van de begrippen Findable = vindbaar, Accessible = toegankelijk, Interoperable = interoperable en Retrievable = herbruikbaar. Deze laatste categorie, 'herbruikbaar' komt het dichtst in de buurt van duurzame toegankelijkheid. De principes die aan bruikbaarheid gerelateerd zijn, gaan voornamelijk in op metadata die aanwezig moet zijn ten aanzien van licenties, de voorgeschiedenis (provenance), dus op eigenschappen van de datasets en niet op eigenschappen van de omgeving waarin de herbruikbaar gerealiseerd moet worden. De FAIR principes¹⁹, gelanceerd door FORCE 11 groep, worden met name in Europa sterk gepropageerd door de Europese Commissie, door onder meer de eis te stellen dat de projectresultaten van de door hen gefinancierde projecten aan deze FAIR principes moeten voldoen.

Helaas lijkt het erop dat er bij de samenstelling van deze principes niet veel aandacht is voor de decennialange ervaring met lange termijn bewaring van digitaal materiaal bij bibliotheken en archieven. Gelukkig zijn datacentra als DANS van meet af aan betrokken geweest, maar er is reëel gevaar dat het wiel van duurzame toegankelijkheid opnieuw wordt uitgevonden als er geen uitwisseling van informatie ontstaat. Op dit moment wordt de OAIS-standaard nog niet gerelateerd aan de FAIR-principes, ondanks het feit dat veel universitaire instellingen hun digitale objecten in een op OAIS gebaseerd systeem opslaan. De basis certificeringsmethode, het Core Trust Seal (CTS), dat ook bij de wetenschappelijke datacenters als primaire certificering wordt gezien, noemt de OAIS wel maar alleen voor instellingen 'with a preservation remit' en deze zinsnede is minder expliciet dan in de voorganger van de CTS, het Data Seal of Approval. Ook datasets die voor zeg 10 jaar bewaard moeten blijven, zullen een omgeving nodig hebben die aan OAIS voldoet.

Als we er niet voor zorgen dat de FAIR-principes aansluiten op de OAIS-principes, lopen we het risico dat er opnieuw concepten worden bedacht, waarvoor reeds een standaard is die daarin voorziet en met een groot toepassingsgebied.

Hoe verder?

De nadruk die op de OAIS-standaard wordt gelegd en het soms als dwingend ervaren karakter ervan stuit sommigen tegen de borst. In de praktijk zullen er niet veel zijn die de standaard van ruim 120 pagina's ook daadwerkelijk gelezen en begrepen hebben. Daarin zit een risico, immers als de standaard niet meer past bij wat men in de praktijk doet, verliest de standaard zijn waarde voor de gemeenschap. Er zullen de facto standaarden gaan ontstaan en menigeen zal zijn eigen oplossing bedenken. Daar zitten nadelen aan, al is het maar dat het de onderlinge communicatie bemoeilijkt maar ook de onderlinge uitwisseling van gegevens en informatie. Aan de andere kant zijn er nu enorm veel instellingen met een digitaal archief die dit volgens de voorschriften van de OAIS-standaard hebben ingericht, ofwel via een

¹⁴ <https://www.nen.nl/> Geraadpleegd 1 oktober 2018

¹⁵ Panel OAIS for all of us (iPRES 2016 Conferentie Bern) http://www.ipres2016.ch/ipr16.organizers-congress.org/frontend/organizers/media/iPRES2016/_PDF/IPR16.Proceedings_4_Web_Broschuere_Link.pdf Geraadpleegd 1 oktober 2018

¹⁶ De special website is te vinden op http://wiki.dpconline.org/index.php?title=OAIS_Community Geraadpleegd 1 oktober 2018

¹⁷ Sierman, Barbara & Wheatley, Paul *Evaluation of Preservation Planning within OAIS, based on the Planets Functional Model* (Zenodo, 2010, May 12) <http://doi.org/10.5281/zenodo.853305> Geraadpleegd 1 oktober 2018

¹⁸ Zierau, Eld, & McGovern, Nancy. *The Outer OAIS - Inner OAIS (OO-IO) Model and use in Distributed Digital Preservation (DDP)*. (Zenodo, 2017) <http://doi.org/10.5281/zenodo.1344530> Geraadpleegd 1 oktober 2018

¹⁹ Meer over de FAIR principles op <https://www.force11.org/group/fairgroup/fairprinciples> Geraadpleegd 1 oktober 2018

gekocht systeem of wel door zelf een omgeving te creëren. Impliciet is er dus veel kennis aanwezig over hoe we digitale duurzaamheid in de praktijk moeten doen. De review van OAIS is als het ware een evaluatie, maar de beperkte bijdrage daaraan vanuit de community zou een andere aanpak vragen om te testen of OAIS nog wel aansluit bij de praktijk. Preservation is nu gebaseerd op gedeelde opinies over hoe je iets aanpakt. Er is een reeks van maatregelen die inmiddels door iedereen als vanzelfsprekend worden aangenomen: fixity checks, file format checks, metadata, context informatie, persistent identifiers. OAIS heeft zeker invloed gehad op de ontwikkeling van al deze zaken. Nu wordt het tijd om eens een praktische gids op te stellen waarin de OAIS-concepten gelinkt worden aan de praktijk: om succesvol te zijn doe je dit zo en dat zo. Zodat duidelijk is waar OAIS teveel een keurslijf is en waar het kansen biedt tot bijvoorbeeld een andere interpretatie bij andersoortig digitale objecten. Dat zou ook ondersteunend werken bij de certificeringsstandaard, waardoor instellingen door een self assessment kunnen vaststellen in hoeverre ze al klaar zijn om de officiële certificering aan te vragen, bijvoorbeeld met het Scoremodel. Maar belangrijker is dat we na 20 jaar eens checken of OAIS ook het heil gebracht heeft dat we ervan verwachten: zijn de objecten echt nog duurzaam toegankelijk als de OAIS-standaard gevolgd is tijdens de bewaring?

The Open Preservation Foundation Reference Toolset

Extracting meaning from stored bits and bytes

Introduction

As the digital revolution continues and we generate more and more content, modern archives and libraries are needing to apply time proven and trusted techniques to content that has been digitised or that has been born digital. The Third Industrial Revolution, or the Digital Revolution, refers to the advancement of technology from analog electronic and mechanical devices to digital technology. Since 1980 content has been moving online and digital.

Catalogues and Metadata

From a simplistic point of view a librarian or archivist is tasked with selecting what needs to be kept, and keeping it. However simply putting a pile of books into a box in the attic is not helpful if you need to rapidly find something in one of the books.

There is a long history and science of techniques and practices to enable librarians and archivists to achieve a well-managed collection. Cataloguing and classification being two of the fundamental areas covered. A good catalogue will tell you where and how to find the information you are looking for. It may tell you where it is physically kept, what form it takes, some history of the source and more. Libraries and archives have different standards and approaches as appropriate to the type of information stored, and the reason for that storage.

Collectively the information about an item (as opposed to the item itself) may be referred to as metadata. Metadata will tell you a range of things about an item and the data that it holds. Metadata may be descriptive, structural or administrative.

- Descriptive metadata describes an item – it may be things like the title, author, or an abstract and key words.
- Structural metadata describes how the item is contained – it may be things like a book, magazine, hardback or pamphlet.
- Administrative metadata describes how the item is managed – it may be revision number, how and when it was created or similar.

It quickly becomes obvious that without a catalogue that gives useful and selective metadata that box in the attic becomes hard to use. Opening such a box put away

many years ago you would need to get each object out and examine it to re-create that metadata. Happily, with a box of books it is fairly clear to see the physical object and to read its title and even glance inside to see what it is about and when it was published.

Once that single box of books becomes a vast warehouse full of boxes of books – the task becomes daunting in scale, but the principle is still valid even though it might take an army of people to read all of the books in the boxes.

Imagine however that your books have no covers and bindings. The box now contains a set of paper sheets. If the sheets remain in series it is still possible to recreate the complete book. And whilst that might turn into a massive jigsaw puzzle – finding the right sheets in the right order for each book – there may be clues such as page numbers, chapter numbers and the like to help. A lengthy and tedious process but still achievable.

Catalogues and Metadata and digital objects (blobs)

In the digital world the problem is similar, but worse. Instead of books, imagine we have a virtual representation of a book. A digital object made up from a collection of bits and bytes.¹

This digital object may be made up of many smaller pieces (much like pages in a book, or pictures in an album) and a term that is used to describe a generic storage object that can take a variety of formats is a blob.

If we want to store that blob, like the books we might put it into a box and place it in the attic. We now have a blob in a box. Indeed, we may have a whole bunch of blobs in a box, left in the attic.

Books have an established form. Book covers and binding have a recognised physical shape and can usually be clearly recognised. Computer storage technology is still in relative immaturity, and has a rate of change that is showing no signs of slowing down. Simple word-processed files from 20 years ago are utterly incompatible with modern programmes, even ignoring the alterations in the physical media upon which it is stored.

So now we have a box of blobs from the attic and we need to look at some data. Unlike before we can't just open the box and read the books. When we look into the box of blobs, we aren't even able to identify any books, covers or pages, let alone the order they should be in. And what is worse the 'language' is something we don't recognise.

¹ Computer data storage and processing architecture is broken down into: bits (each bit being a one or a zero), gathered together in bytes (recent convention is 8 bits make a byte, but historically that may differ in 1995 the DEC PDP8 had a 12 bit architecture), gathered together in words (several bytes handled in one operation).

To find meaning in this data you would need to be a digital archaeologist and to have a modern version of the Rosetta Stone.² Even then, this is a task much less attractive than piecing together the pages of a physical book.

Metadata is key to using digital data on a computer. The structure of files, directories and naming hides a vast amount of data that we take for granted. So much so that modern day personal computers don't often even show the file extension (.PDF, .DOCX etc.) that is critical to how the machine interprets the data.

In this digital world, how does an archivist or librarian determine what is important to store about a digital object?

What is it that may be necessary to provide the information to interpret the collection of bits and bytes that we store away?

Assumptions of today's technology may not be valid nor even rational assumptions in the future based on the rate of change that we have seen over past decades, so how do we make sure that we record the relevant metadata so that we don't need to guess?

Data becomes information only when it has the relevant context. The ultimate context is the entire world at a given date, so in this respect the art of the digital preservation is to determine the appropriate and relevant context for a digital object so that it can be interpreted as information at some future date yet to be determined.

The Open Preservation Foundation

The EU sponsored a four-year (2006-2010) project called PLANETS³ (Preservation and Long-term Access through Networked Services) that aimed to build practical services and tools to help ensure long-term access to digital cultural and scientific assets. Following that project the Open Preservation Foundation (OPF) was set up to provide a long-term sustainable home for such open source digital preservation tools.

The PLANETS consortium, coordinated by the British Library, brought together expertise across Europe from national libraries and archives, leading research universities and technology companies. Most of the members of that consortium remain members of the OPF today. Over the past eight years OPF membership has grown, as has the size of the organisation.

OPF continues to enable knowledge sharing and best practice work in addition to adopting and sustaining digital preservation tools that are used by the international community. OPF relies on its members for guidance and practical feedback on the tools and resources under its stewardship and welcomes new members on a regular basis.

Today OPF manages a reference set of tools for digital preservation, and to put them into context has developed a generic preservation process map.

² The Rosetta stone, found in 1799 and dated back to 196BC, carried three versions of the same decree issued by King Ptolemy V. with one version in Egyptian Hieroglyphics and another in Ancient Greek this allowed interpretation of other hieroglyphic texts.

³ <https://planets-project.eu/> last accessed 3/10/18

Digital Preservation Process Map

Once a librarian or archivist has identified that a ‘thing’ should be preserved, they need to catalogue and classify the item. This is as true with a digital ‘thing’ as a physical thing. One key question is how much context around the thing do you need to record to be able to interpret the information at a later date.

A great advantage with a digital thing is that much of the structural and technical context can be collected programmatically. At the time of wanting to put the thing into a box, we have the computer context to hand.

The process has several distinct stages. Principal of these are unpacking, identifying, validating and characterising the thing and the data in that thing.

One golden rule is that the digital thing to be preserved should not be changed or amended, unless a deliberate decision is made to make a copy of the original. Policies for translation, migration or application of fixes or redactions are highly dependent on the purpose and goal of the preserving organisation. This then implies that each thing has its associated metadata stored alongside it. This second item – a metadata thing – must then be stored away in the box too. The structure and standards⁴ for such a metadata thing is beyond the scope of this article.

The usual high-level model in use in digital preservation is the OAIS⁵ and the process map described here is part of the ingest / pre-ingest function within that model. It accepts a Submission Information Package (SIP) and processes it for storage as an Archival Information Package (AIP) to be later retrieved as a Dissemination Information Package (DIP).

There are many ways that the OAIS model has been implemented in various organisations, and the standards continue to develop. The European E-Ark project is part of that development and standardisation process covering a wider scope of the model than described in this article.

The existing landscape of tools, processes and policies has grown up from practice and research in a number of major organisations across the world. Typically however this has been an under resourced area, and to help it develop there have been a number of European Union funded projects to deliver aspects, and individual tools and elements contributed from individuals and community minded organisations.

The Open Preservation Foundation was born from one such European funded project – PLANETS and has participated in several more along the way and now exists to support a set of open source tools – freely available and long term sustainable – to support a section of the ingest / pre-ingest function of the OAIS model.

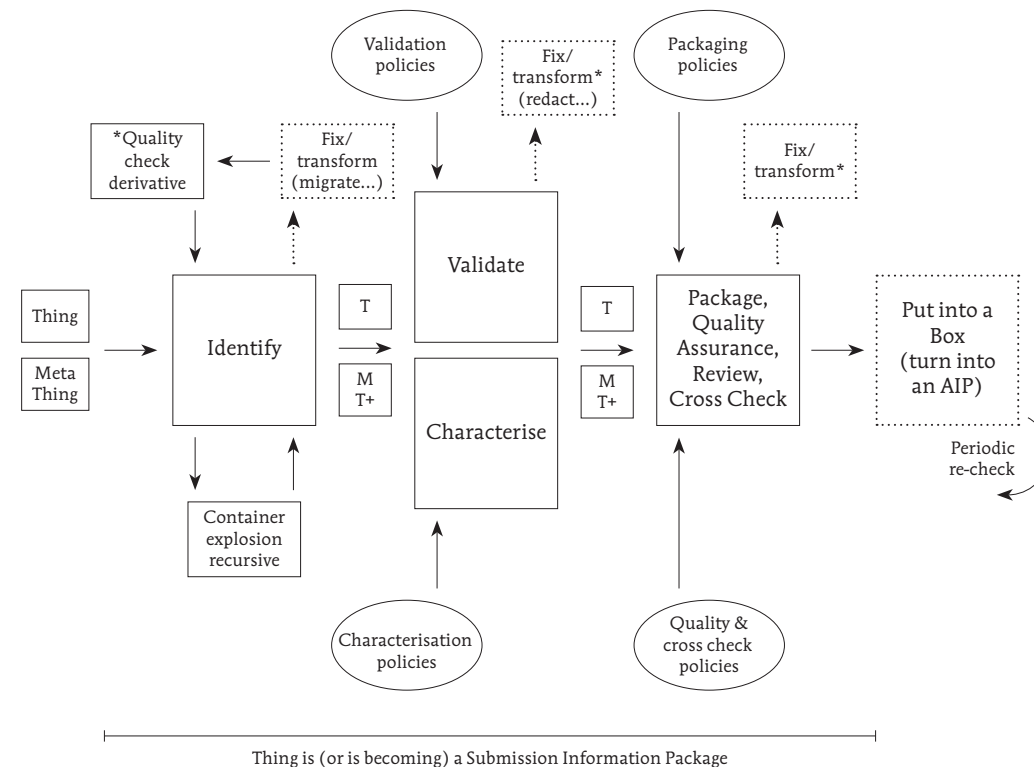


Figure 1 Digital Preservation process map for OPF Reference Toolset

Before we can even start looking at a thing, there is a stage in the process that solves any issues with physical media. For example extraction from a floppy disc, JAZZ drive, micro tape or any other storage media device. There are many challenges in that, but these are not addressed in this article, nor by the work of OPF. We start this process at the point of having a digital thing that we can programmatically examine.

One important factor to point out here is that the sub set of the OAIS process identified by OPF is applicable to virtually every organisation, and doesn't vary much depending on the purpose or specific needs of a digital preservation organisation. Everyone needs to do these common things, although the details, order of processing, policies and choices may well differ.

The typical first step is to *identify* what the thing is. It may be a single file or a complex structure (such as a ZIP file) that contains other discreet structured files and objects. Personal computers today work a very straightforward identification of a file format by use of the filename extension. The reader may like to try an experiment of renaming a file with a filename extension of a different file format type (e.g. rename .docx file as .jpg file) and see that the system is largely unable to cope with such a change (don't forget to rename it back!).

⁴ Metadata standards include PREMIS, METS, MARC, Dublin Core and more

⁵ OAIS – Open Archival Information System is an ISO standard model with six functional stages described in ISO 14721:2012

Simple file extension is a good start, but it has to be confirmed that a file is correctly named, and also what variation, version or sub-type has been used within the structure of the file. Microsoft word files were famously incompatible between versions and yet used the same file extension (.doc⁶)

OPF maintains a tool – *FIDO* – that delivers file identification, and The National Archives of the UK maintains an equivalent tool called *DROID*. The principle way these tools work is to look at the file and identify signature strings of data. For example, a PDF file will contain the tell-tale string as the first character “%PDF – 1.2” or similar, where the number indicates the version of the PDF specification. The main challenge in identification of files using this technique is maintain the list of signature strings. Most of the tools and indeed virtually every organisation relies upon the authoritative database of file signatures called PRONOM which is maintained by The National Archives of the UK.

In software engineering it is generally held true that the standard case is relatively easy, clearly identifiable common cases work well (the so called ‘happy path’). It is when things are unusual, different, malformed or erroneous that causes most of the work. This is very much the case in file identification too.

One significant complication is the idea of a container. A container is a file type that contains other file types, the most common of which would be a zip file. Identification of the container may be straightforward, but not sufficient since much or all of the information is in the files wrapped up inside. The container must be expanded and the individual items contained identified and made available for further analysis. A simple zip may have one or more files of one or more types and a directory of folder structure that may itself provide crucial context and metadata. (Filenames may be duplicated across different branches of a directory inside a zip file).

Probably one of the more complex containers is a complete disk image. The image itself may be the item to be saved, but the identification of all of the sub-items within that image has to be completed before validation and characterisation of each sub-item can be carried out. Clearly some of the items contained on a disk image could be containers – perhaps even more disk images – so the process must be recursive. Happily, the tool *BitCurator* has been developed to assist specifically for disk images as this is a complex form of ‘exploding the container’.

The process of identification will provide an amount of information about a ‘thing’. It may be as simple as a statement of the file type, variety and version, or it may be as much as a complete map of a complex container. All of this information must be recorded as metadata and added to the associated collection of metadata with the digital thing.

If the file is found to be broken at this stage, or indeed of a file type that is undesirable, the process can be extended to transform the format of the thing. Usually this would involve making a copy of the original perhaps in a different format. Once that has been done the original and the transformed copy should

be compared to demonstrate that any differences are deliberate and recorded with the metadata. OPF look after a tool (*xCorrSound*) that does this for certain audio formats.

Once the ‘thing’ has been identified, we move on to check if the thing is well-formed and complies to the rules for that format (Validation), and also to find out more details such as author, creation time, tools used to create the thing and more (Characterisation).

Validate and *Characterise* are logically separate processes that both require the structure of the thing to be opened up and examined, or parsed. That opening up requires an understanding of the low-level structure used in the thing. However, the two tasks are looking at very different aspects. The key tool used for this is OPF’s *JHOVE*. Whilst the principle applies across all file types, each file type will require a specific process. Validating a text based format such as PDF⁷ compared to an image format such as JPEG⁸ (JPG) require very different logic and coding.

Validation is confirming that the file conforms to the current understanding of the format. Typically that means checking every aspect of the format and the files compliance to that standard. OPF recently took part in the PREFORMA project to develop a new PDF/A validator VeraPDF. Developed in association with the PDF Association and adhering tightly to the standards for PDF/A the project developed not only a validation tool that looked into more corners of the specification than ever before, but also developed a comprehensive body of test data (test corpus).

Modern file formats have developed over a number of years and many can be remarkably complex in themselves. While this makes those formats very functional, it does add complexity to the validation since few implementations of any given specification will be identical. Again, the issue of files that comply is straightforward, and the complexity is when there is a real or technical non-compliance. Any given non-compliance may not necessarily be severe enough to prevent a future format reader making sense of the document – but can we be sure of that? Does a ‘format non-compliance’ mean the same for all types of content? For example the definition of a colour palette may be incomplete in a document – no problem for a black and white text page – massive issue for a high definition picture of fine art. It may be that today’s rendering tool automatically compensates for the colour palette, or the local language settings or any one of a range of contexts set in the file.

One format that most of us assume to be safest – PDF – is defined by a particularly complex specification. There are numerous common errors in formats when compared to the specifications that today’s renderers ignore or pass over, but in theory could cause major issues in the future. It comes back to the librarian or archivist to determine the criticalities for the type of content that they are

⁶ DOC format may use one of four different formats on a PC from different versions of Word– Word for DOS, Word for windows, word 6, word97. Since word2007 the .DOCX format is used.

⁷ *Portable Document Format (PDF)* is a file format developed in the 1990s to present documents, including text formatting and images, in a manner independent of application software, hardware, and operating systems. Based on the PostScript language, each PDF file encapsulates a complete description of a fixed-layout flat document, including the text, fonts, vector graphics, raster images and other information needed to display it. Since 2008 PDF has been standardized as a royalty free open format, ISO 32000 (and others)

⁸ *JPEG* Joint Picture Expert Group – an ISO format for images

preserving, and as such the organisation needs a set of policies to determine which validation errors can be ignored and which must be corrected prior to storage.

At the end of the validation process, the metadata that accompanies the digital thing will be updated to carry details of the validation, the policy choices made and the results of the checking. Positive and negative metadata may be included and once again, this is a choice of policy.

Characterisation involves taking information about the thing and the nature of its content to build a richer context beyond the purely technical, structural details. Many formats hold a wealth of information concerning the content – even before looking at the content itself.

For example, a word document usually records the author, the time of creation, the editing time, word count and more. Sometimes there is a specific summary of the document as well. And a JPEG image file may carry information about the camera used to take it, time and day and often even geolocation information for its precise location.

However, the librarian or archivist needs to make a series of decisions about such data. The instinctive response might be to extract everything, going to limits in a jpeg for example finding colour profile, or raw image sensor info. Such information it could be argued is invaluable in the case of checking quality should the file format require migration.

The right level of detail needed depends on the intention and processes in place for the organisation. Once sufficient information has been taken to ensure that the original file can be adequately indexed and examined in future, any requirements for further descriptive data could be extracted at a later date. However, the presence of such data in the file and the policy for data to be extracted and duplicated outside the file is itself a key piece of metadata to be stored alongside the digital thing.

Package and cross check is a process not yet explicitly covered by OPF tools, but usually carried out through workflows and processes and report formatting in each organisation. This is where the metadata gathered is prepared for taking forward to storage in the database. In effect one of the final stages of forming and wrapping the thing to be stored into an appropriate archive information package.

There are a range of metadata formats and standards such as PREMIS⁹ and METS¹⁰ and others that proscribe a structure for metadata. Librarians and archivists need to determine the most suitable form for their purposes and which elements of that to populate.

⁹ PREMIS Data Dictionary: a comprehensive, practical resource for implementing preservation metadata in digital archiving systems; accompanying report (providing context, data model, assumptions); special topics, glossary, usage examples; set of XML schema which was developed to support use of the Data Dictionary.

¹⁰ Metadata Encoding and Transmission Standard (METS) is a metadata standard for encoding descriptive, administrative, and structural metadata regarding objects within a digital library, expressed using the XML schema. The standard is maintained as part of the MARC standards of the Library of Congress, and is being developed as an initiative of the Digital Library Federation (DLF).

Having chosen that format and the required elements, the metadata information gathered so far must be placed into that format in readiness to archive the digital object. However, there is one other aspect considered here. There will be occasions where the found metadata may conflict with provided information, for example sometimes the stated author may differ from the authors name as discovered through characterisation of the digital thing. This aspect of cross-checking prevents future ambiguity.

However, such as cross-check implies yet further decisions as to which data is to be trusted? Or should both be stored? Once again, the librarian or archivist has to make a choice depending on their purposes. Without choosing the right level of metadata to store, the danger is that the volume of metadata can easily exceed that of the original item. Items are usually selected based on a clear preservation need, and a fine balance is needed between minimalist data and keeping everything.

OPF Reference toolset and the digital preservation process map

Tool Name	Part of the process	File types handled
Fido	Identification	Based on list of signatures in PRONOM
Format Sniff	Identification	Based on list of signatures in PRONOM
VeraPDF	Validation, Characterisation	PDF/A
Jpylyzer	Validation, Characterisation	JPEG 2000
Jhove	Validation, Characterisation	PDF, JPEG, TIFF, WAV, PNG, WARC, AIFF, XML, HTML, GZIP, ASCII text, UTF8 Text. MP3, GIF, JPEG 2000
xcorrround	Quality Check	Sound files
DPF Manager	Validation, Characterisation	TIFF
PDF Test suite	Test Corpus	PDF/A

Table 1. OPF Tools and how they map

Practical tool usage, scalability and quality control

OPF monitors the number of downloads of the tools that it supports, and is in discussion with the users to design and implement a more sophisticated monitoring framework. Naturally much of the operations in which the tools are used are sensitive and OPF is mindful of preserving that confidentiality.

However, building a set of usage data will enable OPF and tool users to build statistics on the types of formats processed and format errors found. This will help eliminate systematic errors in generated documents and will quantify the importance of various formats amongst the tool users.

The monitoring framework will also enable OPF to respond to any issues found with the toolset itself. Today OPF runs a comprehensive testing process for each release of a tool, and holds a set of test data of each format to ensure reliability and preserve backwards compatibility wherever possible or practical.

The test data, or test corpus, contains a series of typical documents and specifically designed format error cases to cover the range of situations. This test corpus is added to by the OPF members as new troublesome features of real documents are uncovered in real life usage.

OPF members are actively involved in all aspects of the toolset development. As an example, the OPF Format Interest Group has led the work to provide identifiers for each JHOVE error message and are working towards populating a wiki to document what these errors mean in human readable terms, with links to the relevant standards and specifications.

OPF would like to take this opportunity to call for more members of the digital preservation community to join the effort, either as an OPF member or a technical contributor to the tools, test data or best practice community knowledge. More information is available on the OPF website, <http://www.openpreservation.org>.

OPF Reference toolset future developments

The OPF toolset development roadmap is governed by a Product Board who contributes to and approves individual tool roadmaps. The wider community provides feedback on fixes and features they would like to see, which are then prioritised by the Product Board. Current trends include making the larger tools more modular to accommodate individual workflows better and building up the body of test data for each of the main preservation formats.

OPF delivers all of its tools on an Open Source basis with licences that allow the tools to be incorporated into an organisation's proprietary software or into commercial systems. Major tools typically will have a couple of releases each year, and the roadmap of improvements is discussed, debated and agreed by the OPF members at the annual general meeting.

OPF also co-operates on the Bitcurator project, the E-ARK4ALL project, and the Jisc PAR project.

Why OPF operates cooperative, open source projects

OPF is a membership organisation and an experienced leader in open source digital preservation software development. For more details or to join please see <http://www.openpreservation.org>.

BARBARA SIERMAN, MARGRIET VAN GORSEL,
REMCO VAN VEENENDAAL

Preservering: van Beleid naar Beheer

Inleiding

Dit artikel gaat over de waarde en het gebruik van de catalogus van elementen van duurzaamheidsbeleid uit het Europese project SCAPE (*Scalable Preservation Environments*). Na de toelichting op het belang van dergelijk beleid voor collectie beherende instellingen volgt het ontstaan van de catalogus en de relatie met het referentiemodel OAIS (*Open Archival Information System*). Het artikel sluit af met twee voorbeelden van toepassingen van de catalogus: bij het opstellen van een beleidsplan voor preservering en bij de samenwerking tussen collectie beherende organisaties.

Uitgangspunten en Beleid

Elke organisatie met een digitaal archief (ook wel "collectie beherende organisatie") waarin collecties voor de lange termijn toegankelijk worden gehouden, neemt in de loop der jaren beslissingen over hoe dit gebeurt. Aan deze beslissingen ligt soms wel een reeks uitgangspunten ten grondslag en soms nog niet. Er zijn organisaties die zelf de digitale archivering doen en het referentiemodel OAIS gebruiken als leidraad, inclusief bijbehorende terminologie, functies en informatiemodel. Er zijn ook organisaties, die de digitale archivering uitbesteden aan een andere organisatie maar daarbij wel willen dat dit volgens hun eigen beleid gebeurt.

Nog lang niet alle collectie- en archiefbeheerders passen kennis van preservering en het overkoepelende thema duurzame toegankelijkheid toe. In Nederland zijn kennis en uitgangspunten bekend bij een deel van de organisaties, maar die organisaties stuiten dan op de vraag: "Hoe zetten wij de stap naar het vastleggen van de uitgangspunten en daarna het toepassen door iedereen in de organisatie?" Dit veranderproces kost tijd, vereist managementaandacht en moet actief ondersteund worden. Door de implementatie van preservering uit te stellen of slechts technisch te benaderen, loopt de organisatie risico's. Niet alleen kan de kwaliteit van de duurzame toegankelijkheid bedreigd worden, maar ook de bijbehorende kosten zullen onduidelijk blijven als niet volgens vastgelegde uitgangspunten wordt gewerkt (zie het artikel van H. Uffen over het kostenmodel verderop in dit jaarboek).

Een collectie beherende organisatie maakt onderdeel uit van een groter netwerk en is vrijwel altijd een schakel in een informatieketen. De inrichting van preservering van de collectie vraagt om maatregelen die over de grenzen van de organisatie heen

lopen. Verschillende betrokken partijen willen weten hoe de duurzame toegankelijkheid is geregeld. Bijvoorbeeld depotgevers die hun digitale materialen toevertrouwen aan de organisatie. Zij willen gerustgesteld worden dat hun collecties in goede handen zijn. Steeds vaker wordt dan ook de eis gesteld dat het digitaal archief gecertificeerd is – en daar maakt het hebben van duurzaamheidsbeleid¹ een essentieel onderdeel van uit. Ook subsidiegevers stellen eisen en willen duidelijkheid over aanpak en werkwijze van de organisatie. De buitenwereld is een grote belanghebbende, maar de binnenwereld eveneens. Medewerkers die in hun dagelijkse werkzaamheden geen richtsnoer hebben, zullen zich gedwongen voelen hun best te doen zonder dat zij daarbij rekening houden met andere uitgangspunten. Kortom: duurzame toegankelijkheid vraagt om een samenhangend duurzaamheidsbeleid om ad hoc beslissingen te voorkomen, om depotgevers, subsidieverleners en medewerkers en anderen inzicht te geven in de wijze waarop de organisatie de duurzame toegang realiseert en om langetermijnrisico's te beperken.

Wat is duurzaamheidsbeleid?

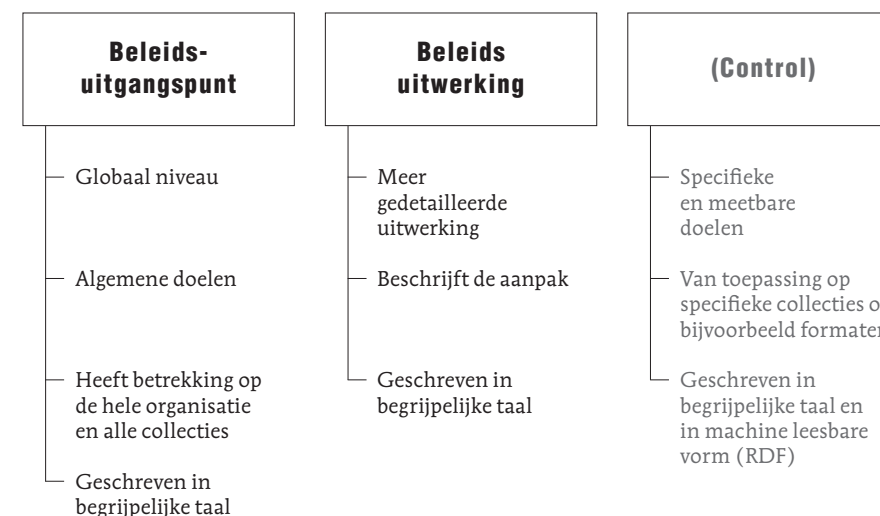
Het referentiemodel voor digitale duurzaamheid OAIS (ISO1721) is vrij beknopt over duurzaamheidsbeleid en geeft aan dat beleid door de functie Management wordt onderhouden. De gelieerde certificeringsstandaard, *ISO 16363 Audit and certification of trustworthy digital repositories*², gaat verder in op duurzaamheidsbeleid en ziet dit als onderdeel van een breder beleidskader, het Policy framework, waarin ook de missie en visie en een strategisch plan van een organisatie een plek vinden, oftewel:

*“The Preservation Policy then declares the range of approaches that the repository will employ to ensure preservation (that is, to implement the Preservation Strategic Plan), and finally the Preservation Implementation Plan translates those into services that the repository must carry out.”*³

Maar ook deze omschrijving is vrij abstract. Binnen het Europese project SCAPE is gewerkt aan een handzame catalogus om dit begrip hanteerbaar te maken.

Het SCAPE Policy Framework

Tijdens het Europese project SCAPE (2011-2014) is een kader voor duurzaamheidsbeleid ontwikkeld. Op hoog niveau zijn in dit kader een aantal onderwerpen vastgelegd: de 'beleidsuitgangspunten' (*Guidance policies*). Hoe deze uitgangspunten in praktijk worden aangepakt, is omschreven op het tweede niveau: de 'beleidsuitwerking' (*Preservation Procedure policies*). Het derde niveau dat in SCAPE genoemd wordt, de *Control policies*, is een gedetailleerde invulling van de beleidsuitwerking, gericht op een specifieke collectie, een specifieke activiteit voor een specifieke doelgroep. Dit niveau is in SCAPE niet nader uitgewerkt. Het SCAPE-beleidskader sluit nauw aan bij het model dat in de certificeringsstandaard ISO 16363 is ontwikkeld. Beide ontwikkelingen liepen indertijd vrijwel parallel. Het in bovengenoemde citaat beschreven *Preservation Strategic Plan* komt overeen met de beleidsuitgangspunten. De *Preservation Policies* zijn te vergelijken met de beleidsuitwerking. En het *Preservation Implementation Plan* beschrijft de gedetailleerde invulling, en komt overeen met de *Control policies*.



Afbeelding 1. SCAPE Policy Framework

Praktische vertaling van dit SCAPE-beleidskader naar het duurzaamheidsbeleid van een organisatie vraagt om concrete voorbeelden. Op basis van bestaande literatuur over duurzaamheidsbeleid van onder andere Neil Beagrie⁴ heeft het SCAPE projectteam een tiental onderwerpen geselecteerd die in de beleidsuitgangspunten benoemd zouden moeten worden. Om te zien wat de bijbehorende beleidsuitwerking zou moeten bevatten, is gekeken naar bestaande en publiek beschikbare documenten met duurzaamheidsbeleid van universiteiten, datacenters, archieven en bibliotheken. Uiteindelijk resulteerde dit in een overzicht van zowel beleidsuitgangspunten als beleidsuitwerkingen in de *Catalogue of Preservation Policy Elements*⁵. Deze Engelstalige catalogus is te vinden op de website van de Open Preservation Foundation.

Nederlands duurzaamheidsbeleid

De kennis van beleidsaspecten helpt bij een goede inrichting van duurzame toegankelijkheid van informatie- en erfgoedcollecties. De oorspronkelijk in het Engels opgestelde beleidsuitgangspunten in de SCAPE-catalogus zijn in 2016 vertaald door de Nationale Coalitie Digitale Duurzaamheid (NCDD, ondertussen opgegaan in het Netwerk Digitaal Erfgoed). Het eerste resultaat⁶ is vormgegeven als een wiki en biedt een totaaloverzicht, definities en use cases. Bij elk onderwerp staat onder het kopje Discussievragen een aantal voorbeeldvragen die de instelling aan zichzelf zou kunnen stellen. De beleidsuitgangspunten zijn opgebouwd vanuit de digitale

¹ In deze tekst gebruiken we duurzaamheidsbeleid als vertaling van de Engelse term Preservation Policy.

² <https://www.iso.org/standard/56510.html> Geraadpleegd op 20-08-2018

³ ISO 16363, pagina 1-4

⁴ Charles Beagrie e.a. *Digital Preservation Policies Study* (JISC, 2008) zie https://www.webarchive.org.uk/wayback/archive/20140615022334/http://www.jisc.ac.uk/media/documents/programmes/preservation/jiscpolicy_p1finalreport.pdf Geraadpleegd op 20-08-2018

⁵ *SCAPE Catalogue of Preservation Policy Elements* (2014) zie <http://wiki.opf-labs.org/display/SP/Catalogue+of+Preservation+Policy+Elements> Geraadpleegd op 20-08-2018

⁶ <http://wiki.ncdd.nl/index.php?title=Duurzaamheidsbeleid> Geraadpleegd op 20-08-2018

informatie zelf, en gaan over relatief kleine onderwerpen, zoals integriteit, en over grotere, zoals certificering. Hierdoor zijn de uitgangspunten bruikbaar vanuit verschillende collectie behorende rollen in een organisatie. Sommige onderwerpen lenen zich meer voor de beleidsmedewerker (strategisch), sommige voor de inrichter van een systeem (tactisch) en sommige voor de preserveringsmedewerker (operationeel). Op dit moment wordt de wiki omgezet naar een nieuwe Wegwijzer Duurzaamheidsbeleid⁷. Daar is een aantal wijzigingen aangebracht ten behoeve van het gebruik.

Schijn en werkelijkheid

De eerder genoemde analyse in het SCAPE-project van bestaande documenten over duurzaamheidsbeleid bracht nog iets anders aan het licht. Veel instellingen zien het document met het beschreven duurzaamheidsbeleid als een gelegenheid te tonen hoe goed ze bezig zijn, dus met name gericht op de buitenwereld. Zoals eerder gezegd is het goed de direct betrokkenen als subsidiegevers en depotgevers gerust te stellen. Maar in menig beleidsstuk werd bijvoorbeeld emulatie als gekozen strategie beschreven, terwijl op het moment van opstellen dit geen algemeen aanvaarde oplossing was (te duur, technisch te ingewikkeld). Anderen streefden naar ISO-certificering, op het moment dat er nog geen officiële instellingen waren die deze certificering konden uitvoeren. Ook stond vaak de intentie genoemd het duurzaamheidsbeleid bijvoorbeeld tweejaarlijks te evalueren, maar die twee jaar waren soms ruimschoots voorbij. Dergelijke loze beloften en irreële intenties schieten voorbij aan het doel van duurzaamheidsbeleid: de interne medewerkers een handreiking te bieden en de externe belanghebbenden te informeren.

Maturity levels

Het werken met *maturity levels* biedt een meer realistische werkwijze. Het is belangrijk dat het duurzaamheidsbeleid van een instelling weergeeft hoe 'volwassen' de instelling is in het uitvoeren van de langetermijnbewaring van de collecties. Geen enkele instelling is meteen perfect. Alle instellingen hebben een groeicurve waarin zij met preservering beginnen op een lager niveau dan zij na een aantal jaren, door ervaringen wijzer geworden, zijn aanbeland. Een goed geschreven duurzaamheidsbeleid reflecteert dit. Vandaar dat dit beleid regelmatig moet worden herschreven, om nieuwe ontwikkelingen weer te geven en aan het groeipad van de instelling recht te doen.

Een organisatie die pas begint met het realiseren van duurzame toegankelijkheid (level 0 in het *Digital Preservation Maturity Model* van Dollar en Ashley⁸, zal zich oriënteren op beleidsuitgangspunten, maar nog niet direct aan de beleidsuitwerking kunnen werken. Pas op niveau 2, wanneer de organisatie ervaring heeft opgedaan met het inrichten van duurzame toegankelijkheid wordt de beleidsuitwerking opgeschreven, zodat medewerkers hiermee bekend raken. In de praktijk echter vindt vaak het omgekeerde plaats: een organisatie begint met de beleidsuitwerkingen zonder dat ergens op papier de uitgangspunten zijn vastgelegd.

Het schrijven van goed duurzaamheidsbeleid hoeft niet vanaf nul te beginnen. Vaak is er binnen de instelling al geschreven en ongeschreven informatie aanwezig, die als basis kan dienen. Door deze losse informatie bij elkaar te brengen via het beleidskader, kan er een expliciet, overzichtelijk en samenhangend duurzaamheidsbeleid tot stand komen.

Toepassing van de catalogus van elementen van duurzaamheidsbeleid

In de catalogus worden de elementen authenticiteit, bit preservering, functionele preservering, digitaal object, metadata, rechten, standaarden, toegang, organisatie, en audit en certificering nader uitgewerkt in procedures. Deze procedures worden gedefinieerd, uitgewerkt in discussievragen en geïllustreerd met voorbeelden. Dat ziet er zo uit:

Het element Integriteit	
Gerelateerd Beleidsuitgangspunt	Authenticiteit
Omschrijving	Integriteitscontrole omvat maatregelen met betrekking tot encryptie, digitale handtekeningen, fixity checks etc. Zie ook de uitleg onder sectie Bit preservering: Integriteitsmaatregelen
Waarom	Een van de belangrijkste doelstellingen van digital preservation is dat de gepreserveerde digitale objecten, wanneer ze eenmaal in het digitaal archief zijn opgeslagen, niet onopzettelijk kunnen worden gewijzigd.
Risico	Als de organisatie niet de maatregelen uiteen zet die zullen worden genomen om onvrijwillig verlies van data te voorkomen, kunnen de organisatiedoelstellingen wellicht niet worden behaald. De risico's kunnen zich voordoen tijdens vele fases in de digitale life cycle. Een digitaal archief dient een set aan maatregelen te beschrijven die worden geïmplementeerd om risico's af te dekken. De voorgestelde maatregelen liggen dicht tegen de operationele IT taken aan en zijn vaak al onderdeel van hun werk, maar de preservation policy moet expliciet maken dat deze maatregelen zullen bijdragen aan de authenticiteit van de digitale objecten. Aandachtsgebieden zijn: Ingest: De compleetheit van het digitale object moet worden gedefinieerd voorafgaand aan de ingest en zou onderwerp van gesprek kunnen zijn met de leverancier van de content of de producer. Gedurende de ingest kunnen de ontvangen checksums worden vergeleken met de checksums die worden gegenereerd tijdens inname. Hiermee kan worden aangetoond of bits tijdens het transport verloren zijn gegaan. Deze maatregel zou moeten gelden voor alle verplaatsingen van data, inclusief het verplaatsen van data binnen het digitaal archief. Storage: voor en na het verplaatsen van data van de ene naar de andere locatie moeten maatregelen worden genomen om te controleren of het digitale object nog steeds volledig en onbeschadigd is. Dit geldt ook voor back up kopieën. Door middel van authenticatie moet worden gewaarborgd dat medewerkers opgeslagen data niet
Fase levenscyclus	Preservation Planning, Ingest proces, Preserveringsactiviteit, Receive, Storage
Belanghebbenden	Management: stelt de algemene maatregelen vast waarmee integriteit wordt gewaarborgd Operationeel management: implementeert deze maatregelen Producer/Depotgevers: draagt bij aan het de integriteits maatregelen (bijvoorbeeld het sturen van checksums met het digitale object)
Relaties	Bit preservering Functionele Preservering [migratie]
Voorbeelden	Parliamentary Records: "The record must be maintained to ensure that it is complete, and protected against unauthorised or accidental alteration. In this Policy, integrity is ensured through the bitstream preservation function [...], and through the provision of metadata to describe all authorised actions undertaken in the course of content and bitstream preservation." URL: http://www.parliament.uk/documents/upload/digitalpreservationpolicy1.0.pdf
Discussievragen	Gaat uw organisatie in gesprek met de producer of de leverancier van het digitale materiaal over gelijktijdig aan te leveren checksums, zodat na aanlevering van het digitale materiaal checksums kunnen worden vergeleken? Accepteert uw organisatie de aanlevering van digitaal materiaal zonder checksum? Creëert uw digitaal archief zelf een checksum bij de ontvangst van digitale objecten? Zijn er procedures geïmplementeerd in uw digitaal archief om regelmatig de checksums te controleren? Heeft uw digitaal archief procedures om met checksums om te gaan gedurende preserverings acties?

Afbeelding 2. Het element Integriteit ⁹

⁷ <https://wegwijzerduurzaamheidsbeleid.nl> Geraadpleegd op 05-10-2018

⁸ Charles M. Dollar and Lori J. Ashley Assessing Digital Preservation Capability Using a Maturity Model Process Improvement Approach (2013) zie: https://www.nycarchivists.org/Resources/Documents/Dollar-Ashley_2013_DPCMM%20White%20Paper_NAGARA%20Digital%20Judicial%20Records_8Feb2013-1.pdf Geraadpleegd 07-10-2018

⁹ <http://wiki.ncdd.nl/index.php?title=Duurzaamheidsbeleid/Beleidsuitwerking-Integriteit> Geraadpleegd op 07-10-2018

De catalogus geeft dus een overzicht van elementen waar een organisatie aan moet denken bij het maken van duurzaamheidsbeleid. Ook inspireert het om te zien hoe anderen met deze elementen omgaan. Toch is er iets extra's nog nodig om van beleid naar implementatie van een beleidsplan te komen. Hieronder volgen twee voorbeelden van hoe het Nationaal Archief (NA) de stap heeft gezet. Het eerste voorbeeld werkt binnen een organisatie. De tweede situatie beschrijft hoe de catalogus gebruikt is in de samenwerking tussen meerdere organisaties.

Voorbeeld 1: Check voor de organisatie, uitwerking van elementen en procedures

Indien een collectie beherende instelling beleid wil opstellen kan de catalogus helpen. Voor de organisatie is bijvoorbeeld het garanderen van de authenticiteit van de digitale objecten, gedurende het beheer- en bewaarproces, een van de belangrijkste doelstellingen. Lezend in de Catalogus staat er onder het kopje authenticiteit: "Authenticiteit: De mate waarin een persoon (of systeem) een object beschouwt als dat wat het beweert te zijn. Authenticiteit wordt vastgesteld op basis van bewijs."

Daaronder staan onderliggende elementen:

"Integriteit: Integriteitscontrole omvat maatregelen met betrekking tot encryptie, digitale handtekeningen, fixity checks etc. Zie ook de uitleg onder sectie Bit preservering: Integriteitsmaatregelen.

Betrouwbaarheid: Transparante en volledig gedocumenteerde preservingsstrategieën en een voorziening van benodigde metadata om inhoud, context en herkomst te beschrijven.

Herkomst: Het documenteren van de processen in de life cycle van een digitaal object. Digitale Herkomst beschrijft over het algemeen de Agents verantwoordelijk voor het bewaren en beheren van Digitale Objecten, belangrijke gebeurtenissen die voorkomen gedurende de levenscyclus van het Digitale Object, en andere informatie die wordt geassocieerd met de creatie, het beheer en de preservering van het Digitale Object."

Om van deze elementen en procedures te komen tot de implementatie van beleid in een beleidsplan, heeft het NA deze elementen verder geoperationaliseerd. Hieronder staat het voorbeeld van het element Integriteit. Het NA volgde de weg van een kwaliteitsmanager of auditeur. Om tot goede indicatoren van een dergelijk begrip te komen, is het handig om vragen te stellen als: "Wat is de definitie van een begrip/norm? Hoe kan worden bepaald of aan die definitie/norm wordt voldaan? Op welke wijze geeft de organisatie dit vorm?"

Het NA heeft op basis van de catalogus de volgende indicatoren gekoppeld aan Integriteit.

Beleidsuitgangspunt	Beleidselement	Indicatoren
Authenticiteit	Integriteit	Procescontroles informatieobjecten en metadata met aandacht voor a. Checksums die vóór instroom in de repository door de producer worden gemaakt en meegeleverd b. Checksum controle in de repository c. Controle op encryptie en de aanwezigheid van digitale handtekeningen
Authenticiteit	Integriteit	Koppeling object en metadata door a. Metadata Integrity check: er wordt nagegaan of alle content files zijn gespecificeerd in de metadata middels de correcte (relatieve) locatie; b. Content Integrity check: er wordt nagegaan of de content files zijn gespecificeerd in de metadata xml en dat dit consistent gebeurt.
Authenticiteit	Integriteit	a. Beschrijving bitstream preservation functie b. Beschrijving logging van alle geautoriseerde acties c. Beschrijving check compleetheit van de aangeleverde files, het voldoen van het formaat aan de standaard en de aanwezigheid en de kwaliteit van de metadata.

Tabel 1. NA-indicatoren voor Integriteit

Door deze uitwerking in indicatoren ontstaat bij de organisatie meer gevoel en kennis over dat wat geregeld moet worden. Hierdoor kan ook makkelijker collegiale afstemming plaatsvinden over de inrichting van systemen, procedures en rollen en taken. Wat moet de organisatie doen om, in dit geval, Integriteit te implementeren, is het makkelijk of moeilijk, wat zijn de risico's, welke middelen heb ik er voor nodig, welke competenties komen erbij kijken en hoe verloopt het proces? Deze operationalisering biedt de organisatie meer houvast bij het maken van een beleidsplan. Wellicht dat niet nodig is om alle elementen tegelijk uit te werken. De organisatie kan kiezen wat voor hen op dat moment belangrijk is om te regelen.

Voorbeeld 2: Check voor de samenwerking, bespreekbaar maken van elementen en procedures

De tweede toepassing gaat over het gebruik van de catalogus indien preservering van een collectie of collectieonderdeel niet bij de eigen organisatie plaatsvindt. Het zal in de toekomst steeds gewoner worden om het beheer van een digitale collectie te *outsourcen* naar een omgeving die daar geschikt(er) voor is. Het aanbod van beheeromgevingen neemt ook toe. Sommigen zijn gespecialiseerd in beheer van specifieke typen informatie, of de samenwerking met een specifiek type producent. Denk in dit geval aan Beeld en Geluid (B&G) voor audiovisuele informatie of aan DANS¹⁰ of 4TU¹¹ voor wetenschappelijke data en software. De ene beheeromgeving (inclusief organisatie) kan meer geschikt zijn voor het beheer van een type informatie dan een andere. Dit kan ervoor zorgen dat een instelling op zoek moet naar een passende beheeromgeving.

¹⁰ <https://dans.knaw.nl/nl> Geraadpleegd op 20-08-2018

¹¹ <https://www.4tu.nl/nl/> Geraadpleegd op 20-08-2018

Een voorbeeld van gebruik maken van elkaars sterke punten is de samenwerking die de afgelopen jaren tussen B&G en het NA is ontstaan met betrekking tot het beheer van audiovisuele informatie. De audiovisuele opnamen van de plenaire zaal van de Tweede Kamer behoren op termijn tot de archieven van het NA. Aangezien B&G haar beheer en raadpleegomgeving heeft gevormd op audiovisuele collecties ligt samenwerking voor de hand. Om tot een goede overbrenging en goed beheer te komen is als eerste een verkenning gemaakt van de collectie. Dit heeft geleid tot een werkverdeling. De tekst-gebaseerde bestanden van de Tweede Kamer worden opgenomen door het NA, de audiovisuele bestanden door B&G. Belangrijk hierbij is het maken van afspraken. Beide partners hebben die nodig om hun verantwoordelijkheid te kunnen nemen en houden. In dit geval geeft het NA een collectie in beheer bij een andere partij. Door middel van afspraken die worden vastgelegd in een dienstenniveau-overeenkomst (DNO) krijgt de samenwerking gestalte als *postcustodial* beheer of *preservation in place*. Ook in dit geval helpt het beleidskader. De uitwerking van elementen in indicatoren en gewenste output voor duurzaam beheer van het NA is gedeeld met B&G. Vervolgens zijn daar de gepubliceerde beleidsstukken van B&G naast gelegd.¹² Voor het element Integriteit leverde al veel informatie op. Een van de opvallende zaken in deze vergelijking is dat, ook al baseren beide organisaties zich op OAIS, de taal en de uitwerking van elementen anders kan zijn. De onderstaande tabel heeft de basis gelegd voor het gesprek tussen NA en B&G over de implementatie van het element Integriteit. Het bleek een goed startpunt voor het overleg tussen NA en B&G over een concrete DNO waarin onder andere de afspraken staan van de partners over het inrichten, monitoren en rapporteren over dit element.

Ter illustratie staat in Tabel 2 (zie hierna) de uitwerking van een deel van het beleids-element 'Integriteit' van het beleidsuitgangspunt 'Authenticiteit'. Omdat nog aan de DNO wordt gewerkt, kunnen we de resultaten nog niet publiceren.

Door lijsten zoals die van Tabel 2 gezamenlijk uit te werken, te bespreken en afspraken te maken, is de basis gelegd voor goede samenwerking en duurzaam beheer. Het gesprek loopt beter, het biedt de mogelijkheid om elkaar aan te spreken op de eigen verantwoordelijkheden. De controles zijn in het proces ingebouwd.

Door op deze manier de beleidselementen uit te werken in indicatoren is samenwerking mogelijk tussen collectie beherende instellingen, waardoor de sterke punten van elke beheer- en raadpleegomgeving optimaal ingezet wordt en beheerders elkaar kunnen helpen.

Beleids-uitgangspunt	Beleids-element	Indicatoren	Output	Partner Bron (in dit geval de beleidsstukken van B&G)
Authenticiteit	Integriteit	Procescontroles informatie-objecten en metadata met aandacht voor a. checksums die vóór instroom in de repository door de producer worden gemaakt en meegeleverd b. Checksum controle in de repository c. Controle op encryptie en de aanwezigheid van digitale handtekeningen	Procesbeschrijving (pre)ingest met daarin de geïmplementeerde controles en frequentie daarvan: - beveiliging (encryptie) - virusscan - digitale handtekening - content en metadata (locatie en xml) - bitstream - logging van Technische controle: - volledigheid collectie - formaat - kwaliteit metadata Basisdocument waarin dit soort keuzes wordt toegelicht en onderbouwd	Deze data-integriteit wordt door Beeld en Geluid o.a. vastgesteld door het gebruiken van checksums. Bij elke bewerking of verplaatsing van een versie van het object voor (hernieuwde) opslag, kopiëring, migratie of voor uitlevering aan gebruikers, wordt met een fixity check teruggegrepen op de checksum, zodat zeker is het object 'qua bits' identiek is aan de content die in een eerdere fase werd overgebracht. Integriteit wordt ook verzekerd door het monitoren van de materialen die reeds in het Digitaal Archief zijn opgeslagen tijdens migraties van het type refreshment en bij toegangsacties. Ingestroomde objecten worden gecheckt en gevalideerd volgens een aantal strikte ingestprocedures waarbij structureel een aantal (kwaliteits)controles plaatsvindt op bijvoorbeeld de compleetheid van de aangeleverde bestanden, het voldoen van het formaat aan de standaard en de aanwezigheid en de kwaliteit van de metadata.
Authenticiteit	Integriteit	Koppeling object en metadata door a. Metadata Integrity check: er wordt nagegaan of alle content files zijn gespecificeerd in de metadata middels de correcte (relatieve) locatie; b. Content Integrity check: er wordt nagegaan of de content files zijn gespecificeerd in de metadata xml en dat dit consistent gebeurt.		Bij de omroepcollecties worden checksums niet meegeleverd door de depotgever. In deze gevallen wordt een andere manier gebruikt om te controleren of een file compleet is, de zgn. header-and-footer check van de file. Optioneel onderdeel (i.c. afhankelijk van de afspraken met de depotgever) van de workflow is vervolgens de geautomatiseerde Quality Assurance waarbij de files worden gecontroleerd op kwaliteit. Ook tijdens dit proces worden technische metadata geëxtraheerd.
Authenticiteit	Integriteit	a. Beschrijving bitstream preservation functie b. Beschrijving logging van alle geautoriseerde acties c. Beschrijving check compleetheid van de aangeleverde files, het voldoen van het formaat aan de standaard en de aanwezigheid en de kwaliteit van de metadata.		

Tabel 2. Op weg naar een dienstenniveau-overeenkomst tussen NA en B&G voor Integriteit

¹² Bijvoorbeeld de publicatie Digitale Preservering van B&G, http://publications.beeldengeluid.nl/pub/387/digitale_preservering_Beeld_en_Geluid.pdf Geraadpleegd op 20-08-2018

Conclusie

Aangezien organisatieontwikkeling en de implementatie van nieuwe technieken niet lineair verlopen en de toekomst laat zich niet voorspellen is het voor collectie-beherende instellingen noodzakelijk om hun activiteiten op het gebied van duurzame toegankelijkheid helder te formuleren. Dit kan in de vorm van een samenhangend duurzaamheidsbeleid voor belanghebbenden zoals depotgevers, subsidieverstrekkers en interne medewerkers. Op deze manier zijn preserverings-activiteiten zo transparant mogelijk en wordt de digitale collectie aan de volgende generatie doorgegeven. Die transparantie kan worden vormgegeven met behulp van de in het SCAPE-project ontwikkelde catalogus. Deze catalogus is vertaald, en wordt in een handzame website omgezet. De onderverdeling in beleidsuitgangspunten en beleidselementen maakt het beleidskader hanteerbaarder voor de collectiebeheerder. De catalogus kan gebruikt worden bij het ontwikkelen van duurzaamheidsbeleid en het uitvoeren van dat beleidsplan. De elementen lenen zich goed voor de analyse van beheer in de eigen en in verschillende omgevingen. Het kader levert hiervoor handvatten, maar vraagt ook nog om operationalisering. Indien een organisatie erin slaagt om deze operationalisering vorm te geven, helpt dit bij haar opdrachtgever-schap en het gesprek met andere collectiebeheerders.

Aan de hand van twee toepassingsvoorbeelden wordt geïllustreerd hoe de catalogus gebruikt kan worden binnen een collectie beherende instelling, en hoe deze tussen instellingen gebruikt kan worden als instrument om het gesprek over *postcustodial* beheer of *preservation in place* vorm te geven. De organisatie kan, vanuit de eigen verantwoordelijkheid, via de catalogus gebruik maken van andermans deskundigheid en de basis leggen voor een gemeenschappelijke taal.

Het beleidskader van de catalogus helpt bij het geven van antwoord op de vraag om een samenhangend duurzaamheidsbeleid om ad hoc beslissingen te voorkomen, om depotgevers, subsidieverleners en medewerkers en anderen inzicht te geven in de wijze waarop de organisatie de duurzame toegang realiseert en om langetermijn-risico's te beperken.

Certificering van digitale archieven in Nederland¹

Inleiding

Nadat in 2002 de Open Archival Information Standard (OAIS) verscheen, ontstonden al snel initiatieven om het gebruik en toepassing van deze standaard te kunnen toetsen. Een van de eerste resultaten was het verschijnen van TRAC (Trusted Repositories Audit and Certification) op initiatief van RLG-NARA in 2005. Gevolgd door verscheidene andere methodieken voor self-assessment en risico-analyse, zoals Drambora en het SPOT model². Hieruit kwamen officiële certificeringsstandaarden voort als de Data Seal of Approval (DSA). Deze is in 2017 samengevoegd met een andere standaard en vanaf dat moment de Core Trust Seal (CTS) genoemd, de Duitse DIN 31644 (Nestor-Seal) de ISO standard 16363³.

Deze drie standaarden zijn officiële, formele certificeringsinstrumenten. De instellingen die hun digitaal archief gecertificeerd hebben op basis van een van deze drie, krijgen een kwaliteitskenmerk voor hun digitaal archief en de wijze waarop zij over de digitale collecties het beheer voeren. De drie certificeringsstandaarden zijn oplopend in de eisen die erin gesteld worden. Daarbij is de Core Trust Seal met 16 richtlijnen de minst veeleisende, terwijl de ISO16363 de meest uitgebreide standaard is met 109 richtlijnen.

Waarom certificering?

In Nederland zijn steeds meer instellingen bezig met digitale duurzaamheid. Er ontstond behoefte om te zien hoe volwassen de instellingen inmiddels waren en of er nog zaken verbeterd moesten worden. Zelfevaluatie en certificering zijn instrumenten om dit proces in gang te zetten. Beide passen binnen de aanpak van het Netwerk Digitaal Erfgoed, waarbij organisaties stapsgewijze aan hun professionaliteit op het gebied van digitale duurzaamheid werken.

¹ Dit artikel is een bewerking en actualisering van een eerder artikel dat verscheen in de iPRES bundel van 2017: Barbara Sierman, Kees Waterman, "How the Dutch prepared for certification". In: *iPRES 2017 Kyoto Zenodo*. (2017). <http://doi.org/10.5281/zenodo.1188246> (Geraadpleegd op 2 oktober 2018). De auteurs van het artikel gedenken Kees Jan Waterman, overleden op 30 oktober 2018, en bedanken hem voor het op de kaart zetten van het onderwerp certificering in Nederland, zijn krachtige bijdragen in de projecten van de NCDD en het Netwerk Digitaal Erfgoed en de totstandkoming van dit artikel.

² DRAMBORA, <https://www.repositoryaudit.eu>. SPOT, <http://www.dlib.org/dlib/september12/vermaaten/09vermaaten.html>

³ Core Trust Seal Certified Data Repositories (<https://www.coretrustseal.org>), Nestor Seal for Trustworthy Digital Archives (http://www.langzeitarchivierung.de/Subsites/nestor/EN/Siegel/siegel_node.html), Audit and Certification of Trustworthy Digital Repositories (<https://public.ccsds.org/Pubs/652x0m1.pdf>). Nederlandse vertalingen van de CTS en de Nestor Seal zijn te vinden via: https://wegwijzercertificering.nl/nl/praktijk/20170225_basiseisen_voor_betrouwbare_archieven_def1.pdf en https://wegwijzercertificering.nl/nl/praktijk/toelichting_nestor_keurmerk_nl2.pdf (Geraadpleegd op 2 oktober 2018).

Een manier om dit te benaderen is het denken in *capabilities* en *maturity-levels*. *Capabilities* beschrijven processen, kennis, technologie en methoden die nodig zijn om een doelstelling te bereiken of een taak uit te voeren. Het maturiteit denken, de mate waarin de organisatie “volwassen” is, voegt daar toetsbare criteria aan toe. Deze criteria helpen een organisatie vast te stellen waar zij staat ten opzichte van haar streven. Het streven hoeft niet altijd het hoogste maturiteitsniveau te zijn, want de investering om op een hoger niveau te komen, moet wel passen bij de doelstelling die de organisatie heeft.

Capabilities en *maturity-levels* kunnen bepaald worden op basis van een zelfevaluatie. Deze evaluatie geeft inzicht in de sterktes en zwaktes van de organisatie, op basis waarvan actie ondernomen kan worden. Zelfevaluatie is een eerste stap. Volgende stappen bestaan uit het invoeren van verbeteringen en het opnieuw evalueren, mogelijk op basis van formele certificeringsinstrumenten. Dit is vaak een complex proces waarvan de meerwaarde voor het management van een organisatie niet altijd duidelijk is.

Werkgroep Certificering

Om nut en noodzaak van het gebruik van instrumenten voor zelfevaluatie en certificering te promoten en erfgoedinstellingen hierbij te helpen, heeft de Nationale Coalitie Digitale Duurzaamheid (NCDD) in 2015 de Werkgroep Certificering opgericht. Doel van deze werkgroep is het verspreiden en verbreden van de kennis over certificering en de wijze van aanpak en de voordelen ervan onder de aandacht van Nederlandse erfgoedinstellingen te brengen. In deze werkgroep nemen een aantal instellingen deel die hun eerste stappen op dit terrein al gezet hebben en bereid zijn om hun kennis en ervaring te delen met anderen. Een aantal van de deelnemers beschikten bij aanvang van de werkgroep zelfs al over een formele certificering op basis van de Data Seal of Approval. Anderen stonden in 2015 op het punt een DSA aan te vragen.

Het eerste wapenfeit van de werkgroep was de publicatie van de factsheet “doe ik het goed?”⁴. Hierin wordt uitgelegd wat certificering is en wat het belang ervan is, zowel voor de erfgoedinstellingen die digitale informatie beheren, als voor de gebruikers van die informatie en de financiers van de erfgoedinstellingen. Daarnaast is er onderzocht welke digitale archieven beschikken over een formele certificering door middel van een survey⁵. Daarbij is er vooral gekeken naar wat het betekent om een certificeringstraject uit te voeren. Hoeveel tijd kost dit, wat kost het, welke kennis is er hiervoor nodig en wat levert het uiteindelijk op? Hieronder zal nader in worden gegaan op deze survey en wat dit betekent voor erfgoedinstellingen die een certificeringstraject willen opzetten.

De volgende stap was het opstellen van een roadmap voor certificering van Nederlandse digitale archieven. In de roadmap wordt een gefaseerde aanpak beschreven met als twee belangrijkste doelen: het vergroten van bewustwording en verbreden

⁴ Netwerk Digitaal Erfgoed, *Doe ik het goed? Certificering van digitale repositories in Nederland*, <https://wegwijzercertificering.nl/nl/praktijk/ncddcertificfolderdigv3def4.pdf> (Geraadpleegd op 2 oktober 2018).

⁵ Deze survey werd uitgevoerd door het Netwerk Digitaal Erfgoed uitgevoerd in 2016. De resultaten van de survey zijn te vinden op de Wegwijzer Certificering: https://wegwijzercertificering.nl/nl/praktijk/201611_de_houdbaar_report_dsasurvey_2016.pdf (Geraadpleegd op 2 oktober 2018).

van de kennis over certificering. Dit zal uiteindelijk leiden tot een vergroting van *capabilities* en *maturity* bij erfgoedinstellingen.

Een nationale aanpak

Om het bewustzijn en de kennis te vergroten werden er een aantal gerichte activiteiten opgezet. Op de eerste plaats was het van belang dat formele instrumenten voor certificering ook voor de kleinere Nederlandse erfgoedinstellingen bereikbaar werden. Daarom is ervoor gekozen om de richtlijnen van de Data Seal of Approval en de Duitse DIN-norm te vertalen in het Nederlands, respectievelijk vanuit het Engels en het Duits. Naast de letterlijke vertaling werd ook een “vertaling” naar een erfgoed-situatie toegevoegd. Beide standaarden zijn oorspronkelijk ontwikkeld vanuit het wetenschappelijke domein en richten zich in hun beschrijvingen sterk op het beheer en behoud van onderzoeksdata en de processen die daarbij komen kijken. Vertaling naar de erfgoedsector betekende dat de terminologie enigszins is aangepast en er voorbeelden uit de erfgoedsector toegevoegd zijn.

Met de Nederlandse vertalingen in de hand is er een workshopprogramma uitgevoerd, waarbij organisaties die zich voorbereiden op de CTS certificering onderling ervaringen uitwisselden en aan de hand van praktijkvoorbeelden de toepassing van de CTS uitwerkten. Onderdeel van de kennisverspreiding was ook het opzetten van de “Wegwijzer Certificering”, een website waarop de verschillende instrumenten worden uitgelegd en voorbeelden uit de praktijk worden gepubliceerd⁶.

Een derde activiteit bestaat uit het doorontwikkelen van het “Scoremodel voor digitale duurzaamheid”. Deze online tool voor zelfevaluatie is ontwikkeld door PACKED en DEN en een instrument waarmee instellingen zelf hun *capabilities* en *maturity* kunnen meten⁷. Dit instrument voor zelfevaluatie wordt door het Netwerk Digitaal Erfgoed gezien als een belangrijke schakel in bewustwording en kennisontwikkeling. Vergroten van volwassenheid ten aanzien van duurzame toegang begint immers bij een analyse van de huidige stand van zaken op dit terrein. In welke mate is mijn beleid ontwikkeld, hoe zit het met het mandaat ten aanzien van het verwerven en bewaren van digitale collecties, hoeveel kennis heb ik in huis, welke kosten zijn hiermee gemoeid en welke software en infrastructuur heb ik hiervoor nodig?

Tot slot is er een model voor gefaseerde aanpak van certificering opgezet. Dit model volgt in grote lijnen de drie opeenvolgende stappen zoals die binnen Europa zijn gedefinieerd⁸; van DSA, via Nestor-seal naar de ISO16363. Hieraan zijn door de werkgroep echter twee stappen toegevoegd.

Een gefaseerde aanpak voor certificering

Hoewel de wens tot het behalen van een certificaat vaak genoemd wordt in het Duurzaamheidsbeleid van een instelling, wordt de inspanning om zover te komen

⁶ www.wegwijzercertificering.nl (Geraadpleegd op 2 oktober 2018).

⁷ www.scoremodel.org (Geraadpleegd op 2 oktober 2018).

⁸ European Framework of Certification <http://www.trusteddigitalrepository.eu/Welcome.html> (Geraadpleegd op 2 oktober 2018).

nog wel eens onderschat. Certificering kost tijd en mankracht en uiteindelijk ook geld en dit kan ten koste gaan van andere noodzakelijke activiteiten. Om het besluit tot certificering beter te onderbouwen hebben we in Nederland twee extra fases toegevoegd, voorafgaand aan het daadwerkelijke traject om het certificaat te behalen: het doen van een initiële self-assessment en het starten van een verkenningsfase. Het resultaat van de verkenningsfase is een “go/no-go” beslissing van het management.

De initiële self-assessment

De initiële self-assessment is bedoeld om vooraf vast te stellen of een organisatie eigenlijk wel klaar is om een certificaat te willen behalen. Hoe volwassen is de organisatie op dit gebied? Er zijn verschillende tools om een self-assessment te doen, maar voor het Nederlandstalige publiek stelt de werkgroep voor gebruik te maken van het Scoremodel. In de online tool worden zeven onderdelen van digitale duurzaamheid bevestigd, zoals de aanwezigheid van missie en beleid, maar ook onderwerpen als toegang, opslag, kwaliteitscontrole en vakkennis komen aan de orde. Het eindresultaat is een overzicht waarop de sterke en zwakkere gebieden zijn weergegeven. Bij elk onderwerp zijn suggesties gegeven hoe de situatie te verbeteren. Bij het ontwerp van het Scoremodel is er aansluiting gezocht met de Core Trust Seal certificering door de vragen af te stemmen op de eisen hieruit. Het invullen van het Scoremodel kan prima een groepsactiviteit zijn met alle betrokkenen van de organisatie, waardoor beter inzicht en discussie ontstaat. Binnen twee tot vier uur heeft men dan een goed beeld of men het certificeringsproces kan starten.

De Verkenningsfase

Omdat het certificeringsproces tijd en mankracht zal vergen, is het van belang vooraf goed in beeld te krijgen hoeveel hiervan nodig is opdat het management voldoende geïnformeerd is, voordat zij een beslissing neemt. Vaak is het zo dat ook uit de self-assessment al is gebleken dat sommige zaken nog verbeterd moeten worden en dat bepaalt mede de tijd die deze fase gaat vergen. Naast de noodzakelijke verbeteringen is het nodig inzicht op de volgende punten te krijgen:

- 1) Welke organisatieonderdelen zullen bij het certificeringsproces betrokken worden? Naast de mensen die zich met digitale duurzaamheid bezig houden, zal ook HR, Financiële Zaken en juridische zaken moeten bijdragen.
- 2) Wie van deze afdelingen zijn beschikbaar tijdens het certificeringstraject en voor hoe lang moeten ze beschikbaar zijn?
- 3) Veelal is documentatie nodig om vragen uit de certificeringsmethodiek te beantwoorden. Hoeveel is daarvan al aanwezig en hoeveel moet er nog geschreven worden? Door wie en hoe lang zal dit gaan duren. Dit is een niet te onderschatten factor in het certificeringsproces.
- 4) Welk deel van de collectie gaat men laten certificeren? Het kan zijn dat een instelling daarin een keuze wil maken.
- 5) En heel belangrijk: is de certificeringsmethode goed begrepen? Vaak is het jargon per domein verschillend en zul je afspreken hoe je een bepaalde term naar je eigen domein vertaalt. Ook een goede discussie over wat de vereisten nu eigenlijk betekenen maakt onderdeel uit van het certificeringsproces, en dit kan tijdrovend zijn.

Het is niet eenvoudig om een inschatting te maken van de tijd die het certificeringsproces nodig heeft. In de DSA-survey is deze vraag wel gesteld, maar veel instellingen hielden niet bij hoeveel tijd zij aan de verschillende activiteiten besteedden. De veelal achteraf gegeven schattingen (waarbij de twee bovengenoemde fases niet apart onderscheiden zijn in de survey) van het hele proces varieerden van 100 tot 1000 uur. Daarbij moeten we er rekening mee houden dat dit erg afhankelijk is van de mate waarin een instelling bijvoorbeeld zijn documentatie al op orde heeft. In de praktijk blijkt dat aan het (her-)schrijven van documentatie de meeste tijd besteed wordt.

Uiteindelijk stelt men op basis van bovenstaande punten een management-rapportage op, zodat er een weloverwogen beslissing genomen kan worden om al dan niet het certificeringsproces op te starten.

Certificering voor de Core Trust Seal

Verscheidene Nederlandse instellingen (en velen daarbuiten) zijn inmiddels gecertificeerd volgens de Core Trust Seal of de voorganger daarvan, het Data Seal of Approval. Het Data Seal of Approval was oorspronkelijk ontwikkeld bij DANS, en kent verschillende versies. In 2017 vond er een samengaan plaats met de World Data System en ontstond een aangepaste certificeringsstandaard, die Core Trust Seal heette. Beide standaarden hebben 16 eisen, maar de CTS versie is iets aangepast en heeft een nadere toelichting bij de eisen welke documentatie men verwacht. De CTS is een assessment op afstand, er is geen auditor die langs komt bij het digitale archief. In dit proces doorloopt de instelling de volgende stappen. Na een aanvraag tot certificering bij de CTS Board, begint men met de interne voorbereidingen. Een speciale tool van CTS vraagt om de benodigde gegevens zodat de ‘peer reviewers’ een oordeel kunnen vormen. De review wordt gemonitord door de CTS Board. De instelling ontvangt terugkoppeling, moet soms een aantal zaken nader specificeren en er kan telefonisch overleg plaatsvinden. Wanneer de reviewer een positief oordeel heeft, ontvangt de instelling het certificaat. Het is een transparant proces, dat blijkt ook uit het feit dat CTS op haar website het rapport van de certificering publiceert: de antwoorden van de instelling, met verwijzingen naar documentatie (mits publiek beschikbaar en niet vertrouwelijk) en het oordeel van de reviewers. De instelling zelf kan het certificaat op haar website plaatsen (het ‘seal’), dit zegel heeft een geldigheidsduur van 3 jaar.

Nederlandse vertaling Core Trust Seal

Om de Nederlandse instellingen te ondersteunen, is er een vertaling gemaakt van de DSA richtlijnen en vervolgens ook van de Core Trust Seal richtlijnen. Bij de vertaling zijn de OASIS termen niet naar het Nederlands omgezet, omdat we ervan uitgingen dat deze terminologie bekend is binnen de beoefenaren van digitale duurzaamheid. Wel is het noodzakelijk vooraf goed de context te begrijpen van deze certificeringsstandaard: ontstaan in een omgeving waarin wetenschappelijke datasets de kern van de collectie vormt, vragen sommige concepten een vertaling naar een andere omgeving. Dit zou in de Verkenningsfase aan het licht moeten komen.

Certificering volgens DIN/nestor

Op dit moment is slechts één Nederlandse organisatie, DANS, gecertificeerd volgens de nestor / DIN 31644 certificeringsmethode. Nestor is het Duitse kennisnetwerk voor digitale duurzaamheid, waarin een 20-tal organisaties samenwerken.

Deze samenwerkende organisaties komen uit verschillende domeinen: archieven, bibliotheken, universiteiten etc. De lijst van eisen is bij nestor echter wel uitgebreider dan bij CTS, er zijn 34 requirements waar men aan moet voldoen. De certificeringsmethode werkt analoog aan die van CTS, ook hier komt niemand langs bij het archief. De instelling stuurt de vereiste documentatie naar de reviewers, die door nestor zijn aangesteld. Na een goedgekeurde review krijgt de instelling het recht om het nestor “seal” op haar website te plaatsen. Op dit moment kent het seal een onbeperkte levensduur maar iedereen weet dat je met een te lang verlopen certificaat niet meer echt betrouwbaar overkomt.

Nederlandse vertaling nestor/DIN standaard

Ook van de nestor/DIN standaard is een Nederlandse vertaling verschenen, omdat er met name in de archiefwereld belangstelling bestond voor deze standaard.

Certificering volgens ISO 16363

In 2012 verscheen de ISO standaard Audit and Certification of Trustworthy Digital Repositories. Deze standaard is een update en uitbreiding van TRAC. TRAC werd samengesteld na de publicatie van het OAIS model in 2002 en de eisen zijn dan ook gebaseerd op dit OAIS model. De beide andere certificeringsinstrumenten zijn ook geïnspireerd op het OAIS model, maar de relatie is was lossier dan bij de ISO standaard.

Wat betekent certificering voor jouw organisatie?

Voor organisaties die zelf een certificeringstraject op willen zetten is het verstandig om te leren van de ervaringen van anderen alvorens men begint. De Werkgroep Certificering heeft dan ook getracht om zoveel mogelijk ervaringsgegevens te verzamelen via een survey onder instellingen die al ervaring hebben met certificering op basis van DSA⁹. Door deze ervaringsgegevens beschikbaar te stellen, zullen Nederlandse instellingen beter voorbereid kunnen zijn op een certificeringstraject. Van belang is het om handvaten te hebben bij hoe het management van een organisatie overtuigd kan worden van het nut van certificering, hoeveel inspanning een certificeringstraject vergt en wat de gebleken voordelen zijn. De vragen in de enquête waren geconcentreerd rondom vijf gebieden, die hierna verder toegelicht worden.

Karakteristieken van de repositories

Om te beginnen is het goed om te weten welke soort repositories en digitale archieven een certificaat behaald hebben, zodat een instelling een goede vergelijking kan maken. Om welke typen organisaties het gaat, hoe groot de omvang van het personeel is dat de duurzame collectie beheert, welke typen collecties ze beheren en wat de omvang daarvan is. Voor een archief is het over het algemeen nuttiger om zich met een collega-archief te vergelijken dan met een wetenschappelijk data-archief. De DSA, en ook de CTS, is opgezet vanuit het perspectief van een data-archief waarin onderzoeksdata beheerd worden. Gebruikte terminologie is dan ook daarop geënt.

⁹ Met instemming van de Board van de Data Seal of Approval is een enquête opgesteld “DSA-Experiences: help your peers!”. Dit artikel bevat een korte samenvatting van de uitkomsten van deze survey. Het volledige rapport is te vinden via: http://www.ncdd.nl/wp-content/uploads/2016/10/201611_DE_Houdbaar_Report_DSA-survey_2016.pdf (Geraadpleegd op 9 oktober 2018). Op het moment dat de survey werd uitgevoerd was er er nog sprake van gebruik van de DSA. De resultaten zijn dan ook gebaseerd op de gegevens van DSA-gecertificeerde archieven. Deze gelden echter ook voor CTS-trajecten.

De Nederlandse instellingen die gecertificeerd willen worden, bestrijken echter een breder scala aan digitale collecties. Dat vraagt om een goede ‘vertaling’ naar de eigen situatie. De survey laat trouwens zien dat je ook met een kleine personeelsbezetting heel goed een certificeringstraject kunt uitvoeren.

Behaalde certificaten en geplande certificeringstrajecten

Interessant is het ook om te weten wanneer een organisatie voor het eerst een certificaat aangevraagd heeft en of dit certificaat ook daadwerkelijk om de 3 jaar vernieuwd wordt, zoals DSA verwacht. Immers, na drie jaar zal de omgeving van het digitale archief zeer waarschijnlijk gewijzigd zijn. Uit de survey van de werkgroep certificering blijkt dat slechts een klein deel van de instellingen die in 2010 een certificaat hadden behaald deze ook na drie jaar vernieuwd hebben. De meeste anderen waren nog niet aan een vernieuwing van hun certificaat toe. Overigens kiezen de meeste instellingen ervoor om hun DSA-certificaat op termijn te vernieuwen en niet voor de hogere stap naar de Nestor-Seal of ISO 16363 te gaan. Inmiddels beschikken 13 Nederlandse instellingen over een CTS¹⁰. De website van de Core Trust Seal biedt een overzicht van alle DSA- en CTS-gecertificeerde archieven. Het vinden van voorbeelden en het ‘gluren bij de burens’ is daarmee heel eenvoudig. Ook al omdat alle certificeringsrapporten openbaar raadpleegbaar zijn.

Wat was de motivatie om het certificaat te willen halen?

Maar wat is nu de motivatie om een formele certificering te krijgen? Deze ligt vooral bij het wekken van vertrouwen bij depotgevers. Het deponeren van digitale collecties in een gecertificeerd e-depot wekt vertrouwen bij de makers van digitale informatie, ‘er wordt goed voor onze data gezorgd’. Daarnaast wordt het als vanzelfsprekend ervaren dat de grotere instellingen hiermee zouden beginnen en een voorbeeldfunctie naar de kleinere instellingen zouden hebben. Een derde reden is de bewijslast ten opzichte van de subsidieverstrekker. Ook hiervoor geldt dat met een formeel certificaat een bepaalde mate van vertrouwen kan worden afgegeven. Daarmee vindt tevens een mate van zelfregulering plaats. De uitkomsten van de survey laten zien een certificering wordt gezien als onderdeel van de erkenning van het belang van professionalisering en kwaliteitsverbetering bij digitale duurzaamheid. De instelling laat daarmee zien dat ze *capable* en *mature* is. Een andere belangrijke reden om gebruik te maken van een certificeringsinstrument is gelegen nog vóór formele certificering zelf, namelijk als instrument voor kennisontwikkeling. Instellingen gebruiken een certificeringsinstrument, met name de CTS, ook regelmatig om zichzelf te toetsen en zichzelf de vraag te stellen “voldoe ik aan de richtlijnen”. Voor veel instellingen is certificering dus een voor de hand liggend en goed instrument voor erkenning en om aan te kunnen tonen dat digitale duurzaamheid in de organisatie op een professionele manier georganiseerd is zodat men een betrouwbare partner is voor het duurzaam behoud van de gegevens. Feitelijk bevestigden de uitkomsten van de enquête wat in de de folder *Doe ik het goed?* als voordelen van certificering werd gezien.

¹⁰ Op de website van de CTS is een overzicht te vinden van CTS gecertificeerde instellingen: <https://www.coretrustseal.org/why-certification/certified-repositories/> (Geraadpleegd op 2 oktober 2018)

Welke inspanning heeft de certificering gekost

Om zo goed mogelijk voorbereid te zijn en een inschatting te kunnen maken hoeveel tijd er voor de verschillende onderdelen in het proces van certificering nodig is (bijvoorbeeld voor de verkenningsfase maar ook het DSA certificeringsproces zelf), is het van belang om naar ervaringsgegevens te kijken met betrekking tot tijdsinvestering en benodigde expertise. Hiervoor kijken we in iets meer detail naar de resultaten van de survey. Hieruit blijkt dat 65% van de respondenten tussen de 10 en 20 uur besteed heeft aan een verkenning van een mogelijk certificeringstraject. De voorbereiding van de certificering zelf kostte aanzienlijk meer tijd, gemiddeld 100 uur. Deze tijd wordt vooral gestoken in het op orde krijgen van de eigen documentatie. Dit wordt gezien als een onderdeel van een certificeringstraject omdat de richtlijn de nodige documentatie eist als bewijs dat er ook daadwerkelijk voldaan wordt aan de richtlijnen. Uiteraard zullen erfgoedinstellingen hun documentatie ook op orde moeten hebben ook als er geen certificering wordt uitgevoerd. Dan het opstellen van het certificeringsrapport. Dit vergt over het algemeen minder tijd, tussen de 50 en 100 uur aldus de respondenten van de survey. De moraal van dit verhaal is dat een goede voorbereiding meer dan het halve werk is. Een kanttekening is hier op zijn plaats. Zoals al eerder opgemerkt, werd door de meeste instellingen geen tijdbesteding bijgehouden tijdens het certificeringsproces. De getallen zijn dan ook vaak een indicatie achteraf.

De inspanning die nodig is voor het uitvoeren van een certificering hangt samen met de duidelijkheid van de certificeringseisen. Zoals eerder is uitgelegd zijn de CTS richtlijnen ontstaan vanuit het wetenschappelijke domein. Toch blijkt dat de meeste erfgoedinstellingen niet al te veel moeite hebben met het maken van een vertaling naar hun eigen situatie. 65% van de respondenten vindt de richtlijnen adequaat en goed leesbaar. Een aantal instellingen had ruim de tijd genomen om de richtlijnen te bestuderen en om te zetten naar hun eigen situatie. Een reden te meer om een verkenningsfase in te bouwen.

Is certificering echt nuttig?

We begonnen dit artikel met de oprichting van de Werkgroep Certificering die zich tot doel stelde het bewustzijn en de kennis met betrekking tot certificering te vergroten om daarmee de mate van volwassenheid van erfgoedinstellingen naar een hoger niveau te kunnen tillen. De voordelen van certificeren zijn beschreven in een factsheet met de vraag “doe ik het goed?”.

Meer vertrouwen van belanghebbenden, betere communicatie, betere processen, transparantie, zich kunnen onderscheiden van anderen en meer aandacht voor de professie van digitale duurzaamheid in de organisatie. Uit de survey bleek dat al deze doelstellingen door de deelnemende instellingen als bereikt resultaat werden genoemd. Dit resultaat is uiteraard niet vanzelf bereikt en een certificeringsinstrument is geen wondermiddel om deze doelstellingen te bereiken. Certificering is het slotstuk van een goede organisatie, zorgvuldig geformuleerd beleid, volledige documentatie en breed draagvlak binnen de eigen organisatie. En het antwoord op de boven gestelde vraag, ‘Ja, ik doe het goed’.

SETT TLL RRS

HOOFDSTUK 2

Ontwerpen, bouwen en verbouwen

Inleiding

Na een periode van ontdekken en pionieren komen er steeds meer settelaars die na een periode van verkenning serieus aan de slag gaan. Het doen, de wijze waarop we de preservering van onze collecties op dit moment zo goed mogelijk proberen in te richten, staat in de hiernavolgende artikelen dan ook centraal. 'Doen' is misschien wel de essentie van alles. Want je mag nog zoveel nadenken als je kan over hoe erfgoedinstellingen idealiter kunnen preserveren: *'the proof of the pudding is in the eating'*. Het is immers dankzij het vallen en opstaan in de praktijk dat we kunnen reflecteren over wat werkt voor welke type van organisatie en collectie.

Kiezen is verliezen is gedachte die de redactie had bij de samenstelling van dit hoofdstuk met praktijkvoorbeelden. Er zijn zoveel preservingsstrategieën als er erfgoedinstellingen zijn om uit te kiezen. Wat de redactie bijeen heeft gebracht zijn zes interessante praktijkvoorbeelden die een beeld geven van wat er op dit moment gebeurt in het veld. De praktijkverhalen gaan over de preservering van audiovisuele archieven in zowel Nederland als in Vlaanderen, de verschillende perspectieven op preservering van architectuurarchieven in het Vlaams Architectuurinstituut en grotere Het Nieuwe Instituut in Nederland, het samenbrengen van internationale collecties in het sociale geschiedenis domein met oog op de duurzame bewaring én de conservering van digitale kunst, een niche met een kleine interdisciplinaire en zeer internationale community.

Het lijkt misschien op een spreekwoordelijke *grand écart* maar deze praktijkvoorbeelden brengen wel enkele opvallende kenmerken en *lessons learned* naar boven die als generieke aandachtspunten kunnen worden meegenomen bij het opzetten van een preservingsstrategieën. Zo is er bij verschillende praktijkverhalen, cf. architectuur- en audiovisuele archieven, veel aandacht voor de voorafgaandelijke afspraken met 'de leverancier' (archiefvormer, content provider). Kortom, een ondersteuning van de gedachte dat preservering niet start op het moment dat de archieven in beheer komen bij de erfgoedinstelling maar dat preservering al in de fase van de creatie moet beginnen.

Misschien de belangrijkste les is die we van al deze praktijkverhalen kunnen meenemen is dat preservering veel meer is dan een technisch verhaal. De duurzame toegankelijkheid heeft een flinke organisatorische component. Er moet bij de ontwikkelen van een preservingsstrategieën heel wat aandacht gaan naar het menselijk aspect, de organisatorische kant van samenwerken, de kennisontwikkeling en het behoud en het beheren van projectresultaten. Settelen is al lang niet meer het uitproberen van technische oplossingen, maar steeds vaker het uitdenken van beleid en organisatie, plannen en vooruitdenken.

Van proefondervindelijk naar gestuurd

Duurzaam preserveren bij het

Nederlands Instituut voor Beeld en Geluid¹

Inleiding

Het Nederlands Instituut voor Beeld en Geluid in Hilversum heeft de wettelijke opdracht tot het verzamelen, bewaren en ontsluiten van audiovisueel erfgoed van nationaal belang: voor de creatieve (media)industrie, voor het erfgoed, het onderwijs en de wetenschap en voor de samenleving als geheel. Uit deze opdracht volgt de verplichting om het materiaal dat aan de organisatie wordt toevertrouwd duurzaam in stand te houden, zodat alle gebruikersgroepen er blijvend toegang toe kunnen hebben.

Al meer dan 12 jaar stromen er grote hoeveelheden *digital born* en gedigitaliseerd materiaal het archief in. De belangrijkste depotgever van Beeld en Geluid, de Nederlandse publieke omroep, startte in 2006 met het *en masse* digitaal produceren van televisie- en radioprogramma's. In dezelfde periode vond de aftrap plaats van het mega-digitaliseringsproject Beelden voor de Toekomst. In dit project is in zeven jaar tijd meer dan 100.000 uur aan legacy collecties gedigitaliseerd, een substantieel deel van alle tussen 1955 en 2007 gemaakte omroepproducties. Inmiddels verzorgt Beeld en Geluid ook al vele jaren digitale archiveringsdiensten voor partijen *buiten* de publieke omroepomgeving, afkomstig uit het erfgoeddomein en de media.

Al deze materialen, omroep en niet-omroep, digital-born en gedigitaliseerd – worden dagelijks ingenomen, beheerd en gepresenteerd vanuit het Digitaal Archief. De gezamenlijke omvang van de audiovisuele files in dit Archief is intussen aangegroeid tot 26 Petabyte. Binnen de technische infrastructuur van Beeld en Geluid zijn geavanceerde interfaces ontwikkeld met de aanbieders van digitale producties binnen en buiten het omroepdomein. De vele tienduizenden gebruikers per jaar, afkomstig uit alle gelederen van de maatschappij, kunnen de audiovisuele collecties online raadplegen en gebruiken, in de vorm en de kwaliteit die ze wensen. Er zijn voor dit alles *high performance* opslag- en toegangsvoorzieningen ingericht.

¹ Voor meer informatie over duurzaam preserveren bij het Nederlands Instituut voor Beeld en Geluid zie A. de Jong *Digitale Preservering Beeld en Geluid: Beleid, Standaarden en Procedures* (2016) via <http://publications.beeldengeluid.nl/pub/387> geraadpleegd 11 oktober 2018. Alle openbare publicaties van Beeld en Geluid met betrekking tot beleid, acquisitie, preservingsplanning, kwaliteitscriteria Digitaal Archief, preservingsmetadata, opslag en technische infrastructuur zijn te raadplegen via <http://publications.beeldengeluid.nl/org/6> sectie 'Governance'.

Ontegenzeggelijk beheert Beeld en Geluid al meer dan een decennium een *full blown* digitaal archief, met alles erop en eraan. De notie 'digitaal archief' is sinds 2006 echter wel wat veranderd. De snel voortschrijdende technologie die digitale formaten en opslagmedia in rap tempo kan doen veranderen, de bizarre volumes in combinatie met een immer complexer wordende procesomgeving, hebben het concept 'digitaal archief' verdiept. Nadruk ligt nu vooral op de kwaliteit en de beheersing van de processen *binnen en rondom* dat archief, als basis voor de bestendigheid en de betrouwbaarheid van de materialen. Werd bij de start van het Beelden voor de Toekomst project het louter digitaliseren nog min of meer gepromoot als ultiem middel tot blijvend behoud van (analoog) materiaal, inmiddels is duidelijk dat er voor onvervalste duurzaamheid meer in stelling moet worden gebracht.

Het gaat nu om het *aantoonbare* behoud van de authenticiteit en de integriteit van de collecties door de tijd en door de technologie heen, ondersteund door het bewust managen van de workflows en de datastructuren. Harde garanties op blijvende bruikbaarheid van de digitale files en hun geassocieerde bestanden zijn prominent geworden. Het besef van verantwoordelijkheid daarin jegens depotgevers en gebruikers is gegroeid. Vóór alles moet een digitaal archief tegenwoordig een *duurzaam* digitaal archief zijn.

Duurzaam word je niet vanzelf. Beeld en Geluid heeft zich dit gerealiseerd en tijdig gehandeld, zowel op strategisch-organisatorisch niveau, als beleidsmatig en technologisch. Vanaf 2013 is de organisatie bewuster gaan werken aan een andere bedrijfscultuur, een cultuur van rekenschap en formele, vastgelegde procedures, op alle niveaus. De rol van de deelnemers aan het conserveringsproces, waaronder die van de organisatie zelf, werd geleidelijk beter afgebakend. Het identificeren en formaliseren van alle processen en gegevens in en rond het Digitaal Archief kreeg meer aandacht, evenals het optekenen van duurzaamheidsbeleid en uitvoeringsregels.

Inmiddels heeft Beeld en Geluid is het begrip 'digitale preservering' expliciet omarmd als primaire business. Aan de bijbehorende operationele consequenties is serieus gevolg gegeven. Het belangrijkste bewijs van deze inspanningen kwam in 2016, met de toekenning van het Data Seal of Approval (DSA) certificaat, een internationaal keurmerk voor betrouwbare archieven ofwel *Trustworthy Digital Repositories* (TDR).² Het instituut werd daarmee officieel bestempeld tot een duurzaam digitaal archief, als eerste nationale AV-archief ter wereld. Het Instituut bereidt zich inmiddels voor op het voldoen aan andere certificaten, zoals het Core Trust Seal (CTS).

Vorbereiding

Het inrichten van een gecontroleerde conserveringsomgeving is uitgebreid voorbereid. Dit is allereerst gebeurd door gedegen bestudering van de verschillende conserveringsstandaarden, waarvan OAIIS de belangrijkste was. Op basis daarvan werd een set normatieve beleidsdocumentatie geschreven. Hierin werd vastgelegd welke partijen *formeel* welke rol spelen in het conserveringsproces, van aanlevering tot en met beschikbaarstelling. Er werd duidelijkheid gecreëerd over de exacte inhoud

² Assesment Beeld en Geluid TDR certificaat Data Seal of Approval (2016)
https://assessment.datasealofapproval.org/assessment_195/seal/pdf/geraadpleegd_11_oktober_2018



Afbeelding 1. Exterieur van het Nederlands Instituut voor Beeld en Geluid. Het Digitaal Archief van het instituut werd in 2016 officieel gecertificeerd als 'trustworthy digital repository' (TDR).

van de afspraken met depotgevers die moesten worden vastgelegd. Van de Designated Communities van Beeld en Geluid (mediaprofessionals, onderwijs en algemeen publiek) werd per groep opgetekend wat ze 'technisch' nodig hebben aan uitlevering en toegang. Er werd een normatief informatiemodel ontwikkeld dat alle *preserveringsevents* benoemde die moeten worden vastgelegd. Bijbehorende datastructuren werden uitgewerkt. De technische eigenschappen van audiovisuele files werden geïnventariseerd en ook de eisen aan opslag- en conserveringsplanning kwamen vast te liggen. Beeld en Geluid beschikte met deze set documenten in een paar jaar tijd over een stevig theoretisch raamwerk voor een integraal digitaal conserveringsproces in de eigen audiovisuele omgeving. Hieraan kon de bestaande praktijk worden getoetst.

Het digitale archiveringsproces van Beeld en Geluid was aan de start in 2006 vooral geënt op de traditionele functie als bedrijfsarchief van de publieke omroepen. Storage en access voor de mediaprofessionals stonden daarbij voorop. Genoemde voorstudies toonden echter aan dat de bestaande IT-infrastructuur een aantal technische voorzieningen bood dat uitstekend zouden kunnen worden ingezet voor het 'bewust' duurzaam conserveren, zoals bepaalde kwaliteitscontroles. Binnen de informatieomgeving bestonden daarbij veel aanknopingspunten om digitale materialen technisch te kunnen karakteriseren en hun bewerkingen door de processen

heen te kunnen volgen, al gebeurde dit niet of niet voldoende. Achteraf gezien was hiermee aan twee belangrijke eisen voor de inrichting van een gecontroleerd preservingsproces feitelijk al voldaan, zij het niet met voorbedachten rade.

Uitgangssituatie

De middelen en methoden die waren geïdentificeerd om digitale duurzaamheid te bereiken moesten gerealiseerd worden in een complex geheel van lopende processen en systemen.

Hart van de technische infrastructuur vormde indertijd het 10 jaar oude, intern gebouwde Media Asset Management systeem (MAM) genaamd iMMix, in duo met het storage management systeem DivArchive. Dit geheel is direct verbonden met De Digitale Voorziening (DDV), de centrale infrastructuur van de omroepproductie-omgeving. Het MAM-systeem stuurt de instroomworkflows aan en is verantwoordelijk voor het management van de metadata, die worden geordend in het catalogusdeel. Het systeem haalt de metadata voor een groot deel op uit de verschillende omroepsystemen. Opslag en toegang tot de bronbestanden werd (en wordt nog altijd) geregeld door DivArchive, als 'go between' tussen het MAM-systeem en de daadwerkelijke storage-units, waaronder de tapelibraries. Werden de files aanvankelijk opgeslagen in het datacentrum op het Mediapark (beheerd door het Nederlands Omroep Productiebedrijf NOB, later Technicolor en nog later Ericsson geheten) met de backup in de Beeld en Geluidkelders, in latere jaren kwam de opslagfaciliteit in bezit van Beeld en Geluid zelf, eerst met uitbestede technische ondersteuning, later in eigen beheer, met de backup bij Ericsson.



Afbeelding 2. Een lading LTO tapes met kopieën van gedigitaliseerd 'legacy materiaal' staat klaar om te worden getransporteerd naar een backup locatie.

Audiovisuele materialen die *niet* afkomstig zijn van publieke omroepen maar van andere depotgevers, zoals erfgoed- en mediapartijen, werden in deze tijd in het Digitaal Archief geïmporteerd met behulp van separate, zelfgebouwde file-

en metadataimportsystemen, die voor iedere nieuwe instroom opnieuw konden worden geconfigureerd, zoals ook gebeurde. Via deze importfaciliteiten stroomden niet alleen uiteenlopende erfgoed- en mediacollecties binnen, maar ook materiaal dat Beeld en Geluid voor partijen opslaat als onderdeel van commerciële diensten. Weer een ander type instroom vormden de materialen die -voorafgaand aan opname in het Digitaal Archief- waren gedigitaliseerd of ge-encodeerd. Dit ging om analoge legacy materiaal van Beeld en Geluid zelf, of om collecties van andere partijen, op film, audio of video. Het betrof ook digitale videoformaten die moesten worden genormaliseerd naar een preserveerbaar archiefformaat.

Toegang tot het Digitaal Archief wordt van oudsher gefaciliteerd d.m.v. verschillende webportals, direct gebouwd op het MAM-systeem. Via de portals kunnen gebruikers zoeken in de catalogus met uitgebreide beschrijvende metadata. Uitlevering van de files in *low* en *high res* formaten aan de mediaprofessionals, inclusief het afhandelen van bestellingen en gebruikslicenties, geschiedt op een speciaal, besloten platform. Voor andere groepen, zoals particulieren en organisaties van eindgebruikers is een algemene portal ingericht. De toegangsinterfaces naar het feitelijke beeld- en geluidsmateriaal in het Digitaal Archief zijn geconfigureerd volgens de copyright-regels, zoals die gelden voor de verschillende gebruikersgroepen.

Procesgegevens en technische kenmerken

Bovengeschetste instroom-, opslag- en uitleveracties genereren in de systemen van Beeld en Geluid (veelal automatisch) procesmetadata en technische gegevens. Zo waren checks en controles van formaten en metadata een vast onderdeel van alle instroomworkflows. Ook tijdens filmscanning- en encodersacties werden de nodige technische en procesmetadata gecreëerd, zoals analysegegevens en de opmerkingen van medewerkers over het bronmateriaal. De uitkomst van al deze acties resulteerde in rapporten die vrij willekeurig werden opgeslagen, ergens in de systeemomgeving. Rechtenmetadata tenslotte, handmatig of automatisch toegevoegd, maakte sinds jaar en dag vast onderdeel uit van de catalogusbeschrijving van ieder ingestroomd object.

Zoals bekend leggen procesmetadata alle bewerkingen aan de objecten vast. Hierdoor kan worden aangetoond dat het object de voorgeschreven acties heeft ondergaan en door alle processen heen – niet ongeautoriseerd is gewijzigd. Technische metadata definiëren de eigenschappen van de objecten. Hierdoor kan worden zeker gesteld dat het object ook na transformaties (zoals migraties) zijn belangrijke kenmerken heeft behouden en afspelbaar blijft. Rechtenmetadata documenteren wat er met een object mag gebeuren en door wie. Hierdoor kunnen objecten worden beschermd op de manier die hun eigenaar en beheerder hebben aangegeven.

Deze eigenschappen maken duidelijk dat een van de belangrijkste sleutels tot een duurzaam digitaal archief ligt in de controle over deze drie categorieën metadata. Samen vormen procesmetadata, technische metadata en rechtenmetadata immers de fundamentele van digitale duurzaamheid, gedefinieerd als het door de tijd en door de veranderende technologie heen- kunnen garanderen van authenticiteit en integriteit, van blijvende toegankelijkheid en van geoorloofd beheer en gebruik van de materialen.

Analyse van de toenmalige situatie wees uit dat de meeste procesgegevens en technische metadata die in de normatieve documentatie als ‘preservation metadata’ waren geïdentificeerd, op de een of andere manier wel ergens in de informatieomgeving konden worden teruggevonden. Probleem was dat – met uitzondering van een deel van de rechtenmetadata – het belang van deze metadata door de organisatie niet of nauwelijks werd onderkend. Waren er voor de beschrijvende metadata en de auteursrechten uitgebreide regels en controlemechanismen, de technische metadata, de procesmetadata en de preservatierechten werden tot dan toe vrijwel niet gemanaged. Ze zaten verborgen in de vele processen en applicaties. Hun creatie was ongestructureerd en niet gestandaardiseerd. Als ze al werden vastgelegd, werd dat ingegeven door de praktische overwegingen van dat moment, die bij ieder nieuw instroom- of digitaliseringsproject weer anders konden zijn.

De Preservation Metadata Dictionary

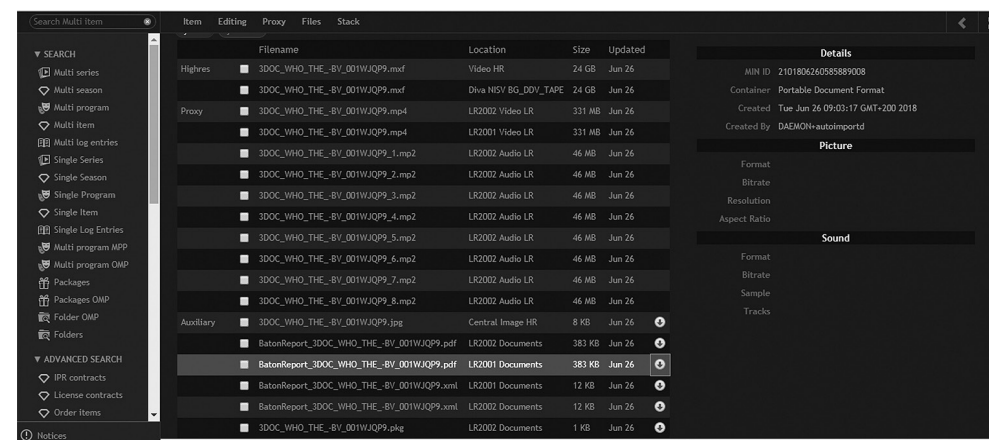
Dat het belang van deze preservation metadata en hun cruciale betekenis voor digitale duurzaamheid zou moeten worden verhelderd voor de organisatie stond al gauw als een paal boven water. Het formeel aanwijzen van bestaande gegevens als preservation metadata en hun standaardisering in de processen en systemen, zou dan ook een onmisbare stap zijn op weg naar de status als duurzaam digitaal archief.

In de documentatieset voor een gecontroleerde preservatieomgeving was om deze reden speciale aandacht uitgegaan naar identificatie, definitie en categorisering van deze metadata. Een overzicht van technische, provenance en (een deel) rechtenmetadata was samengebracht in de eerste versie van de zgn. *Preservation Metadata Dictionary* (PMD), een complete verzameling van alle metadata die mogelijk een rol kunnen spelen bij AV-preserveding. PREMIS 2.0 werd daarbij vrijwel integraal gebruikt voor de provenance metadata. Voor de technische metadata was een keus gemaakt uit attributen in de video- en audiostandaarden van diverse internationale organisaties, zoals de EBU en de Library of Congress.

Om toe te werken naar een versie van de Dictionary, die zowel is gebaseerd op internationale standaarden als op de praktijk, zijn de attributen van alle bij Beeld en Geluid gangbare fileformaten (i.c. MXF, WAV, TIFF, STL, DPX) uiteindelijk gemapt met de velden die daadwerkelijk in de bestaande systemen aanwezig zijn. Hierbij is de nieuwste versie van PREMIS (3.0) betrokken. Door de mapping met gedefinieerde en formaliseerde preservation metadata in de PMD kunnen de gegevens in de processen worden geïdentificeerd om te worden geraadpleegd en gemanipuleerd ten behoeve van preservatie: van het structureren van een nieuwe digitale instroom tot het opstellen van preservatieactieplannen. Logs in het MAM-systeem tonen aan dat iedere transfer-, ingest-, opslag-, preservatie- en access actie die is uitgevoerd op een file of een groep files overeenkomt met het preservatiebeleid. De vooraf gedefinieerde technische eigenschappen van de diverse formaten bieden aldus een controlemechanisme waarmee na transformaties (bijv. migraties naar een nieuw formaat) kan worden bepaald of eigenschappen van het object behouden zijn gebleven dan wel zijn gewijzigd.

Een nieuw Media Asset Managementsysteem

Lang voor de weg naar gecontroleerde digitale preservatie werd ingeslagen was al vastgesteld dat het vigerende MAM-systeem iMMix na vele jaren trouwe dienst aan vervanging toe was. Beeld en Geluid bevond zich tijdens de TDR-voorbereidingen dan ook midden in een organisatiebreed proces waarin de vereisten voor zowel een nieuw MAM-systeem als voor een nieuw registratiesysteem voor analoge dragers werden geformuleerd. In een zgn. RFP-traject (Request for Proposal) brachten alle afdelingen uiteenlopende wensen en verlangens naar voren. De aandacht ging daarbij traditiegetrouw uit naar de functionaliteiten voor instroom, beschrijvende metadata, rechtenmanagement, opslag en beschikbaarstelling. Al snel daagde het besef dat hier een uitgelezen kans lag om ook de normatieve preservatie-eisen mee te nemen. Het ging daarbij niet alleen om het toevoegen van requirements die expliciet voortkwamen uit de OAIS-standaard. Het ging juist ook om het identificeren van reeds benoemde eisen en wensen als *preservatiefunctionaliteit*. Door deze labeling werden bestaande en nieuw in te richten processen expliciet uitgedrukt in preservatietermen.



Afbeelding 3. Overzicht van de niveaus in het nieuwe DAAN systeem (interne interface), waarop files en metadata worden binnengehaald en vastgelegd. Via hetzelfde scherm wordt ook toegang geboden tot de bijbehorende contracten en licentieregelingen met depotgevers.

Het nieuwe MAM-systeem heet DAAN (Digitaal Audiovisueel Archief Nederland) en is sinds medio 2018 in werking. Er is uiteindelijk gekozen voor een beproefd MAM product dat ook in gebruik is bij grote media-archieven zoals CNN, Al Jazeera, Fox Sports, BBC en het Zweedse SVT. De standaardfunctionaliteiten van dit systeem zijn substantieel aangevuld met een aantal op de Beeld en Geluid toegesneden functies. Zo zijn er separate modules voor import, workflowmanagement en kwaliteitscontrole ingericht. Met de ingebruikname van DAAN is de formele identificatie van zowel belangrijke procesgegevens als de technische filekenmerken een feit. Alle files en metadata- uit welke bron ze ook afkomstig zijn – stromen nu hetzelfde systeem binnen, door gestandaardiseerde ingestworkflows. De output aan procesgegevens wordt in de nieuwe situatie gelogd en gestructureerd, en kan aldus beschikbaar komen als overzicht van de levenscyclus van de files.

Geautomatiseerde checks en monitoringmechanismen, samen met nieuwe, geavanceerde analyse- en extractietools, garanderen de kwaliteit en de volledigheid van de metadata en de files, tijdens iedere ingest en na iedere update. In de specificaties van het nieuwe MAM-systeem zijn ook de *rejection* criteria voor metadata gedefinieerd, evenals een workflow voor het repareren, loggen en documenteren van errors. Bij het migreren van de huidige systemen naar het nieuwe centrale MAM-systeem ondergaan alle Beeld en Geluid legacy bestanden bovendien dezelfde checks en controles, zodat straks over zowel oude als nieuwe bestanden dezelfde technische gegevens voorhanden zijn. DAAN biedt uitgebreide rapportagemogelijkheden van alle acties en bewerkingen. Door al deze nieuwe functionaliteiten komen méér technische gegevens beschikbaar voor preserveringplanningsacties dan ooit het geval was.

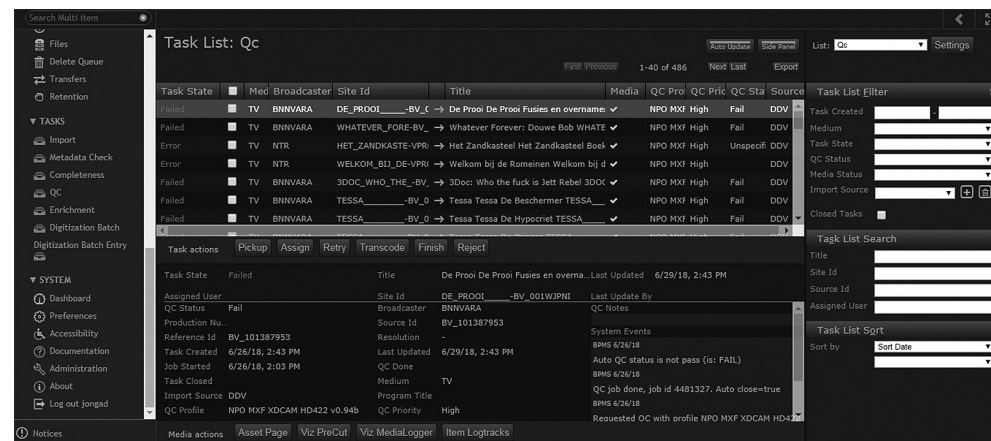
Overige functionaliteiten

Tegelijk met DAAN werd ook het Physical Asset Management systeem (PAM) gelanceerd. In PAM staat alle basisinformatie over de *fysieke dragers* in de collectie, gedigitaliseerd of (nog) niet gedigitaliseerd. Het gaat dan om film, audio en video, objecten en papier zoals omroepgidsen en andere contextinformatie. PAM registreert de gegevens over deze objecten op basis van een apart metadatamodel dat is gebaseerd op de complexe samenstelling van veel van de fysieke dragersets. Er is een directe interface voor dagelijkse updates vanuit PAM naar DAAN. DAAN toont alle beschrijvende en rechtenmetadata alsmede een (beperkte) set technische metadata van de fysieke dragers. Het complete overzicht van alle media, dragers, objecten en papieren collecties én hun herkomst is hiermee uiteindelijk te vinden in DAAN

Met de introductie van DAAN is ook de bescherming van het auteursrecht en de beschikbaarstelling van de collectie aan de Designated Communities verbeterd. DAAN beschikt over een geïntegreerde licentiemodule waarin wijzigingen en updates in het de voorwaarden en condities direct worden doorgezet naar de catalogus. De geschiedenis van de wijzigingen blijft bewaard.

Bij de lancering van DAAN zijn vernieuwde zoekportals opgeleverd: de Media Professional Portal (MPP), de General Public Portal (GPP) en de interne Beeld en Geluid interface Studio. De zoekfunctionaliteit op de catalogus is vernieuwd. Door inzet van technieken zoals sprekerlabeling en trefwoordextractie op basis van teletext (TT888) wordt de beschrijvende metadata in DAAN automatisch aangevuld. Om ze beter te kunnen afstemmen op de wensen van gebruikers analyseert DAAN regelmatig de effectiviteit van de zoekalgoritmes.

De preserveringsservices aan depotgevers uit het veld van het erfgoed, onderdeel van de knooppuntrol van Beeld en Geluid in de nationale digitale infrastructuur, is met de komst van DAAN robuuster geworden. Net als in het iMMix tijdperk is er de mogelijkheid voor depotgevers om collecties in het Digitaal Archief te deponeren, maar het nieuwe systeem kan meer mediaformaten opnemen en uitleveren dan voorheen. ‘Tenancy’ zoals de dienst wordt aangeduid, kan worden gezien als een autorisatiemechanisme dat de toegangsrechten tot de collecties regelt. Een keus uit een aantal generieke diensten moet voorzien in de behoeften van de meeste (potentiële) depotgevers buiten de publieke omroep. In de standaardoplossing wordt de toegang tot de ingebrachte digitale objecten geregeld via de catalogus van Beeld en



Afbeelding 4a. DAAN functie QC. Kwaliteitscontrole per file, automatisch onderdeel van de importfuncties van het DAAN systeem, interne interface.

Container | MXF

Check	Status	Information
Video Track Count	Passed	
Audio Track Count	Passed	
Mismatch in Audio and Video Duration	Passed	00:00:00:003
Operational Pattern	Passed	OP1a
Random Index Pack	Passed	Present
Run-In	Passed	Nct Present
KAG Size	Passed	512
Essence Wrapping Type	Passed	Essence Wrapping Type: Frzme Wrapped
Body Partition Count	Passed	354
Streamable File	Passed	Yes
KLK Fill Items	Passed	Present
Index Table	Passed	Index Table : Present
TimeCode Track (Source Package)	Passed	
TimeCode Track (Material Package)	Passed	{TimeCode Track: Present, Start TimeCode: 00:01:55:000, Start TimeCode S...
Video Track Properties	Passed	
Compare Start TimeCodes	Passed	
Partition Status	Passed	(Header Partition : Closed and Complez)(Footer Partition : Closed and C...

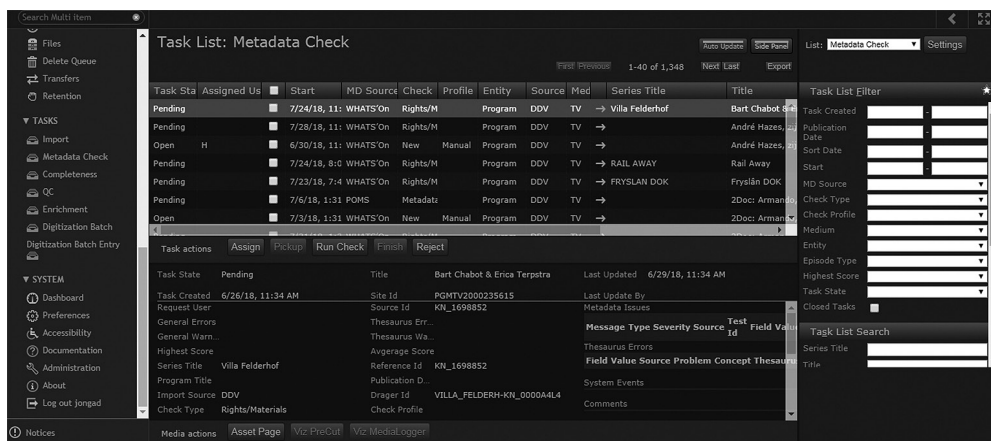
Afbeelding 4b. Daan QC rapport. De afbeelding toont het bijbehorende resultaat: een QC rapport dat per file wordt gegenereerd. Links in dit scherm zijn ook de andere standaardstappen van de ingest workflow te zien.

Geluid. Depotgevers kunnen hun materiaal benaderen via een van de nieuwe zoekportals. In een andere oplossing bouwt depotgever zelf een website op de API van het DAAN-systeem. Met behulp van zo'n API kan alleen door de depotgever zelf worden gezocht en besteld. Een combinatie is ook mogelijk.

OAIS-compliance in techniek en beleid

Het zoveel mogelijk integreren van de normatieve preservatie-eisen zoals die in de voorbereidingsfase waren ontwikkeld, heeft ertoe geleid dat in het nieuwe MAM-systeem functies zijn ingebouwd die rechtstreeks kunnen worden gerelateerd aan een aantal ISO 16363 eisen. Het gaat dan om de specifieke voorzieningen voor de processen ingest, toegang, integriteit- en authenticiteitsbewaking, workflow-management en rechtenafhandeling zoals hierboven beschreven.

Tijdens de ontwikkeling van het nieuwe MAM-systeem werd ook duidelijk welke functies *niet* 'technisch' konden worden opgelost en dus niet in het systeem zelf konden worden ingebouwd. Dit geldt voor processen waarbij strategische, financiële of curatorische afwegingen moeten worden gemaakt, die contextafhankelijk zijn, of die op een andere manier menselijk oordeelsvermogen vereisen. Beleid en procedures voor de uitvoering van deze functies worden beschreven in aparte documenten. Voorbeelden zijn informatiearchitectuurmodellen (i.c. de samenstelling van de SIP, AIP en DIPs en de workflow processen), het beleid t.a.v. *significant properties* en digitale objecttypen, keuzen binnen specifieke migratieacties, het monitoren van technische ontwikkelingen en wijzigingen in de wensen van de Designated Communities, beleid en procedures t.a.v. datachanges, backup- en disasterrecovery en methoden voor het managen van risico's voor het Digitaal Archief. Op een globaal niveau zijn deze onderdelen ook beschreven in het Preservatiebeleidplan van Beeld en Geluid.



Afbeelding 5. De check van automatisch ingestroomde metadata tijdens de ingest (interne interface). De metadata wordt aan de hand van vooraf ingestelde compliance- en rejection criteria gecontroleerd op compleetheit en kwaliteit.

Reflectie

Van een digitale infrastructuur die vooral was gebaseerd op de traditionele omroepdienstverlening heeft het Digitaal Archief van Beeld en Geluid zich in vijf jaar tijd ontwikkeld tot een instelling voor preservatiediensten die depotgevers en gebruikers uit het hele land bedient. Invulling moeten geven aan de nationale opdracht van het Instituut heeft in dit groeiproces ongetwijfeld een rol gespeeld. In feite is het Digitaal Archief en alle processen en data daaromheen nu (her)ingericht vanuit de curatorische verplichtingen van een nationale culturele instelling, die zowel een archief als een museum in zich verenigt.

Onbekend terrein

Duurzaam worden was bij Beeld en Geluid niet meteen een vanzelfsprekende zaak. Het 'waarom' maar vooral het 'hoe' van het toewerken naar een gecontroleerd digitaal preservatieproces stond management en medewerkers niet onmiddellijk voor ogen. Het onderwerp 'digitale preservatie' was in het audiovisuele archiefdomein in 2013 dan ook tamelijk onontgonnen terrein. Voorbeelden van het proces bij vergelijkbare organisaties waren niet voorhanden. Literatuur over het preserveren van dit type collecties bestond niet of nauwelijks en wat er wél was aan referentiemateriaal moest worden 'vertaald' naar een specifieke bedrijfscultuur, naar een zeer grote, dynamische digitale AV collectie én naar een complexe infrastructuur die vooral was ontwikkeld met oog op 'toegang' en minder was gericht op 'behoud'.

Communicatie

Het begrip 'Trustworthy Digital Repository' is bij deze vertaling van aanvang af bewust ingezet als communicatiemiddel. De term maakte management en medewerkers duidelijk welk doel werd nagestreefd: het behalen van een officieel betrouwbaarheidskeurmerk voor het Digitaal Archief met alle business- en imago-voordelen van dien. Het begrip benadrukte de objectiviteit van dat doel, belichaamd door de toetsing van de eigen bedrijfsprocessen aan neutrale, externe normen. Nog belangrijker was dat het ook een reikwijdte beschreef: TDR requirements maken immers duidelijk dat digitale preservatie over veel meer gaat dan alleen techniek. Het begrip gaf daarmee ook richting aan interne organisatorische- en cultuurveranderingen.

Standaarden

De noemer 'TDR' wees ook de weg naar de standaarden en modellen die Beeld en Geluid hebben geholpen om de bestaande technische- en informatieomgeving te doorgronden. De brede scope van OAIS verbond hierbij de domeinen workflows, data-objecten en technologie aan de financiële, wettelijke en organisatorische aspecten van duurzame archivering, zoals de formalisering van afspraken en contracten, de methoden voor risicomanagement en de personele kwalificaties. Met OAIS en andere standaarden in de hand kon de set normatieve documenten worden opgesteld die de gewenste processen en systemen beschreven. In een omgeving die voortdurend bewoog fungeerden deze normen als stabiel raamwerk voor alle technische veranderingen en vernieuwingen die raakten aan het te ontwikkelen digitale preservatieproces.

Semantiek

Zoals gezegd was al bij aanvang van het TDR-traject duidelijk dat veel van de elementen van dit proces al -oversneden of in rudimentaire vorm- aanwezig waren in de

bestaande infrastructuur. Maar ze bestonden feitelijk zonder dat management en medewerkers zich bewust waren van hun belang als voorwaarden voor duurzaamheid. De invulling van de preservingsambitie heeft daarom voor een deel gelegen op het semantische niveau: het benoemen. Natuurlijk moest het businessdoel (i.c. het worden van een duurzaam digitaal archief) op beleids- en operationeel niveau uiteindelijk worden vertaald naar concrete implementatie. Maar het kwam er óók op neer dat de organisatie zich – op alle niveaus – moest leren uitdrukken in preserveringstermen, zowel waar het ging om bestaande als om nieuwe archiveringsprocessen. Ook op dit punt waren standaarden als OAIIS van groot belang.

Inbedding

De organisatie heeft zich de opdracht tot digitaal behoud inmiddels expliciet eigen gemaakt: beleidsmatig, financieel én in de communicatie naar buiten. De medewerkers zijn zich bewuster geworden van de beheersmatige gevolgen van iedere inname van materiaal in het Digitaal Archief voor de processen verderop in de workflow. Er is daardoor meer aandacht voor het documenteren en het werken volgens vaste procedures. Dit betekent dat de voorwaarden en garanties aan de depotgevers en gebruikers worden vastgelegd en dat deze afspraken controleerbaar doorwerken in de opslag, de preservingsplannen en de beschikbaarstelling. De verantwoordelijkheid voor het managen van de preservingsmetadata, de workflows en de digitale objecten is geleidelijk blijvend belegd. Informatiemanagers, metadatabeheerders en de afdeling ICT, samen met de mediamanagers instroom bij wie alle processen beginnen, bouwen gezamenlijk verder aan een instrumentarium van kennis en tools. Zij zullen zich ook buigen over het verder voeren van de inrichting van het integrale preservingsproces. Zo moet het onderdeel Preservation Planning nog vastere vorm krijgen. Andere ‘TDR-onderwerpen’ die op stapel staan zijn: het beter (formeel) beheersen van de risico’s rond het Digitaal archief, zowel op technologisch, beveiligings- als financieel gebied, kostenmodellering en uitbreiding van de preserveerbare formaten.

Preservingsfunctionaliteit

Met het nieuwe MAM-systeem DAAN is een voorziening gecreëerd waarin alle gegevens die nodig zijn voor het behoud van de integriteit en de authenticiteit van de collecties, zijn geïdentificeerd en gestandaardiseerd. Omdat ze op één centrale plek gecreëerd en opgeslagen, zijn deze gegevens beter beheersbaar. Komend vanuit een heel ander doel (nl. de vervanging van het bestaande MAM-systeem onder leiding van de ICT-afdeling) heeft het design en de implementatie van het nieuwe MAM-systeem zich aldus ontpopt tot een van de belangrijkste infrastructurele bijdragen aan de verwezenlijking van de preservingsambities van Beeld en Geluid. Het traject kan tevens worden gezien als een succesvol voorbeeld van een publiek – private samenwerking op het gebied van systeemontwikkeling voor duurzame digitale archieven. Immers, aan de Beeld en Geluid adaptatie van een commercieel MAM systeem is bewust ‘preservingsfunctionaliteit’ toegevoegd.

Nieuwe mediatypen

Nu er voor de bestaande processen, data en metadata kwaliteitscriteria, normen en afspraken zijn neergelegd, vormen zich weer nieuwe aandachtspunten. Een daarvan is de verwerving van nieuwe-mediaobjecten zoals games, sociale media, websites en webvideo. Deze acquisities zijn het logisch gevolg van de uitbreiding van het collec-

tieprofiel van Beeld en Geluid in zijn hoedanigheid als Instituut voor Mediacultuur. Een gecertificeerd TDR zal uiteindelijk ook voor deze mediatypen harde garanties op duurzame toegankelijkheid moeten kunnen afgeven. Het beleid en de technieken voor digitaal behoud worden hiertoe uitgebreid. De hoofdprincipes van duurzame digitale preservering zoals die tijdens de jaren van voorbereiding en implementatie zijn geformuleerd, blijven echter ook voor deze nieuwe mediatypen onverminderd van kracht.

Social History Portal en het project Heritage of the People's Europe (HOPE)

Het preservareren van een gezamenlijke inspanning

De *Social History Portal*¹ is een platform waarop internationale collecties in het sociale geschiedenis domein bijeen worden gebracht. Het is – samen met een Persistent Identifier webservice, een Aggregator en een Shared Object Repository – het resultaat van een 3 jaar durende Europees project *Heritage of the People's Europe* (HOPE) waarin 14 collectiehoudende instellingen (leden van de International Association of Labour History Institutions (IALHI))² samenwerkten tussen 2010-2013.

Heritage of the People's Europe – het HOPE project

In het project HOPE ging het om het bijeenbrengen van en duurzaam toegang bieden tot collecties die inhoudelijk met elkaar verband houden en elkaar aanvullen, maar door de loop van de geschiedenis van Europa verspreid zijn geraakt over verschillende Europese instellingen. Op de Social History Portal zijn tijdens het project uiteindelijk meer dan 2 miljoen beschrijvingen van collectie items bijeengebracht, waarvan 880.000 voorzien zijn van een digitaal object; een afbeelding of scan van het collectie item. In Europeana zijn alleen de beschrijvingen met een digitaal object opgenomen.

Om dit voor elkaar te krijgen zijn er gedeelde voorzieningen³ gebouwd waar de deelnemende instellingen gebruik van konden maken. Voor veel van de instellingen was dit dé manier om de collecties beschikbaar te maken op een ander platform dan de eigen website en voor sommige van de kleinere instellingen was dit destijds de eerste mogelijkheid tot online beschikbaarheid van de collecties.

¹ <https://www.socialhistoryportal.org> geraadpleegd 20181010

² <http://www.ialhi.org> geraadpleegd 20181010

³ The HOPE Federated Repositories; een model van gedeelde voorzieningen die duurzame toegang tot online collecties van afzonderlijke instellingen biedt: <https://socialhistoryportal.org/bestpractices/fedrep> geraadpleegd 20181010

⁴ Vocabularies zijn in dit verband zgn trefwoorden- en of termenlijsten

⁵ <https://pro.europeana.eu/page/available-rights-statements>

⁶ HOPE Common Metadata Structure: <https://socialhistoryportal.org/bestpractices/commondata> geraadpleegd 20181010

⁷ Dit mappingsheet was een excelsheet met twee kolommen waarin in de eerste kolom alle velden stonden uit het gemeenschappelijke metadatamodel en in de tweede kolom door de instelling werd aangegeven hoe het veld in het lokale instituutseigen systeem heette. Bijvoorbeeld het titel veld *Title* uit het gemeenschappelijke model correspondeert met het veld *Titre* uit het systeem van de Franse deelnemende instelling.

De instellingen selecteerden collecties die voor het project in aanmerking kwamen en maakten de beschrijvende en technische metadata van de collecties in orde. Hierbij werden gezamenlijke vocabularies⁴ gebruikt om de collecties te voorzien van trefwoorden en werden er Europeana Rights Statements⁵ toegekend aan de digitale objecten ten behoeve van het toegestane hergebruik.

De deelnemende instellingen maken allemaal gebruik van eigen collectiebeheersystemen en metadata formaten, waarin de collectie items op de instituutseigen wijze worden gecatalogiseerd en beschreven. Het was niet de bedoeling dat alle deelnemende instellingen vanwege het project hun interne beschrijvingspraktijken zouden moeten aanpassen of andere beheerssystemen zouden moeten aanschaffen. De manier om er voor te zorgen dat al deze op verschillende manieren gestructureerde beschrijvingen gezamenlijk konden worden doorzocht als één grote collectie op het terrein van de sociale geschiedenis was door een gemeenschappelijk HOPE metadata model⁶ te ontwikkelen. Instellingen konden hun beschrijvingen aanleveren in de oorspronkelijke structuur uit de instituutseigen systemen en gaven dan in een Excelsheet, het *mappingsheet*⁷ aan waar de waarden uit het lokale metadata formaat terecht moesten komen in het gemeenschappelijke datamodel.

Uiteindelijk werden de aangeleverde beschrijvingen met behulp van de *mapping-sheets* tot één dezelfde structuur gemaakt door de in het project ontwikkelde HOPE Aggregator⁸. Deze aggregator leverde vervolgens de collecties via de HOPE API⁹ in een gezamenlijk metadata model aan de Social History Portal (in HOPE metadata model) en aan Europeana (in het Europeana Data Model (EDM¹⁰)).

Voor de aanlevering naar Europeana was het verplicht om de beschrijvende metadata en de digitale objecten te voorzien van een Persistent Identifier¹¹. Dit werd mogelijk gemaakt door een gezamenlijk Persistent Identifier webservice te realiseren die gebruik maakte van het Handle Systeem¹². Op die manier konden alle beschrijvingen en digitale objecten die door het HOPE project werden geleverd een uniek nummer krijgen.

Een aantal deelnemende instellingen beschikte niet over een manier om de digitale objecten veilig op te slaan en online ter beschikking te stellen en bleken ook niet in staat om dit zelf in hun eentje voor elkaar te krijgen. Hiertoe werd binnen het project HOPE de Shared Object Repository¹³ ontwikkeld, een gedeeld digitaal depot, waarin digitale objecten konden worden opgeslagen en van waaruit de digitale objecten konden worden geleverd.

⁸ <https://socialhistoryportal.org/bestpractices/interoper> geraadpleegd 20181010

⁹ API staat voor Application Programming Interface en is een set van afspraken waarmee de ene computer kan communiceren met de andere computer. In dit geval werden er collecties geleverd van de Aggregator naar de platformen Social History Portal en Europeana.

¹⁰ <https://pro.europeana.eu/resources/standardization-tools/edm-documentation> geraadpleegd 20181010

¹¹ Een persistent identifier (PID) is een permanente verwijzing en uniek label naar een digitaal object die onafhankelijk is van de bewaarlocatie. Het unieke label zorgt ervoor dat het digitale object altijd teruggevonden kan worden op het internet, ook als de naam van het digitale object of de bewaarplaats verandert. Daardoor is een digitaal object altijd en overal eenduidig terug te vinden op basis van zijn PID. Dit is belangrijk voor duurzame opslag (archivering) van digitale objecten in de snel veranderende wereld van internet (bron: Wikipedia)

¹² <https://socialhistoryportal.org/bestpractices/pids> geraadpleegd 20181010

¹³ http://www.peoplesheritage.eu/pdf/D5_1_Grant250549_HOPE.pdf geraadpleegd 20181010

Samenwerking en verloop

Al met al een project waarin behoorlijk technisch aan de weg werd getimmerd, en waarin inhoudelijke collectie experts van de verschillende instellingen en ontwikkelaars met elkaar moesten samenwerken om tot praktische en werkende oplossingen te komen. Het idee was dat de gedeelde voorzieningen juist de kleinere instellingen zonder eigen ICT-personeel in staat zouden stellen mee te gaan in de ontwikkelingen van portals en platformen, duurzaam collectiebeheer en digitale opslag. In de praktijk bleek toch dat in het proces om mee te beslissen over de op te zetten voorzieningen én om deze te gebruiken om de collecties beschikbaar te stellen toch een bepaald kennisniveau vereist is.

Daarbij is een samenwerking van veertien instellingen uit elf verschillende Europese landen geen eenvoudige zaak. De uitdaging was om met het verschil in taal, communicatie, cultuur en omvang ook het verschil in kennisniveau te overbruggen en daarmee de geambieerde voorzieningen en het aantal bijeen te brengen collecties te realiseren zonder dat mensen voortijdig zouden afhaken. Tijdens het project is daarom veel aandacht besteed aan workshops en bijeenkomsten die tot doel hadden om tot een gezamenlijk kennisniveau te komen. *Best practices*¹⁴ werden verzameld en handleidingen werden geschreven en opgenomen in een wiki, die ook nu nog als informatiebron beschikbaar is voor degenen die willen participeren in de Social History Portal.

Met name het gebruik van de Persistent Identifiers bleek een lastige hobbel; naast het bewust worden van het principe bleek het ook nog niet zo simpel dit in de lokale collectiesystemen te (laten) implementeren. Voor een aantal instellingen die om uiteenlopende redenen niet konden voldoen aan de technische voorwaarden moesten andere praktische oplossingen worden bedacht om hen mee te kunnen laten blijven doen aan het project. En het hele traject hield toch meer handmatig voorbereidend werk in voor de instellingen dan van tevoren bedacht.

Het merendeel van de HOPE partners nam tijdelijk personeel aan ten behoeve van het project, die fulltime op het project werkten. Deze medewerkers brachten expertise in en vergaarden in een rap tempo belangrijke en specifieke kennis over de infrastructuur van HOPE. Dit zorgde voor een versnelling in het project, maar deze mensen verdwenen – met hun specifieke kennis – ook weer van het toneel zodra hun taak of het project ten einde was.

Het is belangrijk tijdens een project aandacht te besteden aan het beheer van de projectresultaten na afloop van het project. Dit wordt ook gevraagd door de Europese Unie in de verantwoording en eindrapportage van het project, maar is toch vaak een ondergeschoven kindje tijdens het project. De ervaring leert dat tijdens een project vrijwel alle aandacht uitgaat naar de *deliverables* die dan geleverd moeten worden. Ook in het HOPE project was het halen van de projectdoelstellingen eerste prioriteit. Het maken van een businessmodel voor het beheer van de projectresultaten was in het project een specifieke taak (WorkPackage 7) hetgeen ons gedwongen heeft na te denken over wie op welke wijze zou zorgen voor hetgeen er was opgeleverd. Dit heeft er met name voor gezorgd dat de kennisoverdracht en het technisch beheer van de producten zoals de Shared Object Repository, de Persistent Identifier webservice en

¹⁴ <https://socialhistoryportal.org/bestpractices> geraadpleegd 20181010

¹⁵ <https://pro.europeana.eu/project/europeana-dsi> geraadpleegd 20181010

¹⁶ Content providers is de term die Europeana gebruikt voor de instellingen die collectiebeschrijvingen (de content) aanleveren aan de portal

de Metadata Aggregator goed waren belegd; het inhoudelijke beheer en onderhoud van de projectresultaten zoals de Social History Portal bleek in de praktijk moeilijker te voeren.

Met het einde van het project vielen de financiële middelen weg, waarmee instellingen in staat waren om mensen beschikbaar te stellen voor het uitvoeren van werkzaamheden. De tijdelijk aangestelde projectmedewerkers waren – met de kennis – vertrokken en de vaste medewerkers bij de instellingen waren teruggekeerd naar hun dagelijkse werkzaamheden of aan nieuwe projecten begonnen. Het waren de grotere instellingen binnen IALHI, die als voormalig partner in het HOPE project de verantwoording namen voor het voortzetten van de Social History Portal, het organiseren van workshops en het enthousiasmeren van IALHI instellingen om collecties beschikbaar te stellen via de portal. Deze grotere instellingen kunnen zich het beter dan de kleinere instellingen permitteren om ook zonder een financiële tegemoetkoming een aantal taken uit te voeren.

Grote inspanningen ten behoeve van de ontwikkeling van de portal, het opnemen van nieuwe collecties en de gedeelde voorzieningen bleven echter uit wegens gebrek aan financiële middelen.

Een nieuw project

Twee jaar na afloop van het HOPE project deed zich de kans voor om de bestaande niet gebruikersvriendelijke aggregator infrastructuur te vernieuwen binnen het Europeana DSI project¹⁵ (april 2015 - april 2016). IALHI was een van de 26 partners in dit project, dat tot doel had om Europeana te laten ontwikkelen tot een Digital Service Infrastructure (DSI). Het project bood ruimte aan bestaande *aggregator partners* van Europeana om infrastructurele vernieuwingen uit te voeren om op die manier betere diensten aan te kunnen bieden aan de *content providers*¹⁶ zodat meer instellingen ook daadwerkelijk in staat zouden kunnen zijn om hun collecties aan Europeana te leveren. Naast het werk aan de infrastructuur verplichtte IALHI zich om ook daadwerkelijk 5 nieuwe *content providers* aan te brengen die hun collecties middels de nieuwe infrastructuur aan Europeana zouden leveren.

Dit gaf de nodige financiële ondersteuning voor 4 IALHI instellingen¹⁷ om het nodige werk te verzetten aan de infrastructuur en om uiteindelijk de Social History Portal verder uit te breiden. Een uitbreiding van de Social History Portal maakte geen deel uit van het Europeana DSI project, maar lag met een verbeterde aggregator tool en nieuwe *content providers* wel in het verschiep.

In dit Europeana DSI 1 project was de Stichting IALHI als netwerk tot partner gemaakt hetgeen de uitvoering van het project nodeloos ingewikkeld maakte, gezien alle regels en verantwoording die met een Europees project gepaard gaan. IALHI is immers geen organisatie met een kantoor en medewerkers; het is een netwerkorganisatie, die bestaat uit afzonderlijke instellingen. Het feit dat IALHI geen mensen in dienst kon nemen, en dat de instellingen die het werk namens IALHI uitvoerden over 4 verschillende landen waren verdeeld gaf een enorme administratieve rompslomp. Dit had vermeden kunnen worden als een van de 4 uitvoerende IALHI instituten tot partner was gemaakt, een les voor een volgende keer!

¹⁷ Het werk werd uitgevoerd door Amsab-Instituut voor Sociale Geschiedenis (AMSAB-ISG) www.amsab.be, het Internationaal Instituut voor Sociale Geschiedenis (IISG) iisg.amsterdam, het Archiv der sozialen Demokratie der Friedrich-Ebert-Stiftung (AdsD) https://www.fes.de/archiv/adsd_neu/index.htm en Open Society Archives (OSA) at Central European University <http://osaarchivum.org/index.php>

Desalniettemin zijn we er in geslaagd om in Europeana DSI 1 de nieuwe aggregator tool MINT (Metadata INTERoperability services)¹⁸ te selecteren en te implementeren én vijf nieuwe *content providers* konden bovendien al gebruik maken van deze gebruikersvriendelijke tool bij het aanleveren van de collecties en het *mappen* van de metadata naar het Europeana Data Model (EDM). Tijdens dit project is toen besloten het HOPE Metadatamodel te verlaten en te werken met EDM.

Met de komst van MINT werd het voor potentiële nieuwe content providers veel gemakkelijker om nieuwe collecties aan te leveren en de beschrijvingen te mappen naar het gemeenschappelijke metadatamodel EDM. Voor het gebruik van de in het project HOPE gerealiseerde mapping tool was toch vaak nog een ontwikkelaar nodig hetgeen voor instellingen van kleine omvang zonder ICT kennis in huis toch altijd een te hoge drempel was gebleken om de collecties op de Social History Portal en/of Europeana te plaatsen. Dit is nu verleden tijd.

Met de ervaringen van het post-HOPE tijdperk in ons achterhoofd realiseerden we ons dat er ook na afloop van het Europeana DSI 1 project nog werkzaamheden nodig zouden zijn. In de eerste plaats was het belangrijk de collecties van de vijf nieuwe content providers ook zichtbaar te maken op de Social History Portal. Daarnaast was het technisch noodzakelijk om de oude HOPE aggregator infrastructuur uit te faseren en alle collecties, ook die tijdens het HOPE project waren aangeleverd, met behulp van de nieuwe MINT tools op de Social History Portal aan te sluiten en waar mogelijk te vernieuwen. Ook de Social History Portal zelf zou onder handen moeten worden genomen om te kunnen werken met de API van MINT en een uitbreiding van de zoekfunctionaliteiten was ook hard nodig. Tijdens het Europeana DSI 1 project was hier geen ruimte noch financiering voor, maar is er wel al een projectplan gemaakt voor een vervolgproject waarin genoemde werkzaamheden zouden worden uitgevoerd. De stichting IALHI heeft samen met drie van haar leden financiering gezocht en is met de uitvoering van de werkzaamheden gestart. In september 2018 is de nieuwe Social History Portal gepresenteerd in Milaan tijdens de 49e Annual Conference van IALHI.

Terugblik

Het HOPE project heeft voor vrijwel alle deelnemende instellingen voor ontwikkeling van kennis, collectiebeheer en veel gevallen voor ontwikkeling in infrastructuur gezorgd. Het gezamenlijk ontwikkelen van een datamodel en het leveren van de metadata beschrijvingen aan Europeana in het Europeana Datamodel (EDM) betekende dat de instellingen naar hun eigen systemen en manieren van beschrijvingen moesten kijken. Gegevens werden opgeschoond, genormaliseerd in een verbeteringslag die nodig was om de data gezamenlijk toegankelijk te maken. De Shared Object Repository was voor de deelnemers die daar gebruik van maakten veelal de eerste kennismaking met een digitaal depot en de werking ervan. Het Internationaal Instituut voor Sociale Geschiedenis (IISG) heeft de Shared Object Repository tot de eerste helft van 2018 in gebruik gehad als digitaal depot¹⁹.

¹⁸ MINT services compose a web based platform that was designed and developed to facilitate aggregation initiatives for cultural heritage content and metadata in Europe. It is employed from the first steps of such workflows, corresponding to the ingestion, mapping and aggregation of metadata records, and proceeds to implement a variety of remediation approaches for the resulting repository http://mint.image.ece.ntua.gr/redmine/projects/mint/wiki/Wiki_geraadpleegd_20181010

De introductie van de Persistent Identifier, de bewustwording van het belang ervan en de implementatie ervan in de afzonderlijke instellingen heeft er voor gezorgd dat veel instellingen op dat gebied in de voorhoedes van de nationale ontwikkelingen zaten²⁰.

Voor sommige HOPE partners gold dat ze in de afgelopen jaren verplicht werden gesteld om de eigen nationale ontwikkelingen en voorzieningen te volgen, en zodoende een andere Persistent Identifier Service en een ander digitaal depot gingen gebruiken. De collecties van deze instellingen zijn nog wel toegankelijk via de Social History Portal.

Ook heeft het project een grote meerwaarde gehad voor IALHI en het netwerk van de instellingen. Deze intensieve manier van samenwerking en kennisdelen was zonder het project HOPE niet tot stand gekomen. De onderlinge relaties en het begrip voor elkaars verschillen zijn door het project verbeterd en de jaarlijkse IALHI conferenties spelen nu een belangrijke rol in het onderhouden van contacten.

Deze positieve gevolgen nemen niet weg dat het tegen het einde van het project en daarna heel duidelijk is geworden dat projectresultaten als een collectieplatform een continue aandacht behoeven. Om alle inspanningen van vier jaar samenwerking niet binnen no time te niet te doen is het nodig dat instellingen die met elkaar het project aangaan zich ook bewust zijn van het feit dat het onderhoud en beheer ook na het project om inzet en middelen vraagt.

Naast aandacht voor het technisch beheer is het belangrijk ook oog te hebben voor de continuïteit van het inhoudelijk beheer en daar een onderhoudsplan voor op te stellen. Een dergelijk plan zou helder moeten maken welke inspanningen nodig zijn en wie daar wel of niet aan kan bijdragen. Het duidelijk maken van de verplichtingen van de instellingen omtrent de zorg voor de collectiebeschrijvingen die zij hebben aangeleverd en een goede manier om ook de kleinere instellingen blijvend te betrekken moet onderdeel uitmaken van het onderhoudsplan.

Een van de grootste risico's voor het inhoudelijk beheer van de projectresultaten is het wegvallen van de mensen met kennis. Daarom zou het goed zijn om de vaste staf te laten werken op het project en zodoende in kennisontwikkeling binnen de instelling te investeren. Met het projectgeld kunnen dan tijdelijk mensen worden aangenomen die de dagelijkse taken uitvoeren. Het tijdelijk vervangen van vaste staf lukt vaak niet voor alle taken, maar het is de moeite waard om te onderzoeken welke taken wel kunnen worden overgenomen zodat het instituut op de lange termijn kan profiteren van de kennis die in het project wordt opgedaan.

Het is niet vanzelfsprekend dat het met de opgeleverde producten 'wel goed komt' en het voor elkaar krijgen dat het inderdaad goed blijft gaan is een moeizame weg geweest voor IALHI. Gelukkig is men zich bewust van het belang om de gedeelde voorzieningen blijvend te beheren en de portal toekomstbestendig in te richten. Dat er in 2018 een vernieuwde portal gelanceerd is, is een nieuwe mijlpaal!

¹⁹ Het IISG gebruikt nu Archivematica als systeem voor haar digitale depot, voor zowel born digital als gedigitaliseerd materiaal

²⁰ In het kader van het Netwerk Digitaal Erfgoed zijn er het afgelopen jaar inspanningen verricht om te komen tot een PID wijzer en landelijke gedeelde voorziening voor het uitdelen en beheren van Persistent Identifiers. Het IISG heeft de in het HOPE project opgedane kennis en de inmiddels jarenlange praktijk in het gebruik ervan kunnen delen met het Netwerk Digitaal Erfgoed

10 jaar digitale architectuurarchieven: een stand van zaken

De opdracht van Het Nieuwe Instituut en het Vlaams Architectuurinstituut

Met het Rijksarchief Architectuur en Stedenbouw – eigendom van de Staat der Nederlanden – documenteert het instituut de Nederlandse architectuur en stedenbouw. Het Nieuwe Instituut (HNI) is vanuit deze opdracht ook verantwoordelijk voor verwerving, beheer, preservering en toegang van digitaal geboren archieven.

Het Vlaams Architectuurinstituut (VAi) opereert als sectorinstituut voor hedendaagse architectuur in Vlaanderen en neemt al sinds 2003 een dienstverlenende rol op voor Vlaams architectuurerfgoed via de afdeling Centrum Vlaamse Architectuurarchieven (CVAa). In januari 2018 werd het CVAa uitgebreid tot het Architectuurarchief Vlaanderen.¹ Daarmee krijgt het Vlaams Architectuurinstituut de opdracht om een collectie architectuurarchieven uit te bouwen die de Vlaamse architectuur en stedenbouw documenteert.

Aangezien beide instituten de opdracht hebben om de architectuur in hun land of regio te documenteren, moeten ze natuurlijk oog hebben voor de digitalisering van de ontwerp praktijk binnen de architectenbureaus.

In dit artikel zullen we ingaan op de uitdagingen voor Het Nieuwe Instituut en het Vlaams Architectuurinstituut met betrekking tot preservering van digitale archieven. We kaderen het gebruik van software binnen architectenbureaus en wat dit precies betekent voor onze werking. Vervolgens schetsen we hoe we vanuit onze culturele opdracht werken aan een beantwoording van de problematiek en dit binnen vier thema's: aanpak bij de bron, duurzaamheid en bewaarstrategieën, kennisopbouw en digitaal depot.

Software, de architectuurpraktijk en de uitdagingen

Het eerste gebruik van computers en software voor ontwerp stamt uit de jaren zestig van de twintigste eeuw², toen architecten begonnen te experimenteren met tekenprogramma's die voornamelijk werden ontwikkeld onder impuls van de auto- en

¹ Op 1 januari 2018 integreerde het Architectuurarchief van de Provincie Antwerpen in het Vlaams Architectuurinstituut. Het CVAa en het Architectuurarchief van de Provincie Antwerpen werden daarbij samengevoegd tot het Architectuurarchief Vlaanderen.

² Zie onder meer Mirko Zardini, *Archaeologists of the Digital*, (2013): <https://www.cca.qc.ca/en/issues/4/origins-of-the-digital/2077/archaeologists-of-the-digital> geraadpleegd 11 oktober 2018

luchtvaartindustrie. Dergelijke ontwerpsoftware wordt *Computer-aided design* of CAD genoemd. De echte doorbraak van CAD kwam er in de jaren negentig met de komst van de PC. Vandaag de dag zijn grote delen van het ontwerp- en bouwproces van architecten gedigitaliseerd. In globa kunnen we vier functies onderscheiden waarvoor software wordt ingezet, naast het klassieke gebruik van mail en andere kantoorsoftware.

De eerste functie is die van de uitwerking van het ontwerp. Een ontwerp voor een gebouw moet groeien, worden uitgetest, hertekend... Eenmaal de hoofdlijnen van een ontwerp zijn vastgelegd, werken architecten dit uit tot tekeningen waarmee ingenieurs en aannemers aan de slag kunnen. Deze processen worden bij architectenbureaus tegenwoordig zo goed als altijd met CAD-software uitgevoerd.

Een tweede functie die we identificeren is de creatie en realisatie van complexe ontwerpen. Soms kunnen ontwerpen zo complex zijn dat hiervoor de rekenkracht van de computer is vereist. Het ontwerpen via algoritmes (parametrisch design) heeft geleid tot nieuwe vormelijke expressies binnen de architectuur. Voor dit gebruik zijn er softwaretools op de markt, maar veel architecten ontwikkelen ondertussen hun eigen software.

De creatie van visualisaties is een derde functie. Klanten of buurtbewoners willen graag weten hoe een gebouw er in werkelijkheid uit zal zien. Architecten gebruiken daarvoor klassieke beeldbewerkingssoftware (bv. Photoshop), maar ook 3D-modelling software. Deze software laat onder meer toe om 2D renderings met fotografisch echtheid, interactieve 3D-modellen en walkthroughvideo's te creëren.

Tot slot zien we dat steeds meer software in gebruik genomen wordt die de informatiestromen tussen projectpartners moeten verbeteren. Platformen als Dropbox en Google Drive vormen een belangrijke aanvulling op traditionele kanalen voor informatie-uitwisseling. Daarnaast worden er platformen specifiek voor architecten ontwikkeld en gebruikt. Ze bieden vaak specifieke metadatasets en flows aan voor bouw informatie

Een bijzondere ontwikkeling van de laatste jaren is de opkomst van Building Information Modeling (BIM). Binnen BIM wordt alle informatie met betrekking tot het bouwproject geïntegreerd binnen één multidimensionaal model van het gebouw om de communicatie tussen architecten, opdrachtgevers, ingenieurs, aannemers, leveranciers, facility managers enz. te bevorderen.

De digitalisering van het ontwerpproces is dus voor heel wat aspecten van fundamentele invloed op de architectuurprofessie. Willen HNI en VAi hun opdracht blijven vervullen en architectuur en stedenbouw in hun werkingsgebied blijven documenteren, dan is er een strategie nodig om met de digitalisering van de informatie om te gaan.

Natuurlijk moet iedere archiefinstelling een antwoord zien te vinden om te kunnen omgaan met digitale archieven, maar met oog op architectuur informatie willen we graag de aandacht vestigen op enkele specifieke uitdagingen voor onze context.

Allereerst groeit de vaktechnische complexiteit van de informatievorming binnen architectenbureaus, zodat die informatie voor buitenstaanders ook moeilijker te interpreteren is. BIM-bestanden zitten barstenvols geïntegreerde informatie, opgesteld vanuit technische protocollen. Parametrisch design vereist inzicht in de software en in de algoritmes...

Daarnaast zijn specificaties van software en bestandsformaten meestal eigendom van de bedrijven en niet openbaar gemaakt. Om CAD-bestanden weer te geven, moet de archiefinstelling deze software dus in huis halen, waarbij het vaak gaat om dure, gespecialiseerde pakketten. De complexe technische opbouw van bestanden bevordert de migratie naar meer open formaten niet.

Een laatste uitdaging is ten slotte dat archivering geen core business is van architecten. Behalve wat betreft de aansprakelijkheid van architecten tot tien jaar na de oplevering van een project, is er geen wet die architecten aanspoort om goed te archiveren. Uit surveys blijkt dat de bureaus zich niet bewust zijn van het belang van hun archief, zowel voor het bedrijf als voor de maatschappij. Daarom gebeurt archivering meestal ad hoc, vanuit een praktische noodzaak. De risico's op informatieverlies nemen daardoor toe. Noch het Vlaams Architectuurinstituut nog Het Nieuwe Instituut kan daarbij een beroep doen op een wettelijk kader.

Kennis opbouwen

Enkele initiatieven³ sinds het begin van dit millennium hebben een belangrijke bijdrage geleverd aan de kennisontwikkeling en -verspreiding voor de preservatie van digitale ontwerparchieven:

Naam project	Initiatiefnemer	Doelstelling	Periode	Resultaten
GAUDI				
<i>Website offline, info in: http://www.icam-web.org/data/media/cms_binary/original/1163002339.pdf.</i>				
	Zes internationale archiefinstellingen	Internationale survey naar digitale architectuurarchieven; kennisdeling	2002-2004 / 2006-2008	Survey
FACADE				
<i>Mit.facade.edu (enkel bereikbaar via archive.org)</i>				
	MIT (VS)	Kennis duurzame bewaring 3D CAD	2006-2009	Model workflow; Rapport
The archaeology of the digital				
<i>https://www.cca.qc.ca/en/issues/4/origins-of-the-digital</i>				
	CCA (Canada)	Digitale projectdossiers verwerven; Kennisverhoging digitaal archiveren	2012-2016	Archiefacquisities; Tentoonstellingen; Publicatie

Na afloop van het GAUDI-project in 2007 bleek dat er binnen beide instituten niet de nodige kennis aanwezig om de komst van digitale architectuurarchieven te beantwoorden. Beide organisaties zijn deze uitdaging aangegaan om het kennispeil hierover te verhogen.

Voor het Vlaams Architectuurinstituut, dat opereerde als expertisecentrum, was kennisopbouw steeds de core business. Het onderzoek naar preservatie van digitale architectuurarchieven gebeurde dan ook steeds los van het beheer van een eigen collectie. De werking rond digitale architectuurarchieven ging van start in 2008 en nam de vorm aan van verschillende deelprojecten, vaak in samenwerking met externe partners. Het ging hierbij meestal om case studies, waarbinnen theorieën en methodes voor digitale preservatie getoetst werden aan de praktijk. Op de website van het Vlaams Architectuurinstituut wordt een overzicht up to date gehouden van de projecten.⁴

In 2008 werd bij Het Nieuwe Instituut een bescheiden aanvang gemaakt met het opbouwen van kennis. Zo organiseerde het toenmalige NAI⁵ in 2009 het congres Hybrid Architectural Archives.⁶

In 2015 zette Het Nieuwe Instituut onder leiding van Behrang Mousavi, general manager Erfgoed, een duidelijke ambitie neer. Startpunt was een volwassenheidsstudie naar de situatie voor digital preservation, die de opgave helder maakte en het startpunt was voor een traject voor het realiseren van de voorzieningen en de organisatie voor duurzaam beheer van digitale architectuurarchieven. Gestart werd met het professionaliseren van de organisatie. Er werd een digitale leeromgeving opgezet en een interne training gegeven, waarbij het archief van architectenbureau MVRDV onderzoeksobject was. Daarmee werd Het Nieuwe Instituut ook een belangrijk participant binnen het Netwerk Digitaal Erfgoed (NDE). De leeromgeving en training vormden binnen NDE de case study die leidde tot de ontwikkeling van de blended leeromgeving Leren Preserveren, een samenwerkingsproduct van NDE en Het Nieuwe Instituut. Leren Preserveren werd in 2017 gelanceerd en voorziet sindsdien in een belangrijke behoefte voor erfgoedinstellingen.⁷

Met het Project inrichting digitaal archief maakt Het Nieuwe Instituut een tweede belangrijke stap. Binnen het project wordt de professionalisering en het ontwikkelen van de organisatie verder voortgezet. Met het project gaat Het Nieuwe Instituut de voorzieningen en organisatie voor de digitale collectie op hetzelfde niveau brengen als voor de analoge collectie. Binnen het kader van het project organiseerde Het Nieuwe Instituut de internationale expertmeeting *Archiving Digital Architectural Heritage* (19-21 april 2018). Deelnemers van architectuur- en erfgoedinstellingen, uit de architectuurpraktijk en uit het onderwijs gingen met elkaar in gesprek over hoe zij de uitdaging zien en aangaan. Motieven en mogelijkheden voor het koppelen van duurzame toegankelijkheid aan onderzoek naar born-digital archieven en de gevolgen van het digitaal werken voor ontwerppraktijk, de architectuur en

³ Andere projecten, publicaties of initiatieven die hier kunnen worden vernoemd zijn: DURAARK, Europees project over de archivering van BIM en point-clouds, 2013-2016, (<http://duraark.eu>); Technology Watch Report van de Digital Preservation Coalition "Preserving Computer-Aided Design (CAD)", door Alexander Ball, 2013, (<http://dx.doi.org/10.7207/twr13-02>); Design Records Section van The Society of American Archivists, SAA, (<https://www2.archivists.org/groups/design-records-section>) allen geraadpleegd 11 oktober 2018.

⁴ Een overzicht van alle VAI-projecten m.b.t. digitale architectuurarchieven wordt bijgehouden op: <https://www.architectuurarchiefvlaanderen.be/nl/project/project-digitale-architectuurarchieven> geraadpleegd 11 oktober 2018

⁵ NAI, het Nederlands Architectuurinstituut is in 2013 samen met de Premisela Stichting en Virtueel Platform opgegaan in Het Nieuwe Instituut

⁶ Zie <http://conference.nai.nl/conference> geraadpleegd 11 oktober 2018

⁷ Zie <https://lerenpreserveren.nl/>

innovaties daarbinnen, passeerden de revue. Tim Walsh en Stefana Breitwieser gingen uitgebreid in op de benadering waarvoor het CCA heeft gekozen binnen het project *The Archeology of The Digital* en welke resultaten dat voor digital preservation heeft opgeleverd.⁸

Bewaarstrategieën bepalen en uittesten

Duurzaam toegankelijk houden betekent in eerste instantie *bit preservation*.⁹ Functionele preservatie is de volgende uitdaging. Het ontwikkelen van een strategie voor functionele preservatie¹⁰ vergt onderzoek, waarbij het belangrijk is te weten wat de mogelijkheden zijn van migratie- en emulatiestrategieën en hoe ze elkaar aan kunnen vullen.

In een deelproject van het Vlaams Architectuurinstituut voerde extern onderzoeker Henk Vanstappen in de loop van 2012 een testcase uit om 2D en 3D CAD-bestanden te migreren naar meer open formaten. Het onderzoek leverde interessante informatie op over de belangrijkste CAD-softwarepakketten, -bibliotheken en -bestandsformaten en potentiële archiveringsformaten. Voor enkele programma's en formaten stelde Vanstappen een volledige historiek van versies op met een overzicht van de backward compatibility. Naast deze research werden enkele bestandsmigraties als het ware spelenderwijs uitgetest. Hieruit bleek dat een migratiestrategie problematisch wordt naarmate de complexiteit van de bestanden toeneemt. Uitwisseling tussen CAD-softwarepakketten bleek op dat moment ook niet bijzonder goed ondersteund door de softwareleveranciers.¹¹

Het ontbreken van een ideale oplossing impliceert dat er keuzes moeten worden gemaakt. Deze keuzes worden bepaald door de noden van de 'designated community', de gebruikersgroepen voor wie we archiveren.

Om deze noden te bepalen doorliep het Vlaams Architectuurinstituut in de periode van 2014-2016 een vervolgtraject met een gebruikersonderzoek waarbij we zes hoofd-categorieën van gebruikers identificeerden. Door middel van interviews probeerden we hun noden zo goed mogelijk vast te stellen. Hieruit bleek dat voor onderzoeksdoeleinden de perfecte bewaring van alle eigenschappen van één archiefdocument in de meerderheid van de gevallen van niet zo'n groot belang was, omdat onderzoekers informatie meestal halen uit een aggregatie van documenten. Conclusie was dan ook dat preservatiesacties zich in de eerste plaats dienen te focussen op het niveau van de aggregatie, en niet op die van het individuele document. Op basis van deze bevindingen werkten we een generieke archiveringsworkflow uit, die zowel fysieke overdracht, ordening, beschrijving als archiefpreservatie in scope had. De testcase hiervoor waren de digitale projectdossiers van ontwerper/architect Maarten Van Severen.¹²

Duurzame toegankelijkheid en preservatiestrategieën hebben veel aandacht gekregen binnen het professionaliseringstraject van Het Nieuwe Instituut. Binnen het project *inrichting digitaal archief (2017-2019)* wordt de *preservation policy* uitgewerkt

Aanpak bij de bron

Dat archivarissen niet meer langer passief op de komst van een archief wachten, maar actief ingrijpen op de werkvloer, is al lang een realiteit die wordt toegepast binnen overheden of bedrijven. De snelle evoluties door de digitalisering van de informatiestromen vereisen dit immers. Bovendien zijn de archiefvormers de meest aangewezen personen om documenten in het archief te contextualiseren en te beschrijven, iets wat met de vaak complexe architectuurarchieven zeer welkom is.

Er zijn dus veel argumenten om nauwer samen te werken met architecten. Het Nieuwe Instituut en het Vlaams Architectuurinstituut ontwikkelden hiervoor verschillende initiatieven, waaronder enquêtes of surveys, waarbij we belangrijke informatie vergaarden over de digitale archieven binnen architectenbureaus. Het toenmalige NAI deed in 2008 een enquête bij 4 bureaus en het Vlaams Architectuurinstituut deed dit voor Vlaanderen in 2011 bij 9 bureaus.¹³

In het kader van het *Project inrichting digitaal archief* heeft Het Nieuwe Instituut in 2018 opnieuw onderzoek uitgevoerd bij een zestal architectenbureaus. Uit het onderzoek kwamen de verschillende functies waarvoor software wordt ingezet, en die eerder in dit artikel beschreven werden, duidelijk naar voren. Software wordt op een steeds slimmere manier ingezet, het belang van co-design neemt toe en het gebruik van software wordt vaak bepaald door eisen van projectpartners en door de voorkeuren en knowledge base van zij die binnen het proces samenwerken. De aandacht voor archiefbeheer is laag. Archieven gebruikt voor PR-doeleinden worden vaak goed verzorgd en zijn daarom bij verwerving interessant. De aandacht voor een veilige en vertrouwelijk opslag is doorgaans hoog. Het onderzoek concludeert dat instellingen de uitdaging het best kunnen beantwoorden door ontwikkelingen te monitoren, te blijven leren van archiefvormers en hen hulp te bieden bij hun archiefbeheer. Bij verwerving en opname dienen risico's, zoals incompleetheid, betrouwbaarheid, bruikbaarheid en afhankelijkheden beoordeeld te worden. Het onderzoek heeft voor HNI de opgave duidelijk gemaakt ten aanzien van de bureaus en ontwerpers.

De architectengemeenschap betrekken en mee laten denken rond de verwerving van een digitaal archief creëert betrokkenheid en levert een schat aan informatie op. In oktober 2015 vond er bij Het Nieuwe Instituut een bijeenkomst plaats in het kader van de overdracht van het digitale archief van de eerste 400 MVRDV-projecten. Rond de digitale archieven van MVRDV en Zaha Hadid Architects kwamen vragen aan de orde als: welke items bewaren architecten en welke items vinden erfgoedorganisaties

⁸ Voor de *Archeology of The Digital*, zie paragraaf "Kennis Opbouwen". Voor de website van CCA, zie <https://www.cca.qc.ca/en/> geraadpleegd 11 oktober 2018.

⁹ Bit preservation houdt in dat maatregelen worden genomen om te verhinderen dat de bitsequentie van een bestand wijzigt.

¹⁰ Functionele preservatie houdt in dat maatregelen worden genomen die het behoudt van de eigenschappen van een digitaal object in de toekomst verzekeren.

¹¹ Henk Vanstappen, *Opname en verwerking van born digital objecten uit een architectuurarchief (2013)* https://www.architectuurarchiefvlaanderen.be/sites/default/files/publicaties/cadarchivering_rapport_2013_v0_7.pdf geraadpleegd 11 oktober 2018

¹² Wim Lowet, *Bouwstenen voor de archivering van het digitaal archief Maarten Van Severen, 2016*: https://www.architectuurarchiefvlaanderen.be/sites/default/files/publicaties/mvs_20160907_rapport_v1-3.pdf

¹³ Zie voor de survey van het VAI: <https://www.architectuurarchiefvlaanderen.be/node/23> geraadpleegd 11 oktober 2018. Verslag bureauenquêtes project Digitaal Depot, Frans Neggers, Sanja Lap, NAI, 8-8-2008.

en musea belangrijk om te preserven en op welke nieuwe wijzen kunnen born-digital archieven bevraagd worden? Maar ook andere zaken rond de gevolgen van digitaal werken voor de ontwerppraktijk, de architectuur, selectie en preservering kwamen aan de orde, alsook de belangen van gebruikersgroepen.

Hieruit moge blijken dat betrokkenheid van de architecten bij het beheer van de digitale archieven essentieel is.

Een aanpak om samenwerkingen op te zetten tussen archiefinstelling en architectenbureaus werd in 2016-2017 diepgaand onderzocht en uitgetest via het project *The archives @ the architects*.¹⁴ Dit project werd uitgevoerd door het Architectuurarchief van de Provincie Antwerpen¹⁵ in samenwerking met het Vlaams Architectuurinstituut en met waardevolle input van Het Nieuwe Instituut.

Het ging om een haalbaarheidsstudie waarbij we onderzochten:

- 1) op welke manieren een archiefinstelling een bedrijf als een architectenbureau kan bijstaan in het document- en archiefbeheer.
- 2) of dit zowel voor het architectenbureau als de archiefinstelling tot winsten leidde, zodat een business model kan worden opgesteld.
- 3) of architectenbureaus konden bewogen worden tot vervroegde overdrachten van hun archieven, zodat preservatieacties sneller kunnen gebeuren.

Hiervoor werden vier case studies opgezet met een even groot aantal architectenbureaus van variabele omvang en datum van oprichting. Iedere case bestond uit een voorstelling van het project, een informatieanalyse met interviews en automatische screenings van de serverinhoud, het opstellen van een risicoanalyse voor het digitaal informatiebeheer en een ronde van feedback met de architecten. Tijdens de feedbackronde werden architecten gevraagd naar het nut van mogelijke diensten als een helpdesk, een kennisplatform, informering via nieuwsbrieven, het organiseren van opleidingen, het aanbieden van back-upvoorzieningen, vliegende archivarissen enz.

Een belangrijke conclusie daarvan was dat de informatieanalyse zowel door het architectenbureau als de archiefinstelling als erg positief en nuttig werden ervaren.¹⁶ De bureaus kregen een gratis audit en cursus archiefbeheer, terwijl de archiefinstelling op die manier een vertrouwensband kan opbouwen en waardevolle informatie verwerft over de archiefvorming, wat tijd bespaart bij een effectieve archiefoverdracht.

De risicoanalyse – opgesteld op basis van de categorieën van het scoremodel – werd vooral gewaardeerd door de bureaus van grotere omvang. De risicoanalyse bood een conceptueel kader waarbinnen de bureaus hun informatiebeheer konden uitbouwen. Omdat grotere bureaus beschikking hebben over meer middelen, was hun interesse naar bijkomende praktische ondersteuning eerder lauw. Kleinere

¹⁴ De doelstellingen, het verloop en het eindrapport van het project zijn gebundeld op volgende webpagina: <https://www.architectuurarchiefvlaanderen.be/nl/artikel/project-the-archives-the-architects> geraadpleegd 11 oktober 2018.

¹⁵ Zie voetnoot 1 voor Het Architectuurarchief van de Provincie Antwerpen

architectenbureaus bleken daarentegen veelal interesse te hebben in het tijdelijk ter beschikking stellen van archivarissen en praktisch ingestelde cursussen over digitaal archiefbeheer. Praktische ondersteuning is hier dus vooral het ordewoord, wat voor ons als archiefinstellingen een uitdaging is om dergelijke ondersteuningsmodellen te ontwikkelen.

Algemeen was vastgesteld dat geen enkel bureau de problemen met digitaal archief als dermate groot ervaart, om bereid te zijn voor externe consultancy te betalen. Ietwat verrassend was de vaststelling dat veel architecten mogelijk informatieverlies in hun digitaal archief (bijvoorbeeld een corrupt bestand) niet als een al te grote ramp beschouwden. Met deze aspecten moeten wij rekening houden wanneer we business modellen opstellen die vertrekken vanuit het principe van een win-win situatie.

Een laatste belangrijke bevinding was tot slot dat de bureaus veelal niet bereid waren om digitaal archief vervroegd over te dragen om verschillende redenen, zoals bedrijfsgeheim, een gevoel van controleverlies en omwille van emotionele redenen. Hierop moeten we dus antwoorden zien te formuleren, hetzij door de creatie van preceden-ten waarmee architecten sneller kunnen worden overtuigd, hetzij in de vorm van alternatieve acquisitiestrategieën, zoals verwervingen in het kader van tentoonstellingen of publicaties.

Naar een digitaal depot

Een digitaal depot is het geheel van organisatie, beleid, processen, middelen, personeel, hardware en software, die het archiveren van digitale objecten mogelijk maken. In 2017 werd bij Het Nieuwe Instituut het project inrichting digitaal archief opgestart. Dit beoogt het realiseren van een voorziening, het ontwikkelen van beleid en strategie voor digitale preservering, het inrichten van workflows en de verdere professionalisering van de organisatie. Binnen het project zal een digitaal archief laboratorium als onderzoeks- en ontwikkelomgeving als eerste stap naar een operationeel digitaal archief ingericht worden. Daarbinnen zal onder meer preserveringsoftware getest, ingericht en ontwikkeld worden.

Nu het Vlaams Architectuurinstituut een collectie architectuurarchieven beheert, dient ook deze organisatie een oplossing te vinden voor de bewaring van digitale archieven. Omdat het Vlaams Architectuurinstituut als kleine organisatie op dit moment niet de middelen heeft om op eigen houtje een digitaal depot te ontwikkelen of aan te schaffen, sloot het instituut zich aan bij zes andere culturele archiefinstellingen in Vlaanderen die dezelfde nood delen. Samen onderzoeken deze instellingen van 2019 tot 2024 de haalbaarheid van een gedeelde digitaal depot-oplossing.¹⁷ De expertise die we over digitale archieven ondertussen hebben opgedaan vormt daarvoor een belangrijke basis.

¹⁶ De methodiek die we toepasten voor de informatieanalyse kun je nalezen op: <https://www.architectuurarchiefvlaanderen.be/nl/artikel/the-archives-the-architects-tools> geraadpleegd 11 oktober 2018.

¹⁷ De doelstellingen en stand van zaken worden gebundeld op een projectpagina op de website van CEST: https://www.projectcest.be/wiki/Projecten:Werkplan_Digitale_Preservering geraadpleegd 11 oktober 2018.

Conclusie

In bovenstaande tekst lichten we aan de hand van vier thema's de belangrijkste acties toe die Het Nieuwe Instituut en het Vlaams Architectuurinstituut hebben ondernomen om hun organisatie en competentieniveau te verhogen met betrekking tot de bewaring en ontsluiting van digitale archieven.

In retrospectief valt daarbij wellicht het verschil in aanpak en prioriteiten op, die valt te verklaren door het verschil in opzet van de organisaties. De ambitie van Het Nieuwe Instituut als grote collectiebeherende instelling was vooreerst gericht op professionalisering van de organisatie. De verwerving van het archief van bureau MVRDV zorgde voor veel onderzoeksmateriaal. Vervolgens werd er krachtig ingezet op de ontwikkeling van een digitaal depot.

Het Vlaams Architectuurinstituut is in vergelijking met Het Nieuwe Instituut een kleinere instelling die tot 2018 met betrekking tot digitale archieven vooral een onderzoeks- en expertisefunctie had. De onderzoeken die het VAI deed hadden daarom vooral kennisverspreiding naar andere archiefbeherende organisaties tot doel. Het VAI sloot daarbij zoveel mogelijk aan bij bestaande initiatieven, methodes en denkkaders en testte deze uit op de specificiteit van architectuurarchieven. Nu het VAI ook een collectiebeherende instelling geworden is, wordt een oplossing voor een digitaal depot ook voor deze organisatie een prioriteit.

Samenwerking is van groot belang bij het beantwoorden van de opgave. Samenwerking was er onder meer door de betrokkenheid van het VAI bij de leeromgeving terwijl Het Nieuwe Instituut belangrijke input gaf in de adviesgroep van het project.

In dit artikel is duidelijk geworden wat de opgave is waarvoor de instituten zich geplaatst zien en welke stappen we de afgelopen tien jaar hebben ondernomen. De opgave die er nog ligt werd nog eens duidelijk op de Internationale Expertmeeting *Archiving Digital Architectural Heritage* die Het Nieuwe Instituut maart 2018 organiseerde. Het ligt expliciet in onze bedoeling dat we in de toekomst nauw blijven samenwerken. Wij gaan er hierbij vanuit dat het verschil in de afgelegde weg en de verschillende contexten als een grote verrijking kan uitpakken.

Continue in beweging

Het beheer en behoud van digitale kunst

Digitale kunst

De aanwezigheid en het gebruik van nieuwe media en techniek in de kunst en samenleving neemt exponentieel toe maar gaat helaas even snel verloren. En daarmee ons geheugen aan digitale kunst en een deel van onze culturele geschiedenis. In onze samenleving, waarin media en techniek een belangrijke rol spelen, is (toekomstige) beschikbaarheid van dit innovatieve deel van ons cultureel erfgoed, de digitale kunst, niet vanzelfsprekend. Digitale kunst beweegt zich in een veranderlijk, hoog technologisch veld, een niche met een kleine interdisciplinaire en zeer internationale *community* met een grote uitstraling.

Digitale kunst is een verzamelnaam voor videokunst, geluidskunst, media installaties, born digital art en bijvoorbeeld *online* kunstwerken. De hierin gebruikte analoge, gedigitaliseerde en *born digital* formaten, dragers, hard- en software zijn continu onderhevig aan technologische veroudering. Om het functioneren van het kunstwerk te behouden, is het kunstwerk zelf onherroepelijk aan verandering onderhevig. Dergelijke veranderingen horen intrinsiek bij dit type kunstwerken. Hierdoor zijn, in tegenstelling tot traditionele kunstvormen, begrippen als 'authenticiteit' of 'het origineel' minder leidend. Maar wat moet er gebeuren om de levensduur van een digitaal kunstwerk te verlengen, terwijl de integriteit van het kunstwerk niet verloren mag gaan? De werken zijn multimediaal, interactief, worden live gegenereerd en maken gebruik van meerdere aan elkaar verbonden netwerken. Daarbij vormen ze niet langer een onveranderlijk eindproduct, maar zijn in toenemende mate divergent: ze zijn variabel, vele versies bestaan naast elkaar en kunnen op verschillende tijdstippen en platformen (*online* en *offline*) door verschillende mensen worden gecompileerd. Het borgen van de toegankelijkheid van dergelijke dynamische digitale kunstwerken vraagt om de continue opbouw en deling van specifieke kennis en een gezamenlijke aanpak.

Musea voor hedendaagse kunst worden in toenemende mate geconfronteerd met digitale kunstwerken. Om de werken blijvend een plek te geven binnen het museum is een andere benadering nodig dan bij traditionele kunst of digitaal erfgoed. De Nederlandse musea hebben op dit moment een kleine selectie digitale kunstwerken in hun collectie opgenomen en hebben nog niet veel ervaring in de omgang met deze kunstwerken. Het beheer en behoud ervan is deels nog onbekend terrein, terwijl collecties wel pro-actief moeten opereren en kennis hierover nodig hebben om dit kwetsbare hedendaagse erfgoed op te kunnen nemen en toegankelijk te kunnen

houden voor toekomstige presentaties. Het aantal digitale kunstwerken zal in de toekomst sterk stijgen en naar verwachting ook steeds meer opgenomen worden in museale collecties en archieven. Digitale kunst is zeer gevoelig voor verval, zowel *hardware* als *software* raken op den duur in onbruik en functioneren niet meer. Slechts een heel klein deel van de restauratoren wereldwijd heeft enige training op dit gebied, maar toch ziet een groeiend aantal collectiebeheerders zich belast met het behoud van deze digitale kunstwerken. Zij worden geacht de conditie van deze kunstwerken te monitoren, conservering protocollen op te zetten en wanneer noodzakelijk adequaat te handelen. Door geringe opname in collectie en kennis van conservering en toekomstige presentatie raken deze werken in de vergetelheid, wat leidt tot hiaten in het oeuvre van een kunstenaar, of zelfs in een gehele periode in de kunstgeschiedenis. Er bestaan echter wel degelijk oplossingen voor obsoleete digitale kunstwerken, zelfs decennia later. De technische problemen zijn wellicht complex, maar absoluut niet onoverkomelijk.¹

In Nederland ondersteund LIMA musea, collecties, producenten en kunstenaars bij de conservering, documentatie en digitale duurzaamheid van hun media kunstwerken. LIMA biedt kennis en diensten aan aan zowel collecties als kunstenaars en heeft een domeinspecifieke, digitale repository voor mediakunst, waarin (digitale) mediakunstwerken uit 30 Nederlandse collecties worden bewaard voor de toekomst.

Conserverings strategieën

Conservering begint bij de aankoop van een werk, en niet pas als degradatie wordt geconstateerd of het werk niet meer blijkt te functioneren. Documentatie is daarbij essentieel. Daarvoor documenteren we in het algemeen de significant properties zoals gebruikelijk bij digitaal behoud en de score zoals gebruikt bij de conservering van mediakunst. Significant properties zijn essentiële kenmerken van een digitaal object die in de loop van de tijd behouden moeten blijven zodat het digitale object toegankelijk en zinvol blijft. Een goed begrip van deze eigenschappen is van cruciaal belang voor het vaststellen van de beste werkwijze en helpt bij het beantwoorden van de fundamentele vraag met betrekking tot digitale bewaring: wat te bewaren?² Het is belangrijk om de kern en het concept van het werk, de werking ervan en de uiterlijke kenmerken die daardoor worden bepaald, de intentie van de kunstenaar en haar standpunt ten opzichte van verandering van het werk, te documenteren zodra het werk een collectie binnenkomt. Deze documentatie vormt het noodzakelijk referentiekader voor toekomstige professionals die belast zijn met het behoud ervan en is nodig om te kunnen bepalen welke conservering strategie de beste is. De intake van een kunstwerk is het beste moment om zoveel mogelijk informatie van de kunstenaar te verkrijgen. Hiervoor zijn interview protocollen³ en intakeformulieren⁴ in gebruik.

¹ Nina van Doren en Marleen Wagenaar 'Behoud van digitale kunstwerken grote uitdaging voor collectiehouders. DOCUMENTATIE VORMT EERSTE EN BELANGRIJKSTE STAP' in *Project Transformatie Digitale Kunst* (SMBK/LIMA, 2016) p. 33-45 via <http://www.sbm.nl/pubs/detail/id/23> geraadpleegd 11 oktober 2018.

² <https://www.dpconline.org/events/past-events/significant-properties> geraadpleegd 11 oktober 2018.

³ Lydia Beerkens et al. *The Artist Interview – for Conservation and Preservation of Contemporary Art. Guidelines and Practice* (Jap Sam Books. RCE, SBMK, UvA, 2012)

⁴ Aanleverspecificaties voor videokunst en Software / Computer-based Art, (LIMA 2016).

Er zijn globaal vier conservering strategieën voor digitale kunst die in de afgelopen jaren uitvoerig zijn beschreven.

- OPSLAG: het bewaren van de originele hard- en software.
- MIGRATIE: het 'updaten' van het werk naar een nieuwe standaard.
- EMULATIE: het nabootsen van de originele technische omgeving op een nieuwe computer.
- HERINTERPRETATIE: het steeds opnieuw uitvoeren van het werk (al dan niet op basis van documentatie).

Welke conservering strategie het meest geschikt is, hangt af van het werk, en kan in de loop van de tijd veranderen. Belangrijk hierbij is hoe en wanneer het werk gemaakt is en met welk concept.



Afbeelding 1. Rafaël Rozendaal verklaart de werking van zijn sites op *Transformation Digital Art*, 2016, foto José Biscaya

Wereldwijd wordt sinds ruim 15 jaar onderzoek gedaan naar de conservering van digitale kunst. In diverse samenwerkingsverbanden is nationaal en internationaal in eerste instantie vooral casestudy-based onderzoek gedaan. Het accent lag daarbij op de ontwikkeling van modellen en strategieën voor conservering en documentatie, risicoanalyse, destillatie van intentie en de bepaling van de significant properties. Buitenlandse betrokken partijen daarbij zijn bijvoorbeeld Daniel Langlois Foundation, MOMA, SF MOMA, ZKM, TATE, Smithsonian Institution, Universiteit Bern, Solomon R. Guggenheim Museum en PACKED. Binnenlandse partijen zijn RCE/INCCA (Project Inside Installations/PRACTICS), SBMK en LIMA (Projecten Born Digital Kunstwerken in Nederland en Transformation Digital Art) en

CCDD/NDE (Projecten CD-ROM archiving en Generieke Workflows). Voortdurend worden de richtlijnen bediscussieerd en (aan de hand van casestudies) verder ontwikkeld. De richtlijnen voor elementaire behoud-procedures van digitale kunst, zoals het testen van apparatuur, het vervaardigen van back-ups/disk images, de documentatie van software, opslag en controle, zijn bekend en beschikbaar maar worden slechts door een beperkt aantal instellingen toegepast. Men zou kunnen zeggen dat, hoewel er uitzonderingen zijn, over het algemeen het verzamelen en duurzaam bewaren van digitale kunst nog grotendeels in de kinderschoenen staat. Redenen daarvoor zijn

- 1) Het feit dat conservering en duurzame opslag van digitale kunstwerken om een andere zorg vraagt dan traditionele kunstvormen. Het werk is digitaal, heeft een continue en actieve aanpak en specialistische kennis nodig en daar hangt een prijskaartje aan.
- 2) Musea hebben steeds minder geld te besteden aan conservering.
- 3) De meeste museum professionals hebben doorgaans niet de expertise en apparatuur om deze kwetsbare kunstwerken te conserveren.
- 4) Opbouw van deze expertise en uitbesteding van bijvoorbeeld opslag is in de bekostigingssystematiek bij de musea nog niet (voldoende) doorgevoerd.

De toonaangevende instellingen hebben inmiddels de onderhoudsprocedures voor digitale kunstwerken geïmplementeerd en werken aan protocollen voor het managen van de files en hun levenscyclus. Zo hebben bijvoorbeeld MOMA, Guggenheim en TATE een *repository* en workflow ingericht, voor digitale kunstwerken (videokunst, films en installaties) en doen zij *casestudy based* onderzoek naar de conservering van complexe digitale werken. LIMA heeft een operationele *repository* voor digitale kunst en complexe digitale kunstwerken en werkt aan de uitbreiding van de workflow voor online kunstwerken. Uit lopend onderzoek naar praktijken voor opslag van complexe digitale en online kunstwerken blijkt dat het systematisch verzamelen en duurzaam opslaan alleen bij gespecialiseerde instellingen plaatsvindt en (nog) nauwelijks bij de kunstmusea.

De algemeen geaccepteerde gang van zaken voor de conservering van complexe digitale werken en online kunstwerken is als volgt:

- Het beschrijven van het werk en haar functioneren.
- Het verifiëren en beschrijven van de *files*, besturing en uitvoering.
- Het vastleggen van significant properties en afhankelijkheden.
- Het interviewen van de kunstenaar.
- Het opslaan van de *files*.
- Het maken en opslaan van *diskimages*.
- Het opslaan van de broncode.
- Het monitoren van de *files* en opnieuw afspelen en verifiëren van de disk image.

Op dit moment wordt een enquête uitgevoerd onder 10 internationale instellingen om de stand van zaken met betrekking tot metadata, versie management, dynamisch archiveren, opslag voor complexe digitale kunst en online kunst te inventariseren.⁵ De voorlopige uitkomsten hiervan zijn als volgt: de musea verrichten case study based onderzoek naar complexe digitale kunst en *online* kunstwerken een workflow is nog niet geïmplementeerd.

Emulatie

Het is geen wonder dat emulatie en virtualisatie steeds meer ook in het museale veld onderzocht worden, ze bieden een potentiële bewaarstrategie voor digitale kunstwerken door de afhankelijkheden van het kunstwerk los te koppelen van efemere fysieke computerhardware, met een korte digitale levenscyclus. In 2004 vond in New York het eerste symposium plaats over de praktische en theoretische toepassingen van emulatie bij de conservering van digitale kunst.⁶ Tot voor kort werden EMil, EAAS en andere vormen van emulatie vooral voor CD ROM emulatie ingezet.⁷ Daarbij werd emulatie vooral als documentatie strategie, voor toegankelijk van de werken, en niet zozeer als conserverings strategie gezien omdat de authentieke beleving sterk verandert. Vanaf 2016 kwam daar verandering in, er verscheen een uitvoerig onderzoeksrapport over mogelijkheden in museaal verband.⁸ Het vigerend onderzoek vond plaats in het kader van PERICLES.⁹ Emulatie en software conservering is een van de onderwerpen die ook bij IPRES steeds meer aandacht krijgen. Zo won in 2016 Exhibiting Digital Art via Emulation - Boot-to-Emulator with the EMil Kiosk System van Dragan Espenschied, Oleg Stobbe, Thomas Liebraut and Klaus Rechert de *best paper award*. Waarna Rhizome tools ontwikkelde zoals oldweb.today inclusief emulatie en Webrecorder voor de archivering van websites. Deze tools tonen aan dat de mogelijkheden van emulatie als service en de opname van deze service in de conserverings workflow goede kansen bieden. De volgende stap in de workflow ontwikkeling zal dan ook de implementatie van emulatie zijn.

Herinterpretatie

De voortdurende vernieuwing van media en technologie geeft kunstenaars steeds weer nieuwe expressiemogelijkheden en uitdagingen voor interactie met het publiek. De continu veranderende browsers, tools, operating systems en computers, maken het presenteerbaar houden van digitale kunst tot een uitdaging. Het behoud van deze werken is veelal gericht op technische handelingen. De vergankelijkheid en immaterialiteit van deze werken leiden immers steeds opnieuw tot een reeks handelingen en technieken om de toekomstige uitvoering te waarborgen. LIMA past de verschillende conserverings strategieën toe in haar werk. Alle herinstallaties of herinterpretaties suggereren vertaling en *fidelity*. We zijn daarbij niet objectief of passief, we zijn actief bezig met het construeren en reconstrueren van de (authenticiteit van het) werk gebaseerd op ons begrip van het werk en behoud dogma's. Dat geldt voor al onze diensten en projecten maar geldt in overtreffende trap bij de minst gangbare conservering strategie herinterpretatie. Een artistieke strategie en praktijk die al eeuwenlang gangbare is in muziek, dans en theater, die nu mondjesmaat op academies en in workshops wordt gebruikt maar op professioneel niveau nagenoeg nieuw is in het veld van digitale kunst. Als we echter verder vooruit denken, de technologie voorbij, is het ontwikkelen van een script voor herinstallatie op termijn waarschijnlijk de beste optie voor het voortbestaan van een werk en

⁵ *Survey of digital art storage practices in the media art world* (LIMA). Verschijnt najaar 2018.

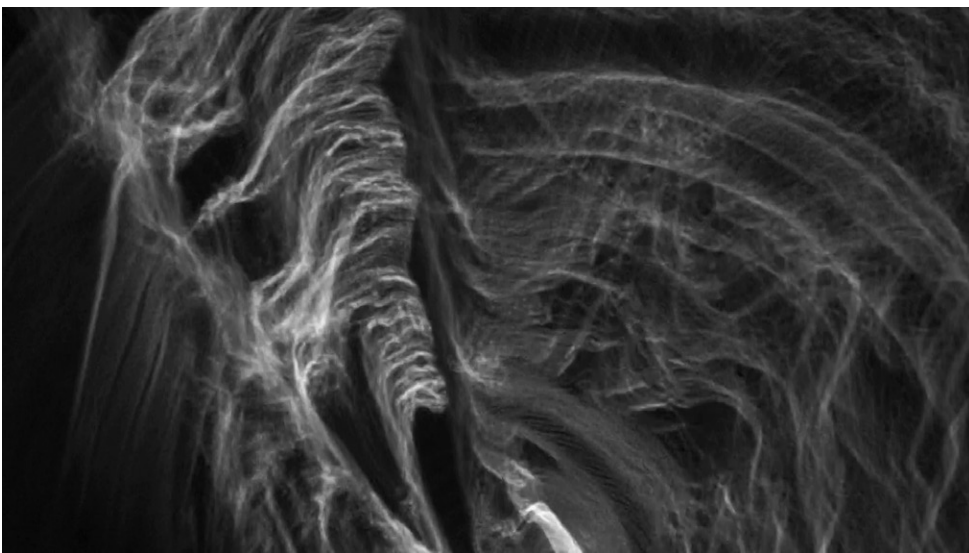
⁶ <http://www.variablemedia.net/e/seeingdouble/> geraadpleegd 11 oktober 2018.

⁷ http://www.ncdd.nl/wp-content/uploads/2015/11/201611_DE_Houdbaar_Final_report_CD-ROM_Archiving_DEF.pdf geraadpleegd 11 oktober 2018.

⁸ Klaus Rechert (University of Freiburg), Patricia Falcao (Tate) and Tom Ensom (Kings College) *Introduction to an emulation-based preservation strategy for software-based artworks*. (13 December 2016) Commissioned by Tate PERICLES.

⁹ <https://www.tate.org.uk/about-us/projects/pericles> geraadpleegd 11 oktober 2018.

om ervoor te zorgen dat het begrijpelijk blijft. Hoe kunnen we creatieve digitale werkprocessen bij productie en presentatie, tussen kunstenaars en publiek, inzichtelijk maken, contextualiseren, documenteren, analyseren en uiteindelijk begrijpen en overdragen in de context van de digitale cultuur en het grotere culturele veld? 'Wat is de kern en productiewijze van een werk', 'Welke technieken worden in welke context gebruikt?' 'Hoe voer je een werk zelf uit?' en 'Hoe hou je een werk levend?' Hoe vertalen we deze artistieke praktijk en kennis naar de volgende generatie? In het project UNFOLD (LIMA 2016-2017)¹⁰ stonden deze vragen centraal.



Afbeelding 2. Joost Rekveld #67, 2017, 17'01", videostill

Dit project bestond uit verschillende onderdelen:

- Een onderzoek naar herinterpretatie als conserveringsstrategie.
- Drie publieke presentaties en drie netwerkbijeenkomsten, met interdisciplinaire en internationale participanten, sprekers en gasten en vertoning van kunstwerken.
- Een productie van Joost Rekveld, waarvan het werk in process tijdens de bijeenkomsten getoond en besproken werd, met de premiere op en in samenwerking met het Internationaal Filmfestival Rotterdam.
- Een workshop in samenwerking met Sonic Acts waarvoor kunstacademies uitgenodigd werden en een presentatie van de workshop resultaten voor een wat technisch publiek.
- Een vooronderzoek naar de mogelijk herinterpretatie van performance based werk van Nan Hoover door dramaturg Fransien van der Putt en choreograaf en danser Vera Sofia Mota.

¹⁰ www.li-ma.nl/site/article/unfold-mediation-reinterpretation geraadpleegd 11 oktober 2018.

Joost Rekveld produceerde gedurende het project zijn werk **#67** een herinterpretatie van de werken **Reminiscence** (1974) en **Telc** (1974) van Steina en Woody Vasulka, en deelde de esthetische, technische en theoretische problemen waarmee hij tijdens de productie werd geconfronteerd. Er werden vooraf geen voorwaarden gegeven waaraan de herinterpretatie moest voldoen, het artistieke proces werd gevolgd. Het creatieve productieproces volgen als herinterpretatie, vestigde de aandacht op het productieproces in zijn geheel en niet alleen op het 'eindproduct'. Drie momenten in het proces waren daarbij essentieel:

- De pre-productiefase: Rekveld's selectie van werken om te herinterpreteren, discussies over het werk Telc en herinneringen aan The Vasulkas, onderzoek naar de technische middelen waarmee de The Vasulkas produceerden, enzovoort. Herinterpretatie functioneert hier als stimulans voor onderzoek.
- De workshop, waarin de nadruk vooral lag op de technische productie van draagbare apparaten voor modulatie, perceptie en verkenning van de mogelijkheden.
- Het uiteindelijke kunstwerk waarbij herinterpretatie het voornaamste doel was.

De inspanningen resulteerden in een werk dat structurele overeenkomsten vertoont met de werken van The Vasulkas, zowel op het niveau van het medium als dat van de ervaring van de toeschouwer. Rekveld creëerde een soortgelijke temporele structuur, met materiaal opgenomen op een wandeling dat hij later manipuleerde. Ook koos hij voor eenzelfde eerstpersoonsperspectief dat lijkt op een bijna abstracte, maar haptische, beeldruimte en de nog herkenbare omgevingsgeluiden. Joost Rekveld beschouwde herinterpretatie als een proces, wat een andere dimensie aan zijn werkwijze gaf. De dimensie van een artistieke methode, aanpak of 'manier van werken'. Zijn onderzoek naar de esthetische en technische kanten van de werken van The Vasulkas gaf veel inzicht in manieren van behoud en in zijn eigen artistiek proces. Rekveld verklaarde dat het project leidde tot het ontwikkelen van een interesse in de *Rutt /Etra Scan Processor* – gebruikt door de Vasulkas en nauwelijks in gebruik vandaag de dag. Dit leidde uiteindelijk tot het bouwen van een dergelijke processor en gebruiken voor de realisatie van het werk #67.

Uit het project werden de volgende conclusies getrokken:

- Zet herinterpretatie in als artistiek instrument, een manier van denken met betrekking tot het digitale. Als methode die het mogelijk maakt om in te zoomen op specifieke aspecten van een werk in zijn eerdere uitvoeringen en in hoeverre deze kunnen worden ingezet in een nieuwe digitale context.
- Herinterpretatie leidt ons niet weg van het denken over de materialiteit van het digitale. Herinterpretatie wordt niet gezien als vervanging van het denken over materialiteit maar als aanvulling.
- Herinterpretatie vertelt ons veel over het werk zelf en geeft ruimte voor creativiteit. Het houdt het werk op een geheel andere manier levend; het zorgt voor de voortdurende actuele presentatie en overlevering.
- Initieer en stimuleer het maken van nieuw werk op basis van herinterpretatie als een strategische activiteit voor het dynamisch behoud van digitale kunstwerken.
- Breidt de multidisciplinaire en internationale samenwerking uit.
- Deze traditionele vorm van overlevering wordt in de beeldende kunst tot alternatieve conservering strategie.

Lessons learned

Het delen van informatie en ervaring binnen en buiten de internationale media-kunst community is belangrijk voor de ontwikkeling van de gemeenschap, zowel voor de praktijk van de makers als voor de duurzame bewaring van digitale kunstwerken bij musea. Binnen het domein is de kennis aanwezig om te bepalen welke digitale objecten, formaten, metadata en documentatie nodig zijn. Ook de technologie-watch en de toegang, in welk geval toekomstige presentatie vraagt om domeinspecifieke kennis van formaten, rechten, klanten, gebruik en meer.

We kunnen veel leren van de processen in digitale archivering, maar standaardisatie en bulk conservering zijn problematisch voor het behoud van digitale kunstwerken. Dit komt mede door het belang van het behoud van de authenticiteit, authentieke uitvoering, installatie en presentatie en ervaring en samenwerking met de kunstenaars. Kunstenaars gebruiken bovendien vaak de nieuwste technologieën en verkennen hun technische eigenschappen en mogelijkheden. Deze experimentele werken zijn bijzonder kwetsbaar. Normalisatie van born digital kunstwerken naar een standaard bestandsformaat heeft niet de voorkeur. Dit kan immers het kunstwerk veranderen door zijn esthetiek en perceptie te beïnvloeden, en daarmee zijn intentie en waarde. Digitale kunst is maatwerk en vereist op maat gemaakte benaderingen voor het behoud en de presentatie ervan. Bovendien is er een grote verscheidenheid aan born-digital kunstwerken.

De keus voor een ‘vergankelijk’ medium maakt dat makers van digitale kunst vaak specifieke ideeën hebben over de toekomst en het voortbestaan van hun werk. Bovendien kan de maker (soms als enige) inzicht verschaffen in het maakproces van zijn werk en technische gegevens verstrekken die noodzakelijk zijn voor de toegankelijkheid. Daarom zet LIMA naast haar ondersteuning en dienstverlening aan musea en collecties nadrukkelijk in op ondersteuning en dienstverlening aan kunstenaars. De kunstenaar is een actieve betrokkene in het conserveringsproces van zijn eigen werk. Beheer en behoud van digitaal erfgoed begint al bij de productie. Wat betreft duurzaamheid en goede en efficiënte archivering van born digital objecten, ligt een deel van het probleem en een deel van de oplossing bij de bron: de formaten en de bijbehorende soft- en hardware, de kwaliteit, ordening en selectie van de bestanden en de informatie over de bestanden. Ook de opslag bij de makers zelf is een punt van aandacht. Training is belangrijk voor alle aspecten van de digitale duurzaamheid. Hoe vroeger in het proces geïmplementeerd, des te effectiever de digitale duurzaamheid.¹¹ Documenteren en bewaren van kunstwerken is ook voor kunstenaars niet gemakkelijk, of wordt bemoeilijkt door gebrek aan kennis, tijd en middelen. Daarom heeft LIMA de Artwork Documentation Tool ontwikkeld om kunstenaars bewust te maken van de problematiek en ook te ondersteunen in de DIY conservering van hun eigen werk.¹²

¹¹ Gaby Wijers en Hannah Bosma *Born digital cultureel erfgoed is bedreigd erfgoed: Op weg naar een generieke workflow voor born digital erfgoed binnen de domeinen kunst, film, fotografie en architectuur* (CCDD, DEN en Lima, 2015) http://www.den.nl/art/uploads/files/Publicaties/Born_Digital_erfgoed_is_bedeigd_erfgoed.pdf geraadpleegd 11 oktober 2018.

¹² <https://www.li-ma.nl/ad/geraadpleegd> 11 oktober 2018.

Casus Universiteit Leiden: Duurzame toegang tot research data

Beleid, protocollen, voorzieningen,
standaarden en het ontwikkelen van vaardigheden
en competenties

De digitale transitie in wetenschappelijk onderzoek

De beschikbaarheid van grote hoeveelheden digitale data leiden ertoe dat e-science in alle wetenschappelijke disciplines een sterke groei doormaakt. Dit heeft niet alleen verstrekkende gevolgen voor de manier waarop onderzoek wordt uitgevoerd, het is ook een belangrijke bijdrage aan de verdere ontwikkeling van Open Science. Voor het beantwoorden van nieuwe wetenschappelijke vragen is het noodzakelijk dat data en metadata toegankelijk zijn, (inter)disciplinair samengewerkt wordt en dat nieuwe methoden, technieken en instrumenten verder ontwikkeld worden. Voor het gebruik hiervan zijn nieuwe vaardigheden nodig, de zogeheten 21st century skills, onder meer samenwerken, sociale en culturele vaardigheden en digitale competenties.¹

E-science beperkt zich dus niet alleen tot de wijze waarop onderzoek uitgevoerd wordt. Het raakt ook aan het publicatieproces, de omgang met onderzoeksdata, het wetenschappelijke evaluatie- en waarderingsproces en de toegankelijkheid op de lange termijn van wetenschappelijke producten zoals publicaties, data en research software.

Het belang van datamanagement

De Europese Commissie en de Nederlandse overheid, financiers van onderzoek, wetenschappelijke uitgeverij, universiteitsbesturen en wetenschappers stellen in toenemende mate eisen aan de omgang met onderzoeksdata. Daarnaast hebben wetenschappers ook wensen om beter gebruik te kunnen maken van de beschikbare hoeveelheden digitale data. Onder datamanagement wordt verstaan het creëren, opslaan, onderhouden, beschikbaar maken, archiveren en langdurig bewaren van onderzoeksdata.

¹ Remco Pijpers *Alles wat je moet weten over 21e eeuwse vaardigheden* (2017) zie <https://www.kennisnet.nl/artikel/alles-wat-je-moet-weten-over-21e-eeuwse-vaardigheden/> geraadpleegd 20180927

Goed datamanagement is om verschillende redenen belangrijk:

- Het komt de kwaliteit van het onderzoek ten goede;
- In het licht van wetenschappelijke integriteit is het belangrijk om zorgvuldig vast te leggen hoe data verworven en beheerd worden en wie wanneer welke rechten heeft om de data aan te passen of te bewerken;
- Met veel onderzoeksdata moet uit het oogpunt van veiligheid en privacy uiterst zorgvuldig worden omgegaan;
- In een wereld waarin beperkt budget beschikbaar is voor wetenschappelijk onderzoek, is het cruciaal dat andere wetenschappers gebruik kunnen maken van bestaande datasets, zowel voor het verifiëren en reproduceren van onderzoeksresultaten als voor het beantwoorden van nieuwe onderzoeksvragen.

De samenleving wordt in sterke mate beïnvloed door wetenschap. Wetenschappers en onderzoeksorganisaties zijn zich hiervan bewust en besteden naast excellent onderzoek eveneens aandacht aan kennisvalorisatie. Doordat onderzoek voor het merendeel uit publieke middelen gefinancierd wordt heeft het publiek ook recht op informatie over wat wetenschappers doen en op welke wijze zij dit doen. Door middel van Open Access publiceren en het toegankelijk maken van onderzoeksdata en research software moeten de resultaten van onderzoek voor iedereen beschikbaar zijn en voor de lange termijn als academisch erfgoed toegankelijk blijven.

In het *Nationaal Plan Open Science*, februari 2017, heeft een groot aantal wetenschappelijke organisaties hun ambities vertaald naar concrete doelstellingen.²

De belangrijkste ambities van het plan zijn:

- 100% Open Access publiceren;
- Onderzoeksdata optimaal geschikt maken voor hergebruik;
- Open Science onderdeel uit laten maken van het evaluatie- en waarderingproces en het stimuleren en ondersteunen van wetenschappers op alle terreinen van Open Science.

Duurzame toegang tot onderzoeksdata

Digitale duurzaamheid is geen nieuw onderwerp voor wetenschappelijke data. Omstreeks 1960 werd in Nederland het Steinmetz archief opgericht. Het Steinmetz stelde digitale data uit sociaalwetenschappelijk onderzoek beschikbaar. Door het toevoegen van beschrijvende en structurele metadata en het beschikbaar stellen van de data door middel van preferred formats zijn deze data nu nog steeds toegankelijk. Decennialang was duurzame toegang tot onderzoeksdata echter een activiteit waar slechts een relatief kleine groep wetenschappers zich mee bezig hield. Tot 2010 waren het voornamelijk de wetenschappelijke data archieven die zich inspanden om onderzoeksdata voor de langere termijn toegankelijk te maken. Zij hebben pionierswerk verricht op het terrein van metadata, voorkeursformaten, persistente identifiers, uitwisselingsprotocollen, citeerbaarheid van datasets en richtlijnen voor databewaarplaatsen.

De toenemende hoeveelheid digitale data (gedigitaliseerd en born digital), nieuwe analysemethoden en -technieken (Data Science) en nieuwe manieren van publiceren (Open Access) maken echter dat datamanagement gemeengoed is geworden in het onderzoek. De omslag naar verantwoord datamanagement is in feite een proces van gedragsverandering. Wetenschappers zullen in toenemende

mate moeten nadenken over de vraag of en hoe zij hun data structureren, delen en voor de lange termijn bewaren. Hiervoor is het essentieel om de workflow van het onderzoek in de discipline in kaart te brengen, om zo protocollen te ontwikkelen die beschrijven aan welke aspecten van datamanagement op welk moment aandacht moet worden geschonken. Alleen bewustwording en gedragsverandering bij wetenschappers is echter niet voldoende. Onderzoeksorganisaties moeten in staat zijn datamanagement door wetenschappers te ondersteunen en te faciliteren, door te zorgen voor adequate opslagvoorzieningen tijdens en na afloop van het onderzoek en door het geven van trainingen en advies. Hierbij moet er oog zijn voor de verschillen tussen wetenschappelijke disciplines: er is geen *one size fits all*-oplossing op het gebied van datamanagement. Ook moeten onderzoeksorganisaties beleid rondom datamanagement nader uitwerken: worden er verplichtingen opgelegd aan wetenschappers met betrekking tot het schrijven van een datamanagementplan, het opslaan van onderzoeksdata, duurzame toegang en/of andere aspecten? En als dergelijke verplichtingen ingevoerd worden, hoe worden die dan gemonitord? In dit artikel wordt met name beschreven hoe de huidige eisen en wensen met betrekking tot onderzoeksdata hebben geleid tot instellingsbeleid en op welke wijze de Universiteit Leiden dit beleid vertaalt naar werkprocessen, voorzieningen, de bijbehorende ondersteuning en de benodigde vaardigheden.

Van (inter)nationaal beleid naar instellingsbeleid

Delen en samenwerken zijn sleutelwoorden in de huidige wetenschapsbeoefening. Met de opkomst van Open Science worden het delen van onderzoeksgegevens en -resultaten en het samenwerken bij het verzamelen, analyseren en publiceren daarvan steeds belangrijker. Ook de kritische kanttekeningen die recent van diverse kanten bij bestaande benaderingen van wetenschapsbeoefening zijn geplaatst, vergen meer transparantie met betrekking tot de werkwijze van wetenschappers en de keuzes die ze daarbij maken.³

De Europese Commissie wil door middel van Open Science wetenschap meer open, mondiaal, collaboratief, creatief en maatschappelijk betrokken maken.⁴ In het Nederlandse regeerakkoord 2017-2021 *Vertrouwen in de toekomst* worden 'Open Science' en 'Open Access' de norm in wetenschappelijk onderzoek.⁵ Steeds meer (inter)nationale subsidieverstrekkingen stellen het opnemen van een datamanagementparagraaf in een onderzoeksvoorstel of het opstellen van een datamanagementplan na toekenning verplicht. In het nieuwe Standaard Evaluatie Protocol 2015-2021, dat wordt gebruikt voor onderzoeksvisaties, wordt onderzoeksgroepen gevraagd verantwoording af te leggen over de manier waarop zij data gebruiken en bewaren.

Om het delen van onderzoeksgegevens en -resultaten en het samenwerken bij het verzamelen, analyseren en publiceren daarvan mogelijk te maken, zijn er twee belangrijke richtlijnen of principes ontwikkeld om de duurzame toegang tot onderzoeksdata te borgen. Het CoreTrustSeal⁶ bevat een aantal richtlijnen waar

² *Nationaal Plan Open Science* (2017) zie <https://www.openscience.nl/> geraadpleegd 20180927

³ KNAW *Wetenschappelijke integriteit* zie <https://www.knaw.nl/nl/thematisch/ethiek/wetenschappelijke-integriteit> geraadpleegd 20180927

⁴ European Commission *Open Science* (2014) zie <https://ec.europa.eu/digital-single-market/en/open-science> geraadpleegd 20180927

⁵ Regeerakkoord 2017-2021 *Vertrouwen in de toekomst* (2017) zie <https://www.kabinetsformatie2017.nl/documenten/publicaties/2017/10/10/regeerakkoord-vertrouwen-in-de-toekomst> geraadpleegd 20180927

⁶ CoreTrustSeal *Data repository certification* zie <https://www.coretrustseal.org/> geraadpleegd 20180927

databewaarplaatsen aan moeten voldoen. Door middel van een peer reviewed self assessment kunnen databewaarplaatsen aantonen dat ze aan de gestelde richtlijnen voldoen. Onderzoeksfinciers stellen in toenemende mate als voorwaarde dat onderzoeksdata na afloop in zo'n zogenaamde Trusted Digital Repository gedeponeerd moeten worden. Wetenschappers behouden zo de mogelijkheid om een databewaarplaats te kiezen die voor hun vakgebied relevant is en de onderzoeksfinciers hebben de garantie dat de data ook na afloop van het onderzoek duurzaam toegankelijk blijft.

Het CoreTrustSeal is ontwikkeld binnen de academische gemeenschap. De richtlijnen zijn echter ook toepasbaar op erfgoedinstellingen. Binnen het samenwerkingsverband van de NCDD zijn de richtlijnen vertaald naar erfgoed en worden er workshops en trainingen gegeven aan erfgoedinstellingen over hoe zij aan de eisen van het CoreTrustSeal kunnen voldoen.

Een Trusted Digital Repository geeft dus een garantie dat onderzoeksdata voor de lange termijn toegankelijk blijft. Om de bruikbaarheid van onderzoeksdata tijdens en na afloop van het onderzoek te bevorderen zijn de FAIR principes ontwikkeld. De FAIR⁷ principes beschrijven hoe onderzoeksdata Findable, Accessible, Interoperable en Reusable gemaakt kunnen worden. De *Digitale Erfgoed Referentie Architectuur* (DERA)⁸ heeft als doel architectuurkaders op te stellen die bijdragen aan de concretisering van de Nationale Strategie Digitaal Erfgoed. De DERA bevat afspraken die samenwerking tussen de sectoren erfgoed, bibliotheek, musea en wetenschap moeten faciliteren. Vaak hebben deze sectoren een eigen referentiearchitectuur. De doelstellingen van DERA 2.0 sluiten aan bij de doelstellingen die gehanteerd worden in de Hoger Onderwijs Referentie Architectuur (HORA) waarin de FAIR principes zijn opgenomen.

Om aan de eisen en wensen van de verschillende belanghebbenden te kunnen voldoen hebben de meeste onderzoeksorganisaties de afgelopen jaren beleid opgesteld voor onderzoeksdata. Grofweg kan dit beleid onderverdeeld worden in richtlijnen voorafgaand, tijdens en na afloop van het onderzoek. De FAIR principes en gecertificeerde data repositories zijn hierin veelal leidend. Onderzoeksdata moet in ieder geval na afloop van het onderzoek voor ten minste tien jaar Findable, Accessible, Interoperable en Reusable zijn. Indien onderzoeksdata voor een langere termijn toegankelijk moet zijn dient deze opgeslagen te worden in een gecertificeerde data repository die minimaal voldoet aan de eisen van het CoreTrustSeal. Er moeten heldere afspraken zijn over de werkprocessen en de datamanagementdiensten die deze werkprocessen ondersteunen. Onderzoeksorganisaties zijn op dit moment volop bezig om het beleid te vertalen naar de bijbehorende werkprocessen, diensten en ondersteuning. Hierbij dient rekening gehouden te worden met de verschillende onderzoeksmethoden en -technieken die er in de verschillende wetenschappelijke disciplines gebruikt worden.

Regeling Datamanagement Universiteit Leiden

De Universiteit Leiden, in 1575 opgericht, is een van Europa's meest vooraanstaande internationale onderzoeksuniversiteiten. De universiteit heeft zeven faculteiten in

⁷ Mark D. Wilkinson e.a. *The FAIR guiding principles for scientific data management and stewardship* (2016) zie <https://www.nature.com/articles/sdata201618> geraadpleegd 20180927

⁸ *Digitaal Erfgoed Nederland Digitaal Erfgoed Referentie Architectuur* (2016) zie <http://www.den.nl/standaard/412/Digitaal-Erfgoed-Referentie-Architectuur> geraadpleegd 20180927

het alfa-, bèta- en gammadomein, is gevestigd in Leiden en Den Haag en heeft ruim 6.700 medewerkers en 28.000 studenten. In april 2016 heeft het College van Bestuur de Regeling Datamanagement Universiteit Leiden vastgesteld.⁹ Deze regeling is bedoeld als universitair raamwerk voor een datamanagementbeleid. Onder onderzoeksdata worden alle gegevens die worden verzameld en gegenereerd tijdens het wetenschappelijk onderzoek verstaan, zowel digitale als niet-digitale onderzoeksdata. De regeling beschrijft aan welke eisen onderzoeksdata moeten voldoen voorafgaand, tijdens en na afloop van het onderzoek. Eveneens wordt er aangegeven wie voor welke handelingen verantwoordelijk is. De voornaamste uitgangspunten zijn dat: vóór de start van de dataverzameling in een onderzoeksproject een datamanagementplan (DMP) wordt opgesteld; gedurende het onderzoek de onderzoeksdata veilig wordt opgeslagen, hetgeen betekent dat de integriteit, beschikbaarheid en -indien vereist- vertrouwelijkheid van de data gegarandeerd moeten blijven; na afloop van het onderzoek moet de onderzoeksdata op een zodanige manier worden beheerd dat zij ten minste vindbaar, toegankelijk, begrijpelijk, herbruikbaar (FAIR) en duurzaam (CoreTrustSeal) opgeslagen zijn. De minimale bewaartermijn van onderzoeksdata is tien jaar.

De regeling is een raamwerk en vergt nadere uitwerking op disciplineniveau. De faculteiten bepalen hoe de regeling binnen de faculteit nader uitgewerkt wordt. Aangezien de implementatie van het beleid dermate veel impact heeft op de kernprocessen van onderzoek is er een programma implementatie datamanagement opgezet om de faculteiten te ondersteunen bij het uitvoeren van de regeling. Het doel van het programma is het implementeren van het beleid. Daarvan afgeleide doelen zijn:

- A. Door coördinatie ervoor zorgen dat de universiteit als geheel efficiënt kan voldoen aan de Regeling Datamanagement;
- B. Door het bieden van voorzieningen gedurende de gehele data-levenscyclus faculteiten en instituten ondersteunen bij de implementatie van datamanagement;
- C. Het beheersbaar houden van het totale aanbod aan voorzieningen door het ontwikkelen en toepassen van een domeinarchitectuur datamanagement;
- D. (Inter)nationaal samenwerken om deze voorzieningen zo efficiënt mogelijk aan te bieden en aan te sluiten bij nieuwe ontwikkelingen;
- E. Met behulp van training en ondersteuning, generiek waar mogelijk en discipline specifiek waar nodig, wetenschappers in staat stellen om hun data op een goede manier te beheren;
- F. Bijdragen aan de bewustwording van het nut en de noodzaak van datamanagement;
- G. Het zichtbaar en meetbaar maken van de wetenschappelijke output aanvullend aan publicaties;
- H. Inzicht geven in de kosten van datamanagement en mogelijke kostenmodellen;
- I. Monitoren, evalueren en mogelijk bijstellen van het beleid.

Het programma wordt uitgevoerd door het ICT Shared Service Center, de afdeling informatiemanagement en de Universitaire Bibliotheken Leiden. Opdrachtgever van het programma is de rector van de Universiteit Leiden.

⁹ *Universiteit Leiden Research Data Management Regulations University Leiden* (2016) zie https://www.bibliotheek.universiteitleiden.nl/binaries/content/assets/ul2ub/research--publish/research-data-management-regulations-leiden-university_def.pdf geraadpleegd 20180927

Grofweg kan er een parallel getrokken worden met de wijze waarop de Universiteit Leiden en het erfgoed in Nederland georganiseerd is. Beiden zijn in sterke mate gefragmenteerd en decentralistisch van aard en hebben de intentie om (academisch) erfgoed voor een zo breed mogelijk publiek beschikbaar te stellen. Het toegankelijk maken van (academisch) erfgoed vergt echter coordinatie op centraal niveau, samenwerking tussen verschillende partijen en veelal nieuwe middelen (financiering, diensten en ondersteuning) om dit te operationaliseren.

Werkprocessen en protocollen

Bij veel onderzoeksorganisaties zijn datamanagementparagrafen en -plannen het instrument om te beschrijven hoe met onderzoeksdata voorafgaand, tijdens en na afloop van het onderzoek wordt omgegaan. Een groot aantal Nederlandse universiteiten heeft templates (sjablonen) ontwikkeld voor datamanagementplannen (DMPs) voor onderzoek. Daarnaast zijn er ook templates gemaakt door data archieven en financiers, bijv. EC/H2020, NWO en ZonMW.¹⁰

Een DMP beschrijft vaak op een generieke manier aan welke richtlijnen onderzoeksdata moet voldoen. De volgende stap die gemaakt wordt, is om de richtlijnen verder te preciseren zodat ze beter passen bij een bepaalde onderzoeksdiscipline. In het kader van de AVG wet en regelgeving moet onderzoeksdata met persoonsgegevens indien nodig een Privacy Impact Assessment ondergaan. Om wetenschappers meer handvatten te geven voor het invullen van de richtlijnen wordt internationaal gewerkt aan het opstellen van Domein specifieke Data Protocollen (DDPs).¹¹ Door het opstellen van DMPs wordt er al voorafgaand aan de data collectie van onderzoeksdata aandacht besteed aan de lange termijn bewaring: Archiving by Design.

Datamanagementplannen en domein specifieke data protocollen bij de Universiteit Leiden

Datamanagementplannen (DMPs) vervullen een centrale rol bij de Universiteit Leiden. Voorafgaand aan het onderzoek wordt in het DMP het datamanagementbeleid van de faculteit of het instituut waar de wetenschapper werkzaam is specifiek uitgewerkt voor het desbetreffende onderzoeksproject. Bij veel onderzoeksprojecten zijn meerdere wetenschappers betrokken, bijvoorbeeld promovendi, postdocs en leden van de vaste staf. In het DMP wordt vastgelegd wie welke verantwoordelijkheid draagt. Speciale aandacht verdienen projecten waarbij wetenschappers van meerdere instituten en/of externe partijen betrokken zijn. Het is belangrijk om ook vast te leggen wat er met de data en de verantwoordelijkheid voor de data gebeurt als één van de betrokken wetenschappers vertrekt. De universiteit heeft een template ontwikkeld die gebruikt kan worden voor het invullen van een DMP¹². De template bestaat uit vijf verschillende onderdelen: data collectie, data opslag en veiligheid, data documentatie, toegang, delen en hergebruik van data en data preservatie en archivering.

¹⁰ Digital Curation Centre *DMPonline* zie https://dmponline.dcc.ac.uk/public_templates geraadpleegd 20180927

¹¹ Science Europe *Presenting a Framework for Discipline-specific Research Data Management* (2018) zie https://www.scienceurope.org/wp-content/uploads/2018/01/SE_Guidance_Document_RDMPs.pdf geraadpleegd 20180927

Het DMP wordt centraal binnen de universiteit opgeslagen voor een periode van minimaal 20 jaar. Na het vernietigen van het DMP kan de wetenschapper of het instituut niet meer aantonen hoe bepaalde beslissingen met betrekking tot het vernietigen van data tot stand zijn gekomen. De bewaartermijn van een DMP is daarom langer dan die van de data (minimaal tien jaar), omdat dan in voorkomende gevallen met behulp van het DMP kan worden aangetoond dat het vernietigen van een bepaalde dataset in lijn is met het DMP.

Het schrijven van een DMP is voor een wetenschapper vaak geen sinecure. De universiteit biedt wetenschappers daarom ondersteuning in de vorm van workshops 'How to write a Data Management Plan' en adviseert wetenschappers bij het opstellen van DMPs. In de workshops worden de voorschriften van de universiteit en subsidieverstrekkers behandeld. Er is ondersteuning bij het beantwoorden van een aantal vragen, die in een datamanagementplan worden gesteld: wat voor soort data worden er verzameld, waar worden ze opgeslagen, met wie en op welke manier worden zij gedeeld, hoe kunnen anderen de data begrijpen en is de lange termijn archivering goed geregeld?

Om wetenschappers verder richting te geven bij het invullen van een DMP wordt er op facultair of instituutniveau gewerkt aan datamanagementprotocollen. Bij de Faculteit Sociale wetenschappen zijn alle stappen in kaart gebracht die leiden van een onderzoeksvorstel tot de uiteindelijke resultaten. Workshops en interviews maakten duidelijk dat er grote verschillen zijn in het verzamelen van data, aard en vorm van de data en het gebruik van persoonsgebonden data. Op basis van deze bevindingen werken alle instituten van de faculteit een eigen protocol uit. Het Instituut voor Psychologie heeft bijvoorbeeld een opslagprotocol opgesteld dat eist dat alle data van gepubliceerde studies onmiddellijk na afronding opgeslagen moeten worden in een centraal archief. Het protocol moet op deze wijze borgen dat alle onderliggende data van een publicatie voor verificatie en replicatie toegankelijk is. Deze data is in niet altijd Open Access beschikbaar maar kan indien gewenst opgevraagd worden conform de vooraf gestelde voorwaarden.

De faculteit der Wiskunde en Natuurwetenschappen heeft een protocol opgesteld voor promovendi. Het protocol bepaalt dat promovendi een dataformulier invullen, niet later dan drie weken voor de verdediging van hun proefschrift. In dit formulier geven ze aan hoe de data, en indien relevant onderzoekssoftware, beschikbaar wordt gesteld. Het formulier moet ondertekend worden door hun promotor en de wetenschappelijke directeur van het instituut.¹³

Erfgoedinstellingen hebben evenals instituten bij universiteiten hun eigen werkwijze en vakspecifieke aanpak. (Inter)nationaal beleid geeft vaak de grote lijn aan van wat men wil bereiken en de algemene principes waaraan moet worden voldaan. Erfgoedinstellingen moeten deze principes vertalen naar verantwoordelijkheden en werkprocessen binnen de eigen organisatie. Dit vergt vaak maatwerk en dient goed aan te sluiten met de wijze waarop werkprocessen in een organisatie zijn ingeregeld. Lange termijn beschikbaarstelling moet daarom een integraal onderdeel zijn van de werkprocessen, bijdragen aan de effectiviteit en efficiency ervan en niet gezien worden als iets wat je na afloop nog moet regelen: Archiving by design.

¹² Universiteit Leiden *Data Management Template* (2018) zie <http://blogs.library.leiden.edu/researchdata/about-this-blog/data-management-plan/>

¹³ Fieke Schoots e.a. *Implementing a Research Data Policy at Leiden University* (2017) zie <https://doi.org/10.2218/ijdc.v12i2.575> geraadpleegd 20180927

Datamanagementvoorzieningen

Onderzoeksorganisaties zijn verantwoordelijk voor het leveren van voorzieningen voor datamanagement zodat wetenschappers in staat zijn te voldoen aan de richtlijnen die gesteld worden voorafgaand, tijdens en na afloop van het onderzoek. In de Nederlandse gedragscode wetenschappelijke integriteit¹⁴ staat dat onderzoeksorganisaties de zorgplicht hebben voor een onderzoeksinfrastructuur waarin goed databeheer de regel is. Onderzoeksorganisaties verschillen echter in de wijze waarop hier invulling aan gegeven wordt. Er kan gekozen worden om alle voorzieningen lokaal aan te bieden, gebruik te maken van externe partijen of een combinatie van lokale diensten en externe diensten. Nationaal en internationaal wordt er gewerkt aan referentie- en domeinarchitecturen en dienstencatalogi. Met behulp van deze instrumenten kunnen wetenschappers kijken welke diensten ze het beste kunnen gebruiken, krijgen onderzoeksorganisaties inzicht in de diensten die ze al aanbieden en waar de witte vlekken zitten en kunnen leveranciers van diensten vaststellen of hun diensten voldoen aan de gestelde eisen. Op basis van de in Leiden ontwikkelde Research Data Information Sheets¹⁵ is door het Landelijk Coördinatiepunt Research Data Management (LCRDM) een eerste proof of concept van een landelijke catalogus van diensten opgezet waarin per onderzoeksfase (voor/tijdens/na) een overzicht wordt geboden van de top-10 van meest gebruikte producten, diensten en faciliteiten in Nederland ter ondersteuning van goed RDM¹⁶.

De sectoren erfgoed en wetenschap delen in toenemende mate niet alleen kennis maar ook voorzieningen met elkaar. De voorzieningenvinder Digitale Duurzaamheid¹⁷ van de Nationale Coalitie Digitale Duurzaamheid (NCDD) bouwt voort op de catalogus die ontwikkeld is in LCRDM verband en geeft een overzicht van diensten die mogelijk door beide sectoren gebruikt kunnen worden.

Datamanagementvoorzieningen bij de Universiteit Leiden

De universiteit heeft als ambitie om haar wetenschappers 'state of the art' voorzieningen te bieden voor het omgaan met onderzoeksdata. Het betreft hier voorzieningen die nodig zijn voorafgaand, tijdens en na afloop van het onderzoek. Uitgangspunt is dat de voorzieningen goed aansluiten bij de behoeftes en de wensen van de wetenschappers en voldoen aan de gestelde richtlijnen met betrekking tot veiligheid, beschikbaarheid en toegankelijkheid. Aan voorzieningen die voorafgaand en tijdens het onderzoek beschikbaar worden gesteld, worden andere eisen gesteld dan aan voorzieningen die na afloop van het onderzoek gebruikt worden om data voor de langere termijn toegankelijk te houden. De periode na afloop van het onderzoek wordt in twee fasen opgedeeld. Voorzieningen die gebruikt worden om data toegankelijk te maken voor een periode van ten hoogste tien jaar moeten voldoen aan eisen die gesteld worden met betrekking tot archivering. Voor de lange termijn bewaring worden naast de archiveringseisen ook eisen gesteld met betrekking tot digitale duurzaamheid. Lange termijn voorzieningen moeten minimaal voldoen aan de richtlijnen van het CoreTrustSeal.

De Leidse wetenschappers mogen hun onderzoeksdata voor de lange termijn dus in de voorzieningen toegankelijk maken die het beste bij hun discipline past. De enige voorwaarde die gesteld wordt is dat de databewaarplaats voldoet aan internationale richtlijnen zoals het CoreTrustSeal. De universiteit wil wel een goed overzicht bewaren van de door haar wetenschappers gecreëerde data. Gepubliceerde

onderzoeksdata moet daarom centraal geregistreerd worden, met behulp van de Persistent Identifier die verwijst naar de gepubliceerde onderzoeksdata. Een Persistent Identifier (PID) is een unieke identificatiecode van een digitaal object die op een afgesproken plaats wordt geregistreerd. Hij blijft gegarandeerd werken, ook al verandert het webadres van een organisatie. De sectoren erfgoed en wetenschap maken voor een deel gebruik van dezelfde PID-systemen. Ze hebben gezamenlijk een PID-wijzer ontwikkeld voor de systemen die in Nederland het meest worden gebruikt: DataCite DOI (via de Technische Universiteit Delft), het Handle System (via SURFsara) en URN:NBN (via de Koninklijke Bibliotheek).¹⁸ Leidse wetenschappers zijn eveneens verplicht om gebruik te maken van de onderzoekersidentificer ORCID.¹⁹ Op deze wijze is ook altijd te achterhalen welke wetenschappers de data gecreëerd hebben ook als zij niet meer verbonden zijn aan de Universiteit Leiden.

Vaardigheden en competenties

Wetenschappers hebben ondersteuning en training nodig bij het opstellen van een DMP, het uitwerken van protocollen, het vinden van de meest geschikte voorzieningen, het gebruik van nieuwe methoden en technieken en het FAIR maken van hun data. Er zijn data stewards en data engineers nodig die de wetenschappers trainen en hands-on ondersteuning bieden.²⁰ De meeste onderzoeksorganisaties ontwikkelen programma's waarin deze ondersteuning aangeboden wordt. Het aanleren van zogenaamde 'transferable skills' en '21st century skills' zijn van belang om wetenschappers de vaardigheden en competenties bij te brengen die nodig zijn voor 'Open Science'.²¹ Deze vaardigheden zijn niet alleen van belang in de wetenschap maar zijn ook van belang in andere sectoren zoals erfgoed, musea, bibliotheken en het bedrijfsleven.

De Leidse universiteit wil het de wetenschappers door middel van state of the art ICT-ondersteuning en -infrastructuur gemakkelijker maken om te delen en samen te werken. Evenmin als bij de didactische innovaties is het mogelijk te voorspellen in welke richting de technologische ontwikkelingen precies gaan. Het is vooral zaak om zicht te houden op de nieuwste ontwikkelingen, en innovaties zo snel mogelijk toegankelijk te maken voor wetenschappers.

De Universiteit Leiden wil de impact van haar onderzoek en onderwijs versterken door wetenschappers te ondersteunen in de transitie naar een meer interactieve

¹⁴ VSNU *Nederlandse Gedragscode Wetenschappelijke Integriteit* (2018) zie <http://www.vsnul.nl/files/documenten/Nederlandse%20gedragscode%20wetenschappelijke%20integriteit%202018.pdf> geraadpleegd 20180927

¹⁵ Peter Verhaar e.a. *Fostering effective data management practices at Leiden University* (2017) zie <http://doi.org/10.18352/lq.10185> geraadpleegd 20180927

¹⁶ Landelijk Coördinatiepunt Research Data Management *Dienstencatalogus* zie https://www.edugroepen.nl/sites/RDM_platform/Dienstencatalogus/SitePages/Dienstencatalogus.aspx geraadpleegd 20180927

¹⁷ Nationale Coalitie Digitale Duurzaamheid *Voorzieningen vinder* zie <http://voorzieningen.ncdd.nl/achtergrond.php> geraadpleegd 20180927

¹⁸ Netwerk Digitaal Erfgoed *Persistent Identifiers* zie <https://www.netwerkdigitaal erfgoed.nl/kennis-en-voorzieningen/digitaal-erfgoed-houdbaar/persistent-identifiers/> geraadpleegd 20180927

¹⁹ ORCID *Connecting Research and Researchers* zie <https://orcid.org/> geraadpleegd 20180927

²⁰ European Commission *Realising the European Open Science Cloud* (2016) zie https://ec.europa.eu/research/open-science/pdf/realising_the_european_open_science_cloud_2016.pdf#view=fit&pagemode=none geraadpleegd 20180927

²¹ European Commission *Providing researchers with the skills and the competencies they need to practice Open Science* (2017) zie <https://doi.org/10.2777/121253> geraadpleegd 20180927

wetenschap en heeft daarom bij de Universitaire Bibliotheken Leiden in 2016 een Centre for Digital Scholarship ingericht.²² Het Centre for Digital Scholarship²³ werkt samen met wetenschappers, faculteiten, nationale en internationale collega's en expertisecentra om Open Access, datamanagement en het gebruik van digitale data te faciliteren en te ondersteunen.

Het Centre for Digital Scholarship organiseert bijeenkomsten en workshops en is dé plaats waar wetenschappers terecht kunnen voor vragen, advies en training. Voor de erfgoedsector is het eveneens van belang om aandacht te besteden aan het stimuleren van het gebruik van het beschikbare digitale erfgoed. Het publiek kan een belangrijke rol spelen in het verrijken en annoteren van erfgoed. Hiervoor moeten ze dan wel beschikken over de nodige vaardigheden. Investerings- en voorlichtingsmateriaal en trainingen kan hierbij instrumenteel zijn.

Lessons learned

Een aantal belangrijke lessen die getrokken kunnen worden, en die mogelijk relevant kunnen zijn voor de erfgoed sector, zijn:

Voor de implementatie van beleid is commitment van het management op het hoogste niveau van een organisatie onontbeerlijk; De betrokkenheid van alle stakeholders is cruciaal (met speciale aandacht voor de pioniers die op de werkvloer activiteiten ontplooiën); (Inter)nationale samenwerking binnen en buiten de eigen sector is noodzakelijk (delen van kennis, best practices, gemeenschappelijke voorzieningen en trainingprogramma's); regulering (beleid) moet door de gebruikers ervaren worden als een middel om de kwaliteit van hun werkprocessen te verbeteren en niet gezien worden als een extra (administratieve) last; Hoewel er grote verschillen zijn per organisatie met betrekking tot de data die gecreëerd of verzameld wordt (formaten, typen, volume etc.) zijn veel van de uitdagingen hetzelfde (betaalbare opslag, de omgang met persoonsgebonden data, embargo's voor het delen van data, licenties voor hergebruik van de data); Door middel van een top-down-bottom-up aanpak moet rekening gehouden worden met verschillen in werkprocessen, methoden en technieken; De behoefte aan voorzieningen neemt toe naarmate de bewustwording van beleid groter wordt; Het gelijktijdig ontwikkelen en implementeren van een groot aantal voorzieningen vormt een enorme uitdaging en vergt een meerjarig implementatie programma; De ervaringen uit de praktijk kunnen ertoe leiden dat het beleid deels moet worden aangepast; Er moeten middelen vrijgemaakt worden voor nieuwe functies zoals data stewards en data scientists die procesondersteuning bieden bij het duurzaam toegankelijk maken van data; Omdat veel van de noodzakelijke diensten ontwikkeld moeten worden terwijl er nog ervaring opgedaan moet worden met best practices moet er naast een waterval projectmethodiek eveneens een agile projectmethodiek gebruikt worden voor het opzetten en uitrollen van nieuwe diensten.

²² Universiteit Leiden *Instellingsplan 2015-2020: Excelleren in vrijheid* (2015) zie <https://www.universiteitleiden.nl/over-ons/profiel/instellingsplan-2015-2020> geraadpleegd 20180927

²³ Universitaire Bibliotheken Leiden *Centre for Digital Scholarship* zie <https://www.bibliotheek.universiteitleiden.nl/onderzoek-en-publiceren/centre-for-digital-scholarship> geraadpleegd 20180927

Een audiovisueel archief voor Vlaanderen

Waar staat VIAA na vijf jaar?

Proloog: wat is de bedoeling? Wat ging vooraf?

Vlaanderen heeft een enorme schat aan audiovisueel materiaal, die tot voor kort voor een groot deel op analoge dragers bewaard werd. Omdat deze dragers bedreigd zijn en gedigitaliseerd moeten worden, werd eind 2012 het Vlaams Instituut voor Archivering (VIAA) opgericht. VIAA begon zijn activiteiten in 2013 en vierde recent zijn vijfjarig bestaan. Begin 2013 bevond VIAA zich in de unieke positie om een grootschalige dienstverlening op te bouwen van de grond af. In dit artikel focussen we op het traject dat VIAA in die vijf jaar doorlopen heeft. Hoe werd de dienstverlening uitgebouwd? Welke beslissingen werden genomen? Welke impact hebben deze beslissingen op de duurzaamheid van de bewaring? We focussen hierbij op de schaal en de specifieke context van audiovisueel archiefmateriaal.

Belangrijk hierbij: VIAA beheert zelf geen collectie, maar levert diensten aan een groep klanten, die *content partners* (CP's) genoemd worden. Deze content partners zijn verspreid over verschillende sectoren, waaronder cultuur en mediabedrijven, erfgoedinstellingen, archieven, organisaties uit de podiumkunstensector en overheden. De groep werd in de voorbije jaren geleidelijk uitgebreid. Voor elk van deze partners digitaliseren we analoge dragers en bewaren we digitaal materiaal. In ruil hiervoor krijgt VIAA licenties voor hergebruik in specifieke contexten zoals onderwijs, onderzoek en via bibliotheken. Selectie, eigenaarschap en het beheer van de collecties blijft echter de taak van de content partner zelf.

Schaal is een sleutelbegrip in de werking van VIAA. Door kennis en investeringen te centraliseren werken we complementair en kunnen we meer partners op een goedkopere en kwalitatieve manier helpen. Van de initiële 40 partners in 2013 is het klantenbestand ondertussen gegroeid tot 140 organisaties.

Het volume werk is aanzienlijk: een inventaris uit 2014 toonde aan dat 650.000 dragers gedigitaliseerd moesten worden. Recente bevragingen omtrent bestaand digitaal materiaal wijzen uit dat ook hier een belangrijke nood is aan duurzame bewaring van grote collecties audio-, video- en fotomateriaal.

VIAA komt evenwel niet uit het niets. Jaren voor de eigenlijke oprichting van VIAA werd de basis gelegd in onderzoeksprojecten zoals BOM-vlaanderen, Archipel en Vlaanderen in Beeld. Tijdens deze projecten werd de blauwdruk getekend van

zowel de technische oplossingen als het organisationele model van het latere VIAA. De projecten hebben ook geholpen om het nodige draagvlak te creëren bij deelnemende organisaties en beleidsmakers voor een brede aanpak van de problematiek. Daarnaast kon VIAA bogen op de ruime ervaring binnen en buiten Vlaanderen, bestaande standaarden en frameworks zoals OAIS. De dienstverlening werd (en wordt) zoveel mogelijk in lijn met deze standaarden ontwikkeld.

Op dit moment archiveert VIAA meer dan 10 petabytes aan data, goed voor ruim 3 miljoen objecten. Daaronder zitten gedigitaliseerde kranten, audiovisueel materiaal en digitale foto's of scans.

Duidelijke afspraken: samenwerkingsovereenkomsten en service agreements

Aangezien VIAA opgezet werd als een dienstverlenende organisatie zijn afspraken tussen VIAA en de CP's van cruciaal belang. In OAIS vinden we een aantal aanbevelingen die betrekking hebben op de relaties tussen de producers en het archief, zoals het verkrijgen van een mandaat om bepaalde acties uit te voeren en het bepalen van formaten waarin data aangeleverd wordt. Deze en andere afspraken worden bij VIAA vastgelegd in twee documenten: een samenwerkingsovereenkomst en een service agreement. Beide documenten zijn contractueel bindend.

In de samenwerkingsovereenkomst wordt ondermeer bepaald dat VIAA het materiaal zal en mag digitaliseren, dat VIAA zal instaan voor duurzame bewaring, dat VIAA zal monitoren tegen *obsolescence* en wat er dient te gebeuren indien de overeenkomst beëindigd wordt. Verder wordt ook aangegeven dat de content partner het materiaal moet voorzien van metadata en ontvangt VIAA een licentie om ontsluiting naar specifieke doelgroepen mogelijk te maken. In de service agreement worden verdere praktische en technische afspraken gedocumenteerd zoals de aan te leveren collectie, de manier waarop de metadata aangeleverd kan worden, maar ook het importproces dat gebruikt zal worden wanneer het materiaal in het archief opgeslagen wordt. Beide documenten samen creëren een kader voor het uitvoeren van bepaalde diensten, maar zorgen ook voor transparantie over de gebruikte processen.

Het feit dat de documenten contractueel vastgelegd en bindend zijn, heeft voor- en nadelen. Enerzijds creëert het een noodzakelijk juridisch kader voor de overeenkomsten, maar anderzijds mag de doorlooptijd van de ondertekening hier niet van onderschat worden. Sommige organisaties hebben hiervoor vrij complexe procedures die gedurende maanden of zelfs langer dan een jaar kunnen lopen. Op die manier wordt een update van een afspraak moeilijker dan wanneer er geen handtekening vereist is: de hele formele procedure moet opnieuw doorlopen worden. Het is dan ook belangrijk om op voorhand te bepalen welke zaken opgenomen worden in een formeel, statisch gedeelte en welke bepalingen bijvoorbeeld in addenda of aanvullende documenten nader omschreven kunnen worden.

VIAA heeft over alle content partners heen zo goed als uniforme afspraken omtrent de import, duurzame bewaring en het beheer van het materiaal. Dit zorgt ervoor dat praktische oplossingen uitgerold kunnen worden over de hele breedte van het klantenbestand wat cruciaal is voor de kostenefficiëntie en schaalbaarheid van

de dienstverlening. Zoals al vermeld: dit soort afspraken organiseren over meer dan honderd content partners is een hele klus. Om het overzicht te behouden werd een platform gebouwd dat dit proces ondersteunt: de afspraken worden – transparant voor de content partner en VIAA – beheerd en bewaard op een gedeelde omgeving. Zo kan elke betrokken partij te allen tijde de meest recente versie van de afspraken consulteren. Tooling, zoals documentmanagementoplossingen of procesondersteunende tools, zijn in deze zeker een dankbaar instrument om het overzicht te bewaren tijdens langlopende processen of wanneer verschillende versies van een document tijdens de onderhandeling ontstaan.

Digitaliseren en archiveren: Uitbouw basisdienstverlening

Digitalisering

VIAA organiseert massa-digitaliseringsprojecten. We lichten een tweetal aspecten toe die cruciaal zijn met het oog op preservatie: de keuze van het formaat en de opslag van PREMIS-informatie *tijdens* het digitaliseringsproces.

Formaten

Het formaat waarin de archiefmasters opgeslagen worden, is een belangrijk element voor de duurzaamheid van het bewaarde materiaal. Bij audiovisuele dragers is dat niet minder het geval. Hier moet aandacht besteed worden aan een aantal elementen: de *wrapper* of container waarin het audiovisueel bestand zich bevindt (zoals MKV, MXF, MOV of AVI) en de encoding van audio (bijvoorbeeld LPCM, FLAC) of video (bijvoorbeeld AVC-I, H.264) zelf. De combinaties en mogelijkheden zijn legio. Net als bij andere types materiaal genieten codecs met lossless compressie de voorkeur. Dit betekent namelijk dat er geen inhoudelijke informatie weggegooid wordt ten voordele van een kleiner bestandsformaat en dat de enige compressie die men uitvoert mathematisch is en steeds omkeerbaar. Tegelijkertijd speelt ook de ondersteuning een belangrijke rol: de openheid van de codec en wrapper, het bestaan van documentatie of standaarden en de mate waarin ondersteunende software voorhanden is, zijn parameters die mee in acht genomen werden.

In overleg met de content partners en experts uit de sector koos VIAA hier voor een aantal archiefformaten: materiaal van de omroepen wordt gedigitaliseerd naar IMXD-10 in een MXF container. Voor de cultuur- en erfgoed-sectoren werd gekozen voor lossless JPEG2000 in MXF. Audio wordt opgeslagen als LPCM in WAV containers. Film wordt dan weer gedigitaliseerd naar DPX. Merk op dat encoding in IMXD-10 geen lossless compressie is: de sector gaf hier voorkeur aan brede ondersteuning door hardware en software en focust op snel en makkelijk hergebruik. In de cultuur- en erfgoedsectoren werd voor een lossless formaat gekozen. De uitdaging hier wordt echter het gebruik van de bestanden. De codec is dermate specifiek dat die met een correcte configuratie in software afgespeeld moet worden. De bestanden stellen bovendien hoge eisen aan de hardware waarmee ze afgespeeld worden. Met een standaard laptop of pc is dit geen evidentie. De creatie van een mezzaninekopie (een gecomprimeerde versie van de archiefmaster aan voldoende hoge resolutie in een formaat dat toch vlot kan gebruikt worden in editeeromgevingen) zou dit kunnen oplossen. Omwille van de impact op opslag doet VIAA dit echter niet systematisch, maar werd na verloop van tijd een on-demanddienst ontwikkeld om een mezzaninekopie te verkrijgen.

Premis:event or it didn't happen

Digitalisering van audiovisueel materiaal is een belangrijke migratie en een ingrijpende transformatie van het bronmateriaal, waar authenticiteit zo goed mogelijk gewaarborgd dient te worden. Digitaliseringsprojecten, en zeker massadigitaliseringsprojecten, volgen daarom best een vastgelegd proces waar vlot op gerapporteerd moet kunnen worden. Bij VIAA loopt dit in grote lijnen als volgt: analoge dragers worden geregistreerd en krijgen een unieke *identifler*, vervolgens worden ze getransporteerd naar een digitaliseringsfirma die een aantal stappen uitvoert zoals inspectie, cleaning van de tape, eventueel voorbereiding en uiteindelijk digitalisering. Aan het eind van het proces worden digitale dragers getransporteerd naar de archiefinfrastructuur. De toepassing waarin de analoge dragers geregistreerd worden is een cruciale schakel in de rapportering op dit proces: een item bestaat omdat het in deze toepassing geregistreerd is. Diezelfde software wordt ook aan het eind van de keten gebruikt om te verifiëren of een bepaald item correct gearchiveerd werd. Ze vormt de bron van waarheid voor onze rapportering.

Om de rapportering fijnmazig te maken en te standaardiseren, worden alle stappen tijdens de digitalisering opgeslagen als PREMIS-metadata. De PREMIS-standaard werd ontwikkeld door onder meer de Library of Congress en bouwt verder op het OAIS-referentiemodel. In essentie stellen deze metadata ons in staat om te weten welke actie op welk moment door welke operator (of applicatie) uitgevoerd werd en wat de uitkomst van deze actie was. De data worden opgeslagen als *events*. Aan de hand van deze events komen we te weten welke werknemer een drager registreerde voor digitalisering of op welk apparaat een bepaalde tape gedigitaliseerd werd. Dit soort metadata wordt praktisch gebruikt bij de controle van de processen en oplevering van de leveranciers, maar is ook onmisbaar om digitale bestanden te kunnen herleiden naar hun oorspronkelijke analoge drager. Ze documenteert de levenscyclus van het audiovisueel materiaal.

De metadata wordt uiteindelijk samen met de digitale video of audiobestanden verpakt als een SIP-pakket ('Submission Information Package') en aangeleverd aan het archief voor verdere verwerking. De aangemaakte identifier tijdens registratie van het analoge materiaal wordt hier opnieuw gebruikt als unieke identificatie van één SIP. In de meeste gevallen gaat dit om audio of video met een metadatabestand in XML. In complexere cases (bijvoorbeeld bij de digitalisering van kranten of film) kan een SIP meerdere scans, audio- of videobestanden bevatten en wordt alles samengehouden aan de hand van metadata in METS. In ons geval wordt het SIP-pakket aangemaakt door de digitaliseringspartner. De technische complexiteit van een dergelijk SIP-formaat, of meer specifiek van METS en PREMIS, mag zeker niet onderschat worden. Digitaliseringsfirma's hebben deze kennis niet altijd aan boord en er moet dus voldoende tijd voorzien worden zodat het hele SIP-creatieproces goed getest kan worden. Praktische voorbeelden van hoe zo'n pakket er precies uit moet zien, bleken in deze instrumenteel. Een goede definitie, documentatie, voorbeelden, testprocedures en planning zorgen ervoor dat een grootschalig digitaliseringsproject vlot kan werken.

Digitaal archiveren

Import workflows

Na digitalisering worden de SIP-pakketten geleverd aan het digitaal archief. Ook hier werd een proces uitgewerkt voor de import van het materiaal, waarbij alle afzonderlijke stappen opnieuw als PREMIS-metadata opgeslagen worden in het archief. In tegenstelling tot het proces bij digitalisering verloopt dit proces helemaal automatisch. Tijdens dit proces worden een aantal checks uitgevoerd: we controleren de integriteit van de bestanden, gaan na of een aangeleverd SIP-pakket wel degelijk verwacht werd in deze aanlevering, controleren de wrapper en codec van het bronbestand en maken bijkomende lage resolutie kopieën aan. Als alles goed loopt, worden metadata en essence naar meerdere opslaglocaties weggeschreven.

De aanlevering van bestanden gebeurt aan een hoog tempo: één levering kan typisch enkele tientallen terabyte aan data bevatten en is snel goed voor meer dan 1000 SIP-pakketten. Batches dienen op enkele dagen verwerkt te worden. Op piekmomenten kan tot 20 TB aan data verwerkt worden per dag. Om over het welslagen van de import te waken wordt de PREMIS-metadata via een zoekinterface en rapporteringstools doorzoekbaar gemaakt. Aan de hand hiervan kunnen operatoren snel inzicht krijgen in mogelijke fouten en bijvoorbeeld escaleren naar de digitaliseringsfirma's of softwareleveranciers.

Dit hele proces wordt geautomatiseerd aan de hand van workflows in het VIAA-archiefsysteem. Samen met de leverancier van het systeem hebben we hier in de loop van de voorbije jaren een hele weg afgelegd: in eerste instantie was de opvolging van de import en ondersteuning voor PREMIS een ontwikkeling op maat van VIAA. Gaandeweg is deze een vast onderdeel geworden van het systeem en worden nu ook alle acties na archivering (zoals export, verrijking van metadata) als PREMIS-events geregistreerd en opgeslagen als onderdeel van het kernpakket. Een nauwe en goede samenwerking met de softwareleverancier is hiervoor noodzakelijk, gezien in onze ervaring concepten zoals PREMIS relatief onbekend zijn (of waren) voor de meeste leveranciers van media-asset-managementoplossingen.

Een andere belangrijke les die we leerden, is niet in te grijpen middenin een dergelijke geautomatiseerde workflow, tenzij die daar specifiek op voorzien is. Het is verleidelijk om zelf bijvoorbeeld een kleine fout aan te passen in een SIP-pakket, maar naar onze ervaring leidt dit stevast tot meer werk dan voorzien. Daarnaast gaan dit soort aanpassingen ook voorbij aan het feit dat (in dit voorbeeld de digitaliseringsfirma) verantwoordelijk is voor de levering van een correcte SIP en zij dus een nieuwe levering moeten organiseren. Wanneer we zelf ingrijpen, wordt die grens vaag. Bij VIAA zijn de workflows overwegend zo georganiseerd dat we bij elke importfout van voraan herbeginnen.

Tot slot mag de tijd die nodig is voor operationele opvolging tijdens de import van grote hoeveelheden data niet onderschat worden. Wanneer dagelijks tientallen terabytes of duizenden items geïmporteerd worden, is het noodzakelijk om hier permanent operationele controle op te voorzien. Ook al worden veel aspecten geautomatiseerd, dit soort werk blijft een complexe keten van leveranciers, stukken software en mensen. Het is daarom nodig dat operatoren uitzoeken of een fout te wijten is aan een foutieve levering, een bug in de software of een hardware error.

Ook eindrapportering of overzichten na oplevering van digitaliseringsprojecten zijn absoluut noodzakelijk om te vermijden dat bepaalde bestanden bijvoorbeeld nooit gedigitaliseerd werden of wel gedigitaliseerd werden maar nooit correct gearchiveerd. De ervaring leert ons hier dat kort op de bal spelen en structurele rapportage voorzien, helpt om snel fouten te detecteren.

Kwaliteitscontrole

Kwaliteitscontrole was en is nog steeds een belangrijk aspect van digitalisering en archivering. Na digitalisering is hiervoor een garantieperiode voorzien. Wanneer de kwaliteit van het digitaal materiaal onvoldoende blijkt, kan VIAA dit kosteloos laten herdigitaliseren. Bij audiovisueel materiaal is kwaliteitscontrole een hele uitdaging: de formaten zijn soms erg complex, de bestanden zijn vrij groot en stellen belangrijke eisen aan de hardware, alleen al om ze te bekijken. Bovendien zijn artefacten in audiovisueel materiaal niet altijd makkelijk te herleiden tot fouten in het digitaliseringsproces, ze konden net zo goed al aanwezig zijn in de originele, analoge bron. Tot slot is het niet altijd makkelijk om de originele, analoge bron nog te bekijken ter vergelijking, door het ontbreken van de correcte afspeelapparatuur.

VIAA bouwde een heel aantal controles uit, zowel tijdens digitalisering als tijdens archivering. Bij de opzet van digitaliseringsprojecten wordt gewerkt met verschillende fases, waar grondig en manueel getest wordt voor een project data kan aanleveren in piloot- en productiefases. Fouten met betrekking tot codec, wrapper of formatering van de metadata in het SIP-pakket worden hier vaak opgevangen. Identieke controles lopen ook automatisch tijdens het archiveringsproces zelf.

De finale, inhoudelijke kwaliteitscontrole dient echter te gebeuren door de content partner zelf: alleen zij zijn in staat om in te schatten of de digitale kwaliteit daadwerkelijk overeenstemt met de verwachtingen. Hier knelt het schoentje soms: kwaliteitscontrole vereist niet alleen technische kennis, het kost ook bijzonder veel tijd. Daarbij komt dat een heel aantal content partners nog steeds druk in de weer zijn met de voorbereiding van nieuwe batches voor digitalisering, waardoor kwaliteitscontrole uitgesteld dreigt te worden. We blijven hier aandacht aan schenken door content partners maandelijks te informeren welk materiaal klaar staat voor kwaliteitscontrole en op welk materiaal de garantie na digitalisering dreigt te vervallen. Via use cases wordt geïllustreerd hoe kwaliteitscontrole verliep bij collega's en met workshops proberen we kennis over digitaal audiovisueel materiaal op te bouwen bij de content partners. Dit soort trajecten lopen in samenwerking met het bredere veld van betrokken partijen en expertisecentra zoals Packed vzw.

Aan de slag: beheer van digitale collecties

Het beheer van het gedigitaliseerde materiaal gebeurt aan de hand van het eerder vermelde archiefsysteem, een media-asset-management-systeem dat niet alleen de processen en metadata beheert maar ook instaat voor het beheer van de onderliggende opslaglocaties. Collectiemedewerkers van content partners krijgen toegang tot dit systeem en kunnen het materiaal verrijken met metadata, het exporteren, bekijken of zelfs verwijderen. Elke content partner van VIAA heeft in dit systeem een eigen omgeving waarin enkel de eigen content zichtbaar is. Zoals gezegd worden ook hier de acties op het materiaal opgeslagen als PREMIS-metadata.

VIAA heeft de voorbije jaren veel geïnvesteerd in de uitbouw van de API's (Application Programming Interfaces) van het archiefsysteem. Aan de hand van deze technische interfaces kan snel nieuwe functionaliteit gebouwd worden bovenop het archiefsysteem zonder dat deze noodzakelijk een vast onderdeel van het archiefsysteem moeten zijn. Zo kunnen applicaties geschreven worden die metadata toevoegen aan bestaande objecten, massametadata-updates uitgevoerd worden of kan materiaal via de API geïmporteerd of geëxporteerd worden. Hierbij blijft het archiefsysteem de centrale component en master van metadata en audiovisueel materiaal, alleen hoeft niet alles noodzakelijk via de user interface te lopen die hierop standaard voorzien is.

Waar enerzijds de technische onderbouw van het archiefsysteem solide moet zijn en de schaal aan moeten kunnen, is anderzijds ook de bruikbaarheid voor de eindgebruiker een aandachtspunt. Interfaces zijn vaak complex en beantwoorden niet aan de hedendaagse eisen van eenvoudig te gebruiken applicaties die we elke dag gebruiken. Bij de aankoop van een systeem dient hier dan ook voldoende aandacht aan besteed te worden. Een belangrijk aantal processen (zoals kwaliteitscontrole, exporteren en metadateren) lopen immers via deze interface en worden vaak uitgevoerd door mensen met vooral inhoudelijke kennis. Ook hier is een nauwe samenwerking met softwareleveranciers en het luisteren naar de eigen gebruikers cruciaal. Op die manier kan ook de interface ingericht worden in functie van de reële noden bij de eindgebruiker. Daarnaast zijn handleidingen en opleidingen instrumenten om de technische kennis op te bouwen bij diezelfde gebruikers. Op dit moment worden door VIAA maandelijks opleidingen georganiseerd en wordt veel aandacht besteed aan ondersteuning via een helpdesk en andere communicatiekanalen.

Toekomstmuziek en nieuwe uitdagingen

Import van digitale collecties

Tot nu toe hebben we in dit artikel vooral gesproken over materiaal uit digitaliseringsprojecten, maar in toenemende mate wordt ook digitaal geboren materiaal geïmporteerd in het archief. Dit brengt enkele specifieke uitdagingen met zich mee, enerzijds operationeel en procesmatig, anderzijds naar preservering toe.

Op het vlak van preservering is vooral de grote verscheidenheid aan digitale archiefformaten een belangrijke uitdaging. Waar de formaten uit digitaliseringstrajecten al uniform zijn, is dat bij digitaal geboren trajecten niet of nauwelijks het geval. Hier staan we als archief voor een keuze: normaliseren we bestanden waar geen goede archiefmaster voor bestaat bij import of doen we dit achteraf? Op dit moment kiest VIAA voor het tweede: we geven voorrang aan de duurzame (bit-level) bewaring van de bestaande formaten zodat de druk op bestaande lokale opslagsystemen snel verlicht kan worden. Tijdens import worden technische karakteristieken omtrent codec en wrapper opgeslagen zodat deze later kunnen gebruikt worden voor een eventuele transformatie naar een duurzamer bestandsformaat. Op die manier wordt de druk op bestaande opslagsystemen alvast verlicht en hoeven we dus maar één probleem (de import van het materiaal zelf) op te lossen.

De dienstverlening rond digitale intake is nu ruim een jaar beschikbaar, maar voor een heel aantal content partners blijkt de bestaande oplossing toch technisch te complex. Het aanleveren van SIPS in een vooraf gedefinieerd formaat is voor veel content partners geen sinecure. Zo blijkt het niet voor alle partners evident om data of metadata in een gestructureerd formaat uit bestaande systemen te exporteren, laat staan deze om te zetten naar een gestandaardiseerd SIP-formaat. In sommige gevallen is daarenboven eerst een datacleaning-stap nodig eer het materiaal gearchiveerd kan worden. Aan de hand van piloten en proefprojecten werd heel wat ervaring opgedaan die nu omgezet wordt in bijkomende gebruiksvriendelijke tools. Tegelijk wordt een projectaanpak voor het aanleveren van digitaal geboren materiaal verder uitgewerkt. We vullen dit aan met opleidingen en documentatie waarin een aantal best practices (aanmaak van checksums en toekennen van unieke ID's op bestaand materiaal, metadatamappings, uniformiseren van de bestaande structuren) uitgelegd wordt. Ons uiteindelijke doel is om ook deze dienstverlening op te schalen, zonder dat dit noodzakelijk leidt tot een evenredige uitbreiding van de benodigde mankracht.

Metadata

Op het vlak van metadata bevindt VIAA zich op dit moment op een keerpunt: de basisinfrastructuur (modellering en implementatie daarvan) bestaat, maar de focus lag tot op heden vooral op de import van materiaal uit digitaliseringsprojecten en digitaal geboren content. De komende jaren zal metadata één van de belangrijkste uitdagingen worden binnen het archief. Een groot aantal collecties hebben op dit moment geen of zeer beperkte metadata (ook in de systemen van content partners) en manuele annotering is duur en tijdrovend. Er zal ingezet moeten worden op de import van bestaande metadataseten en tooling om dit voor de content partners mogelijk te maken. Net als bij digitale intake is dit een belangrijke technische en operationele uitdaging.

Tegelijkertijd worden tools voor de generatie van metadata matuur en betaalbaar. Onderzoek naar de juiste oplossing voor specifieke collecties moet uitwijzen welke geschikt zijn voor onze use cases. Daarnaast zal ook ingezet worden op het gebruik van thesauri of andere, bestaande vocabulaires zoals AAT, GTAA of Wikidata. Aangezien de collecties centraal opgeslagen worden, is het gebruik van gemeenschappelijke metadata bronnen een krachtig hulpmiddel voor de uniforme ontsluiting. Deze trajecten worden verder uitgewerkt in de komende jaren.

Conclusie

Dankzij intensieve samenwerkingen met content partners, expertisecentra, leveranciers, nationale en internationale instellingen is VIAA er in geslaagd om op korte termijn een uitgebreide dienstverlening te realiseren. Standaarden, gestandaardiseerde formaten en een framework als OAIS zijn hierin cruciaal. Frameworks zorgden voor een gemeenschappelijke taal die de hele samenwerking kadert en ondersteunt. De adoptie van standaarden leidt in eerste instantie tot extra werk, maar biedt op langere termijn enkel voordelen. Ze maken uniforme oplossingen mogelijk, die los van technische componenten op grote schaal uitgerold kunnen worden en beheersbaar blijven op lange termijn.

TOWN PLANNERS

HOOFDSTUK 3

Voorbereiden op volgende stappen

Inleiding

Pionieren, ontdekken, settelen en plannen. Deze fasen zijn door velen doorlopen in de afgelopen jaren. Sommigen zijn nu pas in de fase van pionieren en sommigen zijn voorzichtig aan het settelen. Andere organisaties breiden al stevig uit of zijn toe aan hun derde generatie e-depot. In dit derde hoofdstuk wordt er nog kort teruggekeken naar de ontdekkingsreizigers en settelaars van het eerste uur. Ook de nieuwe pioniers worden belicht, want pioniers zullen er altijd nodig zijn in de wereld van digitale preservering.

Zoals de praktijkverhalen uit het tweede hoofdstuk van dit jaarboek hebben laten zien, is preserveren veel meer dan het oplossen van technische problemen. De organisatorische uitdagingen zijn minstens zo belangrijk als de technische, evenals de menselijke component. Omdat geen enkele erfgoedinstellingen de uitdagingen van de digitale samenleving in zijn eentje kan oplossen, is samenwerking noodzakelijk. De uitdagingen waar we voor staan zijn immers te groot en de oplossingen die we nodig hebben te kostbaar en te kennisintensief voor één enkele instelling. Samenwerking brengt echter weer vele nieuwe organisatorische uitdagingen met zich mee.

Om samenwerking op een nationale schaal effectief tot stand te laten komen zijn er in de afgelopen jaren zowel in Vlaanderen als in Nederland initiatieven gestart die hier vorm aan geven. Bijna tien jaar geleden deed de toen net opgerichte Nationale Coalitie Digitale Duurzaamheid (nu Netwerk Digitaal Erfgoed) een onderzoek naar de stand van zaken met betrekking tot digitale preservering in Nederland. De uitkomst was dat de grote instellingen een aardig eind op weg waren, maar dat er een groot gapend gat was bij de middelgrote en kleinere instellingen wat betreft kennis, ervaring en benodigde infrastructuur en financiële middelen. Een landelijke aanpak waarin pioniers, settelaars en planners meer met elkaar zouden optrekken was hard nodig. Dit is gebeurd. Nederlandse organisaties hebben de handen in elkaar gestoken en zijn steeds meer gaan samenwerken met een stip aan de horizon, een netwerk van preserveringsvoorzieningen die met elkaar gedeeld kunnen worden. Dit is niet eenvoudig en heeft tijd nodig. Het netwerk is er nog niet, maar er wordt hard gewerkt

aan de totstandkoming ervan. Zo werd ook geconstateerd dat bij Vlaamse en Nederlandse erfgoedinstellingen nog nauwelijks een bewustzijn was rondom de risico's van informatieverlies van digitale collecties. Nu tien jaar later, kunnen we constateren dat het met dat bewustzijn veel beter gaat. Er zijn nauwelijks nog erfgoedinstellingen die niet te maken hebben met het beheer en behoud van digitale informatie en zich niet bewust zijn van de risico's. Dat wil niet altijd zeggen dat zij ook daadwerkelijk actie kunnen ondernemen. Daarvoor ontbreekt nog steeds de kennis en expertise bij met name de kleinere erfgoedinstellingen.

Ondertussen staan de ontwikkelingen in de samenleving ook niet stil. Er is nauwelijks nog een onderdeel van het dagelijks leven te bedenken, waarop de digitale ontwikkelingen geen invloed hebben. Het conserveringsvraagstuk wordt steeds complexer omdat de digitale objecten die we wensen te verzamelen steeds ingewikkelder en dynamischer worden, we ons voortdurend moeten afvragen wat we willen bewaren en hoe we de authenticiteit van hetgeen we bewaren kunnen waarborgen. Dat betekent opnieuw pionieren, uitproberen, testen, uitproberen, testen Pioniers, settelaars en planners werken daarbij steeds vaker samen en leren van elkaar.

In de afgelopen 20 jaar zijn we na een periode van ontdekken en pionieren steeds meer gaan settelen, uitbreiden en plannen. Er zijn forse stappen gezet ten aanzien van onze kennis, organisatie, bewustzijn en de beschikbaarheid van technische oplossingen. Dat wil niet zeggen dat we nu rustig achterover kunnen gaan leunen. Werken aan de conservering van voor de samenleving belangrijke informatie en kennis blijft een kwestie van het maken van zorgvuldige keuzes, lange termijn commitment van organisaties, doorontwikkelen van technische oplossingen en robuuster maken van infrastructuur. Van voortdurend pionieren, settelen en vooruitdenken. En vooral van het harde werken van professionals. In de afgelopen 20 jaar is er ongelooflijk veel vooruitgang geboekt, mede dankzij dit harde werken.

In de woorden van een beroemde Fransman, "tout cela est bien quelque chose!"¹ ["dit alles is heel wat". Napoleon Bonaparte, Mémorial de Sainte-Hélène.]

Digital Preservation – wat is het niet (of niet alleen)

Tijd voor een andere benadering

Iets wat mij bijgebleven is na het lezen van het boek *Nooit meer te druk* van Tony Crabbe¹ is voor mij van toepassing op organisaties en mensen die zich bezig houden met Digital Preservation of duurzame toegankelijkheid. In een proces van verandering krijgen we – volgens Ronald Heifetz² – te maken met twee soorten problemen, technische en adaptieve problemen. Een technisch probleem – wat overigens niet technisch van aard hoeft te zijn – wordt opgelost door het vergaren van informatie, door oefening en door het ontwikkelen van de juiste vaardigheden. Er zijn helder omschreven manieren voor de oplossing. Hoe complex, cruciaal of lastig een technisch probleem ook is, door te leren en te oefenen kun je deze tot een goed einde brengen. Adaptieve problemen zijn anders. Er is sprake van een adaptief probleem als er geen eenduidige manier is om het op te lossen. Er is geen bewezen uitkomst of gebruiksaanwijzing. Om het probleem te 'tackelen' moeten we onze manieren van denken, onze overtuigingen en onze aannames veranderen.

Digital Preservation heeft adaptieve problemen of liever gezegd adaptieve uitdagingen. We reduceren deze uitdagingen vaak naar enkel technische problemen zoals hierboven beschreven. Het duurzaam toegankelijk houden van informatieobjecten³, de keuze voor een bepaalde bestandsformaat, migratie of emulatie, veranderende hard- en softwareomgevingen. De vraagstukken zijn complex. Maar oplossingen zijn aanwezig en worden binnen verschillende domeinen ontwikkeld. We vergaren informatie, ontwikkelen kennis en vaardigheden en leren bij door pilots en use cases. Dit is positief. Ondanks deze toenemende kennis blijven we zoeken naar oplossingen voor problemen waar we al jaren mee rondlopen. Hoe gaan we om met dynamische content, welke informatie bewaren we, hoe bewaren we deze informatie, in welke vorm en voor hoe lang, welke kenmerken moeten behouden blijven en hoe zorgen we ervoor dat toekomstige generaties de bewaarde informatie nog kunnen raadplegen en interpreteren.

Het proces voor het verkrijgen en duurzaam beheren van informatie bij archiefinstellingen is de afgelopen jaren conceptueel niets of nauwelijks veranderd. Voor het beleid wat we ontwikkelen, de processen die we inrichten en activiteiten die

¹ Tony Crabbe, *Nooit meer te druk. Een opgeruimd hoofd in een overvolle wereld*. Amsterdam 2017

² Leiderschapsdeskundige die Tony Crabbe aanhaalt in zijn boek

³ Een op zichzelf staan geheel van gegevens met een eigen identiteit

https://wiki.nationaalarchief.nl/pagina/DUTO:Informatieobject_geraadpleegd_20181008

we uitvoeren gebruiken we het OAIS model⁴ als referentiekader. Is dit model als uitgangspunt nog steeds de juiste focus? Is het beheren van informatie op een gecentraliseerde plek houdbaar in onze huidige informatiemaatschappij? Wat speelt er op de korte en middellange termijn en wat is de invloed hiervan op de informatie die we willen bewaren? Dit artikel gaat in op veranderingen die plaatsvinden in het vakgebied van Digital Preservation binnen archiefinstellingen. Digital Preservation gaat verder dan het bepalen van standaarden, bestandsformaten en bewaarstrategieën. Het vindt steeds vaker plaats buiten de grenzen van de organisatie wat vraagt om een andere benadering.

Hoe zien we Digital Preservation

Een stapje terug: Digital Preservation, preservering, duurzame toegankelijkheid. Een veelheid aan termen die we door elkaar gebruiken, die elkaar raken en op sommige gebieden elkaar overlappen. De smalle opvatting van het begrip Digital Preservation is het leesbaar en toegankelijk houden van informatieobjecten door de tijd heen. Dit vatten de meeste organisaties breder op. De definitie die het Nationaal Archief hanteert:

*Het op zodanige wijze vastleggen, bewaren, beheren en beschikbaarstellen van digitale archiefbescheiden, dat deze ook na verloop van tijd raadpleegbaar, toegankelijk en authentiek zijn.*⁵

Digital Preservation wordt hiermee gekoppeld aan het bredere begrip van duurzame toegankelijkheid en de organisatie rondom een e-Depot. Dit laatste geldt op dit moment als één van de belangrijkste instrumenten voor het uitvoeren van Digital Preservation. Een e-depot is overigens geen technische voorziening. Het is een geheel van organisatie, beleid, processen en procedures, personeel, databeheer, databeveiliging en aanwezigheid hard- en software, dat de duurzame toegankelijkheid van te bewaren digitale archiefbescheiden mogelijk maakt.⁶ Digital Preservation is daarom, in de brede zin van het woord, niets anders dan digitaal archiveren:

*Archiveren is alles wat nodig is om informatie duurzaam toegankelijk te maken. Zo kunnen niet alleen huidige gebruikers de informatie vinden en gebruiken, maar kunnen toekomstige gebruikers dat ook. Archiveren is niet alleen het bewaren van informatie.*⁷

Ontwikkeling Digital Preservation

Het bewaren en duurzaam toegankelijk houden van informatie gebeurt veel in het kader van cultureel erfgoed en door organisaties die van oudsher opgericht zijn voor het uitvoeren van deze taak. De rede dat we (digitale) informatie bewaren is ruimer.

⁴ Het Open Archival Information System (OAIS) Referentiemodel <https://public.ccsds.org/Pubs/650x0m2.pdf>

⁵ <https://www.nationaalarchief.nl/archiveren/kennisbank/wat-is-preserved-geraadpleegd-20181008>

⁶ <https://www.nationaalarchief.nl/archiveren/kennisbank/wat-is-preserved-geraadpleegd-20181008>

⁷ <https://www.nationaalarchief.nl/archiveren/kennisbank/wat-is-archiveren-geraadpleegd-20181008>

⁸ Informatie is duurzaam toegankelijk als deze vindbaar, beschikbaar, leesbaar, interpreteerbaar en betrouwbaar is.

Informatie – binnen een overheidscontext – is cruciaal voor een goede bedrijfsvoering. Daarnaast moeten publieke organisaties zich kunnen verantwoorden over hun werk. Denk aan WOB-verzoeken, een parlementair onderzoek, managementrapportages en auditrapporten. Overheidsorganisaties hebben belang bij toegankelijke informatie en worden zich daar bewust van. Hiermee komt voor archiefinstellingen het zwaartepunt voor de uitvoering van Digital Preservation steeds meer buiten de grenzen van de eigen organisatie te liggen. Het gaat niet alleen meer over digitale informatie die voor de eeuwigheid bewaard moet blijven of over het verplaatsen van informatie van A naar B. Vanaf het moment van creatie tot het moment dat de informatie niet meer nodig is moet men nadenken over de vindbaarheid, de bruikbaarheid en de duurzaamheid van deze informatie⁸.

Het uitvoeren van Digital Preservation heeft zich binnen een aantal jaar ontwikkeld van een groep experts die – binnen verschillende internationale projecten – de concepten hebben uitgedacht naar opgesteld beleid en procedures binnen (archief) instellingen. Het accent bij het uitwerken van de concepten lag op het inrichten van processen binnen de eigen organisatie:

*“Organizations are focused on capturing and acquiring digital information, rather than preservation or permanent access. Even if they use the term archive or have preservation in their mission, the initial goal is to get a critical mass of material; promote a culture of deposit, submission, harvesting, and sharing, and to provide access to currently collected materials. While many of the institutional repository activities are committed to long term preservation and access, the technical and metadata aspects required are not yet will incorporated into their systems”.*⁹

Dit is één van de conclusies uit het rapport *Digital Preservation and Permanent access to scientific information: the state of practice* uit 2004. Het onderzoek richtte zich op de operationele kant van Digital Preservation en systemen van organisaties die onderzoeksgegevens bewaren. Ontwikkelingen lopen niet gelijk, maar binnen de overheidssector zie je dat de focus de afgelopen jaren inderdaad heeft gelegen op het verkrijgen van informatieobjecten. E-depots werden ingericht en middels vele pilots¹⁰ is er een dienstverleningsmodel ontstaan, waardoor er steeds grotere hoeveelheden informatie vanuit de overheid binnen stroomt bij de archiefinstellingen. Het binnen krijgen van informatie is stap één. De meeste aandacht gaat nu uit naar het verlenen van toegang tot deze digitale informatieobjecten, wat een geheel eigen problematiek met zich meebrengt. Het gevolg is dat het ‘tussenstuk’, namelijk het duurzaam beheren van deze informatie, minder aandacht krijgt en dit in veel gevallen letterlijk als een technisch probleem wordt gezien. Een expert wordt om advies gevraagd. Vragen als: ‘Wat is de beste manier om mijn documenten, mijn e-mails, mijn websites en/of mijn databases te bewaren?’ komen geregeld voor. Vanuit de informatiesoort, de technische vorm of verschijning, wordt gekeken naar de op

⁹ Gail Hodge en Evelyn Frangakis *Digital Preservation and Permanent access to scientific information the state of practice* – (2004) <http://www.icsti.org/IMG/pdf/preservationrpt-3.pdf> geraadpleegd 20181008

¹⁰ Pilots met als onderwerp duurzame toegankelijkheid en/of preservation resulteren in veel gevallen in het opnemen van informatieobjecten in een e-Depot. Het weer beschikbaarstellen van deze informatieobjecten komt soms aan bod. Strategieën of uittesten van verschillende mogelijkheden voor preservation – hoe bewaar ik de informatie – worden vaak buiten beschouwing gelaten. Het zijn technische pilots die zich richtten op de korte termijn.

dat moment beste strategie. Hieruit volgt een advies, een eventuele aanpassing aan het systeem of een nieuwe werkwijze. Een aanpassing binnen de gestelde kaders, een aanpassing die past binnen onze dagelijkse routine. Conceptuele vragen¹¹, die invloed hebben op de gehele organisatie, worden niet beantwoord of blijven hangen in een discussie-fase. Dit zijn de vragen die niet door één persoon opgelost of beantwoord kunnen worden. Dit zijn de vragen die meerdere afdelingen of zelfs meerdere organisaties raken. De manier waarop we onze processen nu hebben ingericht helpt ons niet bij het beantwoorden van deze vragen.

Het lineaire proces

Archieven en andere erfgoedinstellingen werken volgens een lineair proces. Een proces van vooraf gedefinieerde en opeenvolgende stappen, van creatie tot beschikbaarstellen. De stappen worden afzonderlijk van elkaar en door verschillende actoren (organisaties, mensen en systemen) uitgevoerd. Het beleid wat we ontwikkelen, processen die we inrichten en de activiteiten die we uitvoeren gaan allemaal uit van een systeem wat opknipbaar is in verschillende fasen. Het OAIS model gaat hier ook van uit. Functies en workflows worden afzonderlijk uitgevoerd en worden gezien als volgordelijk. Deze aanpak is gebaseerd op de aanname dat het proces een begin- en eindpunt heeft. En laat dit nu net niet van toepassing zijn op Digital Preservation. Het is nooit klaar en dé optimale strategie bestaat niet. Bij de uitvoering van Digital Preservation hebben we te maken met een diversiteit aan doelstellingen, aannames en betekenissen, Organisaties – en zelfs individuele strategieën – verschillen van elkaar vanwege andere eisen, de beschikbare middelen en verschillende datatypen die ze bewaren. Elke aanpak of strategie heeft zijn eigen uitdagingen en afhankelijkheden. Er zijn verschillende richtingen te bedenken met verschillende uitgangspunten die als oplossing kunnen dienen. Ga ik elke e-mail afzonderlijk bewaren of bewaar ik een mailbox in zijn geheel. Twee ogenschijnlijk gelijke oplossingen die beide een totaal andere impact hebben op de inrichting van je processen.

Door de bredere focus op de buitenwereld komen er alleen maar meer stakeholders en mogelijkheden bij. Onze organisaties zijn in transitie. Er is veel uitzoekwerk, herinrichting van processen en infrastructuren nodig. De huidige werkzaamheden gaan gewoon door. Standaarden, metadata, handreikingen geven houvast en structuur in verandering. Bedenk daarbij dat de tools en instrumenten die we hiervoor gebruiken een invulling zijn van onze huidige werkwijze. Wat betekent dat dit anders ingevuld kan worden. Er zijn veranderingen aan te wijzen die vragen om een andere manier van denken en een andere benadering tot de informatie die we bewaren.

Het belang van het document als eenheid neemt af

Een document wordt beschreven als: *vastgelegde informatie die of vastgelegd object dat als een eenheid kan worden behandeld*.¹² Het woord document is gevoelsmatig verbonden met de fysieke wereld. Het begrip informatieobject – *Een op zichzelf staand geheel van gegevens met een eigen identiteit*¹³ – wordt om die rede steeds meer gebruikt

¹¹ Vragen als: Wat is de informatie die we willen bewaren, waarvoor en hoe wordt/werd deze informatie gebruikt, voor hoe lang willen we de informatie bewaren, hoe willen we dit beschikbaar stellen enz.

¹² NEN-ISO 15489-1:2016 nl *Informatie en documentatie – Informatie- en archiefmanagement – Deel 1: concepten en uitgangspunten*.

en past beter bij de grote verscheidenheid aan digitale informatie. Ondanks deze ruimere opvatting vinden we het moeilijk om deze eenheid aan te wijzen of te definiëren. Tot voor kort ging het beheren en bewaren van digitale informatie over documentvormen die we kennen uit de fysieke wereld ondanks de digitale vorm. We bewaren nota's, brieven, verslagen en rapporten gestructureerd en geïnclassificeerd via dossiervorming, opgeslagen in een enkel bestand (sformaat). Dit is de eenheid waarop archiefinstellingen hun processen nu hebben ingericht. Daarbij staat het aansluiten van systemen die deze informatie als eenheid beheren, de documentmanagement systemen, in de meeste gevallen centraal. Dit is duidelijk aan het veranderen. Mensen creëren steeds minder informatie binnen de centrale beheersystemen en systemen zelf generen meer en meer gegevens. Denk hierbij aan Cloud computing, de ontwikkeling van het Internet of Things en de steeds belangrijker wordende toepassingen die gebruik maken van patroonherkenning en zelflerende algoritmen. Wat we op ons scherm zien, voor ons herkenbaar als document of eenheid, wordt virtueel samengebracht. Samengebracht vanuit meerdere systemen en/of meerdere organisaties. Het bestaat enkel in die vorm op het moment dat wij het benaderen voor een bepaald doeleinde, daarna is het vervlogen.

Invloed van de omgeving wordt groter

Onze dominante strategie van het loskoppelen van informatieobjecten uit hun omgeving, oftewel het verplaatsen van informatie van systeem A naar systeem B, staat op losse schroeven. Niet alleen informatie uit de beheersystemen wordt losgekoppeld, ook informatie uit databases, emailsystemen en van websites wordt losgekoppeld van de omgeving door het exporteren en/of het bevriezen van de informatie door 'capturing'. Waar we niet altijd rekening mee houden is dat de omgeving waarbinnen informatieobjecten bestaan van belang is voor onze interactie met en interpretatie van deze informatieobjecten.

*For digital preservation to be succesful, it is not solely the technical elements of digital objects that are taken into account by organisations, but also those of the environments in which digital objects exist.*¹⁴

Het verplaatsen van informatieobjecten van de ene omgeving (een bedrijfsapplicatie) naar een andere omgeving (een beheeromgeving zoals een e-depot) en het meenemen van dit informatieobject door de tijd heen (migratie en emulatie) brengt verandering met zich mee. Dit is niets nieuws en we hebben hier al jaren ervaring mee in de fysieke wereld. Denk aan geluidsbanden of videotapes waarbij het omzetten naar een andere drager na verloop van tijd noodzakelijk is en kwaliteitsverlies geaccepteerd wordt. Bij het archiveren van born-digital materiaal is er de misvatting dat we door de tijd heen een perfecte kopie van het 'oorspronkelijke' informatieobject kunnen leveren. Het zijn per slot van rekening enkel enen en nullen die we exact en met weinig moeite kunnen kopiëren van het ene systeem naar het andere systeem. Discussies over eigenschappen die bewaard moeten blijven gaan niet over de informatie, maar over de algemene verschijningsvorm (moeten attachments in emailberichten bewaard blijven, moeten links op websites beschikbaar blijven enz.),

¹³ DUTO Begrippen <https://wiki.nationaalarchief.nl/pagina/DUTO:Begrippen> geraadpleegd 20181008

¹⁴ PLANETS project (2010)

met als uitgangspunt dat alles te reconstrueren is. Reconstructie heeft meer te maken met hoe informatie gecreëerd en gebruikt is als met de technische vorm waarin deze is opgeslagen. Zolang we informatie blijven verplaatsen en door de tijd meenemen is het belangrijk dat we veranderingen in de informatieobjecten accepteren.

Net als het veranderen van de informatie eenheid zijn de omgevingen volop in verandering. Informatie bestaat uit gegevens die hard- en software voor ons interpreteren. Veel aspecten van digitale informatie bevinden zich niet in het opgeslagen bestanden, maar in regels ingebouwd in de software. Dit kan gaan om layout aspecten die gebaseerd zijn op instellingen van de software (denk aan office documenten), zoekfunctionaliteiten, waarvan we bijna niet meer door hebben dat we ze gebruiken (het zoeken en terugvinden van e-mailberichten in onze e-mail cliënt) en inhoud die verbonden is aan een locatie en een profiel (denk aan dynamische websites die gegenereerd worden op het moment van raadplegen). De afhankelijkheid van externe regels, zoals van scripts, query's en algoritmen, neemt toe door de verandering naar een maatschappij die steeds meer gericht is op gegevens. De focus verlegt zich van het bewaren van documenten (informatieobjecten) naar het bewaren van gegevens (de 'data-driven' organisatie). Afhankelijkheid van technologie voor het uitvoeren van onze werkzaamheden neemt toe en mede als de verwachtingen van wat deze technologie kan. Door google, slimme systemen, persoonlijke assistenten op smartphones en allerlei berichten in de media over AI en zelfrijdende auto's overschatten we wat technologie op de korte termijn kan.

Fundamentele uitdagingen Digital Preservation

Bovenstaande veranderingen hebben invloed op de fundamentele uitdagingen waar we als Digital Preservation community voor staan. Met fundamentele uitdagingen bedoel ik die uitdagingen die altijd aanwezig zijn en gelijktijdig makkelijk te vermijden zijn. Een focus op de acute problemen om ons heen, waarvan er genoeg zijn, is makkelijker. Deze zijn beter traceerbaar en het is duidelijk als we ze opgelost hebben. Zijn de bestanden opgenomen in het e-Depot, loopt het proces, zijn alle metagegevens aanwezig, kan ik de informatie terugvinden en weer tonen. CHECK! Het oplossen van de 'technische' problemen geeft ons meer kennis en vaardigheden, maar er is meer nodig voor het beantwoorden van de eerder genoemde vragen: Hoe gaan we om met dynamische content, wat bewaren we en in welke vorm, welke kenmerken moeten behouden blijven, hoe zorgen we ervoor dat toekomstige generaties de bewaarde informatie nog kunnen interpreteren?

We moeten verder kijken. Ontwikkelde concepten toetsen, oprekken en aanpassen aan de nieuwe werkelijkheid. Het OAIS model is ontwikkeld voordat we er echt mee zijn gaan werken, voordat we digitale informatie in onze organisaties beheerden. Concepten zijn nooit getest op de grote hoeveelheden aan informatie die we nu creëren en zijn niet ingericht op vluchtige en dynamische content. Met het huidige framework hebben we veel bereikt: Digital Preservation is business-as-usual geworden. Er is echter weinig kritisch vermogen. Voor de uitdagingen behandeld in dit artikel is het nodig dat we vaker buiten ons (gelimiteerde) framework gaan kijken. Door te experimenteren en innoveren doen we nieuwe ontdekkingen. We ontwikkelen gevoel voor de grotere vraagstukken en verandering waar we voor staan.

Laten we beginnen met Digitale Preservation niet meer zien als een specifiek vakgebied binnen onze organisatie, maar als de kerntaak van elke archiefinstelling. Iedere medewerker is er in meerdere of mindere mate bij betrokken. Het raakt alle werkzaamheden in de kern. Plaats het in het bredere perspectief:

Archiveren is alles wat nodig is om informatie duurzaam toegankelijk te maken.

Capturing online cultures and storytelling as a method¹

Introduction

Much progress has been made in developing tools, models, strategies and other methods to preserve or document websites², but successful Web preservation also requires comprehending how the dynamic environment in which components thrive can be captured. To enable a future reconstruction of aesthetics, accountability or heritage, it is crucial to understand the context in which these websites functioned. Within the short span of a mere twenty years people have become accustomed to browsing the Web, finding all kinds of information by simply clicking from link to link. While information steams by, the context of how the information surfaces, what strata one search or the click on one link can cause is forgotten immediately since the new is there within milliseconds. The dynamics of the Web have become invisible to many of its users and the way data comes into being is forgotten. Focusing on the preservation of art on the Web, in what follows I will emphasize the importance of capturing the broader environment of platforms and social interactions in which many of these artworks thrive. Next to highlighting some of the difficulties in preserving these contexts, I explore storytelling as a method to develop and enrich a historic understanding of online cultures.

Capturing the Web

Several attempts have been made over the past two decades to document websites. One of the best known is the Internet Archive's Wayback Machine. The mission of the non-profit organization Internet Archive, founded in 1996 by Brewster Kahle, is to provide free access to all kinds of digitized and digital materials, including websites, software, games, music, moving images and books.³ On 24 October 2001 the organization launched the Wayback Machine, a free service allowing people to access and use archived versions of past Web pages, because as they argue:

"Most societies place importance on preserving artifacts of their culture and heritage. Without such artifacts, civilization has no memory and no mechanism to learn from its successes and failures. Our culture now produces more and more artifacts in digital form. The Archive's mission is to help preserve those artifacts and create an Internet library for researchers, historians, and scholars."⁴

Looking more closely at the Wayback Machine shows they only capture time-stamped snapshots of websites. As such, it foregrounds 'single-site histories', which means that single pages in a website can be studied over time.⁵ In some cases, this

works without any problems, as, for example, Jill Lepore, reporter for *The New Yorker*, describes in her article on how to archive the Internet: 'The Cobweb. Can the Internet be Archived'. She references the MH17 Ukraine plane crash in June 2014 to explain the usefulness of the Internet Archive. A mere two weeks before the incident, a curator of the Russia and Eurasia collection at the Hoover Institution, at Stanford, had submitted to the Internet Archive, a list of Ukrainian and Russian websites and blogs that ought to be recorded as part of the archive's Ukraine Conflict collection. They did this and managed to intercept and record a screenshot of a VKontakte (a social network) post by Strelkov (the field commander in Slaviansk) claiming that a plane had been shut down. The original post was removed within two and a half hours after the 'incident', but evidence of the original claim can still be traced in the Wayback Machine.⁶



Figure 1. Screenshot VKontakte, Wayback Machine⁷

- ¹ This article builds on earlier research on the preservation of net art, and some parts were previously published in Annet Dekker, 'Between Light and Dark Archiving', in: *Media Art Histories / Leonardo* (2018) (forthcoming).
- ² For example Niels Brügger 'The Archived Website and Website Philology: A New Type of Historical Document?', in: *Nordicom Review*, Vol. 29, No. 2 (2008) ; Niels Brügger 'Website History and the Website as an Object of Study', in: *New Media & Society*, Vol. 11, No. 1-2 (2009) pp. 115-132 ; Richard Rogers *Digital Methods* (Cambridge (MA), The MIT Press, 2013).
- ³ For more information see, for example, <https://www.uibk.ac.at/voeb/texte/kahle.html>, last accessed 13/10/2018
- ⁴ <https://archive.org/about/faqs.php#21>, last accessed 13/10/2018.
- ⁵ Rogers *Digital Methods*, p. 66.
- ⁶ Jill Lepore, 'The Cobweb. Can the Internet be Archived?', in: *The New Yorker* (2015, 26 January).
- ⁷ See also https://web.archive.org/web/20140717161058/https://vk.com/strelkov_info scroll down to the right time 17/07/2014 17:50 last accessed 13/10/2018.

While this is a very good example of what a large institute can do, in most cases the Wayback Machine proves to be less reliable.⁸ As also argued by Web historian Niels Brügger, an archiving process actively shapes and determines how a website is archived and hence what kind of reconstruction or analysis is possible.⁹ Not only do websites, and their copies, often suffer from temporal or technical inconsistencies, but as Brügger argues ‘the archived website is not an exact copy of the one on the live Web but a unique version as the result of the archival process’.¹⁰ This uniqueness combined with the arbitrary crawls of the Wayback Machine and possible technical inconsistencies make it an insufficient tool for art historical analyses, let alone for the preservation of an artwork. For example, inconsistencies can be part of an artwork, but they may just be glitches, or a case of bitrot. What one does not see or know cannot be traced by looking at random crawls.

To overcome a single-page history the Wayback Machine introduced Memento, an API that allows you to move back in time.¹¹ The application allows users to see the page around the time it was made rather than the present time. In 2011 the Internet Archive began to use Memento, which makes it possible to use the Wayback Machine in an ‘interactive’ mode. In some cases this generates interesting results. For example, looking at the changes over time to the website *mouchette.org* by French artist Martine Neddham and clicking one of the links via the Wayback Machine on any random date shows a standardized answer with a link that redirects to a random other part of the site and not necessarily to the one that it would normally go to (even in the past). In this case the misdirection is interesting because it has always been Neddham’s desire to make navigating the website as convoluted as possible:

“I wanted to get the viewer lost in a very complex navigation, where the placement of the links was invisible or unexpected.”¹²

Another example, for which I used Memento’s ability to recall history, is the work by Slovenian artist Igor Štromajer. In 2011 Štromajer began to delete his earlier net art pieces. After announcing the project *Expunction* on Facebook and other social media, some of the reactions expressed concern: “Igor !!!!!!! Can’t you do something else to go through your mid-life crisis ??? ? !!!!!”.¹³ I tried to trace his deleted works in the Wayback Machine with Memento. With a few exceptions, I was redirected back to his present *Expunction* project page – even in that past I couldn’t revisit earlier instances of his work. I was stuck in a circular present, no past, no memories. Such inconsistencies pose huge challenges to researchers who are unfamiliar with the documentation process, in particular around the nature of absences, redirections and the limitations of the techniques that are applied. Next to a critical assessment

⁸ For more information about the challenges of Web crawling, especially the impact and interactions of contextual factors see, for example, Emily Maemura, Nicholas Worby, Ian Milligan and Christoph Becker, ‘If These Crawls Could Talk: Studying and Documenting Web Archives Provenance’, in: TSpace (University of Toronto, 2018), https://tspace.library.utoronto.ca/bitstream/1807/82840/1/JASIST_IfTheseCrawls-Preprint-20180122.pdf, last accessed 13/10/2018.

⁹ Brügger ‘Website History and the Website as an Object of Study’, p. 126.

¹⁰ Brügger ‘The Archived Website and Website Philology: A New Type of Historical Document?’, p. 156.

¹¹ For more information see, <http://mementoweb.org/about/> and <http://timetravel.mementoweb.org/about/>, last accessed 13/10/2018.

¹² Annet Dekker, ‘In Search of Unexpected’, in: Cathy Brickwood and Annet Dekker eds. *Navigating e-Culture* (Amsterdam, Virtueel Platform, 2008) pp. 66-68.

of the methods and tools that are used, as well as a thorough understanding of the interfaces in which history is viewed, researchers are advised to collaborate with Web archivists and Web users to create a socio-technical perspective.¹⁴ A cultural perspective is just as important, in particular in cases where artists and other users have purposely obfuscated or misused a site’s ‘standard’ functionality. This means involvement of the primary ‘Web users’ such as the artist(s), or creator(s), of a site.

In the case of Štromajer, his process of deleting many of his online works is well documented. All the conversations and the discussions around *Expunction* can be traced on Facebook. However, these discussions are not saved on the Wayback Machine, since Facebook is a closed system and the data of individual users cannot be cached.¹⁵ To avoid the dependency on Facebook, Štromajer took screenshots of the discussions that are still available on his website. Recently, other tools have been developed to document social media platforms. One of them is the Webrecorder created by Rhizome, a non-profit organization based in New York. As Dragan Espenschied, one of the developers, explains:

“Current digital preservation solutions involve complex, automated processes that were designed for a web made up of relatively static documents. Webrecorder, in contrast, can capture social media and other dynamic content, such as embedded video and complex javascript.”¹⁶

Indeed, Webrecorder is a good tool to capture social media platforms, it records the posts, the likes and the comments of other users, and the replay functions as if you’re browsing the live site. But users cannot add anything or make any comments as one normally could do. For example, one of the recordings is of Amalia Ulman’s project *Excellences & Perfections* (2014). For five months and in almost 200 posts, Ulman acted out a scripted performance that culminated into an extreme makeover, which she performed on Instagram and Facebook. Playing on the cosmetics culture and as a comment on the demands social media makes on users’ appearances and experiences, Ulman convinced many of her followers and (artist) friends that what she was doing was real. The Webrecorder team documented the entire Instagram performance, including the Instagram interface, to create a faithful re-performance of the context in which the photos and comments were embedded.¹⁷ At the same time, Ulman reposted everything to Facebook where the discussion and comments were more intense. As she says in an interview:

“People got so mad at me for using fiction. That was the main critique: ‘It wasn’t the truth? How dare you! You lied to people!’”¹⁸

¹³ Annick Bureaud <https://www.facebook.com/intima/posts/144916102244400>, last accessed 13/10/2018. 14 Maemura, et al., ‘If These Crawls Could Talk’.

¹⁴ (Maemura et al. 2018)

¹⁵ See, <http://www.techcomet.com/2011/05/facebook-profiles-alternative-to.html>, last accessed 13/10/2018.

¹⁶ Dragan Espenschied ‘Rhizome Releases First Public Version of Webrecorder. A new Perspective in Web Archiving’, in Rhizome (2016, 9 August) <https://rhizome.org/editorial/2016/aug/09/rhizome-releases-first-public-version-of-webrecorder/>, last accessed 13/10/2018.

¹⁷ However, as stated on the site some elements were adjusted: “The online exhibition version on Rhizome’s Webenact server was modified to contain contemporaneous emoji, mute links that would allow to leave the artwork’s boundary, and to not display items that have been posted before the performance started,” <https://webrecorder.io/despens/amalia-ulman-excellences-perfections>, last accessed 13/10/2018.

¹⁸ Rachel Smal, ‘Amalia Ulman’, in: *Interview* (2015, 14 October) <https://www.interviewmagazine.com/art/amalia-ulman>, last accessed 13/10/2018.

Due to less privacy-friendly settings in Facebook, Ulman felt reluctant to record those statements: “No one really knew I was performing (...) It would be really complicated to archive that and keep the privacy of people”.¹⁹ The decision to record only the actual performance and not all the discussion around it is understandable from a privacy point of view; however, an important part of the work – her comments on the conventions on many social media platforms – can foremost be found on Facebook, but is likely to be lost soon.



Figure 2. Screenshot Amalia Ulman's *Excellences & Perfections*, captured by Webrecorder²⁰

'Amateur' examples

Štromajer documented the context of his project *Expunction* as best as he could by making screenshots and copying the comments that were made on his process; these can now be found on his own website. With initiatives like Webrecorder anyone can now document all kinds of (privacy-sensitive) data. Potentially this could serve as a way to capture the context of online projects as well as the performance, which means a broader and more 'democratic' view on art history.

¹⁹ Vindu Goel, 'A Dynamic New Tool to Preserve the Friendsters of the Future', in: *The New York Times* (2014, 19 October) https://bits.blogs.nytimes.com/2014/10/19/a-new-tool-to-preserve-moments-on-the-internet/?_r=0, last accessed 13/10/2018.

²⁰ <http://webenact.rhizome.org/excellences-and-perfections/20141014150552/http://instagram.com/amaliaulman>.

In the past several attempts were made to capture the context of how users experience the Web. For instance, the *NetArtDatabase* project, initiated by Robert Sakrowski and Constant Dullaart, aims to move beyond the technical specifications and the interaction model of the artwork. They try to capture the reception of net art in the environment where it was originally perceived. As Sakrowski explains, “the context, the private atmosphere, and the hardware interaction defines a large part of the 'net art activity'.”²¹ The project makes clear that documentation needs to move beyond a single method of photography or video and that one should focus on various points of view to illustrate what net art is. Although attention is given to the 'natural' environment where one interacts with net art, the formal requirements for the set up are very static, leaving little room to manoeuvre.

Capturing net art activity also takes place in other ways, for example, in the project *One Terabyte of Kilobyte Age* (2011–ongoing) by Olia Lialina and Dragan Espenschied. With this project Lialina and Espenschied take up the challenge of finding new archival methods that reflect the way archival content was created: the captured universe of Geocities. Geocities was a free Web hosting service founded in July 1995. It soon emerged as one of the most popular and inhabited places on the Web and remained so until the late 1990s. At the peak of the dot.com fever in January 1999 Yahoo! purchased Geocities for 3 billion dollars. However, Geocities soon became synonymous with old-fashioned aesthetics and basic bad taste. At the same time people drifted to social network profiles. In April 2009, Yahoo! announced that it would shut down Geocities in six months. During these months the Archive Team, with the help of about 100 people, managed to rescue almost a terabyte of Geocities pages. And on 26 October 2010, marking the first anniversary of Geocities' closing, the Archive Team released a torrent file archive of 641 GB, containing approximately 1.2 million accounts. As mentioned by digital archivist Jason Scott:

“Geocities arrived in roughly 1995, and was, for hundreds of thousands of people, their first experience with the idea of a webpage, of a full-colour, completely controlled presentation on anything they wanted. For some people, their potential audience was greater for them than for anyone in the entire history of their genetic line. It was, to these people, breathtaking.”²²

As a symbol of the 'amateur' Web, Geocities is a trace of how the Web was used at the time. This was one of the main reasons why on 1 November 2010 Lialina and Espenschied bought a 2-TB disk and began downloading the largest bittorrent file of all time.²³ They started unzipping the first files in January 2011, a process that ended in March 2011. After downloading, storing and sorting the 16,000 archived Geocities sites, which took another year, they started redistributing screen captures of the Geocities homepages through the Web. As Espenschied remarks:

²¹ Kimberley Spreuwenberg ed., *Documenting internet-based Art. The Dullaart-Sakrowski Method. Culture Vortex. Public Participation in Online Collections* (2011) p. 4 http://aaaanet/wp-content/uploads/2015/09/Documenting-Internet-Based-Art_FINAL.pdf, last accessed 13/10/2018.

²² See, <http://ascii.textfiles.com/archives/2720>, last accessed 13/10/2018.

²³ For more information about their research and findings, see Olia Lialina, 'Still There. Ruins and Templates of Geocities', in: Annet Dekker ed. *Lost and Living (in) Archives* (Amsterdam, Valiz, 2017) pp. 193-210.

“‘Content’ that is isolated, de-contextualized and shuffled around in databases of social networking sites is the main form of communication; to be useful an artifact has to work as a ‘post’, it has to become impartible and be brought into a format that is accepted everywhere. And that is a screenshot.”²⁴

The circulation was done in different ways: they opened an automated Tumblr blog that every twenty minutes uploads a new screen shot of a Geocities homepage; the screen captures are liked and reposted by the Tumblr followers, and the most reposted or liked are then presented next to related research on their blog *One Terabyte of Kilobyte Age* while at the same time distributed through Twitter.²⁵



Figure 3. Screenshot of reposting *One Terabyte of Kilobyte Age* on Twitter²⁶

The Geocities archive became a spiral in which Lialina and Espenschied reflect on the Tumblr archive of the torrent archive of the Geocities archive, people reblog, retweet, like and save the posts, and it just keeps going on. While Geocities was

²⁴ Espenschied in Trevor Ownes, ‘Digital Culture is Mass Culture: An interview with Digital Conservator Dragan Espenschied’, in: *The Signal* (2014, 24 March) <https://blogs.loc.gov/thesignal/2014/03/digital-culture-is-mass-culture-an-interview-with-digital-conservator-dragan-espenschied/> last accessed 13/10/2018.

²⁵ For the Tumblr page see <http://oneterabyteofkilobyteage.tumblr.com/>, the blog <http://blog.geocities.institute/>, and the Twitter page https://twitter.com/geocities_txt, last accessed 13/10/2018.

almost a forgotten world on the Web, due to several enthusiasts and thousands of followers and users it became not only visible but also an important marker in the Web’s history, and through liking, sharing and redistribution Geocities keeps circulating and popping up in new contexts. Next, the project introduced a whole new folksonomy through tagging – for example, ‘alive’ and ‘under construction’ – of how this new archival material could be categorized and analyzed.²⁷ Lialina and Espenschied’s project provides all kinds of information on how Geocities was used and misused, in terms of frames, banners, navigation elements, GIF’s etc.

One Terabyte of Kilobyte Age provides a means of archiving over 500,000 screenshots of homepages, and (re-)viewing the home pages through contemporary interfaces says much about the humour that drives online culture – at least in those days. Instead of purely collecting the material for the purpose of preservation the project became about questioning what ‘archive’ could mean in the context of making work accessible. The artists specifically chose to represent Geocities’ history as a dynamic and still evolving project, rather than have it exist as static ‘back log’ of data. As described by Espenschied this form was explicitly selected as a curatorial and conservation method, because

“Digital Culture is Mass Culture; it is also more about practices than objects. In order for artifacts to survive culturally, they need to become useful again in contemporary digital culture.”

Interesting questions arise concerning traditional concepts such as provenance and authenticity. As Espenschied also acknowledges, the screenshots have ‘authenticity issues’, but he goes on to say, ‘this is greatly outweighed by their accessibility and therefore impact’.²⁸ The other way to experience the online archive would be to emulate the Netscape browser, but this would be costly and require complex emulator setups. While accepting the losses, applied in a generative and circulatory way, *One Terabyte of Kilobyte Age* keeps creating new forms and interpretations, at times pushed forward by humans, at other times by bots (Twitter is supposed to have millions of bot accounts, comprising at least 15% of all its accounts). In this way, the digital archive becomes a carefully designed mass re-enactment. The light interface allows for easy re-circulation, and allows ever more forgotten moments and new experiences to resurface.

The project is driven by a desire to use technology as a tool to make visible and open up content or conduct that is neglected, forgotten, discarded or deliberately concealed. In this way, it makes sense – as also pointed out by Bethany Nowviskie, director of the international Digital Library Federation at CLIR – to “take the notion of cultural heritage not as content to be received but also as technology to be used”. This means that artefacts and events are no longer merely about the past, but are tools that can be used to imagine alternative pasts and futures.²⁹ What does this say about the feasibility of preserving online cultures?

²⁶ <https://twitter.com/GIFmodel/status/973663631018774531> post was in 13 maart 2018 (the day will come we will fix the snow).

²⁷ Lialina, ‘Still There. Ruins and Templates of Geocities’.

²⁸ *Ibidem*.

²⁹ Bethany Nowviskie, ‘Speculative Collections’ (2016, 27 October) <http://nowviskie.org/2016/speculative-collections/> last accessed 13/10/2018.

Online cultures and the value of storytelling, or storytelling as method

Jill Sterrett, head of conservation at SFMOMA, proposes an approach to the preservation of complex artworks that is based on ‘planting finds’, which she describes as documents with information value. Since the material of artworks has changed, she suggests that museum professionals need to adapt to a new situation:

“Taking into account the transitory nature of ephemeral materials, built-in physical variability and the performative elements that characterize so much of the art of the last fifty years, the work of a contemporary art museum is not business as usual.”³⁰

Sterrett was inspired by methods in archaeology, where ‘finds’ are constantly and repeatedly placed in new context. However, she suggests using the ‘find’ mechanism in reverse – not as an end point for something new to emerge – but as a method to trace the engagement with an artwork and to reveal its life over time. As Sterrett says, it enables “seeing and seeing anew”.³¹ This is similar to storytelling: relations and recognisable patterns in information create meaning and understanding, and while some things stay constant, other elements may change depending on the time, place and person. It’s important to stress that I’m not only referring to linear storytelling in which a plot develops based on certain events and culminates in a final message. Rather I’m interested in storytelling as an ongoing cyclical structure that links events and actions, which can occur simultaneously, and lacks a clear ending.³² In other words, storytelling in the digital medium provides new modes of conceptualising and ways of thinking. Similar to the transformative force brought about by the invention of writing and print, online storytelling affects modes of understanding.³³ In this sense storytelling accounts for the variables – or malleability and instability – that are inherent qualities of many contemporary artworks. Potentially this could lead to a new situation where museums would need to re-assess their ‘finds’ each time. What does it mean when the preservation of artworks is thought of in terms of (re)production or creation systems instead of ‘fixation’?

The work *The Outage* (2014) by Erica Scourti shows the relevance of storytelling as a method, in particular in relation to preserving context. In an attempt to capture her online footprint Scourti asked a ghostwriter to write her (online) memoirs. It was a first attempt to assess her digital material – from URL histories to Amazon recommendations, Facebook archives and all the other information that is freely available online – and see how her online identity was constructed through various machines. Each millisecond, numerous digital documents are sent between e-mail servers or shared on social platforms. Aided by cheap data storage, easy access

³⁰ Jill Sterrett, ‘Contemporary Museums of Contemporary Art’, in: Alison Richmond and Alison Bracker eds. *Conservation Principles, Dilemmas and Uncomfortable Truths* (Oxford: Butterworth-Heinemann in collaboration with Victoria and Albert Museum London, 2009) pp. 223-228: p. 227.

³¹ Ibidem p. 227.

³² This resembles oral traditions that are characterized by less clear divisions between main and subtopics; details can convey implicit meanings and in the retelling (a process of repetition and reflection) consistency and value perdure. See Annet Dekker, *Collecting and Conserving Net Art. Moving Beyond Conventional Methods* (London/New York: Routledge, 2018) p. 11.

³³ For more information see, among others, Walter J. Ong, *Orality and Literacy. Technologizing of the Word* (London: Methuen & Co/Ltd., 1982) and N. Katherine Hayles, *How We Think. Digital Media and Contemporary Technogenesis* (Chicago: University Press, 2012).

and distribution mechanisms, these acts of blogorrhea – the excessive, compulsive or stream-of-consciousness blogging about trivial things – provide unprecedented access to private lives, and enable the bringing together of large digital collections. Scourti wanted to understand the influence these often-invisible computational systems have over her data. The result was the novel *The Outage*, short blurbs of texts interspersed with screenshots of online material, which combined form a narrative that involves the death of the protagonist. For Scourti the entire process provided a lot of insights into what happens to someone’s online data, while leaving her feeling uncomfortable:

“It was a feeling that I had been objectified, made into an image that I wasn’t in control of; and as the book’s narrative involves a sort of death, there was a feeling that ‘my’ data body had been killed off in some way, an experience that was both exhilarating and stressful.”³⁴

Employing an outsider to speculate on and fabricate their ‘version’ of her biography, also reflects Scourti’s interest in life-writing as essentially a performative rather than a descriptive act:

“We don’t just tell *the story* of our lives, as if there is one singular story that exists prior to its representation in literary or photographic form, but through the telling of that particular story, make it a reality.”³⁵

The act of storytelling as a way to preserve and pass on information, customs and cultures from generation to generation has a long history. Also in preservation, several people have argued to include methods from oral cultures and ethnography in practice, in some cases, to capture the expedience of an artwork,³⁶ in others, to communicate and decide on what strategies to adopt.³⁷ These practices underscore the importance of methods from oral traditions, and more generally, of audience participation in the practice of preservation.

Another way to use storytelling as method to preserve the context of a work is by making the context part the work. To remain with the example of Scourti: inspired by the result of *The Outage*, in her project *Dark Archives* (2015) commissioned by Het Nieuwe Instituut as part of the research project *New Archive Interpretations*, Scourti explored the (im)possibilities and effects of online archiving using various narrative methods. Speculating on what a future, or rather present, online archive could be, for the project she uploaded her entire fifteen-year personal media archive of daily photos, videos and screenshots to Google Photo. Next to archiving users’ photos, Google Photo uses Assistant, an application that searches for similarities

³⁴ Annet Dekker, ‘Archiving Our (Dark) Lives, Interview with Erica Scourti’ *Het Nieuwe Instituut* (2016, January) <https://archieffinterpretaties.hetnieuweinstituut.nl/en/3-erica-scourti/archiving-our-dark-lives-interview-erica-scourti> last accessed 13/10/2018.

³⁵ Ibidem.

³⁶ Lizzie Muller, *Towards an Oral History of New Media Art* (Montreal: Daniel Langlois Foundation, 2008) <http://www.fondation-langlois.org/html/e/page.php?NumPage=2096> last accessed 13/10/2018; Heike Roms, *What’s Welsh for Performance? An Oral History of Performance Art in Wales* (Cardiff: Samizdat Press, 2008).

³⁷ Glenn Wharton, *The Painted King: Art Activism & Authenticity in Hawai’i* (Honolulu: University of Hawai’i Press, 2012).

in someone's photo collection. For example, people often take several photos of the same moment or object; Assistant traces all these instances, and collates them in an animation. It also can detect images that overlap and when possible stitch them together to create panoramas. Of course, nowhere it is explained what is done to detect and browse the images, and none of the other things the programme may have done are made explicit either. The collection is then parsed by auto-editing, classification and tagging software, resulting in many automatically generated videos, collages and animations. Interestingly, *Dark Archives* draws not only on the Scourti's individual collection but implicitly also on the millions of other user media in which her images and videos are tagged or linked.

In general the term 'Dark Archive' is used to indicate a repository for information that can be accessed as a failsafe during disaster recovery – it is a copy of an archive but one that consists only of metadata and is not for public use. However Scourti, is particularly interested in another type of Dark Archive, the information in an archive that cannot be seen. For example, Amazon could be seen as a very 'light' archive. Their business model is based on retrievability, which means that everything can be found and is accounted for. Amazon has to battle against the forces of darkness, which threaten to make things in the archive un-findable. This could be spam or things that have very similar titles; such duplications are rapidly increasing with algorithmically produced content. Thus, there is a need to keep things retrievable otherwise the content of the archive can fall into darkness: items are available but one cannot find – or sell – them anymore.

Scourti is particularly interested in how visibility and invisibility – or darkness – relate to archiving and archives. After producing the automatically generated videos her final step was to involve elements of staging, scripting and fictionalizing. She invited a group of writers to speculate on and caption what they imagined to be the missing set of media that somehow evaded classification within the archive; the false negatives, the misclassifications, the media that fell outside of Google's definition for that search term. By asking the writers to imagine the way an algorithm works, she was trying to get at the core of what perhaps a non-human way of thinking or logic could be. These captions were used to create a new series of videos to feed into the work that visitors can access on their smartphones. This relates to identity and memory and Scourti's interest in what and how others can see things she doesn't, "and how the technologies that we are entangled with are recording and archiving our lives".³⁸ On the one hand, it refers to notions of how identity and memory are constructed, as well as to how knowledge is inscribed in different ways. In other words, "these online platforms offer us new ways of constructing ourselves, [and] they are equally reworking the ways in which it is possible to do so".³⁹

Besides challenging notions of data collecting, shared authorship and individual memory, using her own life and documents as an example, Scourti explored the (im)possibilities of online archiving and how this relates to the way identities

³⁸ Dekker, 'Archiving Our (Dark) Lives'.

³⁹ Jon K. Shaw, Theo Reeves-Everson, 'Introduction', in: Jon K. Shaw, Theo Reeves-Everson eds. *Fiction as a Method* (Berlin: Sternbert Press, 2017) pp 5-72; p. 43.

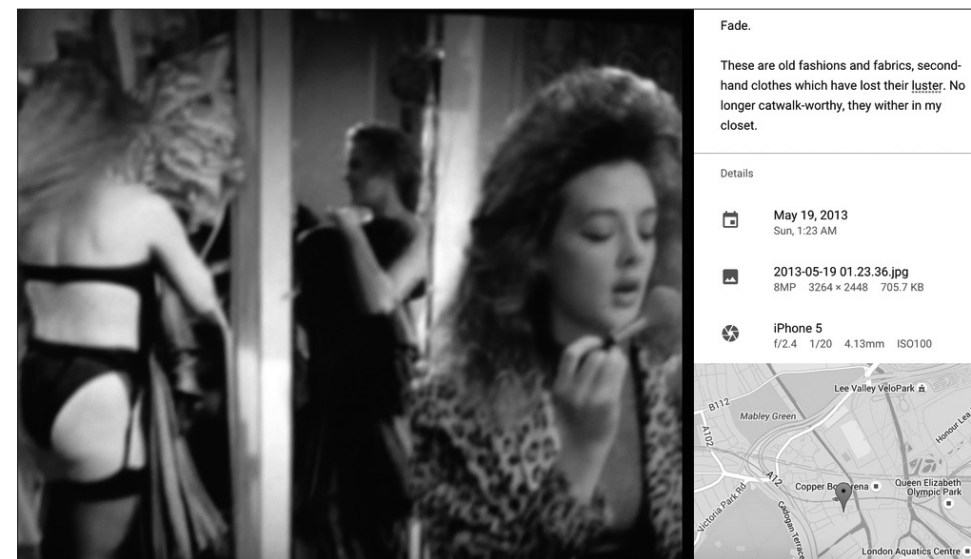


Figure 4. Still from Erica Scourti, *Dark Archives* (2015)

are constructed, while questioning the optimisation of online production and distribution. The project demonstrated how the significance and meaning of identity and memory derive from technical infrastructure and production. It also showed that an online archive is never stable – especially when using automated editing systems or certain platforms. The archive and the potentially limitless constellations within it are now expanded by contextual framings that provide additional unfinished or semi-fictional qualities. The project clearly brings out the challenges preservation is facing. Rather than worrying about missing information and dark holes, such 'loss' may generate a productive quality that focuses on retelling, which, in the process, might also do more justice to the artwork.

To be continued ...

It is clear that preservation is no longer merely an act associated with the power of institutions and authority, as is evident from the shift to artists, audience members and all kinds of technologies that are actively archiving online culture. This new situation necessitates taking seriously the tension between using complex emulation, virtualization or interactive documentation methods and the time and labour required to generate or capture 'original' cultural data. As well as accepting that online culture is no longer object-based, and therefore cannot be preserved as conventional objects, it should also be treated as a network of (inter)connected links and dependencies that are prone to constant change with each archival method that is used. To understand the provenance and context of such an unstable environment means considering online culture as the way through which to preserve it. This requires other forms of knowledge, methods and practices, and, I argue, storytelling as method could facilitate a linking of disparate elements while inciting new

potentials of participation. Such a perspective considers preservation not as following the conventional ideals of completeness and objects, but emphasizes the production of knowledge that appears through acts of copying, circulation, repetition, reflection, reuse and retelling. Preserving online cultures could benefit from disciplines and practices that are already invested in managing change by applying storytelling as method, in addition to the well-known examples of many indigenous cultures, for instance in musicology⁴⁰ or contemporary dance and performance through documentation.⁴¹ Similar to such events, most online cultures are detached from the notion of a discrete and autonomous object, since they are shaped and constructed by (non)human processes that together form a connected network of information with multiple access points. Storytelling provides a dynamic concept that is produced and productive and as such can be thought of as a tool for transformation that extends the radical unknowability of the future.

⁴⁰ Hannah Bosma, 'Canonisation and Documentation of Interdisciplinary Electroacoustic Music, Exemplified by Three Cases from the Netherlands: Dick Raaijmakers, Michel Waisvisz and Huba de Graaff', in: *Organised Sound* Vol. 22, No. 2, (2017) pp. 228-237 <https://www.cambridge.org/core/journals/organised-sound/article/canonisation-and-documentation-of-interdisciplinary-electroacoustic-music-exemplified-by-three-cases-from-the-netherlands-dick-raaijmakers-michel-waisvisz-and-huba-de-graaff/C610AD9070D5F61145F6ACD79394AEA2>, last accessed 13/10/2018.

⁴¹ Sarah Bay-Cheng, 'Theatre is Media: Some Principles for a Digital Historiography of Performance', in: *Theater* Vol. 42, No. 2 (2012) pp. 27-41. https://www.researchgate.net/publication/274350846_Theater_Is_Media_Some_Principles_for_a_Digital_Historiography_of_Performance, last accessed 13/10/2018.

Generieke voorzieningen en samenwerking

Inleiding

Het Netwerk Digitaal Erfgoed is een samenwerkingsverband dat zich richt op de ontwikkeling van een stelsel van gemeenschappelijke voorzieningen en diensten voor het verbeteren van de zichtbaarheid, bruikbaarheid en houdbaarheid van digitaal erfgoed. Het netwerk is in 2014 gestart op initiatief van het Ministerie van Onderwijs, Cultuur en Wetenschap (OCW). Deelnemers zijn grote landelijke organisaties die werken aan professioneel behoud en beheer van digitale data (de Koninklijke Bibliotheek, het Nationaal Archief, het Nederlands Instituut voor Beeld en Geluid, de Rijksdienst voor het Cultureel Erfgoed en de Koninklijke Nederlandse Academie van Wetenschappen), en een groeiend aantal partijen en personen van binnen en buiten de erfgoedsector.

Het Netwerk Digitaal Erfgoed heeft als één van zijn doelen geformuleerd: *“De resultaten van de projecten in het kader van de nationale strategie zullen in 2020 leiden tot een aanbod van sector¹ overstijgende en generieke voorzieningen. De activiteiten van de werkprogramma's zijn gericht op het gebruik van deze generieke voorzieningen binnen de context van de infrastructuur voor een gedistribueerd netwerk van erfgoed-informatie.”*²

Daarbinnen staat het Werkprogramma Digitaal Erfgoed Houdbaar gesteld voor het 'sector overstijgend' delen, benutten en opschalen van voorzieningen voor duurzaam behoud en toegang, met aandacht voor kostenbeheersing en onderlinge rolverdeling. Van 2014 tot 2018 werd deze opdracht ingevuld door de Nationale Coalitie Digitale Duurzaamheid (NCDD). Resultaten van onderzoek verricht in deze periode vormen de basis voor dit artikel.

Doel van dit artikel is inzicht te geven in het begrip generieke voorziening en kort uiteen te zetten waarom samenwerken geboden is. Een model voor 'infrastructuur' biedt inzicht, en praktijkvoorbeelden maken duidelijk hoe samenwerking kan verlopen. Tot slot wordt in gegaan op de behoefte aan samenwerking en de vraag hoe vraag en aanbod bij elkaar kunnen komen.

Wat zijn generieke voorzieningen?

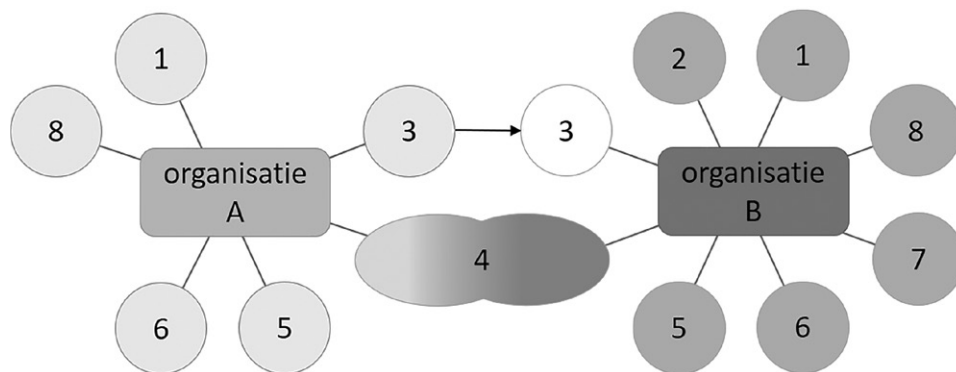
We noemen voorzieningen generiek óf omdat ze door samenwerking tussen bewaarplaatsen worden ontwikkeld, óf omdat ze als dienst door de ene van de andere bewaarplaats worden afgenomen. Een praktijkvoorbeeld van samen ontwikkelen

¹ Het begrip 'sector' is binnen NDE recentelijk in de plaats gekomen van wat eerder 'domein' was. Bedoeld worden de gebieden archief, bibliotheek, audiovisueel, wetenschap, en cultureel erfgoed.

² Overigens vinden ook buiten het Netwerk Digitaal Erfgoed heel veel vormen van samenwerking plaats op het vlak van duurzame toegang.

is de Persistent Identifier voor digitale erfgoed objecten in een digitaal archief. Een voorziening voor het registreren van Persistent Identifiers is in vrijwel alle door Nederlandse erfgoedinstellingen gebruikte collectieregistratiesystemen geïmplementeerd. Dit is gelukt omdat samen is opgetrokken met de betreffende leveranciers daarvan. SURFsara fungeert daarbij als 'nationale resolver service'³. Een praktijkvoorbeeld van het afnemen van een dienst is de relatie tussen het Nationaal Archief en het Nederlands Instituut voor Beeld en Geluid ten aanzien van de duurzame opslag van de 'videotulen' van de Tweede Kamer (zie voor details verderop onder 5, Het model in de praktijk).

In onderstaande afbeelding staan de beide manieren van samenwerking verbeeld:



Afbeelding 1. Twee manieren van samenwerking

Organisatie A levert de dienst (3) aan organisatie B, en organisaties A en B hebben samen voorziening (4) ontwikkeld, en in gebruik. Deze zou weer aan andere organisaties beschikbaar kunnen worden gesteld.

Generieke voorzieningen zouden (onder andere) de volgende kenmerken moeten hebben:

- Interoperabel⁴
Voor samenwerking zijn standaarden, protocollen en procedures nodig. Interoperabiliteit is een vereiste om een functionerend netwerk tot stand te brengen. De rollen en taakverdelingen zijn helder.
- Inclusief
Ze zijn niet specifiek voor één partij. Er worden in principe geen partijen uitgesloten (hoewel in de praktijk deze 'eis' niet hard is: een dienst kan bijvoorbeeld specifiek voor één sector zijn). De dienst is herbruikbaar voor andere afnemers.
- 'Trustworthy'
De afnemer van een dienst moet er op kunnen vertrouwen dat de dienst voldoet aan gemaakte afspraken en standaarden. Certificatie is daarbij een goed middel.
- Schaalbaar
Denk hierbij aan aantallen afnemers en de frequentie van gebruik, maar ook aan aantallen en omvang van bestanden.

- Robuust
Een dienst mag niet gevoelig zijn voor storingen van buitenaf, foutieve input mag niet doorwerken op de output.
- Betaalbaar
En als laatste, maar niet het onbelangrijkste, kenmerk moet de dienst voor de afnemer betaalbaar zijn.

Voor het inzicht in soorten voorzieningen en diensten biedt de huidige *Producten en Diensten Catalogus, Digitale Archiefdiensten* van de Serviceorganisatie Nationaal Archief (versie 1.0 van 11-1-2018)⁵ een goede eerste ingang. Wanneer de indeling van de Serviceorganisatie algemener wordt gemaakt ontstaat het volgende beeld:

- 1) Het beheren van (uitgeplaatste⁶) informatieobjecten van afnemers door de leverancier.
- 2) Het bieden van toegang tot die (uitgeplaatste) informatieobjecten.
- 3) Het beschikbaar stellen van digitale voorzieningen aan afnemers. Afnemers kunnen met deze digitale voorzieningen zelf informatieobjecten beheren. Voor het Nationaal Archief is één van deze voorzieningen het e-Depot. Het ondersteunen van de aansluiting van afnemers op het e-Depot is ook onderdeel van deze diensten.
- 4) De leverancier stelt zijn kennis over digitale archivering beschikbaar aan iedereen die daaraan behoefte heeft. Dat kan zijn door vragen te beantwoorden maar ook door kennisproducten beschikbaar te stellen.
- 5) Daar waar antwoorden op kennisvragen en kennisproducten niet voldoende helpen kan de leverancier op verschillende onderwerpen van digitale archivering gericht advies geven.
- 6) De leverancier biedt opleidingen aan op het gebied van digitale archivering.

Waarom samenwerken?

Eén van de vertrekpunten van het Netwerk Digitaal Erfgoed is dat door samenwerking en het delen van voorzieningen en diensten effectiever en efficiënter het waarborgen van duurzame toegankelijkheid van digitale collecties kan worden bereikt. Dit is een aanname die vooralsnog niet is bewezen.

Effectiviteit of doeltreffendheid. Voor duurzame toegankelijkheid betekent dat: wanneer alle organisaties die hun digitale collecties duurzaam toegankelijk willen (of moeten) bewaren daartoe in staat zijn, dan zijn de betreffende inspanningen en financiële middelen effectief. Vooralsnog hebben niet alle organisaties in Nederland de duurzame toegankelijkheid van hun digitale collecties geregeld. Dat betekent dat nog (heel) veel inspanning en inzet van financiële middelen noodzakelijk is. Kleinere erfgoedinstellingen hebben noch de kennis noch de middelen om de duurzame toegang tot hun digitale collecties te garanderen. Aansluiting bij voorzieningen van anderen kan dit oplossen. Daarnaast zijn veel voorzieningen, tools en diensten

³ Zie <http://www.netwerkdigitaal erfgoed.nl/kennis-en-voorzieningen/digitaal-erfgoed-houdbaar/persistent-identifiers/> geraadpleegd 11 oktober 2018

⁴ Zie voor een uitleg van dit begrip onder andere <https://nl.wikipedia.org/wiki/Interoperabiliteit> geraadpleegd 11 oktober 2018

⁵ Nationaal Archief *Producten en Diensten Catalogus, Digitale Archiefdiensten* (versie 0.1 11-1-2018) zie <https://www.nationaalarchief.nl/archiveren/kennisbank/producten-en-diensten-catalogus-digitale-archief-diensten> (geraadpleegd 11 oktober 2018) alwaar ook de Catalogus zelf is te downloaden.

⁶ 'Uitplaatsen' is een term specifiek voor de sector Archief, maar is wel generiek toepasbaar waar het eigendom bij de vormer blijft en het (duurzaam) beheer door de dienstverlener wordt gedaan.

sterk afhankelijk van kleine ‘communities’ of zelfs van één persoon. Door deze tools gezamenlijk te ontwikkelen en te beheren is de kans groter dat ze in leven kunnen worden gehouden.

Efficiëntie of doelmatigheid. Effectief, maar dan wel tegen zo laag mogelijke inspanningen en kosten. Het is evident dat samenwerken tijd, dus geld kost. De door samenwerking toenemende doelmatigheid wordt door dat samenwerken dus per definitie wat beperkt.

Om onderbouwd beslissingen te kunnen nemen over de consequenties van samenwerken voor effectiviteit en efficiëntie is het nodig om inzicht te hebben in kosten. In het artikel ‘Kosten’ van Herman Uffen elders in deze bundel wordt hieraan aandacht besteed.

Wanneer kosten op een uniforme wijze in zicht worden gebracht is dat een hele stap. Maar óók de baten moeten in zicht komen. Pas dan wordt duidelijk wat schaalvoorwaarden zijn, en wat synergie oplevert. Een correcte kosten – batenanalyse maakt inzichtelijk wat investeringen (in tijd en geld) opleveren bij verschillende opties of scenario’s. Wellicht is zelfs een maatschappelijke kosten – batenanalyse⁷ noodzakelijk. Het kunnen beschikken over huidig digitaal erfgoed in de toekomst, of juist het niet erover kunnen beschikken omdat het verloren is geraakt, zijn welvaartseffecten (of: maatschappelijke effecten) die ook in de vergelijkingen een rol dienen te spelen. En dan kunnen de aannames van grotere effectiviteit en efficiëntie aantoonbaar worden gemaakt, en kunnen beleidsmakers en beslissers tot goede afwegingen komen.

Blokjesmodel als grondvest

Van de duizenden instellingen die cultureel erfgoed of wetenschappelijke data in hun collecties hebben, beschikt een aantal over e-Depot voorzieningen waarmee de langdurige toegankelijkheid van digitale objecten kan worden verzekerd. Daarin is veel (vaak projectmatig) geïnvesteerd. Het werkprogramma Digitaal Erfgoed Houdbaar startte voorjaar 2018 met een inventarisatie van ‘digitale archieven’ in Nederland.⁸ Doel is om, in stappen, een landkaart te maken van de diensten en voorzieningen voor duurzame toegankelijkheid in Nederland. Eén van de voorlopige uitkomsten is dat van de 44 organisaties die in de eerste stap bevestigd zijn, er 27 (60 %) zeggen over een ‘digitaal archief’ te beschikken. Als alle, vaak kleine, organisaties in Nederland zouden zijn meegenomen in de survey dan zou dit percentage zeker aanmerkelijk lager uitvallen.

⁷ Voor informatie over een maatschappelijke kosten- en batenanalyse zie bijvoorbeeld <http://www.mkba-informatie.nl/>

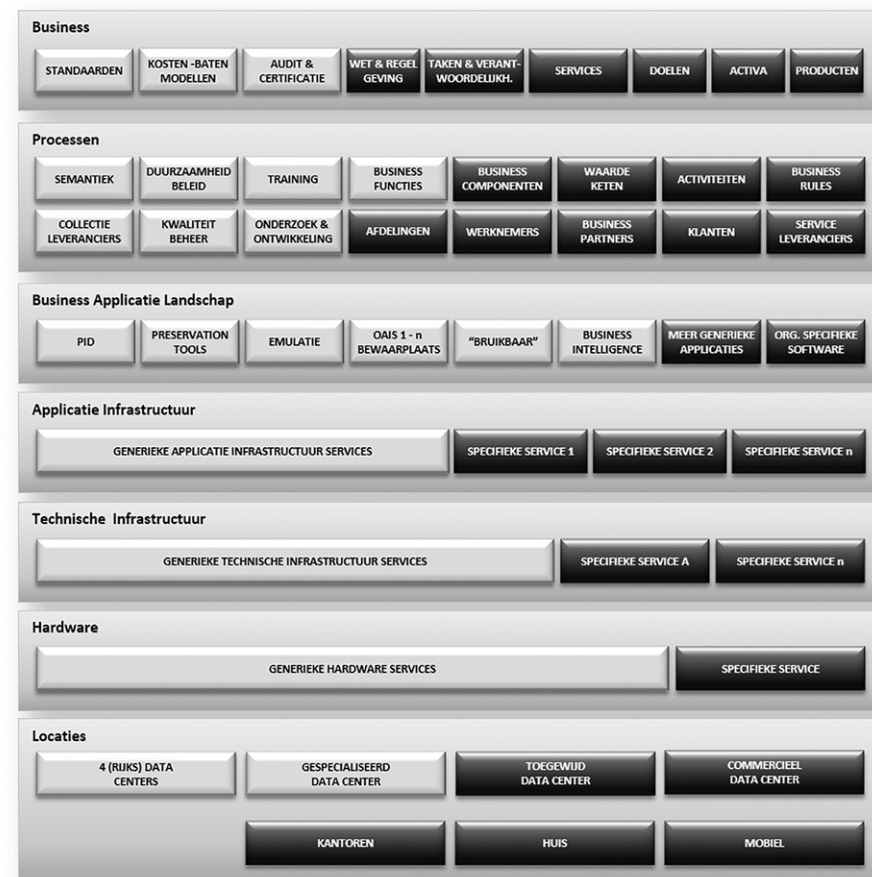
⁸ Joost van der Nat, Marcel Ras *SURVEY Digitale Archieven in Nederland* (7 mei 2018) zie <https://www.netwerkdigitaalergoed.nl/wp-content/uploads/2018/05/20180507-Netwerk-Digitaal-Erfgoed-Survey-Digitale-Archieven.pdf> geraadpleegd 11 oktober 2018. ‘Archief’ is hier gebruikt in de zin van ‘bewaarplaats’ of ‘e-Depot’.

⁹ Joost van der Nat, Marcel Ras *Samen Bouwen aan een netwerk van landelijke voorzieningen. Eindrapport van het onderzoek naar een landelijke infrastructuur voor duurzame toegang tot digitale informatie* (2015) Zie voor het rapport: <http://www.netwerkdigitaalergoed.nl/kennis-en-voorzieningen/digitaal-erfgoed-houdbaar/gemeenschappelijke-voorzieningen/> geraadpleegd 11 oktober 2018.

Een e-Depot of ‘digitaal archief’ is echter maar één van de onderdelen van de infrastructuur voor duurzame toegankelijkheid. In 2014 - 2016 is in NCDD-verband een model ontwikkeld⁹, dat beoogt het geheel van voorzieningen in kaart te brengen dat nodig is om duurzame toegankelijkheid te borgen. Het model is gebaseerd op een zogenoemde ‘Business - IT-stack’:

- ‘Business’ wordt vertaald in....
- Processen, die worden ondersteund door....
- Applicaties en databases, die werken in een....
- Applicatie infrastructuur (servers) die draait op een....
- Technische infrastructuur (o.a. opslag) die bestaat uit....
- Machines & netwerken die....
- Kantoren & internet verbinden met datacenters.

Wanneer nu alle ingrediënten van de infrastructuur voor duurzame toegankelijkheid hierop worden geplaatst, dan ontstaat het beeld zoals weergegeven in afbeelding 2.



Afbeelding 2. Model voorzieningen duurzame toegankelijkheid

Het model is daarmee een soort kookboek, het geeft een overzicht van de ingrediënten voor de infrastructuur voor digitaal duurzame toegang. Het model geeft inzicht in de elementen van die infrastructuur die generiek, en mogelijk deelbaar, zijn. Daarmee is het een nieuwe manier van kijken, die internationaal erkenning heeft gekregen. Het model kreeg in 2016 de *Digital Preservation Award for Research and Innovation* van de Britse Digital Preservation Coalition.

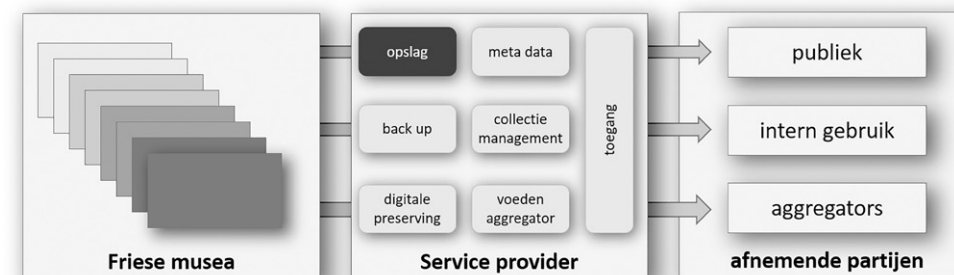
De donkergekleurde blokjes zijn niet, of minder geschikt voor samenwerking. Elke organisatie kent immers bijvoorbeeld zijn eigen doelen, activa en producten, heeft een eigen waardeketen, en heeft eigen gebouwen. De licht gekleurde blokjes¹⁰ zijn kandidaat voorzieningen om te worden gedeeld. Met name in de IT-lagen (Hardware tot en met Business Applicatie Landschap) zijn volop mogelijkheden voor een organisatie om in een gedistribueerd model deel te nemen. Maar ook hogerop, in de Processen en in de Business, zijn veel elementen deelbaar.

Het model in de praktijk

In het begin van dit artikel kwam al het voorbeeld van samenwerken door samen ontwikkelen ter sprake van de Persistent Identifier. Deze is in het Blokjesmodel als element aanwezig (in de laag Business Applicatie Landschap). Op basis van een gedeeld uitgangspunt (alle digitale objecten in een erfgoed collectie dienen te zijn voorzien van een Persistent Identifier), is door gezamenlijk optrekken bereikt dat dit tegen lage kosten kan worden verwezenlijkt.

Als tweede voorbeeld van het afnemen van een dienst werden de 'videotulen' van de Tweede Kamer genoemd. Het Nationaal Archief heeft een wettelijke taak om de notulen van de Tweede Kamer en Commissies daarvan te archiveren. Dat gebeurt al eeuwen (sinds 1576 toen voor het eerst sprake was van de 'Staten-Generaal'), en ook al met het hedendaags digital born materiaal. Een vrij recent nieuw fenomeen is echter dat van de vergaderingen van de Tweede Kamer en de Commissies video opnamen worden gemaakt, ook bekend als 'videotulen'. Ook deze moeten door het Nationaal Archief worden gearhiveerd, na de wettelijke overbrengingstermijn. Het Nationaal Archief beschikte echter noch over de daarvoor vereiste infrastructuur, noch over de vereiste expertise. Het Nederlands Instituut voor Beeld en Geluid beschikt echter wél over de noodzakelijke infrastructuur en expertise voor het duurzaam toegankelijk houden van 'videotulen'. Een succesvolle pilot (2016-17) wordt nu vertaald naar een permanente oplossing. Het Nationaal Archief blijft verantwoordelijk, doet de waardering en selectie van het materiaal, en draagt zorg voor de ontsluiting. Het Instituut voor Beeld en Geluid verzorgt de duurzame opslag van, en toegang tot het materiaal. In het Blokjesmodel maakt het Nationaal Archief dus gebruik van een voorziening ('OAIS Archief', dat in zich meerdere diensten heeft) die het Instituut voor Beeld en Geluid levert.

Als derde voorbeeld gaan we in op de situatie bij de musea in Friesland. Begin 2017 heeft de Museumfederatie Fryslân een onderzoek uitgevoerd onder 30 Friese erfgoed instellingen. Uitkomsten van dit onderzoek geven zicht op de volgende, mogelijke ontwikkeling:



Afbeelding 3. Mogelijke toekomstige situatie Friese musea

Als eerste stap gaan de musea samenwerken met een provider om de opslag van hun digitale objecten zeker te stellen. Een daarbij horende vervolgstap is het realiseren van back up voorzieningen. De data zijn dan in ieder geval voor de korte termijn veiliggesteld ('bit preservation'). Vervolgens kunnen stappen worden gemaakt met het standaardiseren en normaliseren van meta data, en kan worden gewerkt aan een collectie managementsysteem. Parallel daaraan kan worden zeker gesteld dat, op basis van een 'preserveringsbeleid', de duurzame toegankelijkheid van de objecten is geborgd. En kan worden gewerkt aan het ontsluiten van de collecties voor intern gebruik en publiek. Door het juist voeden van 'aggregators' wordt het bereik van de collecties potentieel enorm vergroot. Voor de eindsituatie zouden de musea dan in termen van het Blokjesmodel gebruik maken van een voorziening die wordt geleverd door een Serviceprovider. In die voorziening zijn dan een aantal diensten opgenomen, zoals opslag, back-up, en de uit meerdere diensten bestaande voorziening 'OAIS-archief'.

Behoeft aan samenwerking

In de eerder in dit artikel genoemde, in 2018 gestarte, inventarisatie voor de landkaart van 'digitale archieven' in Nederland is ook gevraagd op welke vlakken de organisaties behoefte hebben aan meer samenwerking en de wijze waarop ze nu al samenwerken met anderen. Alle respondenten geven aan nu op de een of andere wijze in een samenwerkingsverband deel te nemen. Maar liefst 36 instellingen hebben regelmatig overleg met collega-instellingen over het gebruik van hun systemen. Dit betreft zowel digitaal archiefsystemen als andere software die bij de betreffende instellingen in gebruik is. Zo nemen veel instellingen deel aan gebruikersgroepen die ingericht zijn ten behoeve van de systemen die ze gebruiken. Een voorbeeld hiervan is de onlangs opgerichte Vlaams-Nederlandse Archivematica gebruikersgroep.¹¹ 64 % van de respondenten praat met collega's over implementatie issues ten aanzien van een digitaal archief.

Gevraagd naar de behoefte aan meer samenwerking geeft 70 % van de respondenten aan geïnteresseerd te zijn in het uitwisselen van 'best practice'. Daarnaast hebben erfgoedinstellingen grote behoefte aan onder andere samenwerking op het gebied

¹⁰ Zie voor een beschrijving van deze 'blokjes': <http://voorzieningen.ncdd.nl/bouwblokken.php> geraadpleegd 11 oktober 2018

¹¹ Zie https://wiki.archivematica.org/Community/Regional_User_Groups#NETHERLANDS_and_FLANDERS geraadpleegd 11 oktober 2018.

van duurzaamheidsbeleid en het plannen van preservering. Ook vaak worden genoemd file format policies, training, meta data (beheer), preservation watch, en pré ingest.

De uitkomsten van de behoefte aan samenwerking bieden het werkprogramma Digitaal Erfgoed Houdbaar een goed houvast om goed naar 'de klant' te luisteren, en de inspanningen te richten op die zaken waar vraag naar is. Daarmee kan er ook een verdere inkleuring van de Landkaart worden gegeven.

Hoe breng je vraag en aanbod bij elkaar?

Zoals in de inleiding van dit artikel staat geschreven heeft het Programma NDE als een van de doelen gesteld dat in 2020 er een aanbod is van sector overstijgende en generieke voorzieningen. Aan de vraagzijde ervaren met name de kleinere erfgoedinstellingen de doelstelling van duurzame toegankelijkheid als een grote uitdaging. Er is weinig geld, weinig menskracht, en veel van de vereiste stappen vragen behoorlijk wat expertise. De vraag is nu hoe we vraag en aanbod bij elkaar gaan brengen. Eén van de manieren waarop het Netwerk Digitaal Erfgoed dit wil bereiken is met de eerdergenoemde Landkaart van generieke voorzieningen. Als hierboven toegelicht is daarvoor een eerste stap gezet met de inventarisatie bij 44 organisaties welke producten ze gebruiken voor hun 'digitaal archief'. Daarmee is nog niet het aanbod van alle mogelijke elementen van de infrastructuur in kaart gebracht. Het Blokjesmodel dient hier als leidraad. Volgende stappen in de inventarisatie zullen dan ook gericht zijn op het zo breed en diep mogelijk in kaart brengen welke voorzieningen en diensten voor duurzame toegankelijkheid in Nederland bestaan die kunnen worden gedeeld. Die landkaart zal dynamisch zijn.

Aan de vraagkant dienen we te helpen met het formuleren van de vraag, in duidelijke stappen, vanuit verschillende invalshoeken. Antwoorden op vragen als 'Wat ben ik?', 'Wat heb ik?', 'Wat kan ik?', 'Wat zijn mijn ambities?', 'Wat moet ik verbeteren?' moeten uiteindelijk leiden tot een 'gearticuleerde vraag'. Hier komt voor de vragende partij nadrukkelijk het onderwerp 'volwassenheid' van de digitale huishouding aan de orde. Zoals het op de site van Scoremodel van PACKED en DEN treffend staat beschreven: "[Het Scoremodel...] brengt de sterke en zwakke punten van je digitale huishouding in kaart en geeft aan waar je best ingrijpt om de risico's verder te verkleinen."¹²

En zo kunnen we dan vraag en aanbod bij elkaar brengen. In het ideale geval ondersteund door een 'slimme' website die de parameters van de vragende partij kan koppelen aan die van de aanbiedende partij. En daarmee komt een duurzaam toegankelijk digitaal Nederlands Erfgoed dichterbij.

¹² Zie <https://scoremodel.org/site/home>

De kosten van digitale duurzaamheid beheersbaar

Aanleiding / Probleemstelling

Het belang van het toegankelijk maken en het behouden van waardevolle digitale bronnen voor de lange termijn wordt al lange tijd onderschreven en door (erfgoed) instellingen en organisaties onderkend. Echter, de kosten hiervan zijn niet altijd even inzichtelijk. Veelal wordt er onvoldoende rekening gehouden met de structurele kosten van beheer van digitale collecties, doordat deze in veel gevallen onvoldoende in beeld zijn gebracht. Gevolg is dat de beheerskosten vaak niet in de reguliere exploitatie van instellingen en organisaties zijn opgenomen. Ook is de verwachting van veel erfgoedinstellingen dat de structurele kosten zeer hoog zullen zijn, mede gezien de exponentiële toename van beschikbare hoeveelheden gedigitaliseerd en *digital born* materiaal. In veel gevallen worden incidentele gelden (projectmiddelen) gebruikt als dekking voor structurele beheerskosten. Maar zijn deze aannames correct? En hoe zorgen we ervoor dat de kosten voor langetermijnbeheer van digitale collecties beheersbaar gemaakt worden? Om het mogelijk te maken de kosten met betrekking tot digitale duurzaamheid inzichtelijk en beheersbaar te krijgen is een kostprijsmodel ontwikkeld.

Kosten duurzame toegang – ontwikkeling Kostprijsmodel Digitale Duurzaamheid

Doel van het model is om de sturing op de kosten (incidenteel en structureel) van duurzame toegang van digitaal erfgoed te versterken. Inzicht in de kostenstructuur, de opbouw van deze kosten en de kostenbepalende variabelen is daarbij essentieel, niet alleen wat betreft de huidige kosten, maar ook in het begroten van aankomende kosten. Belangrijk is om hierbij dezelfde financiële randvoorwaarden te hanteren, zodat onderlinge vergelijking mogelijk is. Voor instellingen, beleidsmakers en financiers is het dan vervolgens mogelijk om op basis van dit verkregen inzicht beter onderbouwde keuzes te maken ten aanzien van investeringen in en het gebruikmaken van voorzieningen voor duurzame toegang.

Met dit project is voortgebouwd op eerdere (inter)nationale initiatieven, opgestart om organisaties te helpen meer zicht te krijgen op de kosten van duurzame toegang tot hun digitale materiaal. Zo ontwikkelde het project *Collaboration to Clarify the Costs of Curation* (4C) de CCEX-tool (Curation Costs Exchange), waarmee organisaties inzicht krijgen in hun eigen bestedingen en die vervolgens kunnen vergelijken met die van andere instellingen. 4C heeft hierbij de volgende visie geformuleerd¹:

¹ 4C, 2015, *Investing in Curation; A Shared Path to Sustainability*, 20 februari 2015, p. 1-26. <http://www.4cproject.eu>

“In five years time (2020) it will be easier to design or procure more cost effective and efficient digital curation services because the costs, benefits and the business cases for doing so will be more widely understood across the curation lifecycle and by all relevant stakeholders. Cost modelling will be part of the planning and management activities of all digital repositories.”

Bij de ontwikkeling van het Kostprijsmodel Digitale Duurzaamheid is gekeken naar bestaande modellen² en is er een koppeling gemaakt met CCEX. Het model moet organisaties helpen om:

- 1) inzicht te krijgen in de kosten van digitale duurzaamheid, ook om de grip en sturing daarop te vergroten, en in de competenten waaruit deze kosten bestaan;
- 2) de kosten te monitoren en meer grip te krijgen op de ontwikkeling daarvan;
- 3) de kosten van duurzame toegang te vergelijken met die van andere organisaties, om ervan te leren en de grip en sturing verder te vergroten.

Hiermee wordt het mogelijk om keuzes te maken ten aanzien van (collectie)beleid, inzet van mensen en middelen, samenwerking en benodigde infrastructuur.

Activity Based Cost-model

Het Kostprijsmodel Digitale Duurzaamheid is een zogeheten *Activity Based Cost-model* en richt zich op de activiteiten die nodig zijn in het proces van duurzaam bewaren en toegang bieden. Het duurzaam toegankelijk houden van digitale informatie is een actief proces, waarbij het specifiek gaat om de volgende activiteiten: *Selection/Pre-ingest, Ingest, Processing, Documentation, Archive, Access en User Support*. Daarnaast komen in elke activiteit overkoepelende procesactiviteiten terug: *Metadata, Preservation Management, Infrastructure* en *ICT*.

Wat houden de fasen van het proces precies in?

- **Selection/Pre-ingest** bevat activiteiten gerelateerd aan de waardebeoordeling, criteria en selectie van data. Ook kan het gaan over het beschikbaar stellen van datagidsen, training, consulting en communicatie met dataleveranciers over zaken als dataformats, datamanagement, -plannen en -rechten.
- **Ingest** betreft het ontvangen, lezen, controleren van de kwaliteit en categoriseren van ontvangen data, tot het punt van toevoeging in het archief (maar exclusief processing). Ingest kan met de hand of geautomatiseerd gebeuren, met vaak handmatige stappen als kwaliteitschecks et cetera.
- **Processing** bevat het maken van operationele productstromen, non-operationele producten en reprocessing, procescontrole (productieplanning, monitoring, etc.) en het genereren van producten, maar ook: opnieuw processing (reprocessing) van nieuwe versies van eerder gemaakte producten volgens een opgestelde planning of binnen een gespecificeerde ‘reprocessing’-capaciteit. Ook de integratie van verschillende datasets kan eronder vallen, evenals ‘data mining’, waarbij software verschillende datasets doorzoekt en data die aan de gestelde criteria voldoen samenvoegt.
- **Documentatie** betreft het ontwikkelen of verbeteren van ontvangen data en productdocumentatie (zoals gebruikershandleidingen, catalogi, etc.) en gebruikerservaringen, read software, systeeminformatie, onderhoud en vervanging. Exclusief metadata.

- **Achiveren** omvat het invoeren van data in het archief, gegevensbeheer, -management, het duurzaam bewaren van data, metadata en documentatie binnen een archief. Ingevoerde data kunnen data bevatten van externe bronnen of data die door de instelling zelf zijn gegenereerd of aangepast.
- **Toegang en distributie** omvat het ophalen van de gevraagde data uit het archief, het uitvoeren van autorisatie indien nodig, het selecteren van deelselecties, resampling, reformatting/formatconversie, reprojectie van verpakken, en het aanbieden van het eindproduct via elektronische of fysieke media.
- **User support** (gebruikersondersteuning) betreft ondersteuning van gebruikers in direct contact met ondersteuningspersoneel, zoals: vragen beantwoorden, orders aannemen, helpdesk bemannen (personeel wacht op gebruiker met een vraag), het bereiken van potentiële nieuwe gebruikers en educatie/training van huidige of potentiële gebruikers. User support-personeel heeft ook expertise om gebruikers te ondersteunen bij het selecteren en gebruiken van dataproducten..
- **Metadata** betreft beschrijvende of contextuele informatie over een data-object³. Activiteiten gerelateerd aan metadata kunnen ook gerelateerd worden aan andere activiteiten. Metadata betreft een continu proces bij het duurzaam digitaal opslaan van data. Vanwege het belang ervan en het werk dat eraan gerelateerd wordt, is metadata toegevoegd als aparte activiteit.
- **Preservation management** gaat over de diensten en functionaliteiten voor het monitoren van de (digitale) omgeving, het doen van aanbevelingen en het opstellen van duurzaamheidsplannen om de toegang tot en de bruikbaarheid van digitale objecten te bewaren en toegankelijk te houden voor (her)gebruik voor de lange termijn, ook als de originele digitale omgeving verouderd raakt.
- **Infrastructure** omvat de ontwikkeling en het operationeel maken van data- en informatiesystemen en de mogelijkheden die nodig zijn om de doelstellingen van de instelling met betrekking tot digitale duurzaamheid te behalen, inclusief ontwikkeling en het operationeel maken van de data-infrastructuur: hardware, software en back-up-faciliteit(en).
- **ICT** gaat over sustaining engineering: Onderhouden en verbeteren van (op maat gemaakte) applicatiesoftware en alle ontwikkelingen na de implementatieperiode. Engineering support is gericht op de interne operaties van de instelling. Technical coordination, inclusief engineering, is naar buiten gericht: het ondersteunen van de instelling als een deel van een systeem van samenwerkende instellingen.

De keuze voor deze activiteiten en procesactiviteiten is gebaseerd op een studie van verschillende al opgestelde kostprijsmodellen met betrekking tot digitale duurzaamheid.^{4&5} Hierbij is onder andere gekeken naar de variabelen en activiteiten die in deze modellen zijn opgenomen, het doel van de modellen en de complexiteit ervan.

² <http://www.4cproject.eu/summary-of-cost-models/>

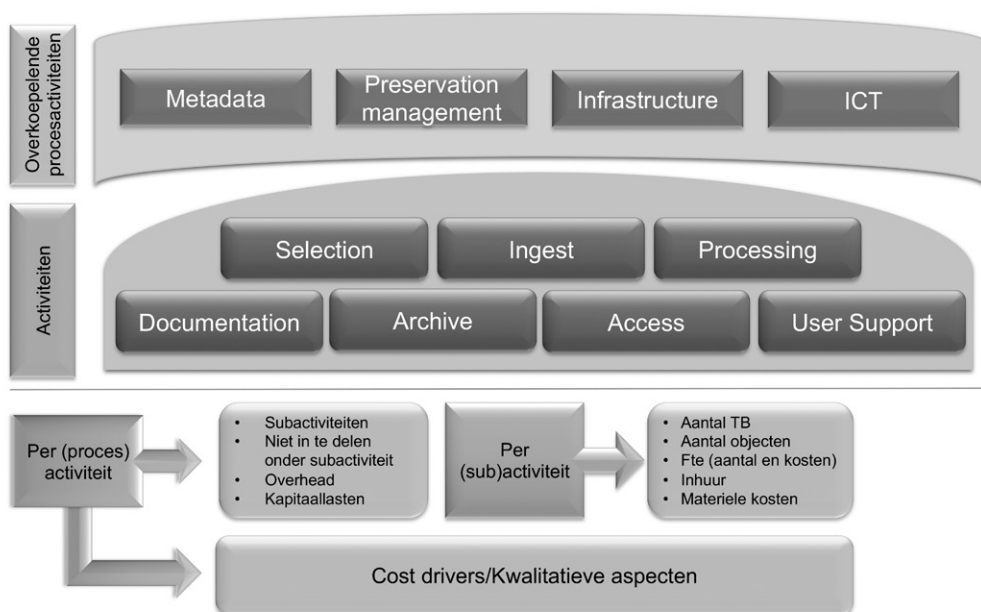
³ ‘Object’ is gedefinieerd met de richtinggevende uitspraak dat er een persistent identifier aanwezig moet zijn.

⁴ <http://www.4cproject.eu/summary-of-cost-models/>

⁵ Keijser, U.B., K.H.E. Johansen, A. Thirifays, A.B. Nielsen, D. Wang, S. Strodl, T. Miksa, J. Davidson, P. McCann, J. Krupp & H. Tjalsma, ‘Evaluation of Cost Models and Needs & Gaps Analysis’, 4C, (30 juni 2014), p. 1-99.

Op basis van de verkregen informatie is het Kostprijsmodel Digitale Duurzaamheid opgesteld. De activiteiten, subactiviteiten en bijbehorende definities zijn daarnaast gebaseerd op het OAIS-model⁶ en door middel van discussie in werkgroepen van deelnemende instellingen aangescherpt. Daar is ook de werking van het model getoetst.

Het Kostprijsmodel Digitale Duurzaamheid geeft inzicht in de kostenstructuren, de opbouw van deze kosten en de kostenbepalende variabelen (*cost drivers*) en toont de financiële cijfers van beheer, behoud en toegang van digitaal erfgoed. Indien het model voor meerdere jaren wordt ingevuld, kan bekeken worden hoe sterk de invloed van deze factoren is op de kosten (= correlaties) en kan het model een statistische forecast geven van de kosten op basis van deze correlaties.



Afbeelding 1. Kostprijsmodel Digitale Duurzaamheid

⁶ Het Open Archival Information System Reference Model is ontwikkeld om de functies te benoemen die je nodig hebt voor duurzaam beheer en om daarvoor een gemeenschappelijke terminologie te ontwerpen. Er worden in dit model zes functies onderscheiden:

- Ingest
- Data Management
- Archival Storage
- Administration
- Preservation Planning
- Access

Lavoie, B., 2014, *The Open Archival Information System (OAIS) Reference Model: Introductory Guide* (2nd edition), DPC Technology Watch Report 14-02, oktober 2014, p. 1-37.

Eerste resultaten invulling kostprijsmodel – fase 1

In 2016 en 2017 hebben 10 instellingen het ontwikkelende kostprijsmodel ingevuld. Deze exercitie was de invulling van de eerste fase uit de Roadmap, die is uitgezet door het Netwerk Digitaal Erfgoed in de ontwikkeling van het kostprijsmodel (zie paragraaf 'En hoe nu verder? Een Roadmap als kompas'). Deze exercitie heeft al tot een aantal interessante conclusies aangaande de kosten van digitale duurzaamheid.

Instellingen hebben nog een beperkt inzicht in de kosten. Gebruik van het model leidt tot inzicht in de kosten en daarmee tot bewustwording binnen de organisatie.

De eerste conclusie is dat instellingen nog geen tot een beperkt zicht hebben op de kosten van digitale duurzaamheid. De meeste instellingen staan in financieel-technische zin nog aan het begin van de ontwikkeling: alleen de grotere instellingen met meer ervaring hebben een relatief goed beeld van de kosten in relatie tot processen en bijbehorende cost drivers.

Daarnaast concluderen we dat het kostprijsmodel zorgt voor een beter inzicht in de kosten binnen de eigen organisatie. Het invullen van het model heeft onder de deelnemers bijgedragen aan bewustwording ten aanzien van het primaire proces en de financiële functie. Ook biedt het model managementinformatie om de grip te versterken.

Kosten in het proces van digitale duurzaamheid

Een groot deel van de kosten wordt gemaakt aan het begin van het verwerkingsproces (Selection, Pre-Ingest en Ingest). Wanneer de aangeboden data beter geselecteerd, geordend en beschreven zijn door de aanbieders of zorgdragers betekent dat ook dat er wellicht minder kosten gemaakt worden door de beherende instellingen.

We zien dat de kosten voor een groot deel betrekking hebben op de overkoepelende procesactiviteiten, wat te verklaren is door het (geconstateerde) feit dat veel instellingen nog aan het investeren zijn in duurzame toegang en de daarbij horende infrastructuur, waarmee afschrijvingskosten gepaard gaan. Hierbij merken we op dat instellingen dit vooral zelf doen en nog maar beperkt de samenwerking zoeken binnen en buiten de eigen sector.

Personele kosten vormen de belangrijkste kostencomponent. Deze bedragen bij de meeste onderzochte instellingen meer dan 50% van de totale kosten. Dit relatief grote aandeel maakt het belangrijk om kennis te ontwikkelen en te delen hoe de personele kosten beheerst kunnen worden en de inzet kan worden geoptimaliseerd.

Het kostprijsmodel en verdere doorontwikkeling

Uit de analyse van de sectoren en de deelnemende instellingen blijkt dat doelstellingen, context en achtergronden van de instellingen sterk verschillen. Voor het interpreteren van de cijfers in het kostprijsmodel is het belangrijk om deze doelstellingen, context en achtergronden mee te nemen en bij onderlinge vergelijking tussen instellingen erbij te betrekken. In de doelstellingen, context en achtergronden van de instellingen ligt de onderbouwing van de *cost drivers* en daarmee ook de verklaring

van de kosten. In de vergelijking wordt nu veelal gezocht naar overeenkomsten.⁷ Het is echter raadzaam om daarnaast ook te focussen op de verschillen binnen en tussen de sectoren en die op te pakken als leerpunten om de samenwerking te versterken.

Op dit moment hebben nog onvoldoende instellingen het kostprijsmodel ingevuld. Hierdoor is het nog lastig om een vergelijking van de kosten van duurzame toegang tussen verschillende instellingen te maken en om de correlaties en regressie te berekenen (relatie tussen kosten en *cost drivers*) en het effect daarvan te kunnen duiden. Hierdoor kan de potentie van het model nog niet volledig worden benut.

Toepassing van het kostprijsmodel door EYE

In deze paragraaf kijken we specifiek naar EYE⁸. EYE is als organisatie nauw betrokken geweest bij de ontwikkeling van het kostprijsmodel, als pilotinstelling en als instelling die het model heeft ingevuld en toegepast.

Terwijl EYE het model aan het toepassen was, speelde binnen EYE de kwestie dat de kosten voor digitalisering en vervolgens het beheer en de ontsluiting van digitaal erfgoed de afgelopen jaren waren toegenomen. Deze toename kwam boven op de kosten voor het beheer van de analoge collecties. In 2016 heeft het ministerie van OCW EYE een activiteitenanalyse laten uitvoeren; deze werd uitgevoerd door de Boston Consultancy Group (BCG)⁹ en was erop gericht om de structurele lasten en baten (en daardoorheen lopende incidentele baten en lasten) weer met elkaar in evenwicht te brengen en daarbij de impact van het groeiende digitaliseringsproces te duiden. Hierbij werd inzicht verkregen in 'de knoppen om aan te draaien' (= *cost drivers*) om de juiste strategische maatregelen te treffen ('*make or buy*'-beslissingen. Zie paragraaf 'En hoe nu verder? – een Roadmap als kompas'). Het Kostprijsmodel Digitale Duurzaamheid heeft EYE bij deze exercitie geholpen. Voor EYE was het project interessant om de kosten voor de gehele keten in kaart te brengen: Ingest-Storage-Access. Het kostprijsmodel verfijnde de door EYE eerder gemaakte berekeningen en kon daarmee BCG faciliteren voor het uitvoeren van de activiteitenanalyse.

Uit de ervaring van EYE met het Kostprijsmodel Digitale Duurzaamheid blijkt dat er inzicht is verkregen, maar ook een aantal kwesties overwonnen dienden te worden:

- De huidige processen waren niet ingedeeld conform Activity Based Costing (ABC); personeel en daarmee gepaard gaande kosten diende nog toebedeeld te worden.
- De financiële administratie sloot niet aan op ABC-costing.
- De projecten en daarmee gepaard gaande kosten dienden terug te komen in de directe kosten van de activiteiten (onderscheid incidenteel - structureel), evenals de bekostiging (onderscheid tussen incidentele (bijvoorbeeld subsidies) en structurele baten (bijvoorbeeld publieksinkomsten)).
- De gehanteerde definities dienden doorvertaald te worden naar de eigen organisatie en vice versa, om aan te sluiten bij de eigen administratie.

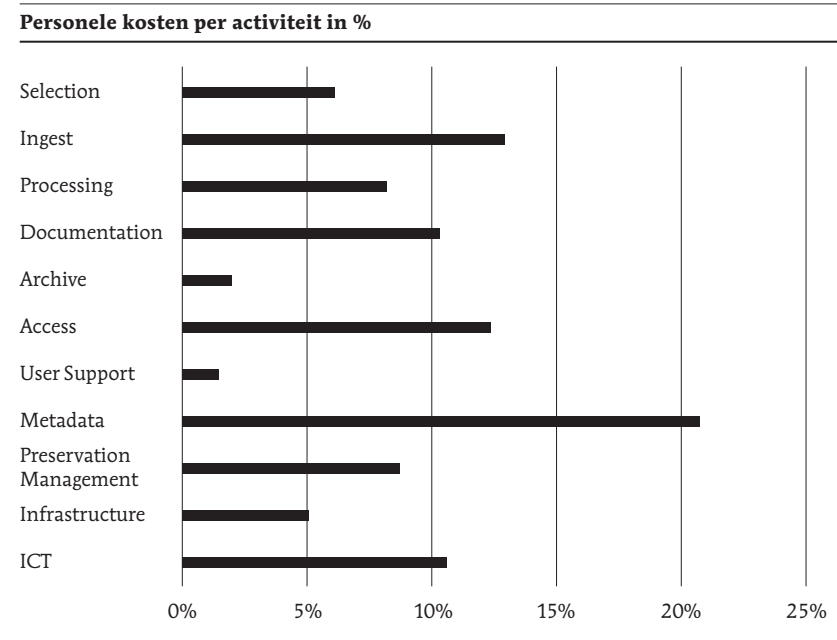
⁷ Van der Nat, J. & M. Ras, *Samen bouwen aan een netwerk van landelijke voorzieningen: Eindrapport van het onderzoek naar een landelijke infrastructuur voor duurzame toegang tot digitale informatie* (NCDD, april 2015) p. 1-80.

⁸ www.eyefilm.nl

⁹ Boston Consultancy Group, *Financieel Onderzoek EYE* (Ministerie van Onderwijs, Cultuur en Wetenschap & EYE, September 2016) p. 1-54.

- Voor overheadkosten (o.a. voor de gebouwen) dienden keuzes gemaakt te worden wat betreft de toerekening aan activiteiten.
- Toekomstige kosten voor toegankelijkheid van objecten zijn nog niet duidelijk; het kostprijsmodel geeft daar nog geen inzicht in.

Wanneer de personele kosten verdeeld worden over de activiteiten binnen het Kostprijsmodel Digitale Duurzaamheid, ontstaat voor EYE het volgende beeld:

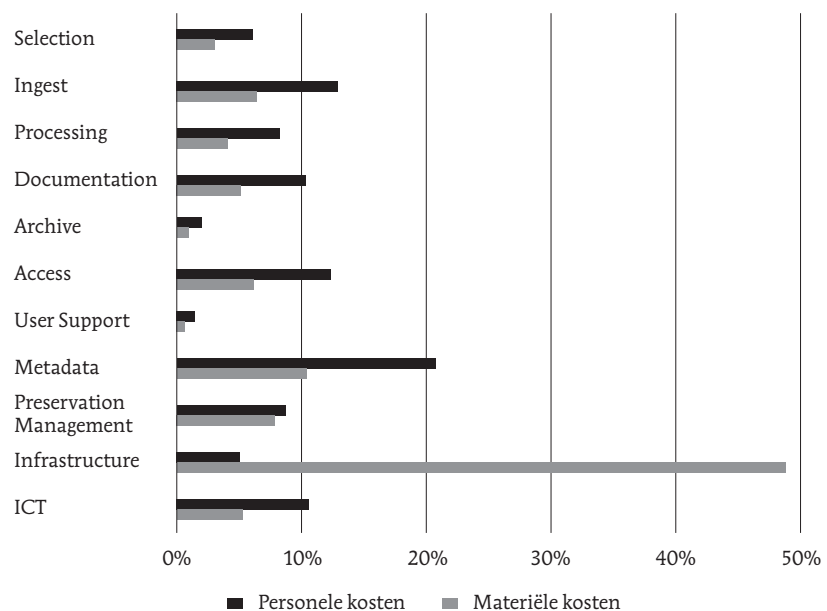


Afbeelding 2. Verdeling personele kosten EYE per activiteit in procenten van het Kostprijsmodel Digitale Duurzaamheid

In het proces van digitale duurzaamheid zitten de personele kosten voor EYE grotendeels in de activiteiten Ingest en Access. EYE investeert veel in de voorkant van het proces om de kwaliteit van het beheer en de ontsluiting van films zo hoog mogelijk te houden. Dit betekent dat er veel tijd gaat zitten in de activiteit Ingest, waarbij regelmatig handmatige handelingen moeten worden uitgevoerd (inzet van mankracht) om de bestanden te verrijken. Dit betekent ook dat de procesactiviteit Metadata qua inzet van personeel hoger ligt dan de andere activiteiten. EYE doet veel aan metadata van films, waarbij deze qua inhoud, creatieve makers, productie en rechtenstatus beschreven worden ten behoeve van ontsluiting en beheer. De technische metadata worden zo veel mogelijk geautomatiseerd aan het informatiesysteem toegevoegd. Daarnaast investeert EYE in de achterkant van het proces, dus in kosten voor het toegankelijk en beschikbaar maken van digitale data (Access).

Wanneer de materiële kosten ook inzichtelijk worden gemaakt, ontstaat het volgende beeld van de totale personele en materiële kosten (excl. inhuur en infrastructuur en ICT)) per activiteit en procesactiviteit van EYE:

Personele en materiële kosten per activiteit in %

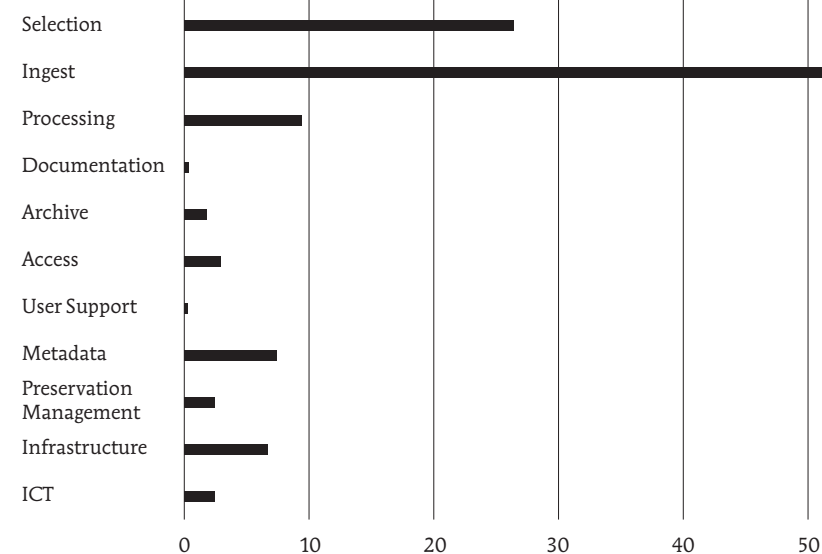


Afbeelding 3. Totale personele en materiële kosten EYE per activiteit en procesactiviteit in procenten van het Kostprijsmodel Digitale Duurzaamheid

Uit afbeelding 3 blijkt dat Infrastructuur relatief weinig personele kosten kent, maar relatief veel materiële kosten aangaande de aanschaf van hardware, software van het operating system, kosten dataverkeer en kosten van softwarelicenties. Hierdoor is Infrastructuur naast Metadata de activiteit met de meeste kosten binnen het gehele proces van Digitale Duurzaamheid. Bij EYE ligt het zwaartepunt van deze kosten dus bij Metadata en bij de technische Infrastructuur. Om de kosten beter beheersbaar te maken is het dan ook belangrijk om deze processen verder te optimaliseren.

Wat betreft de kosten per object zien we voor EYE het volgende beeld per activiteit en procesactiviteit:

Kosten per object EYE



Afbeelding 4. Kosten per object van EYE voor de activiteiten en procesactiviteiten

De verklaring hiervoor is dat de kosten vanaf de activiteit Documentation voor meer objecten van toepassing zijn (gehele collectie) dan tijdens de fasen Selection, Ingest en Processing (deze activiteiten vullen de collectie aan, waardoor ze per object duurder zijn dan de rest van de activiteiten in het proces van Digitale Duurzaamheid).

Gelet op de procentuele verdeling van de kosten constateren we dat 56,8% van de kosten van EYE behoort tot de procesactiviteiten en 43,2% tot de reguliere activiteiten van het proces van digitale duurzaamheid. ICT (8,5%) en Infrastructuur (23,2%) maken 31,8% van de totale kosten uit en Metadata 16,6%. Ingest is binnen de reguliere activiteiten de activiteit met relatief de meeste kosten, namelijk 10,4%. De voorkant van het proces (Pre-Ingest/Selection en Ingest) is in z'n geheel goed voor 15,4% van de kosten. De verhouding tussen personele en materiële kosten bedraagt bij EYE 58,9% - 41,1%. Storagekosten (activiteiten Documentation en Archive) vormen slechts een klein deel (8,3%) van de totale kosten.

Het kostprijsmodel heeft EYE inzicht opgeleverd in de verdeling van de kosten en geeft input voor de keuzes ten aanzien van de bekostiging van deze activiteiten. Ook stelt het EYE in staat om strategisch onderbouwde keuzes te maken aangaande Zelf Doen, Uitbesteden en Samenwerken, bijvoorbeeld ten aanzien van storage.

En hoe nu verder? Een Roadmap als kompas

Met het invullen van het ontwikkelende kostprijsmodel door de 10 instellingen is de eerste fase van de ontwikkeling van het Kostprijsmodel Digitale Duurzaamheid afgerond¹⁰. Het kostprijsmodel is inmiddels aangescherpt naar aanleiding van de ervaring en de eerder genoemde conclusies in dit artikel. De volgende stap is het realiseren dat het kostprijsmodel door meer instellingen wordt gebruikt als middel om de kosten van digitale duurzaamheid inzichtelijk te maken en er op te kunnen sturen. Tevens wordt het kostprijsmodel hierdoor door de instellingen ingevuld, waardoor het Netwerk Digitaal Erfgoed data verkrijgt over de kosten van digitale duurzaamheid bij instellingen. Op deze wijze kunnen (meerjarige) analyses worden uitgevoerd en kan de potentie van het model volledig worden benut. Ten aanzien van dat laatste doelen we op het verbinden van de cost drivers en de kosten. Met andere woorden, wat gebeurt er met de kosten wanneer aan de ‘knoppen’ (beleid) wordt gedraaid? Hierdoor wordt informatie verkregen voor het onderbouwen van strategische keuzes.

De volgende stap in de Roadmap is gericht op de ‘Verbinding Vraag & Aanbod’, waarbij het kortprijsmodel instellingen inzicht geeft in de verschillen en overeenkomsten van kostenstructuren in de sectoren door middel van een analyse per fase van het proces per sector en biedt daarmee tevens de onderbouwing om beslissingen te nemen op het gebied van Zelf Doen, Uitbesteden of Samenwerken en om vraag en aanbod met elkaar te verbinden. Ook worden instellingen en/of (markt)aanbieders uitgedaagd om de juiste proposities te maken om oplossingen binnen en buiten de sector te bieden. In de Roadmap van 4C¹¹ wordt deze fase ‘*Who should do what?*’ en ‘*Market efficiency*’ genoemd.

De laatste stap in de Roadmap is een opmaat voor het maken van MKBA's en het bepalen van het maatschappelijk rendement¹² van digitale duurzaamheid in de verschillende sectoren. Met andere woorden: wat is de toegevoegde waarde voor de

¹⁰ NCDD *Onderzoek naar de kosten digitale duurzaamheid* (BMC Onderzoek, Januari 2017) p. 1-68.

Zie specifiek hoofdstuk 5 ‘Roadmap Dutch Cost Model for Digital Preservation’ waarin de Roadmap wordt geschetst voor de ontwikkeling van het Kostprijsmodel Digitale Duurzaamheid. De volgende vier fasen worden onderscheiden:

- 1) Ontwikkeling van het Kostprijsmodel Digitale Duurzaamheid – A Dutch Cost Model for Digital Preservation;
- 2) Massa vergroten en borging van resultaten;
- 3) Verbinding van Vraag & Aanbod;
- 4) Opmaat voor MKBA's.

¹¹ 4C *Investering in Curation; A Shared Path to Sustainability* (20 februari 2015) p. 1-26.

¹² Maatschappelijk rendement is de toegevoegde waarde van een project voor de hele maatschappij.

Het geeft antwoord op de vraag hoe de kosten van een project zich verhouden tot de maatschappelijke baten, waarbij de baten datgene zijn waar je het, als organisatie in de sector, allemaal voor doet. Dit geeft inzicht in strategische beslissingen, bijvoorbeeld in het kader van Zelf Doen, Uitbesteden of Samenwerken (ZUS). Een MKBA is een methode om dit maatschappelijke rendement te bepalen, waarbij je op een systematische en samenhangende manier alle effecten die een project veroorzaakt in kaart brengt en vergelijkt met de situatie waar het project niet wordt uitgevoerd (het nulalternatief). De baten van het project worden, net als de kosten, uitgedrukt in euro's. Dit geldt ook voor de zaken die geen direct prijskaartje hebben, zoals geluids-overlast, een mooi uitzicht of je veilig voelen. Als de baten vervolgens hoger zijn dan de kosten, kan het project voor een grotere maatschappelijke welvaart zorgen. Het theoretisch karakter en de beperkte betrokkenheid van stakeholders is tegelijkertijd de zwakte van de MKBA. Ook leunt de MKBA in de praktijk nogal eens sterk op kengetallen en theoretische aannames, waardoor de uitkomsten impact en zeggingskracht kunnen missen. Ten slotte is het in zijn volwaardige vorm een arbeidsintensieve en kostbare methode (zie ook Ministerie van Binnenlandse Zaken en Koninkrijksrelaties *Werken aan maatschappelijk rendement; Een handreiking voor opdrachtgevers van MKBA's in het sociale domein* (oktober 2011) p. 1-73).

maatschappij en is het op het moment van besluiten waard om hierin te investeren? De basis voor het maken van MKBA's is het inzichtelijk hebben van de kosten en onderliggende cost drivers. Fase 4 geldt als stip aan de horizon om als Netwerk Digitaal Erfgoed in samenwerking met de gelieerde instellingen in de sectoren naartoe te groeien.

Het belang voor uw instelling

Uit de eerste conclusies rondom het ontwikkelen van het Kostprijsmodel Digitale Duurzaamheid geven aan dat instellingen nog veel kunnen winnen bij het inzichtelijk maken, het beheersen en vervolgens het actief sturen op de kosten van digitale duurzaamheid. Het ontwikkelende kostprijsmodel is een middel om het inzicht, grip en sturing op de kosten te versterken, waarbij een expliciete verbinding wordt gelegd met de beleidskeuzes die een instellingen maakt (=cost drivers) op het gebied van digitale duurzaamheid. Tevens ontstaat er door middel van het kostprijsmodel een mogelijkheid om instellingen onderling te vergelijken en daarbij te kijken naar overeenkomsten en verschillen ten aanzien van cost drivers en kosten. Het gesprek hierover geeft input om de onderlinge samenwerking tussen instellingen, binnen sectoren en bovensectoraal op het gebied van digitale duurzaamheid te versterken met als doel om het digitale erfgoed nog beter te verzamelen, te beheren en te ontsluiten en daarmee beschikbaar te stellen voor de gebruikers. Afgezien van de noodzaak tot verder onderzoek staat één ding vast: het belang van meer financieel inzicht in digitale duurzaamheid zal de komende jaren alleen nog maar toenemen.

Minding the Gaps: Digital Preservation then and now¹

It is a privilege to be invited to contribute to this volume on digital preservation in the Netherlands, not least because of the impressive impact that Dutch contributors and projects have had over the years. This short paper will review the experience of close neighbours at the Digital Preservation Coalition (DPC), identifying themes and gaps from the history of the DPC, and proposing directions for the next decade. In doing so, this paper contemplates common and distinctive challenges; and it renews the longstanding invitation to collaborate, which is a defining characteristic of the digital preservation community around the world.

A Secure Legacy?

In January 2018, the DPC published a new strategic plan.² There is much to commend this plan, establishing six interconnected programmes for the Coalition and providing a mandate from the community on whose behalf the DPC operates. It is not the purpose of this article to report or review that plan. Yet an easily overlooked detail makes a significant statement about the shared challenges and trajectories of digital preservation. The new plan sets out the DPC's mandate to 2022, the twentieth anniversary of our foundation. Therefore, hiding on the very front page is a paradoxical complement and reproach that might well summarize the experience of the global digital preservation community: its enduring commitment to collaboration and the surprising obduracy of the problem it seeks to resolve.

This puzzle might be phrased in more personal terms: the DPC was first proposed at a conference in 2000³ which also heard from the KB about the NEDLIB project⁴. I recall meeting a fellow archaeologist called Marcel Ras who had also found his way into the emerging field of digital preservation. The younger version of myself, not to mention the DPC's founders, would likely be astonished to know that almost twenty years later the digital preservation challenge still persisted, and perhaps would not

¹ I am grateful to Sara Day Thomson who reviewed this prior to release.

² Digital Preservation Coalition 2018 *A Secure Digital Legacy: The Digital Preservation Coalition 2018-2022* (2018) online at <https://www.dpconline.org/docs/miscellaneous/about/1755-dpc-strategic-plan-2018-22/> file last accessed 7/9/2018

³ Day, Michael 'Preservation 2000' in: *Ariadne 26* (2001) online at: <http://www.ariadne.ac.uk/issue26/metadata/> last accessed 7/9/2018

⁴ van der Werf-Davelaar, Titia, 'Long-term Preservation of Electronic Publications, The NEDLIB project' in *DLib Magazine 5.9* (1999) online at <http://www.dlib.org/dlib/september99/vanderwerf/09vanderwerf.html> last accessed 7/9/2018

perceive twenty years of the DPC to be something to celebrate. I certainly had a clear expectation that the digital preservation challenge would be solved in a few years and would have gone back to our day jobs. The question for my younger self and Marcel's younger self then, perhaps as now, was when would we return to the more familiar environment of archaeology?

Minding the gaps?

Digital preservation was very different in 2000 and, from that perspective, there's a lot to celebrate. The DPC's earliest mission was to propose actions that would resolve the problems that were evident. The findings of a needs assessment, published under the title 'Mind the Gap', provided a roadmap of recommendations for the community under eight broad headings which included 'growing awareness', developing 'repositories for all', and the need for the emergence of a 'new discipline'.⁵ Although there is a small but diverse body of literature about digital preservation from the first decade of the century, this report is a good summary of our expectations then and a useful benchmark for the progress we have made since. It is surprising how fresh some of these recommendations seem more than a decade later. For example, recommendation 5 speaks of the need for better tools 'to build a business cases for the long-term preservation of digital materials', while recommendation 11 calls for 'cross-disciplinary forums to allow both experienced individuals and organisations to exchange digital preservation best practice'.⁶ Such needs may be expressed differently today, but they remain true.

In other areas there is clear evidence of progress such as recommendation 10 that calls for digital preservation to be embedded in the training of librarians and archivists, and recommendation 17 which calls for more tools to perform digital preservation activities.⁷ There has been significant expansion of professional development and a transformation in the availability of tools and products. Thus, there are reasons to be positive that we have indeed 'minded the gaps'.

These early years of digital preservation were also marked by a parallel and perhaps less informed narrative that spoke to the consequences of data loss and the dysfunction that could follow. Coming hard on the heels of the Millennium Bug, gloomy predictions of a 'digital dark age' added urgency to our work.⁸ These compelling and eye-catching accounts were useful for generating headlines but were in many cases over-stated and, because they were overstated, may well have been ultimately self-defeating. Typically, the 'digital dark age' narrative did not anticipate the efforts of the digital preservation community and so it will never be evident whether such predictions were always fanciful, or whether they suffered from the 'observer effect', in which the simple act of anticipating a phenomenon changes its outcome.

⁵ Waller, Martin and Sharpe, Rob *Mind the Gap: Assessing Digital Preservation Needs in the UK*, (Digital Preservation Coalition, 2006) online at: <https://www.dpconline.org/docs/miscellaneous/advocacy/340-mind-the-gap-assessing-digital-preservation-needs-in-the-uk/file> last accessed 7/9/2018

⁶ Waller, Sharpe, *Mind the Gap* pp.36-37

⁷ Waller, Sharpe, *Mind the Gap* pp.37-38

⁸ e.g. Bergeron, Bryan *Dark Ages II: When the digital data die*, (Prentice Hall, 2001)

Perhaps more surprising however are the challenges not identified in the report. These might be termed the ‘gaps between the gaps’ that were not recognized and therefore not addressed. Of all the things not stated in the report, the actions of humans seem the most important and impactful. Digital preservation may indeed be a systemic problem and need systemic solutions: but it is also anthropogenic. In the very broadest sense, digital preservation is a problem for people, created by people and addressed through people. If that had been recognized sooner, then perhaps it would be less surprising to find that continued efforts are required. Arguably the digital preservation challenges we face now have been shaped by an emphasis on systems and solutions without fully engaging with the ever-changing dynamics of the human beings that are its cause, its solution and its ultimate purpose.

Obsolescence as a choice

The challenges with digital preservation are typically framed in technical terms, and though the wider context in which we operate is better understood, technology remains the dominant discourse within the literature. Issues like file format obsolescence, media degradation, emulation and the loss of representation information are significant concerns. As early as 2003, Nancy McGovern and Anne Kenny had observed that policy and resources were also critical concerns for digital preservation, which alongside technology, were described as the three legs of the ‘three-legged stool’ of digital preservation.

In the last decade, though, there has been more attention to the role of agency in what might be termed the causes of obsolescence. Working on the assumption that data loss is almost always the outcome of some embedded socio-economic structure that predispose specific historical contexts, then the causes of data loss are also to some extent amenable to control. If the early decades of digital preservation were an overt response to an emerging problem, then there has been a tendency for the digital preservation literature to exhibit ‘solutionism’⁹ in which problem-solving overtakes the subtle appreciation of how problems arise. The contexts of data loss are better understood now than they were in 2006 and with them the processes through which it would be possible to make obsolescence obsolete.

From an historical perspective it is perhaps significant that digital preservation came into focus shortly after the rise of the personal computer and during the rapid expansion of the World Wide Web as a platform for office and home computing in the late 1980’s and early 1990s (there are many accessible accounts, though see Kirschenbaum 2016¹⁰ for a case study of one popular technology). The 1980s and 1990s were the decades that, without which, digital technologies may have taken radically different directions. By extension, the economic forces and business environments that shaped the 1980s and 1990s created the norms of the digital universe. Had these forces been less consumerist, less disposable, more resilient, and more sustainable, then the endemic challenges of technical obsolescence, resource discovery and short-termism may not have arisen in the way that they subsequently

⁹ Morozov, Evgeny *To Save Everything Click Here: Solutionism and the Urge to Fix Problems that Don’t Exist* (Penguin, 2014)

have. By some strange process that is yet to be properly delineated, digital preservation came into existence as a response to the creative destruction implied within neo-liberal economics.

The specific contexts of data loss have come into focus over the last decade and so the scope of the of the digital preservation challenge has grown. The space once occupied by risks of obsolescence or media rot is now crowded also with concerns about ill-managed rights, out of control political interests, failing markets and simple human frailties. As confidence with technical challenges has risen, so there is a greater need to highlight the human behaviours behind data loss. The open nomination process for the ‘BitList’ of Digitally Endangered Species has been a significant milestone for the DPC, considering the nature of the challenge and how to address it (see Kilbride 2018¹¹ for an overview and introduction). Corporate abandonment, malicious deletion and ill-managed encryption have emerged as long-term threats to our digital memory in ways that were not fully recognized in 2006. Barely a week goes by without some new evidence of a duplicitous erosion, deletion or obfuscation by some rich or powerful agent who seeks to sanitise or erase uncomfortable narratives to suppress unwelcome truths or conceal historical facts. Less sensational but perhaps more harmful in the long run, is the impact of corporate failure as a risk to the digital estate, especially in the context of cloud computing. The subtle message in all of this is the need to situate digital preservation quite differently within our organizations and communities: as a counter balance to challenges more sinister than file-naming or bit rot. Obsolescence is not some pre-ordained, invisible force. It is no longer inevitable: and being avoidable, it can only coincide with some form of negligence.

Solutions and plans for their development were very much at the forefront of digital preservation when DPC was established and are laid out very succinctly in the Mind the Gap Report. As these solutions have become available so it has become obvious that the contexts of data loss are always to some extent a choice, and that no solution in the world could ever suffice on its own. The digital preservation community has exposed obsolescence as a convenient excuse and a lucrative business model. It’s not yet clear what can be done with this revelation but in the coming decades our role will be to ensure that responsibilities for preservation are more meaningful and more widely understood.

Digital Preservation as Community

The ‘Mind the Gap’ report was explicit about the need to encourage a new discipline of digital preservation and for training in new skills to become available.¹² It stopped short of calling for a new profession. Understanding how the digital preservation has progressed over the years encourages some thinking about the changing shape of the professional cohort that now delivers digital preservation.

By any measure that cohort is larger and more diverse than ever: a fact represented in some simple statistics from the recent history of the DPC. The Digital Preservation

¹⁰ Kirschenbaum, Matthew *Track Changes; a literary history of word processing* (Belknap, 2016)

¹¹ Kilbride, William *Sic Transit Gloriam Digitalis: BitList Beta* (DPC Blog, 2018) online at: <https://www.dpconline.org/blog/sic-transit-gloria-digitalis-the-bitlist-in-beta> last accessed 3/10/18

Handbook was launched in 2002 as the work of two authors: the 2016 edition credits thirty-three.¹³ The Digital Preservation Awards in 2010 had one winner: in 2018 there were 6. In 2009 the DPC had 2 staff and in 2018 there are 7. In 2009 DPC had 33 members: in 2018 there are 83. This growth is welcome, but it generates a certain amount of disruption. The community is diverse and the use cases for preservation tools are more demanding, the requirements more expansive and expectations more exacting. It may seem like a good time to move into the digital preservation business but only for those able to deal with these disruptive, dynamic and eccentric requirements.

This growth has not been a straight line and there are some significant challenges associated with it. Again, it is worth remembering that practical and concerted digital preservation really began in the late 1990s, coinciding with an unprecedented economic boom. It was such a long boom that economists and bankers congratulated themselves that this was the new norm. They were wrong: it turned out that significant and unsustainable assumptions were hidden in a financial sector that wasn't able or willing to prepare for the crash that would come. The question arises whether casually unsustainable assumptions embedded themselves into our plans for digital preservation in the early 2000's, too?

OAIS set a high standard for digital preservation in 2002.¹⁴ It has provided a shared language and some shared processes. It needs to be read through the prism of reasonable aspiration in 2000's: there is no explicit encounter with values or vision and too little expectation of changing context. *Mind the Gap* anticipated many of the technical and policy challenges but did not foresee the machinations that eviscerated the public sector. It's not a surprise that since 2010 it has become a lot more fashionable to talk about minimal effort ingest¹⁵, parsimonious preservation¹⁶ and 'Preserving Digital Objects With Restricted Resources'¹⁷. This is not simply about lack of resource: it's about competition for resources.

Seen in this context it is worth recognising that the digital preservation community has not been immune from the blunt trauma of economic conditions either. Consider, for example, the widespread redundancies and recruitment freeze that followed the banking crisis in 2008. Many agencies in the UK with an interest in digital preservation simply closed down in the turmoil that followed, including several significant DPC members and founders; and those that survived had little choice but to shed many of their previous complement of archivists, conservators and librarians. The normal employment cycle stalled and new generations of students with new skills were effectively locked out.

¹² Waller, Sharpe *Mind the Gap* p. 37

¹³ Digital Preservation Coalition *Digital Preservation Handbook* (2nd Edition, 2016) online at: <https://dpconline.org/handbook> last accessed 3/10/2018

¹⁴ Lavoie, Brian *The Open Archival Information System (OAIS): Introductory Guide 2nd Edition*, Digital Preservation Coalition Technology Watch Reports 14-02 (2014) online at: <http://dx.doi.org/10.7207/twr14-02> last accessed 3/10/18

¹⁵ Bolette, AJ, Blekinge AA and Christiansen, KF 'Minimal Effort Ingest', in *Proceedings of the 12th International Conference on Digital Preservation* (School of Information and Library Science, University of North Carolina at Chapel Hill, 2016) online at: <https://phaidra.univie.ac.at/o:429591> last accessed 3/10/18

¹⁶ Gollins, Tim *Parsimonious Preservation: Preventing Pointless Processes! A Simple Step that takes Digital Preservation a Long Way*, (Online Information 2009, 2009) online at <https://www.nationalarchives.gov.uk/documents/information-management/parsimonious-preservation.pdf> last accessed 3/10/18

This had a profound effect on individuals and it postponed the emergence of a new 'digital preservation' profession. Existing professional channels were reinforced instead, and salaries were fixed at levels congruent with library and archive posts, which further inhibited the recruitment or retention of in-demand developer skills.

The last decade has confirmed the earlier view that digital preservation is tricky. We've become used to the idea of working without all the resources needed to do all the things which seemed necessary in 2002. In many practical senses this generation are materially worse off now than when 'Mind the Gap' was published. David Rosenthal has observed that "Money turns out to be the major problem facing the future of our digital heritage".¹⁸ In 2006 it was all but impossible to anticipate the scale of the challenges that would arise.

Preservation, Access, Impact

The 'Mind the Gap' report makes almost no mention of users, except in the context of the exploitation of intellectual property. This seems the most significant gap of all.

For digital preservation, considerations about the users are mostly expressed via dissemination efforts, most immediately in the design and delivery of 'Dissemination Information Packages' and the 'Designated Community' (See Lavoie¹⁹ 2014 for fuller explanation). However, access and dissemination alone offer little justification for preservation: access is not an end in itself any more than preservation is. For example, the 'Blue Ribbon Task Force' examined the value of digital preservation insofar as it secures 'depreciable durable assets' which are 'long lasting and produce a flow of value through time'.²⁰ It is hard to think of a single case in which access alone brings benefits because value is constituted after access. Like preservation or interoperability, it is a necessary but not a sufficient condition for impact. There is no doubt that digital preservation needs to be configured around access: but access will only make sense if it is configured to enable the benefits that accrue from timely, dependable and supported use of data.

This may sound like a word play but it has significant implications for the practical digital preservation. It is easy to describe an access function, much harder to translate 'impact' into concrete actions within preservation plans. It gets even harder when we need to demonstrate the links between the implementation of a digital preservation plan and the outcomes in terms of impact and value that depend upon it.

¹⁷ POWRR *Preserving digital Objects With Restricted Resources* (Digital POWRR, 2012), online at: <http://digitalpowrr.niu.edu/> last accessed 3/10/18

¹⁸ Rosenthal, David *Storage Will Be A Lot Less Free Than It Used To Be* (DSHR Blog, 2012) online at: <http://blog.dshr.org/2012/10/storage-will-be-lot-less-free-than-it.html> last accessed 3/10/18

¹⁹ Lavoie OAIS (2014)

²⁰ Blue Ribbon Task Force *Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information, Final Report of the Blue Ribbon Task Force on Sustainable Digital Preservation and Access* (BRTE, 2009) p. 25

The foundations of this user-focussed and value-based approach are set firmly within the digital preservation standards. It is surprising that there was so little practical talk about how to test, predict or report changing user requirements and how to embed these within the ongoing delivery of digital preservation services. This oversight within 'Mind the Gap' seems to be repeated in the subsequent literature too. The energy with which organizational and technical challenges have been researched and resolved does not seem to be matched by an understanding of the user communities that repositories are established to support. Stated more formally, it is hard to track the impact of changing user needs on meaningful re-evaluations of the representation information required to ensure the independent utility of a digital object, nor is it clear that they represent an ongoing relationship with a designated community.

Digital preservation facilities that ignore users end up with two difficulties: their repositories simply won't work and the impact they seek to deliver is lost. If true, it is a significant risk to the relevance and vitality of the community. Is it true that the digital preservation community doesn't care about users enough to embed them into the day-to-day operation of its facilities? The apparent silence can be explained in perhaps two ways.

In some cases, digital preservation is embedded within institutions that already have robust and well-documented user bases, with feedback mechanisms to report changes and trends in user needs. That's certainly true of the larger memory institutions that already have a public profile and therefore need to manage user expectations. Such services may not be reported in the digital preservation literature because they are so well established that they are not of research interest, and partly because when they are described it is at different conferences and with different peers. In other cases, it may be that efforts are so concentrated on submission that digital preservation research just hasn't made it to users yet.

Each of these answers is plausible but neither really accounts for the silence about ongoing tests and assessments. There is a pattern to research but the need to capture and track the requirements of designated communities should not be not an afterthought in digital preservation architectures.

There is a further and perhaps more demanding challenge in the relationship between repositories and their users, embedded deeply within the digital preservation literature and most clearly articulated in the OAI: digital preservation technologies have yet to face up to the significant challenges of privilege and decolonialization which have arisen in the last decade. Again, considering the significant impacts that discourses of inclusion have had since 2006, it's perhaps not surprising that these subtle challenges were not obvious then.

One of the strengths of digital preservation has been its willingness to adopt tools and approaches from many different disciplines. For example, OAI is the product of the Consultative Committee for Space Data Systems. To some extent that origin haunts the language and assumptions of the model, and because OAI is the lingua franca of digital preservation generally, the values and norms of space science lurk below the surface of just about every digital preservation conversation.

This contribution has been immensely welcome and potent, but it also should be set alongside important trends in archival and museological theory which tend to the view that meaning-making can be hard and contradictory. The textual turn of cultural hermeneutics and poststructuralism in particular has been controversial, and it has been argued that the whole genre of post-truth informatics has some origin in the legacy of postmodernism.²¹ It has certainly been a mixed blessing for archives, libraries and museums.

On one hand, the recognition that knowledge production is a fundamental tool in the reproduction of power has transformed memory institutions from the gatekeepers of authoritative resilience to the enablers of progressive narrative(s). Derrida²² equated archives with a sort of house arrest: both as the source and containment of power, arranged to the practical convenience of the authorities, and only shared on asymmetrical terms with the public. It's no small accomplishment to note that for three decades now any number of disenfranchised communities have taken back control of cultural storehouses to establish new and often conflicting histories that subvert established norms and empower those previously excluded. Archives, libraries and museums have largely welcomed these new if at times unruly patrons on the assumption that if the epistemology of the institution is not fundamentally about justice then, by default, its purpose is to sustain injustice.²³

On the other hand, if signifier and signified are in permanent renegotiation, and if context is the last and only arbiter of meaning, then anyone can interpret everything to mean anything. That seems significantly more challenging in the context of a memory institutions where the absence of authorial voice intensifies the impossibility of authoritative meaning-making. In a crisis of relativism and self-congratulatory truth-making, where power is self-creating and context is fluid, what's the use of archives at all? The legacy of postmodernism could be summarized as follows: everyone empowered by their own narrative; and everyone empowered to deny everyone else's.

If the challenge to meaning-making began with postmodernism then it has been turbo-charged by technology. Francophone theorists of the 1960's and 1970's legitimated the challenge to legitimacy while anglophone engineers of the 1970's and 1980's delivered the machinery of change. The result: anyone can assemble their own history from the many ubiquitous sources that they chose not to ignore; they can publish it; and in so doing can find an audience to share their pain; and then all can live secure in self-made echo-chambers of reflexive half-truths repeated so often than they might as well be whole truths. That's becoming the all-too-common experience of social media and it's no wonder it's become fashionable to call it anti-social media.

The difficult history of meaning-making in the late 20th and early 21st century seems strangely at odds with the processes and norms adopted in digital preservation, especially with respect to representation information. Representation

²¹ D'Ancona, Matthew *Post Truth: The New War on Truth and How to Fight Back* (Ebury, London, 2017)

²² Derrida, Jacques and Prenowitz Eric *Archive Fever: a Freudian Impression* (Diacritics 25, 1995, pp 9-63) p. 10.

²³ O'Neill, Mark 'Essentialism, adaptation and justice: Towards a new epistemology of museums' in *Museum Management and Curatorship* 21, (2006) pp 95-116

information is component of an information that holds the semantic and structural components that proved the transparency required to ensure that a 'digital object' can become an 'information object' in the hands of a designated community. Arguably, representation information, is the unique characteristic that distinguishes digital preservation from every other kind of content management. Representation information is interpreted using representation information which, at face value, implies a sort of recursive absurdity. OAIIS avoids this in two ways: by linking representation information into networks; and by accessing an underlying knowledge base which we can take for granted. For example, a person who has a Knowledge Base that includes an understanding of English will be able to read, and understand, an English text. So the extent of representation information is mapped against implicit knowledge between two agents within an information exchange. Information objects are generated therefore through a mix of data object, representation information and implied knowledge.

This has important consequences for anyone attempting to challenge privileged narratives. If the link to an authoritative definition within a representation network is one of the keys to unlocking meaning, then whoever gets to assign that link or manage the end-point is a very important individual, a dependency that is open to abuse. That's even more problematic when one considers the implied knowledge that sits alongside representation information. How might the extent of implied knowledge be established for a preservation process that evolves over decades? There are two important themes here: the designated community and the mechanisms that measure changes within it. OAIIS assumes a special class of consumers that it defines as the designated community, a class that should be able to understand the preserved information. The designated community has a role to define (and thus set boundaries to) the extent of representation information. So long as the OAIIS charts and tracks that community through time, then representation is manageable.

This is self-evidently useful because the alternative is recursive absurdity. But reading that through the lens of three decades of cultural theory, it implies that, if you're not part of the designated community, you're not expected to use or understand the collection and the archive has no explicit responsibility to help you, and no requirement to listen to you. This might be true in the context of academic research where a relatively small but expert group of professionals would be expected to use complex datasets and would be motivated enough to cope with opaque documentation and annotations. It is altogether more concerning when identities, actions or meanings are in dispute: where honest misunderstanding may arise or faux conflicts be engineered and prolonged.

At present, best practice in the digital preservation community means that the digital archivist is empowered, in fact is required, to exercise a kind of intellectual exclusion that is out of step with just about every other kind of memory institution. In summary, archives, libraries and museums have spent 30 years coming to terms with inclusion and polysemy, challenges that were barely considered by the digital preservation community in 2006.

Digital preservation: Community not Process

This article started by wondering why the role and mission of the digital preservation community has remained relevant for so long, and in particular, if anything can be gleaned from the experience of the Digital Preservation Coalition that might inform the Dutch experience of digital preservation. Surely the problem has been fixed? Surely we can get back to our real jobs?

The answer to both questions, for better or worse, seems to be an emphatic no, at least in the context of the UK and Ireland where DPC was founded, and almost certainly in the context of the Netherlands, too. Three themes emerge, which will certainly influence the direction of the DPC and which, because they are universal challenges, will invite ever closer collaboration with colleagues around the world.

Firstly, perhaps most obviously, the nature of the technical preservation challenge continues to change and therefore the tools and approaches that have emerged need constant renewal and revision. This need for renewal and revision creates a tension for information managers who seek to ensure the integrity of their own processes. Understanding how to update and automate procedures without compromising them is a key challenge in the years ahead.

The community has been conspicuously successful in the provision of tools and processes for data management, even if there is more to do. It has been less successful at embedding the outcomes of research into practical workflows and spent too little time considering how data loss arises in practical contexts on a daily basis. These processes, not purely technical, will endure for as long as agencies, corporations and individuals are able to shirk or defer requirements for preservation that enable reasonable expectations of transparency and authenticity. We may never make obsolescence obsolete, but it is certainly possible to identify those for whom data loss is a convenient excuse of lucrative business model.

We have also been less effective in involving users in our discussions, especially on selection, preservation planning and representational information. A sustained and informed dialogue between repositories and their audiences is required. If it is possible to capture and embed the potential of the data then such changes could have a transformative effect. Digital preservation is not done for the good of the data, but for audiences and communities that have hitherto been all but absent from out literature. The coming challenge is to configure our processes around the needs of people and the opportunities they seek to exploit.

In 2006 the DPC conceived of digital preservation as a series of gaps to be avoided through timely interventions and adaptation of process. What's become apparent since then is that digital preservation is not a process, it's a community. It's time to stop minding the gaps: it's time to start seeing the opportunities.

Een wereldbeeld vergruizeld: uitdaging voor archiveringsdeskundigen

Inleiding

In een inleidend artikel heb ik een overzicht gegeven van de belangrijkste ontwikkelingen van de laatste circa dertig jaar op gebied van digitale bewaring/preservering.¹ Daaruit kwam het beeld naar voren van langzame maar zekere vooruitgang in denken en doen op dit gebied. In dit artikel wil ik enkele specifieke ontwikkelingen nader belichten en deze in wat bredere context plaatsen. In deze korte terugblik zal de nadruk (weer) liggen op archieven. Op twee aspecten in het bijzonder wil ik nader ingaan. Dat betreft de vraag of er wat 'preservering' betreft een nieuw vakgebied is ontstaan en wat dat dan zou inhouden, en ten tweede de vraag wat de impact van de nieuwe ontwikkelingen is op de archiefdiscipline ofwel welke reacties heeft dat daar opgeroepen, waarbij ik Nederland als voorbeeld neem.

Nieuw vakgebied?

Eén van de vragen die opkomt, is of er in de afgelopen 30 jaar een nieuw vakgebied is ontstaan? En vervolgens wat dit dan zou inhouden? Het antwoord is niet simpel te geven. Iedereen gebruikt de termen *digital preservation* of *curation* of in dit boek 'preservering', maar het is de vraag of men er hetzelfde beeld bij heeft. Dat is in de regel slechts een impliciete aanname in discussies. Het begrip zelf lijkt nog niet uitgekristalliseerd en kent verschillende benamingen.

¹ Zie Hans Hofman, 'Een wereld van verschil' (2018) Elders in dit jaarboek.

² Die luidt: "field of management responsible for the efficient and systematic control of the creation, receipt, maintenance, use and disposition of records, including processes for capturing and maintaining evidence of and information about business activities and transactions in the form of records" (definitie 3.15).

³ Zie: https://www.nationaalarchief.nl/archiveren/nieuws/preservation-watch-bij-het-nationaal-archief#_ftn2 geraadpleegd 11 oktober 2018.

⁴ In het Nederlandse archiefwezen worden veel Engelstalige termen gebruikt. In het eerste SAP-jaarboek, Peter Horsman, Eric Ketelaar en Theo Thomaasen (red.), *Naar een nieuw paradigma in de archivistiek* ('s-Gravenhage 1999), is een vertaling van Engelse termen opgenomen die voor een belangrijk deel nog geldt. Zie aldaar Peter Horsman, 'Engelstalige archief termen in het Nederlands vertaald', bijlage, p.231-242. Termen als *digital preservation* en *digital curation*, als zijnde geen specifieke archief termen, komen overigens daar niet in voor. Het is sowieso lastig Engelstalige termen in het Nederlands te vertalen. Nederlandse archiefbegrippen kunnen een subtiel andere lading hebben dan die in de Engelse taal.

Terminologie en definities

In dit boek wordt 'preservering' (en niet bijv. 'digitale conservering of bewaring') gebruikt zonder een duidelijke definitie. Zoals al aangegeven levert deze slechte afbakening onduidelijkheid op, want waar gaat dit nu over? De ruime definitie die bijvoorbeeld door het Nationaal Archief (NA) wordt gehanteerd voor 'preservering', overlapt grotendeels met die van ISO 15489:2016 onder de term *records management*: "Het op zodanige wijze vastleggen, bewaren, beheren en beschikbaar stellen van digitale informatieobjecten, dat deze ook na verloop van tijd raadpleegbaar, toegankelijk en authentiek zijn."³ Een niet onbelangrijk verschil is wel dat men in bovenstaande definitie spreekt over digitale informatieobjecten, terwijl de ISO 15489 spreekt over *records*. Dat raakt direct aan de kern van het archiefvak, want zijn die objecten en archiefbescheiden een en hetzelfde? Ik kom daar later nog op terug.⁴ Het begrip 'preservering' wordt bovendien gebruikt zonder het adjectief 'digitaal', waarmee het nog minder zegt en verwarrend kan zijn. Wat is immers het verschil met (papier)restauratie en conservering?

De bovengenoemde definitie komt ook in de buurt van de in sommige kringen gebruikte term 'digital curation', zoals uitgelegd in het handboek van Ross Harvey, waarin het gaat om "...the entire digital curation lifecycle, from creation to appraisal to preservation to organization/access to transformation...".⁵ De term 'digital preservation', die ook veel in Nederland wordt gehanteerd, kent een vergelijkbare definitie: "Digital preservation combines policies, strategies and actions that ensure access to digital content over time".⁶

Wat ontbreekt is de 'entire lifecycle', maar de vraag is of dat werkelijk onderscheidend is? In het Digital Preservation Handbook van DPC staat als omschrijving: "... the series of managed activities necessary to ensure continued access to digital materials for as long as necessary".⁷ Voorlopig is daar geen duidelijk onderscheid te onderkennen. De definities zijn ruim geformuleerd en zeggen in feite weinig.

Het beeld dat hieruit naar voren komt, is dat het begrip *digital preservation* in het algemeen gebruikt wordt voor de digitale versie van het bewaren en beheren (zonder onderscheid) van archiefbescheiden, boeken, datasets en andere te beheren informatie, inclusief alle beheer- en bewaaractiviteiten met betrekking tot digitale informatie. De term *digital curation* lijkt daar niet wezenlijk van af te wijken.⁸

In theorie valt daar ook de creatie onder. Naast de vraag in hoeverre de organisaties, die deze begrippen hanteren (vaak erfgoedinstellingen), de creatie kunnen beïnvloeden, is er de vraag waar men het over heeft of wat men precies bedoelt als men over

⁵ Gillian Oliver and Ross Harvey, *Digital Curation* (2016, 2d edition) zie <http://www.facetpublishing.co.uk/title.php?id=300976#.W3KpDC2xVyo> Zie verder de definitie op de website van DCC: *Digital curation is maintaining and adding value to a trusted body of digital research data for current and future use; it encompasses the active management of data throughout the research lifecycle* Via <http://www.dcc.ac.uk/about-us/dcc-charter/dcc-charter-and-statement-principles> geraadpleegd 11 oktober 2018. Het is hier beperkt tot 'research data'.

⁶ Zie onder andere: <http://www.ala.org/alcts/resources/preserv/defidigpres0408> geraadpleegd 11 oktober 2018.

⁷ Digital Preservation Coalition *Digital Preservation Handbook* (2015, 2d edition) online versie zie <https://www.dpconline.org/handbook/glossary#P>.

⁸ Het Digital Curation Centre (DCC) lijkt de term te reserveren voor digitale bewaring van onderzoeksgegevens (research data). Zie <http://www.dcc.ac.uk/about-us/dcc-charter/dcc-charter-and-statement-principles> geraadpleegd 11 oktober 2018.

objecten of informatie spreekt? Een niet onbelangrijk en vaak onderbelicht aspect in dit alles. Zijn dat de informatie-eenheden, zoals boeken of documenten, die men op het scherm wil zien of zijn dat de digitale bestanden die die informatie in gecodeerde vorm bevatten? Het gaat uiteraard om beide, maar in het alledaagse spraakgebruik lopen deze twee door elkaar en worden de intellectuele eenheden te gemakkelijk gelijkgesteld aan de digitale bestanden. De ‘digitale objecten’ zijn bijvoorbeeld niet hetzelfde als de archiefbescheiden (anders gezegd: wat je hebt of bewaart, is niet wat je ziet).⁹ Dat lijkt nogal eens te worden vergeten.¹⁰ De term ‘object(en)’ heeft al gauw een fysieke connotatie. Het is uiteindelijk echter, wat ik noem het *intellectuele aspect* (dat wat de mens kan lezen en begrijpen), dat uitgangspunt en bepalend is bij het lange termijn bewaren en beheren van informatie. Hoe het fysiek is gecodeerd, is secundair. In dat opzicht is het noodzakelijk een duidelijker onderscheid te maken tussen het feit dat digitale objecten (of beter digitale bestanden) fysiek bewaard worden en dat informatie-eenheden, zoals boeken, digitale kunst of archiefbescheiden, ‘intellectueel’ beheerd worden. Daarbij spelen metadata een cruciale rol en ligt de nadruk op de processen van creatie, beheer en gebruik van archiefbescheiden. De metadata omschrijven wat er bewaard wordt (zowel fysiek als op intellectueel niveau), hoe, waarom en in welke context de informatie ontstaan en gebruikt is (en wordt), welke actoren erbij betrokken waren (en zijn), relaties met andere informatie (structuur), en beheerinformatie over wat er met de digitale bestanden gedaan is en moet worden. Omdat de intellectuele verschijningsvorm van archiefbescheiden (of procesgebonden informatie) niet, zoals bij papier gefixeerd is, maar vluchtig en elke keer opnieuw moet worden ‘geconstrueerd’ en gepresenteerd, is het bovendien noodzakelijk om de essentiële kenmerken, inhoud, structuur, verschijningsvorm, context en eventueel gedrag, vast te leggen of te beschrijven. Met betrekking tot ‘objecten’ kan ook nog opgemerkt worden dat deze niet langer de welbekende documenten zijn maar dat dat allerlei andere vormen van informatie kan omvatten, zoals websites, tweets, whatsapp-jes, sms-jes en wat er verder nog aan variaties hierop bestaat, als ook in toenemende mate data en datasets. Informatievormen waar archivariissen en DIV-ers niet veel affiniteit mee hebben.

Afbakening vakgebied

Terug naar de vraag of er een nieuw vakgebied is ontstaan? Afgaande op functie-benamingen zoals *preservation officer*, *digital archivist*, *digital forensic analyst*, of op het bestaan van een referentiemodel als OAIS, lijkt het er wel op. Hoe verhoudt dit zich dan tot de bestaande disciplines, zoals bibliotheekwetenschap, archiefwetenschap, data management e.d.? Die verdwijnen immers niet. Er is namelijk wel duidelijk onderscheid te maken tussen de methoden, concepten en principes van iedere discipline en dat geldt zeker voor archivering en archiefbeheer. Het volgen van die

⁹ Dit onderscheid wordt evenmin scherp gemaakt in het algemeen gebruikte OAIS-referentiemodel. Wel wordt daar gesproken over *information packages*, waarmee de digitale data worden aangeduid. Van digitale objecten bestaan ook weer verschillende definities. Een uitgebreide is bijvoorbeeld te vinden als: “*Digital objects are marked by a limited set of variable yet generic attributes such as editability, interactivity, openness and distributedness.*” zie Jannis Kallinikos, Alekse Aaltonen en Attila Marton, ‘A theory of digital objects’, in: *Firstmonday*, vol.15, nr.6-7 (June 2010) via <http://firstmonday.org/ojs/index.php/fm/article/view/3033/2564>. Een digital object kan worden opgevat als alle informatie-eenheden die digitaal zijn, zie bijvoorbeeld Margaret Hedstrom en Christopher A. Lee *Significant properties of digital objects: definitions, applications, implications* (Brussel, 2002) via https://ils.unc.edu/callee/sigprops_dlm2002.pdf of als het digitale gegevensbestand en alles er tussenin. Binnen elke discipline lijkt er een (iets) andere definitie van te zijn. Geraadpleegd 11 oktober 2018

concepten en principes is cruciaal. Zonder dat heeft een discipline geen bestaansrecht meer of verdwijnt deze. Die disciplines of functies zijn echter niet fundamenteel veranderd met de komst van computers.

Het gebruik van enkele vergelijkbare termen voor het beheer van digitale informatie dwars door alle disciplines heen geeft niettemin aan dat er een gemeenschappelijk terrein is tussen die disciplines. Voortbordurend op het eerder gemaakte onderscheid tussen fysiek en intellectueel ligt dat gemeenschappelijke voor het grootste deel op het technische vlak, de soft- en hardware omgeving nodig om (fysieke) digitale ‘objecten’ te beheren en in stand te houden. Daarnaast is er de inrichting van de nieuwe processen waarmee het digitale beheer wordt uitgevoerd.

De snelle ontwikkelingen in de informatietechnologie maken dat de digitale bestanden weinig stabiel zijn en actief en continu moeten worden onderhouden. Het grote punt daarbij is dat dat geen effect mag hebben op de intellectuele entiteit. De leesbaarheid en raadpleegbaarheid moeten gegarandeerd zijn. De activiteiten en processen die hierin een rol spelen hebben te maken met o.a. *preservation planning*, toepassen van bewaarstrategieën (emulatie, migratie, normalisatie, replicering), validatie, informatiebeveiliging, calamiteitenplanning, en *bit preservation*.

Als er al sprake is van een nieuw vakgebied dan ligt het voor de hand om het begrip ‘preservering’ (of wellicht digitale conservering) te beperken tot die activiteiten die noodzakelijk zijn om de digitale informatie in technische zin in stand te houden zonder verlies van essentiële informatie en met behoud van authenticiteit, bruikbaarheid en integriteit van de archiefbescheiden (of procesgebonden informatie). Essentieel daarbij is dat al bij creatie rekening wordt gehouden met de onderhoudbaarheid van de informatie in technische zin, omdat het achteraf ‘corrigeren’ van digitale objecten in de regel lastig, zo niet onmogelijk, en altijd kostbaar is. Aangezien er met het aspect onderhoudbaarheid van de digitale informatie tot nu toe vrijwel geen rekening wordt gehouden bij de ontwikkeling van informatiesystemen, kan dat voor de lange termijn bewaring problemen opleveren. Het vraagt soms forensische vaardigheden om informatie weer leesbaar te maken.¹¹ Dit alles overziend zou men kunnen stellen dat voor de bewaring van digitale informatie op lange termijn op dit technologische niveau met e-depots, bewaarstrategieën, opslagtechnieken, controlemechanismen e.d. een apart vakgebied is ontstaan met zijn eigen vaardigheden. Daartoe behoort ook de testbedomgeving, waar o.a. nieuwe migratiestrategieën kunnen worden getest als ook de ‘productie-omgeving’, waar de feitelijke gekozen bewaarstrategie wordt uitgevoerd, gecontroleerd en gevalideerd. Dit alles moet worden ondersteund met curricula, opleidingen en cursussen waar men in de geheimen van ‘digitaal lang leven’ wordt ingewijd.¹²

¹⁰ Een voorbeeld hiervan is de *Preservation policy* (2016) van het Nationaal Archief, zie <https://www.nationaal-archief.nl/archiveren/kennisbank/preservation-policy> p.10, waar digitale objecten (ten onrechte) worden gelijkgesteld met archiefbescheiden zoals genoemd in de Archiefregeling 2009. In de *Nationale strategie digitaal erfgoed* (2015) wordt ‘digitaal object’ omschreven als de volledige content in digitale vorm, zie http://www.den.nl/art/uploads/files/Publicaties/Nationale_Strategie_Digitaal_Erfgoed_MinOCW.pdf p.11. Grote vraag is dan wat precies met ‘volledige content’ wordt bedoeld? Metadata in ieder geval niet, want die vallen onder de noemer ‘beschrijvingen en verwijzingen’ (=informatie over analoge en digitale objecten zoals (proces)metadata en collectieregistraties). Slordig en verwarrend gebruik van termen.

¹¹ Dit wordt ook wel digital archeologie genoemd. Zie Seamus Ross and Anne Gow, *Digital Archeology. Rescuing neglected and damaged data resource*, JISC/NPO study, (Glasgow 1999) via <http://www.ukoln.ac.uk/services/elib/papers/supporting/pdf/p2.pdf> geraadpleegd 11 oktober 2018.

¹² Zie bijvoorbeeld DigCCURR Institute <https://ils.unc.edu/digccurr/institute.html> geraadpleegd 11 oktober 2018.

De nieuwe technologie maakt verder ook automatisering van beheerprocessen mogelijk en vormt daarmee niet alleen een belangrijk hulpmiddel, maar is gezien de snelle technologische ontwikkelingen ook een beheeraspect op zich. De gebruikte beheerssoftware dient, naast de uitgevoerde activiteiten en handelingen, zelf ook goed gedocumenteerd worden om verantwoording over het beheer te kunnen afleggen. Wat zijn de algoritmes waarmee gegevens worden verwerkt en bewerkt? Digitale depots spelen een belangrijke rol, maar anders dan in het verleden toen de archiefbewaarplaatsen voor papier tamelijk statisch waren, zijn digitale 'depots' (in feite een bepaald soort informatiesystemen) permanent aan verandering onderhevig.¹³ Immers de software, waarmee de digitale bestanden en data beheerd worden, ontwikkelt zich steeds verder en dat geldt in feite ook voor de hardware en opslagmedia. Zodra een e-depot wordt opgeleverd, is het vaak al weer aan herziening onderhevig. Hetzelfde geldt voor standaarden voor bijv. opslagformaten. Deze veranderen eveneens in de loop der tijd en dat heeft dan weer invloed op de software. Het maakt het beheer complexer.¹⁴

Om dit alles in goede banen te leiden en te sturen zijn er kaders, standaarden, modellen, methoden en richtlijnen ontwikkeld, die steeds worden aangepast aan de nieuwste ontwikkelingen.

Zeker is dat de uitdaging om digitale informatie in welke vorm dan ook voor de lange termijn op een betrouwbare en authentieke wijze te bewaren en toegankelijk te houden de verschillende disciplines, zoals bibliotheken, archiefdiensten, onderzoeksinstituten, musea, universiteiten, datacentra, nader tot elkaar heeft gebracht en veel samenwerking heeft geïnitieerd, zowel nationaal als internationaal.¹⁵

Ontwikkelingen op gebied van archivering

Het tweede punt dat ik nader wil bekijken is de vraag: hoe is er binnen het vakgebied van archivering en archivalistiek gereageerd op de ingrijpende digitale ontwikkelingen of anders gesteld welke impact hebben die ontwikkelingen op het vakgebied gehad? Duidelijk is dat het vakgebied vanaf ongeveer midden jaren '80 sterk in beweging is gekomen en het eind is nog niet in zicht. Het begon in de tweede helft van de jaren '80 met een relatieve buitenstaander, David Bearman, die de traditionele archivalistische methoden ter discussie stelde.¹⁶ Zijn stelling was dat deze methoden niet langer geschikt waren om de toenemende hoeveelheden informatie te be- en verwerken.

¹³ De term digitaal depot (of e-depot) wordt hier in engere zin dan Trusted Digital Repository (TDR) gebruikt, zie onder andere *Trusted Digital Repositories: Attributes and Responsibilities* (RLG-OCLC report, Mountain View 2002) via <https://www.oclc.org/content/dam/research/activities/trustedrep/repositories.pdf> of zoals beschreven in ISO 16363:2012 *Space data and information transfer systems - Audit and certification of trustworthy digital repositories*. In dat begrip wordt de beheerende instelling als zodanig aangeduid (incl. beheersystemen), terwijl hier het eigenlijke informatiebeheersysteem wordt bedoeld.

¹⁴ Uit een onlangs gehouden enquête door Netwerk Digitaal Erfgoed komt een nogal divers beeld naar voren van instellingen die zeggen 'digitaal archief' te hebben en te beheren. Zie Joost van der Nat en Marcel Ras, *Survey Digitale Archieven in Nederland*, (Den Haag, 2018). 'Digitaal archief' is hier in ruime zin gebruikt en niet beperkt tot archief in enge zin. In het overzicht wordt verder helaas geen onderscheid gemaakt tussen gedigitaliseerd (gescand) en 'digital born' materiaal. Interessant is dat een aantal instellingen zegt alle functies uit OAIS te hebben geïmplementeerd. Gezien de huidige stand van zaken, die nog in ontwikkeling is en alles behalve uitgekristalliseerd, roept dat de vraag op 'is dat echt zo?', en 'in welke mate en hoe is dat gedaan?'

¹⁵ In Nederland is daar het Netwerk Digitaal Erfgoed een voorbeeld van, zie http://www.den.nl/art/uploads/files/Publicaties/Nationale_Strategie_Digitaal_Erfgoed_MinOCW.pdf geraadpleegd 11 oktober 2018

Dat was een steen in de vijver van gezapigheid en leidde tot veel discussie. Al snel kwam daar de toenemende impact van de digitalisering op het vak in de dagelijkse praktijk bij en de betrekkelijke rust in archiefland was voorgoed verdwenen, zodanig zelfs dat sommigen sinds enige jaren spreken van een crisis.¹⁷ De argumentatie daarvoor is dat archiveringsspecialisten, als ze er al zijn, vaak geen rol spelen in de vormgeving van werkprocessen en de informatievoorziening van (overheids)organisaties en dat er sinds de vaststelling van David Bearman (te) weinig is gebeurd om nieuwe methoden te ontwikkelen. En niet alleen dat, om de huidige ontwikkelingen op gebied van informatie en de daaruit voortvloeiende complexe wereld vanuit het gezichtspunt van archivering aan te kunnen, is er een stroming die stelt dat een geheel andere benadering nodig is, die recordkeeping informatics wordt genoemd.¹⁸ Ik kom daar nog op terug.

Een aspect van die crisis is wellicht ook de neiging van DIV specialisten en archivariissen in Nederland (maar niet alleen daar) vanaf ongeveer eind jaren '80 om niet langer over archiefbescheiden te spreken, maar over (het vage en brede begrip) informatie dan wel informatie-objecten. Daarmee wordt in feite de kern van het vak ontkend. Een andere, meer adequate term die in de jaren '90 ontstond, is 'procesgebonden informatie'.¹⁹

In deze paragraaf wil ik allereerst nader ingaan op twee belangrijke 'denkrichtingen' binnen de archiefwereld die een belangrijk stempel op de vele discussies over de gevolgen van de digitalisering hebben gedrukt, en vervolgens op de mogelijke impact ervan op de dagelijkse praktijk van 'preservering' als ook die van archivering in Nederland.

De twee bedoelde stromingen kunnen direct dan wel indirect gezien worden als een reactie op de digitalisering. De eerste grijpt terug op oude, bestaande kennis, zoals in het InterPARES project, waarbij men de methode en principes van de diplomatiek, met als kern authenticiteit van documenten, probeert te vertalen naar de 'digitale wereld'. De focus ligt daarbij op het object: de archiefbescheiden of 'electronic records'.

Het eerste InterPARES project (1999-2001) richtte zich op de lange termijn bewaring van 'statische' (afgesloten) archiefbescheiden met onderzoeksvragen over authenticiteit, selectie en bewaring, waarvan de beantwoording moest leiden tot een beleidskader en richtlijnen. Het project leverde onder andere een functioneel model voor

¹⁶ Zie David Bearman *Archival methods* (1989) http://www.archimuse.com/publishing/archival_methods/geraadpleegd_11_oktober_2018. Bearman heeft daarna nog vele andere vaak provocerende artikelen geschreven met name in relatie tot 'electronic records'. En verder bijvoorbeeld de reflecties van Cassie Findlay in 2014, <https://rkroundtable.org/2014/02/05/reinventing-archival-methods-in-the-hague/> en in 2018 over de herziening van ISO15489: <https://cassiefindlay.com/2018/03/12/crunch-time-the-revised-iso-15489-and-the-future-of-recordkeeping/>

¹⁷ Zie o.a. Frank Upward, Barbara Reed, Gillian Oliver en Joanne Evans, 'Recordkeeping Informatics: Re-figuring a discipline in crisis with a single minded approach', in: *Records Management Journal*, Vol.23, Issue 1 (2013) p. 37-50 via <https://emeraldinsight.com/doi/abs/10.1108/09565691311325013> geraadpleegd 11 oktober 2018.

¹⁸ Met deze nieuwe term wil men een brug slaan tussen records managers (DIV), archivariissen en informatiesystemen zonder het verwarrende gebruik van oude, engelstalige terminologie zoals recordkeeping (één woord) en record keeping (twee woorden). Een Nederlandse vertaling ervan is lastig te vinden. Het dichtst in de buurt komt: *archivistische informatiewetenschap*.

¹⁹ Zie bijvoorbeeld Theo Thomassen, 'Een korte introductie in de archivalistiek', in: P.J. Horsman, F.C.J. Ketelaar en T.H.P.M. Thomassen *Naar een nieuw paradigma in de archivalistiek* (SAP-jaarboek 1999), p. 11-20 zie <https://kvan.courant.nu/periodicals/JB/1999> geraadpleegd 11 oktober 2018.

digitale bewaring ('*preservation function*') op.²⁰ In het vervolgproject (2002-2007) lag de nadruk op verschillende typen 'electronic records', met name '*experiential, interactive and dynamic records*', daarmee de snelle ontwikkelingen in de nieuwe technologieën onderkend.²¹ Op basis van case-studies werden richtlijnen voor het maken en beheren van digitale archiefbescheiden geformuleerd. Daarnaast werd gewerkt aan een register van metadata schema's (MADRAS) en metadata specificaties en werden tenslotte twee functionele modellen uitgewerkt. Eén model richtte zich op de langdurige bewaring, de andere op de integratie van recordkeeping in het bedrijfsproces met ook aandacht voor de bewaarfunctie.²² Beide projecten hadden een sterk multidisciplinair karakter.

Naast deze discussies over traditionele archivalistische methoden en over hernieuwde toepassing van diplomatische methoden op digitale archiefbescheiden ontstond er in de tweede helft van de jaren '90 de records continuum theorie met een meer fundamentele verandering van perspectief, door sommigen al snel een paradigma-verschuiving in de archivalistiek genoemd.²³ Die tweede stroming is niet direct ontstaan als reactie op de gevolgen van digitale technologieën, maar bouwt onder andere voort op de wijze waarop in Australië werd gearchiveerd, het zgn. 'series system'. Met die achtergrond kijkt deze nieuwe denkwijze niet alleen naar de archiefbescheiden, maar tegelijk ook naar de actoren, processen en doelstellingen die erbij betrokken zijn, als mede de verschillende aggregatieniveaus.²⁴ Elk biedt een zelfstandige, andere invalshoek en daarmee beter inzicht in de factoren die komen kijken bij ontstaan, beheer en gebruik van archief.

Deze benadering biedt een rijker en multidimensionaal perspectief.²⁵ De beide stromingen of 'wereldbeelden' sluiten elkaar niet uit. Zoals bij elke paradigma-verschuiving neemt het nieuwe paradigma namelijk ook de bestaande denkwijze in zich op en plaatst die in een breder perspectief en in andere dimensies.

De toepassing van diplomatische kennis in een digitale omgeving is waardevol en kan prima, maar het blijft een methodiek die nog steeds wordt toegepast binnen het traditionele, meer eendimensionale paradigma, dat zich vooral richt op het archief en de onderdelen ervan. Al het andere is context.²⁶

²⁰ Een uitgebreid verslag van dit project is gepubliceerd in: Luciana Duranti (ed.), *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, San Miniato 2005. Voor digitale versie, zie: http://www.interpares.org/ip1/ip1_dissemination.cfm?proj=ip1.

²¹ De resultaten van dit project zijn gepubliceerd in Luciana Duranti en Randy Preston (ed) *International Research on Permanent Authentic Records in Electronic Systems (InterPARES 2). Experiential, Interactive and Dynamic Records* (Padova, 2008) zie <http://www.interpares.org/ip2/book.cfm>. InterPARES kent ook een 3de project (2007-2012) *Implementing the theory of preservation of authentic records in digital systems in small and medium sized archival organizations*. En een 4de project (2013-2018) *Exploring issues concerning digital records and data entrusted to the Internet*. Alle projecten worden gekenmerkt door hun interdisciplinaire karakter.

²² De eerste heette 'Chain of Preservation model', de tweede 'Business-driven Recordkeeping model'. Zie: http://www.interpares.org/ip2/ip2_models.cfm geraadpleegd 11 oktober 2018.

²³ Zie o.a. Theo Thomassen, 'Paradigmatische veranderingen in de archiefwetenschap', in: P.J. Horsman, F.C.J. Ketelaar en T.P.H.M. Thomassen, *Naar een nieuw paradigma in de archivalistiek*, (Den Haag 1999), p.69-79. Of het nu wel of niet om een paradigmaverschuiving gaat, is nog altijd punt van discussie. Zie Luciana Duranti, 'The impact of digital technology on archival science', in: *Archival Science*, Vol. 1, No. 1, 2001, p.39-55, en G.J. van Bussel, 'The theoretical framework for the 'archive-as-is'. An organization oriented view on archives. Part I', in: Frans Smit, Arnoud Gloudemans, Rienk Jonker (red.), *Archives in liquid times* (SAP-jaarboek, 's-Gravenhage 2017), p.16-41.

²⁴ Een belangrijke inspiratiebron voor deze nieuwe benadering vormt het werk van de sociaal theoreticus Anthony Giddens.

Al is het *records continuum* denken dan niet ontstaan als reactie op digitalisering en digitale informatie, het biedt wel een bredere basis om de toegenomen complexiteit in de manier waarop organisaties en mensen met informatie omgaan in de huidige netwerkmaatschappij het hoofd te bieden en 'recordkeeping' in de digitale wereld vorm te geven. De aandacht daarbij gaat vooral uit naar processen van creatie, beheer en toegang, en naar metadata waarin die worden vastgelegd.²⁷ Achterliggende gedachte is dat archief of archiefbescheiden niet statisch is of zijn, maar voortdurend in beweging ('a record is "always in a process of becoming"...'), omdat de context waarin het wordt (her)gebruikt en beheerd, steeds anders is.²⁸ Reden waarom gegevens over die processen steeds moeten worden vastgelegd om de vele contexten (in tijd en ruimte) te kunnen begrijpen en verantwoording te kunnen afleggen.

Sommige nieuwe ideeën bestonden al voordat het *records continuum* vorm had gekregen. Het bleek echter al snel dat met deze nieuwe theorie een ander perspectief en kader werd geboden, waarin die ideeën een meer logische plaats krijgen en beter tot hun recht komen. Voorbeelden van ideeën die al in de lucht hingen zijn bijvoorbeeld het idee dat goed archiveren al begint bij de planning en het ontwerp van informatiesystemen in overheidsorganisaties en dat archivariissen daarbij betrokken moeten zijn; het waarderen en selecteren van archiefbescheiden op basis van functies en taken, *functional appraisal* (macro- of functionele selectie), en niet op basis van aanwezige archiefbescheiden; en het onderscheid tussen fysieke en logische aspecten van (digitale) archiefbescheiden. Tegelijk zijn er obstakels in het denken en doen, zoals het feit dat archivariissen in de praktijk vaak te veel gericht zijn op het te bewaren archief en te weinig naar de informatievoorziening inclusief archiveren in relatie tot de bedrijfsprocessen als geheel kijken of de ook in Nederland nogal sterke scheiding tussen DIV en archiefdienst. Dergelijke zaken bemoeilijken het invoeren van nieuwe werkwijzen.

Nieuwe wegen: recordkeeping informatics

Een voorbeeld van een nieuwe methodiek of benadering is *recordkeeping informatics*, die in Australië is ontwikkeld. Onder deze nieuwe naam beoogt een aantal erkende recordkeeping theoretici het vak van archiveren (*recordkeeping*) een nieuwe basis te verschaffen en daarmee weer onderscheidend te maken.²⁹

²⁵ Zie Frank Upward 'Structuring the Records Continuum. Part One: Postcustodial Principles and Properties', in: *Archives and Manuscripts* Vol 24 No 2 (1996) (eerste publicatie). Zie verder <https://recordscontinuum.info/discover/publications> als ook Barbara Reed, *Reading the Records "Continuum: Interpretations and Explorations*, z.j. https://www.records.com.au/pdf/Reading_the_Records_Continuum.pdf Daarnaast is ook het onovertroffen overzicht van Terry Cook, 'What is past is prologue' van belang om inzicht te krijgen in de verschillende stromingen en ontwikkelingen op het gebied van de archivalistiek in de 20ste eeuw. Gepubliceerd in *Archivaria* (1996) p.17-63. Ook gepubliceerd in *Naar een nieuw paradigma in de archivalistiek* (SAP-jaarboek 1999) p.29-67

²⁶ In digitale vorm zijn archiefbescheiden veeleer logische entiteiten dan fysieke dingen. Het fysieke is het digitale bestand dat de code bevat die de software nodig heeft om het (voor mensen) leesbaar te maken.

²⁷ Hierbij dient aangetekend te worden dat een dergelijke functionele benadering al langer in de lucht hing en toegepast werd. Een voorbeeld is de PIVOT-methode die bij selectie uitgaat van de analyse van de (overheids) handelingen. Ook Theo Thomassen spreekt over functionele archivalistiek in zijn 'Een korte introductie in de archivalistiek' (zie P.J. Horsman, F.C.J. Ketelaar en T.P.H.M. Thomassen, *Naar een nieuw paradigma in de archivalistiek*, (Den Haag 1999), p. 11-20).

²⁸ Zie Sue McKemmish 'Placing records continuum theory and practice' in: *Archival Science* 1 (4) (2001) p. 335.

²⁹ Zie o.a. Gillian Oliver, Frank Upward, Barbara Reed, Joanne Evans *Recordkeeping Informatics: building the discipline base* (DLM, 2014) http://purl.pt/26107/1/DLM2014_PDF/02%20-%20Recordkeeping%20Informatics%20Building%20the%20Discipline%20Base.pdf geraadpleegd 11 oktober 2018.

Deze benadering zegt een reactie te zijn op de 'crisis' in archivering in de huidige netwerkmaatschappij, omdat het onderscheid tussen verschillende informatievormen verloren gaat dan wel vervaagt juist vanwege de enorme hoeveelheid ervan en door de huidige focus op het object of ding dat archiefstuk wordt genoemd.³⁰ De traditionele methoden van selectie, beschrijven, beheren, en beschikbaar stellen voldoen in die dynamische en complexe wereld (door hen *wild frontier* genoemd) niet meer, met als gevolg dat DIV-functionarissen en archivariissen achter de feiten aanlopen en de greep op informatie verliezen. Een andere interessante constatering van deze groep is dat al enige tijd de tendens bestaat om niet langer te spreken over archiefbescheiden (records) als gezaghebbende bronnen, maar over informatie. Daarmee in feite de eigen discipline verloochend. Deze neiging tot eufemistisch en daardoor mistig taalgebruik rond archiveren komt ook al enige tijd in Nederland voor.³¹

Zoals al in het citaat aangegeven omvat *recordkeeping informatics* twee bouwstenen, nl. continuüm denken en metadata, en drie aspecten van procesanalyse, nl. bedrijfsprocessen, organisatiecultuur en toegang³²:

“Recordkeeping informatics is based on two building blocks – continuum thinking and recordkeeping metadata – and the exploration of three facets – information culture, access and business processes – to rethink recordkeeping approaches to better deal with the expanding continuum of recorded information that is a hallmark of the digital and networked age. It aims to take into account the social, cognitive and technological characteristics of our current environment and address the collapse in collective memory brought about by digital disruption.”

Met het pleidooi voor meer nadruk op processen (creatie, beheer en beschikbaar stellen) hoopt men te voorkomen dat we uiteindelijk met rommel zitten en ons 'geheugen' verloren gaat.³³ Of deze andere wijze van denken op basis van het records continuüm model aanslaat, moet nog blijken, maar ook buiten de archiefwereld is er belangstelling voor, zoals in het Pericles-project.³⁴

³⁰ Zie bijvoorbeeld <https://www.ryanstander.com/archival-turn-statement.html> waar onder andere gezegd wordt “Traditionally, an archive has been regarded as a repository of objects and information essential to human history. However, in the last half of the 20th Century it underwent a profound conceptual change shifting its emphases from the objects to the archivist, and from place to process. At the heart of this archival turn is a fundamental skepticism of Modernity’s scientific methodologies.”

³¹ Dat begint al enigszins met het boekje 'Omslag in opslag. Terug naar informatievoorziening' (Den Haag 1991, uitgave ministerie van Binnenlandse Zaken), waarin het begrip *archiefbeheer* als te beperkt wordt omschreven en vervangen wordt door 'informatievoorziening gericht op werkprocessen'. Overigens wordt ook hier al geconstateerd dat bestaande archiveringsmethoden niet voldoen.

³² Zie o.a. Gillian Oliver, Frank Upward, Barbara Reed en Joanne Evans, *Recordkeeping informatics: building the discipline base* (2014) en de website van Monash University <https://www.monash.edu/it/our-research/research-centres-and-labs/rcrg/recordkeeping-informatics>. Geert Jan van Bussel, in zijn kritiek op de records continuüm theorie, vindt deze nieuwe benadering overigens een 'onnodige uitvinding': “The (needless) ‘invention’ of ‘recordkeeping informatics’ to ‘re-figure a discipline in crisis’ (that is: a records management that does not use the records continuum theory and model)”. En passant trekt hij de 'crisis' in twijfel door te suggereren dat de auteurs dat zo stellen, omdat de archiefwereld het records continuüm denken niet volgt. Een boude, niet onderbouwde veronderstelling, die niet is af te leiden uit wat er over *recordkeeping informatics* geschreven is. Zie Geert-Jan van Bussel, 'The theoretical framework for the 'archive-as-is'. An organization oriented view on archives. Part I', in: Frans Smit, Arnoud Gloudemans, Rienk Jonker (red.), *Archives in liquid times* (SAP-jaarboek, 's-Gravenhage, 2017), p.17.

Van organisatie naar functie: archival function (archieffunctie) of recordkeeping function (archiveringsfunctie)

Eerder in dit artikel heb ik al geconstateerd dat er een grote overlap of gelijkenis is tussen de gebruikte definities van *digital preservation* en *recordkeeping*. Nader beschouwd lijkt het erop dat het gebruik van een term mede afhangt van de instelling waar je werkt. Als dat bij een archiefdienst is, wordt het *digital preservation* (of preservering) en als dat bij een overheids- of andere archiefvormende organisatie is, wordt gesproken van *recordkeeping* (archivering). De één bekijkt het vanuit dat wat bewaard moet worden (het ding/object). Hoe het gecreëerd moet worden en goed beheerd (de processen) tot het wordt overgebracht is secundair en wordt min of meer als een premisse gezien. Voor de ander, de medewerkers van archiefvormende organisaties, zijn creatie en beheer juist het perspectief waarin ernaar gekeken wordt. Hier wordt de lange termijn bewaring als een onderdeel van *recordkeeping* gezien, maar het staat niet in het centrum van de aandacht. In feite spreekt men echter over hetzelfde. Er bestaat evenwel een 'kloof' in hoe men er tegenaan kijkt en daarmee ook in hoe men het benadert. Dat leidt in de praktijk tot een onnodige tegenstelling, alsof men verschillende belangen dient.

De vraag is, hoe dit kan worden overbrugd? Een belangrijke stap zou zijn om archivering los te zien van de organisatorische structuur en veel meer te kijken naar het geheel, namelijk de functie, in dit geval de archiveringsfunctie.³⁵

Zoals de voorstanders van *recordkeeping informatics* al stellen gaat het in eerste instantie om processen en pas in tweede om 'het ding' ofwel de records of archief. Dat komt in de buurt bij de ideeën van het ICA *committee on electronic records*, dat midden jaren '90 al pleitte om niet zozeer te kijken naar de organisaties die een rol spelen, maar naar de functie. In dit geval is dat de functie van het voor de lange termijn bewaren van archiefbescheiden om aan bepaalde (historische, maatschappelijke) belangen te voldoen, toen *archival function* genoemd.³⁶ Het verschil met *recordkeeping informatics* is dat hierbij de gehele archiveringsfunctie (*recordkeeping function*) in ogenschouw wordt genomen, dus inclusief ontwerp en creatie, en niet alleen gekeken wordt naar de lange termijn bewaring. Het los te zien van de

³³ Zie Frank Upward, Barbara Reed, Gillian Oliver, Joanne Evans, 'Recordkeeping Informatics: Re-figuring a discipline in crisis with a single minded approach (conceptual paper) 2013, in: *Records Management Journal*, Vol 23 Issue 1 (2013) pp37-50 via <https://www.emeraldinsight.com/doi/abs/10.1108/09565691311325013>. En: Frank Upward, Barbara Reed, Gillian Oliver, Joanne Evans, 'Recordkeeping Informatics for a Networked Age' (Melbourne 2017). Zie <https://www.monash.edu/it/our-research/research-centres-and-labs/rcrg/recordkeeping-informatics>. Zie <https://www.emeraldinsight.com/doi/abs/10.1108/09565691311325013> “The purpose of this paper is to highlight the widespread crisis facing the archives and recordsmanagement professions, and to propose recordkeeping informatics, a single minded disciplinary approach, as a way forward.”

³⁴ Voor Pericles zie <http://www.pericles-project.eu/page/approach>. Anderen vinden het weer onnodig, zoals de reeds aangehaalde Geert-Jan van Bussel. In de masterstudie Archiefwetenschap aan de Universiteit van Amsterdam is het inmiddels in het curriculum opgenomen (<http://www.uva.nl/m-programmas/gsh/dual-masters/archival-and-information-studies-heritage-studies/study-programme/study-programme.html>).

³⁵ Bij het programma Archief 2020 is aandacht besteed aan de archieffunctie en zijn een drietal toekomstscenario's ontwikkeld (zie <https://archief2020.nl/downloads/op-weg-naar-2020-en-verder-het-verhaal-van-de-archieffunctie-geraadpleegd-11-oktober-2018>) als hulpmiddel om een strategische discussie te stimuleren. Daarbij is toch vooral uitgegaan van archiefdiensten en hun mogelijke rol. Gepubliceerd in oktober 2015. Het begrip archieffunctie lijkt sterk op wat hieronder *archival function* wordt genoemd (dus gericht op de lange termijn bewaring). Dat wordt nog eens bevestigd in de *Handreiking toekomstbestendige archiefinstellingen* (Archiefinnovatie Decentrale Overheden, AIDO, 2016), waar gewezen wordt op het belang van modernisering van archiefinstellingen door implementatie van e-depotvoorzieningen. Geen woord over de archivering bij overheidsinstellingen, waar de duurzaamheid toch begint.

³⁶ Zie *Guide for managing electronic records from an archival perspective* (ICA Committee on electronic records, Paris, 1997), p. 25-28.

organisatiestructuur maakt het makkelijker om te zien waar het om draait, los van institutionele belangen. Dit is van groot belang gezien de aard van het internet, de vele nieuwe informatievormen, en de toegenomen complexiteit, waarbij organisaties in toenemende mate in netwerken zijn verbonden.

Het te bewaren object en bijbehorende (beheer)processen zijn in technische zin veranderd, de functie van archiveren is dat op zich niet. Door op deze wijze wat afstand te nemen, is het ook gemakkelijker om te kijken hoe de archiveringsfunctie nu het beste kan worden georganiseerd. 'Archief' is in eerste instantie niet langer een locatie, maar een virtuele omgeving en daarmee een concept.

Waardering en selectie zijn dan ook geen processen die achteraf en eenmalig gebeuren, maar vooraf en iteratief. Het wordt al enige tijd op die manier gepromoot.³⁷ In de nieuwe ISO 15489:2016 heeft het begrip *appraisal* een veel bredere invulling gekregen. Het wordt omschreven als: "...the process of evaluating business activities to determine which records need to be created and captured and how long the records need to be kept."³⁸ Het combineert een analyse van de bedrijfscontext waarbij de eisen voor archiefbescheiden en archiveren worden geïdentificeerd en een risico-analyse. In de praktijk wordt dit evenwel nog niet echt zo toegepast.

De praktijk: stand van zaken in Nederland

Theoretische bespiegelingen zijn misschien aardig, maar wat gebeurde er in de praktijk? Ik neem daarbij de situatie in Nederland als voorbeeld wat betreft digitale archivering en meer speciaal digitale 'preserving'. Die is vergelijkbaar met wat er elders in de wereld gebeurde.

Wat in theorie allemaal samenhangend en logisch lijkt, blijkt in de praktijk allemaal nog niet zo simpel. Theoretici spreken vaak een andere, abstractere taal dan de mensen in de praktijk, al gaat het over dezelfde zaken.³⁹ Zeker als er nieuwe theoretische ideeën of modellen ontstaan, duurt het nog wel enige tijd voor dat echt op de 'werkvloer' is doorgedrongen. Hoe mooi die modellen ook (lijken te) zijn. Dat geldt zeker voor het hoog abstractieniveau van de records continuum theorie. Ondanks het nieuwe denkkader dat het biedt, is het de vraag hoe zich dat vertaalt naar de dagelijkse praktijk van archivarissen en records managers.⁴⁰ Zoals de praktijk er nu uit ziet, is er een grote kloof met nieuw ontwikkelde en aanbevolen werkwijzen. Oude gewoonten en denkwijzen zijn vaak hardnekkig en veranderen is lastig, zeker in een omgeving waarin nog veel onduidelijk is over waar het heen gaat en hoe het moet worden aangepakt. Saillante voorbeelden zijn de lastige en trage aanpak van email- en website-archivering bij de overheid, met als gevolg verlies aan informatie.

³⁷ Charles Jeurgens, 'Threats of the data-flood. An accountability perspective in the era of ubiquitous computing' in: Frans Smit, Arnoud Gloude-mans, Rien Jonker (red) *Archives in liquid times* (SAP-jaarboek, 's-Gravenhage, 2017) p. 196.

³⁸ ISO 15489:2016 *Information and documentation – Records management – Concepts and principles*, par. 7.1. Op dit moment is er een *technical report* in ontwikkeling waarin dit appraisal proces nader wordt uitgewerkt (*Appraisal for managing records*). Deze appraisal vindt niet eenmalig plaats, maar is een steeds terugkerende activiteit. Zie ook <https://cassiefindlay.com/2018/06/05/appraisal-a-strategic-tool-for-modern-recordkeeping/> voor meer uitleg hierover.

³⁹ Charles Jeurgens heeft hier bijvoorbeeld al op gewezen, in het bijzonder met betrekking tot waardering en selectie, in zijn *The target keeps moving! Appraisal in a continuum* (2014) zie <https://openaccess.leidenuniv.nl/bitstream/handle/1887/32027/Jeurgens%20target%20keeps%20moving.def.pdf?sequence=1> geraadpleegd 11 oktober 2018.

⁴⁰ De kritiek van Van Bussel op dit punt heb ik reeds vermeld.

In het algemeen lopen het denken en de praktijk achter bij de snelle veranderingen in de maatschappij en bij organisaties. Dat knelt te meer, omdat zoals gezegd al voordat de digitale informatie is gecreëerd actie moet worden ondernomen om de duurzaamheid en toegankelijkheid te waarborgen. Vergeleken met de oude situatie waar papieren archieven relatief veel minder aandacht vragen en pas na 20 jaar of later bij een archiefdienst belanden, wordt bij digitale archiefbescheiden continu een extra inspanning gevraagd.

Dat vraagt ook afstemming en nauwere samenwerking tussen overheidsorganisaties en archiefdiensten, om te zorgen dat het migreren van archiefbescheiden tussen systemen soepel verloopt en voorkomen wordt dat ze niet meer raadpleegbaar, begrijpelijk of bruikbaar zijn.

De overgang binnen de overheid, maar ook elders, naar automatisering van werkprocessen, waarbij voor zover mogelijk workflow en hopelijk ook archivering wordt geïntegreerd in de informatiesystemen verloopt traag. Immers het werkelijk digitaal afhandelen van taken vergt niet alleen een goede technische omgeving, maar ook organisatorische aanpassingen en niet te vergeten een cultuuromslag. Reeds in 2004 is bij het toenmalige ministerie van Volkshuisvesting, Ruimtelijke Ordening en Milieu (VROM) een werkproces volledig gedigitaliseerd (EVOA)⁴¹, daarnaast waren er pogingen om de documentenafhandeling te automatiseren, zoals bij Digidoc (BZK).⁴² Later zijn er pogingen om het omgevingsloket te ontwikkelen als ook de afhandeling van strafrechtzaken. Dat dit ingewikkeld is, of te ingewikkeld wordt gemaakt, blijkt uit de vele mislukte ICT-projecten die zich de afgelopen jaren hebben voorgedaan. Ambities om bijvoorbeeld de gehele rijksoverheid geheel digitaal te maken in 2017 blijken dan ook te hoog gegrepen en tonen slechts een enorme onderschatting van deze transformatie.⁴³ Een steeds terugkerend fenomeen, parallel daaraan, zijn pogingen om de overheidsinformatiehuishouding op orde te brengen. Met een cyclus van ongeveer 4-6 jaar verschijnt er weer een actieplan of programma met maatregelen en plannen. Voorbeelden zijn sinds 1995: Digitale duurzaamheid, Interlab, Informatie op Orde, Baseline informatiehuishouding, DUTO, Rijk aan Informatie. Vaak gebeurt dat ook onder de noemer van modernisering dan wel vernieuwing van de rijksdienst.⁴⁴

Toch wordt nog grotendeels met document management systemen gewerkt, eventueel in samenhang met records management applicaties. Dat zijn niet de meest ideale middelen om goede archivering te bereiken, ook al niet omdat deze beperkt en erg document-gericht zijn en er steeds meer andere typen informatie zijn ontstaan. Een belangrijk aandachtspunt zijn de vereiste metadata, die zorgen dat de archiefbescheiden (documenten of data) in hun juiste context begrepen of vertrouwd

⁴¹ Zie <https://www.ilent.nl/onderwerpen/afvaltransport-in-europa-evoa-geraadpleegd> 11 oktober 2018.

⁴² Zie bijvoorbeeld <https://www.computable.nl/artikel/achtergrond/infrastructuur/1681886/1444691/bzk-1-eerste-e-ministerie.html> 11 oktober 2018.

⁴³ Het aantal mislukkingen op gebied van ICT bij de overheid laat aan duidelijkheid niets te wensen over, ondanks alle pogingen daar meer greep op te krijgen. Bekende voorbeelden zijn de belastingdienst, de rechtspraak en UWV.

⁴⁴ Zie bijv. Kamerstukken 31490, nr. 122, d.d. 25-06-2013 ('Vernieuwing van de Rijksdienst') en 29362, nr. 260, d.d. 23-03-2017 ('Modernisering van de overheid'). Ter ondersteuning van dit alles is er een Enterprise architectuur Rijksdienst (EAR, zie https://www.earonline.nl/index.php/Wat_is_de_Enterprise_Architectuur_Rijksdienst) en Nederlandse Overheid Referentie Architectuur (NORA), zie https://www.noraonline.nl/wiki/NORA_online. Daarin is ook een hoofdstuk over 'duurzame toegankelijkheid' opgenomen, zie https://www.noraonline.nl/wiki/Duurzame_Toegankelijkheid geraadpleegd 11 oktober 2018.

kunnen worden. Vaak blijken deze onvolledig of ontbreken ze. Bovendien bestaan er grote verschillen in wijze van vastlegging tussen systemen, die ertoe bijdragen dat migratie van archiefbescheiden naar andere systemen moeilijk en ingewikkeld is. Om daarin verbetering te brengen is eind 2009 een Toepassingsprofiel metagegevens Rijk vastgesteld, waarin de belangrijkste typen metadata zijn aangegeven.⁴⁵ Doel ervan was en is om door deze standaardisering betere uitwisseling tussen systemen mogelijk te maken. De invoering ervan verloopt evenwel moeizaam.

Rijk aan Informatie is het meest recente programma dat zich richt op het 'toekomstbestendig maken van de informatiehuishouding van het Rijk' die tevens 'flexibel en robuust' moet zijn.⁴⁶ Gezien het zogenaamde programmaplan (2018) lijkt men de oplossing vooral te zoeken in het automatiseren van archiveringsprocessen ('by design') of beter van 'werkwijzen en regels'.⁴⁷ Daarmee wordt gekozen voor een instrumentele aanpak, al noemt men het een 'paradigmashift'. Het klinkt bovendien alsof het nog steeds aparte processen zijn, terwijl toch integratie in werkprocessen het streven zou moeten zijn. Daarnaast worden als andere onderdelen genoemd, emailarchivering en webarchivering, naast een onderzoek naar DIV-functies en verbeteren van samenwerking. Men spreekt verder over het 'optimaliseren van bestaande werkwijzen', 'vereenvoudiging, modernisering en op orde brengen van kaders en werkwijzen' en experimenteren met *by design*. Met dat laatste wordt bedoeld dat in het ontwerp van informatiesystemen de archivering wordt meegenomen. Het credo is, de medewerker moet 'ontzorgd' worden.⁴⁸

Dat is allemaal mooi, maar het zou interessant zijn te weten op basis van welke principes en concepten dit alles is gebaseerd en wat de samenhang is. Zoals hierboven bij *appraisal* is aangegeven, begint dat al met de analyse van wat een organisatie aan archiefbescheiden zou moeten maken.⁴⁹ Dat wordt niet echt duidelijk. Bovendien ligt de nadruk sterk op document-georiënteerde informatie en wordt een belangrijk deel van de informatie in de vorm van databases en data (nog steeds) vergeten, alsof daar archivering en zaken als integriteit, betrouwbaarheid en authenticiteit, geen rol spelen.

Veel aandacht wordt besteed aan e-depots, emailarchivering (nog steeds een probleem), en websitearchivering. Wat dit laatste betreft is er onlangs (2018) een concept richtlijn websitearchivering opgesteld door het NA.⁵⁰ Het is een thema dat

⁴⁵ Zie https://www.earonline.nl/index.php/Toepassingsprofiel_metagegevens_Rijksoverheid_geraadpleegd 11 oktober 2018. Voor lagere overheden is een soortgelijk profiel vastgesteld Toepassingsprofiel metagegevens lagere overheden (TMLO).

⁴⁶ Zie <https://www.rijkaaninformatie.nl/over-rijk-aan-informatie>, waar als opdracht staat: "Met behulp van nieuwe technologieën het op een geautomatiseerde wijze toewerken naar de "by design" toepassingen in de informatiehuishouding. Enerzijds worden de technische (on)mogelijkheden onderzocht en anderzijds de instructies die ambtenaren hierbij nodig hebben getest." Het programma wordt gezamenlijk aangestuurd vanuit OCW (SG) en BZK (CIO Rijk). Geraadpleegd 11 oktober 2018.

⁴⁷ Zie *programmaplan Rijk aan Informatie*, (2017) p.4. https://www.rijkaaninformatie.nl/over-rijk-aan-informatie/documenten/beleidsnotas/2018/01/02/programmaplan-rijk-aan-informatie-2018_geraadpleegd 11 oktober 2018. In hoeverre zijn of worden DIV of archivarissen dan bij de ontwikkeling van systemen en specifiek de automatisering van archiveringsprocessen betrokken en belangrijker, hebben zij voldoende kennis en deskundigheid om daar een adequate rol in te spelen?

⁴⁸ Ibidem, p. 5.

⁴⁹ Zie hierboven tekst over *appraisal* (en zie de herziene ISO 15489:2016).

⁵⁰ Zie https://www.nationaalarchief.nl/archiveren/nieuws/openbare-review-richtlijn-archiveren-overheidswebsites_geraadpleegd 11 oktober 2018.

al lang speelt, maar eigenlijk nooit van de grond gekomen is. Reeds in 2006 is vastgesteld dat overheidswebsites als archief beschouwd moesten worden en zijn er 'webrichtlijnen' opgesteld.⁵¹ Toepassing ervan in de praktijk was niet verplicht, met als gevolg dat het niet of nauwelijks werd gevolgd. Op die manier zijn veel (versies van) overheidswebsites verdwenen, zoals bij de overgang van de ministeriële websites naar de nieuwe, gezamenlijk website overheid.nl.⁵²

Inmiddels zijn er allerlei e-depots, waarin voornamelijk nog gedigitaliseerd archief is opgeslagen. Digitale archiefbescheiden zijn nog slechts mondjesmaat overgebracht vanuit overheidsinstellingen. Nieuwe ontwikkelingen worden aangekondigd zoals de invoering van de zgn. 'preservation watch' bij het Nationaal Archief.⁵³ Men baseert zich daarbij op het functioneel model dat door het Planets project is ontwikkeld.⁵⁴ Deze monitor-functie zou moeten uitgroeien tot een landelijke dienst en richt zich op ontwikkelingen in technologie, op het beleid en de middelen van de beherende organisatie, de ontwikkelingen bij de archiefvormende instellingen, het beheer van de collectie, en de gebruikers, omdat deze van invloed kunnen zijn voor de wijze waarop de digitale bewaring wordt uitgevoerd. Het is de eerste van een trio van functies. De andere twee zijn *preservation planning* (waardoor de mogelijke risico's worden geïnterpreteerd en zo nodig omgezet in actie) en *preservation action* (uitvoering van feitelijke bewaaractiviteiten, zoals conversie, migratie of emulatie).

Samenvatting

De wijze waarop de automatisering en in het bijzonder de archivering bij de overheid zich in de afgelopen 30 jaar heeft ontwikkeld, laat ons echter geen illusies over de kwaliteit van digitale archiefbescheiden en informatie uit die jaren. Ondanks alle pogingen om daar verbetering in te brengen, zijn veranderingen op gebied van beheer van overheidsinformatie traag en duren ze erg lang, al is er wel vooruitgang. De rapporten van de archiefinspectie sinds ca. 2004 over misstanden en de risico's zoals verlies van informatie, als ook de herhaalde pogingen om binnen de overheid tot gezamenlijke aanpak en richtlijnen te komen (o.a. Interlab, Informatie op Orde, Baseline, DUTO) in de eerste jaren van deze eeuw laten dat zien.⁵⁵ Er was in de loop der jaren geen consistent beleid vanuit de verantwoordelijke ministeries van BZK en OCW, veeleer ad hoc beleid naar aanleiding van incidenten of alarmerende

⁵¹ Brief Algemene Rijksarchivaris d.d. 28 maart 2006 aan de programmamanager Advies Overheid.nl.

De ministeriële websites waren van de ene dag op de andere uit de lucht en zijn niet bewaard. In 2003 was reeds een rapport verschenen in opdracht van ministerie van Verkeer en Waterstaat, getiteld *Blijvend beschikbaar. Onderzoek naar de archivering van websites*, (Erika Hokke, Amsterdam, 2003). In 2005 verscheen een bibliografie van de Archiefschool getiteld *Eerste hulp bij website archivering* (Erika Hokke, Amsterdam, 2005), later gevolgd door een rapport *Richtlijnen voor websites* (Erika Hokke en Peter Horsman, Amsterdam, 2006) als onderdeel van een ICTU project webrichtlijnen.

⁵² Delen ervan zijn misschien nog te vinden op het Internet Archive of wellicht bij de KB.

⁵³ Zie https://www.nationaalarchief.nl/archiveren/nieuws/preservation-watch-bij-het-nationaal-archief?utm_campaign=Nieuwsbrief%20Nationaal%20Archief&utm_medium=e-mailnieuwsbrief&utm_source=email_geraadpleegd 11 oktober 2018.

⁵⁴ Zie https://www.nationaalarchief.nl/sites/default/files/field-file/Planets-Preservation_Watch-functional_model.jpg_geraadpleegd 11 oktober 2018. Hierin ontbreekt overigens de testbed-functie, van belang voor het uitvoeren van experimenten en testen van bewaarstrategieën.

⁵⁵ Zie voor Baseline informatiehuishouding Rijksoverheid, onderdeel van programma 'Informatie op orde' 2006 : https://www.earonline.nl/images/earpub/3/3d/Baseline_Informatiehuishouding_Rijksoverheid_de_7_normen_versie_1.1.pdf, voor DUTO (2016): <https://www.nationaalarchief.nl/archiveren/kennisbank/duto>. Zie voor rapporten Erfgoedinspectie oa. <https://www.erfgoedinspectie.nl/publicaties/rapport/2005/02/01/dementerende-overheid>.

rapporten. Structureel budget voor het Nationaal Archief bijvoorbeeld kwam er pas in 2014. Bij ministeries werd het niet als erg belangrijk gezien. Van belang is hierbij dat archivering als iets aparts werd aangekaart en niet als wezenlijk onderdeel van de taakuitvoering.

In het bovenstaande heb ik voor archivering een stand van zaken proberen weer te geven. Duidelijk is dat er nog sprake is van een overgangsfase met alle onduidelijkheid van dien. Wat is er veranderd, hoe moet erop gereageerd worden en hoe wordt erop gereageerd? Wat is er nodig en welke opties zijn er? Veel zaken lopen door elkaar, terwijl de strategie nog steeds grotendeels reactief is. Dat geldt zowel voor het terrein van 'preservering' als voor dat van archivering. Het zou helpen als daar enige orde in geschapen zou worden.

Kijken we naar de archivistiek en archivering dan zien we een grote kloof tussen theoretische modellen en de dagelijkse praktijk. Een conclusie die wellicht getrokken kan worden, is dat organisaties, zowel archiefdiensten als overheidsinstellingen, zich nu meer bewust zijn van het probleem van digitale duurzaamheid. Dat heeft zo'n 20 jaar gekost.⁵⁶ Die bewustwording vertaalt zich nog vooral in het implementeren van of aansluiten bij een e-depot of het publiceren van richtlijnen m.b.t. website-archivering en email archivering.⁵⁷ De aandacht richt zich op specifieke zaken zoals op applicaties en informatiesystemen, typen informatie, zoals documenten, email en websites, en e-depots, niet op de context waarin zij een rol spelen. Heel veel verder komt men nog niet. Dat is zorgelijk, want de ontwikkelingen op gebied van digitalisering gaan snel. Zo snel dat wat vandaag bedacht wordt voor de problemen van vandaag en gister, al vrijwel meteen achterhaald is, met als gevolg dat men achter de feiten aan blijft lopen. Dat kan alleen veranderd worden door meer proactief te zijn en de bedrijfsprocessen als uitgangspunt te nemen, zoals voorgesteld wordt met de reeds genoemde nieuwe vorm van *appraisal*.⁵⁸ Uiteindelijk gaat het om de goede vastlegging van de uitvoering van taken, zodat de bedrijfsvoering wordt ondersteund en verantwoording kan worden afgelegd. Door deze werkwijze wordt ook de samenhang duidelijk van informatie in verschillende informatiesystemen, variërend van kantoorautomatisering, e-mail, sociale media, specifieke applicaties of databases tot websites, die bij de uitvoering van een taak worden gebruikt. Door het proces als uitgangspunt te nemen worden ook de relaties met systemen buiten de eigen organisatie duidelijk. Een tweede stap is dan betrokkenheid bij de planning en het ontwerp van informatiesystemen en zo meteen de archiveringsfunctie mee te nemen. In de praktijk gebeurt dat niet of nauwelijks. Als redenen kunnen worden genoemd, gebrek aan kennis en deskundigheid, de geringe status van archivering (en DIV),

⁵⁶ Reeds midden jaren '90 van de vorige eeuw is er door de ministeries van BZK en OCW een programma Digitale Duurzaamheid ingesteld met als doel bewustwording te stimuleren en plannen en richtlijnen te maken voor een digitale infrastructuur voor het beheer van digitale informatie. Veel van wat nu gepubliceerd wordt, was toen al bekend.

⁵⁷ Zie o.a. Kamerstukken 29362, nr. 260, Brief van de ministers van BZK en OCW inzake 'Modernisering van de overheid', d.d. 23-03-2017.

⁵⁸ Zie bijvoorbeeld Cassie Findlay *Appraisal: A strategic tool for modern recordkeeping* (2018) via <https://cassiefindlay.com/2018/06/05/appraisal-a-strategic-tool-for-modern-recordkeeping/> geraadpleegd 11 oktober 2018.

⁵⁹ De poging tot het ontwikkelen van een kennisfunctie door het NA is te beperkt en te vrijblijvend en lijkt in goede bedoelingen gesmoord. Op dit moment zijn er, naast het bredere Netwerk Digitaal Erfgoed (NDE), in Nederland twee kennisplatforms die zich bezighouden met het delen van kennis en informatie op gebied van archivering en preservering. Dat zijn BREED (vooral lagere overheden) en Kennisnetwerk Informatie en Archief (KIA). Zie o.a. <https://kia.pleio.nl/cms/view/52664862/over-kia> en <http://www.breednetwerk.nl>.

en niemand voelt zich verantwoordelijk of is als verantwoordelijk aangewezen. Ook de huidige rol en positie van archiefdiensten geeft daarvoor te weinig basis en handvatten, terwijl ook daar de nodige kennis en deskundigheid onvoldoende is.⁵⁹

Kunnen of moeten we nu spreken van een crisis op gebied van archivering (record-keeping) zoals sommigen suggereren?⁶⁰ Feit is dat veel informatiesystemen bij de overheid zijn of worden ontwikkeld en/of geïmplementeerd waar DIV of archivariissen niet of nauwelijks bij betrokken zijn.⁶¹ Is daarmee de goede archivering en indirect de betrouwbare overheid in gevaar? Heel veel bewijs is daar nog niet voor. Dat zou onderzocht moeten worden. Al sinds het begin van de jaren '90 wordt gewezen op het risico van verlies aan informatie en een 'gat in de geschiedenis'.⁶² Tot op heden lijkt dat mee te vallen. Als het zonder DIV of archivariissen ook loopt, is de vraag wat hun rol dan nog is? Zijn zij overbodig geworden? In die zin lijkt er wel sprake van een crisis in de archiefdiscipline. Dit beperkt zich overigens niet tot Nederland, maar is een wereldwijd voorkomend verschijnsel. Kortom, hier zijn wel wat vragen te onderzoeken en te beantwoorden. Deze situatie mag ook niet te lang duren, want anders staat de archiveringsspecialist buiten spel.

Welke rol willen, moeten of kunnen archiefdiensten in deze context spelen en wat betekent dat voor hun positie? Waar ligt dan het zwaartepunt? Op de recordkeeping functie, lange termijn bewaring/beheer, toegang, kennisfunctie? Wat betekent dat in de praktijk? Datzelfde geldt voor de rol (en positie) van DIV. Ook daar zijn veranderingen nodig die verder gaan dan nu gebeurt. Tenslotte zijn er de plannen voor aanpassing van de Archiefwet 1995 in samenhang met verdere verkorting van de overbrengingstermijn.⁶³ Dit laatste geeft aan dat de erkenning dat er verandering nodig is, aanwezig is. De vraag is of men het probleem voldoende heeft geanalyseerd en het niet te veel benadert vanuit de huidige situatie? Een belangrijk punt is dat de overgang naar een meer op de moderne situatie gerichte aanpak tijd zal vergen. Veel tijd die er eigenlijk niet is. De vraag is dan wat gebeurt er in de tussentijd?

Hoe verder?

Constateren dat er nog veel mis is, is één, maar wat kan eraan gedaan worden? In welke richting moeten we dat zoeken? Om de geconstateerde kloof tussen praktijk en theorie en nieuwe ontwikkelingen te overbruggen is een belangrijke rol weggelegd voor opleidingsinstellingen. Het begint immers met een goede basis in kennis en deskundigheid op gebied van zowel de moderne overheid met de bijbehorende informatiehuishouding, informatietechnologie (incl. blockchain⁶⁴) als archivering

⁶⁰ Zie bijvoorbeeld Frank Upward, Barbara Reed, Gillian Oliver en Joanne Evans, 'Recordkeeping Informatics: Re-figuring a discipline in crisis with a single minded approach', in: *Records Management Journal*, Vol. 23, Issue 1 (2013) p. 37-50.

⁶¹ Via programma Rijk aan Informatie probeert men wel DIV medewerkers beter op te leiden en beter te positioneren, zie <https://www.rijkaaninformatie.nl/dashboard/div-foto>, maar dat is nog toekomst.

⁶² Zie Rekenkamerrapport uit 1991, de rapporten van MLG-project en in 2005 nog het rapport van de Erfgoedinspectie over 'De dementerende overheid'. Opvallend is hoe snel men zaken vergeet. In de brief van de minister voor Basis- en Voortgezet Onderwijs en Media aan de Tweede Kamer inzake het aanpassen van de Archiefwet, d.d. 11-06-2018 wordt terugverwezen naar het rapport van de Archiefinspectie inzake 'Dementerende overheid', terwijl er al 10 jaar eerder allerlei rapporten waren gepubliceerd waar dit al aan de orde kwam en zelfs een programma is opgezet.

⁶³ Zie brief van de minister voor Basis- en Voortgezet Onderwijs en Media aan de Tweede Kamer inzake het aanpassen van de Archiefwet, d.d. 11-06-2018. Interessant is dat archieven nu zijn ondergebracht onder de Directie Media en Creatieve Industrie van OCW. Dat zegt wellicht ook iets over hun positie.

en de nieuwste ontwikkelingen daarin. Een goed begrip van de concepten en principes van archiveren biedt een baken om in de verwarrende nieuwe wereld koers te houden. Ook kunnen opleidingen en trainingen helpen de nieuwe theorieën te begrijpen en vervolgens handvatten te bieden om deze in de praktijk toe te passen. Dat geldt dan met name het *appraisal* proces, om de pro-actieve betrokkenheid van archiveringsspecialisten een goede basis te geven en te stimuleren.

Daarbij zal in die praktijk nauw samengewerkt moeten worden met andere disciplines, zoals systeemontwikkelaars, datamanagers, business analisten, informatiebeveiligers en informatie-architecten. Dat vraagt enige kennis over en van die vakgebieden. Dat zou aangevuld kunnen worden met betere kennis van de praktijk en vooral hoe organisaties tegenwoordig hun taken uitoefenen in een toenemend gedistribueerde en netwerk omgeving en hoe informatie (in welke vorm dan ook) daarin een rol speelt. Dat deze organisaties sterk aan verandering onderhevig zijn, is duidelijk, maar wat zijn die veranderingen dan precies? Daar is tot op heden onvoldoende inzicht in. Dit vraagt om een goede analyse van de bedrijfscontext en bedrijfsprocessen. In hoeverre de huidige opleidingen daar al goed op zijn toegesneden, is even de vraag.⁶⁵ Het aantal aangeboden opleidingen en trainingen alsmede de capaciteit ervan zijn bovendien nog beperkt.

Vervolgens zou dan de eerste stap moeten zijn om bij het vormgeven van archivering in de verschillende bedrijfsprocessen uit te gaan van de archiveringsfunctie als geheel. De bestaande organisatiestructuren komen dan op de tweede plaats. Hoe kan de (maatschappelijke) archiveringsfunctie in een bepaalde context het beste vorm gegeven worden.⁶⁶

Het zou tenslotte helpen als dit alles ondersteund werd door een consistent beleid, dat voortbouwt op wat reeds bereikt is.

⁶⁴ Zie bijv. https://www.researchgate.net/publication/309414363_Blockchain_for_Recordkeeping_Help_or_Hype en <https://rkroundtable.org/2016/02/12/blockchain-technology-applications-and-implications/> en <https://cassiefindlay.com/2017/06/15/appraisal-blockchain/>

⁶⁵ Veelbelovend is dat in de masterstudie archiefwetenschap aan de Universiteit van Amsterdam reeds *recordkeeping informatics* wordt gedoceerd (2018/2019) zie <http://studiegids.uva.nl/xmlpages/page/2018-2019-en/search-course/course/64709>.

⁶⁶ Dit gaat verder dan de drie scenario's die beschreven worden in *Op weg naar 2020 en verder... Het verhaal van de archief functie*, (programma Archief2020, oktober 2015), waarin de archiefdiensten nog centraal staan en de nadruk op permanent te bewaren archief ligt.

De overbrenging voorbij. Het institutionele geheugen van de overheid

'Er is één ding erger dan vervangbaarheid: onvervangbaarheid'.
Arnon Grunberg, *De Volkskrant* 15 mei 2018.

Inleiding

Preserveren is vooral organiseren. Digitale duurzaamheid, zo stelt de Nationale Wetenschapsagenda, gaat niet alleen over data, maar evengoed over organisatiestructuur,¹ of om met de woorden van Trevor Owens hoofd van de afdeling *Digital Content Management* van de *Library of Congress* te spreken: '[n]othing has been preserved, there are only things being preserved. Preservation is the result of ongoing work of people and commitments of resources'.² Owens stelt dat het instituties zijn die preservering mogelijk maken, omdat ze een robuuste infrastructuur vormen om preservering op lange termijn te realiseren.³ De Nationale Wetenschapsagenda signaleert daarentegen een ontwikkeling waarbij de geprivilegieerde rol van archieven en musea in het erfgoed domein aan het veranderen is: '[d]e duurzame productie, toegang en opslag van gedigitaliseerd en in digitale vorm ontstaan erfgoed komen ook steeds meer in handen van burger' waardoor de rollen van consument, producent, maker, beheerder en gebruiker versmelten.⁴

Archiefinstellingen, musea en bibliotheken houden naar hun eigen zeggen het culturele geheugen van de samenleving in stand om iedereen die er behoefte aan heeft toegang tot dat culturele geheugen te geven. Ze zijn als producten van de Verlichting en Franse Revolutie sterk verweven met de structuren van de natie staat en ontwikkelden in de loop van de negentiende eeuw een 'culture of preservation' die, zoals Susan Crane beweert, tot op de dag van vandaag dominant is gebleven.⁵ In deze bijdrage verdiep ik me in de 'culture of preservation' van archiefinstellingen om een beeld te krijgen hoe die 'culture of preservation' zich verhoudt tot de grote veranderingen die de 'digital turn' teweeg heeft gebracht. Of om het anders te stellen: met welke dilemma's op preservingsgebied hebben de gevestigde archiefinstellingen te maken in hun streven om een betekenisvolle rol te blijven spelen in

¹ De Nationale Wetenschapsagenda deel II, *Portfolio voor onderzoek en innovatie. Samenwerking, creativiteit, vernieuwing*, (2016) 45

² Trevor Owens, *The theory and craft of digital preservation* (John Hopkins University Press 2017) p. 7. (Geraadpleegd 29 januari 2018) <https://osf.io/preprints/lissa/5cpjt/>

³ Ibidem.

⁴ Nationale Wetenschapsagenda, *Portfolio*, 46.

⁵ Susan A. Crane, 'Writing the individual back into collective memory' in *American Historical Review*, December 1997, 1372-1385, aldaar 1374-75.

het digitale tijdperk? Hoe toekomstvast is het op verplichte overbrenging⁶ gebaseerde model dat de Nederlandse overheid hanteert om haar eigen archieven duurzaam toegankelijk te houden?

Dat doe ik door eerst te bekijken waar die ‘culture of preservation’ uit bestaat en hoe deze in de loop der tijd tot stand is gekomen. Ik richt me daarbij hoofdzakelijk op de instellingen die tot het openbaar archiefwezen worden gerekend omdat zij deel uitmaken van een samenhangend stelsel van instituties die op basis van wet- en regelgeving verantwoordelijk zijn voor en vorm geven aan de archiveringsfunctie van de overheid. Dit samenhangende stelsel wordt ook wel het archiefbestel genoemd. Wat mij in het bijzonder interesseert is hoe de ‘culture of preservation’ en de inrichting van het archiefbestel elkaar hebben beïnvloed en wellicht nog steeds beïnvloeden. Kan, of op welke wijze kan het huidige archiefbestel, dat is gebaseerd op de overbrenging van overheidsarchieven naar archiefbewaarplaatsen, een doeltreffend antwoord blijven bieden op de behoeften om het digitale heden ook voor de toekomst veilig te stellen?

Veilige havens

De termen ‘preservering’ en ‘preserveren’ verschijnen eind jaren negentig van de twintigste eeuw voor het eerst in het *Archievenblad* in de betekenis van een actieve strategie die nodig is om digitaal gevormde informatie beschikbaar te houden.⁷ Het zijn ruimere termen dan het tot dan toe gebruikte begrip conserveren. Bij conserveren gaat het om het behoud van het object met de daarop vastgelegde informatie. Bij preserveren gaat het primair om het behouden van de informatie zonder dat de originele drager daarbij noodzakelijkerwijze behouden blijft.⁸

In de periode voor de Tweede Wereldoorlog betekende preserveren volgens Paul Conway niet veel anders dan objecten bij elkaar brengen in een goed geoutilleerde ruimte: ‘[t]he sheer act of pulling a collection of manuscripts from a barn, a basement, or a parking garage and placing it intact in a dry building with locks on the door fulfilled the fundamental preservation mandate of the institution’.⁹ De instellingen die zich uitsluitend gingen richten op behoud van collecties werden niet bij toeval opgericht op de drempel van de 18e en 19e eeuw. Dat wil niet zeggen dat daarvoor niet voor archieven werd gezorgd, maar dat gebeurde door de instellingen die de archieven hadden gevormd. Alleen al het Franse koninkrijk van rond 1770 telde in totaal meer dan 6000 plekken waar archieven werden beheerd.¹⁰ Tussen 1789 en 1815 veranderde de samenleving in heel Europa fundamenteel van karakter. Eeuwenoude bestuursstructuren en de bijbehorende instituties uit het verleden werden met

⁶ Overbrenging dient meerdere doelen tegelijkertijd, o.a. het toepassen van het openbaarheidsregime van de Archiefwet. Ik besteed aan dat aspect in deze bijdrage geen aandacht.

⁷ Charles Jeurgens, ‘Het eeuwige leven...’ *Archievenblad* 102, nr. 2 maart 1998, p. 19-20. De term komt in het online gepubliceerde (Nederlands(ch)) *Archievenblad* van 1892 tot 2012 in totaal 23 keer voor.

⁸ B.N. Westerbrink, ‘Wie heeft er recht op de kopie’ *Archievenblad* 102, nr 8. Oktober 1998, p 7-8. De Engelse term ‘preservation’ daarentegen is al veel langer courant en duikt al honderd jaar eerder op in het *Nederlandsch Archiefblad*, maar wordt al die tijd bijna uitsluitend gebruikt in de betekenis van het Nederlandse ‘conserveren’. Voor een compact overzicht van preserveren van archieven zie Michèle V. Cloonan, ‘Preserving Records of Enduring Value’ in Terry Eastwood and Heather MacNeil (eds) *Currents of Archival Thinking* (Santa Barbara, Denver, Oxford 2010) 69-88.

⁹ Paul Conway, ‘Overview: Rationale for Digitization and Preservation’ in *Handbook for Digital Projects* (Andover 2000) 5-20. <https://deepblue.lib.umich.edu/bitstream/handle/2027.42/111828/BC11%20Conway%20Overview%20Rationale%202000.pdf%3Bsequence=1> geraadpleegd 11 oktober 2018

¹⁰ Judith M. Panitch, ‘Liberty, equality, posterity?: Some archival lessons from the case of the French revolution’ *American Archivist* 59 (1996) 30-47, aldaar 38.

horten en stoten vervangen door een nieuwe maatschappelijke orde met nieuwe structuren en instellingen. De vaak gewelddadige regimewisseling ging op tal van plaatsen in Europa gepaard met vernieling van archieven en kunst- en cultuurschaten in een poging om op die manier definitief af te rekenen met het verderfelijke ancien regime en zo herinneringen aan het verleden uit te wissen. Het vandalisme leidde echter ook tot een sterke tegenbeweging die wilde verhinderen dat onherstelbare schade werd toegebracht aan ‘le patrimoine de tous’, ofwel het erfdeel van allen.¹¹ Er werden *plekken* gesticht die het ‘erfgoed’ een veilige haven konden bieden. Hoewel in Nederland de regimewisseling als een ‘fluwelen omwenteling’ kan worden gekarakteriseerd, vielen ook hier de bezittingen van de oude machthebbers ten prooi aan de dreiging van verkoop, afstoting en soms vernieling. Ook in Nederland vormde deze revolutionaire periode de betekenisvolle context voor de oprichting van de Nationale Bibliotheek (1798; voorloper van de Koninklijke Bibliotheek), Nationale Konst-Gallerij (1800, voorloper van het Rijksmuseum) en de aanstelling van Hendrik van Wijn als archivarius van de Bataafs Republiek (1802, basis voor het Nationaal Archief).¹² Kunsthistoricus Wessel Krul noemt het instituties ‘waar voorwerpen uit het verleden aan de stroom van de tijd worden onttrokken. Zij worden van vernietiging uitgezonderd en vanwege hun esthetische of historische waarde geconserveerd (...)’.¹³

Misschien is het succes van deze instellingen nog wel opmerkelijker dan de oprichting ervan. Dat succes kan niet los gezien worden van de andere kijk op en beleving van het verleden na de ingrijpende veranderingen die in de revolutionaire periode van de late achttiende en vroege negentiende eeuw hadden plaatsgevonden en die door de opkomende industriële revolutie nog eens werden versterkt. Het kunnen beschikken over sporen uit het verleden was een voorwaarde om de brug naar dat verleden te kunnen slaan. In de loop van de negentiende eeuw is geleidelijk het besef ontstaan dat de overheid een taak had om objecten van cultuur in stand te houden en veilig te stellen.¹⁴

De culture of preservation in de archiefsector

Archieven worden, niet in de laatste plaats door de archiefinstellingen zelf, graag vergeleken met het geheugen van de samenleving.¹⁵ De website *Geheugen van Nederland* biedt toegang tot bijna 800.000 beelden uit 132 collecties van 92 instellingen¹⁶, Het Nationaal Archief noemt zich het nationaal geheugen¹⁷, het Brabants Historisch Informatie Centrum het geheugen van Brabant¹⁸ en in de Universele Verklaring over Archieven van de *International Council on Archives* lezen we dat archieven een

¹¹ Armand-Guy Kersaint sprak hiervan in een ‘Discours sur les monuments public’ op 15 december 1791. Op cit Frans Grijzenhout, *Erfgoed. De geschiedenis van een begrip* (Amsterdam 2000) 7.

¹² N.C.F. van Sas, *De metamorfose van Nederland. Van oude orde naar moderniteit, 1750-1900* (Amsterdam 2005) 360-370.

¹³ Wessel Krul, ‘Tegen het erfgoed: over vooruitgang en vandalisme’ in Grijzenhout (ed), *Erfgoed*, 267.

¹⁴ F.J. Duparc, *Een eeuw strijd voor Nederlands Cultureel Erfgoed* (’s-Gravenhage 1975); Roel Pots, *Cultuur, koningen en democraten. Overheid & Cultuur in Nederland* (Nijmegen 2000).

¹⁵ Al in 1894 werd die vergelijking gemaakt door het archief aan te duiden als ‘het geheugen van den juridischen persoon’ en enkele jaren later werden de archieven als het ‘geheugen’ van stedelijke besturen aangeduid. Toespraak door S. Muller ter opening van de derde jaarvergadering van de Vereniging van Archivarissen in Nederland, *Nederlandsch Archiefblad* (1894/95) nr 2, p. 57; Kolonel F. De Bas, *Nederlandsch Archiefblad* (1896/97) nr 3, p. 52.

¹⁶ <http://www.geheugenvannederland.nl> (geraadpleegd 2 mei 2018)

¹⁷ <https://www.nationaalarchief.nl/over-het-na> (geraadpleegd 2 mei 2018)

¹⁸ <https://www.bhic.nl/het-geheugen-van-brabant> (geraadpleegd 2 mei 2018)

wezenlijke rol spelen in de ontwikkeling van de samenleving 'door het veilig stellen van en het bijdragen aan individueel en gemeenschappelijk geheugen'.¹⁹ Het zijn mooie, aantrekkelijke en op het eerste gezicht ook onschuldige metaforen waar iedereen zich wel wat bij kan voorstellen. Echter, zoals alle metaforen, verhuult ook deze metafoor een aantal essentiële kenmerken van het echte geheugen.

De grondlegger van memory-studies, Maurice Halbwachs, maakte een scherp onderscheid tussen 'memory' aan de ene kant en geschiedenis aan de andere kant. Terwijl het individuele en collectieve geheugen deel uitmaken van levende en dus door mensen beleefde en doorleefde geschiedenis, is het historische geheugen daarentegen kunstmatig en gebaseerd op fragmenten van het verleden in de vorm van artefacten.²⁰ Archieven en musea zijn gericht op het in standhouden van een historisch geheugen en doen dat door de artefacten een welhaast sacrale status te verlenen. Het geheugen is dan ook in hoge mate een 'archieffgeheugen' geworden.²¹ Pierre Nora noemt het zelfs de terreur van het gehistoriseerde geheugen: 'the indiscriminate production of archives is the acute effect of a new consciousness, the clearest expression of the terrorism of historicized memory'.²² Terwijl het collectieve en individuele geheugen door herinneringen tot leven komt en daarmee levend, veranderlijk en altijd in beweging zijn, is het historische geheugen statisch, onveranderlijk en gericht op objecten waarmee reconstructie plaatsvindt. Of zoals Douwe Draaisma schrijft: 'Herinneringen hebben het vermogen achteraf iets anders te worden. De aangetaste herinnering is nog steeds een herinnering, maar niet meer aan wat oorspronkelijk herinnerd werd en daarom is het óók een vorm van vergeten. Het is beide tegelijk'.²³

Archieven en musea opereren binnen de kaders van het historische geheugen. Susan Crane spreekt van een 'culture of preservation' die in de loop van de negentiende eeuw, onder invloed van het historisme, in hoge mate gedomineerd is geraakt door professionele instellingen als archieven en musea.²⁴ Als we de geheugenmetafoor willen gebruiken, moeten we ons realiseren dat het archieffgeheugen een in hoge mate selectief, door instituties georganiseerd objecten-geheugen is, gericht op het onveranderlijk maken van objecten die uit hun oorspronkelijke omgeving zijn gehaald en overgebracht worden naar een kunstmatige omgeving waarmee het natuurlijk proces van herinneren en vergeten wordt verstoord.

Wessel Krul, die zich in zijn geschriften richt op musea, beschrijft het paradoxale karakter van deze instituten omdat museale conservering onherroepelijk leidt tot vernietiging van oorspronkelijke functie en context. In feite hebben musea, en dat geldt evengoed voor archieffinstellingen, hierdoor met een duivels dilemma te maken. Niet opnemen van objecten in een museum of archieffinstelling betekent dat objecten in omloop blijven en in het dagelijkse gebruik zullen slijten of worden aangepast aan de functionele behoeften van het moment. Wel opnemen betekent evenwel evenzeer verlies aan functionaliteit en betekenis door de genoemde selectie en de- en recontextualisering.

¹⁹ Universele Verklaring over Archieven, aangenomen tijdens de algemene vergadering van de ICA in Oslo, september 2010.

²⁰ Crane, 'Writing the individual back', 1376.

²¹ Paul Ricoeur, *Memory, History, Forgetting* (vertaald door Kathleen Blamey en Paul Pellauer) (Chicago-Londen 2006), 403.

²² Pierre Nora, 'Between Memory and History: Les Lieux de Mémoire' in *Representations* 26 (1989) 7-24, aldaar 14.

²³ Douwe Draaisma, *Vergeetboek* (Groningen 2010) 197.

De keuzes die zijn gemaakt met betrekking tot de vraag hoe met dit dilemma om te gaan, is bepalend geweest voor de door Crane genoemde 'culture of preservation'. De keuzes die gemaakt zijn in het verleden werken nog steeds door in het heden. Toen de voorzitter van de *Vereeniging van Archivarissen* Samuel Muller in 1896 de opgave uitzette waarvoor het archiefwezen stond, deed hij niets anders dan herhalen wat vijftig jaar eerder ook al was vastgesteld: "[e]ene vaste organisatie van ons archiefwezen, een centraal toezicht, controle over de plaatselijke archieven, waarborgen tegen gedachteloos vernietigen van archieffstukken, geleidelijke overbrenging der *belangrijke stukken* [mijn cursivering CJ] naar de archieffdepots (...) uitgave van de inventarissen van alle archieven, vaste regelen voor de uniforme inventarisatie (...)'.²⁵ Het is het raamwerk dat ook in de meer dan honderd jaar nadat Muller zijn uitspraken deed, onveranderd is gebleven. Archieffinstellingen proberen sinds de negentiende eeuw veilige bewaaromstandigheden te scheppen door belangrijke archieffbescheiden over te hevelen 'naar een centraal depôt, waar gelegenheid tot goede berging en daarmee behoorlijke bewaring is' en daarmee werd overbrenging in de ogen van de *Vereeniging van Archivarissen* 'een onafwijsbare eisch des tijds'.²⁶ Overbrenging van archieffbescheiden van de administratie naar de archieffinstelling werd het belangrijkste geloofsartikel waarop de 'culture of preservation' van het archiefwezen was gebaseerd. Alleen op die manier, zo was de stellige overtuiging, zou de authenticiteit en integriteit van archieffbescheiden kunnen worden gewaarborgd. De theoretische onderbouwing hiervoor werd nog eens op duidelijke wijze uiteengezet door Luciana Duranti in een artikel waarin ze ageert tegen het postcustodiale idee dat de archieffbewaarplaats een minder relevante positie gaat innemen in het digitale informatiedomein. Duranti gebruikt de 'archival threshold' als beeld van de demarcatielijn tussen de buitenwereld waar documenten gemuteerd kunnen worden en de binnenwereld van de archieffinstelling waar ze onveranderbaar zijn. Paradoxaal genoeg zorgt het passeren van de 'archival threshold' juist voor een fundamentele transformatie. Michel Foucault hanteert in *The Archaeology of Knowledge* het concept van de 'threshold' (*seuil*) om een breuk, een discontinuïteit, een verandering aan te duiden. De 'threshold' in Foucauldiaanse zin markeert de transformatie naar iets anders.²⁷ Het statisch en onveranderlijk maken en het canonerende effect dat uitgaat van het opnemen van objecten in de bewaarplaats van de erfgoedinstelling betekent dat ze in een uitzonderingspositie worden geplaatst ten opzichte van de 'wild frontier'²⁸ van alle dag. Met het passeren van de archieffdrempel krijgen documenten voor Duranti feitelijk een andere status: ze verwerven de kwaliteit van bewijs en geheugen en worden daarmee referentieobjecten. Immers, 'for the transparency of its preservation, its security and its stability, it is necessary that the record pass the archival threshold, the space beyond which no alteration or permutation is possible, and where every written act can be treated as evidence and memory'.²⁹ Ieder document dat de drempel van het archief passeert zal, 'for as long as it exists, [be] truly a permanent monument to it's creator's actions'.³⁰

²⁴ Susan A. Crane, 'Writing the individual back into collective memory' in *American Historical Review*, December 1997, 1372-1385, aldaar 1374-75; Susan A. Crane, *Collecting and historical consciousness in early nineteenth-century Germany* (Ithaca-Londen 2000) 163.

²⁵ Jaarverslag van de voorzitter, *Nederlandsch Archieffblad*, 1896/97 nr 2, p. 25.

²⁶ Brief van het bestuur van de Vereeniging van Archivarissen in Nederland aan de Tweede Kamer der Staten-Generaal, november 1893, gepubliceerd in *Nederlandsch Archieffblad*.

²⁷ Michel Foucault, *The archaeology of Knowledge* (Londen - New York 2004) (reprint), 34, 206

²⁸ John McDonald, *Managing records in the modern office: Taming the Wild Frontier* *Archivaria* 39 (Spring 1995).

²⁹ Luciana Duranti, 'Archives as a place' in *Archives & Manuscripts* 24, 2 (1996): 242-255.

³⁰ *Ibidem*.

Waar Duranti dit mechanisme positief waardeert, en zelfs als noodzakelijk beschouwt, zijn er met name vanuit de hoek van erfgoedstudies ook kritische geluiden te horen. Het archief is net als het museum, de bibliotheek en de archeologische site een plaats waar betekenis wordt gegeven, gemaakt en in stand gehouden. De objecten die op deze plekken worden bewaard, ontlenen een belangrijk deel van hun betekenis aan het feit dat ze door deze instituties worden beheerd. Laurajane Smith spreekt in dit kader van een westers ‘authorized heritage discourse’ waarbij professionals in gevestigde instituties, die hun gezag voortdurend baseren op een omstreden begrip als ‘authenticiteit’, de standaard bepalen welke materiële overblijfselen uit het verleden op welke wijze gepreserveerd behoren te worden.³¹ Verabsolutering en onveranderlijk maken van het materiële object is een belangrijk kenmerk in die benadering en Smith illustreert dit met een treffend voorbeeld uit Australië waar in de jaren tachtig van de vorige eeuw Aboriginal-beheerders van een site beschuldigd werden van het vernielen van traditionele rotsschilderingen (Aboriginal rock-art) door toe te staan dat ze werden overgeschilderd. De beheerders van de site wierpen tegen dat vanuit een Aboriginal perspectief de culturele praktijk van ‘re-painting was vital in keeping alive certain values and meaning in a way that the simple existence of the sites could not’.³²

Preserveringsscenario's

Hoewel de invulling van het begrip authenticiteit cultuur-, plaats- en tijdgebonden is en verschillende disciplines verschillende definities hanteren, bestaat binnen de archiefwereld consensus over het belang van het kunnen garanderen van de authenticiteit van archiefbescheiden.³³ Alleen dan kan archief fungeren als betrouwbaar geheugen en bewijs van hoe personen en instellingen hebben gehandeld. Tegenwoordig wordt het veelomvattender begrip ‘duurzame toegankelijkheid’ gebruikt om een aantal kwaliteiten aan te duiden waaraan archief moet voldoen om te kunnen dienen als betrouwbaar geheugen en bewijs. De erfgoedinspectie omschrijft duurzaam toegankelijke overheidsinformatie als informatie die gedurende de in een selectielijst vastgestelde termijn ‘toegankelijk, vindbaar en in contextuele samenhang raadpleegbaar [moet] zijn. Daarbij moet het zeker zijn dat de informatie juist, volledig en authentiek en leesbaar is’.³⁴ In de definitie van DUTO (de in opdracht van de minister van OCW door het Nationaal Archief ontwikkelde eisen die ervoor moeten zorgen dat overheidsinformatie duurzaam toegankelijk blijft) wordt duurzaamheid omschreven als ‘bestand tegen veranderingen van elke aard’. Waar in het analoge tijdperk de nadruk werd gelegd op het voorkomen van veranderingen in vorm en inhoud van archiefbescheiden, gaat het bij DUTO om

een effect zeker te stellen: toegankelijkheid in de tijd. Duurzame toegankelijkheid betekent in DUTO-termen dat de ‘vindbaarheid, beschikbaarheid, leesbaarheid, interpreteerbaarheid en betrouwbaarheid van informatieobjecten gegarandeerd kunnen worden zolang dat noodzakelijk is’.³⁵

Dat authenticiteit, en vanuit een hedendaags perspectief duurzame toegankelijkheid, alleen veiliggesteld zou kunnen worden door archiefbescheiden de drempel van een archiefbewaarplaats te laten passeren wordt al lang niet meer als enige of zelfs beste optie beschouwd. Theoretici als Terry Cook, Frank Upward en Sue McKemmish wezen er al in de jaren negentig van de vorige eeuw op dat in een digitale omgeving overbrenging naar een archiefbewaarplaats geen wezenlijke bijdrage levert aan het kunnen garanderen van de authenticiteit als die niet vanaf het moment van creatie geborgd is.³⁶ Denkend vanuit het theoretisch raamwerk van het *records continuum* nemen archiefinstellingen niet langer een centrale en allesbepalende positie in maar zijn ze onderdeel van een veel groter archivistisch multiversum waarin verschillende tradities, percepties en organisatievormen van archiveren naast elkaar bestaan.³⁷ David Bearman en Margaret Hedstrom stelden zelfs dat digitale archieven beter in hun oorspronkelijke omgeving kunnen blijven omdat het overbrengen van digitale archieven naar een archiefbewaarplaats geen garantie biedt voor kwalitatief goed beheer. Onderhoud van soft- en hardware en de continue migratie van data en applicaties vereisen veel geld en kennis en ‘[t]his puts records in archival custody at relatively greater risk than those whose on-going management is regulated by archivists but which remain in the physical custody of agencies that created them’.³⁸ Sheila Powell, werkzaam voor het *archival electronic records program* van het Canadese Nationaal Archief, beschreef al in 1995 drie verschillende benaderingen om lange termijn preservering van archieven vorm te geven: ‘custodial approach, distributed custody approach, combination of custodial and distributed custody approaches’.³⁹ Bij de custodiale benadering worden de kwaliteitskenmerken van archieven gegarandeerd door ze uit hun oorspronkelijke omgeving naar de archiefinstellingen over te brengen. De functies van waardering, acquisitie, ordening, beschrijven, preserveren, beheer en beschikbaar stellen worden allemaal uitgevoerd door de archiefinstelling waarnaar de archieven worden overgebracht. Bij de ‘distributed custody approach’, die ook wel bekend staat als de non-custodiale benadering, heeft de archiefinstelling een adviserende en monitorende functie ten aanzien van archivering, en is zij betrokken bij waardering en eventueel bij het beschrijven van archieven. Alle andere functies zoals beheer, preservering en beschikbaar stellen worden uitgeoefend door de archiefvormende instelling of de rechtsopvolger ervan. Tot slot onderscheidt Powell de combinatie variant die, zoals de naam al zegt, aspecten van de custodiale en de gedistribueerde benadering

³¹ Laurajane Smith, *Uses of Heritage* (London-New York 2006) 125.

³² *Ibidem*, 54.

³³ Een veel gebruikte definitie is te vinden op de InterPares website: ‘The trustworthiness of a record as a record; i.e., the quality of a record that is what it purports to be and that is free from tampering or corruption’ http://www.interpares.org/ip2/display_file.cfm?doc=ip2_glossary.pdf&CFID=15883021&CFTOKEN=73665030. Richard Pearce-Moses definieert authenticiteit in de *Glossary of Archival and Records Terminology* (Chicago 2005) als ‘the quality of being genuine, not a counterfeit, and free from tampering, and is typically inferred from internal and external evidence, including its physical characteristics, structure, content, and context’ (p. 41).

³⁴ Erfgoedinspectie, *Wel digitaal, nog niet duurzaam. Informatiebeheer bij de departementen* ('s-Gravenhage 2018).

³⁵ DUTO Wiki, <https://wiki.nationaalarchief.nl/pagina/DUTO:Begrippen> (geraadpleegd 18 juli 2018).

³⁶ Adrian Cunningham, ‘Archival Institutions’ in Sue McKemmish, Michael Pigogott, Barbara Reed and Frank Upward (eds), *Archives: Recordkeeping in Society* (Wagga Wagga 2005) 21-50.

³⁷ Ann J. Gilliland, ‘Archival and recordkeeping traditions in the multiverse and their importance for researching situations and situating research’ in Anne J. Gilliland, Sue McKemmish and Andrew J. Lau (eds), *Research in the Archival Multiverse* (Melbourne 2017) 31-73, aldaar 50. De archival multiverse wordt gedefinieerd als ‘encompassing the pluralism of evidentiary texts, memory-keeping practices and institutions, bureaucratic and personal motivations, community perspectives and needs, and cultural and legal constructs with which archival professionals and academics must be prepared, through graduate education, to engage’.

³⁸ David Bearman and Margaret Hedstrom, ‘Reinventing Archives for Electronic Records: Alternative Service Delivery Options’ in Margaret Hedstrom (ed.), *Electronic Records Management Program Strategies* (Pittsburgh: Archives and Museum Informatics, 1993), p. 87.

³⁹ Sheila Powell, *Elements of an Archival Electronic Records Program* (Ottawa: National Archives of Canada, 1995).

combineert: sommige archieven zullen worden geacquireerd door een archiefinstelling en andere blijven beheerd door en kunnen worden geraadpleegd via de archiefvormer.⁴⁰

Dominant custodiaal perspectief

Als we de drie scenario's als uitgangspunt nemen om naar de Nederlandse situatie te kijken, kunnen we vaststellen dat meer dan twintig jaar na verschijnen van dit rapport de custodiale benadering niet alleen nog steeds springlevend is, maar ook dat archivarissen en archiefinstellingen de meeste tijd en energie stoppen in het operationaliseren van een digitale variant. Dat is wereldwijd het geval. De archiefomgeving, zo is de overtuiging, zorgt voor de blijvend te bewaren archieven immers voor 'adequate controls to guarantee that records will be preserved and that they will not be altered'.⁴¹ De samenwerking met de archiefvormers is in deze benadering vooral gericht op 'to ensure that all of the archives' needs for custodial purposes are met, i.e. that migration to the archives' system can be accomplished as efficiently and inexpensively as possible (...). [T]ransfer agreements must stipulate the transfer specifications, and the archives must keep these conditions up to date in order to avoid costly delays and problems upon transfer'.⁴²

Het Nederlandse openbare archiefwezen heeft zijn pijlen gericht op de ontwikkeling van een voorziening die is gebaseerd op het overbrengen van informatieobjecten uit bedrijfssystemen van overheidsorganen naar een e-depot.⁴³ Overbrenging is immers stevig verankerd in de Nederlands archiefwetgeving en de archiveringsketen eindigt, zoals de Erfgoedinspectie in haar onlangs verschenen rapport *Wel Digitaal, Nog niet Duurzaam* benadrukte, vanuit dat perspectief logischerwijze in het e-Depot. In genoemd rapport neemt de Erfgoedinspectie de staat van het digitaal informatiebeheer bij de departementen onder de loep en heeft daarbij onder meer naar de aansluiting op het e-Depot gekeken. Zij concludeert dat in de afgelopen jaren weliswaar vooruitgang is geboekt door de duidelijkheid die het Nationaal Archief heeft gegeven over de eisen voor aansluiting op het e-Depot en door een aantal pilots uit te voeren, maar stelt tegelijkertijd vast dat 'het verplaatsen [mijn cursivering CJ] van digitaal archief van een departement naar het e-Depot (...) nog niet voorbij het experimentele stadium (is)'.⁴⁴

In het Nederlandse archiefbestel wordt de discussie om een ander model dan het overbrengingsmodel te hanteren nauwelijks gevoerd.⁴⁵ De aangekondigde modernisering van de archiefwet zou het uitgelezen moment zijn om dit model kritisch onder de loep te nemen al is de kans dat dit gebeurt niet erg groot. De roep om moderni-

⁴⁰ *Ibidem*.

⁴¹ *Ibidem*. Het citaat is afkomstig van Ken Thibodeau, hoofd van de US National Archives and Records Administration, Center for Electronic Records, die het onderscheid benadrukte tussen 'operational environment' en 'archival environment'.

⁴² Sheila Powell, *Elements of an Archival Electronic Records Program*.

⁴³ <https://www.nationaalarchief.nl/archiveren/kennisbank/aansluiten-op-het-e-depot> (geraadpleegd 27 juli 2018).

⁴⁴ Erfgoedinspectie, *Wel digitaal, nog niet duurzaam. Informatiebeheer bij de departementen* ('s-Gravenhage 2018) 33.

⁴⁵ Heel voorzichtig wordt in de brief waarin minister Slob de Tweede Kamer informeert over het voornemen om de archiefwet aan te passen bespreekbaar gemaakt dat bepaalde datacenters (bijvoorbeeld CBS, DANS, kadaster, RWS) die ook data bewaren die voor permanente bewaring in aanmerking komen, misschien als archiefbewaarplaats in de wettelijke zin kunnen gaan functioneren zodat data dus niet 'fysiek' overgebracht zouden hoeven worden. Zie Kamerbrief, d.d. 11 juni 2018.

sering van de archiefwet komt op uitdrukkelijke wens van de politiek nadat de onderzoeken van de commissie Oosting en de Erfgoedinspectie naar de gang van zaken rond de ontnemingschikking tussen het Openbaar Ministerie en Cees H. er veel schortte aan de informatiehuishouding bij de overheid. Op initiatief van Kamerlid Segers werd een motie aangenomen waarin de regering werd verzocht de archiefwet aan te passen aan de digitale ontwikkelingen en eisen van transparantie, en gaven ook meteen een suggestie voor een maatregel om dit te bewerkstelligen: verkorting van de overbrengingstermijn.⁴⁶ Deze motie is voor het openbaar archiefwezen een belangrijke katalysator geworden om de discussie over modernisering te voeren. Tegelijkertijd is de motie, vanwege de nogal dwingende inhoudelijke suggestie om de overbrengingstermijn te verkorten, erg sturend voor de wijze waarop de discussie wordt gevoerd. Het idee dat het eerder overbrengen van digitale archieven naar een veilige omgeving de duurzaamheid en toegankelijkheid van deze archieven garandeert, wordt binnen de archiefsector breed gedeeld.⁴⁷

Alternatieven

Ronald Rommelse heeft in zijn met de Duparcprijs bekroonde scriptie uit 2015, waarin hij de grote e-depotprojecten in Nederland onderzocht, laten zien dat ze allemaal lijden aan het zogeheten *horseless carriage syndroom*: het digitale archief wordt begrepen als het analoge archief maar dan gevormd met een nieuwe techniek.⁴⁸ Ze zijn een logisch uitvloeisel van de dominante *culture of preservation*. De e-depots zijn een continuering van het custodiale denken maar dan met gebruikmaking van digitale technieken. Hij concludeert dat het e-depot een voortzetting is van het levensloopmodel; dat het uitgaat van beheer op één centrale plek op basis van generiek opgelegde kwaliteitseisen; dat het is gericht op statische informatiebestanden en dat duurzame toegankelijkheid gezien wordt als een eigenschap van het archief.⁴⁹ Innovatie is een complex, iteratief proces dat met horten en stoten verloopt. Het loslaten van bestaande kaders – zeker als het voortbestaan van bestaande organisatorische structuren op het spel staat – is geen gemakkelijke opgave. Het custodiale, op overbrenging gebaseerde levensloopmodel vormt immers de bestaansbasis van het huidige archiefbestel. Het is, ondanks de archieftheoretische beginselen van *respect des fonds*, in essentie een stelsel dat gericht is op het beheren en veiligstellen van objecten door ze onveranderlijk te maken. Die benadering zien we terug in de archiefwetgeving. Het object van de archiefwet is immers niet 'archief' in de zin van 'het geheel van archiefbescheiden'⁵⁰, maar 'archiefbescheiden' in de zin dat ze door overheidsorganen zijn ontvangen of opgemaakt.⁵¹ Het lijkt op het eerste oog

⁴⁶ Motie Segers, 8 juni 2016, Kamerstukken 34 362, nr. 21. De motie spreekt overigens van 'overbrengingstermijnen'.

⁴⁷ Gezamenlijke brief van de besturen van BRAIN en KVAN aan de minister van OCW over de aanpassing van de archiefwet aan digitale ontwikkelingen, dd 10 oktober 2017. http://kvan.nl/images/Nieuws/171010_brief_aan_minocw_verkorting_overbrengingstermijn_kopie.pdf (geraadpleegd 27 juli 2018) en Natascha Stroeker, Guido Bummelkamp, Freek Kuiperi, Enno Gerdes, *Impact verkorting overbrengingstermijn. Onderzoek naar de impact van een substantiële verkorting van de overbrengingstermijn (Archiefwet 1995)* (Zoetermeer 2017). Dat is overigens veel minder het geval ten aanzien van de verwachting van de indieners van de motie dat vervroegde overbrenging ook een positief effect zal hebben op de transparantie en openbaarheid. Overbrenging is immers niet hetzelfde als openbaar maken.

⁴⁸ Ronald Rommelse, *Het e-depot, een horseless carriage. Een onderzoek naar duurzame toegankelijkheid van digitaal overheidsarchief door middel van e-depotvoorzieningen* (UvA 2015).

⁴⁹ *Ibidem*.

⁵⁰ *Archiefterminologie voor Nederland en Vlaanderen* (2003).

⁵¹ Archiefwet 1995, artikel 1.

misschien een triviaal onderscheid, maar zet bij nadere beschouwing de deur open naar de genoemde objectgerichte inrichting van een aantal archiveringsfuncties. Dat blijkt bijvoorbeeld uit de wijze waarop de in de archiefwet vastgelegde bepaling van goede, geordende en toegankelijke staat is uitgewerkt. In de archiefregeling zijn de artikelen 3 tot en met 26 gewijd aan duurzaamheid en aan geordende en toegankelijke staat van archiefbescheiden. De in deze artikelen genoemde eisen en maatregelen zijn echter alleen van toepassing op archiefbescheiden die volgens een selectielijst naar een archiefbewaarplaats moeten worden overgebracht. Het is een duidelijk voorbeeld van bovengenoemde stelling dat niet het archief als geheel maar geselecteerde archiefbescheiden hier het uitgangspunt vormen. De archiefwetgeving hanteert hiermee een perspectief van wat Frank Upward cs het ‘beheren van archiefbescheiden als eindproducten’ (‘manage objects as end products’) aanduidt.⁵² Het is een benadering die voor tal van digitale toepassingen niet langer voldoet om de eenvoudige reden dat zij volledig voorbijgaat aan de eigenschap van beweeglijkheid die in veel informatietoepassingen geldt.⁵³

Een andere en in veel opzichten vruchtbaarder benadering zou kunnen zijn niet langer uit te gaan van een selectie van archiefbescheiden (als eindproducten van handelen) die duurzaam toegankelijk moeten blijven maar het perspectief te verschuiven naar transacties die aan de basis liggen van het vastleggen, verwerken en uitwisselen van gegevens die een betrouwbare, gezaghebbende en duurzame vorm van informatiemanagement vereisen.⁵⁴ Een dergelijke benadering wint vanuit ontwerpgericht onderzoek met een gerichtheid op verbeteren van praktijksituaties geleidelijk (ook binnen het informatiedomein) terrein.⁵⁵

Het grote verschil tussen beide benaderingen zit vooral in het gekozen uitgangspunt: geen generiek opgelegde kwaliteitseisen die toegepast dienen te worden om objecten volgens een van tevoren vastgelegde standaard in een digitaal magazijn te kunnen beheren, maar per toepassing kijken welke functionele duurzaamheidseisen nodig zijn om de informatie betekenisvol te houden. Vanuit dat perspectief kan het in sommige gevallen verstandiger zijn om informatie niet over te brengen maar de pre-serveringsinspanningen te richten op de informatiesystemen en applicaties bij de archiefvormers. Niet het beheer van objecten maar veiligstellen van functionaliteit die nodig is om de informatie betekenisvol te houden zou prioriteit moeten krijgen. Richard Rinehart en Jon Ippolito pleiten in hun boek *Re-Collection* behalve voor het integreren van formele en informele pre-serveringsstrategieën, ook voor het met andere ogen naar pre-servering te kijken. Hoewel zij zich in hun boek op digitale kunst richten, zijn veel van hun conclusies ook relevant voor de archivering in het administratieve domein. Fixatie op het ‘origineel’ is in hun ogen volstrekt zinloos omdat er niet langer een origineel is. Onveranderlijkheid is bij veel digitale mediatoe-passingen juist een corrumperende kracht die leidt tot vernietiging van de essentie; een gevaar dat bij bestaande strategieën van webarchivering, databasearchivering en archivering van bouwwerkinformatiemodellen (BIM) in het administratieve domein evenzeer aanwezig is. ‘Curators and archivists’ zo concluderen ze, ‘must be able to dust themselves off and become futurists and reenactors. They must be able

⁵² Frank Upward, Barbara Reed, Gillian Oliver and Joanne Evans, *Recordkeeping Informatics for a Networked Age* (Clayton 2018) 97 en 120.

⁵³ Ibidem. Zie ook Frans Smit, Arnoud Glaudemans, Rienk Jonker (eds) *Archives in Liquid Times* ('s-Gravenhage 2017).

⁵⁴ Upward cs, *Recordkeeping Informatics*.

⁵⁵ Voor ontwerpgericht wetenschappelijk onderzoek zie Joan van Aken en Daan Andriessen (eds) *Handboek ontwerpgericht wetenschappelijk onderzoek. Onderzoek met effect* (2011).

to describe the behaviors of an artwork with the same rigor, authority, and even linguistic specificity that they now use to describe their form (...) they need to become expert at embracing and managing change in addition to fixity’.⁵⁶ De noodzaak om de ‘culture of preservation’ te bevrijden van de dominante benadering waarbij objecten in de tijd worden bevroren zien we ook terug bij Upward, Reed, Oliver en Evans. Zij stellen ‘[f]ixing in the sense of freezing something in aspic is not the requirement. Rather a more flexible interpretation would be to enable the ‘replaying’ of sequences of actions in order to prove or disprove the state of something or the conclusions reached’.⁵⁷ Dat zou bereikt kunnen worden door naar analogie van de wetenschappelijke wereld de herhaalbaarheid van manifestaties centraal te stellen. Niet alleen de dataset maar ook de instructies (toegepaste methode) zijn nodig om een manifestatie uit het verleden weer tot leven te kunnen wekken. In de *performing arts* zijn instructies in de vorm van partituren en dansnotaties altijd al noodzakelijk geweest om het per definitie efemere karakter van de uitvoering herhaalbaar te maken. Ook in het administratieve domein zullen steeds meer ‘events’ ofwel manifestaties niet meer te vangen zijn in over te brengen statische representaties.⁵⁸

Conclusie

In deze bijdrage heb ik benadrukt dat de dominante ‘culture of preservation’ binnen het Nederlandse openbare archiefwezen nog steeds is gebaseerd op het veiligstellen van objecten in -weliswaar digitale- archiefmagazijnen. Het Nederlandse archiefbestel is gebaseerd op en georganiseerd rond het overbrengingsmechanisme. Dit model dient om verschillende redenen geflexibiliseerd te worden. In de eerste plaats houdt het model een kunstmatige kloof in stand tussen het administratieve domein en het culturele domein waardoor de overbrenging zowel een eindpunt als een beginpunt markeert. Veel energie, geld en middelen lekken weg om informatieobjecten van het ene domein naar het andere domein te kunnen overhevelen. De niet aflatende pogingen en frustraties om de achterstanden in selectie en bewerking van archieven weg te werken is daarvan wel het beste voorbeeld. Dat laat meteen een tweede belangrijk reden zien waarom flexibilisering van dit model noodzakelijk is. Het bestaande model werkt in de hand dat de nadruk is komen liggen op objecten (archiefbescheiden, bestanddelen). De objecten die de ‘threshold’ tussen beide domeinen met succes weten te passeren worden voor altijd duurzaam toegankelijk gehouden. Het leidt tot wat ook wel denigrerend ‘eindproductenbeheer’ is genoemd. Denigrerend omdat specifieke functionaliteiten en relevante kenmerken vanuit de specifieke toepassingen (transacties) verwaarloosd of zelfs geheel genegeerd worden. Door het informatieobject als voorwerp van pre-servering centraal te stellen en niet de specifieke informatietoepassing (transactie) wordt de eerdergenoemde kloof in stand gehouden, worden objecten altijd uit hun oorspronkelijke systemen en applicaties gehaald en is er veel capaciteit nodig om in de nieuwe omgeving een aantal kenmerken uit de vroegere toepassingen kenbaar te houden.

⁵⁶ Richard Reinhart, Jon Ippolito, *Re-collection. Art, New Media, and Social Memory* (Cambridge 2014) 233.

⁵⁷ Upward cs, *Recordkeeping Informatics*, 236.

⁵⁸ In mijn artikel ‘Threats of the data flood. An accountability perspective in the era of ubiquitous computing’ verschenen in Frans Smit, Arnoud Glaudemans, Rienk Jonker (eds) *Archives in Liquid Times* ('s-Gravenhage 2017) 196-210 geef ik daarvan de nodige voorbeelden. Zie ook mijn Amsterdamse oratie *Het archief is dood. Leve het archief!* (Amsterdam 2017) http://cf.bc.uva.nl/download/oraties/Oraties_2017/Jeurgens_Charles.pdf

Ik pleit dan ook voor flexibilisering van het overbrengingsmodel. Flexibilisering betekent dat in bepaalde gevallen overhevelen van objecten (archiefbescheiden) naar een veilige omgeving nog steeds functioneel en relevant kan zijn, maar dat het niet langer de enige vorm is. Veel belangrijker nog is dat een meerdimensionale organisatievorm van conserveren mogelijk wordt die gebaseerd is op de functionele behoeften en eisen die gesteld worden aan het kenbaar houden van overheidstransacties. Het betekent ook, en daarvoor is een veel grotere transformatie nodig, dat het archiefbestel zich weet te bevrijden van de institutionele logica (en dwangmatigheid) dat overheidsinformatie alleen duurzaam toegankelijk gehouden kan worden door het beheer ervan in handen van archiefinstellingen te geven. Het bouwen van een duurzame infrastructuur, waarvan beveiligde digitale 'depots' nog steeds onderdeel uitmaken maar niet langer het monopolie hebben, met meer aandacht voor functionaliteiten die duurzaam beschikbaar blijven, vereist niet alleen een groter repertoire aan technische en organisatorische scenario's maar ook een flexibeler archiefbestel dat in staat is vanuit een continuüm perspectief te opereren.

Over de auteurs

Annet Dekker is Assistant Professor Media Studies: Archival Science at the University of Amsterdam, and Visiting Professor and co-director Centre for the Study of the Networked Image at London South Bank University. Her recent publications include 'Collecting and Conserving Net Art. Moving Beyond Conventional Methods' (Routledge 2018) and 'Lost and Living (in) Archives' (Valiz 2017).

Afelonne Doek is Directeur Collecties en Digitale Infrastructuur bij het Internationaal Instituut voor Sociale Geschiedenis en tevens Directeur Collecties en Diensten bij NIOD Instituut voor Oorlogs-, Holocaust- en Genocidestudies, beiden gevestigd in Amsterdam. Zij is betrokken bij initiatieven en (Europese) projecten op het snijvlak van collecties en digitale infrastructuur, waaronder het HOPE project in 2010-2013.

Mette van Essen is als Adviseur Digitale Innovatie verbonden aan het Nationaal Archief. Zij is sinds 2005 werkzaam in de archiefsector. Vanaf het begin heeft zij zich beziggehouden met Digital Preservation en ontwikkelingen rondom een E-Depot. De focus van haar werkzaamheden ligt tegenwoordig op onderzoek naar nieuwe technologieën en hoe deze in te zetten binnen de informatie- en archiefwereld.

Margriet van Gorsel is Hoofd Toezicht, Beleid en Advies bij het Stadsarchief Amsterdam. In de twintig jaar die Margriet werkzaam is in het vakgebied van informatie- en archiefbeheer heeft zij verschillende verantwoordelijkheden op zich genomen. In 1998 is zij gestart bij het Stadsarchief Rotterdam als inspecteur en eDepot ontwikkelaar. Daarna is zij kort werkzaam geweest als manager collectiebeheer om zich daarna als zelfstandige zonder personeel te wijden aan toezicht, onderwijs, voorlichting, beleid en (programma)management. In 2013 heeft zij haar kennis van preservering verdiept als preservation officer bij het Nationaal Archief en vanaf januari 2018 is zij werkzaam als Hoofd Toezicht, Beleid en Advies bij het Stadsarchief Amsterdam.

Hans Hofman was tot 2013 werkzaam bij het Nationaal Archief als senior adviseur. In die hoedanigheid was hij o.a. betrokken bij verschillende Europese en internationale projecten op het gebied van digital preservation en was hij actief op het gebied van standaardisatie.

Erika Hokke is vrijgevestigd archivaris. Sinds 1997 houdt zij zich bezig met de vraag hoe digitale informatie duurzaam toegankelijk gehouden kan worden. Aanvankelijk als adviseur bij het programma Digitale Duurzaamheid, later als docent-onderzoeker bij de Archiefschool. Sinds 2010 is zij directeur-eigenaar van STROOM in, advies en opleidingen. In haar vrije tijd werkt zij aan een proefschrift over de ontwikkeling van het beroep archivaris in de 19e-21e eeuw.

Charles Jurgens is hoogleraar archiefwetenschap aan de Universiteit van Amsterdam en als adviseur verbonden aan het Nationaal Archief. Hij was eerder onder andere werkzaam als gemeentearchivaris van Schiedam en Dordrecht en als hoogleraar archivaliek aan de Universiteit Leiden.

Annemieke de Jong is als Digital Preservation Officer verantwoordelijk voor het ontwikkelen van (strategisch) beleid voor het toekomstvast beheer en het duurzaam toegankelijk houden van de digitale media collecties van het Nederlands Instituut voor Beeld en Geluid.

William Kilbride is Executive Director of the Digital Preservation Coalition (DPC), an international not-for-profit membership organization based which provides community engagement, advocacy, workforce development, capacity, good practice and standards in digital preservation. William started his career as an active researcher in archaeology in the 1990s when the discipline's enthusiasm for new technology outstripped its capacity to manage the resulting data.

Tamar Kinkel is Senior Adviseur bij BMC Advies. Vanuit die functie adviseert zij organisaties in het publieke domein op het gebied van financiën en bedrijfsvoering. Zij is als adviseur via het Netwerk Digitaal Erfgoed betrokken bij de ontwikkeling en uitrol van het Kostprijsmodel Digitale Duurzaamheid.

Wim Lowet is archivaris en werkt sinds 2013 voor het Vlaams Architectuurinstituut. Zijn werk focust voornamelijk op de preservatie van born digital archieven van architecten, waarvoor hij verschillende onderzoeksprojecten leidde. Sinds de creatie van Architectuurarchief Vlaanderen in 2018 is hij verantwoordelijk voor het preservatiebeleid voor de digitale collecties van de organisatie.

Becky McGuinness is Community Manager Open Preservation Foundation. Becky leads the Foundation's programme of outreach and knowledge sharing activities including webinars, interest groups, workshops, and internal and external communications. She holds a Professional Postgraduate Diploma in Marketing, and a degree in English Literature.

Joost van der Nat is sinds 2014 (deeltijd) onderzoeker voor het Werkprogramma Digitaal Erfgoed Houdbaar van het Netwerk Digitaal Erfgoed. In 2015 verscheen het onderzoek "Samen bouwen aan een netwerk van landelijke voorzieningen" dat in opdracht van de Nationale Coalitie Digitale Duurzaamheid werd uitgevoerd. Het onderzoek ontving in 2016 de Digital Preservation Award in de categorie 'Research and Innovation'. Dit onderzoek vormt mede de basis voor de activiteiten die nu in het Werkprogramma Digitaal Erfgoed Houdbaar worden uitgevoerd. Joost is medeoprichter van het organisatieadviesbureau Singel & Partners (2007). Daarvoor werkte hij bij o.a. de Verenigde Naties, Consultdata / Bakkenist Management Consultants, en Deloitte Consulting.

Frans Negers is digitaal archivaris bij Het Nieuwe Instituut in Rotterdam, de beheerder van de Rijkscollectie Architectuur en Stedenbouw. Binnen zijn rol als projectleider en coach voor de blended leeromgeving Leren Preserveren brengt hij erfgoedcollega's bij elkaar rond de opgave van duurzame toegankelijkheid.

Bart De Nil is sectorcoördinator archieven bij FARO, Vlaams steunpunt voor cultureel erfgoed vzw. Bart is historicus en werkt sinds 2008 voor FARO. Hij was van 2007 tot en met 2009 projectleider van het IBBT-GBO innovatieproject Erfgoed 2.0 dat zich heeft beziggehouden met digitaal erfgoed en mobiele devices én van 2009 tot en met 2011 van het IBBT-ICON innovatieproject EPICS dat onderzoek heeft verricht rond cultureel erfgoed en digitale leerplatformen. In zijn vrije tijd engageert Bart zich voor verschillende erfgoed- en buurtprojecten. Over al deze interessevelden publiceerde hij reeds talrijke monografieën en artikelen.

Matthias Priem is manager archivering bij VIAA. Matthias is verantwoordelijk voor alles wat archivering betreft. Daaronder vallen infrastructuur, software, processen en metadata. Hij zet het beleid uit, coacht het team en volgt lopende projecten op. Verder houdt hij vinger aan de pols omtrent technologie en zet waar opportuun in op innovatieve ontwikkelingen.

Marcel Ras is programmamanager digitale duurzaamheid bij het Netwerk Digitaal Erfgoed, een samenwerkingsverband van Nederlandse erfgoedinstellingen gericht op het ontwikkelen van een stelsel van landelijke voorzieningen en diensten voor het verbeteren van de zichtbaarheid, bruikbaarheid en houdbaarheid van digitaal erfgoed. Marcel is afgestuurd als historicus en archeoloog en al ruime tijd werkzaam op het terrein van digitale duurzaamheid, als projectleider, afdelingshoofd en programmamanager.

Laurents Sesink is hoofd van het Centre for Digital Scholarship (CDS) bij de Universitaire Bibliotheken Leiden (UBL). De doelstelling van het CDS is om 'Digital Scholarship' binnen de Universiteit Leiden te ondersteunen en te faciliteren. Onderzoekers kunnen bij het CDS terecht voor informatie, vragen, advies, training en workshops met betrekking tot datamanagement, Open Access, Data Science en digitale duurzaamheid. Laurents is onder meer co-auteur van het Data Seal of Approval en co-auteur van het activity-based costing model for long-term preservation.

Barbara Sierman werkt sinds 2005 bij de Koninklijke Bibliotheek als digital preservation manager op de Onderzoeksafdeling, van waaruit zij de organisatie adviseert op het gebied van digitale duurzaamheid. Heeft sindsdien deelgenomen aan verschillende Europese projecten (o.a. Planets, SCAPE) en was co-auteur van de ISO 16363 standaard op het gebied van certificering. Internationale netwerken zijn belangrijk voor de ontwikkelingen in digitale duurzaamheid, vandaar dat ze namens de KB zitting heeft in de Stuurgroep van de IIPC en Chair of the Board of Directors is van de Open Preservation Foundation.

Herman Uffen is senior Adviseur Financiën en Bedrijfsvoering bij BMC Advies. Hij adviseert organisaties in het publiek domein op het gebied van financiën, bedrijfsvoering en organisatie-inrichting. Tevens is hij projectleider bij het Netwerk Digitaal Erfgoed voor het Kostprijsmodel Digitale Duurzaamheid.

Remco van Veenendaal is preservation officer bij het Nationaal Archief. Hij heeft meer dan 15 jaar werkervaring op de snijvlakken van taal, technologie, data en archieven, en van onderzoek, onderwijs, onderneming en overheid. Hij is een ervaren adviseur, bestuurslid, projectmanager, taaltechnoloog en onderwijzer. Sinds maart 2015 werkt hij voor het Nationaal Archief, eerst als preservation adviseur, en tegenwoordig als preservation officer. In 2018 trad hij toe tot de board of directors van de Open Preservation Foundation.

Donald Weber is Hoofd Onderzoek bij Amsab-Instituut voor Sociale Geschiedenis en tevens verbonden aan de Vakgroep Geschiedenis van de Universiteit Gent. Hij was betrokken bij meerdere Europese projecten waaronder het HOPE-project in 2010-2013.

Gaby Wijers is oprichter en directeur van LIMA. Voorheen was zij hoofd collecties en onderzoek bij het Nederland Instituut voor Mediakunst (NIMk) en het Theater Instituut Nederland (TIN), beide in Amsterdam.

Carl Wilson is the Technical Lead for the Open Preservation Foundation, overseeing all of OPF's technical activities. He is an experienced software engineer with a focus on software quality through testing. He is an open source enthusiast, both as a user and developer.

Martin Wrigley is Executive Director Open Preservation Foundation. Martin is responsible for working with OPF members and the Board to develop the OPF's strategy and enhance the Foundation's portfolio of open source digital preservation tools. He brings over 30 years experience of working in the software industry, including software development, mobile telecoms and phone apps.

E-depot?



MAIS-Flexis beheerst

- digitaal atelier
- ingest van digitaal archief en metadata (TMLO/Topx)
- duurzame digitale opslag
- beheer van digitaal archief
- preserveren
- ontsluiting

MAIS-Flexis het veelzijdige en complete archief- en collectiebeheersysteem, ook voor E-depot



Erg goed in Erfgoed

DE REE archiefsystemen
de-ree.nl | archieven.nl | 050 311 26 00