



ANNEMIEKE DE JONG

Annemieke de Jong is beleidsadviseur bij de Unit Infrastructuur, Beheer en IT van het Nederlands Instituut voor Beeld en Geluid. Zij is momenteel speciaal belast met het ontwikkelen van strategisch beleid voor het duurzaam behoud van de digitale collectie. De Jong heeft een lange staat van dienst in het AV-archiveringsdomein en werkte o.a. als manager Informatiebeleid, projectleider van innovatieprojecten en hoofd Catalogus. Zij is actief in meerdere (inter)nationale expertgroepen op het gebied van mediamanagement, digitaal AV-erfgoed en digitale conservering. Daarnaast werkt zij sinds drie jaar als projectleider Kennisdisseminatie van Beeld en Geluid. De Jong was een van de oprichters van AVA_Net en fungeert sinds 2008 als hoofdredacteur voor de Kennisbank AV Archivering, het kennisplatform van AVA_Net en Beeld en Geluid. Van haar hand verschenen diverse publicaties over theoretische en praktische ontwikkelingen met betrekking tot het beheren en ontsluiten van digitale AV-collecties. Annemieke de Jong studeerde Nieuwe Media en Digitale Cultuur aan de Universiteit van Utrecht en Informatiemanagement aan de Erasmus Universiteit in Rotterdam.



ERNST VAN VELZEN

Ernst van Velzen vervult de functie Unitmanager Infrastructuur, Beheer en IT bij het Nederlands Instituut voor Beeld en Geluid. Daarvoor was hij hier ruim drie jaar werkzaam als IT-manager. Eerder werkte Van Velzen tien jaar in de ICT, o.m. als projectleider en informatiespecialist. Hij verrichtte voorheen sociaalwetenschappelijk onderzoek voor het Sociaal en Cultureel Plan Bureau en de Erasmus Universiteit. Bij Beeld en Geluid leidt Van Velzen alle IT-operaties die nodig zijn voor het uitvoeren van de archivale kernactiviteiten alsmede voor de ondersteunende processen. Hij is ook verantwoordelijk voor de ICT-voorzieningen binnen het Media Experience en de uitvoering van de massadigitalisering in het kader van het project Beelden voor de Toekomst. Speciaal aandachtsgebied binnen de Unit vormt de ontwikkeling en het beheer van de technische infrastructuur voor het digitale archief, dat inmiddels ruim 4 petabyte aan audiovisuele content bevat. Van Velzen's focus hierbij is het realiseren van de condities voor de status van Beeld en Geluid als Trusted Digital Repository voor het digitale audiovisuele domein in Nederland. Ernst van Velzen studeerde politieke wetenschappen aan de Universiteit van Amsterdam.



WAT ALLES OVERHEERST
IS DE GROTE
AFHANKELIJKHEID VAN
DE TECHNOLOGIE, DIE
ZELF VOORTDUREND IN
BEWEGING IS...

HET AUDIOVISUEEL ARCHIEF ALS DIGITAAL CURATOR

HET NIEUWE COLLECTIEMANAGEMENT, EEN INTRODUCTIE

Annemieke de Jong en Ernst van Velzen

Digitalisering kan worden gezien als een belangrijke nieuwe stap in het vastleggen en beschikbaarstellen van het audiovisueel geheugen. AV-archieven gaan beter doen wat ze altijd al deden: het verzamelen en bewaren van erfgoedcollecties. Door digitale netwerken kunnen beeld en geluid op grote schaal worden aangeboden. De archiefmasters worden in optimale omstandigheden opgeslagen terwijl het publiek en de professionele gebruikers versies krijgen voorgeschoteld op een voor hen bruikbare norm. Het wordt gemakkelijker materiaal te bewaren. Het opslaan gebeurt in feite automatisch, op het moment van productie. Digitale opslag leidt daarbij tot een enorme reductie van de benodigde fysieke ruimte. De opslag wordt ook efficiënter. Alle gegevens, ongeacht hun oorsprong, worden op dezelfde manier 'weggezet'. Hierdoor kunnen grote bestanden makkelijker worden georganiseerd en doorzocht. Starre, hiërarchische classificatiesystemen maken plaats voor een flexibele en rijke indexering, automatisch gegenereerd. Semantische netwerken brengen onverwachte en creatieve dwarsverbanden aan, tussen talloze verspreide archiefcollecties, waar ook ter wereld. Door de grote toegankelijkheid van de informatie zullen de archieven belangrijker worden in het maatschappelijk leven. De betrokkenheid van het publiek groeit. Digitalisering bevrijdt archieven en hun gebruikers van het probleem van slecht toegankelijke fysieke dragers. In principe kan nu iedereen kennisnemen van de inhoud van de collecties.

Aldus de legio voordelen van digitalisering voor de archieven en hun werkprocessen. De gunstige effecten van deze technologie kunnen echter moeilijk los worden gezien van een groot aantal problemen. Digitalisering genereert nieuwe archivale zorgen rond selectie, ontsluiting en behoud van collecties. Er zijn veel kritische 'digitale' issues: de heterogeniteit van files en formaten, de schaal van hun productie en verspreiding en de wijze waarop digitale

items gewaardeerd moeten worden. Minstens zo kritisch zijn de dynamische levenscyclus van digitale informatie en de duurzaamheid van bestanden. Allesoverheersend is de grote afhankelijkheid van de technologie, die zelf voortdurend in beweging is.

In de nieuwe omstandigheden moeten archieven kunnen blijven fungeren als betrouwbare en duurzame opslagplaats voor audiovisueel erfgoed, als trusted digital repository. Digitale archieven met een conserveringsopdracht moeten hiervoor een nieuwe rol op zich nemen: die van digitaal curator. Dit artikel biedt een visie op de taakgebieden van zo'n digitaal curator. Vragen die aan de orde komen zijn: voor welke specifieke uitdagingen ziet een curator van digitale AV-collecties zich gesteld? Gaat er iets wezenlijk veranderen aan de manier waarop archieven selecteren, bewaren en beschikbaarstellen? Wat zijn de kenmerken van het beheren van een digitale opslagvoorziening en hoe alle oude en nieuwe gebruikers te bedienen? Ook de kennis en kunde die nodig zijn voor het vervullen van een digitaal curatorschap worden behandeld, evenals de noodzaak tot het vormen van nieuwe relaties met de omgeving van het archief: de producenten, de depotgevers en de afnemers van de materialen. Vóór alles wil dit artikel duidelijk maken dat de cruciale rol van de technologie gaat leiden tot nieuwe strategieën en andere methoden en technieken voor een betrouwbare overlevering van audiovisueel erfgoed ofwel: tot een nieuw soort collectiemanagement.

ARCHIEFWAARDIGHEID, ARCHIVEERBAARHEID, BESCHIKBAARHEID

Ook – en hoogstwaarschijnlijk juist – in het digitale tijdperk moeten archieven zorgen voor een 'archiefwwaardige' collectie. Dat wil zeggen dat er grondig moet worden nagedacht over de inhoud en de samenstelling van de materialen die instromen, in relatie tot historische waarde en gebruik. Daarbij is het van groot belang de collectie te allen tijde 'archiveerbaar' te houden, in de zin dat ze adequaat kan worden bewaard, voor de korte of voor de lange termijn. De gearchiveerde items moeten vervolgens op ieder moment te benaderen zijn, beschikbaar voor eenieder die ze zou willen raadplegen of gebruiken.

Uit deze drie verantwoordelijkheden, voor de archiefwaardigheid, de archiveerbaarheid en de beschikbaarheid van de collectie, vloeien de belangrijkste taakgebieden voort van een digitaal curator: selectie en acquisitie, het duurzaam opslaan en het blijvend verschaffen van toegang. Het ene taakgebied kan niet los worden gezien van het andere. Er moet worden geselecteerd wat op te nemen in de archieven. Van de collecties die zo worden opgebouwd



Audiovisuele archieven hebben al heel lang te maken met formatprogressie en migratie. Op de foto de technische ruimte waar BCN naar VHS en digi-beta kon worden overgezet.

FOTO: L. BOUMA

moet weer worden bepaald welke onderdelen na verloop van tijd worden overgezet naar courante formaten, om toegankelijkheid te garanderen. Anders gezegd: na verwerving van de digitale files staan collectievormers voor de opdracht ze technisch zodanig up-to-date te houden dat de gebruikers er blijvend bij kunnen.

FUNDAMENTELE VERANDERINGEN

What's new? zou je haast vragen. Audiovisuele archieven zijn immers van oudsher met deze taken belast. Juist deze archieven waren door de aard van hun collecties altijd al technisch georiënteerd. Beeld en geluid vereisen nu eenmaal apparatuur om ze zichtbaar en hoorbaar te maken. Om gelijke tred te houden met de vernieuwing van deze apparaten moesten ook in het verleden de fysieke tapes en banden bij herhaling worden overgezet. Regelmatige migratie van dragers naar eigentijdse formaten en standaarden is voor AV-archieven dan ook bepaald geen nieuw verschijnsel.

Deze tegenwerping is zowel waar als niet waar. In de kern is de verantwoordelijkheid voor een toegankelijke AV-collectie ontegenzeggelijk gebleven. De instabiele aard van digitale content, de volumes waarin deze wordt geproduceerd én de netwerkomgeving waarbinnen het archief opereert, zorgen echter voor fundamentele veranderingen.

Digitale formaten kunnen namelijk niet worden gezien als de zoveelste nieuwe drager, na de nitraatfilm, de acetaatfilm, de 2-inch en de 1-inch banden, de U-Matic en de Digi Beta. Ook het reproduceren van digitale items is op weinig manieren te vergelijken met het viewen of beluisteren van fragmenten en programma's vanaf een filmtafel, een televisiescherm of een taperecorder. Het beheren en conserveren van digitale AV-objecten is daarbij technisch veel ingewikkelder dan het opbergen van banden en blikken in stellingen en het overschrijven van de inhoud van fysieke dragers naar andere fysieke dragers.

HET DIGITAAL ARCHIEF ALS BLACK BOX?

Anders dan we misschien hebben gehoopt en verwacht is een digitale opslagvoorziening geen black box, waarin een oneindige hoeveelheid content automatisch ordentelijk kan worden opgeslagen en als vanzelf voor de eeuwigheid wordt bewaard. Waar fragmenten, programma's en beschrijvingen naar believen kunnen worden uitgehaald en hergebruikt, door iedereen die dat wil, op elk gewenst moment. Digitale archieven zijn daarbij niet langer 'stand-alone'. Ze maken deel uit van een groter technologisch geheel: de netwerkomgeving. In de complexe keten van creatie, bewerking en beschikbaarstelling van digitale content moeten audiovisuele archieven kunnen fungeren als bron en als schakel, als begin- en als eindpunt. Dit gaat op voor archieven die opereren in een professionele productieomgeving, maar geldt evenzeer wanneer archiefcollecties worden verspreid via de bredere context van het web.

In de beheersing van de digitale archiveringsprocessen is de technologie alomtegenwoordig. ICT dringt door tot in de kleinste onderdelen van de operatie, speelt een essentiële rol in elke actie, bij elke uitlevering en elk transport, hangt aan en tussen alle files en is verbonden aan elk stukje metadata.

De mogelijkheden voor opslag en toegankelijkheid worden met behulp van de techniek sterk vergroot. Archieven zullen zeer grote, veelzijdige collecties kunnen herbergen en die kunnen aanbieden aan gebruikers over de hele wereld. Keerzijde is dat de technologie vele nieuwe risicofactoren met zich meebrengt. Digitalisering maakt van het archief immers een soort IT-systeem, met alle technische problemen die daarbij horen. Risico's zijn bijvoorbeeld het uitvallen van software, opslagmedia en netwerken, of het kwijtraken of beschadigen van data. Ook kunnen archieven snel achterop raken bij de ontwikkeling van de opslag- en reproductietechnologie. Al deze risico's raken het duurzaam behoud van de collectie en dat is precies de reden waarom een digitaal archief in helemaal niets kan lijken op een black box. Juist in een



Een digitaal AV-archief kan beter in helemaal niets lijken op een black box : er is overzicht en transparantie nodig op alle niveaus.

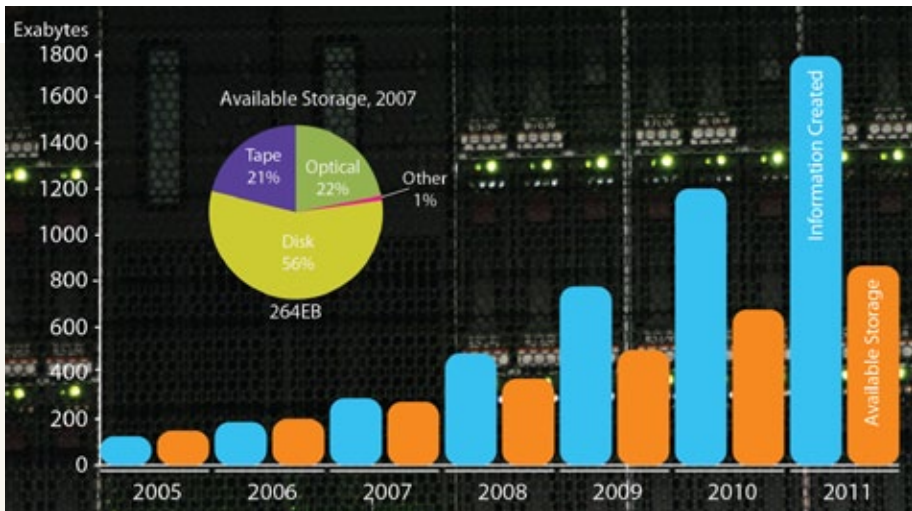
FOTO: A. VAN DER VALK

omgeving die volledig wordt beheerst door de technologie is inzicht, overzicht en transparantie nodig: inzicht op het niveau van de files, de metadata en de media waarop alles is opgeslagen en overzicht van alle bewegingen en bewerkingen van de data in de netwerk omgeving.

SELECTIE EN ACQUISITIE VAN DIGITALE AV-COLLECTIES

De mogelijkheden en het gebruiksgemak van digitale techniek leiden tot een grote productie. Veel daarvan is in principe archiefwaardig, vanuit erfgoeddoogpunt, vanuit hergebruiksmotieven of vanuit een combinatie daarvan. Het is verleidelijk om alles op te nemen; het idee is immers dat het opslaan steeds goedkoper wordt. Feit is natuurlijk dat digitale technieken tegelijkertijd zorgen voor een sterk toenemende hoeveelheid archiefwaardige materialen. Ten eerste worden de bestanden groter. Ten tweede wordt er meer geproduceerd. Deze toename houdt geen gelijke tred met de verwachte verdubbeling van opslagcapaciteit van harde schijven en datatape, iedere 18 maanden. Ze verloopt aanzienlijk sneller.

Alles opnemen is daarom voor archieven met een preservingstaak niet wenselijk en realistisch. Los van wat er aan technologie, kosten en mankracht nodig is om grote digitale hoeveelheden audiovisueel materiaal alleen al op te slaan, betekent meer volume ook: méér gebruik en méér migratie, met



De ontwikkeling van de verhouding tussen de hoeveelheid geproduceerde content en de beschikbare opslagcapaciteit op wereldschaal, een prognose uit 2008.

BRON: EMC / GRAFIEK: A. VAN DER VALK

alle technische complicaties van dien. Juist in het digitale tijdperk zullen audiovisuele archieven moeten selecteren, eenvoudigweg om de volumes en de processen werkbaar te houden en zo de kosten van duurzaam behoud te beteugelen.

ACQUISITIE EN BEHOUD

Selectie is een meervoudig proces. Archieven staan voor de keuze wat op te nemen van het materiaal dat digitaal is ontstaan, de digital borns. Van het oude, analoge materiaal dat reeds onderdeel vormt van de collectie moet vervolgens worden bepaald welke onderdelen gedigitaliseerd moeten worden. In een derde selectieslag moet worden gekozen uit de nieuwe aanwas en de gedigitaliseerde legacy collecties. Welke onderdelen daarvan worden geconserveerd op nieuwe opslagformaten en wat kan eventueel worden afgestoten? Deze derde slag is niet eenmalig. Iedere keer dat de techniek wordt vernieuwd, zal wéér en op meerdere niveaus- moeten worden gekozen. Selectie in digitale archieven is daarom ook een continu proces, dat nauw samenhangt met langetermijnpreserving.

Tijdens het acquireren en selecteren van materiaal doen zich nieuwe complicaties voor die vrijwel allemaal draaien om de noodzaak tot 'onmiddellijk handelen', afgedwongen door de aard van digitale bestanden. Al bij de acquisitie van digitaal materiaal moeten beslissingen worden genomen die (onomeerbare) gevolgen hebben voor later gebruik en preserving.

Bij de inname van traditionele, analoge materialen bestaat een scheiding tussen de beslissing om te verwerven en de keus om te behouden. Het een kan eerst, het ander later. Een doos met filmblikken kan - samen met de ordner met de bijbehorende gegevens - tijdelijk in het archiefdepot worden opgeslagen, alvorens de materialen definitief op te nemen dan wel af te stoten, bijvoorbeeld op het moment dat er beter zicht bestaat op het belang van de items voor de collectie. Bij digitaal materiaal vallen acquisitie en behoud feitelijk samen: het besluit om items blijvend te behouden wordt veel eerder genomen en vraagt om direct antwoord op complexe vragen.

FORMATEN

De actieve beslissing om een item te archiveren op een bepaald formaat moet al op het moment van verwerving worden genomen, dat wil zeggen vlak na uitzending, publicatie of vertoning, en soms zelfs daaraan voorafgaand. Een andere kans komt er namelijk niet meer: een formaat later nog upgraden kan niet. Een van de grootste problemen is dan de veelheid aan dataformaten waaruit moet worden gekozen. Voor een archief is het van groot belang te kunnen garanderen dat gekozen formaten ook op de langere termijn toegankelijk blijven. Lastig is dat op het moment van acquisitie nog onvoldoende kennis bestaat over toekomstige mogelijkheden en beperkingen van formaten. Bekend is eigenlijk alleen dat alle digitale formaten een veel kortere levenscyclus zullen hebben dan de voormalige 'analoge' formaten.

Een andere complicatie schuilt in het begrip toegankelijkheid. Toegankelijk voor wie? En op welk kwaliteitsniveau? In de keus voor het formaat wordt immers vastgelegd door wie en hoe het materiaal in de toekomst kan worden gebruikt. Dit impliceert dat er al op het moment van verwerving een beeld moet bestaan van de mogelijke cultuurhistorische- en hergebruikswaarde van items, en daarmee van de toekomstige gebruikers. Dit beeld moet worden gevormd zonder de reflectie en het noodzakelijke overzicht dat - door een zekere afstand in tijd - doorgaans wel mogelijk was bij het selecteren van analoge materialen.

DE ONMIDDELIJKE WORKFLOW

Digitale files dulden ook geen uitstel waar het gaat om de verwerking. De luxe om materiaal tijdelijk even te laten staan om het op een later moment uit te zoeken, bestaat niet langer: files moeten meteen een workflow in. Dit betekent dat de vorm en de inhoud, maar ook de herkomst en bestemming van iedere afzonderlijke file al bij instroom bekend, gecheckt en gevalideerd moeten zijn. Dit geldt ook voor de relaties met andere digitale bestanden,



In het analoge tijdperk kon geacquireerd materiaal zonder probleem tijdelijk worden weggezet, om later te worden uitgezocht. Ingestroomde digitale files moeten onmiddellijk een gestructureerde workflow in.

FOTO: L. BOUMA

nodig om het programma technisch en inhoudelijk te kunnen reproduceren.. Is deze controle er niet, dan valt iedere garantie op betrouwbaarheid en reproduceerbaarheid weg: files kunnen zoekraken, worden ongecontroleerd gekopieerd of zijn gaandeweg corrupt of onafspeelbaar geraakt. Ook kan de noodzakelijke koppeling met de metadata (als die er tenminste bij instroom al was) in de loop van de tijd teniet gegaan zijn. Cruciaal bij het managen van digitale files is namelijk de strakgedefinieerde, stevige relatie tussen de essence (het digitale beeld en geluid) en de technische en administratieve metadata waarin alle bovengenoemde checks, validaties en kenmerken van het materiaal in zijn vastgelegd.

Materiaal dat niet al bij acquisitie op deze wijze wordt gecontroleerd, is ontoegankelijk en kan niet worden gebruikt. Voor een archief dat functioneert in een productieomgeving, maar ook voor andere digitale collectiehouders, is dit onacceptabel. Gebruikers, of het nu gaat om depotgevers, interne gebruikers of eindgebruikers, verwachten immers direct te kunnen beschikken over het materiaal vanaf het moment dat het in het digitaal archief is gestroomd.

NIEUWE DIMENSIE

Bij selectie en acquisitie was de vraag vooral: is de inhoud interessant? De archiefwaardigheid van materiaal werd met name bepaald vanuit het inhoudelijke belang van audiovisuele producties. Dit inhoudelijk belang is ook nu weer leidend en zal overigens vele nieuwe afbakeningsproblemen opleveren. Oorzaak hiervan is de vervlochten, crossmediale productie waaraan - naast de traditionele professionele producenten- ook talloze consumenten hun content toevoegen. Het begrip 'archiefwaardig' krijgt er in het digitale domein echter een hele nieuwe dimensie bij, die vooral samenhangt met de noodzakelijke voorwaarden voor gebruik en langetermijnpreservering. Om te kunnen voldoen aan deze voorwaarden moet het selectie- en acquisitieproces worden uitgebreid met een reeks ingewikkelde beslissingen op technisch vlak. De afbreukrisico's in deze fase van het archiveringsproces nemen hierdoor sterk toe. Bij de selectie uit zeer grote hoeveelheden digital borns moet al direct voor een formaat worden gekozen, zonder enige distantie in de tijd en zonder kennis van toekomstige toegangsvoorwaarden. Hiermee ontstaat een risico op verkeerde inschattingen. Bij het selecteren voor de migratie van legacy collecties blijft dit probleem bestaan. En hier geldt iedere keer opnieuw: wanneer wordt besloten een item niet te migreren, is het nog maar de vraag of het straks –op een niet meer courant format- nog wel 'bestaat'.

DUURZAAM BEHOUD VAN DIGITALE AV-COLLECTIES

Nieuwe technologieën en formaten voor registratie en reproductie hebben archieven altijd al voor preserveringsproblemen gesteld. De Australische archiefdenker Ray Edmondson formuleerde in zijn *Audiovisual Archiving: Philosophy and Principles*, het zogeheten 'loss principle'. Dit houdt zoveel in als: wanneer we niet alles kunnen bewaren, zal alles niet uit zichzelf bewaard blijven. Edmondson refereerde hiermee aan de voortdurende noodzaak tot het overzetten van audiovisuele dragers naar nieuwe formaten (zoals film naar video, video naar nieuwe videoformaten) om ze afspeelbaar, en dus toegankelijk te houden. Dit 'loss principle' is in het digitale domein geldiger dan ooit.

Digitale bestanden kunnen worden gekopieerd zonder kwaliteitverlies. Dat is een groot voordeel. Maar digitale formaten verouderen snel en hun houdbaarheid is daardoor beperker dan die van analoge formaten. Daardoor is ook de migratiecyclus veel korter: in plaats van iedere 15 tot 20 jaar is de verwachting dat digitale opslagformaten gemiddeld iedere vijf jaar tot zeven moeten worden vernieuwd. Ook opslagmedia, software en hardware moeten



Interieuroopname van een gloednieuwe, nog vrijwel lege taperobot voor de opslag van datatapes. Ook dit apparaat zal binnen afzienbare tijd weer vervangen moeten worden.

FOTO: G. BAKKER

iedere keer weer worden geüpgrade en vernieuwd. De standaardafschrijving van de IT-componenten zoals de databases en de servers waarin alles wordt verwerkt, is zelfs maar hooguit vijf jaar. En hoe lang of hoe kort opslagmedia als harde schijven en datatapes meegaan, is nog vrijwel onbekend. Er zijn inmiddels datatapes op de markt met een vermeende levensduur van 30 jaar. Maar geen enkel archief zal een file op zo'n 'duurzame' datatape willen laten staan als de kans bestaat dat de apparatuur om het bestand af te spelen op een gegeven moment niet meer voorhanden is.

Feit is dat het geheel van deze snel verouderende componenten uit de mainstream IT-technologie zowel de betrouwbare opslag als het duurzaam behoud van AV-materiaal zal moeten kunnen faciliteren. Er zijn immers geen alternatieven. Archieven kunnen vanuit hun conserveringsopdracht nu eenmaal niet kiezen voor experimentele oplossingen, waar garantie op blijvende ondersteuning al te vaak ontbreekt.

EIGENSCHAPPEN AUDIOVISUELE MEDIA

Net zo min als een analogo archief alleen maar tapes en banden opbergt, bestaat het werk van een digitaal archief uit het louter opslaan van files en metadata. Digitale opslag gaat weliswaar over het bewaren van bytes, duurzaam behoud is het kunnen begrijpen (en dus kunnen reproduceren) van de

informatie in de opgeslagen bytes op een later moment. Behoud draait dus om het veiligstellen van de mogelijkheid het materiaal in de toekomst te kunnen gebruiken. De specifieke complicaties rond dit behoud hebben vooral te maken met de eigenschappen van de AV-media zelf. Anders dan bijvoorbeeld tekstbestanden zijn AV-media time-based: ze bestaan uit temporele, sequentiële objecten. Het bewaren en reproducieren van deze media is hierdoor technisch lastiger dan het behouden van statische documenten, zoals bijvoorbeeld PDF's. Voor AV-bestanden moet een speciale opslagtechnologie worden ontwikkeld die het AV-signaal zogezegd 'begrijpt'.

Digitale AV-bestanden zijn vaak simpelweg te groot om ongecomprimeerd op te kunnen slaan. Als bestanden vanuit het productieproces al niet gecompri-meerd worden aangeleverd (zoals bijvoorbeeld het geval kan zijn wanneer op locatie wordt geschoten) wordt voor het opslaan van deze objecten vrijwel altijd gebruikgemaakt van datacompressie en datareductie. Compressie en reductie gebeurt met behulp van speciale software: codecs. Voor duurzame toegang is een correcte reproductie van het oorspronkelijke signaal een absolute eis. Om het beeld adequaat te kunnen weergeven moet gebruik worden gemaakt van dezelfde codec als waarmee het materiaal is gecompri-meerd. Dat kan tot problemen leiden, omdat ook codecs snel verouderen. Bestanden die nog met oude codecs zijn geëncodeerd kunnen dan niet meer gedeco-deerd worden, tenzij deze software backwards compatible is. Zaak is dus ook de codecs te conserveren. Het helpt dan als een codec is gebaseerd op open standaarden. Daarnaast is het belangrijk dat oude codecs blijvend worden ondersteund in de afspeelapparatuur.

Het time-based karakter van AV is ook van invloed op de metadata die de samenstellende onderdelen (shots, fragmenten) van dit materiaal technisch en inhoudelijk identificeren. Er zullen mechanismen moeten worden geïm-plementeerd om deze metadata op de juiste manier, op de juiste plaats in de file, te blijven linken. Toegang tot de afzonderlijke fragmenten in een bestand wordt anders erg problematisch.

DYNAMISCHE OMGEVING

Toegang tot AV-collecties stelt sowieso ingewikkelde voorwaarden aan de inrichting van de opslagvoorziening van materiaal. De (productie)omgeving waarin AV-collecties worden gebruikt, is vaak zeer dynamisch. Files en delen van files worden voor allerlei doeleinden constant geraadpleegd, gekopi-eerd, hergebruikt, gemuteerd en opnieuw gepubliceerd. Het materiaal moet zodanig opgeslagen kunnen worden dat de master veilig is gesteld en tegelijk voor allerlei doeleinden getranscodeerd kan worden. Vaak is het wenselijk

om meerdere versies of varianten van dezelfde file te bewaren, bijvoorbeeld als high res kopie, als browsefile én als serie keyframes. Daarnaast moeten voorzieningen worden getroffen om op verzoek bestanden te kunnen transcoderen naar specifieke formaten, bijvoorbeeld voor publicatie op internet. Tegelijkertijd zal materiaal dat minder vaak wordt gebruikt nearline of diep moeten kunnen worden opgeslagen, op een ander type storage. Dit is nodig opdat het archief niet de (zeer hoge) kosten draagt van het voortdurend online beschikbaar houden van minder courant materiaal. Het storagestelsel moet dan ook nog geschikt zijn om grote hoeveelheden materiaal regelmatig te kunnen migreren naar een andere (vernieuwde) storageomgeving zónder dat het stelsel zodanig zwaar wordt belast dat andere processen hinder ondervinden. Door dit alles lijkt een digitale 'audiovisuele' opslagvoorziening in heel weinig op een statisch depot. Het is een stelsel waarin de files permanent in beweging zijn.

In het digitale AV-domein gaat het bovendien niet alleen om het bewaren en toegankelijk houden van enkelvoudige files, net zo min als een tape of een film het enige was wat er in het analoge tijdperk van een audiovisueel product behouden bleef. Behalve de video-essence is er ook nog de bijbehorende audio-essence. Bij de essence horen de gegevens die deze technisch identificeren. Essence en technische metadata moeten gekoppeld blijven en worden samengebracht in een zogeheten wrapper of datacontainer. Ook zijn er ondertitels, logo's, contracten, productiemetadata, rechtengegevens en dergelijke in digitale vorm, die deel uitmaken van een bepaalde productie. Dit soort gegevens wordt veelal in een database opgeslagen of apart bewaard als ongestructureerde data. Vaak echter worden ze – samen met de essence – weer in een ander type datacontainer gestopt. Het is een gecompliceerde zaak om al deze verschillende wrappers duurzaam te archiveren. Wrapperformaten zijn namelijk zelf ook permanent onderhevig aan verandering en vernieuwing. Bovendien worden sommige typen in verschillende (productie) omgevingen ook nog eens anders geïnterpreteerd.

AUTHENTICITEIT EN INTEGRITEIT

Speciale aandacht in een archivale context verdient de zekerstelling van de authenticiteit en de integriteit van de objecten. Kenmerk van een digitale archiefomgeving zijn, zoals uiteengezet, de snelle, opeenvolgende technologische veranderingen, in combinatie met voortdurende verplaatsing en migratie van (versies van) objecten. Als niet gegarandeerd kan worden dat digitale items ondanks deze omstandigheden hun authenticiteit en integriteit hebben behouden, raakt materiaal zijn erfgoedwaarde en zijn zeggingskracht kwijt en



LANGS_DE_LIJN-AEN55718212.wav	Archive new content [DDV]	2011-06-23 16:52:41
LANGS_DE_LIJN-AEN55718211.wav	Archive new content [DDV]	2011-06-23 16:52:41
VONNISTRIBUNA-HRE0005AA7E.mxf	Archive new content [DDV]	2011-06-23 16:52:41
LANGS_DE_LIJN-AEN55718216.md5	Archive new content [DDV]	2011-06-23 16:52:41
LANGS_DE_LIJN-AEN55718215.md5	Archive new content [DDV]	2011-06-23 16:52:41
LANGS_DE_LIJN-AEN55718214.md5	Archive new content [DDV]	2011-06-23 16:52:41
LANGS_DE_LIJN-AEN55718213.md5	Archive new content [DDV]	2011-06-23 16:52:41
LANGS_DE_LIJN-AEN55718212.md5	Archive new content [DDV]	2011-06-23 16:52:41
PRGTJJDSEIN2_-HRE0005AEDF.mxf	Archive new content [DDV]	2011-06-23 16:48:37
MODELNAOMICAM-HRE0005AECF.md5	Archive new content [DDV]	2011-06-23 16:48:37
MODELNAOMICAM-HRE0005AECF.mxf	Archive new content [DDV]	2011-06-23 16:48:37
PRGTJJDSEIN2_-HRE0005AEDF.md5	Archive new content [DDV]	2011-06-23 16:48:37
LOUJOURNAAL_-HRE0005ACEA.mxf	Archive new content [DDV]	2011-06-23 16:48:37

Het monitorscherm toont WAV en MXF files die in realtime het digitale archief binnenstromen. Te zien is dat er voor iedere file ook een checksum wordt meegeleverd om de integriteit te kunnen waarborgen: een zgn. MD5.

FOTO: A. VAN DER VALK

verliest het archief zijn autoriteit. Er is dan al helemaal geen sprake van een trusted digital repository.

Integriteit wordt in een digitale setting vertaald naar de specifieke samenstelling op het niveau van de bits van een object: de bitconfiguratie. Een digitaal archief moet voldoen aan de eis van de persistentie. Dat houdt in dat het materiaal er exact zo uit moet komen als het erin is gegaan, dat wil zeggen compleet en in de juiste datastructuur. Deze data-integriteit kan bijvoorbeeld worden vastgesteld door middel van checksums: het vergelijken van de file tegen de originele bitconfiguratie, zoals die is vastgelegd bij de instroom van het materiaal. Bij elke migratie van een versie van het object voor opslag, preservering of uitlevering aan gebruikers, kan met een fixity check worden teruggegrepen op de checksum, zodat zeker is het object 'qua bits' identiek is aan de content die wordt overgebracht.

Een digitaal object is weliswaar discreet in de zin dat het bestaat uit een aanwijsbare combinatie van bits en bytes, de specifieke bitconfiguratie kan echter niet dienen om het ene van het andere object te onderscheiden, oftewel om objecten te identificeren. Dit onderscheid is noodzakelijk om door alle veranderde omstandigheden heen te kunnen garanderen dat een object 'authentiek' is gebleven. Authenticiteit wordt hier opgevat in zijn archivale betekenis, wat inhoudt dat een object werkelijk is wat het voorgeeft te zijn.

Dat kan zekergesteld worden door het vastleggen van gegevens over dat object in de context waarbinnen het is gecreëerd en wordt gebruikt. De mate van authenticiteit in een archiveringsproces hangt dan samen met een gedocumenteerde en vaardige beoordeling van wat de content van een digitaal object precies is.

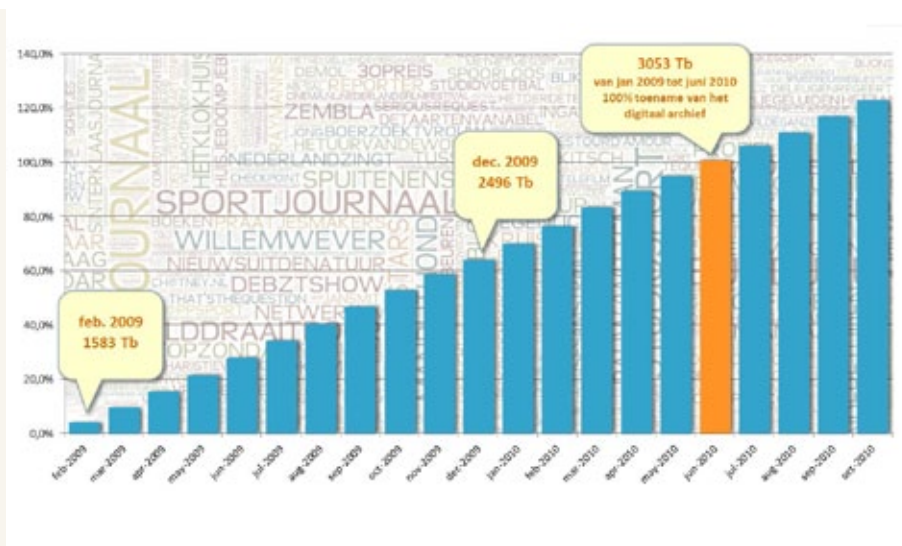
SIGNIFICANT PROPERTIES

Hiervoor moeten de kenmerken die een digitaal object identificeren en betekenis verlenen, de significant properties, worden gedefinieerd. Significant properties zijn te zien als die eigenschappen van een digitaal object die duurzaam overgeleverd moeten worden, zodat het object ook op de lange termijn toegankelijk, bruikbaar en begrijpelijk blijft. Deze kenmerken kunnen de herkomst en de geschiedenis van een object betreffen, maar ook de (technische) omgeving waarin een object werd gemaakt, zoals de gebruikte software en hardware. De relatie met objecten in de context van het betreffende object wordt erin vastgelegd, en ze beschrijven de technische eigenschappen en de structuur van de essence. Er wordt ook in gedocumenteerd wie het object heeft gebruikt en mag gebruiken, en voor welke handeling. Samen vormen deze gegevens het bewijs van de 'geloofwaardigheid' van een object zoals dat bewaard is gebleven door de tijd. Preserveringsacties, nodig om digitale objecten op een later tijdstip zowel technisch te kunnen reproduceren als inhoudelijk en contextueel te kunnen 'plaatsen', kunnen niet worden uitgevoerd als dit soort eigenschappen niet gedetailleerd in de systemen wordt vastgelegd

Vastlegging van fixity en van significant properties als herkomst, (technische) context en gebruiksrechten gebeurt al bij de ingest van materiaal. Dit soort gegevens wordt meegeleverd vanuit de productieomgeving en in alle volgende stadia van de levensloop van een digitaal object bijgehouden en aangevuld door het archief, zowel automatisch als handmatig. Het is duidelijk is dat het beschermen van authenticiteit en integriteit in een dynamische audiovisuele archiefsetting veel nieuwe documentatieprocedures met zich mee zal brengen. Ook vereist het andere typen metadata, die de voortdurende migratie en verandering van objecten kunnen beschrijven en ondersteunen.

VOLUMES

Het tijdig migreren van formaten, datacontainers, opslagmedia en software, het gedifferentieerd opslaan van files en metadata, het vastleggen van hun significant properties en het managen van de levenscyclus van objecten: al deze activiteiten zijn onontbeerlijk om audiovisuele collecties duurzaam te



In de beheersing van de snel toenemende volumes schuilt een van de grootste uitdagingen voor digitale AV-archieven. De digitale collectie van Beeld en Geluid verdubbelde in niet meer dan anderhalf jaar tijd.

GRAFIEK: A. VAN DER VALK

kunnen behouden en beheren. Het timebased karakter van de AV-media en de metadata, alsmede de complexe samenstelling van de audiovisuele wrappers maakt deze processen nog eens extra ingewikkeld.

Maar de echte uitdaging schuilt zonder enige twijfel in het feit dat de processen moeten worden uitgevoerd op onwaarschijnlijke hoeveelheden materiaal. De volumes die omgaan in het audiovisuele domein zijn onvoorstelbaar groot. Middelgrote tot grote audiovisuele archieven herbergen uiteindelijk petabytes tot -in de toekomst- zelfs exabytes aan materiaal, dat permanent moet worden gemanaged in een dynamisch proces van inname, opslag, preservering en beschikbaarstelling. Al dit materiaal moet een veelheid aan gecontroleerde processen ondergaan: voor het valideren en verifiëren van de wrappers en hun inhoud bij instroom en uitstroom, voor het extraheren en mappen van de 'significant properties' van de files naar de informatiesystemen en voor het beheersen van de lifecycle van alle items. Om de data proactief te kunnen beschermen en te behouden is het nodig om alle opgeslagen items permanent gedetailleerd te monitoren. De opslagsoftware zal daarbij het materiaal moeten kunnen bewaren als een collectie nauwkeurig gedefinieerde 'objecten', elk bestaande uit een set files. Al deze objecten zullen permanent moeten worden gescreend op hun duurzaamheidsrisico, zodat de noodzakelijke preservingsacties kunnen worden toegepast op hele groepen materiaal. Omdat het niet mogelijk is om de bewegingen van tienduizenden

files per dag handmatig te managen, zal een groot deel van deze processen geautomatiseerd moeten verlopen.

AFBREUKRISICO

In zo'n technologische omgeving liggen veel single points of failure op de loer: risicovolle en bedrijfskritische processen zijn niet langer gespreid maar geconcentreerd op een plek. Het afbreukrisico is hierdoor zeer groot: een complete collectie is met één druk op de knop te verwijderen, één verkeerde naamgeving en een hele groep files is niet meer te identificeren, één lelijke scheur in een LTO5-tape en 50 programma-uur zijn voorgoed verloren. Even niet optellen kan, kortom, leiden tot dataverlies van catastrofale omvang. Om dit soort risico's van digitale massaopslag terug te dringen tot een acceptabel niveau en ook om eventuele uitval van systemen en netwerken op te vangen, wordt vaak gebruikt gemaakt van redundancy. Back-upbestanden en een back-uparchief moeten echter ook permanent worden gemonitord en gemanaged. Wanneer archieven petabytes of zelfs exabytes gaan beheren, wordt deze oplossing zeer kostbaar.

Feit is dat digitaal beheer van zeer grote volumes AV-materiaal eigenlijk nog nergens ter wereld gemeengoed is. Het is op dit moment dan ook niet mogelijk om precies aan te geven aan welke technologische requirements een dergelijke grote en complexe infrastructuur precies moet voldoen om een zo groot mogelijke mate van veiligheid, betrouwbaarheid en duurzaamheid te garanderen. Research en innovatie op alle onderdelen van managed storage en digitale preservatie zijn dan ook aan de orde van de dag. Dit plaatst grote audiovisuele archieven momenteel in de voorhoede van de technische ontwikkelingen, met alle trial & error risico's van dien. De vraag is of dit een natuurlijke positie is voor deze erfgoedinstellingen, die uiteraard van meet af aan moeten kunnen fungeren als trusted digital repository.

GEBRUIK EN GEBRUIKERS VAN DIGITALE AV-COLLECTIES

Het wezen van een digitaal archief bestaat niet uit het kunnen bewaren van files. Archieven moeten vooral in staat zijn om hun gebruikers blijvend toegang te verlenen. Opslag en toegankelijkheid zijn in het digitale tijdperk echter niet te scheiden. Een achterhaald opslagformaat is niet reproduceerbaar en dus niet te benaderen door gebruikers. Naast de eis van de persistentie moet een digitaal archief dus ook voldoen aan de eis van de courante formaten, die zoveel inhoudt als: opgeslagen items zijn daadwerkelijk af te spelen door de gebruikers van het archief. Dit betekent dat het niet zo kan

zijn dat er obsoleete (technisch achterhaalde) formaten worden aangeboden, die niet meer af te spelen zijn. Voldoet een archief blijvend aan de eis van courantheid, dan is er sprake van duurzame opslag. Digitale duurzaamheid van archieven en collecties wordt daarom ook wel omschreven als duurzame toegankelijkheid.

Met de digitalisering neemt het online gebruik van audiovisuele collecties sterk toe. Dit geldt ook voor het type gebruik en de soorten gebruikers. Naast een groeiend aantal 'traditionele' gebruikers van av-archieven (professionals uit de productiewereld, studenten, onderzoekers) komen er steeds meer nieuwe online gebruikers bij, zoals het algemeen publiek, digitale televisiekijkers en speciale groepen zoals de educatieve sector. Al deze oude en nieuwe gebruikers willen er op kunnen vertrouwen dat de items die zij zoeken aanwezig zijn en kunnen worden afgespeeld. Het creëren van toegang aan deze groepen begint daarom al bij de instroom van materiaal. Files moeten al in deze fase worden gecheckt op de mogelijkheid ze te kunnen uitleveren. Afgezien van mogelijke belemmeringen door auteursrechtelijke restricties, is er namelijk de variëteit aan de technologische voorwaarden voor toegang, die per gebruikersgroep kunnen verschillen. De standaard voor opslagformaten moet zó worden gekozen dat het materiaal naar meerdere gebruiksnormen kan worden getranscodeerd.

EISEN EN VERWACHTINGEN

Het materiaal moet bovendien te allen tijde kunnen worden ingezet voor het doel dat de gebruiker voor ogen heeft, of het nu gaat om professioneel hergebruik, raadpleging, informatievoorziening of entertainment. Om aan de uiteenlopende gebruikersverwachtingen tegemoet te kunnen komen, worden er - naast de technische toegangsvoorwaarden - complexe eisen gesteld aan de inhoud, de structuur, de kwaliteit en de presentatie van de collectie. Er spelen hier tal van wensen en belangen. Zo zal een programmamaker uit de omroepwereld het belangrijk vinden dat actueel materiaal na uitzending meteen in een hoge kwaliteit beschikbaar is. Zo'n gebruiker wenst de collectie op fragmentniveau te benaderen en wil kunnen zoeken in gestructureerde, time-aligned metadata, browsekopieën en keyframes. Het grote publiek, de 'Google-generatie', zal liever intuïtiever zoeken, bijvoorbeeld via een tagcloud. Deze gebruikers verwachten in ieder geval op een eenvoudige manier door de beelden heen te kunnen browsen en snel door te kunnen klikken naar interessante contextinformatie in de netwerkomgeving. Ze zullen er verder voetstoots van uitgaan dat zij eigen commentaar en tags toe kunnen voegen. Specialistische doelgroepen, bijvoorbeeld vanuit het onderwijs, wensen de collectie op hun beurt te kunnen raadplegen binnen



Weergave van het zenuwcentrum van de technische infrastructuur : de zgn. serverfarm. Alle bewegingen van de content worden hier vanuit gestuurd. Links de racks met de verwerkingservers, rechts alle servers die de opslag regelen.

eigen, gesloten netwerken, waarbij beeld en geluidsfragmenten ook nog eens moeten kunnen worden voorzien van eigen 'educatieve' metadata.

PARTIAL RETRIEVAL

Met het verlenen van enkelvoudige toegang heeft dit soort digitale dienstverlening nog weinig te maken. Het beschikbaarstellen van AV bestaat dan ook niet uit het simpelweg versturen van een file naar gebruikers en is in die zin in niets te vergelijken met het versturen van een PDF of een fotobestand. Belangrijkste kenmerk hier is het faciliteren van partial retrieval van metadata en essence. Hierdoor wordt – afhankelijk van de gebruikersvraag- bij iedere uitleveringsactie gericht ingegrepen in de content, met alle technische en beheersmatige complicaties van dien.

De toenemende gebruikersverwachting is bovendien, dat de collectie-items en hun beschrijvingen gecontextualiseerd kunnen worden aangeboden. Dat wil zeggen dat de eigen catalogusgegevens van het archief kunnen worden gekoppeld aan additionele informatie op het web of in andere, externe bestanden, met behulp van semantische technieken. De te distribueren digitale 'pakketten'- bestaande uit wisselende samenstellingen van beeld, geluid en technische en beschrijvende metadata uit in- en externe bronnen - moeten vervolgens kunnen worden aangeboden op totaal verschillende strea-

ming platforms, zoals intranet- en extranetsystemen, websites, editsuites, mobiele telefoons en digitale themakanalen, alle met eigen, toegesneden toegangsautorisaties.

In technische zin betekent beschikbaarstelling van digitaal audiovisueel materiaal het gelijktijdig kunnen bedienen van grote gedifferentieerde gebruikersgroepen met specifieke wensen wat betreft de beeldkwaliteit, de granulariteit, browse- en navigatievoorzieningen, metadata en interoperabiliteit. Het intensieve dataverkeer en het veelvoud aan interfaces en uitwisselingsprotocollen dat deze distributie met zich meebrengt, vormen een enorme belasting voor de technische infrastructuur, zeker als wordt gewerkt binnen dynamische (productie)omgevingen. De leidingen van en naar een digitaal archief staan in zo'n omgeving roodgloeiend van alle items, fragmenten, complete programma's, metadata en contextinformatie, die als browseskopie, tekstbestand of high res file permanent worden verzonden naar een veelvoud aan bestemmingen.

WIE ZIJN DE GEBRUIKERS?

Zoveel producten, diensten en publicaties, op zoveel verschillende niveaus kunnen uitleveren aan zoveel soorten gebruikers, vereist het maximum aan capaciteit en flexibiliteit van de technische infrastructuur. Om de technologische, logistieke, organisatorische en financiële gevolgen van online beschikbaarstelling in bedwang te houden, moet een duidelijk antwoord zijn geformuleerd op één belangrijke vraag: wie zijn de huidige en toekomstige gebruikers en wat verwachten zij van het archief? Gedetailleerd inzicht in gebruik en gebruikers kan archieven helpen doelgericht en daardoor kosten-efficiënt te acquireren, te conserveren en aan te bieden. Gebruikers van een digitaal archief kunnen geen vage entiteit meer zijn zoals 'het nageslacht', 'het grote publiek' of 'de professionele gebruiker'. Het archief kan ook niet langer werken vanuit eigen aannames omtrent de beste samenstelling van de collectie en de beste manieren om deze beschikbaar te stellen. Voor just-in-case scenario's, waarbij alles duurzaam wordt bewaard en op alle denkbare wijzen beschikbaar wordt gesteld, zijn de volumes eenvoudigweg te groot.

Er zal een veel directere relatie moeten worden gelegd tussen de archiefpraktijk en het gedrag en de verlangens van degenen die het archief gebruiken. Deze groepen zullen heel precies moeten worden gedefiniëerd waar het gaat om hun wensen wat betreft de inhoud van het aanbod, de technische faciliteiten en de gebruiksmogelijkheden. Hoe wil de programmamaker uit de omroepwereld het materiaal exact voorgeschoteld krijgen? Op welke norm, met welke begeleidende metadata? Geldt dit voor alle programmamakers? Uit



Via een webcam die alle bewegingen van de taperobot op het scherm laat zien, kan het dynamische gebruik van items uit de digitale AV-collectie 'live' worden gevolgd.

FOTO: A. VAN DER VALK

welke doelgroepen bestaat het online publiek en hoe willen zij de collectie precies kunnen benaderen? Wat zijn de inhoudelijke en technische wensen van domeingebonden gebruikers zoals het onderwijs, de wetenschap en de diverse commerciële klantgroepen?

Alle requirements voor collectie-aanbod, metadata, distributie en presentatie, van alle gebruikersgroepen tezamen, vormen het fundament waarop het archief zijn collectie, zijn infrastructuur en zijn diensten inricht, binnen de aanwezige financiële kaders. Deze requirements zijn niet statisch. Wensen, zoekvragen en technische omstandigheden kunnen door de tijd veranderen. Gebruikersgroepen moeten daarom blijvend worden gevolgd, zodat het archiefbeleid permanent wordt getoetst en kan worden verantwoord. Het vastleggen en analyseren van kennis over gebruikers en gebruikspatronen moet hiertoe structureel worden ingebed in het archiveringsproces.

DIGITAAL CURATOR IN HET HIER EN NU

De zorg voor de archiefwaardigheid, de archiveerbaarheid en de beschikbaarheid van materialen behoort tot de kerntaken van het audiovisuele archief in zijn rol als digitaal curator. Door de tijd en door alle technologische vernieuwingen heen zal de authenticiteit en de integriteit van ieder item moeten worden bewaakt en beschermd. Feit is dat op de ontwikkelingen

bijna niet kan worden geanticipeerd. Bestaande AV-archieven staan namelijk voor de opgave de transformatie hier en nu in te gaan. Voor theoretische reflectie en een gedegen voorbereiding op de nieuwe rol is eigenlijk geen tijd. Hun collecties staan nu op het netwerk, en ze hebben nu te maken met nieuw, veeleisend gedrag van hun gebruikers. Verse digitale aanwas stroomt dagelijks het archief binnen, in grote hoeveelheden en in vele vormen en versies. In de kasten en stellingen staan grote groepen legacy materialen te dringen om ook te worden gedigitaliseerd. Archieven worden daarbij ook nog eens geacht al op dit moment te kunnen fungeren als trusted digital repository, in een situatie waarin nog maar weinig informatie voorhanden is over de levensduur van formaten, software en opslagmedia.

BEHEERSING EN CONTROLE

Toch lijkt de grootste bedreiging voor digitale AV-archieven niet te liggen in de snel verschuivende standaarden en formaten, ook al zullen collectiebeheerders hier op een bewuste, gecontroleerde manier mee om moeten leren gaan. Een veel groter risico bij het beheren van (middel)grote digitale AV-archieven - zeker als ze functioneren in een dynamische productieomgeving - vormt het zoekraken of corrupt worden van items en zelfs hele collecties. Dit gevaar is er niet alleen nu, in de overgangssituatie van analoog naar digitaal, maar zal altijd bestaan wanneer grote hoeveelheden data en metadata zó vaak en zó veel worden bewerkt en verplaatst. Een dwingend gevolg van de digitale technologie is dan ook de noodzaak tot een totale beheersing van het archiveringsproces. Het managen van een digitaal archief zal gelijk moeten staan aan het voortdurend en zeer precies weten en volgen wat er binnenkomt, hoe het is opgeslagen, van wie het is, waar het vandaan komt, waar het naartoe moet en wie het heeft gebruikt. Regie en controle op alle bewegingen van de data op dit detailniveau is van levensbelang voor een rationele, kosteneffectieve uitvoering van de processen.

PROCEDURES

In de praktijk betekent digitaal collectiemanagement vooral het werken volgens procedures, die nauwgezet moeten worden gedocumenteerd. Op basis daarvan worden de vele checks en controles uitgevoerd die garant moeten staan voor het behoud van de authenticiteit en integriteit van het materiaal door het hele proces van inname, opslag, migratie en beschikbaarstelling. Metadata spelen hierbij een cruciale rol. Alle technische, administratieve, beschrijvende en preservation metadata moeten precies worden gedefinieerd en blijvend worden gekoppeld aan de essence.



Het 'total cost of ownership': een weergave van alle procesonderdelen van digitale lange termijn archivering die tijd en geld kosten en die dus moeten worden gepland en begroot.

BRON: PRESTO PRIME / VORMGEVING: A. VAN DER VALK

Bij het beheersen van de processen hoort ook het beheersen van de kosten. Archieven zullen greep moeten krijgen op de kosten van digitale langetermijnarchivering op een 'audiovisuele' schaal. Hiervoor zal een grondig inzicht moeten worden opgedaan in het zogeheten TOC ('total costs of ownership') van een digitaal archief. Dit TOC bestaat niet alleen uit kosten voor de opslag (ook al zijn die in het AV-domein hoog), maar moet worden berekend voor het behoud van een digitaal object gedurende zijn hele levenscyclus, inclusief de kosten voor het verlenen van toegang. Het TOC is dan samengesteld uit kosten voor de acquisitie en selectie, de ingest, de extractie en toevoeging van metadata, de opslag, de preservatie en de streaming aan gebruikers. Al deze onderdelen moeten structureel worden begroot. Daarbij zullen digitale archieven moeten leren om trade offs te maken tussen kwaliteit (normen) en kwantiteit (volumes) met als doel blijvend in staat te zijn om grote hoeveelheden AV-materiaal duurzaam te bewaren en aan te bieden op een manier die voor alle belanghebbenden acceptabel is.

GEDEELDE VERANTWOORDELIJKHEID

Het maken van de afwegingen tussen het 'hoe' en het 'hoeveel' moet een gezamenlijke verantwoordelijkheid zijn van het archief en van de producenten of andere depotgevers en de afnemers van de collectie: de gebruikers. De kostbare opdracht tot het duurzaam overleveren van audiovisuele materialen kan zo gemeenschappelijk en meer kosteneffectief worden gemaakt. Het archief zal actief met de partijen moeten onderhandelen over de inrichting van de archivale processen, de keuze voor bepaalde standaarden en formaten

en de voorwaarden voor beschikbaarstelling. Met producenten moet worden afgesproken hoe en op welke norm het materiaal gepreserveerd gaat worden. Met gebruikers moet formele overeenstemming worden bereikt over de gewenste manier van distribueren.

In de praktijk betekent dit dat de conserveringsopdracht van het archief bestuurlijk moet worden vastgelegd, evenals de mensen en de middelen waarmee deze opdracht wordt vervuld. Ook zullen de procedures voor inname, opslag en beschikbaarstelling in overleg met de belanghebbenden toetsbaar en transparant moeten worden gemaakt. Om misverstanden te voorkomen zullen de overeenkomsten tussen de partijen moeten worden gebaseerd op een gemeenschappelijke terminologie voor de gebruikte systemen, de procedures, de files en de metadata en de uitwisselingspakketten. Belangrijk onderdeel vormt ook het vastleggen van het recht van het archief om de gearchiveerde essence op de lange termijn te bewaren. Het archief moet bovendien formeel garanderen dat het op elk moment toegang kan verlenen tot 'authentieke' kopieën van de gearchiveerde files in originele vorm. Daarvoor zal het nodig zijn om de inhoudelijke en technische wensen van de gebruikersgroepen blijvend te monitoren, zodat zeker wordt gesteld dat de opgeslagen files ook steeds daadwerkelijk door hen kunnen worden benaderd.

Gedeelde verantwoordelijkheid maakt het archief wellicht in sommige gevallen minder 'autonoom' dan voorheen, zeker waar het gaat om het maken van bepaalde conserveringskeuzen en het vaststellen van het niveau van beschikbaarstelling. Groot voordeel is dat de grote uitgaven en investeringen van langetermijnpreserving – door de keuzen in lijn te brengen met de wensen van producenten en gebruikers – doelmatig zijn en rationeel kunnen worden verantwoord.

KENNISONTWIKKELING

Audiovisuele archieven zijn door hun traditionele afhankelijkheid van de voortschrijdende techniek van oudsher vertrouwd met het vastleggen van technische en administratieve gegevens over dragers, mutaties en opslag- en conserveringshandelingen. Een deel van deze kennis kan worden ingebed in de nieuwe digitale situatie. De meeste regels en procedures behorend bij digitaal collectiemanagement, kunnen echter niet zomaar uit de analoge conventies worden afgeleid. Daarvoor heeft de bescherming van authenticiteit in een digitaal domein teveel technische en technologische kanten. Door het hoge ICT-aandeel in de opslag- en conserveringsprocessen zullen archieven vooral in aanraking komen met kennisgebieden waarmee ze nog niet vertrouwd zijn. Er ontstaan ook nieuwe afhankelijkheden van bijvoor-



LTO tapes die klaar staan om te worden vervoerd naar een backup- locatie. Deze bescheiden stapel bevat ruim 13.000 uur aan audiovisueel erfgoed.

FOTO: A. VAN DER VALK

beeld broadcast-IT, serviceproviders en hardware- en softwareleveranciers. Met deze werelden – die vaak nog weinig oog hebben voor de specifieke noden en wensen van archieven – moet intensief worden samengewerkt en gecommuniceerd.

Dit betekent allemaal niet dat het ontwikkelen van ingestprocedures, pre-serveringsworkflows, preservation metadata en lifecycle management aan IT-consultants moet worden overgelaten. Archieven zullen – vanuit hun eigen verantwoordelijkheid- hun digitale collecties zelf onder controle moeten leren krijgen en dus ook intern de benodigde kennis opbouwen. Dat is overigens niet hetzelfde als het ook eigenhandig uitvoeren van alle processen: (delen van) de opslag, de migratie en beschikbaarstelling kunnen bijvoorbeeld worden uitbesteed aan commerciële partijen. Zo kan een archief besluiten de hardware zelf in huis te hebben, maar bepaalde distributie- of opslagoperaties te outsourcen bij gespecialiseerde bedrijven. Maar om dit soort keuzes überhaupt te kunnen maken en te kunnen overzien wat de consequenties zijn voor de eigen preserveringsopdracht, moet grondige kennis aanwezig zijn van de digitale materialen en alle processen daaromheen. Alleen dan kan worden bepaald welke taken aan anderen kunnen worden overgelaten en hoe ze uitgevoerd dienen te worden. Om wat voor service level agreement (SLA) dan ook te kunnen afsluiten met een serviceprovider, een datastore of een andere externe partij, zal ieder detail van die service gedefinieerd moeten kunnen worden. Dat is niet mogelijk zonder de onderlig-

gende technologie volledig te doorgronden. Deze fundamentele kennis moeten archieven zich eigen maken, hetzij om het digitaal beheer zelf te kunnen uitvoeren, hetzij om gespecialiseerd IT-personeel – van binnen of buiten de organisatie- te kunnen managen en aan te kunnen sturen. Door structureel te investeren in eigen kennis en competenties kan ervoor worden gezorgd dat de digitale archiefvoorziening geen black box wordt, ook niet wanneer die box elders staat en door anderen wordt bestuurd.

SAMENWERKING

Voor veel kleine archieven is het lastig zelf te voorzien in gespecialiseerd personeel, in de ontwikkeling van in-house skills en in de faciliteiten voor digitale preservering. Kennis, expertise en best practices hoeven echter niet in isolement te worden ontwikkeld. Archieven en andere collectiehouders kunnen hier goed samenwerken en zoeken naar modellen die verder reiken dan het eigen instituut. Samenwerking kan niet alleen leiden tot efficiency en schaalvoordelen, maar ook tegemoetkomen aan de grote gemeenschappelijke kennisbehoefte. Het ligt voor de hand aan te sluiten bij initiatieven op nationaal niveau, zoals de Nationale Coalitie Digitale Duurzaamheid, waar alle grote collectiehouders in Nederland deel van uitmaken. Kennis en best (en worst) practices kunnen ook met met elkaar worden gedeeld binnen AVA_Net, de netwerkorganisatie van AV-archieven. In een Europese context werkt het Presto Prime researchproject aan gemeenschappelijke oplossingen voor duurzame opslag en digitaal AV-collectiemanagement. AVAN (Audiovisual Archive Network) en het National Digital Information Infrastructure and Preservation Program (NDIIPP) in de VS zijn als netwerkorganisaties ook bezig met een grootschalige, gezamenlijke aanpak van het beheer van grote volumes AV. Binnen al deze verbanden wordt onderzocht hoe services voor digitale opslag van AV-collecties eruit moeten zien, wat ze mogen kosten en hoe ze zoveel mogelijk centraal en generiek gemaakt kunnen worden. Alles is erop gericht om te leren hoe de zaken beheersbaar en betaalbaar te houden, opdat niet iedere audiovisuele collectiehouder voor zich -via de kostbare weg van trial & error - is gedwongen uit te zoeken hoe een digitale AV-collectie moet worden gemanaged.

TENSLOTTE

AV-archief en technologische vernieuwing zijn per definitie onscheidbaar. Houders van audiovisuele collecties kennen de regels en de voorschriften die een ordentelijke overgang naar nieuwe technische omgevingen begeleiden. In zekere zin bestaat het duurzaam bewaren van digitale collecties uit het

opnieuw benoemen van deze aloude conventies en het explicieter en bewuster toepassen ervan.

Met deze omschrijving wordt het nieuwe collectiemanagement echter bij lange na niet gekarakteriseerd. De technische veranderingen van formaten, opslagmedia en software dienen zich immers niet alleen veel sneller, maar ook veel dwingender aan dan voorheen. De bewegingen en bewerkingen van data en metadata zijn vele malen dynamischer dan vóór het digitale online tijdperk. Maar wat alles fundamenteel zal veranderen is de schaalgrootte van de collecties en processen.

De technologie die te pas komt bij het duurzaam bewaren van grote, dynamische AV-collecties kan nauwelijks nog worden beschouwd als slechts een middel dat instellingen helpt beeld en geluid op te slaan, te ontsluiten en aan te bieden. In feite wordt digitale technologie allesbepalend, voor de vorm en voor de inhoud van het archief. Voor wat er kan worden bewaard en voor hoe het kan worden bewaard. Voor de manier dus waarop met huidige en toekomstige gebruikers wordt gecommuniceerd. De taak van een digitaal curator is dan om heer en meester te blijven over dit proces en daarmee garant te staan voor een bewuste, gecontroleerde overlevering van ons audiovisueel erfgoed.

GERAADPLEEGDE LITERATUUR

ADDIS, MATTHEW, WRIGHT, RICHARD, 2010

Audiovisual preservation strategies, data models and value-chains (Presto Prime Deliverable D2.1.1.) <http://www.prestoprime.org/project/public.en.html>
[15 januari 2011]

BESSER, HOWARD, 2001

Digital Preservation of Moving Image Material?
<http://besser.tsoa.nyu.edu/howard/Papers/amia-longevity.html>
[1 juni 2011]

CCSDS, 2009

Reference Model for an Open Archival Information System
<http://public.ccsds.org/publications/archive/650x0b1.pdf>
[1 maart 2011]

EDMONDSON, RAY, 2004

Audiovisual Archiving: philosophy and principles
<http://unesdoc.unesco.org/images/0013/001364/136477e.pdf>
[1 januari 2011]

GANTZ, JOHN F. ET AL, 2008

The diverse and exploding digital universe
<http://www.emc.com/collateral/analyst-reports/diverse-exploding-digital-universe.pdf>
[18 juli 2011]

LACINAK, CHRIS 2010

A Primer on Codec for Moving Image and Sound Archives
http://www.avpreserve.com/wp-content/uploads/2010/04/AVPS_Codec_Primer.pdf
[29 juli 2011]

NG, YVONNE ET AL , 2010

Strategies for sustainable preservation of born digital public television
http://www.thirteen.org/ptvdigitalarchive/files/2009/10/PDPTV_SustainabilityStrategies.pdf
[1 maart 2011]

LIBRARY OF CONGRESS, 2009

Understanding Premis
<http://www.loc.gov/standards/premis/understanding-premis.pdf>
[15 februari 2011]

OCLC/NARA, 2007

Trustworthy Repositories Audit and Certification (TRAC): Criteria and Checklist http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf [1 maart 2011]

SCHÄDLER, NENA, 2010

Preservation Process Modelling (Presto Prime Deliverable D2.1.1)
<http://www.prestoprime.org/project/public.en.html>
[15 januari 2011]

VELZEN, ERNST VAN, 2010

Must Archives become IT-organisations?
<http://www.fiatifta.org/index.php/archives/2263>
[22 juli 2011]

WRIGHT, RICHARD, 2006

What Archives want, the requirements for digital technology
http://tech.ebu.ch/docs/techreview/trev_308-archives.pdf
[22 juli 2011]

