

**DE NIEUWE
TOEGANG
TOT** Ontwikkelingen in de netwerkcultuur
**AUDIOVISUELE
CONTENT** door
Annemieke de Jong

BEELD EN GELUID



COLOFON

De Nieuwe Toegang is een uitgave van het Nederlands Instituut voor Beeld en Geluid / Project Kennisdisseminatie.

Auteur **Beeld en Geluid, Annemieke de Jong**

Tekstcorrectie **De Schrijverij**

Productiebegeleiding **Afdeling Communicatie Beeld en Geluid**

Illustraties **Beeld en Geluid, Karen Drost**

Fotografie ?

Vormgeving **Fabrique Communicatie en Design**

Lithografie en druk **De Toekomst**

Copyright **Creative Commons (+ logo)**

Dit werk is een Creative Commons
Naamsvermelding-gelijkdelen 3.0 licentie van toepassing

De gebruiker mag:

- Het werk kopiëren, verspreiden en doorgeven
- Remixen – afgeleide werken maken

Onder de volgende voorwaarden:

Naamsvermelding. De gebruiker dient bij het werk de door de maker of de licentiegever aangegeven naam te vermelden (maar niet zodanig dat de indruk gewekt wordt dat zij daarmee instemmen met uw werk of uw gebruik van het werk).

Gelijk delen. Indien de gebruiker het werk bewerkt kan het daaruit ontstane werk uitsluitend worden verspreid krachtens dezelfde licentie als de onderhavige licentie, een gelijksoortige of een compatible licentie.

- Bij hergebruik of verspreiding dient de gebruiker de licentievoorwaarden van dit werk kenbaar te maken aan derden. De beste manier om dit te doen is door middel van een link naar <http://creativecommons.org/licenses/by-sa/3.0/nl/>.
- De gebruiker mag afstand doen van een of meerdere van deze voorwaarden met voorafgaande toestemming van de rechthebbende.
- Niets in deze licentie strekt ertoe afbreuk te doen aan de morele rechten van de auteur, of deze te beperken.

Volledige licentietekst:

<http://creativecommons.org/licenses/by-sa/3.0/nl/>

INHOUDSOPGAVE

1	Inleiding	5
2	Beeld en geluid ontsluiten	7
2.1	De beschrijving als representatie	7
2.2	Semantische niveaus	7
2.3	Details en relevantie	10
2.4	Metadatacreatie nu	11
3	Beschrijven in de workflow	13
3.1	Digitale productie	13
3.2	Media asset management	15
3.3	Authenticiteit	18
3.4	De organische beschrijving	19
4	Zelfgenererende toegang	23
4.1	Automatisch indexeren	23
4.2	Beeld-, audio- en tekstanalyse	24
4.2.1	Gecombineerde technieken	26
4.2.2	Datamining	29
4.3	Het semantisch web	30
4.4	De (on)mogelijkheden van software	32
5	De gebruiker beschrijft	37
5.1	Van consument naar producent	37
5.1.1	Expertbeschrijvingen	37
5.1.2	Crowdsourcing	38
5.1.3	Social tagging	40
5.2	Kunnen gebruikers beschrijven?	41
5.3	Social taggen en de audiovisuele beschrijving	44
6	De trends en hun gevolgen	47
6.1	Nut en noodzaak	47
6.2	Een nieuwe focus voor de professional	48
6.2.1	Context en kwaliteit	50
6.2.2	Waardeoordelen	52
6.2.3	Structurering en thematisering	55
7	Ten slotte	57
	Geraadpleegde literatuur	27

1 INLEIDING

Toegang tot audiovisuele collecties wordt van oudsher verkregen door de onderdelen te ontsluiten d.w.z. door er beschrijvende gegevens aan toe te voegen zoals samenvattingen, *shotlists* en trefwoorden. Via deze gegevens kan kennis genomen worden van de inhoud van films en video- en audioproducties *zonder* dat deze integraal hoeven te worden bekeken of beluisterd. Traditioneel wordt het beschrijven van audiovisuele materialen verricht door professionals, zoals documentaristen en archivariissen werkzaam bij een audiovisueel archief. Gegevens over de producties worden door hen opgesteld volgens welomschreven regels en overzichtelijk gerangschikt in een catalogus. Deze catalogi vormen voor de gebruiker de eerste ingang tot de collectie, als geheel en op het niveau van de onderdelen.

Beeld en geluid komen inmiddels steeds vaker tot ons in een geïntegreerde, digitale vorm. Het gaat daarbij om twee groepen materiaal: de analoge films, audiobanden en videotapes die op een bepaald moment zijn omgezet in digitale files en de producties die altijd al digitaal waren: de *born digitals*. De samenstellende delen van al deze nieuwe digitale objecten zijn onderling niet langer gescheiden. Audio, video, maar ook tekst hebben nu een gemeenschappelijke digitale code, d.w.z. dat ze zijn opgeslagen in dezelfde discrete eenheden, in nullen en enen. De onderdelen van een audiovisuele product kunnen hierdoor apart én als eenheid worden benaderd. Deze modulaire structuur maakt materialen manipuleerbaar en makkelijk uit te wisselen. Door het koppelen van systemen kunnen digitale objecten moeiteloos worden overgenomen in een andere omgeving. Audiovisuele bronnen en hun beschrijvende gegevens worden door deze connectiviteit integraal onderdeel van grote virtuele netwerken. Dit kan het gesloten netwerk zijn van één organisatie, maar ook de grote, openbare delen van het internet. Tekst, audio en video-objecten worden binnen deze netwerken niet alleen geraadpleegd maar ook gecreëerd, verrijkt en vermengd.

Digitalisering en connectiviteit hebben grote gevolgen voor de manier waarop audiovisuele collecties ontsloten gaan worden i.c. voor het tot stand komen van de toegang tot deze materialen. In de digitale netwerkcultuur zal de toegang vanuit meerdere locaties tegelijk worden gecreëerd. Het beschrijven van materialen blijft daarbij niet voorbehouden aan de professionals: zowel de producenten als de consumenten van digitaal beeld en geluid gaan een substantieel deel van de gegevens aanleveren. De modulaire structuur van digitale files biedt vervolgens mogelijkheden voor het automatisch analyseren van hun inhoud door computers en het aanbieden van het resultaat als doorzoekbare gegevens.

Deze werkuitgave wil een overzicht bieden van deze ontwikkelingen. Gepoogd is bovenstaande, nieuwe praktijken te verbinden aan de bestaande ontsluitingsconventies voor beeld en geluid. Daartoe zijn de effecten van digitalisering en connectiviteit op het toegankelijk maken van institutionele audiovisuele collecties in kaart gebracht. Vragen die aan de orde komen zijn: welke grote trends kunnen worden onderscheiden? Welke nieuwe vormen van ontsluiting leveren ze op? Wat is hun invloed op de totstandkoming, de vorm en de kwaliteit van beschrijvende informatie? Wat zijn de gevolgen voor de professionele documentarist, zijn werkwijze en zijn ethiek? Om de effecten op het domein te kunnen begrijpen, wordt allereerst ingegaan op de algemene beschrijvingsmethoden voor beeld en geluid. Vervolgens worden de verschillende trends uiteengezet en gewaardeerd. Aan het eind wordt een voorzichtige balans opgemaakt. Bij het samenstellen van deze publicatie

fungeerde de omroeparchie omgeving – waar de impact van de veranderingen al duidelijk zichtbaar is – als belangrijke informatiebron. Uitgangspunt was echter dat uiteindelijk alle audiovisuele archieven in min of meerdere mate met de beschreven ontwikkelingen zullen worden geconfronteerd.

2 BEELD EN GELUID ONTSLUITEN

2.1 DE BESCHRIJVING ALS REPRESENTATIE

Audiovisueel materiaal is opgebouwd uit informatie die wordt overgebracht door beeld en geluid. Het is lastig om dit materiaal zijn eigen index te laten vormen, zoals dat wel kan met tekstuele documenten. Dit komt omdat beeld en geluid niet bestaat uit eenduidige symbolische eenheden zoals woorden, die kunnen worden gebruikt als directe toegang tot de inhoud. Bovendien zijn audiovisuele documenten samengesteld uit sequentiële en temporale objecten: ze zijn *time-based*. Dit alles maakt een snel en synthetisch begrip van de inhoud problematisch. Om te kunnen weten wat er op de film, de tape of in de file staat moet een audiovisuele productie beschreven worden. Een beschrijving biedt een gebruiker de mogelijkheid zich een idee te vormen van de structuur en de 'verhaallijn' van een programma *zonder* dat hij/zij het programma daadwerkelijk in zijn geheel hoeft te bekijken of te beluisteren. De beschrijvingen kunnen bij wijze van shotlist in tijd gelijk oplopen (*time-aligned*) met het beeld- en geluidsmateriaal. Wanneer eenmaal een relevant element van de beschrijving is geselecteerd, kan het corresponderende fragment op een filmtafel, een videorecorder of een computer worden afgespeeld.

Professionele audiovisuele documentaristen en catalogiseerders schrijven op wat er in het programma te zien is. Ze doen dit door het bekijken van het materiaal, vaak aan de hand van bijgeleverde gegevens afkomstig van de maker, producent of distributeur. Hun beschrijving fungeert als een vervanging of representatie van de audiovisuele productie. Het definiëren van de inhoud van een beeld teneinde zijn representatie vast te stellen is een gecompliceerde taak. Een beschrijving moet voldoende kenmerken bevatten om een audiovisuele productie te kunnen identificeren en deze te onderscheiden van soortgelijke producties. De verschillende elementen en betekenislagen in radio, televisie en film moeten worden uiteengerafeld en benoemd. De documentalist wordt daarbij geacht te bewerkstelligen dat hetzelfde programmafragment uiteindelijk kan worden benaderd door gebruikers met volledig verschillende zoekvragen, zoals programmamakers, journalisten, beeldresearchers, studenten, onderzoekers en het algemeen publiek.

2.2 SEMANTISCHE NIVEAUS

Een formele beschrijving legt de eigenschappen en de uiterlijke verschijningsvorm vast van het materiaal. Door er formele kenmerken aan toe te kennen, wordt een productie terugvindbaar gemaakt op zaken als nummer, titel, datum, lengte, namen van de makers, tijdcodes, auteursrecht en genre. Onder de formele gegevens vallen ook sommige 'fysieke' kenmerken: vindplaats van de drager, bepaalde technische informatie, formaateigenschappen, publicatiegegevens en mediatype. Formele gegevens zijn doorgaans aan het materiaal en de bijbehorende productiegegevens zelf te ontleen en vergen geen interpretatie. Dit is anders met de inhoudelijke ontsluiting. Een inhoudelijke beschrijving geeft het onderwerp en de deelonderwerpen weer middels een samenvatting en/of een shotbeschrijving. De onderwerpen worden ontsloten door gestandaardiseerde, geüniformeerde trefwoorden. Inhoudelijke ontsluiting vergt analyse: er moet worden bepaald



Professionele documentaristen werkzaam bij een audiovisueel archief viewen en beluisteren audiovisuele producties. Het beschrijven geschiedt volgens vaste, uniforme regels.

waar het programma over gaat en de meest karakteristieke elementen moeten worden geselecteerd en vastgelegd. Een zekere mate van subjectiviteit is hierbij onvermijdelijk. Er is immers sprake van selectie (wat zijn hoofd- en bijzaken en ga ik die in mijn beschrijving allemaal noemen?) en interpretatie (duiding van de onderwerpen en het aangeven van de context waarin e.e.a. zich afspeelt). Richtlijnen voor dit selectie- en interpretatieproces zijn maar ten dele te geven.

De inhoudsbeschrijving van een audiovisuele productie valt uiteen in drie onderdelen of niveaus: de informatiecontent, de audiovisuele content en de *stockshots*. Het eerste niveau, de informatiecontent, geeft antwoord op het wie, wat, waar en wanneer van een productie: de feiten. Op dit niveau wordt het centrale onderwerp of object benoemd, evenals de context en de gezichtspunten die aan de orde komen. Namen en functies van personen, groepen en organisaties worden geïdentificeerd, evenals geografische locaties. De onderwerpen worden

Informatiecontent

Wekelijkse uitzending waarin presentator Wilfred Kemp spreekt met een hoofdgast over spirituele en religieuze thema's op een voor hem of haar belangrijke locatie. Deze aflevering: Imca Marina op Ameland.

Productie: KRO. Uitzonden: 27 oktober 2004. Tijdsduur: 25'13"

Audiovisuele content

Kemp laat zich door zangeres Imca Marina rondleiden in haar authentieke huis in Nes. Samen bezoeken ze achtereenvolgens het dorpscentrum, het katholieke kerkje, het kerkhof, het strand en natuurgebied Het Oerd. Al wandelend wordt gesproken over het leven op Ameland, Marina's scheiding en haar belangstelling voor theosofie. Als afsluiting declameert Imca Marina een van haar eigen gedichten.

Stockshots

00:06:14 - 00:07:02 Algemene beelden historisch dorpscentrum Nes.

00:19:08 - 00:19:19 Mediumshots vuurtoren in de avond en overdag.

00:21:22 - 00:21:59 Jutters op het strand aan het begin van de dag;
shots opkomende zon.

Fig. 1. Een voorbeeld van een catalogusbeschrijving waarin onderscheid is aangebracht tussen informatiecontent, audiovisuele content en een beschrijving van stockshots.

zowel beschreven op het niveau van het programma als geheel, als op het niveau van de onderdelen.

Centrale vraag bij het tweede semantische niveau, de audiovisuele content, is: wat is er in het programma te zien en/of te horen? Hiertoe worden de scènes, de shots, de uitspraken en de interviewgedeelten chronologisch weergegeven en zo nodig getimed. Ook stilistische en compositie-elementen (montage, shottypes, cameraposities, muziek en effecten) horen tot de audiovisuele content. Door het beschrijven van deze vormkenmerken wordt de relatie tussen het onderwerp en zijn weergave, ofwel de conceptuele en de visuele kenmerken, expliciet gemaakt. Aandacht voor de vorm verleent een beschrijving een soort 'objectiviteit'. Daarnaast helpt het bij het onderscheiden van verschillende items over hetzelfde onderwerp. Toevoegingen als wide shot, mediumshot, close-up, pan, tilt, luchtopname, kikkerperspectief e.d. vergroten de onmiddellijke herkenbaarheid van bepaalde beelden.

Stockshots zijn een integraal onderdeel van de audiovisuele content en vormen het derde semantische niveau. Het gaat hier om shots of – in het geval van audio – om quotes die geschikt zijn om autonoom te kunnen worden hergebruikt in een andere context, bijvoorbeeld een nieuw te maken productie. Stockshots moeten aan bepaalde formele vereisten voldoen: ze mogen niet te kort zijn en er mag geen zichtbare contextinformatie in

beeld zijn, zoals logo's, titels, tekstblokjes of een reporter ter plekke in een zgn. stand-upper. Bij film en video is onderscheid te maken tussen algemene, neutrale stockshots, zoals beelden van 'een drukke markt', of 'spelende jonge dieren' en meer specifiek stockmateriaal, bijv. 'exterieur van het Witte Huis in de sneeuw' of 'recente bijeenkomst van een goed gevulde Tweede Kamer'. Om snelle terugvindbaarheid te bevorderen worden stockshots bij voorkeur apart benoemd en uitgebreid getimed en beschreven.

2.3 DETAILS EN RELEVANTIE

Het onderscheid tussen de verschillende semantische lagen van een inhoudsbeschrijving – informatiecontent, audiovisuele content en stockshots – zegt nog niet direct iets over het wel of niet opnemen van bepaalde inhoudelijke elementen in een beschrijving. Ook de mate van detaillering van de beschrijving zelf wordt er niet duidelijker door. Een tweetal algemene catalogiseerprincipes biedt doorgaans een eerste houvast.

Allereerst is er de drempel van het detail ('seuil de finesse'). Dit principe houdt in dat *niet* wordt beschreven wat onmiskenbaar een integraal deel is van een groter geheel. Dus: de mensenmenigte wordt beschreven, maar niet de individuele figuren en gezichten. Ander voorbeeld: het landschap wordt in de beschrijving vermeld, echter niet de afzonderlijke struiken, bomen en heuvels in dat landschap. Het tweede principe, de drempel van de relevantie ('seuil de pertinence'), betreft het opnemen van objecten, gebeurtenissen en personen die binnen het betreffende beeld van betekenis zijn en duidelijk kunnen worden geïdentificeerd. Dus: de onderwerpen en de gasten van een talkshow, maar niet de willekeurige elementen uit het decor. En ook: het gebouw, maar niet de toevallig passerende auto's. Uitzonderingen zijn er overigens altijd, zoals bij atypische elementen in het beeld, in het geval van bruikbare stockshots en wanneer er achteraf bepaalde (nieuwe) contextinformatie voorhanden is. Bijvoorbeeld: tussen de mensenmenigte loopt een 'bekende Nederlander' of: de auto die toevallig het gebouw passeert is er net een met een buitenissig, historisch model.

Er is nog een aantal andere factoren dat direct van invloed kan zijn op de 'diepte' c.q. de detaillering van een beschrijving. Allereerst is er de complexiteit die wordt bepaald door de formele eigenschappen van het te beschrijven materiaal. Deze eigenschappen kunnen worden gezien als intrinsiek en zijn in zekere zin autonoom. Zo kan een programma een meer of minder gelaagde structuur bezitten die zich uit in de opbouw van beeld- en geluidselementen, het dooreenlopen van shots en interviews en een rijk en gevarieerd gebruik van archiefmateriaal uit verschillende bronnen met rechteninformatie die per onderdeel verschilt. Als intrinsieke eigenschap zou ook de informatiewaarde of informatiedichtheid van een programma kunnen worden bestempeld. Over een uitgebreid nieuwsbulletin met meerdere reportages en studiogesprekken is nu eenmaal meer te zeggen dan over een kort en eenvoudig quizprogramma.

Niet intrinsiek is de schaarste en/of de uniciteit van beelden of van audio binnen de collectie die wordt beschreven of, breder, in de context van nationale en internationale audiovisuele collecties. Dit criterium geldt in de regel voor bepaald historisch materiaal, maar kan ook betrekking hebben op unieke contemporaine beelden of geluidsopnamen. Hoe schaarser of

'unieker' het materiaal, hoe belangrijker het is om terugvindbaarheid te garanderen op een gedetailleerd niveau.

Een vierde factor wordt gestuurd vanuit het (her)gebruik van het materiaal. Hier is het de combinatie van het te verwachten gebruikerstype, het soort gebruik en de dynamiek van de vraag, die bepaalt welke beschrijvingsdiepte het meest geschikt is voor een bepaald genre, programma of bepaald fragment. Voor audiovisueel materiaal kunnen zich diverse typen klanten aandienen. Belangrijke categorieën zijn: de professionele gebruikers uit de mediawereld en de creatieve industrie, gebruikers uit het domein van wetenschap en onderwijs en het algemeen publiek. Deze groepen hebben uiteenlopende wensen waar het gaat om de graad van detaillering van de beschrijving, de identificatie van shots en citaten en de beschikbaarheid van contextinformatie. Wat ze vaak ook onderscheidt is de tijdsdruk achter hun vraag: een redacteur van een actualiteitenrubriek heeft meestal niet de tijd om shots zelf te viewen en is dus afhankelijk van een uitgebreide, getimede beschrijving. Dit ligt anders voor een onderzoeker die wél in de gelegenheid is materiaal zelf te bekijken. Deze gebruiker is dus minder afhankelijk van een uitgebreide beschrijving. In het algemeen kan worden gesteld dat de diepte (detaillering) en breedte (hoeveelheid gegevens die over één programma wordt vastgelegd) van de beschrijving van bepaalde programma's of fragmenten, wordt bepaald aan de hand van de mate waarin de te verwachten klantgroepen afhankelijk zijn van de (snelle) beschikbaarheid, de uitgebreidheid en de betrouwbaarheid van de gegevens.

2.4 METADATACREATIE NU

Traditioneel is het maken van formele en inhoudelijke beschrijvingen een exclusieve taak van professionele documentalisten, archivariissen en catalogiseerders, werkzaam bij audiovisuele archieven en andere collectiebeherende instellingen. Zij bepalen welke keuzen worden gemaakt en derhalve of en hoe de inhoud wordt ontsloten. Het zijn de professionals die het laatste en beslissende woord hebben over de mate van toegang tot de collecties, de programma's en de items. Zij werken daarbij zo uniform en consistent mogelijk, met inachtneming van de belangen van de gebruikers en de rechthebbenden én de principes van authenticiteit en integriteit. Een uitgebreid professioneel instrumentarium van beschrijvingsregels, ontsluitingsmethodieken en trefwoordsystemen dat is opgebouwd gedurende vele decennia, is ze daarbij behulpzaam.

Deze situatie is hard op weg te veranderen. Inmiddels worden namelijk ook *buiten* de professionele omgeving veel formele en beschrijvende gegevens gegenereerd over audiovisuele archiefcollecties. Deze ontwikkeling is een rechtstreeks gevolg van de digitalisering die unieke audiovisuele producties, vastgelegd op fysieke dragers, verandert in files die vanaf meerdere locaties kunnen worden gekopieerd, bewerkt en uitgewisseld. Met de vorm van de 'drager' van audiovisuele producties zijn ook de traditionele benamingen van materiaal en gegevens gewijzigd: digitale audiovisuele items – of het nu gaat om hele programma's of fragmenten – heten *content*. De labels en gegevens – maar ook de complete beschrijvingen en de trefwoorden – worden inmiddels algemeen aangeduid als *metadata*, een term die uitdrukt dat het gaat om informatie over *digitale* bronnen die *uitwisselbaar* zijn in een *netwerk*.

Een deel van deze metadata wordt onwillekeurig gegenereerd. De computersystemen waarin materiaal wordt gedigitaliseerd, opgeslagen of bewerkt, creëren immers automatisch gegevens zoals datum, grootte en versie van een document. In veel gevallen gaat het echter om metadata die heel doelbewust worden gemaakt. Zo leggen steeds meer makers en producenten maar ook gebruikers van audiovisuele materialen zélf beschrijvende gegevens vast in centrale of eigen systemen. Daarnaast worden er softwaretoepassingen ontwikkeld die beeld en geluid op een betekenisvolle manier kunnen analyseren en 'beschrijven'. In al deze gevallen gaat het om het creëren van formele en/of inhoudelijke gegevens, tot voor kort het exclusieve domein van de professionele audiovisuele documentalist, om 'archivale' metadata dus, bedoeld om toegang te creëren tot de inhoud.

3 BESCHRIJVEN IN DE WORKFLOW

3.1 DIGITALE PRODUCTIE

Tot voor kort – en in veel gevallen nog steeds – bevond het audiovisuele archief zich aan het eind van de productieketen. Een productie op film, video of audiotape werd binnen of buiten de moederinstelling vervaardigd om na vertoning of uitzending te worden opgenomen in het omroeparchief, het bedrijfsarchief of het filmarchief. Samen met de fysieke drager kwamen er vaak papieren gegevens mee naar het archief. Deze informatie over de productie, de maker(s) en de inhoud, zoals bijvoorbeeld logboeken, scripts, presentatieteksten en rechtenggegevens, werd door de documentalist of archivaris gebruikt als bron bij het maken van de catalogusbeschrijving.

Momenteel maken de meeste audiovisuele catalogi en collecties deel uit van een digitaal netwerk. Dat kan een grote, complexe netwerkomgeving zijn zoals het archief van een omroep dat integraal onderdeel vormt van de productieomgeving van radio- en televisieprogramma's. Het kan ook gaan om een eenvoudig netwerk zoals dat waarin de catalogus van de audiovisuele collectie van een universiteitsbibliotheek of gemeentearchief is verbonden met de andere bedrijfsonderdelen. Feit is dat veel archiefsystemen niet langer standalone zijn, maar zich weten ingebed in organisatiebrede informatiestromen en workflows.

De productie en distributie van veel audiovisuele materialen verloopt meer en meer *server-based*. Steeds vaker is er sprake van een 'tapeloze' omgeving die de complete productieketen omvat: van preproductie via productie en acquisitie, tot postproductie, publicatie en archivering. Door het linken van alle systemen en subsystemen waarbinnen digitale content tot stand komt worden de onderdelen van het productieproces verbonden. De koppeling tussen individuele units vormt een digitale workflow. Materialen en bijbehorende gegevens bevinden zich dan niet langer in een lineair proces maar circuleren over het netwerk om simultaan te worden opgeslagen, gedistribueerd, geraadpleegd, hergebruikt en gemuteerd. Het totaal aan content en metadata vormt een virtuele collectie van mediaobjecten: losse video- en audiofragmenten, *intermediates*, teksten, foto's, grafiek en complete programma's en series. Vanuit deze virtuele verzameling wordt gepubliceerd op verschillende platforms zoals televisie en radio, websites, dvd's en cd's of via mobiele telefonie. Deze meervoudige publicatie wordt *divergentie* genoemd.

Uiterste consequentie van zo'n geïntegreerde productieomgeving is het naar de achtergrond verdwijnen van het archief als fysieke entiteit, afdeling of gebouw met een afgebakende collectie en professionele documentaristen. Het archief, of liever: 'het archiveren', wordt in de *tapeless* omgeving een dynamische, decentrale activiteit, een integraal onderdeel van de digitale workflow, uitgevoerd tijdens het hele productieproces. Het verschijnsel archief is in zo'n omgeving misschien nog het best te kwalificeren als een specifieke functie van een systeem. Van exclusief eindpunt en uitgifteloket van geauthenticeerde materialen is 'het archief' veranderd in een gedistribueerd, dynamisch gebruiksdépot van digitale mediaobjecten.

3.2 MEDIA ASSET MANAGEMENT

Tijdens de productieprocessen worden er door de diverse systemen automatisch metadata gegenereerd. Zo levert het creëren van digitale items identificatiegegevens, formatinformatie en tijdcodes op. Een digitale camera registreert locatie, tijdcode en weersgesteldheid tijdens een opname. Editingsystemen leggen wijzigingen vast in de montage. Er is ook manuele invoer, en wel door iedereen die aan de producties meewerkt. Regisseurs, technici, video-editors, programma-assistenten en beeldresearchers voeren de metadata in die horen bij het item of het programma waar ze aan werken. Op deze manier worden door de hele organisatie heen administratieve, juridische, technische en inhoudelijke en publicatiegegevens gegenereerd.

In een volledig geïntegreerde productieomgeving is er feitelijk geen centrum meer waar deze metadata en de content waarop ze betrekking hebben centraal worden opgeslagen.

PRAKTIJKVOORBEELD

DIGITALE WORKFLOW OP HET MEDIA PARK

Op het Media Park draait sinds een aantal jaren De Digitale Voorziening (DDV), een gezamenlijke digitale infrastructuur tussen de publieke omroepen, Technicolor en Beeld en Geluid. DDV vormt de eerste stap op weg naar een volledig geïntegreerd productie-, uitzendings- en archiveringsproces. Binnen DDV wordt het geproduceerde omroepmateriaal 10 dagen vóór en 15 dagen na uitzending opgeslagen. De bijbehorende metadata worden door de programmamakers, technici en productiemedewerkers ingevoerd in een gemeenschappelijke database. Naast centrale opslag is De Digitale Voorziening ook een 'verkeersplein', geschikt voor het versturen en ontvangen van video in uitzendkwaliteit tussen alle aangesloten partijen. Audiovisuele content kan via DDV makkelijk en snel worden gevonden en bekeken. De gegevensafhandeling is verbeterd, waardoor ook de betrouwbaarheid van het uitzendproces is vergroot. Via DDV kunnen omroepen hun programma's niet alleen op de traditionele manier uitzenden, maar ook als interactieve tv en via mobiele communicatie en (breedband)internet. Dit alles is hoogstbevorderlijk voor het hergebruik van audiovisuele content. Voor DDV is een gezamenlijke metadatalijst samengesteld van ca. 150 velden. Deze velden worden tijdens de verschillende productiestadia ingevuld. De metadata zijn zodanig gedefinieerd dat ze na creatie automatisch IMMIX, de archiefcatalogus van Beeld en Geluid, kunnen binnen stromen. Documentalisten van de afdeling Mediadocumentatie corrigeren en verrijken de gegevens zo nodig. De content zelf wordt na publicatie opgeslagen in het Digitale Archief. Door de digitale workflow kan Beeld en Geluid – ten opzichte van het analoge tijdperk – een veelvoud van radio- en televisieprogramma's in zijn collecties opnemen.

www.dedigitalevoorziening.nl

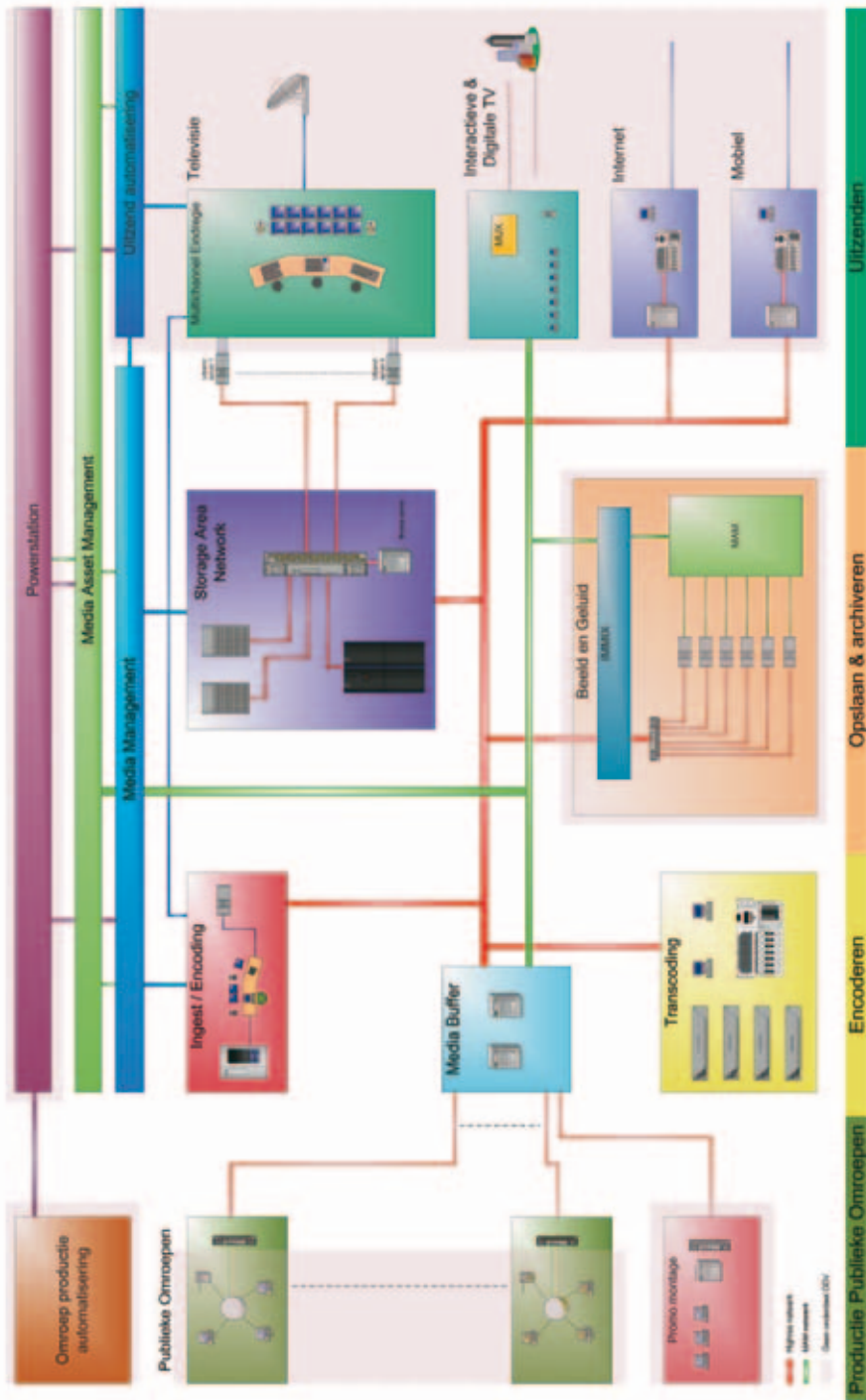


Fig. 2. De aangesloten systemen, processen en organisaties in De Digitale Voorziening.
Bron: Video & Audio Report

BESCHRIJVEN VOOR EXPLOITATIE

Binnen ROOS (Stichting Regionale Omroep Overleg en Samenwerking), de overkoepelende organisatie van de Nederlandse regionale omroepen, is een gezamenlijk metadatamodel ontwikkeld waarin audiovisuele content wordt beschreven om te kunnen worden verspreid naar afnemers en distributieplatforms. Een cruciale rol spelen de metadata die het materiaal moeten identificeren om het automatisch naar het juiste kanaal of doelgroep te kunnen laten versturen, zoals formele en inhoudelijke trefwoorden, rubrieken en mediacategorieën. De metadata worden op vier manieren gegenereerd: handmatig door verslaggevers en programmamakers en automatisch door camera's en geluidsrecorders op het moment van opname en acquisitie. Ook software voor de analyse van beeld en geluid produceert beschrijvende metadata. Ten slotte worden bestaande metadata in de productieomgeving, zoals gegevens in newsroomsystemen en draaiboekapplicaties, hergebruikt. De documentalisten in het archief toetsen de uiteindelijke beschrijvingen en stellen deze eventueel bij. Het ROOS-metadatamodel kent verschillende varianten waaronder een archiveringsmodel waarin alle voorkomende gegevens zijn opgenomen, die bovendien per omroep kunnen verschillen. Een andere variant vormt het wat compactere uitwisselingsmodel. De verwachting is dat deze doordachte metadatastructuur, gecombineerd met een goed opgebouwde digitale workflow – zonder veel personele inspanning – nieuwe kansen creëert voor (her)exploitatie van content.

www.mediamanage.nl/project/metadatamodel-roos

Daarvoor in de plaats staat een proces centraal: *media asset management*. MAM definieert de parameters voor het uitwisselen tussen systemen, afdelingen en medewerkers en schept aldus de voorwaarden voor de geïntegreerde werkprocessen. MAM is geen systeem, geen afdeling en ook geen archief. Het verschijnsel moet eerder worden gezien als een centraal concept voor het adequaat managen van media vanaf het moment van productie via publicatie naar opslag en hernieuwde publicatie.

De digitale productie verloopt niet lineair en is niet transparant. Het gaat hier om grote hoeveelheden digitaal én (vaak ook nog) analoog materiaal, om gestructureerde en ongestructureerde onderdelen opgeslagen in onderdelen en als geheel, in meerdere versies en tijdens verschillende productiestadia. Vanuit het oogpunt van informatiemanagement is er sprake van een meervoudige relationele context van creatie, bewerking en distributie. Om de processen soepel te laten verlopen zijn regels en protocollen nodig die interoperabiliteit en compatibiliteit garanderen tussen alle verspreide systemen, files en databases. De content – of het nu gaat om een hele serie programma's of om een enkele clip – moet door iedere deelnemende partij in het productieproces te allen tijde eenduidig kunnen worden geïdentificeerd. Sleutel tot terugvindbaarheid van de content ligt in het gebruik van uniforme metadata. Het fundament van de gedistribueerde digitale productie wordt dan ook gevormd

END-TO-ENDMODEL ALS SLEUTEL TOT INTEGRATIE

De BBC onderscheidt drie verschillende benaderingen van content en metadata-management. In het Media Data Exchange Project ligt de nadruk op het beheersbaar maken van het digitaal geproduceerde BBC-radio- en -televisiemateriaal *zelf*. Speciale labels gearrangeerd in een organisatiebreed model moeten alle verspreid opgeslagen digitale shots, items, geluid, tekst, stills en graphics (de zgn. mediaobjecten) terugvindbaar maken. Invoer geschiedt uniform. Een initiatief dat de systeemverbindingen op een andere manier gebruikt is het BBC Digital Media Asset Management Centre Project. Hierbij worden digitale archiefcatalogi (muziek, beeld en geluid) gelinkt aan de systemen die managementinformatie verschaffen over conserverings- en overschrijfprojecten. Geïntegreerd zoeken kan via een webbrowser op het BBC-intranet. Zoeken en vinden was ook het voornaamste doel waarmee BBC-journalisten het geavanceerde Electronic News Production System (ENPS) hebben ontwikkeld. Dit systeem verschaft snel en overzichtelijk toegang tot allerlei interne en externe multimediale bronnen. Centrale rol in al deze metadatakoppelingen speelt het Standard Media Exchange Framework (SMEF™), een (meta)datamodel dat is ontwikkeld om het managen van de media-assets als een end-to-endproces door te kunnen voeren. Het SMEF Data Model biedt een set definities voor alle metadata-elementen in productie, distributie en archivering en de relaties daartussen. Dit allesomvattende model moet de sleutel zijn tot de uiteindelijke integratie en harmonisatie van alle systemen in alle BBC-afdelingen en -locaties. SMEF kan als referentiemodel gratis worden gedownload.

www.bbc.co.uk/guidelines/smf/

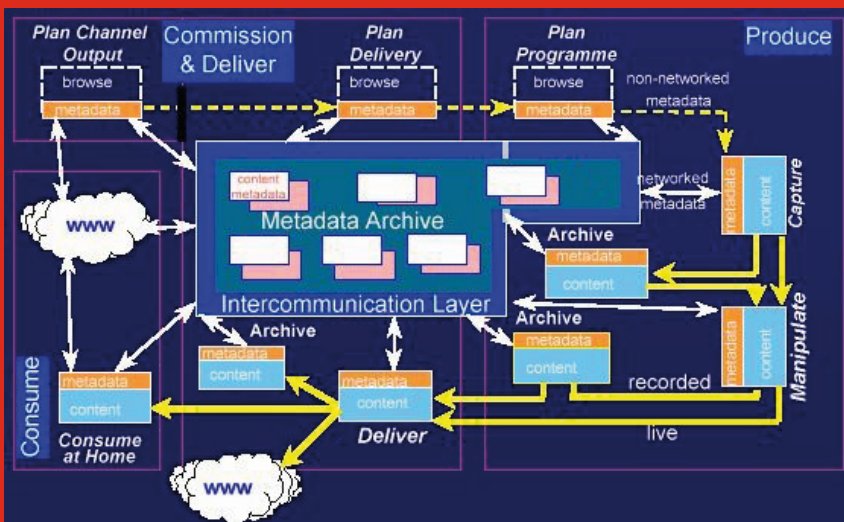


Fig. 3. Een model van de integrale productie, distributie en opslag van content en metadata bij de BBC.

Bron: Richard Hopper, BBC

door een gemeenschappelijke datastructuur waarin alle metadata zijn opgenomen die te pas komen bij het maken, publiceren en archiveren. Er wordt ook in aangegeven op welk moment in het proces metadata kunnen of moeten worden toegevoegd, verwijderd en/of gemuteerd. De vorm en inhoud van individuele gegevens is in zo'n datamodel gestandaardiseerd en ook de relaties tussen groepen metadata worden erin vastgelegd. Gebruik en hergebruik van de mediaobjecten of *assets* staan in de audiovisuele productieomgeving centraal. Granulariteit – het op detailniveau kunnen identificeren van digitale content – is dan ook een belangrijke eis bij het modelleren van de metadata. Tegelijkertijd moeten alle onderdelen zijn ingepast in de bredere context van een programma, een serie en een collectie. Via een dergelijk organisatie-breed datamodel ontstaat aldus samenhang tussen verschillende mediatypen (tekst, audio en video), tussen taken in het proces (preproductie, editing, archiveren en publiceren) en tussen informatieniveaus (van programma naar fragment, van serie naar aflevering).

3.3 AUTHENTICITEIT

In de dynamische productieomgeving wordt digitaal materiaal overgebracht in ruimte, d.w.z. tussen mensen, afdelingen en systemen, en in tijd (van format naar format en binnen verschuivende technologische contexten). Dit brengt risico's met zich mee voor de authenticiteit en de integriteit van het materiaal. Is een digitaal item wel wat het voorgeeft te zijn? Is er in de loop van de productiegang niet mee geknoeid? Professioneel beheer van digitale content impliceert dat hier te allen tijde uitsluitel over kan worden gegeven. Hiertoe moet de levensloop van een audiovisuele productie in de metadata worden vastgelegd.

Onmiddellijk na creatie wordt de content actief bewerkt en gebruikt door de makers. Op het moment dat het directe operationele nut eindigt – vlak voor of na vertoning of uitzending – wordt ze voor langere tijd opgeslagen, centraal of gedistribueerd. Wanneer fragmenten of programma's worden opgevraagd om te worden hergebruikt, begint een nieuwe cyclus. Het is van belang al deze bewegingen van aanvang af te documenteren. Hoe nauwgezetter de gebruiks- en bewerkingsgeschiedenis wordt bijgehouden, hoe 'archiefwaardiger' – in de zin van authenticiteit en integriteit – de content. In een digitaal productienetwerk gaat het beheren van audiovisuele materialen zo eigenlijk een beetje lijken op *recordmanagement*, met zijn nadruk op de levensloop en de organische groei van documenten, en zijn onderscheid tussen courante en niet-courante materialen.

Metadata brengen ook het onderscheid aan tussen dynamisch productiemateriaal en die items die zijn bedoeld om duurzaam te bewaren: het erfgoed. Dit erfgoeddeel – een selectie uit de losse mediaobjecten en hun representatie als eindproduct voor web, televisie of anderszins – zal ergens in het proces moeten worden geormerkt om te worden opgenomen in een *trusted digital repository*. Speciale metadata, bestemd voor identificatie en authenticatie, begeleiden de overgang naar zo'n betrouwbare omgeving. Dit *trusted repository* kan zich zowel centraal als gedistribueerd voordoen en fungeert in zekere zin als de digitale tegenhanger van het fysieke archief aan het eind van de analoge productieketen. Een digitaal omroepproductienetwerk kan worden gezien als een continuüm met twee transitionele domeinen: het domein van de *media assets*, i.c. de dynamische objecten die worden gecontroleerd door het MAM-systeem en vooral een directe operationele waarde bezitten. Dit MAM-domein loopt vloeiend over in de tweede omgeving waarin dát

materiaal toegankelijk is gemaakt dat kan worden bestempeld als erfgoed: de radio- en televisieprogramma's, de websites en de andere publicaties die zijn samengesteld uit de losse mediaobjecten. Via de metadata wordt toegang gecreëerd tot beide domeinen.

3.4 DE ORGANISCHE BESCHRIJVING

Het zal duidelijk zijn dat een geïntegreerde productie- en archiefomgeving voor zowel de inhoud als de totstandkoming van de beschrijving van materialen belangrijke gevolgen heeft. Ten eerste gaat de invulling van een beschrijving meer gelijk oplopen met de wording van een 'fysieke' productie. De kiem van de beschrijving wordt al gelegd in de allereerste fase, tijdens het bedenken van een programma. Gedurende de opeenvolgende productiestadia worden de gegevens steeds verder aangevuld. Beschrijvende gegevens komen zo feitelijk tot stand lang voordat de content als afgemonteerd, integraal eindproduct wordt opgeslagen. In een omgeving waar digitaal wordt geproduceerd, is het beschrijven dan ook niet meer los te zien van de processen waarin het beschrevene, de audiovisuele content, wordt gemaakt en (her)gebruikt. Net als bij de content zélf komt de nadruk hier ook te liggen op het proces van het creëren, bewerken en muteren.

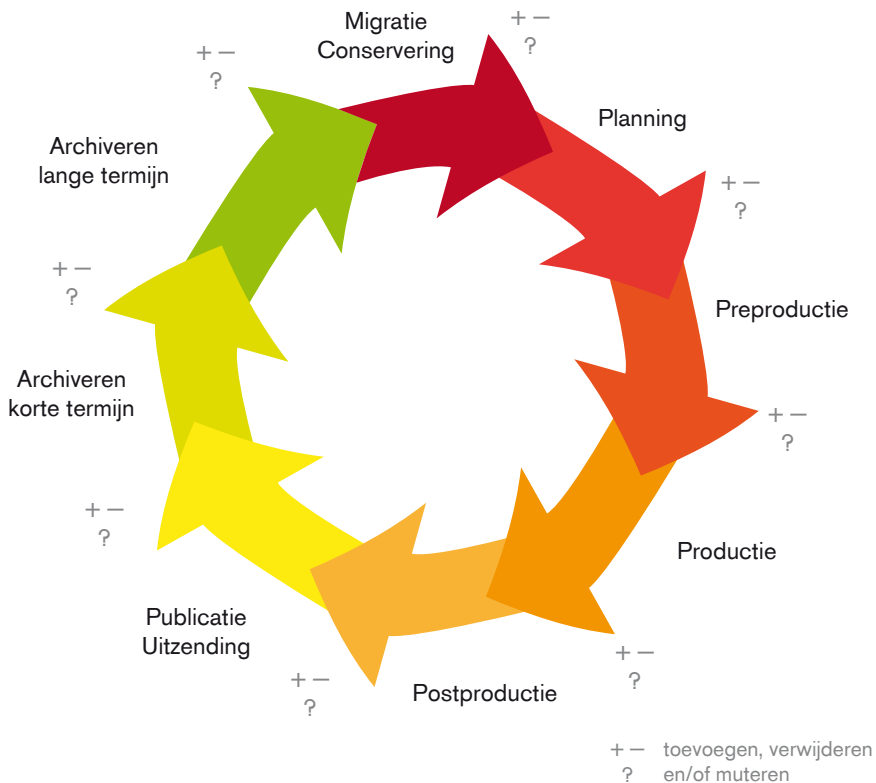


Fig. 4. De levenscyclus van metadata in het digitale productieproces: in elke fase worden er gegevens toegevoegd, gemuteerd en/of verwijderd.

METADATAKWALITEITSCRITERIA

Het archief van de Zweedse publieke omroeporganisatie SVT is nauw betrokken bij de opzet van het 'Coordination System' dat een groot aantal bestaande eilandsystemen binnen de omroep aan elkaar koppelt tot één groot productienetwerk. De bestaande catalogusinformatie over de SVT-programma's en -items is opnieuw gestructureerd en de beschrijving van digitale content verloopt nu volgens een organisatiebreed format en een uniforme terminologie. Het archief heeft het informatieontwerp mede bepaald en kreeg de algehele supervisie over de regelgeving en de kwaliteit van de metadata. Hiertoe is een set criteria ontwikkeld waaraan alle gegevens die worden gecreëerd en gemuteerd door deelnemers aan het productieproces moeten voldoen. Als grote inspiratiebron fungeerde het gedachtegoed van Diane Hillmann zoals verwoord in het hoofdstuk 'The Continuum of Metadata Quality' in haar werk *Metadata in Practice*. De criteria zijn achtereenvolgens: *completeheid*: het programma wordt zo gedetailleerd mogelijk beschreven als het budget toelaat; *precisie*: metadata moeten accuraat, correct en feitelijk zijn; *herkomst*: levenscyclus van de metadata (i.c. alle mutaties en aanvullingen) moet zijn gedocumenteerd; *conform aan gebruikersverwachtingen*: een zoekinterface met voor gebruikers nuttige metadatavelden; *logische consistentie en coherentie*: consequent gebruik van metadata in verschillende productiesystemen; *promptheid*: metadata moeten op tijd geleverd worden en tijdig worden aangepast aan iedere update van de content; en *toegankelijkheid*: alle gebruikersgroepen moeten toegang hebben tot het archiefsysteem zonder extra kosten.

Diane I. Hillmann/Elaine L. Westbrooks, *Metadata in Practice*
Chicago American Library Association 2004, 285 p., ISBN 08389 08829
(hoofdstuk 'The Continuum of Metadata Quality' is gepubliceerd op Google Books)

Het moment van voltooiing van de beschrijving wijzigt. Want wanneer is een beschrijving af? Voorheen was dit duidelijk. Het archief fungeerde als eenduidig eindstation en met het opnemen van de beschrijving in de catalogus was de kous af. In de digitale productieomgeving bestaat er al centraal toegankelijke informatie vóórdat het programma is voltooid. Deze gegevens worden gaandeweg uitgebreid en gemuteerd om uiteindelijk terecht te komen in een centraal of gedistribueerd virtueel (archief)systeem. Na uitzending of publicatie houdt het documenteren niet op: door hergebruik van fragmenten en programma's in nieuwe publicaties zullen aan de oorspronkelijke beschrijving steeds weer andere gegevens worden toegevoegd. In de digitale productie- en archiefomgeving verandert de beschrijving van 'af en gesloten' naar 'open en dynamisch'. Ze kan niet langer worden gezien als statisch eindproduct.

In een digitaal netwerk verandert ook de aard van de gegevens over een productie. Naast formele en inhoudelijke informatie gaat een beschrijving ook dynamische kenmerken

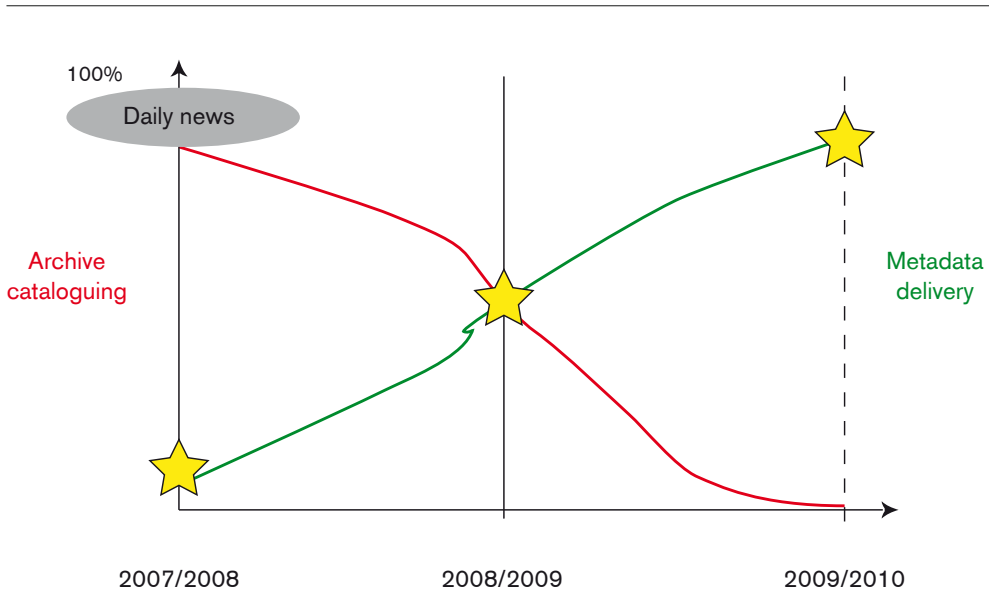


Fig. 5. In een paar jaar tijd zullen de professionele beschrijvingen van nieuwsitems geheel zijn vervangen door de metadata die worden gegenereerd tijdens het productieproces: dit geldt althans bij de Zweedse omroep SVT.
Bron: SVT

bevatten over de productiestadia, over technische bewerkingen en over gebruik en verspreiding. Er kunnen gaandeweg selectie-, conserverings- en digitaliseringsgegevens aan worden toegevoegd. Op sommige momenten tijdens de het maken van een programma worden er tijdelijke, technische metadata aan de content gekoppeld of informatie die bepaalde functionaliteiten beschrijft. Het kan daarbij gaan om editing- en compositiemetadata, zoals een *electronic decision list* (EDL), of om metadata die nodig zijn voor een ordentelijk transport van de content en de metadata over het netwerk, de zgn. 'wrappers'.

Zoals uiteengezet wordt beschrijven als integraal onderdeel van de digitale workflow een gemeenschappelijke, organisatiebrede activiteit die tijdens het hele productieproces wordt uitgevoerd. In de digitale netwerk omgeving zijn het dan ook niet langer alleen de professionele documentaristen van het archief die zich hiermee bezighouden. Dit betekent dat bepaalde *informatieskills* nu over de hele breedte van het productieproces aanwezig moeten zijn. Soms zal dit problemen geven in de uitvoer. Informatieprofessionals zijn zich van oudsher bewust van de waarde van nauwkeurige en geüniformeerde gegevens; een programmamaker of technicus is vanzelfsprekend meer gericht op het bewerken van de onder handen productie dan op het consciëntieus invoeren van gestandaardiseerde metadata.

Gevolg is hoe dan ook dat – in het geval van een (archief)systeem waar items en producties na afmonteren en publicatie worden opgenomen – professionele documentaristen de beschrijving niet langer vanaf nul hoeven te beginnen. Het zijn extern gecreëerde metadata

die de basis vormen waarop zij verder werken. Dit betekent dat er veel controle- en correctiewerk moet worden verzet om de aangeleverde metadata aan te passen aan de 'archivale' kwaliteitseisen. Een positiever effect is dat in routinematige onderdelen van de beschrijving (zoals sommige formele, administratieve, technische en inhoudelijke gegevens) al is voorzien, zodat men zich kan richten op de noodzakelijke verdieping en contextualisering. Het zijn dan de professionele documentalist die – na uitzending of publicatie – zorg dragen voor de uiteindelijke kwaliteit van de organisch gegroeide beschrijving: door het toevoegen van het bredere, historisch perspectief en door het aanbrengen van het semantisch onderscheid tussen informatiecontent, audiovisuele content en stockshots. Deze professionele controle en verrijking garandeert alsnog de toegang aan zoveel mogelijk gebruikersgroepen, ook buiten de productieomgeving.

4 ZELFGENERERENDE TOEGANG

4.1 AUTOMATISCH INDEXEREN

Er wordt steeds meer software ontwikkeld die kenmerken en eigenschappen van gedigitaliseerd audiovisueel materiaal automatisch extraheert en presenteert. Deze software analyseert de beelden en het geluid, waarna de resultaten worden weergegeven in de vorm van metadata, als transcripten, als serie keyframes of als fragmenten (storyboards). De gebruiker kan zijn zoekresultaat ook laten presenteren in een chronologische, geografische of historische context, bijvoorbeeld in de vorm van een grafiek, een tijdlijn of een landkaart. Ging het eerst vooral om automatisch gegenereerde *formele* kenmerken van film en video, zoals kleur, shotwissels en contrast, inmiddels is er software op de markt die de 'president van Amerika' herkent, eigenhandig aangeeft waar hij het over heeft en of hij spreekt op een persconferentie die binnen wordt gehouden of in de openlucht. Films en televisieprogramma's hoeven straks ook niet langer integraal geviewd te worden: ten behoeve van een snelle voorselectie van fragmenten kunnen de meest betekenisvolle delen uit een film of programma als audiovisuele samenvatting aan de gebruiker worden gepresenteerd.

Technologie voor automatische indexerings werkt op basis van algoritmen, rekenprogramma's met instructies voor het uitvoeren van bepaalde taken door een computer. De technologie is gebaseerd op machinaal leren (*machine learning*), een onderzoeksveld binnen de kunstmatige intelligentie. Methodes voor machinaal leren kunnen gecontroleerd zijn of ongecontroleerd. In het eerst geval krijgt het algoritme voorbeelden van invoer en bijbehorende uitvoer. Op basis van deze voorbeelden leert het hoe de eigenschappen van de invoer bepalend zijn voor de uitvoer en kan het zelf voor nieuwe invoer zelfstandig de juiste uitvoer produceren. Bij het ongecontroleerd leren ontdekt het algoritme zelf een structuur in de gegeven invoer. Dit kan bijvoorbeeld door de invoer te verdelen in groepen van kenmerken die op elkaar lijken door ze te clusteren en te classificeren.

Een aantal van de automatische technieken baseert zich uitsluitend op de inhoud van een audiovisuele productie: deze zijn *content-based*, d.w.z. dat ze de informatie analyseren die in het beeld en/of het geluid zélf aanwezig is. Dit zijn bijvoorbeeld beeldherkenning en spraakherkenning. Deze technieken kunnen worden gebruikt voor semantische analyse of voor een kwantitatieve benadering. Technieken die zich richten op taal en tekst – gesproken, in beeld of in geassocieerd tekstmateriaal – vallen onder de noemer *natural language processing* (NLP). Een andere vorm van automatische ontsluiting extraheert informatie die bestaat *buiten* het audiovisuele product. Ze benut tekstmateriaal dat direct of indirect met de productie verbonden is zoals trefwoorden, beschrijvingen en contextinformatie in de netwerkomgeving.

Technieken die content-based zijn onderscheiden klassen van kenmerken die variëren in de mate van semantische inhoud: *low-level features*, *mid-level features* en *high-level features*. Het detecteren en groeperen van de eenvoudige kenmerken vormt vaak de basis voor het extraheren van meer betekenisvolle kenmerken, de *high-level features*. *Low-level features* zijn bijvoorbeeld kleur, vorm of contour, (camera)beweging en textuur. Bij textuur kan

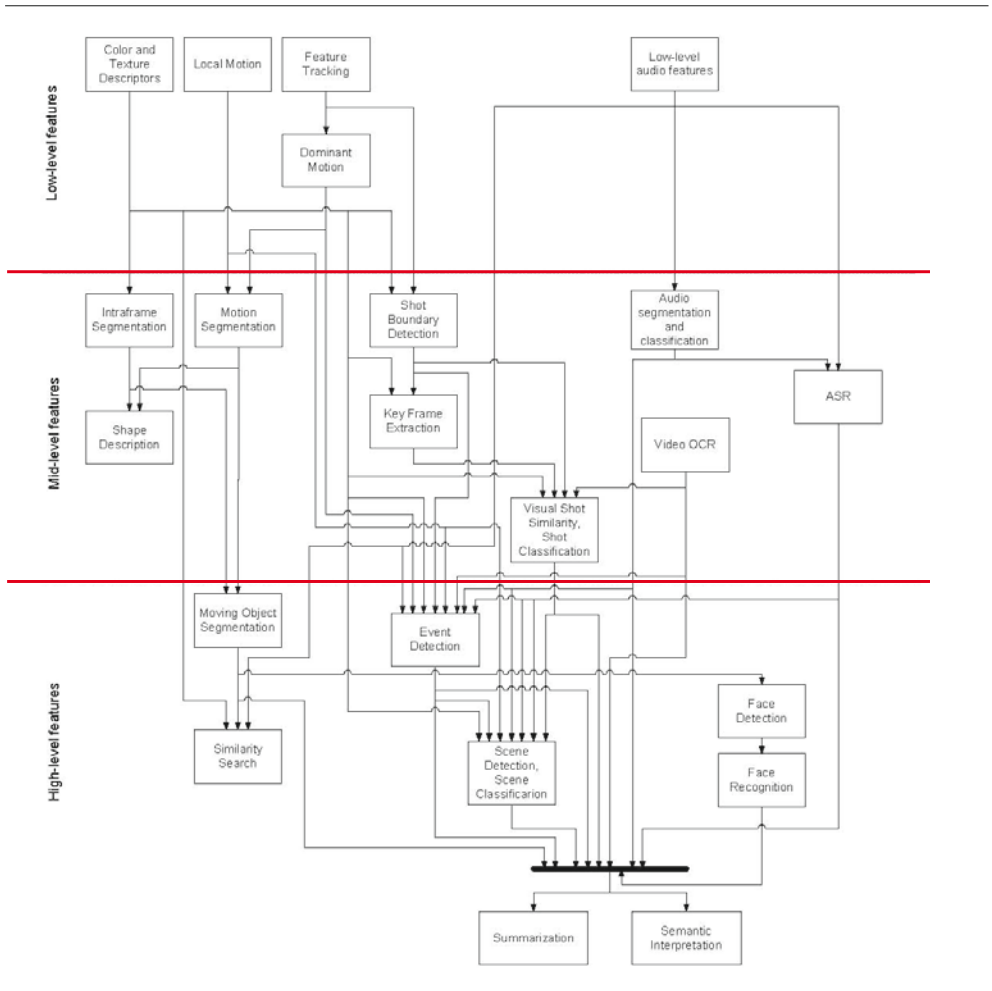


Fig. 6. Visualisatie van de onderlinge afhankelijkheid tussen contentanalyse-tools. Welke lagere kenmerken vormen de basis voor de detectie van de high-level features met meer semantische informatie?

Bron: PrestoSpace

onderscheid worden gemaakt tussen natuurlijke en kunstmatige structuren. Als de textuur 'huid' wordt herkend kan automatisch worden geconcludeerd dat er mensen in beeld zijn. Ziet de computer de textuur van trottoirtegels, dan wordt aangegeven dat het een opname betreft in de gebouwde omgeving. Low-levelkenmerken kunnen ook *direct* (d.w.z. niet via trefwoorden of andere metadata) worden gebruikt voor puur visueel zoeken, het zgn. *query by image*.

4.2 BEELD-, AUDIO- EN TEKSTANALYSE

Videoanalyse gebeurt in drie stappen: het segmenteren van de beelden (ook wel *parsing* genoemd), het classificeren en/of clusteren van de kenmerken en het samenvatten

waarmee de gestructureerde presentatie op het scherm wordt bedoeld. Om de visuele inhoud op te delen kunnen scène- of shotwissels worden gebruikt. Het detecteren van beweging of een bewegende achtergrond wordt ingezet om te bepalen of er een object in het beeld aanwezig is. Beweging vormt samen met kleur de basis van het herkennen van menselijke gezichten. De algoritmen die voor deze toepassing nodig zijn werken uitsluitend onder strenge voorwaarden: een bepaalde belichting bij de opname, er mag niets 'in de weg staan' en het gezicht mag niet al te ver zijwaarts gedraaid zijn. Het maken van modellen voor gezichtsdetectie gebeurt vanuit bestaande kennis van gezichten, vanuit de gezichtskenmerken zelf en vanuit statische analyse van gezichtstrekken. Op dit onderdeel is de laatste tijd veel vooruitgang geboekt, met name met de detectie van zgn. vphoofden. Maar in een audiovisuele collectie van enige omvang is het nog steeds lastig om een verzameling willekeurige gezichten van elkaar te onderscheiden. Er wordt ook nog flink gesleuteld aan herkenningssystemen voor complexe objecten zoals dieren en de meer 'rigide' auto's en vliegtuigen.

Bij audio zijn het de signalen en (technische) waarden en frequenties in de stream die behulpzaam zijn bij stiltdetectie, spraak-muziekclassificatie en het detecteren van stemcomponenten. Door segmentatie wordt het gesproken woord onderscheiden van de andere geluiden en kan het worden omgezet in tekst, de zgn. transcripten. Taaltechnologische analyses (NLP) filteren de semantisch meest 'belangrijke' woorden eruit. Omdat geïndexeerde natuurlijke taal de zoekmogelijkheden enorm vergroot, wordt momenteel gewerkt aan sprekeronafhankelijke spraakherkenners met een onbeperkte vocabulaire. Audio en video hebben een groot aantal kenmerken gemeen. Low-level-kenmerken in de audio worden dus ook bij video gebruikt voor segmentatie, clustering en patroonherkenning (*query by example, sound spotting*). Tijdelijke veranderingen in het geluidsspoor kunnen gemakkelijk gemeten worden. Dat geldt ook voor stiltes en pauzes. En harde geluiden in een film of televisieprogramma impliceren doorgaans een verhoging van de emoties en daarmee een wijziging in de semantische inhoud.

Taaltechnologische technieken (NLP) kunnen worden ingezet om tekstelementen in het beeld te analyseren. Het gaat hier om zgn. *printed characters* zoals ondertitels en tekstblokjes, maar ook om bijvoorbeeld opschriften en uithangborden die in de opname zélf voorkomen. Ook tekstvormen die als signaal met het beeld en geluid worden meegestuurd, zoals ondertitels, teletekst en grafische elementen als logo's en statistieken vallen binnen deze categorie. Een wijziging of verschuiving van deze elementen in het beeld is meestal een indicatie van een verandering in de semantische inhoud. De tekstuele elementen hebben een link naar de temporele structuur van de video en kunnen dus worden omgezet naar een tijdcode-index die kan worden gebruikt bij het zoeken. Nadat de teksten en grafische vormen door het systeem zijn gedetecteerd wordt een *optical character recognition* (OCR)-toepassing ingezet waarna ze als doorzoekbare *full text* aan de content worden toegevoegd. Het is mogelijk om deze teksten te koppelen aan thesauri of andere gecontroleerde woordsystemen. Als er een vertaalfunctionaliteit is ingebouwd kan dat ook in een andere taal (*cross-language retrieval*). Met behulp van taaltechnologie kunnen ook geassocieerde teksten in de netwerkomgeving worden geanalyseerd wanneer die eenmaal zijn getraceerd en gekoppeld aan het te ontsluiten materiaal. Contextinformatie rond de video- of audioproductie (zoals websites, persberichten, recensies, logboeken, presentatieteksten) kan aldus worden gebruikt als input voor de uiteindelijke 'beschrijving' van een audiovisueel product.

4.2.1 GECOMBINEERDE TECHNIEKEN

Door low- en mid-levelkenmerken in video en audio te combineren kunnen sommige *genres* automatisch worden herkend. De opbouw van bepaalde televisieprogramma's is erop gericht onderliggende narraties en boodschappen over te brengen. Vaak gaat het om vaste formats waarin het materiaal op een standaardmanier is gerangschikt. Het succes van genredetectie hangt af van de mate waarin audio- en videokekenmerken corresponderen met de inhoud van de opgenomen gebeurtenis. In het algemeen geldt dat vooral bij genres met een duidelijke relatie tussen de low-level features (kleur, beweging, achtergrond) en de inhoud, de resultaten van classificatie goed genoeg zijn om in de praktijk toe te passen. Genredetectie lijkt daarom vooral perspectieven te bieden voor (televisie)programma's met eenvoudig herkenbare formats, zoals nieuwsbulletins en weerberichten, sommige sportprogramma's en andere uitzendingen met een vaste opbouw (shows, quizzen e.d.).

A-priorikennis over een bepaald genre kan worden gebruikt voor het lokaliseren van fragmenten met een relevante actie, zoals bijvoorbeeld een goal in een voetbalwedstrijd. Bij veldsporten is er altijd (veel) gras te zien, lijnen en hoeken van het veld, close-ups van spelers en opnamen van bewegingsacties en menigten. Bij dit soort wedstrijden volgt de camera ook bijna altijd de beweging van de bal en is er veel dynamiek in de audio (stem-volume commentator en gejuich, applaus en boegeroep van het publiek). Als al deze kennis wordt gecombineerd, kan worden gedetecteerd wanneer er een goal wordt gescoord of een andere significante actie plaatsvindt. In speelfilms en dramaproducties zijn er groepen te



Fig. 7. Zoekresultaat gebaseerd op de gecombineerde toepassing van beeldherkenning en spraak- en taaltechnologie. Bron: Informedia

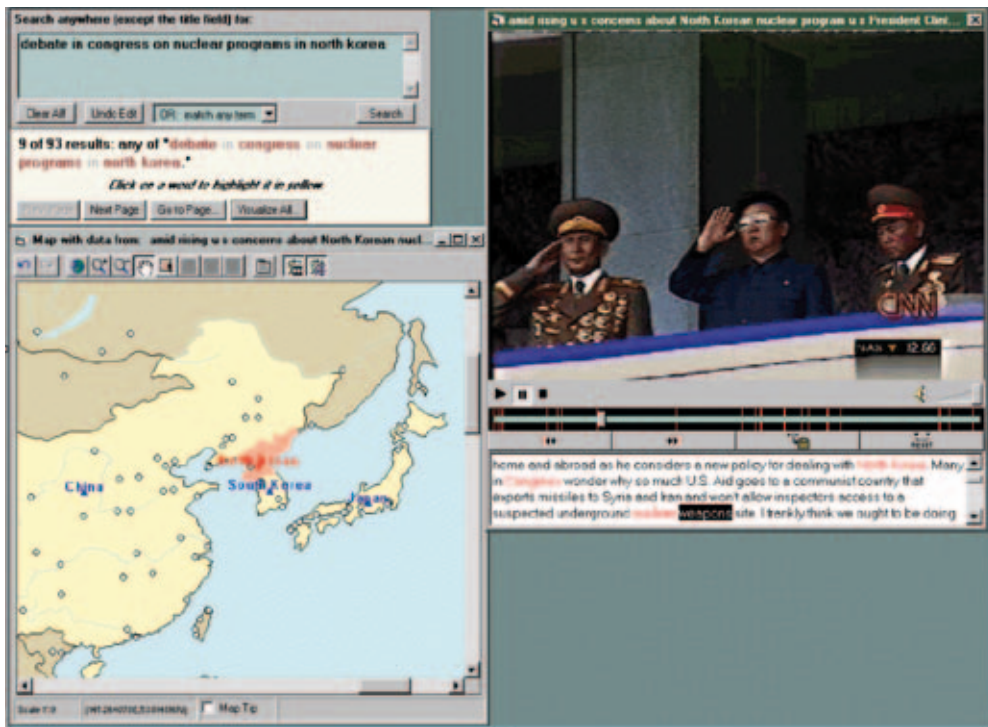


Fig. 8. Automatisch gegenereerde contextinformatie: onderdeel van het zoekresultaat is een kaart van het geografische gebied dat in de zoekvraag werd genoemd. Bron: Informedia

maken van dialogen met actie, dialogen zonder actie en actie zonder dialoog. Aan de hand van shotlengte, bewegingsintensiteit, volume, lichtmetingen en spraak en stilte identificeert het algoritme bijvoorbeeld actie, drama en dialoog. Het vinden van de meest grappige scènes in een comedy of de meest enge scènes in een horrorfilm is een verfijning van deze techniek en heet affectieve contentanalyse. Het resultaat is uiteraard subjectief, maar is wel gebaseerd op de algemene ervaring, de verwachting van de maker en typische reacties van het publiek.

Een andere geavanceerde combinatie van meerdere technieken, toegepast op low- en high-levelkenmerken, is de zgn. *visual summary*, een automatisch gegenereerde weergave in (bewegend) beeld en geluid van de semantisch meest belangrijke momenten van een programma. Middels het zgn. *skimmen* kan de essentie van de visuele inhoud in een fractie van de normale speelduur worden overgebracht. Gecombineerde taal-, audio- en beeldherkenningstechnieken genereren hierbij een synopsis van het origineel, in feite een soort trailer. Voor het audiodeel worden de gebruikte woorden 'gewogen' om de meest relevante termen te selecteren en deze als doorzoekbare index bij de content op te nemen. Ook andere aanwijzingen, zoals overgangen tussen sprekers en items, worden gedetecteerd. Het videodeel wordt gescreend en geanalyseerd op scènewisselingen en op relevante objecten en beweging. Wanneer de video aldus is gesegmenteerd, wordt statistisch het relatieve belang van elk stukje content uitgerekend. Levendige (en dus veelzeggende, informatieve) scènes worden eruit gelicht aan de hand van optische analyse

van camerabeweging en de aanwezigheid van bewegende objecten. Gezichten, tekst en grafische elementen worden geïdentificeerd en gebruikt als basis voor de samenvatting. De uiteindelijke representatie van het document kan door de gebruiker zelf worden ingesteld en indien gewenst variëren in lengte en soort content. Het belangrijkste probleem schuilt in de selectie van het materiaal voor de samenvatting. Er moet rekening worden gehouden met het proportioneel representeren van de originele content. Tegelijk moeten de meest opvallende *events* eruit gelicht worden. Het kan echter zijn dat deze maar een klein deel uitmaken van de gehele inhoud. Uiteraard is het niet alleen lastig, maar ook heel subjectief om de menselijke cognitie te vertalen naar een geautomatiseerd abstractieproces. Niettemin zijn er – vooral bij gelimiteerde hoeveelheden documentair materiaal met expliciete spraak- en tekstelementen – met deze techniek al indrukwekkende resultaten geboekt.

Op basis van low-level en mid-levelkenmerken zoals kleur, camerabeweging, textuur en geluidselementen kan van audiovisueel materiaal ook een zgn. *fingerprint* worden gemaakt. Deze fingerprints fungeren voor elke individuele sequentie beelden als een soort DNA waarmee digitale items door alle bestanden, kanalen en systemen heen kunnen worden geïdentificeerd. De techniek werkt niet met metadata die in het materiaal worden geïntegreerd, zoals bij het zgn. digitale watermerken. Identificatie gebeurt door de intrinsieke beeldkenmerken van bepaalde shots te vergelijken met de fingerprints zoals die eerder in een database zijn opgeslagen. Zo kan bijvoorbeeld worden achterhaald wie de auteursrechten bezit van een bepaald fragment. Fingerprinting kan ook van dienst zijn bij het traceren van de herkomst van archieffragmenten in nieuwe producties, de zgn. *dupes*.

4.2.2 DATAMINING

Bovenbeschreven toepassingen van beeld- en spraaktechnologie zijn voornamelijk gericht op het indexeren van de semantiek van audiovisueel materiaal, i.c. op het creëren van toegang tot de inhoud van individuele fragmenten of programma's. Dezelfde technologie kan ook worden ingezet voor het maken van kwantitatieve analyses van complete archieven of collecties met als doel trends en relaties bloot te leggen: *datamining*. Datamining werd tot voor kort vooral gebruikt in bepaalde (bèta)wetenschappen waar men uit grote hoeveelheden onderzoeksmateriaal bijvoorbeeld economische trends naar voren haalt of klimaatanalyses maakt. Ook in marketing en andere commerciële domeinen werkt men veel met deze statistische methoden, teneinde consumentengedrag te kunnen analyseren en voorspellen. De techniek wordt inmiddels meer en meer ingezet voor de analyse van digitale data uit de beeldende kunst, de internetcultuur en de film-, foto- en televisiegeschiedenis. Een relatief jonge toepassing van datamining is het genereren en presenteren van (nieuwe) contextuele kennis over culturele artefacten in de vorm van historische patronen over lange tijdsperiodes. Deze kennis wordt inzichtelijk gemaakt door middel van geavanceerde grafische visualisaties. In het geval van digitale audiovisuele materialen meet men de visuele structuur aan de hand van formele low-levelkenmerken zoals bijvoorbeeld kleurdistributie, shotwissels, montage, snelheid en camerabeweging. Voor het uiteindelijke resultaat worden deze kenmerken gecombineerd met de metadata die bij het materiaal horen. Het is ook mogelijk alléén de bestaande metadata over een bepaalde collectie of groep programma's statistisch te meten en daar patronen uit te destilleren. Kwantitatieve analyses kunnen plaatsvinden vanuit zowel de catalogus als de collectie om vervolgens grafisch te worden weergegeven per periode, per producent, per genre, per regio, per doelgroep of per soort gebruik en gebruiker. Voorbeelden van toepassingen zijn: de ontwikkeling van bepaalde genres; het

CULTURELE ANALYTIEK

Lev Manovich, schrijver, denker en hoogleraar op het gebied nieuwe media, digitale cultuur en visuele kunsten is oprichter van het Software Studies Initiative, verbonden aan de Universiteit van Californië in San Diego. In dit laboratorium worden computergebaseerde technieken ingezet voor datamining: kwantitatieve analyse en interactieve visualisatie van culturele digitale data. Men werkt in dit researchdomein – door Lev Manovich inmiddels ‘cultural analytics’ gedoopt – met eigentijdse (online)artefacten van de visuele en mediacultuur zoals video, audio, games, visuele kunst, mediadesign, film, animaties, foto’s etc. Om trends en ontwikkelingen aan te kunnen tonen over langere tijdspannen richt de analyse zich op omvangrijke datasets: hele verzamelingen of substantiële delen daarvan. Het is daarbij noodzakelijk dat de te analyseren onderdelen op de een of andere manier samenhangen en een bepaalde homogeniteit vertonen, zoals bijvoorbeeld bij meer exemplaren van één genre het geval is. Meting van de visuele structuur van audiovisuele content gebeurt op basis van een aantal formele low- level-kenmerken zoals bijvoorbeeld kleurdistributie, shotwissels, montage, snelheid en camerabeweging. Men combineert deze kenmerken met metadata over het materiaal die hetzij handmatig, hetzij automatisch zijn gecreëerd . De resultaten worden weergegeven in indrukwekkende grafische visualisaties. In het Software Studies Lab wordt momenteel – als onderdeel van een pilot met Beeld en Geluid – kwantitatief onderzoek gedaan naar vormkenmerken van leaders en stationcalls van Nederlandse omroepen in de periode 1960-1980.

lab.softwarestudies.com/

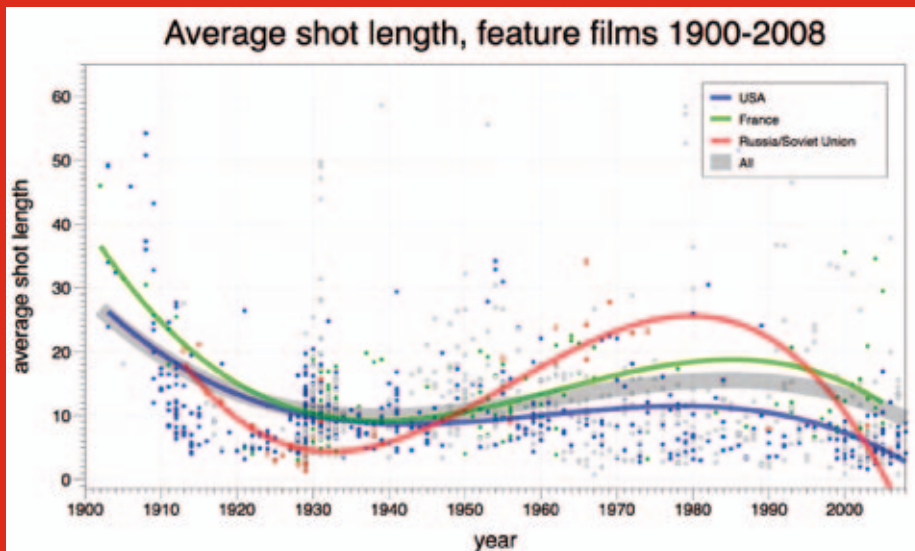


Fig. 9. Resultaat van datamining toegepast op een collectie speelfilms geproduceerd in drie verschillende landen over een periode van meer dan een eeuw. Bron: Software Studies Lab

visualiseren van vormgeving, snelheid en presentatie van programma's; het plaatsen van de receptie in een historisch perspectief; de visualisatie van aantallen en hoeveelheden in collecties en catalogi etc.

Datamining is niet direct gericht op het ontsluiten van de semantische inhoud van afzonderlijke items. Gebruikers van de resultaten van audiovisuele datamining moeten dus niet zozeer worden gezocht in kringen van de programmamakers, producenten en anderen die fragmenten willen hergebruiken. De analyses en visualisaties van culturele en historische trends zijn vooral van belang voor mediawetenschappers, historici, culturele en sociale wetenschappers en het algemeen publiek. Met andere woorden: doelgroepen die geïnteresseerd zijn in de (langetermijn)ontwikkelingen en effecten van vorm en/of inhoud van televisie, film en internet.

4.3 HET SEMANTISCH WEB

Een heel ander type research richt zich op het ontwikkelen van methoden en technieken die gebruik kunnen maken van beschrijvende metadata die al aanwezig zijn op het web. Hier gaat het opnieuw *niet* om de analyse van beeld en geluid binnen één programma of groep programma's. Doel is het creëren van uniforme toegang tot meerdere collecties: semantische interoperabiliteit. Interoperabiliteit is het vermogen om vanuit verschillende invalshoeken te communiceren, waarbij de partijen blijven uitgaan van hun eigen systemen. Technische interoperabiliteit gaat over het koppelen van informatiesystemen; over interfaces en andere technische standaarden. Semantische interoperabiliteit richt zich op het geven van een gedeelde *betekenis* aan de uit te wisselen gegevens. Informatie in een netwerk wordt daarbij op zo'n manier gerepresenteerd dat niet alleen mensen maar ook computers haar kunnen verwerken, linken en interpreteren. Het zgn. *semantisch web* is een initiatief van de standaardorganisatie W3C dat erop gericht is om het world wide web uit te breiden met deze technologie. Het concept gaat ervan uit dat alle informatie (personen, plaatsen, objecten, zaken en gebeurtenissen) op het web op een betekenisvolle manier geïdentificeerd kan worden. De gegevens zijn daarbij uitbreidbaar zonder dat ze aan de bron veranderen. Op deze manier kunnen verschillende gemeenschappen onafhankelijk van elkaar gegevens uitwisselen.

Om semantische interoperabiliteit te bereiken wordt de kennis en structuur benut van de geüniformeerde metadata zoals die zijn vastgelegd in bestaande trefwoordenlijsten, thesauri en taxonomieën, in gebruik bij bijvoorbeeld archieven en bibliotheken. Voor het louter uitwisselen tussen informatiesystemen kunnen deze woordsystemen worden beschreven in een standaardtekstformaat: XML (Extensible Markup Language). Om termen te kunnen toevoegen en/of om ze op een zinvolle manier te kunnen koppelen aan gegevens in andere bestanden moeten computers ook de *betekenis* van woorden en begrippen kunnen begrijpen. Hiervoor is RDF (Resource Description Framework) ontwikkeld. RDF voegt eigenschappen toe aan de termen. Een voorbeeld: een maker is een persoon, een productie heeft een maker en een productie kan worden gepubliceerd. De volgende stap naar semantische interoperabiliteit bestaat uit het modelleren van de termen en hun eigenschappen in een semantisch model dat ook alle *relaties* tussen die eigenschappen beschrijft: een ontologie. Ontologieën worden beschreven in OWL (Web Ontology Language), een standaard die is ontwikkeld om de samenhang tussen begrippen en onderwerpen te beschrijven. XML,



Fig. 10. Semantische interoperabiliteit: gegevens over objecten uit verspreide collecties zijn benaderbaar via één interface. Hiertoe zijn de metadata uit verschillende woordsystemen met elkaar gemapt.
<http://e-culture.multimedial.nl/demo/session/search>

RDF en ontologieën in OWL staan los van alle (lokale, organisatiegebonden) methoden die bij archieven en andere instellingen gebruikt worden om gegevens geüniformeerd vast te leggen.

Het inzetten van RDF en OWL biedt mogelijkheden om via intelligent redeneren nieuwe kennis te destilleren uit bestaande gegevens en die als zoekresultaat aan te bieden. Voortbouwend op bovenstaand voorbeeld kunnen de termen ‘programma’, ‘onderwerp’, ‘maker’ en ‘persoon’ zijn gemodelleerd in een ontologie van audiovisuele begrippen. Daarin is ook vastgelegd dat een programma altijd een onderwerp heeft, een maker een persoon is en ieder programma een maker heeft. Uit deze gestructureerde begrippen, eigenschappen en relaties kan naar aanleiding van een zoekvraag een nieuw resultaat worden gedestilleerd, bijvoorbeeld over andere programma’s die een bepaalde maker heeft gemaakt of over meer programma’s met eenzelfde onderwerp. Op deze manier ontstaan – geheel automatisch – nieuwe, betekenisvolle gegevens. Veel zoekmachines kunnen dit niet bieden, omdat daarmee alleen de metadata worden gevonden die exact in de zoekvraag zijn ingevoerd.

Wanneer het intelligent redeneren wordt gecombineerd met taaltechnologie kunnen

SEMANTISCH ANNOTEREN

Doel van het researchproject CHOICE (Charting the Information Landscape Using Context Information) – uitgevoerd bij Beeld en Geluid door het Max Planck Instituut, de VU en Telematica – was het kunnen doen van behulpzame suggesties voor trefwoorden aan documentalisten tijdens het catalogiseren. De trefwoordsuggesties worden met behulp van natural language processing (NLP) en semantische webtechnieken ontleend aan bestaande contextuele tekstbronnen over radio- en televisieprogramma's zoals websites, onlinetelevisiegidsen, recensies en kijkcijferonderzoek. De methode werkt als volgt: algoritme no. 1 detecteert die segmenten van contextdocumenten in de netwerk omgeving die mogelijk relevante termen zouden kunnen bevatten over een bepaalde te beschrijven productie. Deze termen worden gelinkt aan de trefwoorden in de Gemeenschappelijke Thesaurus Audiovisuele Archieven (GTAA). Algoritme no. 2 gebruikt vervolgens de relaties tussen de gevonden trefwoorden om semantische clusters te maken, deze te wegen en de meest centrale termen als suggestie te presenteren aan de documentalist. Deze kiest op basis van deze automatisch gegenereerde, semantische annotatie de juiste thesaurusterm. In het project werd gekozen voor een testset van nieuws- en actualiteitenprogramma's en documentaires. Dit omdat er juist rond dit soort programma's veel digitale contextbronnen voorhanden zijn in de netwerk omgeving. Experimenten met de CHOICE-methode hebben inmiddels uitgewezen dat ongeveer de helft van de eerste 10 suggesties een goede is. Deze 10 bestreken 65% van de trefwoorden die ook door documentalisten zouden zijn toegekend.

<http://ems01.mpi.nl/CHOICE/>

metadata ook automatisch worden toegevoegd aan documenten of audiovisuele materialen. Er ontstaat dan een zgn. 'semantische annotatie' met een inhoudssamenvatting van de film of video in een voor zoekmachines begrijpelijke taal. In feite zijn deze semantische annotaties te vergelijken met een door een documentalist gemaakte beschrijving, in die zin dat ze bestaande kennis gebruiken om andere, nieuwe kennis te genereren. Kennis van de wereld en kennis van de context van een bepaald audiovisueel programma die elders in de netwerk omgeving bestaat, wordt opgespoord en via een ontologie en taaltechnologie omgezet naar een beschrijving in de eigen catalogus van een instelling.

4.4 DE (ON)MOGELIJKHEDEN VAN SOFTWARE

Automatische technieken creëren andersoortige ingangen waarmee op audiovisueel materiaal kan worden gezocht. Gebruikers krijgen toegang via objectieve identificatie en groepering van (formele) beeld- en geluidskwaliteiten. De ontsluiting van audiovisuele producties wordt hierdoor niet alleen meer bepaald door selectie en (subjectieve)

DE MULTIMODALE AANPAK

MUNCH (Multimedia Analysis for Cultural Heritage) is een researchproject van de UvA, de VU en Beeld en Geluid dat zich richt op de automatische analyse van digitale bewegende beelden. Er is gekozen voor een multimodale benadering: aangestuurd door een ontologie worden beeldeigenschappen geanalyseerd samen met de (geschreven en gesproken) taal die dat beeld beschrijft. Hiertoe is een multimediathesaurus gebouwd. Deze bestaat uit een set high-leveldetectoren in een onderlinge structuur. Met deze detectoren kunnen automatisch beeldconcepten worden opgespoord, zoals 'opstijgende vliegtuigen', 'George Bush sr.' of 'de Amerikaanse vlag'. De detectoren worden getraind door de automatische methode te combineren met handmatig beschreven beelden uit het televisienieuws. Dit lexicon is verrijkt met semantische beschrijvingen en structuur, verkregen vanuit het onlinewoordbestand WordNet. Ook de termen en relaties binnen de Gemeenschappelijke Thesaurus Audiovisuele Archieven (GTAA) van Beeld en Geluid en het Filmmuseum worden ingezet. Om het zoeken in transcripten van het gesproken woord te verbeteren experimenteert MUNCH met het gebruik van taaltechnologie. Men werkt daarbij aan het probleem van de temporele mismatch tussen het noemen van objecten in spraak en hun verschijning in beeld. Er is een oplossing gevonden door de spraak die hoort bij shots vóór en na het shot met het betreffende object mee te nemen in de analyse. Deze shots worden gewogen: hoe verder weg van het 'hoofdshot', hoe lager het gewicht. Het betrekken van de spraak in de naburige shots verbetert het zoekresultaat aanzienlijk.

<http://ilps.science.uva.nl/munch/>

interpretatie vooraf. *Content-based retrieval* genereert resultaten met een hoge formele precisie, zoals exact getimed shots en scènes, bepaalde vormkenmerken en verschillende shottypen. Vanwege de tijdrovendheid werd juist dit type gegevens allang niet meer handmatig gegenereerd. Deze techniek kan dus van belang zijn bij het ontsluiten van sommige typen stockshots. Er worden daarbij ook nog eens low- en mid-level-elementen geïndexeerd die feitelijk nooit beschreven werden, zoals kleur, textuur, (camera)beweging, gesproken woorden in transcripten en geluidskenmerken. Al deze ingangen bij elkaar gaan zorgen voor nieuwe, creatieve zoekopties.

Content-based retrieval heeft wat betreft de hogere semantische niveaus (i.c. de inhoud en interpretatie) nog flinke beperkingen. Alleen tekst in beeld geeft een min of meer accuraat beeld van wat er te zien is en waar een programma of fragment over gaat. Veel semantische informatie zal dan ook nog handmatig moeten worden toegevoegd. De (on)mogelijkheden van automatische indexering kunnen echter niet kunnen worden teruggebracht tot een soort vorm-inhoudtegenstelling, waarbij low-levelvormkenmerken door de software, en betekenisvolle informatie door mensen wordt ingevoerd. Daarvoor is de relatie vorm en

inhoud hier te complex. Zoals de zaken er nu voor staan zal de computer én (intelligente) hulp kunnen verlenen bij de manuele annotatie én (beperkt) kunnen interpreteren. Het helpt in alle gevallen wanneer beeld en geluid dezelfde boodschap overbrengen, zoals bij het gros van de nieuws- en sportfragmenten. Om collectiebreed te kunnen worden toegepast zullen er veel soorten algoritmen nodig zijn. Modellen die worden ontwikkeld voor de televisiebeelden van nu kunnen immers niet straffeloos worden toegepast op historisch materiaal. Film- en opnameconventies verschillen bijna per decennium en ook *wat* er gefilmd wordt (objecten, automerken, gezichten, gebouwen) is steeds weer anders.

Met behulp van datamining kan nieuwe informatie worden gegenereerd die zonder die techniek niet gevonden zou zijn. Kwantitatieve analyses maken het eenvoudiger om trends en patronen naar de oppervlakte te brengen in de productie, de vormgeving, de inhoud en het gebruik van film en televisie op basis van grote hoeveelheden materiaal, over zeer lange tijdsperiodes. Filmanalytisch en mediahistorisch onderzoek kunnen van deze mogelijkheden enorm profiteren. Datamining vormt daarbij een mogelijk antwoord op de toenemende digitale productie, of het nu gaat om *digital born*-materiaal of om gedigitaliseerde objecten, om collecties van archieven of om voortbrengselen van privépersonen. Zonder de hulp van kwantitatieve technieken zijn deze gigantische heterogene hoeveelheden immers niet meer in kaart te brengen.

Een groeiende rol is er ook voor de semantisch-webtechnieken. Op het niveau van collecties en instellingen zal het semantisch web de komende jaren veel nieuwe toegangsmogelijkheden genereren. Door het koppelen van ontologieën uit verschillende domeinen zal uiteindelijk een zeer grote kennisbank worden gecreëerd, een web van gekoppelde, uitbreidbare metadata tussen vele culturele collecties en archieven. Omdat deze technieken gebruikmaken van bestaande, gestructureerde metadata kan deze benadering zowel de precisie als de relevantie van zoekresultaten verbeteren. De formele kenmerken van audiovisueel materiaal zoals gegenereerd vanuit beeld- en spraaktechnologie kunnen zo worden ingebed of gecontextualiseerd. Het combineren van deze content-based technieken met het redeneren vanuit ontologieën lijkt tegemoet te komen aan de complexe, gelaagde semantiek van een audiovisuele productie. Cees Snoek, onderzoeker aan de Universiteit van Amsterdam, bedacht in dit verband het concept van de *semantic pathfinder*: in de eerste fase van de automatische zoekactie, de contentanalyse, worden de low-levelkenmerken in de video, de audio en de tekst, gebruikt; de tweede stap, de *style analysis stage*, is er om het 'concept' van een beeld te vangen door het detecteren van gezichten, locaties, objecten en beweging. Voor de laatste fase, de semantische contextanalyse, wordt een gecontroleerd woordsysteem ingezet. In een ononderbroken *semantic path* worden in één actie alle bronnen aangeboord: zowel de kenmerken van de video en audio als de woorden in het geluidsspoor, de handmatig toegevoegde trefwoorden en annotaties alsmede de contextinformatie in de netwerkomgeving.

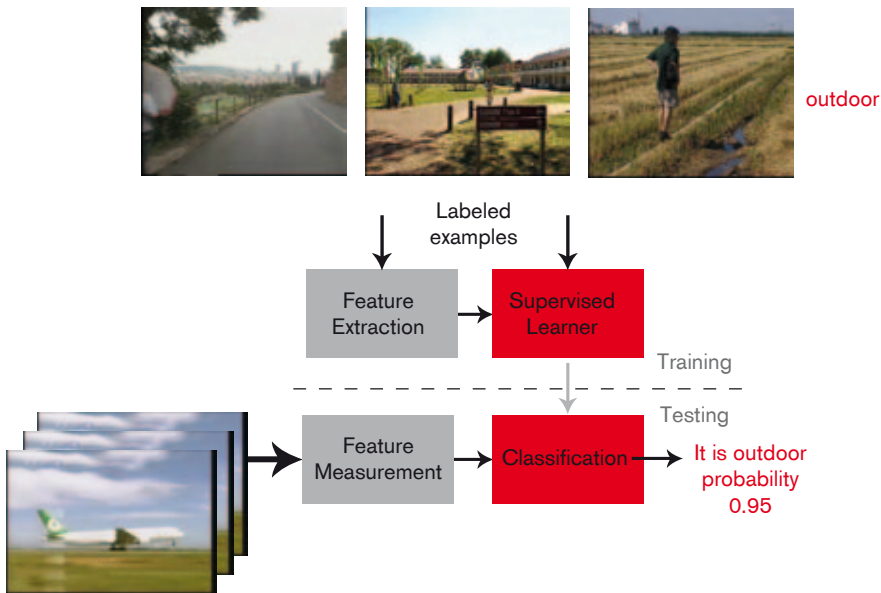


Fig. 11. Het detecteren van één bepaald concept (in dit geval het concept 'outdoor' ofwel buiten) en voorbeelden van andere concepten die momenteel automatisch kunnen worden gegenereerd. Er wordt hierbij gebruikgemaakt van beeldanalyse gecombineerd met semantische technieken. Bron: Cees Snoek, Universiteit van Amsterdam <http://staff.science.uva.nl/~cgmsnoek/index.php/demonstrations/semantic-pathfinder/>

5 DE GEBRUIKER BESCHRIJFT

5.1 VAN CONSUMENT NAAR PRODUCTENT

Eerder is beschreven hoe deelnemers aan een audiovisueel (omroep)productieproces langs een digitale workflow metadata invoeren over de programma's waar ze aan werken. Er is ook uiteengezet welke kenmerken van beeld en geluid op een automatische manier kunnen worden geëxtraheerd om te worden omgezet in beschrijvende informatie. Er is nog een derde nieuwe manier in opkomst waarop audiovisueel materiaal wordt ontsloten. Ook hierbij speelt de netwerkomgeving – i.c. het feit dat audiovisueel materiaal digitaal is en online beschikbaar wordt gesteld – een belangrijke rol. Het gaat hier om het beschrijven van fragmenten door groepen mensen die van oudsher behoren tot de afnemers van audiovisuele collecties: door gebruikers. Deze groepen veranderen in de netwerkomgeving van rol. Van consument worden ze producent, in dit geval van beschrijvende metadata.

5.1.1 EXPERTBESCHRIJVINGEN

Audiovisueel materiaal kan op veel manieren worden gebruikt en hergebruikt. In de beschrijving van professionele documentaristen wordt daar zoveel mogelijk rekening mee gehouden. Allereerst is er de algemene informatie over de inhoud en makers van een programma die moet worden opgetekend. Vervolgens worden die delen van een programma beschreven, die mogelijk geschikt zijn voor hergebruik in een productiecontext. Maar dit zijn niet de enige invalshoeken vanwaaruit fragmenten en programma's kunnen worden



The screenshot shows the Teleblik website interface. At the top, there is a navigation bar with 'Home', 'Mijn media', 'Nieuws', and 'Forum'. A search bar contains the text 'zoek een uitzending' and a 'zoeken' button. Below the navigation bar, there is a section titled 'Uitzending bekijken' with the subtitle 'Egel op jacht (expr_id: 395...)'. A video player shows a hedgehog in a grassy field. To the right of the video player is a 'Details' section. The details include: 'Lengte: 01:44', 'Teleblik nummer: 1203420', 'Onderwerp: Egel, insect, jachtgedrag, panter, prooi, reserve, winterslag', 'Waardering: ★★★★★ (2 stemmen)', 'Vakgebied: Kunst en cultuur (1), Mens en natuur (1)', and 'Niveau: peuters (1), primaire onderwijs (2)'. At the bottom of the video player, there are two buttons: '+ toevoegen aan mijn media' and '< fragment maken ...'.

Fig. 12. Voorbeeld van een 'expertbeschrijving' van een audiovisueel fragment. In het project Teleblik van Beeld en Geluid worden vorm en inhoud van de beschrijving op professioneel-didactische wijze toegesneden op de informatiebehoefte van scholieren. www.teleblik.nl/

ontsloten. Zo is er inmiddels de niet meer weg te denken inzet van audiovisueel materiaal voor onderwijsdoeleinden waarbij beeld en geluid dient om lesprogramma's te illustreren. Audiovisueel materiaal is ook steeds meer object geworden van cultuurwetenschappelijk en mediahistorisch onderzoek.

Voor het maken van beschrijvingen voor dit soort doeleinden zoeken collectiehouders vaak domeindeskundigen aan. Deze beschrijven het materiaal vanuit de optiek van specifieke doelgroepen, i.c. vanuit vooronderstellingen over de wijze van raadpleging en hergebruik. De trefwoorden die worden toegekend zijn dan bijvoorbeeld gerelateerd aan onderwijscurricula en samengesteld vanuit gangbare woordsystemen uit de educatieve wereld. Een 'mediahistorische' beschrijving kan meer zijn gericht op contextinformatie: zo kunnen er metadata worden toegevoegd over de receptie van een productie, over de kijkcijfers, prijzen die het programma heeft gehad en over de maatschappelijke effecten. In deze gevallen gaat het dus niet in de eerste plaats om het maken van een neutrale beschrijving van de informatiecontent en de audiovisuele content zoals dat wordt gedaan ten behoeve van raadpleging en hergebruik in de productiewereld.

Een andere vorm van annotatie door expertgebruikers vindt plaats binnen zgn. *peer-to-peer* netwerken. De groep mensen die – vanaf hun eigen pc – (audiovisuele) informatie verzamelt en opslaat over lokale en nationale geschiedenis en erfgoed is inmiddels zeer groot. Binnen *peer-to-peer* netwerken wordt het verzamelde materiaal van deze privépersonen verbonden aan grote culturele instituten. De kennis die deze groep heeft over de inhoud van bepaalde audiovisuele bronnen wordt algemeen toegankelijk gemaakt door ze te koppelen aan de officiële catalogus, website en/of de inventaris van een archief of collectiebeheerder.

Annotaties door experts worden gemaakt buiten de archivale omgeving maar staan doorgaans wel onder controle van het archief of de collectiebeherende instelling. Deze is dan verantwoordelijk voor het organiseren van de workflow van en naar de domeindeskundigen, voor de correctie en voor het inpassen in de bestaande catalogi en informatiebestanden.

5.1.2 CROWDSOURCING

"It's clear that the Web is structurally congenial to the wisdom of crowds," vindt James Surowiecki, auteur van het boek *The Wisdom of Crowds*. Surowiecki meent dat de kwaliteit van informatie die door miljoenen mensen wordt opgeschreven, gecontroleerd en herschreven zich zonder meer kan meten met de inhoud van de Encyclopaedia Britannica. *Crowdsourcing* zou omschreven kunnen worden als een neologisme voor een traditioneel door professionele werknemers verrichte taak, die wordt overgelaten aan een grote groep mensen, na een open oproep om bijdragen. De taak waar het om gaat kan liggen op het gebied van consultancy, (software)ontwikkeling en onderzoek. Het ontsluiten van data kan ook verlopen via crowdsourcing. Het gaat in alle gevallen om het inzetten van kennis en expertise die bestaat bij een publiek van vrijwilligers en geïnteresseerden.

Ook voor het ontsluiten van digitale erfgoedcollecties wordt inmiddels gebruikgemaakt van crowdsourcing. Vooral fotocollecties lijken zich hier goed voor te lenen. Een van de eerste Nederlandse voorbeelden was het initiatief 'Zoekplaatje' van het Gemeentearchief Den Haag. Gebruikers wordt hier gevraagd behulpzaam te zijn bij het identificeren van de gebeurtenissen en de personen op de foto's en bij het verschaffen van contextuele details

zoals straatnamen en namen van gebouwen. Sommige andere websites bieden gebruikers de mogelijkheid om ook eigen materiaal te uploaden. Behalve foto's komen er ook steeds meer beeld- en geluidsfragmenten online beschikbaar om door gebruikers te worden geïdentificeerd, becommentarieerd en gewaardeerd. Het aanleveren van gegevens gebeurt via e-mail, met onlineformulieren of door het rechtstreeks werken in de database.

Kenmerk van deze vormen van gebruikersannotatie is dat het archief of de collectiehouder de coördinatie, de regelgeving en de eindcontrole zelf in handen houdt. Wat betreft de presentatie van de ingebrachte informatie wordt in alle gevallen een duidelijke scheiding aangebracht tussen wat gebruikers aanleveren en de officiële informatie van het archief of de collectiebeheerder. De kennis van de gebruikers fungeert weliswaar als belangrijke bron, maar de uiteindelijke selectie en de vorm waarin de gegevens eventueel deel gaan uitmaken van de 'officiële' catalogus, blijft strak geregisseerd door de professionals.

Haags Zoekplaatje 51



Deze foto is afkomstig van H.J. Tinnenbroek (1880-1982), die zelf rechts op de traprede staat. De foto van rond 1910 is door zijn kleindochter, mevrouw Schoonens-de Gans, aan het Haags Gemeentearchief geschonken. De heer Tinnenbroek vertelde haar dat het aan de Groot Hertoginnelaan was. De deurpost links met een bovenlicht met ronde hoeken, gecombineerd met de blokbeploistering links ervan, viel echter tot op heden niet aan de Groot Hertoginnelaan terug te vinden.

Wie weet waar dit is?

naam/datum	waar?	hoe zeker?	toelichting
Toon/verberg reacties			
Martin Snuverink (martin snuverink @planet nl) vr 15-2-2008 11:19:12	Nog geen idee	niet van toepassing	Het lijkt me wijs om eerst eens goed te analyseren wát er verbouwd wordt. Zulke deurposten met ronde hoeken komen voor na verbouwing tot open portieken, waarin 2 voordeuren, bijv. Laan v NOI 32 en 34. Hier is het niet!
Nicole (nicole @caret net) di 1-4-2008 10:29:21	Niet waar maar wanneer...?	niet van toepassing	Het is waarschijnlijk de BOUW van een pand, niet de verbouwing: de takken/bloemen die er hangen wijzen (waarschijnlijk) op het zojuist gereedkomen van de dakconstructie. Het gaat dus om een pand met bouwjaar 1910-1911.
H de Haas. (tab av @wanadoo nl) za 5-4-2008 01:46:16	Geen idee.	niet van toepassing	Toch zijn er een aantal opmerkelijke feiten aan deze foto. Wat opvalt is, dat de muren nog niet zijn afgewerkt, dus geen verbouwing. Verder zitten er nog geen plafonds in en schijnt er licht door de tengels. Ook is er geen spoor van steigers te bekennen. Ik denk aan laagbouw en waarschijnlijk een aanbouw aan een bestaand pand.
H de Haas. (tab av @wanadoo nl) zo 15-6-2008 18:43:01	Groot Hertoginnelaan?	het is maar een idee	Op www. HTMfoto.net staat een foto van een pand met gepleisterde blokken en een laagbouw uitbouw. Zie: http://www.htmfoto.net/Rurup/GrootHertog

Fig. 13. Aan de oproep om informatie over een bepaalde historische foto is door meerdere gebruikers gevraagd. Het Haagse Gemeentearchief lanceerde met zijn website Zoekplaatje een van de eerste Nederlandse crowdsourcinginitiatieven. www.zoekplaatjes.nl/haags/index.htm#11

WISDOM OF THE GRONINGSE CROWDS

Het Poparchief Groningen is een voorbeeld van crowdsourcing door een bestaand audiovisueel archief: GAVA/RHC Groninger Archieven, dat daarmee als een van de eerste regionale archiefinstanties ervaring opdeed met dit fenomeen. De website heeft een autonome uitstraling maar wordt door GAVA/RHC Groninger Archieven in technische zin ondersteund, d.w.z. dat gebruikgemaakt wordt van dezelfde digitale infrastructuur. Ook in organisatorisch opzicht zijn er nauwe banden: GAVA/RHC Groninger Archieven biedt het Poparchief een fysiek onderkomen en is verantwoordelijk voor de financiering. Inmiddels bestaat het Poparchief uit een omvangrijke, gestructureerde verzameling materiaal over de lokale muziekcultuur van weleer, maar er is ook informatie te vinden over hedendaagse muzikanten en poppodia uit stad en provincie. De collectie wordt regelmatig uitgebreid door de hulp van liefhebbers, experts en popmuzikanten. Deze leveren hun bijdragen aan via e-mail of een formulier. Controle en eindredactie is in handen van professionals. Er is een redactie gevormd van tien redacteuren die de bijdragen verwerken onder leiding van een eindredacteur. De redacteuren vertegenwoordigen ieder een genre. Gebruikers kunnen ook fysiek materiaal deponeren, dat op vakkundige wijze wordt opgeslagen in de archiefdepots. De koppeling van een website aan een officiële archiefinstantie biedt gebruikers het gevoel van veiligheid en betrouwbaarheid waar het gaat om de verwerking van hun gegevens en materialen. Voor GAVA/RHC Groninger Archieven vormt het Poparchief een manier om haar collecties uit te breiden met het onderwerp 'popcultuur' en op eigentijdse wijze te communiceren met gebruikers.

www.poparchiefgroningen.nl/

5.1.3 SOCIAL TAGGING

Een tag is een beschrijvende term die wordt toegekend aan een stukje informatie, zoals een tekst, een link, een foto of een videoclip. *Social tagging* is simpelweg te beschrijven als het toekennen van deze trefwoorden of labels binnen een informeel collectief netwerk. Het verschijnsel wordt – samen met de resultaten van crowdsourcing – wel aangeduid als *user generated metadata* (UGM). Momenteel vormt social tagging de meest in het oog springende wijze waarop er buiten de professionele omgeving metadata worden geproduceerd. Social tagging is begonnen vanuit *social software*: webplatforms waarop gebruikers eigen materialen, bookmarks, commentaren e.d. kunnen uploaden, om ze vervolgens zelf te ontsluiten met een of meer voor hen relevante termen. Tags worden in eerste instantie toegekend per enkele foto, object of fragment. Er zijn ook platforms waar hele collecties van objecten worden getagd, zoals Flickr, de bekendste fotowebsite. Andere voorbeelden van grote platforms waarop wordt opgeslagen, geordend en gedeeld zijn: Technorati voor weblogs, Delicious voor websites en natuurlijk YouTube voor videoclips.

Inmiddels wordt social tagging meer en meer ingezet bij de ontsluiting van *bestaande* culturele collecties zoals die van archieven, musea en bibliotheken. Zo laat Steve Museum uit



Fig. 14. De BBC deed ervaring op met het online laten taggen van radiofragmenten door luisteraars. De proef is inmiddels afgerond en op basis van de resultaten worden nieuwe taggingprojecten opgezet. www.plasticbag.org/archives/2005/10/on_the_bbc_annotatable_audio_project/

Washington, een pionierproject op dit gebied, gebruikers online afbeeldingen van schilderijen taggen. Naast de officiële, klassieke ingangen die zijn gemaakt door de professionals zoals de naam van de kunstenaar, de stijlperiode en de gebruikte materialen, voegen gebruikers allerlei vrije termen toe die tot stand komen vanuit hun eigen interpretatie en associaties. Hierdoor wordt het materiaal toegankelijk gemaakt op tal van 'onorthodoxe' kenmerken zoals de kleuren, de achtergrond en onderwerpen en details van het schilderij. Voor audiovisueel materiaal lopen er inmiddels ook initiatieven. Zo heeft de BBC een audiotaggingproject opgezet waarbij gebruikers radiofragmenten kunnen afspelen om daar vervolgens hun eigen, vrije trefwoorden aan toe te kennen. Beeld en Geluid lanceerde samen met de KRO het *video labeling game* Waisda? waarbij deelnemers historische items kunnen taggen.

5.2 KUNNEN GEBRUIKERS BESCHRIJVEN?

Veel metadatacreatie door niet-professionele documentalisten bezit feitelijk een soort parallel in de analoge offlinewereld. Vanuit de productiestadia, van makers en producenten kwamen er immers altijd al gegevens mee naar het archief, zij het op papier: de scripts, de

PEER-TO-PEER IN BRABANT

Van de foto's van Bob van Brunschot tot de filmcollectie van de Zuidelijke Land- en Tuinbouworganisatie: het is allemaal te vinden op de Film- en Fotobank Noord-Brabant, een beeldbank met historische films en foto's van Noord-Brabant, ontwikkeld door de afdeling Beeld en Geluid van de Brabant-Collectie. De publiekscant van de Film- en Fotobank is toegankelijk voor iedereen. De beheeromgeving is alleen bestemd voor personen en organisaties die historisch film- en/of fotomateriaal over Brabant beheren. Deelname aan de beeldbank is gratis. Onder de participerende collectiehouders vallen archieven, heemkundekringen, gilden, harmonieën en andere institutionele beheerders van cultuurhistorische beelden. Medewerkers van de Brabant-Collectie, onderdeel van de bibliotheek van de Universiteit van Tilburg, beheren de Film- en Fotobank en zien toe op het goed functioneren. Ook kan men bij hen terecht voor aanmelding, informatie en advies. De fysieke collecties blijven waar ze al waren: bij de rechtmatige eigenaar of beheerder. In de onlinebeheeromgeving kunnen collectiehouders hun gedigitaliseerde films en foto's zelf uploaden, beheren en beschrijven. Er is een op Dublin Core gebaseerd beschrijvingsformat ontwikkeld waarin formele en inhoudelijke gegevens kunnen worden ingevoerd. Collectiehouders die meedoen aan de beeldbank beslissen zelf of, en in hoeverre, hun content online wordt getoond. Gebruikers op hun beurt kunnen het materiaal voorzien van eigen gegevens en commentaar. De collectiebeheerder van de desbetreffende film of foto beslist hoe met dit commentaar om te gaan. Deze Film- en Fotobank vormt een unieke combinatie van crowdsourcing, expertbeschrijvingen en social tagging, samengebracht in een peer-to-peer netwerk dat wordt gefaciliteerd vanuit een professioneel audiovisueel archief.

www.filmenfotobank-nb.nl/

productieformulieren en de presentatieteksten. Het gebruikmaken van kennis van externe experts voor het toegepast verrijken van beschrijvingen is uiteraard ook geen compleet nieuw verschijnsel. En dat het grote publiek wordt opgeroepen informatie aan te leveren over foto's en andere objecten, het *crowdsourcen*, is misschien wel nieuw, maar de eindcontrole van deze gegevens is vaak nog altijd in professionele handen. Dit gaat niet automatisch op voor het fenomeen social tagging. Voor professionele documentalisten en catalogiseerders vertegenwoordigt deze trend dan ook een heel nieuwe uitdaging. Dit geldt in het bijzonder voor die vormen waarbij de tags van gebruikers fungeren als autonome zoekingen. De mogelijke implicaties van deze ontwikkeling voor de professionele beschrijvingsconventies zijn dan ook groot. Het dilemma ligt in de mate van 'professionele' correctie en controle van de tags. Moeten documentalisten en catalogiseerders de tags standaard corrigeren en classificeren, zoals ze dat ook doen met andere extern geproduceerde metadata? Of kunnen ze vertrouwen op een autonoom proces zonder centrale richtlijnen, waarbij tags van gebruikers zonder meer kunnen worden gebruikt als volwaardige zoekingen?

TAKING PICTURES TO THE PUBLIC

In oktober 2008 zette het Nationaal Archief samen met Spaarnestad 800 foto's op Flickr. Het publiek kon de foto's taggen en van commentaar voorzien. In zes maanden tijd werden de foto's meer dan een miljoen keer bekeken door mensen over de hele wereld. Er werden 6852 tags toegekend en 1916 commentaren achtergelaten. Flickr bleek een uitstekend middel voor het vergroten van het publieksbereik. Zo steeg het aantal pageviews op de eigen NA-beeldbank in deze periode van 200.000 naar 300.000 pageviews per maand. De bedoeling van de pilot was allereerst: onderzoeken of de zoekfunctionaliteit verbeterd kan worden met behulp van gebruikerstags. In een vervolgonderzoek zal het zoekresultaat dan ook worden vergeleken met dat van 'professionele' trefwoorden. Een andere vraag ging over de mogelijkheid van verrijking van kennis rondom fotocollecties door eindgebruikers. Het bleek dat uiteindelijk maar 3% van alle *comments* is toegevoegd aan de eigen beeldbanken van het NA en Spaarnestad, wat neerkomt op het aanpassen van 23 records. Slechts een klein deel van de comments bevat namelijk relevante gegevens over locatie, datum en personen op de foto's, i.c. feitelijke informatie waarvan de juistheid gecheckt kon worden. Conclusie van de pilot is dat gebruikers op het punt van kennisverrijking niet kunnen worden ingezet als vervanging van professionele archivariissen en ervaren vrijwilligers. De door gebruikers ingevoerde tags zullen echter naar verwachting wél gaan bijdragen aan het verbeteren van de doorzoekbaarheid, omdat zij aansluiten bij de eigen meer intuïtieve terminologie van gebruikers.

www.flickr.com/photos/nationaalarchief

Deze vraag raakt direct aan het debat tussen de voorstanders van het vrije social tagging en de representanten van de strakkere lijn zoals dat nog in volle hevigheid wordt gevoerd op web- en andere platforms. Voor de eerste groep – informatieprofessionals die voornamelijk afkomstig zijn uit de wereld van de bibliotheken en de papieren archieven – is het volledig vrijlaten van social tagging bepaald onwenselijk. Deze volgelingen van Aristoteles en Melvil Dewey, de 'vaders' van de traditie van het ordenen en classificeren, baseren zich op het leidende principe van het klassieke catalogiseren: de bedoeling van de auteur of de maker. Hun redenering luidt: de bottom-up methode van het vrije taggen laat ieder standpunt toe, dus ook meerdere standpunten. Wat voor de ene gebruiker zwart is, is voor de andere wit, waar de één een ezel waarneemt, ziet de ander een paard. Betekenis en 'waarheid' worden aldus relatief. Deze groep ziet social tagging als belangrijk en welkom, maar toch vooral als *aanvulling* op de officiële, professioneel gemaakte ontsluiting.

De aanhangers van het vrije taggen menen dat de nieuwe fenomenen social tagging en crowdsourcing het aristotelische denken zélf op zijn kop zetten. De catalogus is geen statische, voltooide bron meer maar wordt een proces, een *open source* in een netwerk van informatie, waarin de meerderheid gelijk heeft. Het genereren van informatie wordt geliberaliseerd en is niet langer uitsluitend in handen van professionele instituten. De

BASALE, SOCIALE EN LUIE TAGGERS

Het Telematica Instituut heeft een researchproject afgerond naar taggedrag. Men wilde vaststellen of het vooraf voorstellen van bepaalde termen het taggingproces zou versnellen. Ook werd nagegaan of social tags geschikt zijn als autonome zoekingang. De proeftaggers werden opgedeeld in 3 groepen die elk 150 clips moesten taggen: de Basic Taggers die hun tags zelf moesten bedenken, de Social Taggers die suggesties kregen aangeboden gebaseerd op de tags die andere taggers hadden toegevoegd en de Lazy Taggers, die ook nog eens de originele metadata van de professionals als suggestie ontvingen. Bij het invoeren van tags bleek dat veel social tags onderling werden overgenomen, iets dat de kwaliteit van de tags zou kunnen aantonen. De groep Social Taggers had daarbij het laagste aantal unieke tags. Tags van de Lazy Taggers beïnvloedden derhalve de Social Taggers, die daarmee indirect ook de professionele metadata overnamen. Andere bevinding was dat taggers die suggesties ontvingen in het algemeen niet sneller werkten. En ondanks de suggesties bedachten de Social Taggers en de Lazy Taggers nog steeds een hoop tags zelf (i.c. 57% overgenomen versus 43% eigen tags). Echter, de suggesties zorgden wél voor meer consistentie in de set tags. Een belangrijke conclusie t.a.v. het zoeken was dat gebruikte zoektermen zeer vaak correspondeerden met social tags. Conclusie: social tags sluiten kennelijk aan bij de belevingswereld van gebruikers en zullen dan ook in bepaalde gevallen beter werken dan professionele metadata. Aanbeveling van het Video Tagging-project: er is vervolgonderzoek nodig naar manieren om gebruikers te motiveren, want het is duidelijk dat social tagging een goede manier is om metadata te creëren!

www.informatik.uni-trier.de/~ley/db/indices/a-tree/v/Veenstra:Mettina.html
http://tech.ebu.ch/docs/techreview/trev_2008-Q2_social-tagging.pdf

voorstanders beamen dat verschillende gebruikers hetzelfde object wellicht bezien vanuit meerdere gezichtspunten en het mogelijk nét weer anders zullen interpreteren. Dat er zo misschien niet langer één 'waarheid' is, wordt toegegeven. Maar is dat erg? Tot chaos hoeft social tagging in elk geval niet te leiden. Semantische technieken zijn in staat de verschillende tags te filteren en bundelen. De impliciete informatie die zich bevindt in de relaties tussen de afzonderlijke tags, gebruikers, objecten en collecties wordt statistisch geanalyseerd en bewerkt en creëert aldus verband en betekenis. De 'waarheid' komt dan vanzelf wel bovendien: hoe groter de verzameling foto's, fragmenten of objecten, hoe meer eenstemmigheid over de tags.

5.3 SOCIAL TAGGEN EN DE AUDIOVISUELE BESCHRIJVING

Wat betekent dit debat tussen de traditionele catalogiseerders en *free tagging*-adepten voor het ontsluiten van audiovisueel materiaal? Allereerst valt op dat veel socialtagging-

TAGGEN ALS SPELLETJE

Beeld en Geluid lanceerde in 2009 in samenwerking met de KRO het video labeling game Waisda? Internetgebruikers taggen hierbij spelenderwijs Polygoonfragmenten en afleveringen van de tv-programma's *Boer zoekt vrouw*, *Memories* en *Spoorloos*. Spelplezier en competitiedrang moeten voldoen als motivatie om deel te nemen. Doel van de pilot is om de meerwaarde van gebruikersannotaties te bepalen bij het ontsluiten van audiovisueel materiaal. Via deze vorm van serious gaming wordt ook geëxperimenteerd met een nieuw soort publieksdienst als basis voor een mogelijk verdienmodel. Op de Waisda?-site kiest de gebruiker een van de 2000 video's en vult hij de woorden in die omschrijven wat er te zien en te horen is. Er kunnen punten worden verdiend wanneer een andere gebruiker bij dezelfde clip exact dezelfde term invoert. De bedenkers van Waisda? gaan dus uit van het consensusprincipe: hoe meer mensen dezelfde tag gebruiken, hoe groter de kans dat het gaat om een correcte, zinvolle term. Er wordt in Waisda? samengewerkt met de VU die onderzoekt of en hoe gesproken woord in de video's de keuze van de tags beïnvloedt. Dit gebeurt o.m. door de toegekende tags te matchen met woorden in de Teletekstondertitels en de spraak-naar-teksttranscripten. Ook wordt bekeken of de kwaliteit van de tags verbeterd wordt door het aanbieden van gecontroleerde wordsystemen, zoals WordNet en de GTAA (Gemeenschappelijke Thesaurus Audiovisuele Archieven). In de tweede fase van Waisda? wordt een koppeling gemaakt tussen de gebruikerstags, automatisch gegenereerde metadata en professionele catalogusbeschrijvingen: van social tagging naar semantic tagging...

www.waisda.nl

initiatieven worden genomen vanuit culturele collecties die van oudsher niet gedetailleerd op inhoud of onderwerp werden ontsloten, zoals foto's, archiefdocumenten en objecten uit de (beeldende) kunst. Van dit soort materialen werden traditioneel m.n. veel formele gegevens (maker, bron, vindplaats e.d.) vastgelegd. Voor veel officiële archiefdocumenten geldt dat ze vooral worden ontsloten op herkomst en originele gebruikscontext. Boeken, tijdschriften en veel andere papieren documenten worden daarbij meestal niet op hoofdstuk of paragraaf ontsloten, zoals dat wel gebeurt met de items en onderdelen van een audiovisuele productie. De tendens naar onlinepresentatie en -exploitatie van culturele collecties leidt echter tot de noodzaak van ontsluiting van de inhoud: er moet ook door leken op gezocht kunnen worden. Social tagging kan dan uitkomst bieden. In de 'officiële' catalogus of inventaris blijft alles bij het oude, maar voor de webpresentatie van de collectie aan het grote publiek, wordt datzelfde publiek ingeschakeld bij het inhoudelijk beschrijven.

Audiovisuele archieven hebben een eigen beschrijvingstraditie. Allereerst moet de beschrijving hier dienen als substituut van het audiovisuele document dat immers niet

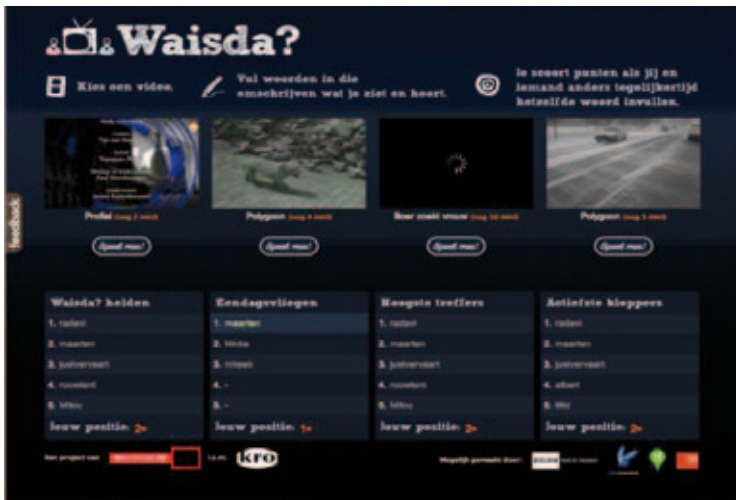


Fig. 15. Startscherm van het serieuze taggingspel Waisda?.

als zodanig te bevatten is. Bij veel audiovisuele collecties is daarnaast de mogelijkheid van hergebruik op fragmentniveau belangrijk. Beide zaken vereisen een chronologische beschrijving van de scènes en shots. Zoals eerder betoogd kent de audiovisuele beschrijving verschillende semantische lagen: de informatiecontent, de audiovisuele content en het stockmateriaal (zie 2.2). Het is duidelijk dat hierdoor sprake is van meerdere gezichtspunten op het materiaal. De beschrijving wordt immers deels vervaardigd met het oog op mogelijk hergebruik van de onderdelen – de shots, de scènes en de fragmenten – in een andere context. Dit betekent dat voor dit doel de intentie van de maker wordt losgelaten. In deze zin lijkt het beschrijven van veel audiovisueel materiaal sterk op de gedecontextualiseerde en clipgerichte werkwijze van social taggers.

Het creëren van metadata op shot- en clipniveau is zeer arbeidsintensief. De immer toenemende vloed aan binnenstromend materiaal belet veel collectiehouders al hun materiaal op clipniveau te ontsluiten en zo te werken aan het opbouwen van grote collecties stockshots. Er is simpelweg geen tijd om ieder potentieel herbruikbaar fragment te voorzien van trefwoorden. Op termijn kunnen beeld- en spraakalgoritmen hier ongetwijfeld bij gaan ondersteunen, maar op dit moment zijn ze archieven en collectiehouders op een hoger semantisch niveau nog nauwelijks van dienst. Op dit onderdeel zouden – onder bepaalde voorwaarden – de social taggers kunnen worden ingezet. 'Free tagging' door het algemeen publiek kan audiovisuele collectiehouders helpen hun collecties in elk geval op *fragmentniveau* toegankelijk te maken en ze aldus beter te exploiteren. De professionele documentalist – met zijn classificatiesystemen, thesauri en gecontroleerde woordenlijsten – dient echter zoveel mogelijk de feiten te blijven controleren: het wie, wat, waar en wanneer van de productie als geheel met inachtneming van de intentie van de maker en de bredere context van het document. Zo blijft de informatiecontent de verantwoordelijkheid van de professional, terwijl voor de overige semantische niveaus – de audiovisuele content en de stockshots – de gebruiker kan worden ingezet.

6 DE TRENDS EN HUN GEVOLGEN

6.1 NUT EN NOODZAAK

De vernieuwingen zoals hierboven beschreven hebben grote economische, culturele en maatschappelijke voordelen. Deze voordelen hangen samen met een sterk verbeterde beschikbaarheid van content en metadata aan vele gebruikersgroepen. De automatische metadatadoorgifte binnen de digitale productieworkflow bevordert de snelheid en de efficiency van de werkprocessen. Een groot aantal (formele) gegevens over audiovisuele producties bestaat al voordat ze een archiefomgeving in stroomt en is dus ook al vóór uitzending of publicatie te benaderen. Door het *write once read/edit many*-principe worden er minder invoerfouten gemaakt.

Met behulp van automatische indexersoftware kan véél meer materiaal in elk geval op een minimaal ontsluitingsniveau (transcripten en visuele low-levelkenmerken) worden teruggevonden. Dat is een niet te onderschatten verbetering in iedere audiovisuele productieomgeving zoals nu al blijkt bij organisaties als CNN waar dagelijks met deze toepassingen wordt gewerkt. Het indexeren van beelden waarvan het begrip van de context niet op de eerste plaats staat (i.c. de audiovisuele content, in het bijzonder de stockshots) zal op niet al te lange termijn waarschijnlijk grotendeels automatisch verlopen. Automatisch indexeren van beeld en geluid creëert daarbij ingangen die tegemoetkomen aan nieuwe vragen van een nieuw onlinepubliek.

Het betrekken van gebruikers kan helpen bij het ontsluiten van grote hoeveelheden digitale data. Door het inzetten van experts, crowdsourcing en social tagging wordt gebruikgemaakt van kennis die elders bestaat. Dit bespaart kosten omdat de documentalist zijn research niet zelf hoeft te doen en het materiaal veelal niet meer hoeft te bekijken of te beluisteren. Gebruikersinput kan er ook toe bijdragen dat een collectiehouder zijn waren beter kan toesnijden op gebruikers. *Profiling*- en *recommendation*-technieken analyseren alle impliciete kennis achter de gegevens die gebruikers toevoegen. ‘Amazon-achtige’ toepassingen (‘...andere gebruikers die dit materiaal raadpleegden, raadpleegden ook...’) maken een gepersonaliseerde, interactieve presentatie mogelijk. Het laten taggen en becommentariëren van materiaal van buitenaf kan archieven en collectiehouders ook nog eens belangrijke feedback geven op de eigen ontsluitingsconventies. Collectiebeheerders kunnen social tagging op meerdere manieren inzetten. Ze kunnen zelf digitale platforms inrichten waarop gebruikers kunnen taggen, waarna deze informatie in gecontroleerde vorm beschikbaar komt in hun officiële catalogi en informatiesystemen. Archieven kunnen ook bestaande initiatieven zoals Flickr en YouTube gebruiken om delen van hun collecties te presenteren en te laten taggen. Social tagging leidt aldus tot gebruikersengagement, extra *exposure* én nieuwe zoekingen.

Door deze ontwikkelingen kunnen audiovisuele archieven hun bereik vergroten en hun services verbeteren. Dit legitimeert hun bestaansrecht en versterkt hun culturele en maatschappelijke rol. De coherentie tussen gebruikers van de collecties onderling wordt ook bevorderd. Crowdsourcing- en socialtagging-initiatieven brengen immers niet alleen groepen clips, programma’s, documenten en objecten samen maar vormen rond bepaalde collecties

of onderwerpen automatisch *user communities*. Hierdoor worden gemeenschappelijke interessesferen van individuen gelokaliseerd en wereldwijd verbonden. Het virtueel samenvoegen van particuliere en professionele materialen in de zgn. peer-to-peer netwerken doet dit ook en verbetert tevens de kennisuitwisseling tussen gebruikers en instituten.

Het is vooral de combinatie van de trends en ontwikkelingen die straks de echte winst gaat opleveren. Administratieve, formele en bepaalde inhoudelijke gegevens worden gegenereerd tijdens de digitale workflow, software indexeert de low- en mid-levelkenmerken van beeld en geluid en tags van gebruikers leveren inhoudelijke gegevens. Al deze extern toegekende metadata bevatten impliciete informatie, die door semantische technieken naar boven kan worden gehaald. Semantische interoperabiliteit legt verbanden tussen collecties, tussen classificaties en thesauri en tussen objecten en gebruikers. Door uit deze verbanden intelligente conclusies te trekken komt de context van een programma of fragment naar de oppervlakte. Machinaal leren kan worden toegepast wanneer de automatische analyse van gegevens wordt gekoppeld aan interactie met de gebruiker. Er kan dan worden geleerd van de vragen die de gebruiker stelt en van de patronen in diens zoekgedrag. Elke zoekactie levert zo weer nieuwe kennis op, die weer kan worden gebruikt voor het genereren van nieuwe zoekresultaten.



Steeds meer gebruikers creëren vanuit hun luie stoel zelf metadata: het beschrijven van audiovisuele content is niet langer voorbehouden aan de professionele documentalist.

De traditionele, manuele inhoudsbeschrijving van de documentalist is ongetwijfeld nog steeds het meest precies en doelmatig. De reden voor het gebruik van de nieuwe middelen ligt dan ook in de toenemende hoeveelheden digitaal materiaal en het tekort aan mensen om die hoeveelheden adequaat te ontsluiten. Het gezamenlijk inzetten van de hierboven beschreven methoden en technieken kan niet alleen snel maar ook breed toegang verschaffen tot audiovisuele collecties. Er wordt daarbij een zeer hoog niveau van granulariteit bereikt, iets dat in het digitale domein door zowel professionele gebruikers als door het grote publiek wordt verwacht. Centrale vraag die direct opdoemt: waar laten deze ontwikkelingen op termijn de professionele archivaris en documentalist?

6.2 EEN NIEUWE FOCUS VOOR DE PROFESSIONAL

Tot dusver waren het de documentalist en archivariissen die de gegevens over audiovisuele collecties creëerden en vastlegden. Zij maakten de beschrijvingen, kozen de trefwoorden en bepaalden zo welke formele en inhoudelijke informatie beschikbaar kwam aan de buitenwereld. In de netwerkomgeving gaat de informatiestroom niet meer alleen naar buiten: zij komt nu ook *naar binnen*, in de vorm van productiemetadata, automatisch gegenereerde gegevens en gebruikerstags. Hiermee zijn documentalisten niet langer van begin tot eind verantwoordelijk voor het tot stand brengen van beschrijvende gegevens. Deze ontwikkeling noopt tot het anders gaan inzetten van hun traditionele kennis en vaardigheden.

Documentalisten blijven echter verantwoordelijk voor een aantal essentiële keuzen waar het gaat om het niveau en de kwaliteit van de toegang. De relevantie van alle extern gecreëerde gegevens zal namelijk blijvend moeten worden getoetst. Dit betekent dat steeds weer moet worden bepaald voor welke onderdelen van de content, op welke semantische niveaus (i.c. informatiecontent, audiovisuele content en stockshots) de extern geproduceerde metadata volstaan en waar zij moeten worden gecorrigeerd, verplaatst, aangepast en/of verrijkt. Dit is een gecompliceerde taak die vooralsnog alleen kan worden verricht vanuit een combinatie van kennisdomeinen: begrip van de structuur van audiovisuele materialen, *knowhow* op het gebied van ordening en classificatie met vooral: kennis van de vormen van (her)gebruik van programma's, fragmenten en collecties en inzicht in de diverse gebruikersgroepen en hun zoekgedrag. Op basis van deze specifieke vakinhoudelijke en contextuele kennis kan (en vooralsnog: *moet*) manueel worden ingegrepen waar nodig. Door de metadata die binnenkomen te *screenen* en zinvol in te passen behoudt de documentalist zijn bepalende invloed op de wijze waarop informatie over een programma of fragment uiteindelijk weer 'naar buiten' gaat.

6.2.1 CONTEXT EN KWALITEIT

In de digitale (omroep)productieomgeving verandert het statische concept 'archief' in het werkwoord 'archiveren', een dynamisch gedistribueerd proces dat zijn beslag krijgt langs de hele productieketen. Ongemonteerd en gepubliceerd materiaal, of het nu gaat om clips, hele programma's of complete series, moet tijdens het productieproces terugvindbaar blijven. Gestructureerde en gestandaardiseerde metadata – van oudsher het domein van de professionele audiovisuele documentalist – zijn hierbij de meest kritische ingrediënten. Kennis en vaardigheden op het gebied van informatiebeheer zijn dan ook niet alleen meer nodig aan het eind van het proces wanneer het materiaal de archiefomgeving binnen stroomt, maar zullen moeten worden overgenomen door andere deelnemers aan het proces.

MEEWERKEN AAN INNOVATIE

Al ruim tien jaar lopen er grote en kleine researchprojecten tussen academische en technische partijen en archieven, audiovisueel of anderszins. Op Europees niveau is er het programma Information Society Technologies (IST) voor de ontwikkeling van nieuwe producten en diensten in het voordeel van culturele verscheidenheid. Op nationaal niveau worden binnen de programma's MultimediaN en CATCH geavanceerde technologieën bedacht voor het ontsluiten, presenteren en aan elkaar koppelen van Nederlands erfgoed. Doel van al deze initiatieven is het vergroten van de bijdrage van de *memory institutes* aan de zgn. cultuureconomie. Belangrijk middel is steeds weer de realisatie van geïntegreerde, (semi)automatische toegang tot grote hoeveelheden content, het liefst zowel crossmediaal als multilinguaal. Partners binnen deze projecten zijn universiteiten en researchinstututen, technische en commerciële bedrijven en collectiebeheerders. Deze laatste groep levert vooral de inzichten en ervaring op het gebied van het ontsluiten en beschikbaar stellen. Projectmedewerkers zijn veelal (oud-)documentalisten en archivarissen. Hun praktische kennis blijkt hier onmisbaar. Hoe anders moeten de hoogtechnologische prototypes worden gekoppeld aan *requirements* van de eindgebruikers? En hoe – zonder deze inbreng – maken de projectresultaten ooit kans daadwerkelijk te worden geïmplementeerd? Actief bijdragen aan innovatieve projecten betekent voor medewerkers van archieven dat nieuwe kennis moet worden opgedaan van informatietechnologie. Ook moeten deze projectmedewerkers in staat zijn hun praktijkervaring op het gebied van hergebruik, zoekgedrag en catalogisering te vertalen naar een researchomgeving en vice versa. Communicatieve vaardigheden zijn in zo'n proces heel belangrijk. De resultaten van innovatieve projecten zijn immers niet meteen zichtbaar, terwijl de noodzakelijke personele en dus financiële inspanning van de eigen organisatie wél moet worden 'verkocht' aan management en collega's.

cordis.europa.eu/home_en.html

www.nwo.nl/nwohome.nsf/pages/NWOP_5XSKYG

www.multimedien.nl/nl/home.php

Vanuit de professionele kennis van het identificeren, ordenen en toegankelijk maken van audiovisuele documenten worden de randvoorwaarden gecreëerd voor een organisatiebrede metadata-invoer. Deze randvoorwaarden bestaan uit definities, formats en modellen voor de metadata en thesauri of *controlled vocabularies* voor de geüniformeerde invoer. Als materialen en metadata eenmaal in een 'archivaal' stadium zijn beland – na uitzending of publicatie – volgt aanpassing van de ingestroomde metadata. Op dat moment kunnen de gegevens worden verrijkt en gecontextualiseerd. Anders gezegd: professionals en professionele kennis gaan de noodzakelijke 'kwaliteit' toevoegen aan het beheer van de gegevens tijdens de digitale workflow, opdat terugvindbaarheid gegarandeerd blijft. Een tweede belangrijke taak is het toevoegen van context en verdieping aan de metadata die

BEROEP: MEDIAMANAGER

Bij de BBC zijn zgn. mediamanagers werkzaam, vaak als lid van het team dat verantwoordelijk is voor de productie van een bepaald radio- of televisieprogramma of van een serie. Hun werk bestaat uit het toezien op de ordentelijke creatie, mutatie en opslag van metadata. Deze mediamanagers – veelal oud-documentalisten van de afdeling Information & Archives van de BBC – zijn ooit begonnen zonder definitieve functieomschrijving. Bij het opzetten van het digitale productieproces een aantal jaren geleden, was er al wel de behoefte aan professioneel metadata-beheer, maar nog geen zicht op het hoe en wat, waar en wanneer van dit beheer. Inmiddels is de functie helemaal ingeburgerd. Een mediamanager bij de BBC moet allereerst beschikken over de vaardigheden van een documentalist en dus vertrouwd zijn met richtlijnen en regels als basis van informatiesystemen. Maar de vereisten zijn ruimer: hij of zij moet ook kennis hebben van de complete productieworkflow en van de verschillende systemen voor bijvoorbeeld het ingesten, het editen en het publiceren van content. Veel van het werk bestaat uit dingen uitleggen aan en uit het onderhandelen met medewerkers van productieafdelingen. Dit om een zo goed mogelijke organisatiebrede metadata-invoer te garanderen. Mediamanagers dienen dus ook te kunnen optreden als trainer. Hiervoor zijn communicatieskills belangrijk. Deze skills komen ook van pas bij het uitleggen waar een mediamanager *niet* voor is: het oplossen van computerproblemen bijvoorbeeld. Mediamanagers vormen bovenal de zo belangrijke verbinding tussen de productie en het archief en daarmee tussen het dynamische dagelijkse uitzendproces en de voorwaarden voor langetermijnbewaring van de BBC-programma's.

www.bbc.co.uk/guidelines/dq/contents/archives.shtml

tijdens de digitale workflow is ingevoerd: het aanbrengen van het langetermijnperspectief. Hiermee wordt het erfgoedbelang van materiaal behartigd dat in de dynamische productieomgeving anders wellicht verloren zou gaan.

6.2.2 WAARDEOORDELEN

Automatisch indexerend is *quick* maar vooralsnog *dirty*. De techniek is op dit moment nog geen serieuze vervanging van de professionele, manueel gegenereerde *counterpart* met zijn precisie en zijn onderscheiden semantische niveaus. De tools voor beeld-, spraak- en taaltechnologie zullen het werk van de audiovisuele documentalist wel belangrijk kunnen ondersteunen, waarmee voorlopig een situatie ontstaat van *automatische* precatalogisering en *manuele* contextualisering.

Voorlopig blijft er behoefte aan menselijke *beoordelingen* van audiovisuele fragmenten en producties ten behoeve van bepaald gebruik. Waardeoordelen zijn nuttig om te midden van de digitale overvloed beelden en geluiden met een bijzonder gebruiks- dan wel cultuur-historisch belang te oormerken. Deze oordelen zijn contextgebonden, kunnen per tijdsperiode



Vanuit hun praktische kennis kunnen documentalisten en archivariissen actief bijdragen aan researchprojecten op het gebied van toegang en ontsluiting.

wisselen en zijn dus moeilijk in algoritmen om te zetten. Dit soort manuele toevoegingen door de documentalist kan zorgen voor de noodzakelijke 'subjectiviteit' in een beschrijving: de resultaten van automatisch indexeren alléén zijn feitelijk nog te 'objectief'.

Een belangrijke *nieuwe* taak voor audiovisuele documentalisten is het aanleveren van kennis uit de praktijk bij de ontwikkeling van de softwaretoepassingen voor automatische indexering. Documentalisten hebben een manier van werken ontwikkeld die is gebaseerd op veel ervaring met catalogisering ten behoeve van meerdere typen gebruik. Door samen te

BEROEP: MEDIAREDACTEUR

Het Institut National de l'Audiovisuel (INA), het nationale audiovisuele archief van Frankrijk, herbergt miljoenen uren materiaal van Franse producenten, waaronder veel omroepen. Een substantieel deel daarvan is intussen gedigitaliseerd. Talloze nieuwe programma's stromen dagelijks in digitale vorm het archief binnen. Het catalogiseren van dit materiaal wordt inmiddels niet meer gedaan door de INA-documentalisten maar door de producenten zelf. Hun metadata worden direct geïmporteerd in de archiefdatabase. Documentalisten verrichten nu vooral redactioneel werk en zijn actiever geworden in de klantenservice. Een belangrijke nieuwe taak heet: *Thématisation*. Hierbij wordt het historisch en actueel archiefmateriaal opgedeeld in hiërarchisch gestructureerde collecties die op de INA-website beschikbaar worden gesteld. De onderwerpen zijn gerangschikt naar thema, (actueel) event en publieke persoonlijkheid en kunnen worden benaderd via beschrijvingen en bewegende beelden. Oud en nieuw INA-materiaal kan zo door gebruikers in gecontextualiseerde vorm worden geraadpleegd. Documentalisten zijn verantwoordelijk voor de selectie en de structurering van de thema's. Ze zijn ook betrokken bij de publicatie. De nieuwe redactionele taken vereisen veel kennis van de inhoud van de collecties. Deze kennis is ook belangrijk bij het ondersteunen van klanten tijdens het zoekproces, een andere nieuwe focus van documentalisten. Gedetailleerde kennis van de collecties vormt tenslotte de basis voor gemotiveerde keuzen in conserveringsprojecten. Voor het opbouwen, delen en verrijken van deze kennis onder documentalisten wordt binnen INA een bewust beleid gevoerd. 'Thématisation' van de INA-collecties heeft inmiddels geleid tot een sterk verhoogd gebruik van archiefmateriaal.

inatheque.ina.fr/
www.ina.fr/



Fig. 16. Thema Mei 1968, onderdeel van de historische online themacollecties van het Institut National de l'Audiovisuel.

werken met de ontwikkelaars kunnen zij de behoeften van (eind)gebruikers inbrengen in de specificaties en een belangrijke rol spelen in het testen en valideren van de nieuwe tools.

6.2.3 STRUCTURERING EN THEMATISERING

Het betrekken van gebruikers bij het ontsluiten van collecties levert voordelen op wanneer wordt voldaan aan een aantal randvoorwaarden: de informatie heeft een bepaalde massa en kwaliteit, en het belang van een dergelijk interactieproces voor beide partijen is bewezen. Als informatieprofessionals kunnen documentalisten en archivariissen helpen deze voorwaarden te creëren. Crowdsourcing en socialtagging-processen dienen te worden georganiseerd. De behoeften van verschillende gebruikersgroepen moeten worden geïnventariseerd en gestructureerd. Een professionele collectiehouder is in staat de gemeenschappelijke thema's en onderwerpen aan te bieden die buiten het bereik liggen van de individuele deelnemers. Aanpassing van de eigen professionele systemen ten behoeve van communicatie met de buitenwereld kan gebruikers een veilige en betrouwbare structuur bieden, waarin professionals de menselijke ondersteuning vormen.

Social tagging genereert doorgaans een grote hoeveelheid subjectieve, inconsistente, relativistische en onnauwkeurige zoekresultaten. Gebruikers beschikken niet over de professionele instrumenten om hun eigen wereld te beschrijven en te ordenen. In veel gevallen kennen zij tags toe die vooral voor henzelf van belang zijn. Door het aanbieden van professionele trefwoordenlijsten en thesauri aan social taggers kunnen gegevens in een bredere, meer uniforme context worden geplaatst en wordt de algemene terugvindbaarheid bevorderd. De combinatie van de professionele ontsluiting met de social tags leidt zo tot een verhoging van het aantal zoekingen. Tegelijkertijd kunnen professionals zorgen voor een zichtbaar onderscheid tussen de persoonlijk getinte gebruikerstags en de professionele metadata. Door *folksonomy* op deze wijze te paren aan *taxonomie*, voegen professionals de noodzakelijke objectiviteit toe aan de vaak subjectieve gebruikerstags.

7 TEN SLOTTE

De bruikbaarheid van traditionele documentatie- en informatiemanagementvaardigheden wordt ook in de digitale netwerkgeving alom erkend. Hierboven werd beschreven op welke manieren deze professionele kennis en vaardigheden zouden kunnen worden ingebed in het proces van digitale productie, distributie en ontsluiting. Hierbij werd duidelijk dat er verschuivingen gaan optreden in het werk en in de focus van de audiovisuele documentalist en archivaris. De onlinewereld zal de professionele praktijk nog op een aantal andere punten veranderen. Sommige vertrouwde uitgangspunten ten aanzien van de ordening en de kwaliteit van informatie zullen daarbij moeten worden losgelaten. Dit geldt niet alleen voor de netwerkkultuur in zijn algemeenheid, maar evenzeer voor de archief- en andere collecties, die daar onvervreemdbaar deel van gaan uitmaken. Het beschrijven van audiovisueel materiaal blijft immers niet voorbehouden aan de professional. In de netwerkkultuur zijn het veelal niet-professionals die zelfstandig gegevens creëren over beeld- en geluidsfragmenten, en andere objecten en documenten. Ook softwaretools leveren een toenemend aandeel in de ontsluiting. Metadatacreatie wordt een gemeenschappelijke activiteit waarbij de catalogus als statische verzameling gecontroleerde gegevens transformeert tot een proces, een 'open source' die vanaf meerdere plekken wordt gevoed en gebruikt. De van oudsher strikt gescheiden domeinen van de 'invoer' en het 'zoeken' (bevolkt door respectievelijk de professional en de gebruiker) groeien hierdoor naar elkaar toe.

Gegevens en beschrijvingen afkomstig uit velerlei bronnen gaan op het netwerk dooreenlopen. Professionele beschrijvingen van audiovisuele items worden vermengd met automatisch gegenereerde gegevens, met 'user generated metadata' en met professionele beschrijvingen uit andere domeinen. De eigen catalogusbeschrijving van een collectiebeheerder zal in wisselende, vaak oncontroleerbare presentaties op het web terugkomen. Veel informatie op het netwerk ontbeert context. Werkelijke semantische interoperabiliteit is immers vooralsnog toekomstperspectief. Maar ook al zou deze techniek nu al op grote schaal kunnen worden toegepast: steeds meer materiaal wordt zonder verband aangeboden. Er ontstaat een informatiewereld waarin een groot deel van de materialen en documenten los in de digitale ruimte zweeft, zonder onderlinge relaties. Zo werken deelnemers aan een digitaal productieproces gezamenlijk aan een grote verzameling losse mediaobjecten, fragmenten, teksten en clips waarvan veel nog niet is ingepast in een programma, serie of collectie – en dat ook nooit zal worden. Automatische annotatie ontsluit vooralsnog alleen veel lagere semantische kenmerken. Zonder manueel ingrijpen krijgen items geen inbedding in een betekenisvol geheel en worden ze niet geïnterpreteerd. Gebruikers in hun rol als social taggers kennen weliswaar inhoudelijke trefwoorden toe, maar doen dat vooral vanuit hun eigen belevingswereld. Ze brengen tussen de fragmenten en objecten geen verbanden aan waardoor een bredere context in eerste instantie niet wordt geboden.

Precisie, context en details waren altijd de kracht van het professioneel catalogiseren. De gestandaardiseerde manier van beschrijven van documentalisten en archivariissen leidt tot krachtige zoekmogelijkheden. Een professionele beschrijving toont reflectie, historische afstand en objectiviteit. Er wordt zinvol onderscheid aangebracht tussen de verschillende semantische niveaus van een productie. Hierin ligt het begrip besloten van de complexe

relatie tussen vorm en inhoud van een audiovisueel document. Tegelijk garandeert dit onderscheid meervoudig (her)gebruik. Onlinegebruikers en deelnemers aan een digitaal productieproces kunnen echter niet zomaar wat professionele documentaristen kunnen. En computers kunnen niet wat mensen kunnen: onmiddellijk begrijpen en duiden wat ze zien en horen. In het algemeen zal het gebruik van professionele standaarden, met de bijbehorende precisie en de juiste doseringen objectiviteit en subjectiviteit, sterk verminderen. Gevolg is dat de nieuwe informatieomgeving als geheel veel *metanoise* gaat bevatten: slechte, onnauwkeurige of niet ter zake doende metadata.

Toegang tot audiovisuele collecties wordt in de netwerkcultuur op meerdere manieren bewerkstelligd. De kwaliteit van de toegang, ofwel de aangeboden metadata, loopt daarbij uiteen: van tags die direct zijn aangebracht door gebruikers via automatische annotaties door computers naar beschrijvingen die zijn gemaakt dan wel gecontroleerd door professionals. Die laatste categorie kenmerkt zich nog altijd door de grootste nauwkeurigheid, betrouwbaarheid en consistentie. Dit soort waarden zal in het digitale informatietijdperk alleen nog maar in belang toenemen. Het is daarom te overwegen om – te midden van alle andere metadata op het netwerk – beschrijvende gegevens die zijn *gescreend* door een documentalist of een andere informatieprofessional standaard te voorzien van een kwaliteitsmerk.

Hilversum, september 2009

Annemieke de Jong

Nederlands Instituut voor Beeld en Geluid

GERAADPLEEGDE LITERatuur

- Bailer, Werner et al., 2005. State of the Art of Content Analysis Tools for Video, Audio and Speech, Deliverable D15.3 MDS3, FP6-IST-507336. PrestoSpace.
- Borgman, Christine L., 2000. From Gutenberg to the Global Information Infrastructure: MIT Press/Library of Congress.
- Castells, Manuel, 1996. The Information Age Volume I: The Rise of the Network Society. Blackwell Publishers Inc.
- Declercq, Brecht, 2009. De digitale ontsluiting van audiovisueel erfgoed. Erfgoed 2.0: nieuwe perspectieven voor digitaal erfgoed, pp. 159-177. Pharo Publishing.
- Film- en Beeldbandarchief NOB, 1993. Regels voor de beschrijving van audiovisuele documenten.
- Gazendam, Luit et al., 2009. Automatic Annotation Suggestions for Audiovisual Archives: Evaluation Aspects. Interdisciplinary Science Reviews, Volume 34, Numbers 2-3, pp. 172-188 (17). Maney Publishing.
- Hollink, Laura et al., 2009. A Multidisciplinary Approach to Unlocking Television Broadcast Archives. Interdisciplinary Science Reviews, Volume 34, Numbers 2-3, pp. 253-267 (15). Maney Publishing.
- Jong, de Annemieke, 2005. Het audiovisueel domein als kringloopwinkel. Jaarboek Stichting Archiefpublicaties.
- Jong, de Annemieke, 2003. Metadata in the audiovisual production environment. Werkuitgave Nederlands Instituut voor Beeld en Geluid.
- Jong, de Annemieke, 2000. Het woord, het beeld en de algoritmen. Tijdschrift voor Mediageschiedenis. Uitgeverij Het Spinhuis.
- Keen, Andrew, 2007. The Cult of the Amateur: How Today's Internet Is Killing Our Culture and Assaulting Our Economy. Nicholas Brealey Publishing.
- Kennisland, 2008. Inventarisatie Social Tagging voor de ontsluiting van fotocollecties van archieven. Project Beelden voor de Toekomst.
- Lauwers, Mieke (ed.), 2004. Changing sceneries, changing roles Part II. Proceedings FIAT Seminar Media Management Commission Amsterdam. Netherlands Institute for Sound and Vision.
- Lévy, Pierre, 1997. Collective Intelligence: translated by Robert Bonnono. Perseus Books, Cambridge, Massachusetts.
- Lusenet, de Y., 2008. Geven en nemen, archiefinstellingen op het sociale web. Taskforce Digitale Toegankelijkheid Archieven.
- Manovich, Lev, 2009. Introduction to Cultural Analytics: from new media to more media. Paper FIAT Seminar Media Management Commission, Changing sceneries, changing roles Part IV, Stockholm.
- Manovich, Lev, 2009. Cultural Analytics: Visualizing Cultural Patterns in the Era of More Media. <http://lab.softwarestudies.com/2009/06/publications.html> [augustus 2009]
- Millan, Douglas et al., 2007. Changing sceneries, changing roles Part III. Proceedings FIAT Seminar Media Management Commission Wenen. http://www.fiatifta.org/conferences/seminars/past/vienna_2007/report.light.html [augustus 2009]
- Moortgat, Judith, 2009. Taking Pictures to the Public. Evaluatieverslag Nationaal Archief & Spaarnestad Photo op Flickr The Commons. <http://office.imagesforthefuture.org/wp-content/uploads/2009/07/evaluatie-nationaal-archief-op-flickr-commons.pdf> [augustus 2009]
- Nederlands Instituut voor Beeld en Geluid, 2009. Regelgeving t.b.v. documentalisten iMMix 2006-2009.
- Peterson, Elaine, 2006. Beneath the Metadata, Some Philosophical Problems with Folksonomy. D-Lib Magazine, November 2006, Volume 12, Number 11, ISSN 1082-9873. <http://www.dlib.org/dlib/november06/peterson/11peterson.html> [augustus 2009]
- Smeulders, Arnold et al., 2005. Multimedia information technology and the annotation of video. Jaarboek Stichting Archiefpublicaties.
- Snoek, Cees en Worring, Marcel, 2008. Concept Based Video Retrieval. Foundations and Trends in Information Retrieval, Volume 2, No 4, pp. 215-322. <http://staff.science.uva.nl/~cgmsnoek/pub/snoek-concept-based-video-retrieval-fntir.pdf> [augustus 2009]

- Snoek, Cees, 2008. The Semantic Pathfinder: Using an Authoring Metaphor for Generic Multimedia Indexing. <http://staff.science.uva.nl/~cgmsnoek/pub/snoek-pathfinder-pami.pdf> [augustus 2009]
- Snoek, Cees en Smeulders, Arnold, 2009. Video Search Engines. <http://staff.science.uva.nl/~cgmsnoek/pub/slides/CVPR-2009-Snoek-Smeulders-VideoSearch-web.pdf> [september 2009]
- Surowiecki, James, 2004. The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations. Doubleday, ISBN 0-385-50386-5
- Tan, E. en Müller, H., 2003. Integration of Specialist Tasks in the Digital Image Archive. In H. van Oostendorp (ed.), Cognition in a digital world, pp. 47-73, Mahwah, NJ: Erlbaum.
- Uricchio, William, 2005. Moving beyond the Artefact. Preserving the Digital Heritage: Principles and Policies, selected papers. Netherlands National Commission for UNESCO. ISBN 978-90-6984-523-4.
- Velsen van, Lex en Melenhorst, Mark, 2008. User Motives for Tagging Video Content. <http://www.lexvanvelsen.nl/VanVelsen&Melenhorst2008.pdf> [augustus 2009]
- Vroemen, Lucas, 2009. ROOS Metadata. http://www.lucasvroemen.nl/?page_id=36 [augustus 2009]
- Weinberger, David, 2005. Tagging and why it matters. <http://cyber.law.harvard.edu/sites/cyber.law.harvard.edu/files/07-WhyTaggingMatters.pdf> [augustus 2009]
- Weinberger, David, 2007. Everything is Miscellaneous: The Power of the New Digital Disorder. Times Books.



Nederlands Instituut voor Beeld en Geluid
Media Park
Sumatralaan 45
Postbus 1060
1200 BB Hilversum
T 035 – 677 3434
www.beeldengeluid.nl