

BOM-VL. BEWARING EN ONTSLUITING VAN MULTIMEDIALE DATA IN VLAANDEREN

Werkpakket 1

Deeltaak 4

GEBRUIKERSNODEN

GECONTEXTUALISEERD AANBOD (~CONTEXTUALISERING)

Versie

18.03.2009

Auteurs

Brecht Declercq, H el ene Verreyke, mmv Jeroen Walterus
FARO. Vlaams steunpunt voor cultureel erfgoed

Feedback versturen naar:

H el ene Verreyke

helene.verreyke@faronet.be

Jeroen Walterus

jeroen.walterus@faronet.be



INLEIDING.....	4
1. SITUERING	5
2. CONTEXTUALISERING	8
2.1. METADATA	9
2.1.1. <i>Klassieke metadata-annotatie en metadatastandaarden</i>	10
2.1.2. <i>Expert generated tags</i>	14
2.2. VERBINDING MET ANDERE CONTENT EN RANDDOCUMENTATIE	16
2.2.1. <i>Inhoudelijke contextualisering</i>	16
2.2.2. <i>Doelgroepgerichte contextualisering</i>	17
2.3. CONTEXTUALISERING DOOR DE GEBRUIKER.....	18
2.3.1. <i>User generated tags</i>	19
2.3.2. <i>Folksonomies</i>	20
2.3.3. <i>Rating</i>	22
2.3.4. <i>Fora</i>	22
2.3.5. <i>Collecties</i>	24
2.3.6. <i>Social bookmark managers</i>	25
2.3.7. <i>Externe gebruikersplatformen</i>	27
3. CASES.....	28
3.1. ONDERZOEKSPROJECTEN	28
3.1.1. <i>CATCH (Continuous Access to Cultural Heritage)</i>	28
3.1.1.1. <i>CHIP (Cultural Heritage Information Presentation)</i>	28
3.1.1.2. <i>CHOICE (Charting the Information Landscape Employing Context Information)</i>	29
3.1.1.3. <i>MUSEUM (MULTiple collection-SEarching Using Metadata)</i>	29
3.1.2. <i>Beelden voor de toekomst – Pilot OMP (Onderwijs Media Platform)</i>	29
3.1.3. <i>BIRTH (Building an Interactive Research and delivery network for Television Heritage)</i>	30
3.2. PRAKTIJK	31
3.2.1. <i>Video Active</i>	31
3.2.2. <i>Nederlands Erfgoed: Digitaal!</i>	31
3.2.3. <i>Klara.be</i>	32
3.2.4. <i>Teleblik en Academia</i>	32
3.2.5. <i>INA</i>	33
3.2.6. <i>Smithsonian Photography Initiative</i>	34
3.2.7. <i>EUROPEANA</i>	34
4. AANDACHTSPUNTEN UIT DE SECTORFEEDBACK	36
5. CONCLUSIES EN AANBEVELINGEN	41

5.1. CONCLUSIES	41
5.2. AANBEVELINGEN	42
6. BIBLIOGRAFIE.....	46

Inleiding

Werkpakket 1 Gebruikersnoden bestaat uit de deeltaken open en dynamisch archief (1), creatief atelier (2), wetenschappelijk gebruik (3) en gecontextualiseerd aanbod (4). Contextualisering wordt beschouwd als een gebruikersnood omdat, zo vermeldt het BOM-VL-projectdossier, een multimedia-archief voor bepaalde gebruikers vooral zin heeft wanneer de content in een betekenisvolle context wordt aangeboden.

Ten eerste wordt gezocht naar een degelijke afbakening van de begrippen *contextualisering* en *gecontextualiseerd aanbod* (1. Situering). Ook gaat de nodige aandacht naar enkele met contextualisering gerelateerde begrippen. Vervolgens wordt aan de hand van deskresearch en via een survey van buitenlandse cases een analyse gemaakt van de verschillende vormen van contextualisering (2. Contextualisering). Hoe en door wie wordt de contextualisering tot stand gebracht, op welke manieren kunnen collectiebeheerders de collectie contextueel aanbieden? Welke context levert de gebruiker zelf aan en kan dit ook van nut zijn voor andere gebruikers? Dat zijn de centrale vragen van de analyse. Ten derde (3. Cases) worden een aantal relevante internationale en nationale cases besproken, gaande van onderzoeksprojecten over contextualisering tot praktijkvoorbeelden van betekenisvol aanbieden van content. Vervolgens worden de aandachtspunten opgenomen die onderzoekers en erfgoedbeheerders uit de verschillende sectoren (onderwijs, cultuur, kunsten, erfgoed en media) formuleerden op basis van dit rapport en een vragenlijst (4. Aandachtspunten uit de sectorfeedback). Het rapport sluit af met de belangrijkste conclusies en aanbevelingen (5. Conclusies en aanbevelingen).

1. Situering

Tijdens het onderzoek is al snel gebleken dat aan de term contextualisering de meest uiteenlopende interpretaties worden gegeven. Daarom wordt eerst gezocht naar een algemene situering van het begrip contextualisering. Daarbij wordt een onderscheid gemaakt tussen het proces van *contextualisering* en het product, het *gecontextualiseerd aanbod*.

Letterlijk betekent contextualisering *het voorzien van een context*. Het *voorzien* duidt op het toekennen, het verbinden, in dit geval van de context aan de *essence*. Het begrip *essence* slaat op elke mogelijke inhoud van een collectie, in het geval van het project BOM-VI zijn dit multimediale data, zoals digitale audiovisuele bestanden of digitale afbeeldingen. Naast deze *essence* bevatten collecties (doorgaans) ook metadata. Dit zijn alle gegevens die op de collectiedelen slaan: *data over data*. De *essence* en de metadata samen noemen we *content*.¹

Context staat voor de totale omgeving waarin iets (i.c. de *essence*) zijn betekenis krijgt. Binnen een online aanbod kan die omgeving bestaan uit:

- andere content
- metadata

Contextualisering slaat dus op:

- het toekennen van metadata aan *essence*
- de verbinding van content met andere content

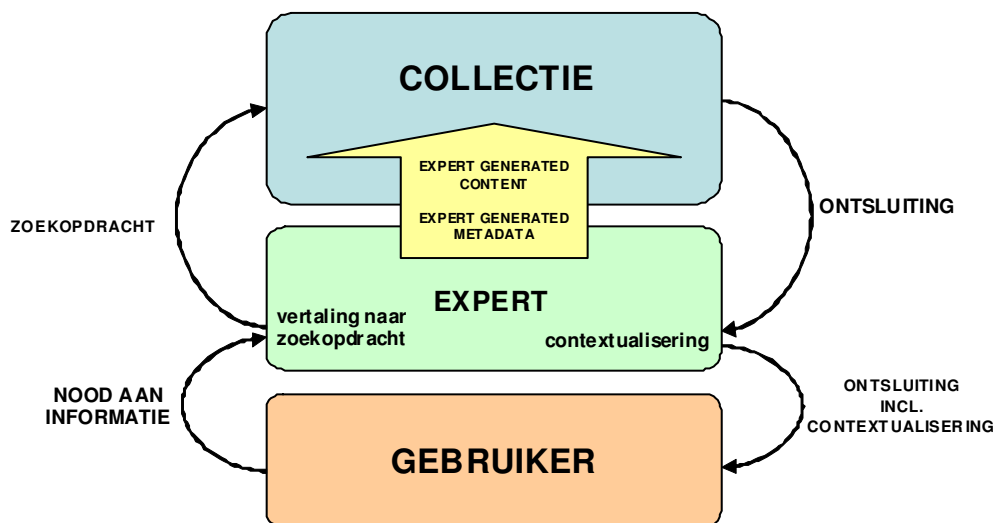
De content wordt dus door de context verrijkt. Door de verbinding met andere items en metadata krijgt een archiefitem bovendien zijn betekenis. In die zin kan contextualisering ook geïnterpreteerd worden als het aanbrengen van *betekenisvolle relaties*, al dan niet afgestemd op een bepaalde doelgroep. Contextualisering is één van de mogelijke wegen waarlangs de inhoud van het archief toegankelijk, vindbaar en begrijpbaar gemaakt wordt. Het groeiende belang van de contextualisering wordt duidelijk aan de hand van de evolutie van het begrip *ontsluiting*. Ontsluiting duidt op elke manier waarop content bij de gebruiker geraakt, als proces of als het eindresultaat van dit proces. Een collectie kan bijvoorbeeld ontsloten worden via een boek of televisiereeks waarin de collectie besproken wordt. Contextualisering slaat in dat geval op de verwijzingen, in het boek of de televisiereeks, naar allerlei soorten van context, binnen of buiten de collectie. In dit rapport gaat het enkel over de contextualisering van een collectie die ontsloten is via een database en een online zoekmachine die daarop zoekvragen formuleert.

Vóór de ontwikkeling van de netwerkcultuur, web2.0 en de digitaliseringsgolf had het collectiebeheer ook al (ten minste gedeeltelijk) een publieksgerichte finaliteit. Maar de

¹ OOMEN & VAN KRANENBURG 2003. We verwijzen hier ook naar de rapporten die in het kader van het BOM-VI *Werkpakket 3 Metadatastandaarden en uitwisselingsformaten* werden uitgewerkt: <https://projects.ibbt.be/bom-vl/>

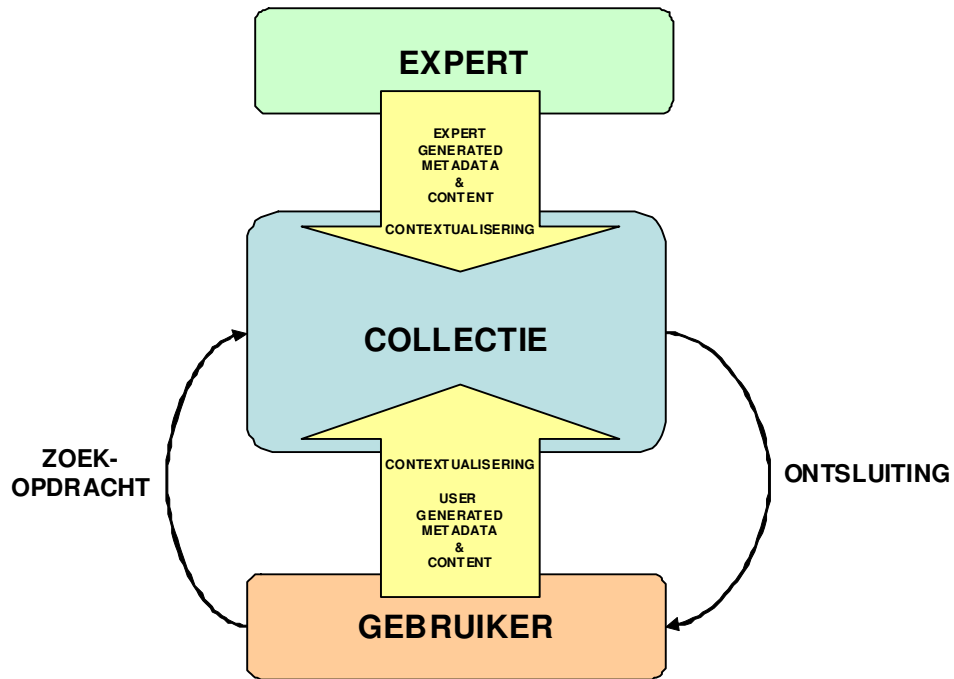
metadata-annotatie was afgestemd op doorzoekbaarheid door experts: zij ondervroegen ten dienste van de gebruikers de databanken via metadata die ze vaak zelf hadden ingevoerd en brachten de inhoud vervolgens naar het grote publiek via *one-to-many* kanalen, inclusief contextualisering. De expert bevond zich tussen gebruiker en collectie, de collectie was met andere woorden (zo goed als) enkel via de expert toegankelijk (fig. 1).

Tegenwoordig worden collecties rechtstreeks online aangeboden en geconsulteerd. De intermediaire rol van de expert inzake vertaling van zoekvragen van de gebruiker en contextualisering van zoekresultaten naar de gebruiker is in deze weggevallen. Deze evolutie heeft drie duidelijk zichtbare gevolgen. Ten eerste is het belang van een gebruiksvriendelijke zoekinterface veel groter geworden.² Bovendien kreeg de gebruiker een veel actievere rol: hij krijgt de kans zelf metadata (*user generated metadata*) en/of content (*user generated content*) toe te voegen. Het derde gevolg staat centraal in dit rapport: de contextualisering van de collectie gebeurt niet louter meer door de expert, maar (bewust of onbewust) ook door de gebruiker. Omdat de collectie permanent bevraagd kan worden, gebeurt de contextualisering niet meer *ad hoc*, maar sluit ze steeds dichter aan bij de andere ontsluitingsmogelijkheden. Zo kan ook de metadata-annotatie steeds meer beschouwd worden als een deel van de contextualisering (fig. 2).



Figuur 1: relatie expert-collectie-gebruiker en de plaats van de contextualisering hierin in een pre-digitale omgeving.

² Over deze interfaces, zie: KOUWENHOVEN 2006.



Figuur 2: gevolgen van digitalisering, online aanbod en web2.0 voor de relatie expert-collectie-gebruiker en de plaats van de contextualisering hierin.

2. Contextualisering

Contextualisering kan worden opgedeeld aan de hand van een aantal eigenschappen:

- Intern - Extern. Waar bevindt zich de contextualisering, waar leidt ze naar toe?

Contextualisering kan zich zowel binnen als buiten de collectie bevinden. Interne contextualisering wil zeggen dat de context afkomstig is of toegekend wordt binnen de eigen collectie. Bij externe contextualisering wordt de content verbonden met metadata of andere content afkomstig van buiten de collectie.³

- User generated – Expert generated. Door wie gebeurt de contextualisering?

Het onderscheid tussen *user generated* en *expert generated* biedt een antwoord op de vraag door wie de contextualisering gebeurt: de gebruiker of de expert. Dit nieuwe onderscheid is een rechtstreeks gevolg van de web2.0-ontwikkelingen, waarbij de gebruiker een centrale plaats inneemt.⁴

- Actief, bewust – Passief, onbewust. Hoe komt de contextualisering tot stand?

De meeste vormen van contextualisering gebeuren bewust en actief. De gebruiker of de collectiebeheerder voegt wetens en willens context toe aan de collectie. Maar zeker in een web2.0-omgeving kan er ook onbewust of toch zeker passief context worden toegevoegd. Het gedrag van de gebruikers bij het exploreren van de collectie vormt een immense informatiebron voor contextualisering. De gebruikers laten als het ware hun sporen achter, die door de collectiebeheerder dankbaar worden gebruikt om het aanbod door middel van contextualisering af te stemmen op diezelfde (identificatie door login) of een gelijkgezinde (identificatie door gelijkaardig gedrag) gebruiker.

Een voorbeeld van contextualisering op basis van onbewust aangeleverde gegevens is de toepassing **Muziekadvies** van de Centrale Discotheek Rotterdam. Deze toepassing suggereert artiesten en groepen die de gebruikers zouden kunnen interesseren. De gebruiker moet daarvoor eerst zelf drie namen van artiesten of groepen opgeven. Op basis van statistieken uit de uitleendatabank wordt er dan gezocht naar profielen van gebruikers met een gelijkaardige muzieksmaak. Muziekadvies suggereert vervolgens de groepen waarvan de gelijkgezinde gebruikers muziek hebben uitgeleend. Elke uitlening kan dus, zonder dat de lener ervan bewust is, andere gebruikers helpen om nieuwe artiesten te leren kennen.⁵

³ De mogelijkheid om een eigen selectie samen te stellen binnen de collectie is dus een vorm van interne contextualisering, social bookmark managers zijn een vorm van externe contextualisering.

⁴ Zo gebeurt tagging meestal *user generated*, terwijl metadata die volgens een thesaurus gestructureerd zijn, meestal *expert generated* zijn.

⁵ Zie <http://www.muziekweb.nl/menu/shared/muziekadvies/>.

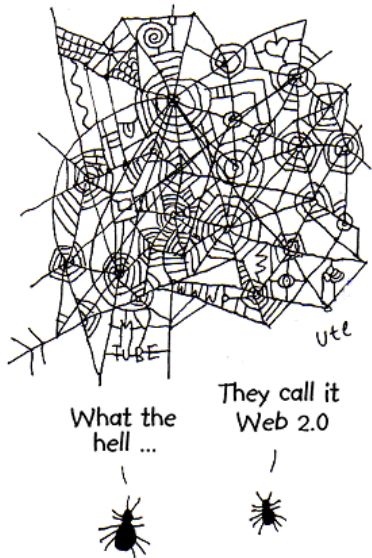


Fig. 3: Ute Hamelmann, www.schnutinger.de.

Het is duidelijk dat er een ruime waaier aan vormen van context en contextualisering bestaat. In de praktijk is een erg strikte onderverdeling misschien zelfs onmogelijk. Dat is eigen aan de chaotische netwerkgeving van het fenomeen web2.0.

Tijdens de interdisciplinaire conferentie 'Strategies for multimedia archives', dat werd georganiseerd in het kader van BOM-VI, waren de deelnemers van de sessie 'How is context generated and maintained in interdisciplinary archives' het eens dat er nog geen duidelijke definitie bestaat van contextualisering.⁶ Sommigen zien metadata als context, voor anderen bestaat contextualisering enkel uit de verbanden tussen de objecten. Wij hebben ervoor gekozen om metadata mee op te nemen in het rapport, omdat daar de contextualisering start. Vervolgens bespreken we het verbinden van objecten zodat een inhoudelijke meerwaarde ontstaat. Ten slotte worden alle vormen van contextualisering besproken waar de gebruiker een rol in speelt.⁷

2.1. Metadata

De meest bekende toevoeging van context zijn ongetwijfeld de metadata (beschrijvende, structurele, administratieve, ...), die al eeuwen gebruikt worden als ontsluitingsvorm. De *data over data* beschrijven de *essence* in velden met als einddoel deze zoek- en vindbaar te maken.

⁶ Verslagen van de individuele bijdragen van deze conferentie zijn momenteel beschikbaar op de conferentiewebsite: <http://www.conference.bom-vl.be/>. Het rapport met het algemene verslag van de conferentie dat binnen het BOM-VI Werkpakket 2 Archivalisering en selectie werd gemaakt, is beschikbaar via de BOM-VI website: <https://projects.ibbt.be/bom-vl/>

⁷ Zie ook het rapport dat binnen BOM-VI Werkpakket 5 werd opgesteld, waar in functie van het onderzoek naar distributiemodaliteiten een onderscheid werd gemaakt tussen structurele contextualisering, inhoudelijke contextualisering en gebruikerscontextualisering: MOSTMANS, VAN PASSEL & NULENS 2008: https://projects.ibbt.be/bom-vl/fileadmin/user_upload/frontendfiles/BOM-VI%20WP5.2%20Distributiemodaliteiten.pdf

In het web2.0-tijdperk is de visie op metadata echter geëvolueerd, want niet alleen de experts, maar ook leken willen de collectie kunnen bevragen via metadata die op hen zijn afgestemd, die desnoods door henzelf zijn ingevoerd.

In audiovisuele collecties nemen de beschrijvende metadata een bijzonder belangrijke plaats in omdat audiovisuele elementen op zich (nog) geen zoekterm zijn. Alvorens bestanden via tekstuele zoektermen doorzoekbaar zijn, moet de inhoud dus eerst in tekstuele vorm beschikbaar zijn. In zekere zin geldt hetzelfde voor teksten op papier, die eerst digitaal als tekstuele tekens herkend moeten worden.⁸ Het komt er voor de ontsluiting van audiovisuele archiefitems dus op aan om ten eerste efficiënte tekstuele beschrijvende metadata te annoteren als representatie van de audiovisuele content en ten tweede niet-tekstuele indexeringen en zoekstrategieën te ontwikkelen.

Niet-tekstuele indexeringen en zoekstrategieën staan nog in de kinderschoenen en worden nog maar zelden in de praktijk gebracht⁹, al zijn er al websites waar muziekcollecties doorzoekbaar zijn via *query by humming* (QbH, zoeken door te neuriën).¹⁰ Een bespreking van deze technieken valt echter buiten het doel van dit rapport. Wat de metadata-annotatie van audiovisuele archiefitems betreft, is er de laatste jaren sprake van enkele belangrijke nieuwe ontwikkelingen zoals automatische indexering, annotatie en data mining, digitalisering van het metadatabeheer en tagging en folksonomies.

2.1.1. Klassieke metadata-annotatie en metadatastandaarden

Klassieke metadata kunnen beschouwd worden als de meest basale manier van contextualisering. Na de opname in een archief is de toekenning van deze metadata door de collectiebeheerder de eerste vorm van ontsluiting. Aan de inhoud van het archief worden namelijk de eerste eigenschappen verbonden met de essence. Dit kan op verschillende niveaus gebeuren: van de collectie als geheel tot het itemniveau.

De klassieke metadata worden onderverdeeld in een aantal groepen. De belangrijkste zijn daarbij de beschrijvende, de structurele en de administratieve. Beschrijvende metadata zijn die metadata die een bepaald object beschrijven voor identificatie, zoals titel, auteur, ... Structurele metadata zijn de eigenschappen die aangeven hoe het archiefitem zich tot een groter deel of de gehele collectie verhoudt: een kopie, een onderdeel, ... Administratieve metadata zijn de metadata die gebruikt worden om het object te beheren, zoals de auteursrechterlijke gegevens en de nummering van het archiefitem.

⁸ Dit kan bijvoorbeeld door Optical Character Recognition (OCR, optische tekenherkenning).

⁹ Bijvoorbeeld in WITCHCRAFT en CHORAL, delen van het Nederlandse CATCH-programma worden de mogelijkheden hiervan onderzocht. Zie ook http://www.nwo.nl/nwohome.nsf/pages/NWOP_6CCC3L.

¹⁰ Bij QbH fungeert een al dan niet zelf opgenomen digitale melodie als zoekterm. De geluidsgolf wordt daarbij via verschillende parameters vergeleken met geluidsgolven van de muzikale archiefitems uit de database, zoals gebeurt bij <http://www.midomi.com>. Musipedia heeft een enigszins primitievere interface: hier moeten de muzieknoden via een piano-toepassing als zoekterm worden ingevoerd: <http://www.musipedia.org>.

Om de opzoekbaarheid van alle items via de metadata te garanderen, moeten de metadata zo veel mogelijk systematisch en uniform zijn opgesteld. Dat is ook nodig voor de hanteerbaarheid van grote hoeveelheden, bijvoorbeeld bij de migratie van data of metadata, de opbouw van ontsluitingstoepassingen of de interoperabiliteit.¹¹ Structureringen zoals metadata-annotatie volgens standaarden of thesauri¹² maken dit alles mogelijk. Het metadatamodel of de metadatastandaard beschrijft de regels waarin de metadata (moeten) zijn opgesteld. Hieronder worden ter illustratie drie belangrijke metadatastandaarden kort besproken.¹³

- Dublin Core

Dublin Core is een metadatastandaard die een grootste gemene deler wil vormen tussen verschillende systemen. Binnen Dublin Core wordt gesproken over *resources*, *elementen*, *qualifiers* en *schemes*. *Resources* zijn de te beschrijven objecten. Ze worden beschreven met behulp van 15 verschillende *elementen* gaande van *creator* tot *rechten*. *Qualifiers* en *schemes* duiken op in Dublin Core Qualified. *Qualifiers* worden gebruikt om elementen te verfijnen. *Creator* kan bijvoorbeeld verfijnd worden als *fotograaf*, maar niet elke software zal de *qualifiers* begrijpen. Software die niet weet wat de term *fotograaf* uit het voorbeeld betekent, zal dit interpreteren als een creator. Aan die software die er mee overweg kan, bieden de *qualifiers* dus extra informatie aan, evenwel zonder aan interoperabiliteit in te boeten. *Schemes* laten toe aan te geven op welke manier elementen moeten worden ingevuld. Zo kan worden aangegeven dat het subject een trefwoord is uit een bepaalde thesaurus en niet een willekeurig woord. De *scheme* vermeldt dan welke thesaurus er is gebruikt. Naast de *qualifiers* en *schemes* voegt Dublin Core Qualified aan Dublin Core overigens ook nog drie elementen toe. En naast die standaardset kunnen ook andere elementen worden toegevoegd. Hierbij wordt aangeraden elementen te kiezen die in andere metadatastandaarden wel worden gebruikt. Zo'n nieuwe set elementen vormt een *application profile*.

- P/Meta

P/Meta is de Europese standaard die gebruikt wordt bij de *business-to-business* uitwisseling van radio- en televisieprogramma's. Bij de ontwikkeling van dit model wenste de European Broadcasting Union niet te raken aan de structuur, de werkmethode of de algemene concepten van de participerende organisaties. Een essentieel voordeel van P/Meta is dan ook dat organisatiespecifieke opslagschema's

¹¹ Interoperabiliteit is de mogelijkheid om data uit te wisselen tussen verschillende systemen met verschillende hardware, software, datastructuur en interfaces met een minimaal verlies van gegevens en functionaliteiten.

¹² Een thesaurus is een geordende woordenlijst, waarbij de betekenis van de opgenomen termen kan worden afgeleid uit de plaats binnen de lijst. Die betekenis legt op die manier spelling vast en de criteria op basis waarvan een term ter ontsluiting (bv. als trefwoord) aan een archiefitem moet worden toegekend.

¹³ We verwijzen hier nogmaals naar de rapporten die in het kader van het BOM-VI *Werkpakket 3 Metadatastandaarden en uitwisselingsformaten* werden uitgewerkt, waar verschillende standaarden uitvoerig worden uiteengezet:
<https://projects.ibbt.be/bom-vl/>

voor data gefilterd en afgebeeld kunnen worden *bovenop* het P/Meta-schema, voor een vlotte metadata-uitwisseling, onafhankelijk van de onderliggende infrastructuur voor datatransport. P/Meta is een verzameling van definities voor het semantisch raamwerk voor de uitwisseling van informatie bij audiovisueel materiaal.¹⁴ Concreet bevat P/Meta de exacte definitie en betekenis van elk beschrijvend element (attribuut) dat van belang is bij de meeste productieprocessen. Deze definities berusten op de professionele betekenis en interpretatie van elk concept en worden uitgedrukt in normale menselijke taal. P/Meta is hoofdzakelijk van toepassing in een B2B omgeving, al blijft de consument natuurlijk het einddoel. P/Meta focust op de interoperabiliteit tussen het archief en de *content creators* (mediamakers) en verder tussen die content creators en de *content distributors* (uitzendverantwoordelijken).¹⁵ Om de mogelijkheid tot uitwisseling te garanderen beschikt P/Meta over een aantal functionele groeperingen van attributen. Er kunnen zelfs nieuwe logische sets gecreëerd worden om aan de eigen specifieke transactievereisten te voldoen, ook los van de technische implementatie. Een bijkomend maar erg belangrijk voordeel is dat het model geïmplementeerd kan worden onafhankelijk van bepaalde platformen, of andere beperkende omstandigheden.

- ISAD (G)

ISAD (G) of International Standard for Archival Description (General) is zoals de naam het zegt een metadatastandaard die speciaal voor de beschrijving van archieven is ontwikkeld, door de International Council of Archives. Zoals P/Meta zich richt op het vergemakkelijken van de uitwisseling van metadata (al dan niet in combinatie met de essence), zo richt ISAD(G) zich op het vaststellen en verklaren van de context en de inhoud van het archiefmateriaal, om zo de raadpleegbaarheid ervan te bevorderen. ISAD(G) bevat algemene regels voor archivistisch beschrijven, die kunnen worden toegepast ongeacht de vorm of het medium van het materiaal. De regels worden uitgedrukt in zesentwintig elementen (waarvan zes essentiële) die gecombineerd mogen worden om een archiefdeel te beschrijven. ISAD(G) definieert dus deze elementen, maar de inhoud en de structuur moeten verder worden ingevuld op basis van nationaal verschillende regels. Er wordt wel aangeraden om voor de ontsluitingstermen gebruik te maken van de International Standard Archival Authority Record for Corporate Bodies, Persons and Families (ISAAR(CPF)). De regels hebben betrekking op zeven gebieden: identificatie van het archiefstuk, de context ervan, de inhoud en structuur, voorwaarden voor raadpleging en gebruik, aan het archiefstuk verwant materiaal, aantekeningen (gegevens die niet in een van de andere gebieden

¹⁴ P/Meta is van toepassing op de zogenaamde *definitie laag*. Daarnaast kunnen metadataseten ook nog een *technologielaag* (over het behoud van de betekenis van de originele informatie-elementen bij data-uitwisseling tussen verschillende technologieën) en een *data-uitwisselingslaag* (over hoe en via welk medium de gecodeerde informatie uitgewisseld wordt) bevatten.

¹⁵ HOPPER 2002. Overigens voorziet Hopper in zijn artikel wel al de uitbreiding naar toepassingen rechtstreeks voor de consument.

kunnen worden ondergebracht) en beschrijvingsbeheer (over de archivalische beschrijving zelf).

Metadata-extractie en -annotatie is erg vatbaar voor automatisering. Enerzijds omdat het een erg arbeidsintensieve bezigheid is, anderzijds ook omdat ze in haar klassieke vorm aan strikte regels moet beantwoorden. Sinds enkele decennia spitst nogal wat onderzoek zich daarom toe op (semi-) automatische indexering en metadata-annotatie. Een eenvoudig voorbeeld hiervan is spraakherkenning, waarbij de computer een gesproken tekst herkent (extractie) en transcribeert (annotatie). De achterliggende technologie, die naar harde, door de computer constateerbare criteria zoekt om kenmerken van beeld of audio te herkennen, is heel wat ingewikkelder.¹⁶ Digitale audiovisuele bestanden worden daarbij beschouwd als een opeenvolging van data. Hierbinnen wordt bijvoorbeeld gezocht naar patronen die gekoppeld worden aan waarneembare kenmerken, eventueel met een kansberekening als tussenstap (*data mining*). Zo wordt bijvoorbeeld de toonhoogte (mediale frequentie) van de audio geanalyseerd om automatisch het geslacht van de spreker (met een zekere waarschijnlijkheid) vast te stellen. Indien het bijvoorbeeld om audio bij beelden gaat kan gezichtsherkenning in steun worden gebruikt. In de praktijk worden deze recente technieken doorgaans gezien als complementair met een menselijke vorm van metadata-extractie en -annotatie, in die gevallen is de automatisering dus slechts gedeeltelijk. Het ligt in de lijn van de verwachtingen dat de automatische extractie (en annotatie) een zeer interessante basis voor contextualisering van audiovisuele collecties zal gaan vormen. Nu al worden in professionele omgevingen zoals omroeparchieven enkele van deze technieken toegepast.¹⁷

Door de arbeidsintensiviteit van de metadata-extractie en -annotatie en de grote investeringen die gemeoid zijn met de automatisering daarvan, zijn steeds meer beheerders van audiovisuele collecties gaan beseffen dat de metadata die in de loop van het audiovisuele productieproces in digitale vorm ontstaan, best zorgvuldig bewaard kunnen worden en hergebruikt tijdens de archivering. Dit principe wordt *write once, read/edit many* genoemd. Op die manier zijn er minder inspanningen nodig om deze metadata tijdens de archivering opnieuw samen te stellen. Het feit dat het productieproces van heel wat audiovisuele data vandaag digitaal verloopt, maakt deze evolutie nog meer voor de hand liggend. De archiefdiensten vragen daarom van de hele organisatie om de basisprincipes van het metadatabeheer over te nemen, zodat binnen het documentbeheer een echt *lifecycle-management* ontstaat. Voor tekstuele documenten wordt dit principe bijvoorbeeld al succesvol toegepast in de administratie van de stad Antwerpen. De VRT past dit al toe bij de archivering van haar audiovisuele collecties. Ook waar productie en archivering niet binnen dezelfde organisatie gebeuren, zoals bij het Nederlands Instituut voor Beeld en Geluid, kan dit worden toegepast, mits de nodige integraties van software en opslagsystemen en een degelijk kader van afspraken.

¹⁶ Een gestructureerd overzicht van deze technieken zowel voor audio als voor beelden is te vinden bij BAILER, et al. 2005.

¹⁷ Zie bijvoorbeeld VAN DER MAAS et al. 2006.

Vermits alle metadata fungeren als data over data, geven ze meer informatie over het archiefitem, situeren het in zijn context en contextualiseren het dus. Maar metadata kunnen ook de basis betekenen van een contextualisering via kunstmatige intelligentie, meer bepaald de zgn. *zwakke kunstmatige intelligentie* en de toepassing daarvan in *dynamic scripting*. Deze *lerende systemen* baseren zich op de automatische extractie van metadata uit de digitale items (*data mining*, bv. het aantal beats per minuut in een muziekstuk, zie hierboven), op de generalisatie van ervaringen die het systeem opdoet telkens de gebruiker zich over dit bestand uitsprekt (bv. door een track in een reeks systematisch over te slaan), of bijvoorbeeld door te indexeren welke bestanden vaak in elkaars buurt opduiken.¹⁸ Op die manier *leert* het systeem uit het menselijke gebruik en past het zijn eigen gedrag daaraan aan, of suggereert het aanpassingen, oplossingen, verwante items enzovoort. Vanzelfsprekend bieden deze lerende systemen enorme mogelijkheden wat contextualisering betreft.

Een voorbeeld van de toepassing van zo'n lerend systeem op een audiovisuele collectie is de website **Last.fm**. Zoals de webextensie al aangeeft, kunnen gebruikers op deze site luisteren naar *zenders* die niets anders dan hun favoriete muziek laten horen, of zelfs hun eigen online muzikzender beginnen. De metadata waarop dit systeem steunt, zijn automatisch gegenereerde tags. Theoretisch kan achter elke verzameling van tags een muzikzender zitten, die enkel muziek uitzendt met dezelfde metadata en dus dezelfde muzikale kenmerken, afgestemd dus op de smaak van de gebruiker. Terwijl de zender speelt, kan de gebruiker ook zijn mening over elk nummer uitspreken, nieuwe tags toekennen of een link naar het nummer doormailen. Het lerende aspect zit onder andere in het feit dat de door de gebruikers uitgedrukte meningen (*nummer overslaan* of *ik hou van dit nummer*) onthouden worden en dat er verder mee rekening gehouden wordt bij de selectie van de volgende nummers. Als de gebruiker bijvoorbeeld systematisch de nummers met een bepaald kenmerk (tag) overslaat, dan zal de zender de nummers met dezelfde kenmerken (tags) steeds minder spelen.

2.1.2. Expert generated tags

Zoals reeds vermeld moeten metadata, om de opzoekbaarheid van alle items via die metadata te garanderen, zo veel mogelijk systematisch en uniform zijn opgesteld. Een thesaurus kan hierbij als hulpmiddel dienen. Althans, dit is de klassieke logica. In web2.0-toepassingen worden dergelijke strenge vormvereisten als tegengesteld aan het gebruikersbelang beschouwd. Een vorm van metadata-toekenning die helemaal aan die voorwaarde van eenvoud beantwoordt, is *tagging*. Bij *tagging* worden de zogenaamde *tags* toegekend aan de data die bij een zoekopdracht naar boven gekomen zijn. Het gaat om een sleutelwoord dat zonder vormvereisten of andere beperkingen (zoals metadatastandaarden, thesauri) aan de

¹⁸ Uiteraard kunnen de toepassingen van kunstmatige intelligentie zich ook enten op *social tags*.

data wordt toegekend, veelal door de gebruikers zelf (*social tagging*).¹⁹ Ook tags zijn een vorm van contextualisering door metadata-annotatie, aangezien het klikken op een tag de gebruiker doorverbindt met een aantal items waaraan diezelfde tag is toegekend. Collectiedelen worden dus gelinkt via een tag, wat verrassende resultaten kan opleveren: er ontstaat een context van items die op welk vlak dan ook een eigenschap delen, die door de toekenners ervan relevant werd geacht.

In de praktijk worden de termen *tagging* en *social tagging* vaak door elkaar gebruikt. Toch is tagging niet noodzakelijk *user generated*. Een voorbeeld van *expert generated tags* zijn te vinden op **Klara.be**.²⁰ De redacteurs van deze site baseren zich voor de toekenning van hun tags losjes op de (thesaurusgebonden) trefwoorden van de VRT-archieven, maar maken wel gebruik van de vrijheid die tagging biedt. Niettegenstaande dit vrije concept is er ook hier sprake van zelfregulering: er zijn binnen de redactie vaste afspraken omtrent soorten tags die zeker vermeld moeten worden: een categorie (bv. muziek, literatuur, expo, ...), wat specifieker (bv. schilderkunst, fictie, ...), de geïnterviewde, de vorm waarin de essence wordt voorgesteld (bv. interview, documentaire, reportage), enzovoort.



Fig. 4: screenshot met *expert generated tags* op Klara.be.

¹⁹ Voor een bredere kadering van het fenomeen *tagging*, zie het rapport dat binnen het Werkpakket 1 Gebruikersnoden – open en dynamisch archief werd gemaakt: https://projects.ibbt.be/bom-vl/fileadmin/user_upload/frontendfiles/BOM-VI%20WP1.1%20Open%20en%20dynamisch%20archief.pdf

²⁰ Zie <http://www.klara.be>.

2.2. Verbinding met andere content en randdocumentatie

De volgende stap bestaat uit het verbinden van de archiefitems met andere content of randdocumentatie, zoals uitgeschreven teksten van audio-opnames, documentatie over het ontstaan van het collectiedeel, verduidelijkingen over de rechten enzovoort.²¹ Die andere items en randinformatie kunnen zich zowel binnen als buiten de eigen collectie bevinden. Zo kunnen interessante dwarsverbanden worden gelegd tussen collecties van verschillende musea en instellingen. Dit is onder andere de doelstelling van het project **Nederlands erfgoed: digitaal!**, dat content van tien Nederlandse erfgoedinstellingen wil beschikbaar maken via één portaal.

2.2.1. Inhoudelijke contextualisering

Met het verbinden van content voegt men inhoudelijke betekenis toe. Content en randdocumentatie kunnen thematisch of chronologisch worden aangeboden. Op de website van **Video Active** bijvoorbeeld, krijgt men via de webpagina *'Hoofdthema's uit de Europese geschiedenis'* beeldfragmenten per thema aangeboden, zoals *'Nationale feestdagen'* of *'Mode'*. **L'Institut national de l'audiovisuel** (INA) biedt op de webpagina voor het grote publiek content aan per thema, periode en volgens de visueel aantrekkelijke *jalons* of *fresques interactives*. Via de *'jalons pour l'histoire du temps present'* worden de beeldfragmenten volgens verschillende thema's, zoals *'la première guerre internationale'*, op een tijdbalk weergegeven.

Audiovisuele content wordt ook vaak illustratief gebruikt op websites die bijvoorbeeld historische en erfgoedberichten ontsluiten en op die manier die content - die ze weliswaar van andere platformen halen - contextualiseren op hun eigen site. Een treffend voorbeeld van deze werkwijze is **www.anno.nl**. Deze website verzamelt historische berichten uit de actualiteit van op verschillende bronnen. Ze breidt deze berichten uit met (geschreven) historische achtergronden en gaat ook actief op zoek naar allerlei audiovisuele illustraties, die dan bij het bericht geplaatst worden. Dit illustreert dat de collectiebeheerder niet altijd zelf verantwoordelijk hoeft te zijn voor de contextualisering, maar dat door die content eenvoudig beschikbaar te maken, de contextualisering ook door anderen op een gedegen manier kan gebeuren.

²¹ Andere voorbeelden: een samenvatting, een playlist, medewerkers, contactgegevens, een bredere collectieomschrijving, enz.

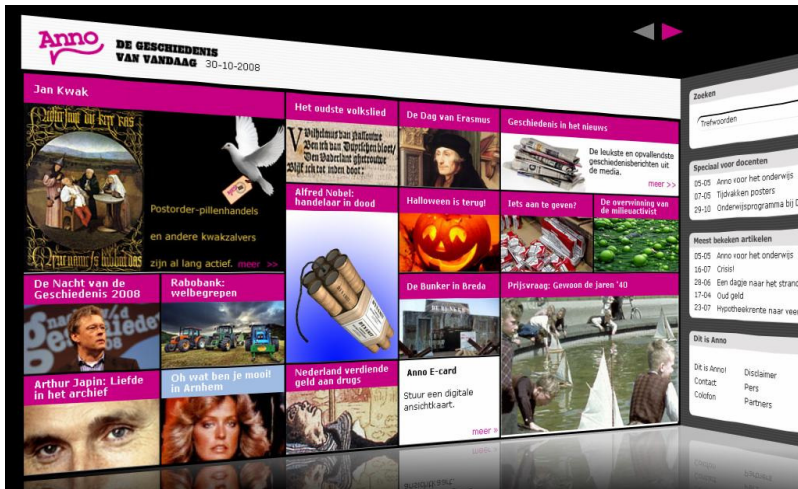


Fig. 5: Berichten in verband met geschiedenis in de actualiteit, onder andere geïllustreerd met audiovisueel archiefmateriaal op www.anno.nl.

2.2.2. Doelgroepgerichte contextualisering

Met doelgroepgerichte contextualisering verstaan we het betekenisvol ontsluiten van content met aandacht voor de noden van een specifieke doelgroep. Verschillende gebruikersgroepen hebben andere noden, wensen en verwachtingen. Het inspelen op deze noden, wensen en verwachtingen kan gebeuren ter hoogte van de collectie zelf, maar ook granulair, ter hoogte van de losse archiefitems dus.²²

In principe kan elke vorm van contextualisering doelgroepgericht worden ingevuld. Zo moeten bijvoorbeeld de beschrijvende metadata de context waarin de informatie wordt gebruikt reflecteren. Dat zorgt voor verschillen in onder meer beschrijvingsstandaarden, uitgebreidheid, tijdcodering, contextinformatie, schrijfstijl. Een goed voorbeeld van doelgroepgerichte metadata-annotatie is de **iMMix-catalogus van het Nederlands Instituut voor Beeld en Geluid**. Deze vertrekt vanuit het principe dat het vaststellen en verklaren van de context en inhoud van archiefmateriaal nauwkeurige en adequate representaties vereist. De normen voor archivalisch beschrijven zijn bovendien gebaseerd op algemeen aanvaarde theoretische beginselen. Toch acht men het mogelijk en zelfs wenselijk om de metadata (gedeeltelijk) af te stemmen op specifieke doelgroepen en in die zin te contextualiseren. De iMMix-catalogus maakt voor die contextualisering gebruik van zgn. *intenties*: andere versies van een beschrijving, waarin een aantal velden op een afwijkende manier wordt ingevuld. Zo ontstaat de mogelijkheid om domeinspecifieke en/of additionele gegevens aan de eerst gemaakte beschrijving te hangen en gegevens autonoom in te voeren en te presenteren. Het doel is uiteraard om in te spelen op specifieke informatiebehoeften van verschillende

²² Zie supra: 2.1.1. Klassieke metadata-annotatie en metadatastandaarden. Bijvoorbeeld de modules (collectieniveau) en de zgn. *intenties* in het metadatamodel van Beeld en Geluid.

gebruikersgroepen. Gebruikers mogen daarbij niet meer of minder gegevens op hun scherm krijgen, dan waaraan ze behoefte hebben. Dit is wat de Jong de *functionele granulariteit* noemt.²³ Hoewel er in dit geval niet enkel metadata worden toegevoegd, maar ook weggelaten, kunnen doelgroepgerichte metadata ook als een vorm van contextualisering worden beschouwd.

Doelgroepgerichte contextualisering kan gebeuren door de collectiebeheerder zelf, die de content op een webportaal aanbiedt aan één of meerdere doelgroepen, anderzijds kan een doelgroepgericht webportaal content uit één of meerdere collecties verzamelen. **Teleblik** is een mooi voorbeeld van dergelijke doelgroepgerichte contextualisering. Teleblik stelt audiovisuele bronnen uit de archieven van de Nederlandse publieke omroepen en het Polygoon Journaal (bioscoopjournaal) online ter beschikking van het primaire en voortgezet onderwijs in Nederland. Per onderwijsniveau krijgt men thema's aangeboden (zoals WOI, de winter, evolutie, enz) die een verzameling aan beeldfragmenten bevatten. Docenten en leerlingen kunnen het audiovisueel lesmateriaal downloaden, bewerken, opnieuw samenstellen, versturen enz. Teleblik put haar voorraad aan filmpjes uit de collecties van het Nederlands Instituut voor Beeld en Geluid. Bovendien kan iedereen in de catalogus van Beeld en Geluid ook nog op zoek gaan naar nieuwe items die nog niet gedigitaliseerd zijn en zo opname in Teleblik suggereren.

Het webportaal van het **Nederlands Instituut voor Beeld en Geluid** is een mooi voorbeeld van een audiovisuele collectie die via verschillende webpagina's (*Experience, Educatie, Webshop, Archieven* en *Instituut*) wordt aangeboden, die centraal toegankelijk zijn via de portaal-site. De zoekfuncties van de verschillende webpagina's zoeken weliswaar in dezelfde database, maar geven de resultaten anders (eventueel gefilterd) weer, en wel op een manier die specifiek tegemoet komt aan de noden en wensen van een bepaalde doelgroep. Aan de basis hiervan ligt de identificatie van de gebruikerstypes en de rollen die zij tijdens hun bezoek aan de website spelen.

2.3. Contextualisering door de gebruiker

In de klassieke archiefkunde wordt van de annotator zo veel mogelijk objectiviteit verlangd. Het wordt niet als zijn taak beschouwd om het archiefmateriaal te interpreteren of er zijn mening over te geven. In het web2.0-tijdperk staat de mening van de gebruiker echter centraal. Deze subjectieve meningen worden op allerlei manieren verwerkt in de ontsluiting.

De contextualisering van archiefmateriaal via de meningen van de gebruikers gebeurt onder andere door *tagging*, *rating* en op *fora*. Deze contextualiseringsvormen betekenen voor de gebruiker een manier om zijn interpretatie en appreciatie van het collectiedeel te kennen te geven aan anderen. Bij kwantitatieve formuleringen van deze appreciatie zoals rating, is het

²³ DE JONG 2007, 26.

voor de gebruiker eenvoudig en aantrekkelijk, indien hij hiervoor de kans krijgt, om zijn zoekresultaten te rangschikken op de gemiddelde rating-score die door anderen gegeven is.

2.3.1. User generated tags

Zoeksystemen zochten vroeger enkel naar exacte overeenkomsten, waardoor het belangrijk was om de metadata gestructureerd en eenduidig, volgens een omvangrijke set van regels aan te brengen. In het web2.0-tijdperk wordt aan de exacte overeenkomst een nieuwe manier van zoeken toegevoegd, namelijk *tagging*. Via *tagging* kan de gebruiker de collectie zelf omschrijven en krijgt hij/zij de kans zijn/haar mening te uiten. Bovendien kan er beroep worden gedaan op de geïnteresseerde gebruiker om de metadata te verrijken. Tags geven de gebruikers de kans om, voor zichzelf of ook ten dienste van anderen, de bestaande metadata aan te vullen, zodat ook een zoekopdracht die zich niet aan de beperkingen van de klassieke metadata houdt, toch resultaat kan opleveren.²⁴

De functionaliteiten rond toekenning van en zoekmogelijkheid via *tags* komen tegemoet aan de veel voorkomende klacht dat zowel metadata als zoeksystemen te complex en te ontoegankelijk zijn opgesteld. De oorzaak daarvan wordt gelegd bij de databaseontwikkelaars zelf, die er als deskundigen onterecht van uit zouden gaan dat (de occasionele en zelfs de geofende) gebruikers in de eigen collectie even ingewikkelde zoekvragen weten te stellen als zijzelf. Het succes van *tagging* steunt op de vooronderstelling dat de beste metadata degene zijn waarop de gebruikers het meest zoeken en dat de gebruikers dan ook deze metadata zullen toekennen als ze daartoe de kans krijgen. In de klassieke archiefkunde staat de juistheid van de metadata centraal. De voorstanders van *tagging* komen hieraan tegemoet door te wijzen op de zogenaamde *Wet van Linus*, die stelt dat als iedere gebruiker de kans heeft verbeteringen aan te brengen, op de duur alle fouten verbeterd zullen worden.²⁵

Voorbeelden van *user generated tags* (ook *social tagging*, *collaborative tagging* genoemd) bij audiovisuele collecties zijn legio, denk maar aan youtube.com en haar talrijke klonen. Uit de museum- en archiefwereld is het Amerikaanse *steve*-project een goed voorbeeld.²⁶ Dit project liet de gebruikers tags toekennen aan de items op de website, waarna de resultaten werden onderzocht. Het bleek dat 86% van de tags niet voorkwamen in de aanwezige museumdocumentatie. De overgrote meerderheid van de tags, namelijk 88.2%, werden als *useful* omschreven door het museum personeel.²⁷

We moeten hier wel de kanttekening bij plaatsen dat rond de meeste web2.0-gerelateerde vormen van contextualisering discussies zijn ontstaan. Daarbij is de centrale vraag of deze nieuwe vormen de content beter toegankelijk maken en of daarbij de voordelen wel opwegen

²⁴ In de vele definities van *social tagging* blijven de factoren *informeel* en *user-defined* steeds overeind. Zie ook: MELENHORST et al. 2008.

²⁵ Zie ook: MELENHORST et al. 2008.

²⁶ Zie <http://steve.museum>.

²⁷ TRANT 2009.

tegen de nadelen. Bovendien kan men zich vragen stellen over de duurzaamheid van user generated tags in tijden van snel evoluerend taalgebruik. In de marge van deze discussie rijst ook de vraag of de verdeling *experts* versus *gebruikers* geen te eenvoudige voorstelling van zaken is. Ook gebruikers kunnen immers een inhoudelijke expert zijn in het onderwerp waarrond ze zoeken. Het is hier niet de bedoeling opnieuw de argumenten pro en contra op te sommen. Hiervoor wordt verwezen naar het werk van Laura Gordon-Murnane en naar deeltaak 1 van WP1 (open en dynamisch archief).²⁸ De discussie heeft echter geleid tot een vrij breed gedragen consensus over de status van de *social tags* en andere vormen van contextualisering uit de web2.0-sfeer.²⁹ Deze moeten gezien worden als aanvullend bij de *expert generated metadata*, niet als de vervangers ervan. Op die manier zouden databases en zoekmethoden het beste van beide werelden kunnen combineren.

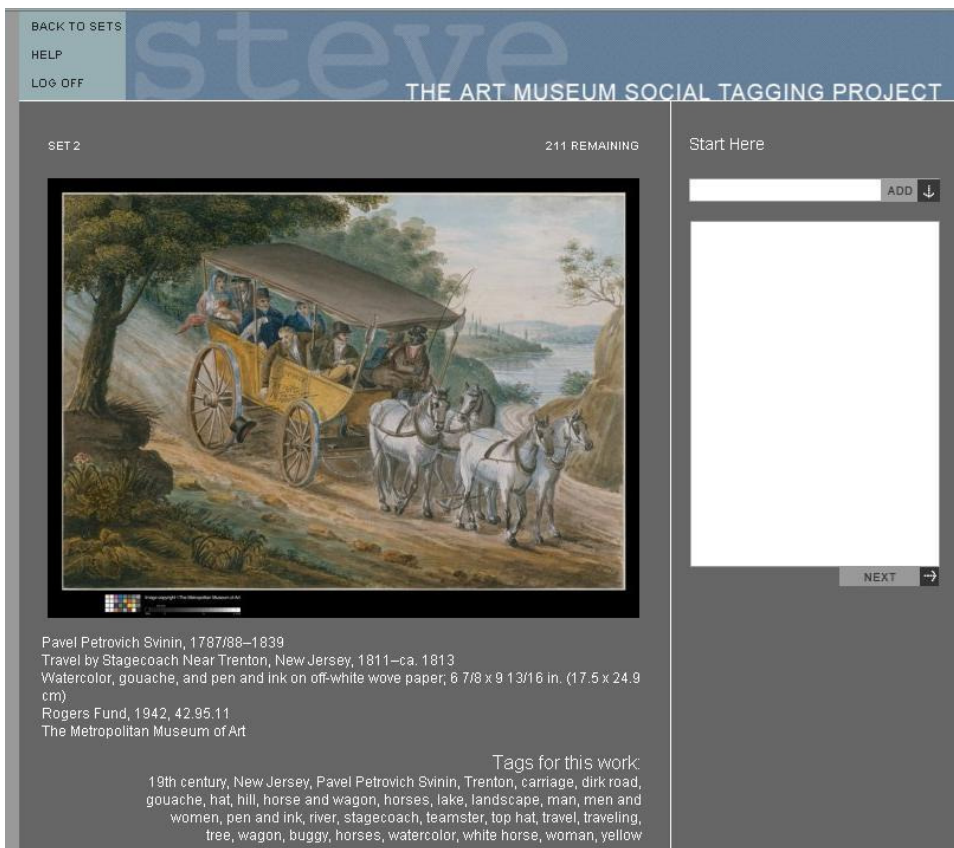


Fig. 6: screenshot van de toepassing voor *user generated tagging* op www.steve.museum.

2.3.2. Folksonomies

²⁸ GORDON-MURNANE 2006.

²⁹ Stilaan bewijzen meer en meer studies op een wetenschappelijke manier het nut van *social tagging* voor de effectieve opzoekbaarheid. Zie bijvoorbeeld VAN HOOLAND 2006 of MELENHORST et al. 2008.

Klassiek zijn databases en thesauri geordend volgens taxonomieën. Dit zijn subtype-supertype-hiërarchieën die steunend op gestandaardiseerde definities van de database-onderdelen stambomen vormen. De web2.0-sfeer heeft op deze taxonomieën haar klassieke kritiek van (het gebrek aan) gebruiksvriendelijkheid toegepast en als alternatief de *user generated taxonomies* voorgesteld, beter bekend als *folksonomies*. In die zin zijn *folksonomies* eigenlijk een gevorderde versie van *user generated tags*: er wordt niet enkel *getagged*, de *tags* zelf worden ook automatisch organisch geordend, zodat er als het ware een nieuwe structuur ontstaat. Veelal werken verschillende gebruikers samen aan dezelfde *folksonomy*. Opnieuw uitgaand van de Wet van Linus wordt het gebrek aan kwaliteitscontrole daarbij omzeild.

Bij websites die de mogelijkheid tot *tagging* hebben, verschijnt de folksonomy vaak als een zogenaamde *tag cloud*. Daarbij komen de vaakst toegekende *tags* samen in een schijnbaar ongestructureerde groep (*wolk*) bij de zoekvelden. Deze tag clouds zijn erg populair, maar er is op het internet een levendige beweging ontstaan die zoekt naar meer creatieve voorstellingen van folksonomies. Die nieuwe voorstellingen trachten via steeds meer visuele kenmerken, elk gekoppeld aan criteria, het belang van de respectieve tags voor te stellen, zodat een steeds beter beeld van het onderliggende semantische landschap ontstaat. In de eerste tag clouds varieerde de grootte van het lettertype volgens het aantal keer dat de tag was toegekend. Al snel kwam daar de afstand tot het midden van de wolk bij. Die duidde meestal aan hoe vaak er op die term werd gezocht. Kleur en de helderheid gingen vervolgens variëren en tegenwoordig behoort zelfs een driedimensionale voorstelling, bijvoorbeeld als een ingekleurd berglandschap, tot de mogelijkheden.

Folksonomies zijn een vorm van contextualisering omdat ze de meest voorkomende –en soms zelfs alle- tags groeperen. Daarmee ontstaat een visualisatie van trefwoorden die gelden voor de hele collectie, in de mate natuurlijk dat de tags op een representatieve manier zijn toegekend. De manier waarop deze folksonomies de collectie beschrijven, wordt weleens vergeleken met een *camera obscura*: het zijn gefilterde visualisaties van een multidimensionale wereld.³⁰ Daarnaast brengt een muisklik op een tag de gebruiker naar een lijst van alle items die deze tag dragen. De gebruiker krijgt via folksonomies dus enerzijds een zicht op de hele collectie als context van één item en anderzijds op groepen van items die via eenzelfde tag met elkaar verbonden zijn.

Een voorbeeld van een visuele variant van de *tagcloud* is de Aquabrowser-toepassing van het Nederlandse bedrijf Medialab. Onder meer de **Vlaamse portaalsite van de openbare bibliotheken** maakt hiervan gebruik.³¹ Deze toepassing stelt de gebruiker een waaier aan termen voor die gerelateerd zijn aan de eerste zoekterm: associaties, spellingsvarianten (woorden met grotendeels dezelfde letters), thesaurustermen en vertalingen. De afstand tot

³⁰ Naar de woorden van informatie-architect Joe Lamantia, http://www.joelamantia.com/blog/archives/ideas/tag_clouds_a_new_user_interface.html.

³¹ Zie <http://zoeken.bibliotheek.be>.

het centrum van de waaier duidt bovendien aan hoe vaak de gerelateerde termen voorkomen in de buurt van de oorspronkelijke zoekterm.

2.3.3. Rating

Net als *tagging* is *rating* een bijzonder eenvoudige en klantgerichte vorm van metadata-annotatie. Ook *rating* gebeurt nagenoeg steeds door de gebruikers zelf. Het steunt *mutatis mutandis* op dezelfde assumptie als *tagging*: de interessantste items uit een database zijn deze die door de gebruikers als dusdanig worden aangewezen. Dat aanwijzen gebeurt door het (al dan niet) toekennen van een schaalgebaseerde score aan één of meerdere collectiedelen, in de eenvoudige visualisatie van een zoekresultaat. Zo kan de gebruiker bij een volgende zoekopdracht zien welke items (door hemzelf of door anderen) het meest gewaardeerd zijn. Samengevat komt *rating* neer op een spontane populariteitspoll van items uit de database. Waardering is daarbij veruit het meest voorkomende criterium, maar theoretisch is het uiteraard mogelijk de gebruiker te laten oordelen op eender welk ander criterium. Contextualisering aan de hand van *rating* kan gebeuren door de ordening volgens populariteit van de zoekresultaten. Als kwaliteit en populariteit werkelijk overeenkomen, dan wordt het zoekresultaat aangeboden in de context van gelijkwaardige zoekresultaten. Het meest bekende voorbeeld van een website die gebruikers aanbiedt collecties te raten is Youtube.³² Aangemelde gebruikers krijgen hier de kans om elk filmpje dat ze bekijken een score te geven van maximaal vijf sterren. Een lijst met zoekresultaten kan ook op deze "waardering" gesorteerd worden.

2.3.4. Fora

Een andere vorm van contextualisering die in het web2.0-tijdperk een hoge vlucht heeft genomen, zijn de fora waarop hetzij op itemniveau, hetzij op collectieniveau commentaren worden toegekend. Fora bieden de gebruiker dus de gelegenheid de content te becommentariëren. Voorbeelden van websites die een forum aanbieden zijn legio. **Radio1.be** bijvoorbeeld biedt bij elk onderwerp uit elk programma de mogelijkheid om het audiofragment te herbeluisteren en te reageren.³³ Registreren of aanmelden is hier in tegenstelling tot bij Youtube.com niet nodig. Fora zijn een vorm van contextualisering omdat ze de essence koppelen aan de commentaren van de gebruikers. Zij geven hun persoonlijke mening over de inhoud, ze geven extra informatie, ze corrigeren de bestaande metadata (weliswaar niet in de metadatavelden), stellen vragen of suggereren interne of externe links. Al deze informatie plaatst de content in een breder kader. In die zin kunnen fora multimediale erfgoedcollecties contextualiseren. Fora zoals dat op Radio1.be leggen de nadruk op de mening van de gebruiker. Dit soort van fora is expliciet bedoeld om de betrokkenheid van de gebruiker bij het aanbod te vergroten. Maar fora kunnen ook een doelstelling hebben die uitgaat van de

³² <http://www.youtube.com>.

³³ <http://www.radio1.be>.

collectie zelf. Collectiebeheerders beschikken niet altijd over gedetailleerde metadata en manuele metadata-annotatie is ook erg arbeidsintensief. Daarom kunnen fora ingezet worden om de metadata te verrijken met de kennis van de gebruiker. Een voorbeeld hiervan is de **beeldbank van het Nederlands Nationaal Archief**,³⁴ die de kennis van de gebruiker met groot succes heeft ingezet. Zij verantwoordt deze manier van werken als volgt: *“Alhoewel de kwaliteit van de beschrijvingen van de foto's niet optimaal is - soms zijn deze niet correct, soms onvolledig - heeft het Nationaal Archief er toch voor gekozen zijn fotocollectie online te plaatsen. De foto's zijn voor te velen interessant om nog langer alleen in de depots te bewaren. Wel wordt inmiddels gewerkt aan het verbeteren van de beschrijvingen. U kunt ons daarbij helpen: indien u een aanvulling heeft op een beschrijving, dan kunt u die aan ons doorgeven via de link 'commentaar geven' (naast elke afbeelding). Ook kunt u vrijwilliger worden bij de fotocollectie van het Nationaal Archief en meewerken aan het optimaliseren van de toegankelijkheid.”*³⁵ We moeten hierbij wel vermelden dat het hier niet gaat om een forum in de strikte zin van het woord, maar om een UGM-annotatie die in sommige gevallen is uitgegroeid tot wat je een forum zou kunnen noemen (post en wederpost, met commentaar niet op de essence maar op de post van de anderen).

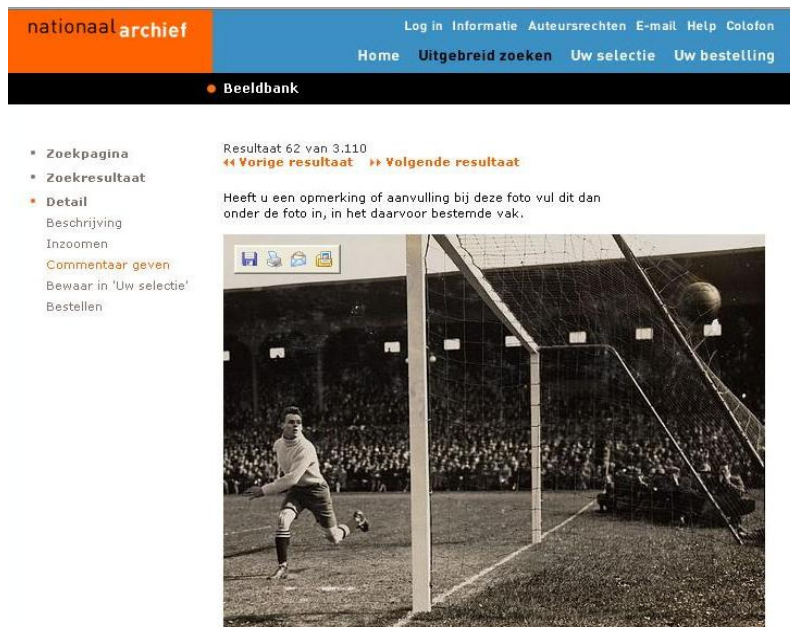


Fig. 7: formulier waarmee gebruikers commentaren kunnen invoeren bij de foto's van de beeldbank van het Nederlands Nationaal Archief.

³⁴ Zie <http://beeldbank.nationaalarchief.nl>. De mogelijkheden van dit initiatief zijn ook wetenschappelijk geëvalueerd in VAN HOOLAND 2006, 8.

³⁵ http://beeldbank.nationaalarchief.nl/index.php?option=com_memorix&mr_x_mod=content&mr_x_item=verantwoording.

Het is overigens opvallend dat er ook nogal wat websites die audiovisuele data ontsluiten er juist niet voor kiezen een forum aan te bieden, dat forum goed afschermen³⁶, of de geposte commentaren slechts na controle publiceren. Deze keuze heeft in veel gevallen te maken met de nadelen van het erg publieke karakter van fora. De collectie zelf loopt immers het risico geassocieerd te geraken met het negatieve imago van het forum, indien dit te veel geplaagd wordt door beledigingen, schuttingtaal, vandalisme, commerciële berichten, enz.

2.3.5. Collecties

Met een collectie wordt hier de mogelijkheid bedoeld die de gebruiker krijgt om zelf een deelselectie samen te stellen uit de volledige collectie. Zo kan die gebruiker helemaal volgens eigen smaak items groeperen en bewaren. Hij krijgt als het ware de kans om zelf curator te spelen van zijn eigen hoogst persoonlijke tentoonstelling van filmpjes, liedjes, foto's, speeches of andere audiovisuele content. Collecties zijn de eenvoudigste vorm van het bundelen van onderdelen van grotere verzamelingen. Het staat de samensteller volledig vrij het onderling verband tussen de items te bepalen. Vaak gaat dat verband nauwelijks verder dan een zekere appreciatie bij de samensteller voor alle items die van de verzameling deel uitmaken.

Een interessant voorbeeld hiervan is de site van de Londense **Tate Gallery**.³⁷ Op de deelsite *Your collection* krijgt de bezoeker eerst een aantal voorbeelden te zien van deelcollecties die de Tate zelf suggereert. Ze hebben meestal een originele invalshoek, die niet vertrekt bij de inhoud, maar bij de gebruiker: geen pre-rafaëlieten-collectie dus, maar een *'I haven't been there for ages-collectie'*. Na die voorstelling krijgt de gebruiker de kans om zijn eigen deelcollectie samen te stellen uit een voorlopig nog vrij beperkt aanbod aan werken uit de Tate Gallery. Paradoxaal genoeg vertrekken de collecties die de gebruikers zelf samengesteld hebben, wél vaak van de inhoud zelf: de *history of hair*-collectie bijvoorbeeld. Gebruikers blijken dus niet altijd even creatief als de *web2.0-minded* collectiebeheerders wel zouden willen. De verrassende verbindingen zoals die bij *social tagging* soms ontstaan, blijven voorlopig dus wel eens uit. Eens de collectie is samengesteld kan de gebruiker zijn collectie ook annoteren en aangeven waar hij zelf het verband tussen de geselecteerde werken legt. De bedoeling is dat dit tot interessante invalshoeken leidt. Ten slotte kan er ook een kleine *catalogus* van de collectie worden afgedrukt. Het concept *Your collection* is weliswaar een bescheiden toepassing van een mogelijkheid die nog veel verder uitgebouwd zou kunnen worden, maar niettemin weerspiegelt het een nieuwe manier van gebruikersgerichte erfgoedpresentatie. Niet voor niets staat er op het achterflapje van de hoogstpersoonlijke collectiecatalogus: *"Yes it's a museum. But it's also a living room. All those works of art are yours."* Op de website van *Click! Photography changes everything* van het **Smithsonian**

³⁶ Door middel van een verplichte registratie en aanmelding bijvoorbeeld, of een Captcha. Dit laatste is een enkel door de mens herkenbare tekenreeks die de gebruiker moet invoeren als sleutel tot het forum. Dit om te voorkomen dat er automatische (bijvoorbeeld commerciële) berichten op het forum geplaatst zouden worden.

³⁷ Zie <http://www.tate.org.uk/britain/yourcollection>.

Photography Initiative nodigt men onderzoekers én bezoekers uit om foto's en verhalen te posten over hoe fotografie onze cultuur en ons leven beïnvloedt.

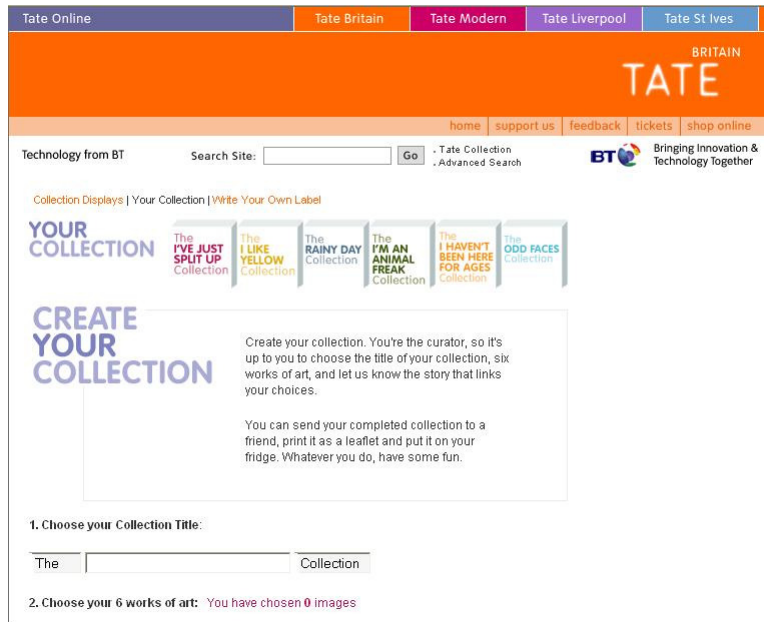


Fig. 8: screenshot van Your Collection op de website van de Tate Gallery.

2.3.6. Social bookmark managers

Steeds meer websites bieden de gebruikers de mogelijkheid aan om binnen één collectie content te groeperen en zo als het ware eigen deelcollecties aan te leggen. De *social bookmark managers* doen eigenlijk hetzelfde, maar over verschillende collecties heen. Zij laten de gebruiker toe zijn eigen portaalwebsite te creëren met toegang tot alle websites of items die hijzelf interessant vindt. In feite zijn de social bookmark managers in die zin niet meer dan een online versie van de favorietenlijst zoals die in elke internetbrowser kan worden aangelegd. Het *social*-aspect bestaat erin dat die favorietenlijst online ook voor andere gebruikers toegankelijk is. Elke link kan bovendien worden getagged, zodat ze gemakkelijk kan worden teruggevonden. Op die manier kan elke gebruiker kennismaken met de links die een andere gebruiker met dezelfde interesses gelegd heeft en zo mogelijk nieuwe websites ontdekken.

Social bookmark managers zijn een vorm van contextualisering aangezien ze links naar audiovisuele content groeperen en verspreiden. Het is met andere woorden een vorm van contextualisering die door de gebruiker gebeurt en volledig buiten het veld van de collectiebbeerder valt. Eén van de bekendste social bookmark managers van dit ogenblik is del.icio.us:

De collectiebeheerder kan de toegankelijkheid van zijn collectie ook vergroten door de gebruiker toe te laten het audiovisueel materiaal zelf te contextualiseren. Websites zoals **addthis.com** bieden *social bookmark sharing buttons* aan. Dit zijn knoppen (webware zichtbaar als een knop) die de collectiebeheerder zelf onder zijn content kan zetten en die de gebruiker toelaten snel naar deze content te linken van op zijn eigen website. Bovendien – en

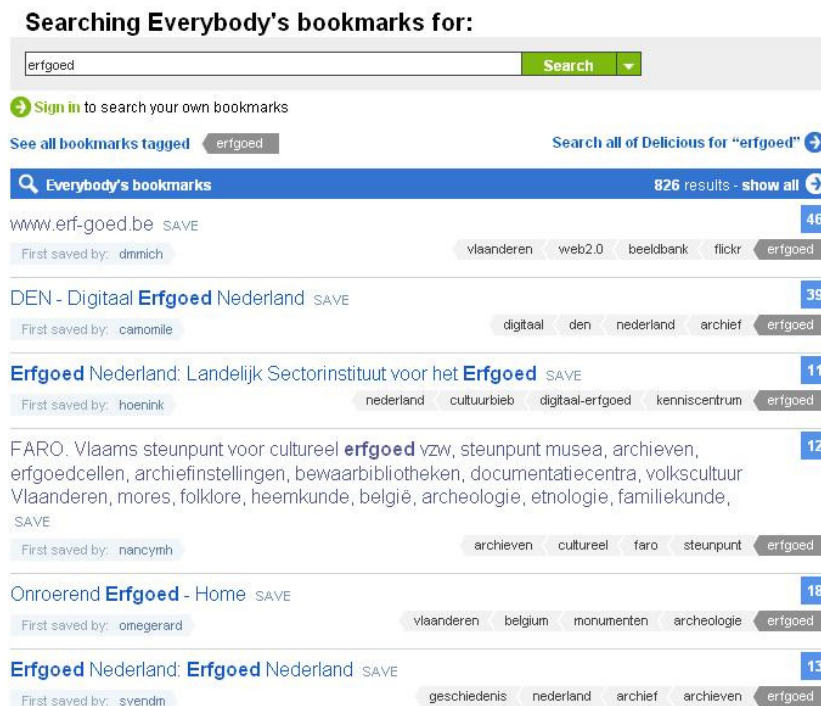


Fig. 9: screenshot van de resultatenlijst na de zoekopdracht *erfgoed* op del.icio.us.

dat is voor collectiebeheerders nog interessanter – houdt addthis.com op elke centrale gebruikersaccount bij wanneer, door wie en hoe vaak en tussen welke sites er gelinkt wordt. De collectiebeheerder krijgt dus als het ware een zicht op de verspreiding van de content en dus ook op de wegen waarlangs surfers zijn collectie bereiken. Gebruikers die via de addthis-knoppen naar een collectie-item linken, zorgen met hun link elke keer opnieuw voor een nieuwe contextualisering.³⁸

³⁸ www.addthis.com.

2.3.7. Externe gebruikersplatformen

Externe gebruikersplatformen bieden content aan buiten de context van een klassieke archiefcollectie. Enerzijds kan de gebruiker zelf langs deze weg de toegankelijkheid vergroten door, wanneer het door de collectiebeheerder wordt toegelaten, het collectie-item ergens te posten en zo andere gebruikers op de hoogte te brengen. Anderzijds kan de collectiebeheerder zelf (delen van) zijn eigen collectie op deze platformen plaatsen om ze toegankelijk te maken, bijvoorbeeld om zo user generated tags te verzamelen.

Deze laatste manier van gebruik van externe gebruikersplatformen wordt momenteel onder andere in praktijk gebracht door het **Nederlands Nationaal Archief**. Deze biedt als testcase delen van haar fotocollectie aan op Flickr the Commons.³⁹ Met deze pilot willen ze ten eerste hun collectie internationaal beschikbaar stellen, de publieksparticipatie bevorderen, de fotocollecties verrijken met informatie van gebruikers en kennis delen rondom 'social tagging' voor archieven. Na de testperiode zal worden nagegaan of de metadata die werden toegevoegd door de gebruikers bruikbaar zijn voor de eigen beeldbank.

Uiteraard kan deze manier van ontsluiten ook voor andere doelen worden gebruikt. Een onderwijsplatform zou bijvoorbeeld audiovisuele content kunnen verzamelen om ze in een educatieve context ter beschikking te stellen aan leerlingen en leerkrachten. Of een toeristisch audiovisueel platform zou op die manier alle interessante content in verband met het gebied dat toeristisch gepromoot wordt kunnen bundelen. Een audiovisueel digitaal repository moet ook op dit soort toepassingen kunnen inspelen door de voorwaarden te bepalen welke content op welke manier voor welke soort van *delende* toepassingen gebruikt mag worden.

*The only way to control your content is to be the best provider of it.*⁴⁰ Het is een omstreden uitspraak, maar vast staat dat als audiovisuele archiefcontent vrijelijk verspreid wordt, de controle over die content voor de collectiebeheerder snel verloren is. De archiefitems beginnen als het ware een eigen leven op het internet. Door de eigen content op een degelijke en goed gecontextualiseerde manier aan te bieden, kan de eigen collectiewebsite zich manifesteren als de beste aanbieder van deze content, wat de gebruiker automatisch naar deze website zal lokken. Wie in deze uitdaging slaagt, kan net zoals NRK Media een eigen Youtube-kanaal openen, zonder het risico te lopen dat de eigen collectie daardoor nooit meer online wordt bezocht.

³⁹ <http://www.flickr.com/photos/nationaalarchief>.

⁴⁰ Naar het devies van NRKbèta, het experimentele mediaplatform van de Noorse openbare omroep NRK, zoals uitgesproken door Eirik Solheim, developer bij NRK Media.

3. Cases

In dit hoofdstuk bespreken we kort enkele onderzoeksprojecten die bepaalde aspecten van contextualisering onderzoeken en overlopen we enkele praktijkvoorbeelden. We beperken ons hierbij niet tot online audiovisuele archieven maar nemen ook relevante voorbeelden op van initiatieven die beeldende kunst gecontextualiseerd aanbieden.⁴¹

3.1. Onderzoeksprojecten

3.1.1. CATCH (Continuous Access to Cultural Heritage)

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

CATCH (Continuous Access to Cultural Heritage) is een NWO onderzoeksprogramma (Nederland) dat methoden en technieken wil ontwikkelen die collectiebeheerders in staat moeten stellen hun collecties beter digitaal toegankelijk te maken. CATCH is een samenwerkingsverband tussen zes grote Nederlandse culturele erfgoedinstellingen, zoals het Rijksmuseum en de Koninklijke Bibliotheek, en de onderzoeksgebieden Exacte Wetenschappen en Geesteswetenschappen. Het onderzoeksprogramma wil het mogelijk maken om nieuwe verbanden en relaties zichtbaar te maken tussen verschillende collecties en wil de interactie tussen gebruikers en bronnen optimaliseren. Er lopen momenteel tien projecten die zich richten op de verbetering van kruisbestuiving tussen wetenschappelijk onderzoek en cultureel erfgoed, waaronder CHIP, CHOICE en MUSEUM.

3.1.1.1. CHIP (Cultural Heritage Information Presentation)

<http://www.chip-project.org/>

CHIP (Cultural Heritage Information Presentation) is een samenwerkingsverband tussen de Technische Universiteit Eindhoven, Rijksmuseum Amsterdam en het Telematica Institute. Het doel van CHIP is gepersonaliseerde toegang tot gecombineerde cultureel erfgoed content mogelijk te maken, met andere woorden artefacten en achtergrondinformatie van verschillende collecties en bronnen combineren en doorzoekbaar maken voor de gebruiker. Daarvoor integreert CHIP bestaande technologieën zoals *semantic browsing*, *semantic search* en *ontology*-gebaseerde *user modeling*. De voornaamste case die wordt uitgewerkt is de museumwebsite van het Rijksmuseum Amsterdam (Chip @ Rijksmuseum Amsterdam). In deze demonstrator zijn tools uitgewerkt zoals de *art recommender* en de *tour wizard*, die er respectievelijk voor moeten zorgen om de gebruiker items uit te collectie te laten terugvinden die overeenkomen met zijn interesse en om de gebruiker een gepersonaliseerde museumtour te laten samenstellen. Dit gebaseerd op een web-gebaseerde *rating dialog*.

⁴¹ Sommige van deze cases werden reeds besproken in het kader van Werkpakket 1, Deeltaak 1 – Open en dynamisch archief: zie VAN PASSEL & BEYL 2008: https://projects.ibbt.be/bom-vl/fileadmin/user_upload/frontendfiles/BOM-VI%20WP1.1%20Open%20en%20dynamisch%20archief.pdf

3.1.1.2. CHOICE (Charting the Information Landscape Employing Context Information)

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

Aan CHOICE (Charting the Information Landscape Employing Context Information) wordt gewerkt door onderzoekers van het Max Planck Instituut, de Vrije Universiteit en het Telematica Instituut en het wordt uitgetest op het archief van het Nederlands Instituut van Beeld en Geluid. Het belangrijkste doel van dit project is het uittesten van de mogelijkheden van semi-automatische semantische annotatie van gearchiveerde objecten aan de hand van informatie uit bijvoorbeeld thesauri of contextuele tekstbronnen over radio- of televisieprogramma's zoals websites, online televisiegidsen, recensies en kijkcijferonderzoek. Het CHOICE-project wil uiteindelijk een raamwerk creëren voor het trefwoordsuggestiesysteem: de CHOICE Documentalist Support-omgeving. Hiertoe wordt momenteel gebouwd aan prototypes voor een database voor contextdocumenten en voor een geïntegreerde zoekomgeving voor metadata, teksten, semantische annotaties en audiovisueel materiaal.

3.1.1.3. MUSEUM (MULTIPLE collection-SEARCHING Using Metadata)

<http://staff.science.uva.nl/~kamps/museum/>

MUSEUM (MULTIPLE collection-SEARCHING Using Metadata) is een project van de Universiteit van Amsterdam, het Gemeentemuseum Den Haag, het Rijksbureau voor Kunsthistorische Documentatie en het Gemeentelijk Archief Rotterdam. Het MUSEUM project wil een antwoord bieden op de problemen die culturele erfgoedinstellingen ondervinden wanneer zij hun content via één systeem willen ontsluiten. MUSEUM wil onderzoeken hoe men moderne *information retrieval* methoden gebaseerd op *statistical language modeling* kan combineren met verschillende soorten metadata. De voorwerpen uit het Gemeentemuseum Den Haag worden bijvoorbeeld op allerlei manieren beschreven met aanvullende informatie. In totaal beschikt het museum over ruim 100.000 objectbeschrijvingen, zo'n 250.000 bibliografische beschrijvingen, zoals boeken, artikelen en multimediaobjecten, en driekwart miljoen aanvullende archiefbeschrijvingen, zoals stukken uit het tentoonstellingsarchief. Veel van deze informatie is nu nog moeilijk toegankelijk. Niet alleen zou de museumprofessional graag met één zoekstelsel door alle drie soorten beschrijvingen zoeken, het is ook interessant voor de geïnteresseerde leek. MUSEUM wil dit mogelijk maken.

3.1.2. Beelden voor de toekomst – Pilot OMP (Onderwijs Media Platform)

<http://www.pilotomp.nl/webtv2/home>

Beelden voor de Toekomst is een project van Stichting Nederlands Filmmuseum, Nederlands Instituut voor Beeld en Geluid, Centrale Discotheek Rotterdam, Nationaal Archief, Vereniging Openbare Bibliotheken en Stichting Nederland Kennisland. Eén van de vijf doelstellingen van het project Beelden voor de Toekomst is het inhoudelijk toegankelijk maken (contextualiseren)

van het materiaal, in het bijzonder ten behoeve van educatief gebruik, maar ook voor de creatieve industrie en het brede publiek. Het onderwijs is één van de belangrijkste doelgroepen van Beelden voor de Toekomst. In 2007 en 2008 werd een Onderwijs Media Platform (OMP) opgezet voor het voortgezet onderwijs. Het OMP is een multimediaal onderwijsplatform, waarin collecties worden ontsloten in context met didactisch materiaal. Het OMP heeft als uitgangspunt dat het geheel groter is dan de som der delen en zal alle beschikbare bronnen van de archiefinstellingen aangevuld met bronnen van andere aanbieders centraal ontsluiten via één plek. In het pilot OMP komen op een innovatieve manier drie diensten samen: een online omgeving om werkstukken en presentaties met video te maken, een lessensbank en -bouwer, en een zoekfunctie die in meerdere bronnen tegelijk zoekt. Er werd ook onderzoek verricht naar methodes om collecties inhoudelijk te beschrijven en voor het brede publiek toegankelijk te maken via specifieke thema's. Beelden voor de toekomst is zich zeer bewust van het belang van contextualisering: *"Een belangrijke stap richting het succesvol beschikbaar maken van materiaal, is de contextualisering. De gedigitaliseerde audiovisuele content wordt daarbij zodanig voorbereid, dat zinvolle en gerichte toepassing mogelijk wordt. Voor het onderwijs wordt het materiaal bijvoorbeeld in een thematische, didactische context geplaatst die aansluit bij leermethodes en elektronische leeromgevingen."*

3.1.3. BIRTH (Building an Interactive Research and delivery network for Television Heritage)

<http://www.birth-of-tv.org/birth/>

BIRTH (Building an Interactive Research and delivery network for Television Heritage) was een door de EU gefinancierd project in het kader van het MEDIA plus programma en bestaat uit BBC (British Broadcasting Corporation), Joanneum Research (Oostenrijk), Instituut voor Beeld en Geluid (Nederland), Noterik Multimedia (Nederland), ORF (Österreichischer Rundfunk), SWR (Sudwestrundfunk, Duitsland) RTBF en Hagenberg University of Applied Science. De doelstelling van BIRTH was om digitaal materiaal uit de begindagen van televisie in Europa via een webportaal op een aantrekkelijke manier beschikbaar te maken voor een doelpubliek van media, wetenschap en het brede publiek. Binnen dit onderzoeksproject was er ook plaats voor gebruikersonderzoek naar de verwachtingen van de doelpublieken. Wat contextualisering betreft verwachtte vooral de academische wereld dat de context betekenisvol werd aangeboden in het licht van de educatieve mogelijkheden van audiovisueel materiaal. Zo vroegen zij bijvoorbeeld zoekmethodes via een tijdslijn, voldoende informatie bij het materiaal en de mogelijkheid om feedback te geven op de content.⁴²

⁴² HECHT 2004.

3.2. Praktijk

3.2.1. Video Active

<http://www.birth-of-tv.org/birth/>

<http://www.videoactive.eu/VideoActive/TopicSearch.do>

Het project is gesubsidieerd binnen het eContentplus programma van de Europese Commissie en loopt tot september 2009.⁴³ Video Active bouwt voort op het BIRTH of TV onderzoeksproject (*supra*), dat audiovisueel materiaal uit de begindagen van de televisie beschikbaar maakte. Met Video Active werd de collectie uitgebreid met recent materiaal. De content werd aangeleverd door nationale audiovisuele instituten uit Griekenland, Zweden en Hongarije en publieke omroepen zoals BBC (British Broadcasting Corporation), ORF (Österreichischer Rundfunk) en SWR (Südwestrundfunk). Ook het Nederlands Instituut voor Beeld en Geluid en de RTBF/VRT participeren. De universiteiten van Londen en Utrecht zijn verantwoordelijk voor het scheppen van de wetenschappelijke kaders. Het doel is deze (grotendeels) gesloten archieven beschikbaar te maken voor educatieve en academische doeleinden, voor het lager en voortgezet onderwijs, academici, erfgoedinstellingen en een breder publiek. De collecties worden gecontextualiseerd aangeboden. Fragmenten worden aangevuld met gedetailleerde beschrijvingen, de videobestanden worden gekoppeld aan artikelen, aan foto's, ingescande tijdschriften en omroepgidsen. Op deze manier wil Video Active een basis aanbieden voor onderzoek naar de representatie van culturele en historische gebeurtenissen en naar de ontwikkeling van het medium televisie op een crosscultureel niveau, zowel nationaal als transnationaal. De programma's zijn met een thesaurusmodule doorzoekbaar in verschillende talen. De gebruiker krijgt de keuze uit subportalen. Via de twee eerste portalen 'ontdek de Europese televisiegeschiedenis' en 'hoofdthema's uit de Europese geschiedenis' krijgt de gebruiker vrij toegang tot oud en recenter materiaal. Het derde portaal is een webwinkel, waar de leveranciers van content tegen betaling series en programma's aanbieden.

3.2.2. Nederlands Erfgoed: Digitaal!

<http://www.nederlandserfgoeddigitaal.nl/>

Tien Nederlandse erfgoedinstellingen hebben zich verenigd in het project Nederlands Erfgoed: Digitaal! De ambitie van het project is het creëren van een gezaghebbende digitale collectie Nederlands erfgoed, voor een breed doelpubliek (onderwijs, wetenschap, toerisme en creatieve industrie). Er zal veel aandacht worden besteed aan het in context presenteren van het cultureel erfgoed door het leggen van interessante dwarsverbanden tussen collecties van verschillende musea en instellingen. Daarvoor zal een informatie-architectuur in verschillende lagen worden opgezet. De eerste laag zal bestaan uit de data (bijv. digitale objecten) en metadata zoals beschikbaar bij de instellingen. Naar deze data zal worden verwezen in de

⁴³ Na september 2009 blijft het materiaal online beschikbaar.

centrale index. De index- en (semantische)verrijkingslaag zal er voor zorgen dat deze data en metadata samen zullen worden gebracht en kunnen worden opgevraagd, uitgewisseld en gebruikt. De bovenste laag zal bestaan uit een aantal publieksinterfaces, producten en diensten. In het project zullen enkele intelligente gebruikersapplicaties worden ontwikkeld en zal in 2009 een overkoepelende *portal* worden aangeboden, waarmee gebruikers op een aantrekkelijke interactieve wijze informatie kunnen opvragen en gebruiken.

3.2.3. Klara.be

<http://www.klara.be/>

Klara.be was al sinds 2001 online als website van de radiozender Klara van de Vlaamse Radio en Televisieomroep (VRT). Sinds 17 april 2008 werd het concept van de site grondig gewijzigd, in het kader van het driesporenbeleid dat de VRT voor haar cultuuraanbod toepast (de zgn. "Cultuurdelta": signalering in nieuws- en duidingsprogramma's, verdieping in algemene programma's en verrijking op *on demand*-platformen en in specifieke programma's). In plaats van een radio-website werd het een crossmediale site, die het culturele aanbod van alle VRT-kanalen opvolgt, samenbrengt en verrijkt met materiaal uit de beeld- en geluidsarchieven en links naar kwalitatief hoogstaande cultuurwebsites. Op die manier wordt audiovisueel erfgoed gecontextualiseerd met actueel materiaal en omgekeerd. Voor de digitalisatie van het archiefmateriaal werd binnen de VRT een speciaal budget vrijgemaakt. De navigatie op de site gebeurt per medium of per culturele discipline. Bovendien is de inhoud onderling verbonden via *expert generated tags*, in dit geval toegekend door de redacteurs. Doordat de volledige inhoud steeds op de site aanwezig blijft, ontstaat er langzaamaan een heel netwerk van culturele audiovisuele content die actueel is of dat ooit geweest is. In zijn korte bestaan is Klara.be al uitgegroeid tot één van de voornaamste online aanbieders van audiovisueel erfgoed in Vlaanderen.

3.2.4. Teleblik en Academia

<http://www.teleblik.nl/>

<http://www.academia.nl/>

Om het gebruik van de archieven in het onderwijs op grote schaal mogelijk te maken, biedt Het Nederlands Instituut Beeld en Geluid diensten aan die specifiek zijn gericht op het onderwijs in Nederland. Teleblik is het onderwijsproject voor het primair en voortgezet onderwijs in Nederland. Via de Teleblik website worden duizenden uren archiefmateriaal van het Nederlands Instituut voor Beeld en Geluid beschikbaar gemaakt, die speciaal geselecteerd zijn (door Teleac/NOT) en gecleared zijn voor gebruik in het onderwijs.

Academia is een vergelijkbare site bestemd voor het hoger onderwijs. Deze onderwijsgerichte sites bieden een selectie van direct bruikbaar materiaal, in een toegankelijk low-res browse formaat, aangeboden in thematische pakketten. Het succes van deze strategie blijkt uit het

feit dat tegenover de duizenden traditionele “bestellingen” door scholen, nu inmiddels honderdduizenden bekeken filmpjes op Teleblik staan.

Eind januari 2009 werd ED*IT gelanceerd. Een nieuw onderwijsplatform dat werd uitgewerkt in het kader van Beelden voor de Toekomst door het Nederlands Instituut voor Beeld en Geluid, de Stichting Nederlands Filmmuseum, Centrale Discotheek Rotterdam, Nationaal Archief, de Vereniging Openbare bibliotheken en de Stichting Nederland Kennisland. Het onderwijs is ook voor Beelden voor de Toekomst één van de belangrijkste doelgroepen.

3.2.5. INA

<http://www.ina.fr/>

L’Institut national de l’audiovisuel (INA) stelt ongeveer 20.000 uur materiaal van de Franse openbare radio- en televisiezenders ter beschikking van het grote publiek. De archiefbeelden worden via thematische portals en rubrieken ontsloten en van de nodige context voorzien in de vorm van interviews, foto’s en tekstbestanden. Er wordt ook gewerkt met gebruikersgroepen, namelijk het grote publiek (waaronder ook educatie), de professionele sector (*inamediapro*) en onderzoekers (*inathèque de france*). Voor het grote publiek worden archiefbeelden per thema, tijdperk of programma aangeboden (*explorer*) en wordt de bezoeker uitgenodigd om bepaalde thema’s te verkennen (*découvrir*). In deze laatste rubriek kan men ook het journaal van zijn/haar geboortedatum bekijken. Hier wordt ook op de actualiteit ingespeeld door actuele thema’s uit te werken rond bijvoorbeeld rechten van de mens naar aanleiding van het zestigjarig bestaan van de universele verklaring van de rechten van de mens. De bezoeker wordt ook uitgenodigd om deel te nemen (*partager*). In deze rubriek wordt elke maand niet alleen de persoonlijke keuze van een bekende persoonlijkheid en de selectie van de redactie getoond, maar wordt ook de selectie van een internetgebruiker gepost. Ten slotte is er ook een rubriek gericht op educatie (*apprendre*) met aparte pakketten voor het onderwijs (*collège, lycée en supérieur*) en de zogenaamde ‘jalons’. In deze tijdsfresco’s worden in feite tijdslijnen per thema (vb.: politiek, cultuur, WOII&II) gerangschikt. Wanneer men een beeld uit de fresco opvraagt, dan verschijnen in een kader de metadata met een uitgebreide historische situering.



Fig. 10 screenshot van de tijdsfresco thema cultuur op de website van het INA.

3.2.6. Smithsonian Photography Initiative

<http://photography.si.edu/>

Met het Smithsonian Photography Initiative wil dit instituut hun uitgebreide collectie foto's ontsluiten en via interactieve online programma's zoals *Click! Photography changes everything* de gebruikers laten bijdragen tot de dialoog over de integrale rol die foto's spelen in ons leven. Op *Click!* nodigt men onderzoekers én bezoekers uit om foto's en verhalen te posten over hoe fotografie onze cultuur en ons leven beïnvloedt (bv. '*photography changes how family history is sustained*', '*Photography changes how we experience history*').

3.2.7. EUROPEANA

<http://dev.europeana.eu/home.php>

Europeana is een Thematic Network gefinancierd door de Europese Commissie onder het eContent^{plus} programme, als deel van het i2010 policy. Europeana (European digital library network) is een samenwerking tussen 90 culturele erfgoedinstellingen, kenniscentra en IT experts uit Europa. De doelstelling van Europeana is om via een cross-domain portal toegang te verlenen tot het culturele en wetenschappelijke erfgoed van Europa, te werken aan de duurzaamheid van zo'n gezamenlijk portaal, het bestaande digitale materiaal te verenigen en digitalisatie te stimuleren. Vanaf december 2008 krijgen bezoekers toegang tot Europeana, de multimedia online-bibliotheek met meer dan twee miljoen boeken, kaarten, foto's,

documenten, schilderijen en films van nationale bibliotheken en culturele instellingen uit de 27 lidstaten van de EU. In 2010 zal de inhoud van Europeana verder worden uitgebreid én zullen interactieve zones worden opgestart voor thematische groepen. In de huidige versie kan men via *Mijn Europeana* zoekopdrachten of onderwerpen als bookmark vastleggen, onderwerpen markeren en toevoegen aan de persoonlijke map.

4. Aandachtspunten uit de sectorfeedback

Dit rapport biedt een theoretisch kader voor contextualisering. Blijft de vraag wat ter zake de noden zijn van zowel de aanbieders, die de audiovisuele data van een contextualisering moeten voorzien, als van de gebruikers, die op dit gecontextualiseerd aanbod een beroep willen doen. Daarom organiseerden we een informatieronde per mail. Actoren uit het onderzoeksveld en de sector konden schriftelijk reageren op enkele stellingen rond contextualisering.

- Stellingen

1. Bent u het eens met onze afbakening van het begrip contextualisering?
2. Bent u het eens met de afgebakende vormen van contextualisering en ziet u eventueel nog andere soorten contextualisering?
3. Heeft een multimedia-archief enkel zin als het in een betekenisvolle context wordt aangeboden?
4. Hoe ziet u in de toekomst de rol van de expert en expert generated metadata, expert generated tags evolueren?
5. Wat is het potentieel van innovatieve toepassingen rond automatische indexerings, annotatie en data mining in uw sector?
6. Wie contextualiseert het aanbod momenteel in uw sector (experts, gebruikers) en welke intermediairen kunnen daar in de toekomst een rol in spelen?
7. Welke voor- of nadelen ziet u aan het toekennen van metadata door de gebruiker (user generated metadata)?
8. Wat is voor uw sector de meerwaarde van de mening van de gebruiker, via folksonomies, rating en fora?
9. Welke voordelen ziet u voor uw sector aan het verbinden van content met andere content en randdocumentatie binnen hetzelfde archief, dezelfde instelling?
10. Welke voordelen ziet u voor uw sector aan het verbinden van content met andere content en randdocumentatie van verschillende collecties, verschillende sectoren?
11. Wat zijn de voordelen en valkuilen van externe gebruikersplatformen en hoe ziet u dit in de toekomst evolueren?
12. Welke meerwaarden ziet u voor uw sector in een doelgroepgericht, gecontextualiseerd multimedia-archief?
13. Wat zijn momenteel de meest gebruikte vormen van contextualisering in uw sector?
14. Wat zijn (in de toekomst) de meest relevante vormen van contextualisering voor uw sector? Motiveer uw antwoord.
15. Wat zijn de meest haalbare vormen van contextualisering voor uw sector? Motiveer uw antwoord.

- Respondenten:

- Kristin Verboven, producent cultuur VRT
- Stijn Lybeert, verantwoordelijke digitalisering archief Focus - WTV
- Rony Vissers, coördinator Packed
- Cees Klapwijk, directeur Digitale Bibliotheek voor de Nederlandse Letteren (DBNL)
- Tom Brink, manager onderwijs, Nederlands Instituut voor Beeld en Geluid
- Dirk Terryn, Canon cultuurcel

- Aandachtspunten:

Over het algemeen gaat men akkoord met de **afbakening van het begrip contextualisering** zoals het is beschreven in het rapport. Zowel het toekennen van metadata aan essence als het verbinden van content met andere content wordt dus beschouwd als contextualisering.

Wat betreft de vraag of **een multimedia-archief** enkel zin heeft als het in een betekenisvolle context wordt aangeboden halen de archiefbeheerders aan dat een multimedia-archief ook nuttig kan zijn als opslagplaats, voor het bewaren en beschermen van digitaal materiaal. Wanneer de objecten online raadpleegbaar zijn kan ontsluiting op alleen formele kenmerken soms afdoende zijn. Het multimediaal archief wordt namelijk in de eerste plaats ook gebruikt door de expert, die via zoekfuncties naar bronnenmateriaal op zoek gaat. Langs de andere kant zijn de respondenten van mening dat voor de eindgebruiker, of niet-expert doelgroepen, een zeker vorm van contextualisering noodzakelijk is. Anders bestaat het gevaar dat de gebruiker afhankelijk wordt van de aangeboden zoekstrategieën om zijn/haar weg te vinden in het archief.

Ook wat betreft de rol van de **experts** zijn er verschillende opmerkingen. De respondenten zijn het er over eens dat de inbreng van experts noodzakelijk blijft én dat er voorrang moet worden gegeven aan de informatie die wordt aangereikt door de experts. Men blijft er ook voorstander van om de inbreng van de experts duidelijk te onderscheiden van de inbreng van de gebruiker. De inbreng van de experts blijft noodzakelijk omdat zij een grondige kennis van het terrein hebben en omdat contextualisering deel uitmaakt van hun taak. Bovendien is voor het uitwisselen en linken van content een zekere uniformiteit van de metadata gewenst. De klassieke metadata met standaarden, ingevuld door de expert blijven hier de basis. Wanneer het mogelijk zal worden om op grote schaal gebruik te maken van stem- en beeldherkenning ten behoeve van metadatering, dan kan de rol van de expert wel meer veranderen in de richting van een procescontroleur. De respondenten zijn het er wel over eens dat consumentenbijdragen zoals *user generated metadata* en *user generated content*, *folkonomies* en *tags* een nuttige aanvulling kunnen zijn. De gebruiker kan soms over

bepaalde kennis beschikken die nuttig kan zijn voor andere gebruikers én de collectiebeheerder.

Wat betreft **innovatieve toepassingen** zoals automatische indexering, annotatie en data mining zijn de respondenten het erover eens dat dit interessante, complementaire aanvullingen zijn of kunnen zijn in de toekomst (zoals GAMA, Gateway to Archives of Media Art). Voor het archiefmateriaal met een tekort aan metadata kan in de toekomst de verdere ontwikkeling van spraakherkenning een mogelijkheid bieden om een deel van de metadata aan te vullen. In de mediasector laten de doorgedreven digitalisering van de productieprocessen toe om items te annoteren en te contextualiseren vanuit het productieproces zelf (scripts, programmagegevens). In deze sector ziet men ook een enorm potentieel aan (passieve) gebruikersinformatie die over verschillende media, instellingen of sites geregistreerd worden. De respondenten zijn het er wel over eens dat deze toepassingen de expert nooit volledig zullen kunnen vervangen.

Momenteel **wordt in hoofdzaak een beroep gedaan op experts** voor het contextualiseren van digitaal multimediaal aanbod. Vooral in de mediasector ziet men in de toekomst welkome bijkomende contextualisering door intermediairen en gebruikers. In Nederland voegen gebruikers op vrijwillige basis reeds contextinformatie toe (bv. In de Beeld en Geluid wiki en op het onderwijsplatform).

Wanneer het gaat over mogelijke voordelen van het toekennen van metadata door de gebruiker (**user generated metadata**) wordt vaak aangehaald dat dit interessante aanvullende kennis kan opleveren en dat de gebruiker correctieve feedback kan geven. De meest aangehaalde nadelen zijn dat de bijdragen van de gebruiker niet altijd relevant zijn en dat steeds de nodige redactie moet gebeuren op de reacties van de gebruiker. In de mediakunst sector is men geen voorstander van al te grote inbreng door de gebruiker. Daar gebeurt de contextualisering meestal in samenspraak met de kunstenaar. Deze behoudt over het algemeen liefst de controle over hoe zijn/haar werk wordt gecontextualiseerd. Langs de andere kant heeft een gebruikersinitiatief zoals Ubuweb (<http://ubuweb.com>), een onafhankelijke online bron voor allerlei soorten avant-garde kunst, zijn nut al bewezen.

De **mening van de gebruiker** via folksonomies, rating en fora is volgens de respondenten slechts van aanvullende meerwaarde. De meest positieve kijk op de bijdrage van de mening van de gebruiker komt uit de richting van de mediasector. Daar is men van mening dat reportages of beeldmateriaal dat vaak wordt opgevraagd indicatief kan zijn voor de interesse van de gebruiker. Bovendien werkt het de aanmaak van communities in de hand, wat dan weer een nauwe verbondenheid met het publiek tot stand brengt. De respondenten uit de mediasector zijn van mening dat fora ook inspiratie en/of feedback kunnen bieden op de producties van de collectiebeheerder. Het instituut voor Beeld en Geluid in Nederland waardeert de input van de gebruiker, omdat het niet altijd mogelijk is om alle content adequaat en volledig gericht op een bepaalde doelgroep te beschrijven. Daarvoor is hulp van

gebruikers nodig. Rating en folksonomies zijn in de mediakunst sector echter niet wenselijk, omdat veel kunstenaars zelf zo veel mogelijk controle willen behouden. Fora zijn in deze sector dan wel weer gebruikelijk.

Het **verbinden van content met andere content en randdocumentatie binnen hetzelfde archief**, dezelfde instelling wordt gezien als essentieel. Bijvoorbeeld in de digitale bibliotheek DBNL is het met elkaar in verband brengen van primaire en secundaire literatuur van belang omdat het een organische samenhang en clustering van data oplevert. Het verbetert mogelijk ook de interne werking voor het maken van reportagereeksen of het opzoeken van relevant archiefmateriaal. In de mediakunst sector zou het verbinden van content met andere content en randdocumentatie de hedendaagse en mediakunst toegankelijker kunnen maken voor de gebruiker.

Ook **het verbinden van content met andere content en randdocumentatie van verschillende instellingen en sectoren** wordt gezien als belangrijk. Het zou ook een nauwere samenwerking tussen de sectoren kunnen bevorderen. Ook waardevol is dat verschillende media (beeld en audio bijvoorbeeld), met elkaar kunnen worden verbonden. Een concreet voorbeeld binnen Beelden voor de toekomst is het onderwijsplatform ED*IT, waarin bronnen van Beeld en Geluid (audio en video), het Filmmuseum (video), het Nationaal Archief (foto, documenten), Naturalis (artikelen, foto's), Rijksmuseum van Oudheden (artikelen, afbeeldingen, dossiers), Museum Volkenkunde (artikelen, afbeeldingen, dossiers) en Nationaal Centrum voor Wetenschap en Technologie (artikelen, afbeeldingen) gekoppeld worden waardoor ze gezamenlijk doorzoekbaar zijn.

Eén van de **voordelen van externe gebruikersplatformen** die wordt aangehaald is de zichtbaarheid van de instelling buiten het eigen webportaal. In de mediakunst sector kunnen de externe gebruikersplatformen, zoals Ubuweb, ook een geschikte context bieden voor kunstenaars om zélf hun materiaal te posten. De grootste **valkuilen** van externe gebruikersplatformen volgens de respondenten zijn de rechtenkwesties zoals het onrechtmatig gebruik van archiefmateriaal.

Doelgroepgerichte contextualisering is vooral voor het onderwijs en onderzoek van belang. Daar zijn de respondenten uit alle sectoren het over eens. Binnen de sectoren is veel archiefmateriaal ontoegankelijk voor bepaalde doelgroepen, de doelgroepgerichte contextualisering is dé oplossing. Canon cultuurcel ziet zowel voordelen voor de leerkracht (*e-learning, user generated content*) als voor de leerling (pakketten, mediageleerdheid) in een webportaal met multimediale content voor het onderwijs.

De **meest gebruikte vorm van contextualisering** momenteel is nog steeds het toekennen van metadata door experts en het hergebruiken van archiefmateriaal binnen een andere context. Voor de mediasector is het crossmediaal en intersectoraal gebruik van archiefmateriaal de meest relevante vorm van contextualisering voor de toekomst. Met andere

woorden een combinatie van contextualisering door de expert, door intermediairen uit de culturele en educatieve sector, én de gebruiker. Voor de mediakunst sector zal in de toekomst nog een grote inhaalbeweging nodig zijn op het vlak van het toekennen van metadata en de verbinding van content met randdocumentatie. Deze basisvorm van (gecontroleerde) contextualisering is ook datgene waar experts én kunstenaars vertrouwd én akkoord mee zijn.

5. Conclusies en aanbevelingen

5.1. Conclusies

Bij aanvang van het onderzoek werd eerst gezocht naar een afbakening van het begrip contextualisering. Het werd echter al snel duidelijk dat er *no such thing* is als een internationale definitie van contextualisering, laat staan dat er een vaste begripsbepaling zou bestaan in de context van de ontsluiting van multimediale data. Daarom hebben we ervoor gekozen eerst te verkennen waar contextualisering zich bevindt tussen andere, gerelateerde termen die tijdens het literatuuronderzoek naar voren kwamen zoals *ontsluiting*, *context*, *verrijkte inhoud* of *betekenisvolle relaties*.

Het is daarbij belangrijk gebleken te schetsen hoe de ontsluiting van erfgoed is geëvolueerd en nog evolueert bij de overgang naar een digitale omgeving. Vóór de ontwikkeling van de netwerkcultuur, web2.0 en de digitaliseringsgolf had het collectiebeheer ook al (ten minste gedeeltelijk) een publieksgerichte finaliteit. Maar de metadata-annotatie was afgestemd op doorzoekbaarheid door experts: zij ondervroegen ten dienste van de gebruikers de databanken via metadata die ze vaak zelf hadden ingevoerd en brachten de inhoud vervolgens naar het grote publiek via *one-to-many* kanalen, inclusief contextualisering. De expert bevond zich tussen gebruiker en collectie, de collectie was met andere woorden (zo goed als) enkel via de expert toegankelijk.

Tegenwoordig worden collecties online aangeboden en geconsulteerd. De intermediaire rol van de expert inzake vertaling van zoekvragen van de gebruiker en contextualisering van zoekresultaten naar de gebruiker is in deze weggevallen. Deze evolutie heeft drie duidelijk zichtbare gevolgen. Ten eerste is het belang van een gebruiksvriendelijke zoekinterface veel groter geworden. Ten tweede kreeg de gebruiker een actievere rol: hij krijgt de kans zelf metadata (*user generated metadata*) en/of content (*user generated content*) toe te voegen. Ten derde gebeurt de contextualisering van de collectie niet louter meer door de expert, maar (bewust of onbewust) ook door de gebruiker.

De eerste stap in contextualisering is de toevoeging van metadata. Dit is de meest primaire vorm van contextualisering, waarop bovendien een aantal andere vormen steunen. Het verbinden van de metadata als contextualiseringsvorm met de essence kan gebeuren door klassieke *expert generated* annotatie (manueel, semi-automatisch en automatisch). Daarnaast is *expert generated tagging* ook een vorm van contextualisering via metadata-annotatie.

Een tweede stap is het verbinden van content met andere content en randdocumentatie. Deze verbindingen kunnen gebeuren binnen dezelfde collectie maar kunnen ook tot stand gebracht worden tussen verschillende collecties. Het betekenisvol verbinden van content zorgt voor toevoeging van inhoudelijke context. Dit laat toe pakketten samen te stellen rond bepaalde thema's of volgens een tijdslijn. Dit betekenisvol aanbieden van content kan doelgroepgericht

gebeuren. Collectiebeheerders kunnen hun content op maat naar verschillende doelgroepen ontsluiten via aparte webpagina's, anderzijds kunnen doelgroepgerichte webpagina's materiaal van verschillende collecties verzamelen.

Ten derde bestaat er een waaier aan mogelijkheden waardoor de gebruiker (bewust en onbewust) items kan contextualiseren. Heel wat web2.0-toepassingen hebben tot doel om meningen bij wijze van contextualisering met de essence te verbinden, zoals *user generated tags*, *user generated taxonomieën* of *folksonomies* en *rating*. De *expert generated metadata* blijven de kern van de contextualisering, dit werd ook bevestigd door de resultaten van de sectorfeedback. De respondenten gaven echter ook aan dat de gebruikerscontextualisering als aanvulling een meerwaarde kan betekenen voor het archief. *Fora*, collecties en *social bookmark managers* laten toe om online discussie op te starten, eigen collecties samen te stellen en favorieten te verzamelen en te delen met anderen. Op externe platformen zoals Youtube.com en haar vele klonen worden niet de favoriete links, maar de items zelf verzameld.

De hierboven toegelichte vormen van contextualisering kunnen ook gecombineerd worden toegepast. Zo kan een totaalpakket ontstaan, waar de content op verschillende wijzen betekenisvol wordt aangeboden. Er ontstaat als het ware rond elk archiefitem een omgeving van allerlei soorten contextualisering, die het item langs alle kanten met gerelateerde gegevens, hetzij metadata, hetzij andere content, verbinden.

5.2. Aanbevelingen

Hoe moeten beheerders van multimediale data / audiovisueel erfgoed hun materiaal contextualiseren om daarmee tegemoet te komen aan de noden van de verschillende gebruikersgroepen? En welke voorzieningen moet een centraal depot treffen om die contextualisering ten volle tot haar recht te laten komen?

In deze voorlopige aanbevelingen verwerken we de conclusies op basis van het onderzoek tot nu toe. We delen deze aanbevelingen op in de adviezen voor de korte termijn, die ingepast kunnen worden in de demonstrator waaraan gewerkt wordt in WP6, en in de adviezen voor de lange termijn, met een uitgebreide contextualisering, ingepast in de ontsluiting van een grote *repository* van multimediale data.

Korte termijn: demonstrator

De demonstrator die in WP6 ontworpen wordt dient nog niet alle mogelijkheden omtrent contextualisering ten volle te benutten. Door toch al enkele aanbevelingen ter harte te nemen kan eventueel wel een vooruitzicht worden gegeven op de richting waarin het ontwerp evolueert. Een basiscriterium daarbij is uiteraard dat er een kritische massa aan kwalitatieve

content (essence én metadata) aanwezig is die als voorbeeld van contextualisering kan dienen.

Door niet zomaar de essence, maar ook de bestaande, *expert generated* metadata in het zoekresultaat te laten zien kan aangetoond worden hoe deze bij de essence meer context geven. Verder is het zeker interessant ook een bredere contextualisering mogelijk te maken op basis van *user generated metadata* zoals tags. Deze contextualisering kan verder uitgewerkt worden door de content via links op basis van de metadata te verbinden met allerlei randdocumentatie, of met andere multimediale data van binnen (door bijvoorbeeld de mogelijkheid aan te bieden een eigen collectie samen te stellen) of buiten (door bijvoorbeeld *embedding* mogelijk te maken op andere gebruikersplatformen, interoperabiliteit van data en metadata met andere platformen) de collectie. Uiteraard moeten de mogelijkheden daarbij verder gaan dan enkel het verkennen (*exploring*) van de verschillende opeenvolgende zoekresultaten. Op basis van de zoekresultaten moet het ook mogelijk zijn verder te verfijnen of te verdiepen (*drill down*).

Een geïntegreerde contextualisering in de vorm van een module of een gepersonaliseerde interface is voor de demonstrator wellicht een grote technische uitdaging, maar ze kan toch aantonen hoe tegemoet gekomen kan worden aan de noden van de verschillende gebruikersgroepen wat betreft de zoekfunctionaliteit, de presentatie van de zoekresultaten, het samenstellen van eigen collecties enzovoort. Het zou bijvoorbeeld verhelderend kunnen zijn om meerdere zoekinterfaces te ontwerpen, van een eenvoudige tot een geavanceerde.

Lange termijn: repository

In het algemeen is de aanbeveling voor de contextualisering in een groot Vlaams multimediaal *repository* natuurlijk voor de hand liggend: het platform moet zo goed mogelijk uitgerust zijn om zo veel mogelijk vormen van contextualisering op een degelijke manier aan te bieden én te ontvangen. In de praktijk hangen de vereisten echter niet alleen af van het ontwerp van het *repository* en de toegang daartoe. Voor heel wat vormen van contextualisering zijn de metadata namelijk de basis en dienen de aanbevelingen zich dus meer te richten tot de toeleveranciers hiervan.

Vanuit WP3 weten we dat er wordt gewerkt aan een gelaagd metadatamodel, met Dublin Core als basis en sectoraal verschillende uitbreidingen. Maar voorafgaand hieraan is het essentieel dat er kwalitatief zeer hoogstaande metadata zijn toegekend én een zo goed mogelijke afstemming (*mapping*) van de erg verschillende metadatavelden uit de praktijk met de velden zoals die worden beschreven in de metadastandaarden. Dit is essentieel voor een goed zoekresultaat, zeker bij geavanceerde zoekopdrachten. Aangezien dit voor een niet te onderschatten deel aan collectiebeheerders nog steeds een groot probleem vormt, hangt de haalbaarheid van een goede contextualisering ook hiervan af. Kwalitatief sterke metadata zijn dus een kritische succesfactor.

Een tweede aanbeveling houdt verband met de compatibiliteit van de presentatie van de zoekresultaten met allerhande web2.0-toepassingen. *Tagging, rating, fora, communities, folksonomies*, de mogelijkheid tot *embedding* (uiteraard rekening houdend met de beperkingen inzake IPR) ... vragen allemaal een specifiek ontwerp van de presentatie van de zoekresultaten. Wat hiervan de technische implicaties zijn moet in andere werkpakketten verder worden onderzocht.

De derde grote aanbeveling spitst zich toe op de zoekfunctionaliteit. De manier waarop de zoekmogelijkheden worden aangeboden is van groot belang. De uitwerking van een aangepaste zoekinterface is niet opgenomen in het onderzoeksproject, maar aangezien dit duidelijk deel uitmaakt van het onderzoek rond gebruikersnoden moeten we deze problematiek toch even aanhalen. Het ontwerp van degelijke zoekinterfaces is bijna een wetenschap op zich. Een goede zoekinterface verenigt in zich twee paradoxale kenmerken. Ze moet allereerst voldoende eenvoudig zijn, zodat van de gebruiker niet te veel voorkennis en inschattingsvermogen geëist wordt omtrent de draagwijdte van de zoekvelden. Ze moet anderzijds ook voldoende geavanceerd zijn, om de gebruikers die dat wensen de kans te geven ingewikkelde, meervoudige en strikt afgelijnde zoekopdrachten te laten uitvoeren. Features als *fuzzy search* bijvoorbeeld, oorspronkelijk bedoeld om spelfouten te omzeilen en thesauri overbodig te maken, kunnen handig zijn voor leken, maar voor experts zijn ze soms juist erg onhandig. Daarom is het aan te raden ze optioneel te maken. Het is aan te raden de moeilijkheid van een eenvoudige en toch geavanceerde zoekinterface te omzeilen door de collectie bevrraagbaar te maken via modules. Bij modules zijn de zoek- en resultaatsschermen afgestemd op de vereisten van gebruikersgroepen. Ze vormen als het ware verschillende toegangspoorten tot dezelfde collectie. Dat geeft de verschillende doelgroepen niet alleen de kans te zoeken op een manier die hen het beste past, maar ook de resultaten gepresenteerd te krijgen op een manier die specifiek op hun noden is afgestemd, bijvoorbeeld door ze te rangschikken op een volgorde die voor die doelgroep specifiek van belang kan zijn.

Een vierde grote aanbeveling in verband met contextualisering is die van de technologische toekomstgerichtheid van het platform. Dit moet voorbereid zijn op nieuwe evoluties inzake contextualisering, bijvoorbeeld op het vlak van automatische metadata-extractie. Dit moet mogelijk maken dat er automatisch links worden gelegd op basis van metadata die nog niet in de klassieke set aanwezig zijn. Ook van de zoekmethoden valt te verwachten dat ze zullen evolueren, misschien nog sneller dan dat de afgelopen jaren al het geval is geweest. *Query by humming* bijvoorbeeld is nu al een mogelijkheid op enkele vooruitstrevende muziekwebsites.

Ten slotte moeten we wijzen op de nood aan gebruikersonderzoek. Dit rapport schetst de verschillende verschijningsvormen van contextualisering, maar de eigenlijke noden van het beoogde publiek kunnen enkel door een grootschalig gebruikersonderzoek worden achterhaald. Professionele mediamakers en collectiebeheerders uit de kunsten, de culturele en de erfgoedsector willen hun collecties betekenisvol aanbieden aan de respectieve doelgroepen

en zijn bereid het publiek te laten participeren. Het publiek kan daarbij zowel bestaan uit de particuliere eindgebruiker als uit 'intermediaren', die het op hun beurt naar de eindgebruiker brengen. Die 'intermediaren', zoals het onderwijs of de toeristische sector beheren zelf meestal geen collectie, maar ze kunnen uiteindelijk wel een functie hebben als doorgeefluik naar een groter publiek of naar specifieke doelgroepen. Enkel door gebruikersonderzoek in deze verschillende sectoren kan bij het ontsluiten worden ingespeeld op de noden van de eindgebruikers.

6. Bibliografie

BAILER, W., HÖLLER, F. & A. MESSINA, *State of the art of content analysis tools for video, audio and speech. Deliverable D15.3 MDS3 of PrestoSpace*, Rome, 2005.

DE JONG, A. Metadatamodel Beeld en Geluid biedt gebruiker van digitale content diepte en structuur. In: *InformatieProfessional*, 6, 2007.

GOLDMAN, J., RENALS, S., BIRD, S. & F. DE JONG, Accessing the spoken word. In: *International journal on Digital Libraries*, 5(2005)4.

GORDON-MURNANE, L. Social bookmarking, Folksonomies and Web 2.0 Tools. In: *Searcher – the magazine for Database Professionals*, 6, 2006.

HECHT, A. 2004, The BIRTH project, *Viewfinder* 57, 16-17.
http://www.birth-of.tv.org/birth/content/viewfinder_BIRTH.pdf;jsessionid=2B4BD5F60D4E003E47A512CA6079FB3F

HOPPER, R., EBU Project Group P/Meta: Metadata exchange scheme v1.0. In: *EBU Technical Review*. 290 (2002, 2), Genève, 2002.

KOUWENHOVEN, T. Zoeken + navigeren = vinden! Over zoekers, zoekgedrag, zoekmachines en hun interfaces bij het zoeken naar audiovisuele content. In: LAUWERS, M. & B. HOGENKAMP, *AudioVisueel. Van emancipatie tot professionalisering. Jaarboek 2005*. Den Haag, 2006.

MELENHORST, M., GROOTVELD, M. & M. VEENSTRA, Tag-based information retrieval for educational videos. In: *EBU Technical Review*. 214 (2008, 2). Genève, 2008.

MOLS, B. Het virtuele museum opent alle deuren. In: *InformatieProfessional*, 9, 2008.

OOMEN, J. & R. VAN KRANENBURG, *Internet en het nieuwe leren. De toepassing van streaming media*. Hilversum, 2003.

MOSTMANS, L., VAN PASSEL, E. & G. NULENS, *Werkpakket 5. Architectuur digitale bewaring en ontsluiting. Deeltaak 5.2. Distributiemodaliteiten*. Brussel, 2008.

TRANT, J. *Tagging, Folksonomy and Art Museums: Results of steve.museum's research*, Archives & Museum Informatics, 2009.

VAN DER MAAS, P., HEEREN, W. & A. VAN HESSEN, Het Radio Oranje-project. In: *DIXIT - Tijdschrift over Toegepaste Taal- en Spraaktechnologie*, 4 (2006)2, p. 6-7.

VAN HOOLAND, S. From spectator to annotator. Possibilities offered by User-Generated Metadata for Digital Cultural Heritage Collections. In: *Immaculate Catalogues: Taxonomy, Metadata and Resource Discovery in the 21st Century*, Proceedings of CILIP Conference, University of East Anglia, UK, 13-15 September 2006.

VAN PASSEL, E. & J. BEYL 2008, BOM-VI. Werkpakket 1 – Gebruikersnoden. Deeltaak 1 – Open en dynamisch archief.

https://projects.ibbt.be/bom-vl/fileadmin/user_upload/frontendfiles/BOM-VI%20WP1.1%20Open%20en%20dynamisch%20archief.pdf