



VIAA: Digitaliseringsgolf 2

Audio-cd-r

Eindrapport
v.1.0 Maart 2016



Inhoudsopgave

I.	Inleiding.....	4
II.	Start van het project	6
2.1	Scope	6
2.2	Timing.....	6
III.	Resultaten van het project.....	7
3.1	Scope	7
3.2	Timing.....	2
IV.	Overzicht resultaten registratie en digitalisering.....	3
4.1	Resultaten van de registratie	4
4.1.1	Batches	4
4.1.2	Materiaal	5
4.1.3	Content Partners	6
4.1.4	Merk.....	6
4.1.5	Datum	7
4.1.6	Duur	8
4.1.7	Deterioratiefenomenen.....	9
4.1.8	Bewaard door een andere VIAA partner	10
4.2	Resultaten van de digitalisering	11
4.2.1	Het digitaliseringsproces	11
4.2.2	Duur	12
4.2.3	Niet-gedigitaliseerde cd-r's.....	13
V.	Conclusies en Lessons Learned	14
5.1	Transport	14
5.2	Verpakking en Barcodering	14
5.3	Registratie en digitalisering	15
5.3.1	Merk.....	15
5.3.2	Datum	15
5.3.3	Duur	15



5.3.4	Deterioratiefenomenen.....	16
5.3.5	Overige velden.....	16
5.3	Algemeen.....	17

I. Inleiding

Het Vlaams Instituut voor Archivering (VIAA) werd opgericht in december 2012. De eerste stap was het in kaart brengen van het aandeel audiovisueel materiaal in de collecties van een afgebakende kring partners. Dit waren de cultureel-erfgoedinstellingen die gesubsidieerd worden binnen het Vlaams cultureel-erfgoeddecreet, de regionale omroepen en de openbare omroep. VIAA heeft daarop samen met FARO en PACKED een globale inventaris opgemaakt van audiovisueel materiaal dat aanwezig was, door de inventarislijsten, ingevuld door de instellingen zelf, te verzamelen. Op basis van deze cijfers is beslist om te starten met een eerste Digitaliseringsgolf van 2 videoformaten (Betacam SP en U-matic) en 2 audioformaten (kwartduims audiotape en compact audiocassette).

In 2014 hebben VIAA en PACKED een tweede kring partners bevroegd: de stadsarchieven van de centrumsteden en de cultureel-erfgoedinstellingen die erkend zijn binnen het Vlaams cultureel-erfgoeddecreet. Tegelijkertijd bereidde VIAA haar tweede Digitaliseringsgolf voor met 6 dragerformaten: Betamax, VCR, open reel video, **audio-cd-r**, wasrol en draadspoel.

Bij de aanbestedingsprocedure kwam het Nederlandse **Picturae** als beste uit de bus voor de van de audio-cd-r. In dit project waren enkel partners uit de cultureel-erfgoedsector betrokken.



Fig. 1: een voorbeeld van een cd-r van het merk Sony.

Bij de aanvang van het digitaliseringsproject werd uitgegaan van bepaalde cijfers en voorspellingen wat betreft scope, timing en budget. Deze cijfers waren vaak gebaseerd op schattingen omwille van het ontbreken van absolute cijfers bij de Content Partners (CP). Tijdens het verloop van het project werd duidelijk dat de vooropgestelde cijfers qua aantal te digitaliseren dragers zouden wijzigen.

Het project krijgt mogelijk later nog een uitbreiding. Bij de Content Partners van VIAA kunnen immers nog steeds cd-r's binnenkomen. Daarnaast voorziet VIAA ook om vanaf 2016 zijn groep van Content Partners

uit te breiden, die op hun beurt nog over cd-r's kunnen beschikken. De cd-r's die op die manier nog aangemeld worden, kunnen gebundeld worden in één of meerdere extra batches, die later gedigitaliseerd kunnen worden door dezelfde Service Provider (SP), Picturae.

In dit rapport vindt u een bondige samenvatting van VIAA's project van cd-r's, grotendeels gebaseerd op gegevens uit de registratie door de CP's en de digitalisering door de SP, uit het registratiesysteem AMS, maar ook gestoeld op eigen ervaringen. We hebben ervoor geopteerd om in deze fase noch de CP's noch de SP uitgebreid te bevragen naar hun ervaringen bij dit traject, maar we houden dit voor een latere en meer globale evaluatie van Digitaliseringsgolf 2. Hier zal u dus enkel vaststellingen en vergelijkingen terugvinden van de registratie- en de digitaliseringgegevens, gekoppeld aan de bevindingen van VIAA zelf.

Ten slotte delen wij nog graag drie belangrijke punten die u in het achterhoofd dient te houden bij het lezen van dit rapport:

- 1) Het gaat in dit project enkel over **zelfgeschreven audio-cd-r** (Orange Book), maar ook de cd-rw valt hierbinnen. Wat duidelijk niet in de scope van dit project valt zijn commerciële (geperste) cd's, data-cd's allerhande, cd-rom en dvd-r(w). Hoe een CP deze formaten uit elkaar kon houden, werd toegelicht in de handleiding voor de Content Partners. In dit document zijn de termen 'audio-cd-r' en 'cd-r' inwisselbaar.
- 2) Een cd-r bevat zelf al digitale bestanden, we spreken hier dus eigenlijk over **migreren of rippen** van de inhoud i.p.v. over digitaliseren in de strikte zin van het woord. Voor de gemakkelijkerheid wordt in dit rapport echter systematisch gesproken over digitalisering.
- 3) Dit rapport behandelt enkel de digitalisering van de audio-cd-r's uit de cultureel-erfgoedsector (Kring 1 en 2). **De 22.000 andere audio-cd-r's van de VRT** worden *inhouse* gedigitaliseerd, in een project dat nog loopt tot de zomer van 2016. Hiervoor werd gekozen omdat de VRT enerzijds zelf over alle benodigde apparatuur beschikt, en anderzijds omdat het verpakken en transporteren van de audio-cd-r's een onzinnig grote inspanning zou betekenen, vergeleken met de geringe inspanningen die nodig zijn om een cd-r te digitaliseren.



II. Start van het project

2.1 Scope

De scope van het project werd op 30/08/2014 als volgt vastgelegd:

- We digitaliseren **1816 cd-r's** naar .wav, LPCM, 16bit/48kHz.
- Aangezien de duurtijd van de cd-r's door de Content Partners in de inventarisatietabellen meestal niet opgegeven werd, werd er in de tender uitgegaan van een gemiddelde van **0,8 uur** of 00:48:00 per schijf, dus in totaal ongeveer **1500 uur**.
- Van **9 Content Partners** uit de cultureel-erfgoedsector.
 - ADVN
 - AMSAB-ISG
 - In Flanders Fields Museum
 - KC Antwerpen
 - Letterenhuis
 - M HKA
 - Provinciale Bibliotheek Limburg
 - SMAK
 - UGent

2.2 Timing

Als timing voor het project werd het volgende voorstel uitgewerkt:

- Registratie van 1 november 2014 tot 11 maart 2015
- Digitalisering van 20 januari tot 1 juli 2015

III. Resultaten van het project

3.1 Scope

Er werden door Picturae in totaal **1209 cd-r's** verwerkt. Dat is ongeveer een derde minder dan oorspronkelijk werd ingeschat. Dit is deels toe te schrijven aan het feit dat de 200 door het ModeMuseum Hasselt in de inventarisatie opgegeven cd-r's, uiteindelijk andere optische formaten dan audio-cd-r's bleken te zijn. Ook bij de andere CP's waren de uiteindelijke aantallen wat lager dan oorspronkelijk opgegeven.

Daarbij moet nog een belangrijke **kanttekening** gemaakt worden: er werden oorspronkelijk 1200 dragers naar Picturae getransporteerd. Daar werd echter opgemerkt dat er enkele doosjes waren waarin zich meerdere schijfjes bevonden waarvan Picturae geen metadata had ontvangen. Er zat dus niets anders op dan extra records aan te maken in AMS per extra schijf. Die negen cd-r's werden dan verzameld in een extra batch (CDRB02A), maar hoefden dus niet meer fysiek getransporteerd te worden. Zo kwamen we op een totaal van 1209 records, en dus 1209 digitale files.

Uiteindelijk konden 219 cd-r's niet gedigitaliseerd worden om verschillende redenen (zie verder). Van **990** cd-r's kregen we dus een digitale file terug.

De gedigitaliseerde cd-r's waren uiteindelijk afkomstig van **16 Content Partners** uit de cultureel-erfgoedsector:

- ADVN
- AMSAB
- *Cultureel Erfgoed Annuntiaten*
- *Den AST*
- IFF
- KC Antwerpen
- Letterenhuis
- M HKA
- *Museum Dhondt-Dhaenens*
- *Museum Dr. Guislain*
- Provinciale Bibliotheek Limburg
- SMAK
- *Stadsarchief & HIDOC Dendermonde*
- *Stadsarchief Gent*
- *Stadsarchief Waregem*
- UGent

Ondanks dat er dus 7 Content Partners bij kwamen, ging het aantal cd-r's toch naar beneden. Deze 7 extra partners behoren tot Partnerkring 2, die in 2014 werd bevestigd en in 2015 Content Partner werden bij VIAA.

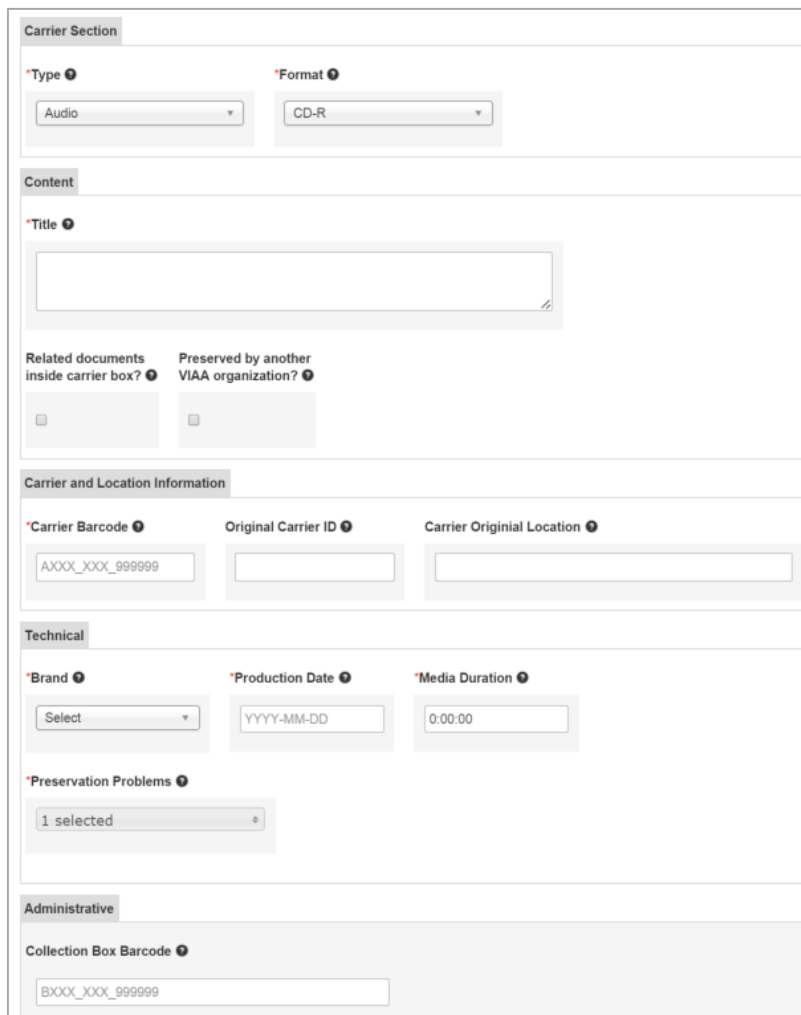


3.2 Timing

De effectieve registratie liep van 20 oktober 2014 tot 1 juli 2015. De eerste cd-r werd geregistreerd door IFF, de laatste door SMAK. De digitalisering van de cd-r's liep van 1 maart 2015 tot 30 september 2015. De digitalisering is dus iets later gestart dan voorzien. Dit was te wijten aan het tempo van de registratie bij de Content Partners, vooral bij de CP's uit Kring 2, die er in een later stadium bij kwamen en dus nog een registratie-opleiding moesten krijgen.

IV. Overzicht resultaten registratie en digitalisering

Tijdens de voorbereiding en de registratie werd metadata verzameld in het registratiesysteem AMS. Dit registratiesysteem werd op maat ontwikkeld voor VIAA door het Amerikaanse *AVPreserve*. De metadata die nodig was voor de digitalisering en het transport werd ingegeven door de Content Partners zelf. Hieronder wordt een overzicht gegeven van de belangrijkste bevindingen na afloop van het project, zowel van de registratie door de Content Partner als van de digitalisering door de Service Provider.



The screenshot displays a registration form for CD-Rs in the AMS system, organized into several sections:

- Carrier Section:** Includes dropdown menus for ***Type** (set to 'Audio') and ***Format** (set to 'CD-R').
- Content:** Features a ***Title** text input field. Below it are two checkboxes: **Related documents inside carrier box?** and **Preserved by another VIAA organization?**, both currently unchecked.
- Carrier and Location Information:** Contains three input fields: ***Carrier Barcode** (with placeholder 'AXXX_XXX_999999'), **Original Carrier ID**, and **Carrier Original Location**.
- Technical:** Includes ***Brand** (a dropdown menu set to 'Select'), ***Production Date** (a date input field with placeholder 'YYYY-MM-DD'), and ***Media Duration** (a time input field set to '0:00:00'). Below these is a ***Preservation Problems** dropdown menu showing '1 selected'.
- Administrative:** Contains a **Collection Box Barcode** input field with placeholder 'BXXX_XXX_999999'.

Fig. 2: het invulformulier voor de registratie van cd-r's in AMS



4.1 Resultaten van de registratie

Er werd door alle Content Partners samen gedurende **40 werkdagen** geregistreerd, en hierbij waren **15 registratoren** aan het werk.

Om de registratie zo vlot mogelijk te laten verlopen, voorzag VIAA een gedetailleerde handleiding voor de Content Partners. Hierin stonden onder andere de verschillende kenmerken van een cd-r beschreven, alsook een toelichting bij alle velden die in AMS in te vullen waren. Er werd ook aan de registratoren ter plaatse een opleiding gegeven door een VIAA-medewerker.

4.1.1 Batches

De cd-r's werden door VIAA verzameld in **2 aparte batches** om telkens ongeveer dezelfde hoeveelheden te kunnen aanbieden aan de Service Provider, en tegelijkertijd de Content Partners voldoende tijd te geven om de registratie af te ronden.

De cd-r's van één CP in een batch noemt VIAA een shipment of verzending. In één batch kunnen dus verschillende shipments of verzendingen zitten, en dus ook verschillende CP's.

De eerste batch vertrok op 16 februari 2015, de tweede (en voorlopig) laatste batch op 1 juli 2015. De gemiddelde grootte van een batch was 605 cd-r's. Hier wordt de extra batch waarvan hierboven sprake (zie 3.1 Scope) niet meegeteld omdat die niet fysiek werd getransporteerd. De transporten zijn als volgt in cijfers samen te vatten:

Batches	2
Shipments/verzendingen	20
Verpakkingsdozen	33
Transportbakken	10

Fig. 3: Aantal batches, shipments, verpakkingsdozen en transportbakken in het project

4.1.2 Materiaal

Voor de verpakking en het transport van de cd-r's werden barcodes voorgedrukt en dozen op maat gemaakt. In tegenstelling tot eerdere projecten werd enkel een barcode op de dragerdoos geplakt en niet op de schijf zelf. De reden hiervoor is dat het plakken van labels op de schijf problemen kan veroorzaken bij het afspelen en dus digitaliseren van de cd-r, maar ook de lijm van de labels kan de cd-r beschadigen. We hebben daarom gekozen om enkel een barcode op het doosje te kleven. Aan de SP werd dan ook gevraagd bijzondere aandacht te besteden aan het voorkomen van vergissingen tussen de drager en zijn doos.

Ook de verpakkingendozen werden voorzien van een barcode, zodat er in elke stap van het proces snel nagegaan kon worden in welke doos een drager zich bevond. Er werden in totaal:

- **1200 barcodes** voor de dragers gebruikt. Oorspronkelijk werden er 2145 besteld, dat is 44% te veel en is te wijten aan het verminderde aantal aangeboden cd-r's. Er wordt ook telkens iets te veel besteld om het opduiken van extra materiaal op te vangen, in dit geval was dat niet nodig.
- **33 barcodes** voor de dozen gebruikt. Oorspronkelijk werden er 59 besteld, dat is eveneens 44% te veel.
- **33 verpakkingendozen** werden gebruikt. Sinds de opstart van Digitaliseringsgolf 2 werken we met een doostype systeem waarbij verschillende audio- en videoformaten in dezelfde dozen passen. Hieronder staat het overzicht van de verschillende doostypes. Doostype 5 is geschikt voor het verpakken van de cd-r's, maar ook voor kleine formaten ½" open reel video, en in de toekomst ook voor andere optische schijven, die vaak in gelijkaardige jewel-cases zitten als de audio-cd-r's.

Type nummer	Dragerformaat
Doostype 1	Betacam (SP) (groot) Digital Betacam (groot)
Doostype 2	U-matic ½" open reel video (groot)
Doostype 3	Audiocassettes
Doostype 4	¼" audiotape 1" open reel video
Doostype 5	Audio-cd-r ½" open reel video (klein)
Doostype 6	VCR
Doostype 7	Betamax Digital Betacam (klein) Betacam (SP) (klein)

Fig. 4: tabel van door VIAA gebruikte dozen en de dragers waarvoor ze gebruikt worden.

4.1.3 Content Partners

Er werden cd-r's van **16 verschillende Content Partners** geregistreerd in dit digitaliseringproject.

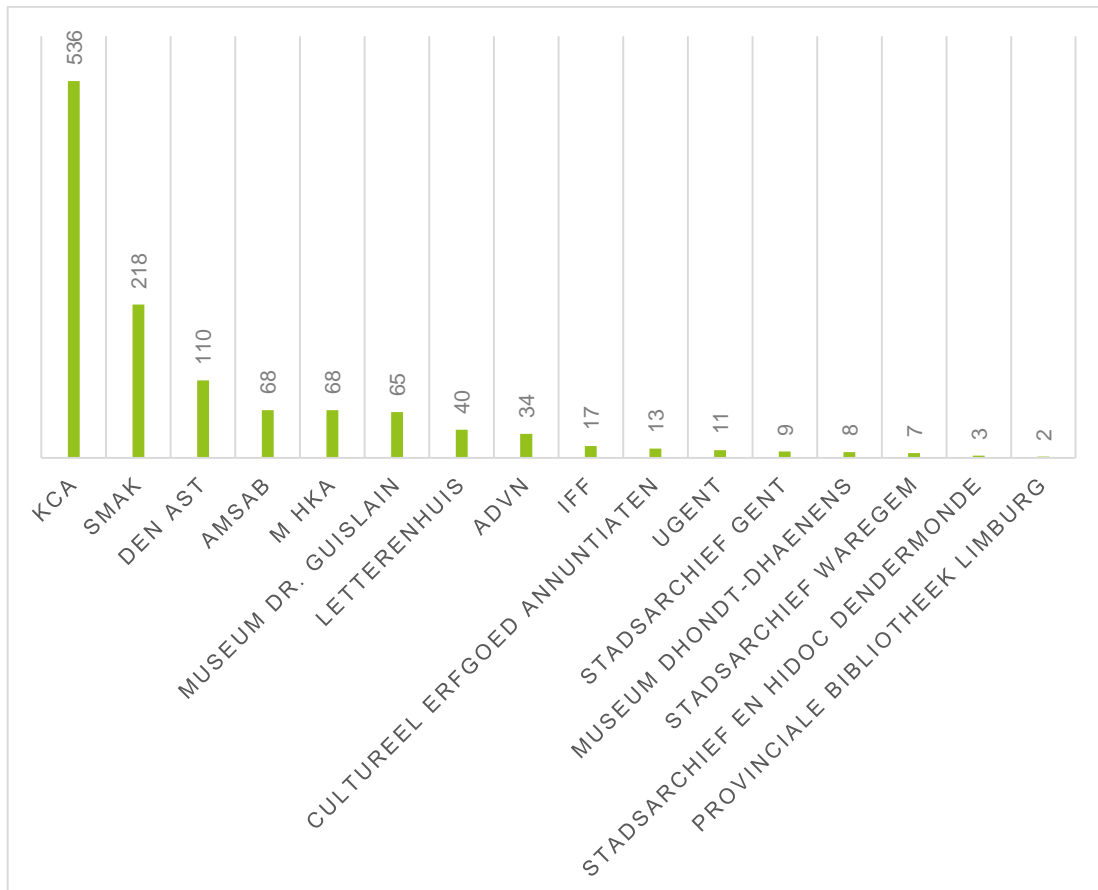


Fig. 5: de aantallen van de verschillende Content Partners in aflopende volgorde.

KC Antwerpen heeft met 536 het grootste aantal cd-r's. PBL bezit met 2 stuks het kleinste aantal.

4.1.4 Merk

Tijdens de registratie werden **8 verschillende merken** opgegeven. **TDK** kwam van de gekende merken het meest voor: 156 keer, of 13% van het totaal. 'Onbekend' en 'Andere' werden opvallend veel aangeduid, resp. 395 en 253 keer. In totaal vertegenwoordigen die 54 %, dat is meer dan de helft.

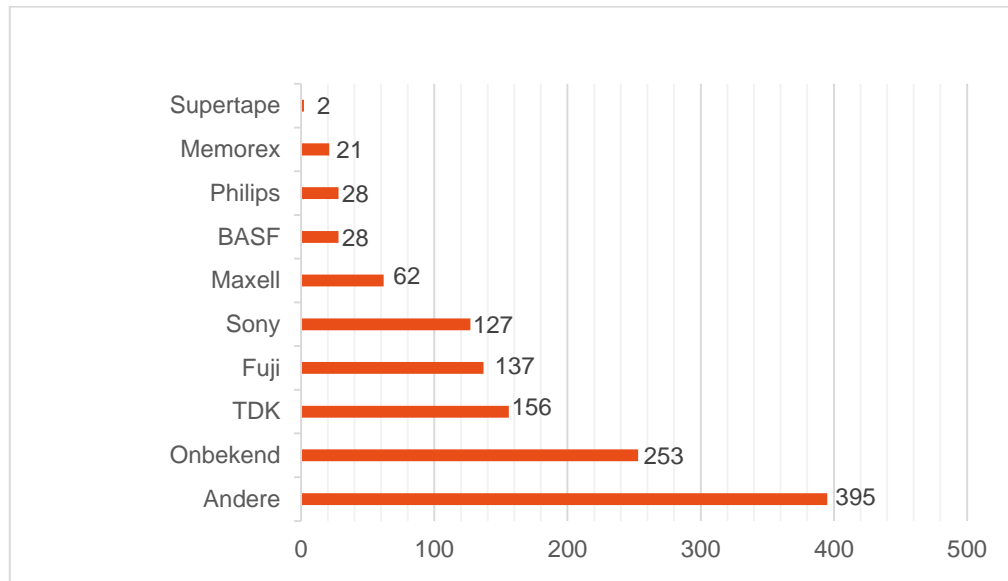


Fig. 6: de merken van de cd-r's in oplopende volgorde.

4.1.5 Datum

Om de ouderdom van de cd-r's in te schatten, werd naar de 'productiedatum' gevraagd van de inhoud. Doordat de productiedatum van de cd-r's soms moeilijk te achterhalen is, vroegen we om de dichtstbij liggende datum in te geven, hetgeen meestal de datum van opname was. In totaal werden er **32 verschillende jaartallen** genoteerd.

De jongste dateert uit 2015, dat is vorig jaar. Er zijn ook enkele cd-r's aangemeld met een datum die voor het op de markt komen van de cd-r zelf ligt: 1950, 1957, 1963, 1964, 1965, 1978, 1979 en 1987. Dit kan er op wijzen dat de inhoud een kopie is van een oudere drager.

De meeste van de gedateerde cd-r's dateren uit het jaar **2006**: 158 stuks of 13% van het totaal. In die tijd was het uiteraard al duidelijk dat de audio-cd-r geen geschikt medium is voor duurzame langtermijnopslag, maar de CP's hebben de meerderheid van deze dragers dan ook niet zelf geschreven, maar verworven van hun archiefvormers.

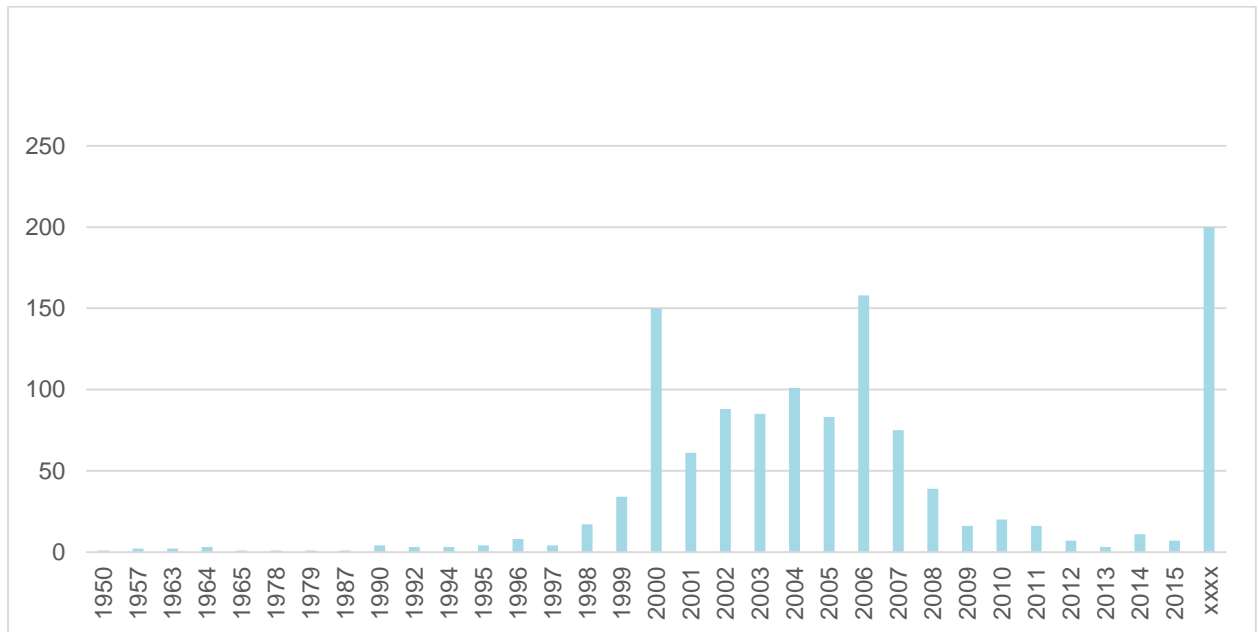


Fig. 7: de geregisteerde productiedatum van de inhoud van de cd-r's per jaar.

Deze statistiek heeft ook een inhoudelijke waarde en geeft dus een eerste indicatie in verband met de annotatie, als basis voor de toekomstige ontsluiting. Van ongeveer 17% van de opnames afkomstig van de cd-r's is er zelfs geen jaar bekend en zullen archivariissen en collectiebeheerders dus volledig moeten trachten te dateren op basis van de inhoud en/of hun archiefcontext.

- Enkel jaar en maand gekend: 38 stuks
- Enkel jaar gekend: 561 stuks
- Jaar, maand en dag onbekend (xxxx-xx-xx): 200 stuks

Dit betekent dat 410 cd-r's (34%) tot op de dag gedateerd zijn. Gezien het feit dat dit een recente drager is (die nu nog steeds frequent gebruikt wordt), is dat een vrij klein percentage.

4.1.6 Duur

Volgens de registratie vertegenwoordigden de cd-r's in het totaal **874:47:25 uur**, dat is iets meer dan de helft van de inschatting in de tender. De kortst geregisteerde duurtijd was 00:00:08; de langste 06:02:44. Dit laatste is vermoedelijk een fout, vermits er niet meer dan 80 minuten op een audio cd-r kan staan.

De duurtijd die het meest werd aangeduid (na 00:00:00) is **01:20:00** (79 keer). Dit is de maximum duurtijd die werd meegegeven in de handleiding en werd vermoedelijk door de CP's aangeduid omdat ze niet over extra informatie beschikten over de duurtijd. Het feit dat 00:00:00 zo vaak werd aangeduid (80 keer), heeft te maken met dezelfde reden. De gemiddelde duurtijd van de geregisteerde cd-r's is **00:43:25**.

4.1.7 Deterioratiefenomenen

Er werd aan de Content Partners gevraagd om in AMS aan te geven indien er bepaalde deterioratiefenomenen voorkwamen.

Bij cd-r kunnen er verschillende fenomenen waargenomen worden. Volgende opties konden aangeduid worden door de CP's tijdens de registratie (meerdere mogelijk):

- Verkleuring
- Krassen
- Stift
- Schilfers
- Audio speelt niet af
- *Disc rot*

Deze fenomenen zijn ofwel met het blote oog waar te nemen, ofwel door de schijf af te spelen.

Een fenomeen dat specifiek is voor cd-r is *disc rot*. Disc rot betekent dat de schijf onleesbaar is, door toedoen van fysische of chemische factoren die de reflectieve laag van de schijf aantasten.

Men kan disc rot op 2 manieren herkennen:

- als men de schijf tegen een sterk licht houdt, ziet men licht schijnen door gaatjes de grootte van een speldenprik
- als men een verkleuring waarneemt op de schijf, die lijkt op een koffievlek.

Ook 'Andere' en 'Geen' waren opties die konden aangeduid worden als er respectievelijk een deterioratiefenomeen zichtbaar was dat niet in de lijst stond, of als er geen deterioratiefenomenen zichtbaar waren. Dit laatste staat altijd standaard aangeduid, zodat in het geval er zich geen deterioratie voordoet, de Content Partner dit veld niet telkens moet invullen.

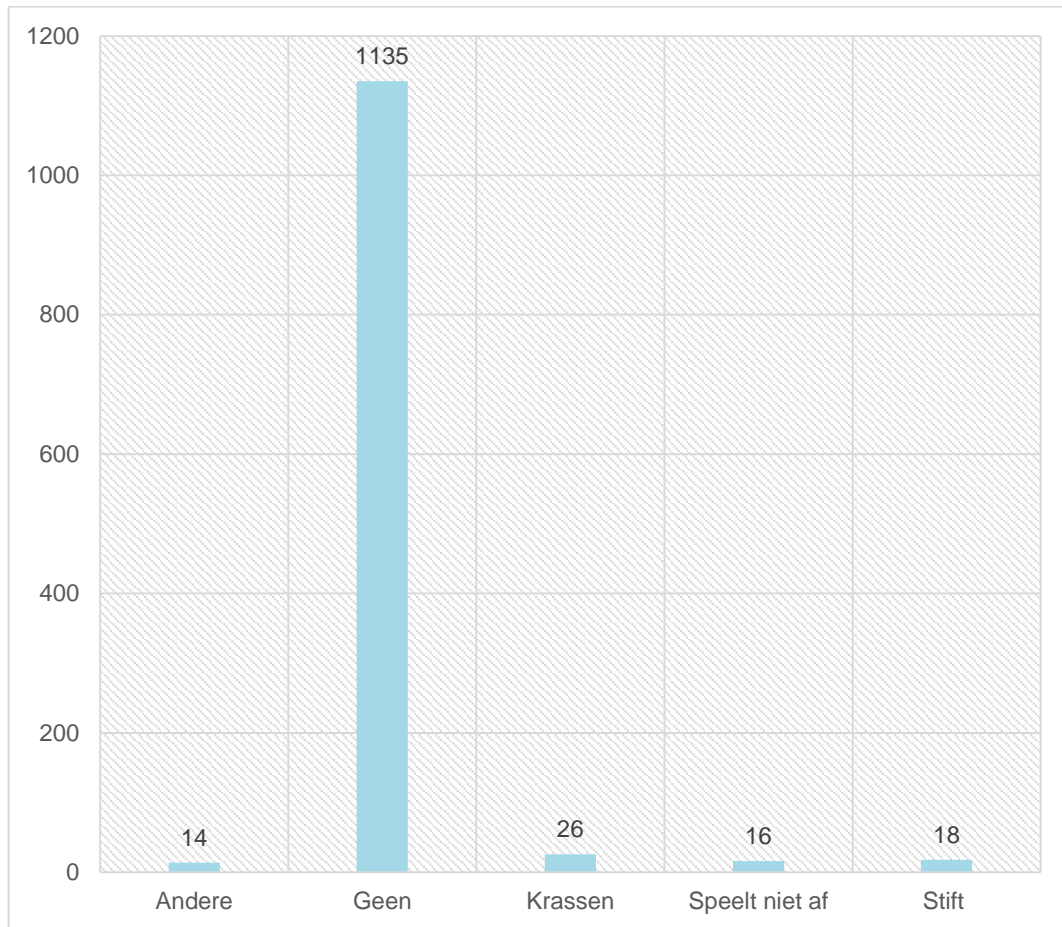


Fig. 8: de deterioratiefenomenen zoals aangeduid door de CP's in de registratie, met 'Geen' als duidelijke uitschieter.

Volgens de Content Partners waren de cd-r's in redelijk goede staat, slechts **74 schijven** vertoonden deterioratiefenomenen zoals Krassen (26), Stift (18), Speelt niet af (16) of Andere (14). De rest vertoonden volgens de Content Partners geen tekenen van deterioratie.

4.1.8 Bewaard door een andere VIAA partner

Er werd aan de Content Partners ook gevraagd of zij weet hadden van kopies of originelen die bij andere partners van VIAA bewaard zouden kunnen worden. In 35 gevallen of 3% werd deze vraag positief beantwoord. Deze partner was in 34 gevallen VRT, in 1 geval werd niet aangeduid welke partner dat was.

4.2 Resultaten van de digitalisering

Er werd aan Service Provider Picturae gevraagd om bepaalde informatie over de cd-r's te verzamelen voor, tijdens en na de digitalisering. Deze gegevens worden hieronder geanalyseerd.

4.2.1 Het digitaliseringsproces

Picturae gaf in zijn offerte meer informatie over de behandeling en de digitalisering, of in dit geval het rippen van de cd-r's.

4.2.1.1 De voorbereiding

Het door Picturae gebruikte opvolgingsysteem ondersteunde het volledige proces en maakte het mogelijk alle onderdelen en handelingen van het proces vast te leggen. De software garandeerde een transparant productieproces en bood een grote flexibiliteit om een groot aantal verschillende taken te realiseren. Dit is een belangrijke eigenschap aangezien audio- of videocollecties onderling vaak van elkaar verschillen. Voor elk nieuw project kunnen nieuwe eisen gelden en andere parameters worden ingesteld in het systeem. Na invoer en controle van de metadata (in xml-formaat), werden er taaklijsten gegenereerd voor de operators.

Tijdens het proces werden met grote regelmaat zogenaamde *health checks* uitgevoerd. Elke vijftigste te migreren cd-r was de referentie cd-r. Picturae gebruikte de data en metadata van deze cd-r om vergelijkingen te maken en een eventuele degradatie van de lees kwaliteit vast te stellen, bijvoorbeeld door een vuil wordende lens van de laser.

Tijdens de voorbereiding werd de metadata met de bevindingen door de operators in de workflow manager ingebracht. Deze metadata bevat ook informatie over de fysieke conditie van elk object, de inspectie en het eventuele herstelwerk dat werd uitgevoerd.

Elke cd-r werd door de operator voorzichtig uit de *jewel case* of andere verpakking gehaald. De operator droeg hierbij altijd handschoenen. Als eerste stap werd de barcode van de drager gescand en de beschikbare metadata in het opvolgsysteem van de SP opgeroepen. Vervolgens werd de cd-r aan een inspectie onderworpen en werden verrichte handelingen en bevindingen in het opvolgsysteem genoteerd.

Elke cd-r werd met (olievrije) perslucht schoon geblazen waarbij de grootste zorg in acht werd genomen om de disc niet te beschadigen. In het geval dat er vuil kleefde aan de schijf werd deze gereinigd met gedestilleerd water en een zachte katoenen doek. Bij het reinigen van de disc met een doek werd er altijd

vanuit het midden naar de rand bewogen en werd dus geen cirkelvormige beweging gemaakt. Dat verhoogt namelijk het risico op krassen. Ernstig beschadigde discs, met daarin bijvoorbeeld diepe krassen, werden apart gelegd en aan VIAA gerapporteerd.

4.2.1.2 De digitalisering

Picturae gebruikte voor het rippen van de audio-cd-r's een opensource software. Het is een Compact Disc Digital Audio (CDDA) Digital Audio Extraction (DAE) tool. De applicatie leest de audio van de audio-cd-r direct als data, zonder analoge stap ertussen, en schrijft de data naar een .wav, 16 bit lineaire PCM file.

De applicatie concentreert zich volledig op het rip-proces en daarnaast op het uitlezen van zo veel mogelijk informatie over de CD-drive zelf. Het kan zelfs de audiodata die ooit op goedkope CD's is gebrand corrigeren. Deze schijven zijn namelijk gevoelig voor bijvoorbeeld verkeerde uitlijning en frame jitter. De applicatie repareert ook de gegevens van CD's die op één of andere manier zijn beschadigd.

De applicatie wordt via *command line* aangestuurd en is een zeer flexibele, maar ook betrouwbare oplossing. In de *library* bevinden zich allerlei opties die het mogelijk maken het proces en de voortgang te volgen, errors te corrigeren en te loggen. De uitleessnelheid kan worden ingesteld. VIAA stond een maximale uitleessnelheid van *8x real time* toe. De tool kent een '*analyse CD drive*' commando waarbij een volledige analyse van de caching, timing en het leesgedrag van de drive wordt gelogged.

Nadat de operator de barcode heeft gescand en de drager correct in de lade van de drive heeft ingelegd, begint de tool automatisch met het extraheren van de data. Gedurende het extractieproces werden de .wav-bestanden lokaal opgeslagen. De operator bediende tussen 6 en 8 drives tegelijkertijd.

Gedurende de extractie kon de operator de voortgang en status van het proces aflezen. Na afloop van het proces werd duidelijk met welk resultaat de extractie is uitgevoerd. De status van de extractie werd in het opvolgingssysteem opgenomen. Een log van het proces (resultaten en correcties) werd automatisch gegenereerd en in het juiste formaat opgeslagen.

4.2.2 Duur

Tijdens de digitalisering zelf werd de duur van de opname op de cd-r's opgemeten. In de regel werd de hele cd-r van begin tot einde gedigitaliseerd. Van de 990 gedigitaliseerde cd-r's was de kortste 00:00:07, de langste duurde 01:19:57.

- De gemiddelde duurtijd van de gedigitaliseerde cd-r's is **00:38:34**
- De totale duurtijd van alle gedigitaliseerde cd-r's is **636:14:39**

Met dit laatste cijfer kunnen we ook een vergelijking maken tussen de duurtijd zoals ingeschat tijdens de inventarisatie (een cijfer dat ook gebruikt werd bij de opmaak van de aanbesteding), de duurtijd zoals die uit de registratie naar voren kwam en de duurtijd zoals die werkelijk bleek te zijn tijdens de digitalisering. Dat leidt tot deze conclusie:

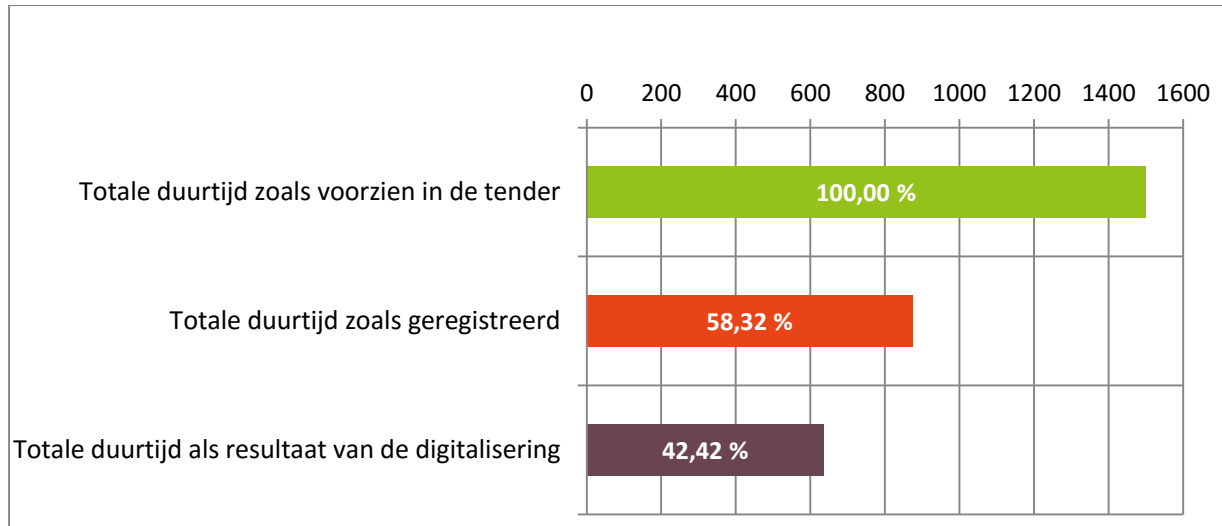


Fig. 9: vergelijking tussen de totale duurtijd van alle audio-cd-r's samen, zoals voorzien in de aanbesteding, zoals geregistreerd en als resultaat van de digitalisering.

4.2.3 Niet-gedigitaliseerde cd-r's

Tegen het einde van het project kreeg VIAA een rapport van Picturae met 217 cd-r's die niet gedigitaliseerd konden worden en de reden waarom. Deze problemen konden over het algemeen in vier categorieën worden ondergebracht:

- Afgebroken na 20min geen signaal (11)
- Foutmelding of blijft hangen (3)
- Niet leesbaar (200)
- Niet leesbaar, geen cd-r (3) maar wel:
 - Mini CD
 - DVD-r
 - MP3-cd

Uiteindelijk zijn bij deze 217 nog 2 niet te digitaliseren cd-r's bij gekomen, wat het totaal op **219** of 18% zet.

V. Conclusies en Lessons Learned

5.1 Transport

VIAA verwacht in haar digitaliseringsprojecten dat de Service Provider het transport verzorgt van de dragers heen en terug. Tijdens het transport worden de kartonnen dozen met dragers geplaatst in plasticen **transportbakken**, die voorzien worden door de SP. Normaliter haalt de SP alle dragers op bij de verschillende CP's die in één batch zitten. Vermits Picturae echter het transport van en naar Heiloo doet, en dus niet zoveel plaatsen op één dag kan aandoen, is met VIAA overeengekomen dat VIAA zelf de dozen op één centrale locatie verzamelde per batch. Dit betekende extra werk en verantwoordelijkheid voor VIAA en was een tegemoetkoming van VIAA aan de SP.

Achteraf gezien heeft VIAA, bij kleine projecten als dit, het werken met transportbakken voorzien door de SP als minder positief ervaren, omwille van:

- De extra transporten van lege bakken (heen en terug)
- Geen controle over de kwaliteit van de bakken

VIAA kiest ervoor om in toekomstige projecten eigen transportbakken te voorzien.

5.2 Verpakking en Barcodering

Voor de cd-r werden **dozen op maat** gemaakt. Bij het kiezen van de maat werd rekening gehouden met:

- de grootte en het gewicht van de cd-r's zelf
- de maat van de transportbakken waarin de dozen werden geplaatst tijdens transporten
- eventueel hergebruik van bestaande doostypes uit de eerste Digitaliseringsgolf

De doos had volgende afmetingen:

Type nummer	Afmetingen (hxbxl in mm)	Dragerformaat	Aantal cd-r's per doos
Doostype 5	150 x 150 x 450	cd-r	45

In tegenstelling tot eerdere digitaliseringsprojecten van audio- en videomateriaal, werd er op deze drager geen **barcode** gekleefd. Dit omwille van het schadelijke effect van lijm op de optische schijf en het eventueel blokkeren van de afspeelapparatuur door het label tijdens het afspeelen.

Omwillen van de kleine oppervlakte op de dragerdoos en het niet overplakken van metadata die aanwezig zou zijn op de doos, hebben we ervoor geopteerd om de kleinste barcodes te gebruiken (50 x 12 mm). Het gebruik van de barcodes is zeer efficiënt gebleken, er is geen enkele drager verloren gegaan.

5.3 Registratie en digitalisering

Voor het bepalen van de velden die moesten worden ingevuld door de Content Partner in het registratieformulier, werd uitgegaan van de volgende drie criteria:

- 1) Het veld is **nuttig** voor de logistieke opvolging (doorheen het volledige traject van de drager) of het is nuttig voor de SP in het uitvoeren van de digitalisering.
- 2) De inhoud van het veld is met **enige zekerheid** aan te leveren. Het heeft geen zin een veld aan te leveren waarvan de inhoud zo onbetrouwbaar is dat het de SP niet helpt.
- 3) De inhoud van het veld invullen is **technisch en qua timing haalbaar** voor de CP. Anders kost dit te veel tijd, zeker als men de deadline van de registratie wil halen.

5.3.1 Merk

Uit de registratie bleek dat 54% van de cd-r's een onbekend of ander merk had, dan tijdens de registratie was opgegeven. Dit is onder andere te verklaren doordat de CP's voor een relatief groot aantal dragers geen merk had opgegeven.

5.3.2 Datum

Er werd maar 34% van de cd-r's tot op de dag gedateerd, wat vrij weinig is voor zo'n recente drager. De meeste cd-r's dateren volgens de registratie uit de periode 2000-2010. Uit de digitaliseringsmetadata kunnen we niets afleiden over de productiedatum. Op basis van de digitale kopie in het VIAA archief zullen de Content Partners verder moeten bepalen wanneer de content werd geproduceerd, zodat deze informatie meer volledig en betrouwbaar wordt.

5.3.3 Duur

In de tender werd uitgegaan van een gemiddelde van 00:48:00 per cd-r, uit de registratie werd een gemiddelde duurtijd van 00:43:25 afgeleid en uit de eigenlijke digitalisering blijkt dat de cd-r's in dit project gemiddeld 00:38:34 duurden.

In de tender werd er dus door ons een overschatting gemaakt van ca. 10 minuten of ongeveer 20%, terwijl de CP's er dichter bij zaten maar alsnog de duurtijd gemiddeld 5 minuten langer inschatten dan de realiteit.

5.3.4 Deterioratiefenomenen

Op een totaal van 1209 cd-r's werden door de Content Partners maar 74 deterioratiefenomenen waargenomen. Er konden uiteindelijk echter 219 cd-r's niet gedigitaliseerd worden, dat is bijna drie keer zoveel. Hier past wel de opmerking, dat het eigen is aan audio-cd-r's dat men er vaak niet aan kan zien of ze afspeelbaar zijn of niet. Een zichtbare beschadiging is een aanwijzing, maar vormt geen zekerheid. Anderzijds kunnen ook audio-cd-r's waaraan geen schade of deterioratie te zien is, toch onafspeelbaar blijken.

De digitaliseringsmetadata zoals aangemaakt door Picturae geven niet meer informatie over de reden waarom de cd-r's niet gedigitaliseerd konden worden behalve de standaardformulering 'Niet leesbaar'. Om de precieze reden te kennen, zou een diepgaandere analyse nodig zijn, bijvoorbeeld door een bit-per-bit uitlezing van de schijf, waarbij de locatie van de bits wordt uitgezet op een schematische voorstelling van de schijf. Een dergelijke analyse viel buiten de scope van het project, maar gezien het hoge aantal dragers dat niet gedigitaliseerd kon worden (18%), zou VIAA dit wel in een zijproject meer in detail kunnen onderzoeken.

5.3.5 Overige velden

- **Gerelateerde documenten in de dragerdoos?**

Slechts in 108 gevallen (9%) werd deze checkbox aangevinkt. Dit lijkt weinig, maar voor de content providers is het wel interessant snel te kunnen zien bij welke cd-r's er papieren documenten zaten, zodat ze indien gewenst snel weer bij elkaar gebracht kunnen worden na de digitalisering.

- **Bewaard bij een andere VIAA partner?**

Ook dit veld werd slechts in 3% van de gevallen aangeduid, in bijna alle gevallen was dit VRT. Deze checkbox werd toegevoegd zodat men hierop in het archiefsysteem kan filteren om eventuele dubbels eruit te halen of andere collecties te verrijken. Het moet nog blijken of deze filter effectief zal gebruikt worden. Toch is dit gegeven ook op een andere manier interessant:

- Het biedt de meest aangeduide Content Partner (VRT) de kans om in deze collecties op zoek te gaan naar eventuele archiefstukken die zij zelf niet meer hebben.
 - Het biedt de oorspronkelijke partner de kans om deze stukken eventueel over te dragen aan de VRT of, indien het om dubbels gaat, te schrappen uit hun eigen collectie.
 - Het biedt VIAA de kans om dubbels in kaart te brengen en hierrond een procedure te ontwikkelen.
- **Oorspronkelijk dragernummer**

Dit vrije veld werd voorzien zodat de CP hier het inventarisnummer van de drager kon invullen, zoals gekend binnen de organisatie. Op die manier is er altijd een link tussen de VIAA-nummering (PID, i.e. unieke nummering automatisch toegekend aan het record in AMS, en de barcode) en het inventarisnummer van de CP. VIAA heeft steeds aangemoedigd om, in het geval er nog geen oorspronkelijk dragernummer bestond bij de CP, deze gelegenheid aan te grijpen om dit te toe te kennen.

Dit veld werd in 997 gevallen ingevuld (82%), en werd dus nuttig bevonden door de CP en door VIAA.

- **Standplaats drager**

Dit vrije veld werd toegevoegd in de registratie zodat de CP na de de dragers makkelijk en snel terug kan plaatsen in het archief. Dit veld werd in 1069 gevallen (88%) ingevuld, wat relatief veel is. Aangezien het geen verplicht veld was en slechts een minimale inspanning vraagt van de CP om het in te vullen, kunnen we aanwezigheid van dit veld als positief evalueren.

5.3 Algemeen

Het niet correct weergeven van bepaalde gegevens tijdens de registratie kunnen we op twee manieren verklaren.

Eenzijds kan VIAA er niet van uitgaan dat elke registrator dezelfde technische kennis heeft wat betreft het identificeren en het inspecteren van audiovisueel materiaal. Om de registrator bij te staan heeft VIAA echter wel een zeer gedetailleerde handleiding geleverd (zowel op papier als digitaal), waarin de meeste kenmerken van de cd-r staan uitgelegd en hoe deze te herkennen. We kunnen niet achterhalen of deze handleiding effectief gebruikt werd.

Hoewel de digitaliseringsprojecten van VIAA een grote hulp zijn voor de Content Partners om de digitaliseringachterstanden in te halen, mogen we anderzijds ook niet vergeten dat de registratie, barcodering en verpakking van de dragers moet gerealiseerd worden naast hun dagelijkse werk. Voor Content Partners met grote hoeveelheden is dit een groot deel extra werk dat erbij komt. Beide zaken kunnen dus worden aangehaald als reden waarom de registratie niet altijd even grondig gebeurde.

Van de kant van VIAA waren er enkele veranderingen in de scope, de timing en het budget. Wat betreft **de scope** valt het verschil op in aantallen en duurtijd, tussen het uitgangspunt van het project en het uiteindelijke resultaat.

Het moet daarbij worden opgemerkt dat VIAA geen enkel zicht had op de correctheid van de tijdens de inventarisatie opgegeven aantallen. Het feit dat het aantal cd-r's een derde lager lag dan oorspronkelijk ingeschat, lag aan een verkeerde opgave van cijfers bij één CP in het bijzonder, maar ook aan een lichte overschatting bij de meeste andere CP's. VIAA vertrouwt hier op de kennis van de collectiebeheerders bij het invullen van de inventaris en heeft niet de middelen noch het personeel om dit te gaan controleren. Daarnaast bleek de duurtijd per CD een zesde lager te zijn dan oorspronkelijk ingeschat, en tot slot was ruim 18% van de dragers onleesbaar.

Alle factoren bij elkaar opgeteld zorgen ervoor dat het project uiteindelijk maar ruim 42% opleverde van de oorspronkelijk ingeschatte duurtijd, en 55% van het oorspronkelijk ingeschatte aantal bestanden.

Wat **de timing** betreft, bleef het project grosso modo binnen de vooraf opgestelde grenzen.

Tot slot moeten we **enkele bijzondere slotopmerkingen** maken over dit eerste digitaliseringsproject van optische schijven van VIAA. Enerzijds werden door de CP's slechts enkele schijven foutief als audio-cd-r herkend en meegegeven voor digitalisering. Dit is een bijzonder succes te noemen, aangezien verwarring met andere optische schijven bijzonder makkelijk kan voorvallen. Anderzijds lag de *dropout rate* (het aantal dragers dat niet gedigitaliseerd kon worden) bijzonder hoog in vergelijking met alle andere digitaliseringsprojecten van VIAA.

Dit heeft er ons toe gebracht een speciaal project rond de problematiek van de diverse optische schijven te organiseren in de zomer van 2016, waarin bewustmaking rond gebruik in de archiefwereld en preservatie centraal zullen staan.

We danken alle Content Partners voor hun inzet en flexibiliteit om dit project samen met VIAA tot een goed einde te brengen.