



TULIPANA PROJECT



Fotocollectie digitaliseren – Handleiding, Standaards,
Richtlijnen voor scannen, opslaan en metadateren

Door: Marco Roling (CGHD)

Document Versie: 9/16/2016

TULIPANA PROJECT

FOTOCOLLECTIE DIGITALISEREN – HANDLEIDING,
STANDAARDS, RICHTLIJNEN VOOR SCANNEN, OPSLAAN EN
METADATEREN

Table of Contents

1. INTRODUCTIE	3
2. ORDENING EN NAAMGEVING.....	3
De archief/collectie-instelling	3
Naamgeving collectie(s).....	4
3. HET DIGITALISERINGSPROCES.....	5
Selecteren en prioriteren	5
Vorbereiden van de te digitaliseren foto's.....	6
Scannen	6
4. DIGITALE OPSLAG.....	7
Locaties en media.....	7
Mappen structuur.....	8
5. METADATA VASTLEGGEN.....	9
Metadata elementen voor foto-beschrijvingen.....	10
6. VRAGEN EN ANTWOORDEN	12
Taalkeuze en vertaling	12
Archivistisch indelen foto-collectie.....	12
Het gebruik van thema's	12
Foto's op volgorde verwerken.....	13
Bestaande scans	13

TULIPANA PROJECT

Meerdere exemplaren van een foto.....	13
Verzamelingen scans.....	13
Volgorde van metadata elementen.....	14
Het invullen van specifieke metadata elementen en hun betekenis	14
Voor- en achterkant van een foto scannen en beschrijven	15
Gebruik van trefwoorden	16
Opslag in de cloud	17
Het vermelden van personen	18

1. INTRODUCTIE

Deze handleiding is bedoeld als leidraad voor het digitaliseren, digitaal opslaan en metadateren van foto-collecties en is bedoeld voor de personen die in de praktijk het digitaliseren begeleiden of zelf uitvoeren.

Getracht is om deze handleiding zo eenvoudig mogelijk te houden om ook zonder voorkennis aan de slag te kunnen gaan. De handleiding is gebaseerd op voorbeelden uit de praktijk, waarbij er niet één de facto standaard is voor het digitaliseren en beschrijven van foto-collecties. Daarom zijn er ook diverse keuzes gemaakt bij de totstandkoming van deze handleiding en worden concrete voorstellen gedaan om niet voor moeilijke keuzes te komen staan.

Gebruik van Microsoft Excel wordt hier aanbevolen om de metadata in eerste instantie vlot te kunnen invoeren en hiervoor is er een template gemaakt. Verder is er uiteraard een scanner en opslag op een interne of externe harde schijf nodig voor de digitale bestanden zelf.

Deze handleiding is niet bedoeld als leidraad voor de collectie-beheerder als het gaat om het conditioneren, ordenen en acquireren van collecties. Deze handleiding is niet gericht op het daadwerkelijk online publiceren en beschikbaar stellen van de gedigitaliseerde bestanden, maar slechts op de digitale voorbereiding daarop.

2. ORDENING EN NAAMGEVING

Voordat de foto-collectie(s) gedigitaliseerd worden is het zaak om een ordening aan te brengen in de collecties en voor een gestandaardiseerde naamgeving te kiezen. Dit zijn voorbereidende activiteiten voordat de digitalisering start, en die vereisen dat er ook keuzes gemaakt worden ten aanzien van bewaren, selecteren, en prioriteren.

De archief/collectie-instelling

Om de digitale bestanden uniek te kunnen aanduiden is het verstandig om een gestandaardiseerde naamgeving te gebruiken. Deze naamgeving is niet alleen nuttig voor de digitale bestanden maar kan ook worden gebruikt in het vastleggen van de archief- en collectiebeschrijvingen volgens internationaal geaccepteerde richtlijnen.

TULIPANA PROJECT

Als eerste stellen we vast dat elke instelling die een archief en/of collectie beheert een unieke identificatie nodig heeft. Dit kunnen we doen volgens de ISIL-standaard¹.

Hierbij kiezen we voor de combinatie van een landcode en een instellingscode. Het voorstel is om de volgende codes te gaan gebruiken:

- Brazilië, Museum Holambra = BR-HOLMBRMUS
- Brazilië, Museum Castrolanda = BR-CASTRLMUS

Naamgeving collectie(s)

Vervolgens hebben beide musea één of meerdere foto-collecties, die elk moeten worden aangeduid met een unieke code. Deze code kan dan ook weer worden gebruikt bij het digitaliseren.

De oorsprong van een collectie ligt vaak bij een persoon of een groep van personen die een verzameling bij elkaar heeft gebracht. Een foto-collectie kan dus door het museum zijn gevormd, maar ook door een particuliere persoon of familie zijn aangelegd en aan het museum zijn overgedragen.



Het kan een keuze zijn om foto's chronologisch in series op te nemen, maar dat stuit al snel op problemen, dat de foto's inhoudelijk moeten worden beoordeeld en later wellicht verplaatst moeten worden als blijkt dat de datering niet klopte. Foto's kunnen ook thematisch worden geordend, maar het is vaak zo dat meerdere thema's van toepassing zijn op dezelfde foto, zodat een foto niet zomaar eenduidig kan worden gegroepeerd. Beter is veelal om de foto-collecties zoveel mogelijk gegroepeerd te houden rond de vorm van de collectie, maar dit is geen strenge eis.

¹ International Standard Identifier for Libraries and related organizations (ISIL), zie ook: <http://www.nationaalarchief.nl/informatiebeheer-archiefvorming/isil>

Eenvoudige voorbeelden:

- De fotocollectie van het Museum Holambra als geheel kan worden aangeduid als bijvoorbeeld FCOL–Mus01
- Als de Familie Jansen zijn particuliere collectie aan het museum overdraagt, dan kan het museum dit gaan aanduiden als bijvoorbeeld FCOL–Jansen01
- Als de foto–collectie van de Coöperatie aan het museum wordt overgedragen, dan kan dat worden aangeduid als bijvoorbeeld FCOL–Coop01

3. HET DIGITALISERINGSPROCES

Selecteren en prioriteren

Indien de hoeveelheid foto's groot is, is het verstandig om eerst vast te stellen in welke volgorde het digitaliseren plaatsvindt. Daarbij kan ook worden besloten om een deel van de collectie wel te beschrijven, maar niet te digitaliseren. De volgende overwegingen kunnen daarbij een rol spelen:

- De oudste foto's eerst digitaliseren, omdat die nu eenmaal sneller qua conditie achteruit gaan, en het feit dat de oudere migranten nu nog aan inhoudelijke beschrijving van de foto's zouden kunnen bijdragen
- De foto's die niet te digitaliseren zijn, vanwege een te slechte conditie, wel beschrijven, en verder zo goed mogelijk bewaren
- Hoewel misschien niet alle foto's direct openbaar gemaakt mogen worden, deze wel digitaliseren en in de beschrijving aangeven dat deze niet openbaar is
- Met de eigen museumcollecties beginnen en daarna pas met de collecties die het museum van derden heeft ontvangen
- Originele foto's die in tentoonstelling hangen eerst digitaliseren en vervolgens in de tentoonstelling vervangen door een kopie.
- Het kan zijn dat er meerdere exemplaren zijn van dezelfde foto. Het is aan de collectiebeheerder om te bepalen of beide foto's in de collectie moeten blijven. Elke foto is in principe uniek en heeft zijn eigen metadata record. Hierbij wordt afgeraden om bijvoorbeeld 'aantal foto's' als metadata element op te nemen. Raakt er een foto toch zoek, dan klopt meteen de metadata niet meer. Dus zijn er twee gelijke foto's en wordt er gekozen, dan degene met de meeste informatie (op de achterkant geschreven tekst misschien) en de beste kwaliteit opnemen in de collectie en de ander niet.

Vorbereiden van de te digitaliseren foto's

Het is aan te raden om voorafgaande aan het daadwerkelijk scannen, de foto's van een collectie op een gekozen volgorde te leggen om reden van efficiency. Veel vergelijkbare foto's hebben vergelijkbare metadata en dat is mogelijk handiger en sneller bij vastlegging van metadata. Op volgorde leggen kan bv. chronologisch, thematisch, of op volgorde van beschrijven.

Daarna kan op elke foto een unieke aanduiding worden geschreven, zodat deze later met de scan en de metadata kan worden gekoppeld. Breng bij voorkeur met potlood op de achterkant van de foto een unieke aanduiding aan, bestaande uit bijvoorbeeld de code van de collectie, en een volgnummer aanduiding.

Bijvoorbeeld: FCOL-Mus01_0003

Scannen

Zowel Holambra als Castrolanda werken met een HP scanjet vlakbed, A4 formaat (situatie: eind 2014). Hiermee kunnen de masterscans worden gemaakt.

De volgende richtlijnen worden geadviseerd om in acht te nemen bij het scannen van foto's naar een masterscan formaat:

- Scannen in JPG formaat
- Scannen op 300 ppi (pixels per inch, soms ook wel aangeduid als dpi)
- Scannen met een kleurdiepte van 24 bits
- Scan inclusief de randen van de foto, dus niet 'croppen' (afsnijden)
- Indien de achterkant ook tekst, informatie of relevante content bevat, dan deze apart scannen, en als een afzonderlijke scan beschouwen (hierover later meer bij de metadatering)
- Geen beeldbewerkingen uitvoeren na het scannen, zoals wijzigen van kleur, contrast, helderheid, scherpte. De scan in principe onbewerkt laten, maar zo goed mogelijk het origineel laten benaderen.
- De naam van het scanbestand is als volgt opgebouwd:
 - code van de instelling (bv BR-HOLMBRMUS), gevolgd door een '_'
 - code collectie (bv FCOL-Mus01), gevolgd door een '_'
 - item identificatie, veelal een volgnummer (bv 0001), gevolgd door .jpg

Voorbeelden:

TULIPANA PROJECT

BR-HOLMBRMUS_FCOL-Mus01_0002.jpg

BR-HOLMBRMUS_FCOL-Mus01_0003.jpg

- Het scannen kan tijdrovend worden, dus het is zaak om het proces zo efficiënt en effectief mogelijk te houden. Overbodige handelingen worden vermeden.

Er wordt overigens niet gescand volgens Metamorfoze standaarden. Dit is met de beschikbare apparatuur, mensen en middelen niet realistisch.²

Er wordt ook niet gescand in TIF formaat, vanwege de grootte van de bestanden die dan ontstaat en bijbehorende kosten van opslag en backup die dit met zich meebrengt.



Van de originele masterscans kunnen met een fotobewerkingsprogramma eventueel kopieën en andere afgeleiden worden gemaakt voor tentoonstellingsdoeleinden, publicatie of beeldbewerking, maar dat zijn dan duplicaten geworden, die de masterscan niet vervangen of vergezellen. Afgeleide duplicaten worden dan ook beslist niet bij de masterscans opgeslagen om verwarring te voorkomen!

4. DIGITALE OPSLAG

Locaties en media

De masterscans moeten op minstens twee plaatsen worden opgeslagen, om het risico op verlies of digitale schade te beperken. De eerste opslag (primaire) kan lokaal zijn, bijvoorbeeld in de PC van het scan-werkstation, op een aparte Hard Disk, of 'in de cloud'. De tweede opslag is de back-up en dient elders te worden opgeslagen.

Het heeft voordelen om de eerste opslag meteen 'in de cloud' te doen. De bestanden worden automatisch aldaar veilig gesteld, en er kan ook eenvoudig toegang aan anderen worden gegeven. Eventueel kan er vanaf verschillende locaties aan het scannen of meta-dateren worden gewerkt. Het gebruik van Dropbox werkt alsof het

² Metamorfoze is het Nederlandse programma voor het behoud en digitalisering van het papieren cultureel erfgoed. Zie: <http://www.metamorfoze.nl/digitalisering>

lokaal is, met een dropbox-map op de computer, die automatisch synchroniseert met de opslag in de cloud.³



Advies is om een betaald Dropbox abonnement in te zetten voor de eerste opslag. Voor 10 euro per maand is er dan 1TB aan opslag beschikbaar, hetgeen in de praktijk zeker genoeg zal zijn voor zo'n 800.000 JPG scans.

Advies is verder om een losse 1TB hard disk als externe backup lokaal te houden, en regulier (bv maandelijks) de extra bestanden van de eerste opslag vanuit Dropbox naar de backup te kopiëren. Deze harde schijf kan elders op een veilige plek worden bewaard.

Mappen structuur

Om de bestanden ook digitaal makkelijk terug te kunnen vinden is het handig om een overzichtelijke mappenstructuur te kiezen, die aansluit bij de indeling van het archief en de collecties. Dezelfde mappenstructuur kan zowel lokaal, als in de backup worden gebruikt.

Het eerder genoemde voorbeeld krijgt dan bijvoorbeeld een mappenstructuur als volgt:

- 📁 Digitaal Achief Holambra BR-HOLMBRMUS
 - 📁 Foto collecties (FCOL)
 - 📁 FCOL-Coop01
 - 📁 FCOL-Jansen01
 - 📁 FCOL-Mus01
 - BR-HOLMBRMUS_FCOL-Mus01 metadata V20150313.xlsx
 - BR-HOLMBRMUS_FCOL-Mus01_0001.jpg
 - BR-HOLMBRMUS_FCOL-Mus01_0002.jpg
 - BR-HOLMBRMUS_FCOL-Mus01_0003.jpg

³ Zie: <http://hoewerktdropbox.nl/> en <https://www.dropbox.com/>

- 📁 Archieven ARCH
 - 📁 Bestuur
 - 📁 Periodiek

5. METADATA VASTLEGGEN

Nadat er een masterscan is gemaakt van de foto, en deze is opgeslagen in de juiste map, dan kan er ook aanvullende metadata worden vastgelegd over de scan.

Metadata beschrijft de scan op drie mogelijke gebieden, namelijk:

- Aanvullende beschrijving over het scannen
- Aanvullende beschrijving van de originele foto
- Inhoudelijke beschrijving van de foto afbeelding (wie, wat, waar, wanneer)

De metadata kan in eerste instantie worden ingevoerd in een Excel-templete dat voor TULIPANA is gemaakt⁴. Deze templete kan per foto-collectie worden ingevuld en bij de collectie in de betreffende map bewaard. Zo blijven scans en metadata in de digitale opslag tenminste bij elkaar. De templete is in het Nederlands gemaakt, omdat we er op dit moment toch vanuit gaan dat de archieven en collecties in de Nederlandse taal zijn vastgelegd. Uiteraard is het Portugees de andere taal die we moeten ondersteunen om een breder publiek te bedienen, maar we moeten een effectieve keuze maken voor de eerste taal en daarna vertalen. Advies is om de inhoud van de metadata velden in eerste instantie ook in het Nederlands te gaan vastleggen.

De drie metadata gebieden hebben elk een groep van elementen, die per fotoscan kunnen worden ingevuld. In principe zijn alle elementen facultatief, ook voor de eerste twee metadata gebieden. Het kan namelijk zijn dat er óf geen origineel is, óf geen scan. De metadata van de content kan eventueel ook op een later tijdstip worden aangevuld, en kan afhankelijk zijn van de informatie die nog levende personen aan de foto kunnen bijdragen. Dit vormt aanleiding voor de organisatie van ‘crowd-sourcing’ en ‘social evenings’.

⁴ De basis templete is: TULIPANA fotocollectie metadata V20150325.xlsx

Metadata elementen voor foto-beschrijvingen

Hieronder worden de TULIPANA metadata elementen in het kort beschreven (geactualiseerd 20160916):

<i>Identificatie</i>	Identificatie van de digitale scan, volgens een vast formaat, zoals eerder beschreven. Bv: BR-HOLMBRMUS_FCOL-Mus01_0002.jpg
<i>Formaat scan</i>	Digitaal bestandstype, veelal een vaste waarde: image/jpeg
<i>Resolutie scan</i>	Aantal dots of pixels per inch zoals ingesteld tijdens het scannen, veelal een vaste waarde: 300dpi
<i>Datum scan</i>	Datum waarop de digitale scan is vervaardigd (YYYY-MM-DD) Bv: 2015-03-13
<i>Maker scan</i>	Naam van de digitaliseerder/instituut, zal voor alle scans van de collectie veelal hetzelfde zijn, tenzij de scans deels door iemand anders zijn aangeleverd.
<i>Gerelateerd aan</i>	Indien de scan hoort bij een andere scan van dezelfde foto, dan kan hier een referentie worden opgenomen naar de eerste scan. Bv: BR-HOLMBRMUS_FCOL-Mus01_0002.jpg is een scan van de achterkant van bijvoorbeeld scan 0001 In dat geval zetten we hier: BR-HOLMBRMUS_FCOL-Mus01_0001.jpg
<i>Originele collectie</i>	Een unieke aanduiding (code, tekst, afkorting) van de originele fysieke collectie. Deze aanduiding zal voor alle scans in dezelfde collectie dezelfde waarde bevatten
<i>Originele identificatie</i>	Identificatie, zoals op het origineel aangebracht. Het is aan te bevelen dat als er geen unieke identificatie op het origineel staat, om die aan te brengen. Anders is er naderhand geen verbinding meer te leggen tussen het origineel, de scan en de metadata. Het is acceptabel met potlood op de achterkant een identificatie aan te brengen als archivering. De aanduiding van de collectie en een volgnummer ligt het meest voor de hand.
<i>Afmeting</i>	Hoogte en breedte in cm, Bv. 7 x 15 cm
<i>Materiaal soort</i>	Eén van de volgende: foto, dia, negatief, drukwerk
<i>Kleur</i>	ja=kleur, nee=monochroom
<i>Publicatie herkomst</i>	Indien afbeelding afkomstig is uit een tijdschrift of anderszins, kan hier een referentie worden opgenomen naar die publicatie, bv Naam van de publicatie, nummer en jaar van uitgifte
<i>Donateur</i>	Naam van degene die de foto heeft afgestaan aan de collectie
<i>Titel</i>	Een korte woordelijke beschrijving

TULIPANA PROJECT

<i>Opname jaar</i>	Jaar van vervaardigen originele opname
<i>Opname maand</i>	Maand van vervaardigen originele opname
<i>Opname dag</i>	Dag van vervaardigen originele opname
<i>Periode van</i>	Begin van de periode, indien bv exacte datum opname onbekend is
<i>Periode tot</i>	Eind van de periode, indien bv exacte datum opname onbekend is
<i>Locatie</i>	Namen van plaatsen, locaties, aanduidingen etc. die op de foto zichtbaar zijn, of relevant worden geacht, zo specifiek mogelijk. Bv. Holambra, kruising straat A met straat B
<i>Personen</i>	Namen van afgebeelde personen, zo compleet mogelijk, eventueel met genealogische gegevens. Bv. Jan Jansen, geb. 12-3-1923, ovl. 12-12-1968; Maria Thesesia Jansen; Bertus Jansen
<i>Opschriften</i>	Tekst geschreven op de voor- of achterkant
<i>Beschrijving inhoud</i>	Een woordelijke beschrijving van de inhoud (max. ca 100 woorden). Dit is een stuk vrije tekst waarin kan worden beschreven wat er te zien is op de afbeelding, of wat er gebeurd.
<i>Trefwoorden</i>	Eén of meer termen uit een onderwerpenlijst. Deze thematische lijst kan tijdens het metadateren uitgebreid worden. Het is verstandig om deze onderwerpenlijst apart bij te houden, om overzicht te houden. Bv. Groepsportret, Nieuwjaar, Landbouw, Behuizing ...
<i>Naam fotograaf</i>	Naam van de fotograaf
<i>Publicatie rechten</i>	Eén van: openbaar, toekomstig openbaar, privaat. Dit element geeft aan of de foto of afbeelding publiek toegankelijk gemaakt kan worden.
<i>Exemplaren</i>	Het aantal exemplaren dat in de collectie aanwezig is als dat er meer dan één is (dus standaard is dat er één exemplaar aanwezig is, en dat noteren we niet)
<i>Afdruk</i>	Alleen 'nee' invullen als er alleen een digitale afbeelding is en geen fysiek origineel (dus standaard is dat er wel een afdruk aanwezig is, en dat noteren we dus niet)

6. VRAGEN EN ANTWOORDEN

Over een aantal praktische zaken is in onderling overleg nagedacht en daarbij zijn vragen (V) en antwoorden (A) geformuleerd over wat het advies is om wel of niet te doen. Dit om het gebruik van standaarden, scannen en het vastleggen van metadata zoveel mogelijk op een weloverwogen consequente en consistente wijze te doen. De volgende zaken zijn besproken (in willekeurige volgorde):

Taalkeuze en vertaling

V: Hoe gaan we om met taal in de metadata? Zowel de benamingen van de metadata elementen, als de inhoud van de metadata?

A: Het is in elk geval handig om de metadata elementen zoals die in de Excelsheet als standaard zijn gedefinieerd, om die van de Braziliaans/Portugese vertaling te voorzien, zodat we geen onduidelijkheid hebben over de terminologie van de elementen in beide talen. Dus in de excelsheet zal een vertaallijstje gezet worden met in 'Elementnaam PT' de vertaling zoals die in Brazilië kan worden gebruikt. De metadata kan in één taal worden ingevoerd en later in een kopie van de metadata worden vertaald in een andere taal. Het is niet eenvoudig om twee talen bij te houden, maar het is een verstandige keuze om twee talen uiteindelijk te doen, als de doelgroepen zowel Nederlandstalig als Portugeestalg zijn.

Archivistisch indelen foto-collectie

V: Hoe delen we onze foto-collectie in? Het is nogal arbitrair als er meerdere thema's en groeperingen mogelijk zijn. Moeten we de collectie wel indelen eigenlijk?

A: Alle foto's als één collectie beschouwen is natuurlijk prima en je vermijdt discussie over hoe je een bestaande collectie opnieuw moet indelen in thema's die toch niet eenduidig zijn. Dus alle foto's komen dan bv in BR-HOLMBRMUS_FCOL

Het gebruik van thema's

V: Hoe ga je om met thema's en neem je een thema op in de identificatiecode?

A: Foto's eerst meteen thematisch indelen stuit op het probleem dat er altijd wel meerdere thema's te bedenken zijn voor een foto. In de fysieke collectie zou ikzelf allereerst indelen op de vorm van de collectie, en daarbinnen misschien ordenen op thema maar dat is een zeer arbitraire keuze. Voor het vastleggen van metadata is het eigenlijk niet van belang, want alle thema's die je bedenkt zet je in de rubriek trefwoorden. Daar zou later digitaal op gegroepeerd moeten kunnen worden, zo moet je het zien. Dus je kunt later digitaal een deel van de collectie opvragen behorende bij een bepaald thema (trefwoord). Dus heb je een foto met een huis, met bloemen, met een familie, dan zou je bv thema's kunnen maken: woning, bloemen, familie, portret. Deze woorden zet je dan in de rubriek Trefwoorden.

Ik zou geen thema(s) in de naamgeving van de scan opnemen. Daar krijg je lange bestandsnamen van, en wat kun je daar in de praktijk mee. In een folder met bestanden kun je nauwelijks gestructureerd zoeken op thema, daar is juist de uitgebreidere metadata voor.

Foto's op volgorde verwerken

V: Heeft het zin om voorafgaand aan scannen of metadateren de foto's op een bepaalde volgorde te leggen?

A: Het is op zich zinnig om de foto's op een volgorde te leggen voorafgaande aan het scannen en beschrijven. Dat is geen regel, maar een advies, want als je een aantal vergelijkbare foto's hebt met dezelfde thema's en andere vergelijkbare metadata dan lijkt het handig ivm kopiëren en plakken van gegevens in de excelsheet, om dat zo efficiënt mogelijk te doen nietwaar. Maar elke foto staat op zichzelf in principe, dus een volgorde is niet echt noodzakelijk, hooguit dat bij digitaal bladeren deze volgorde de standaard volgorde wordt waarschijnlijk waarin foto's worden getoond.

Bestaande scans

V: Wat doen we met scans die we al hebben, maar niet voldoen aan de voorgestelde standaards?

A: Zijn foto's al gescand dan zou ik dat ook niet meteen overnieuw gaan doen, dat kan altijd nog als je dat wilt. Ik zou me dan inderdaad eerst op het beschrijven concentreren en later op het evt opnieuw scannen.

Meerdere exemplaren van een foto

V: Wat doen we als we meerdere exemplaren hebben van dezelfde foto? Alle exemplaren scannen en metadateren?

A: De vraag aan de collectiebeheerder is eigenlijk of beide foto's even interessant zijn. Als ze echt identiek zijn, zou je er één kunnen wegdoen. Als ze niet identiek zijn, en er is bv op de ene iets geschreven en de andere niet, dan zou ik degene bewaren met de meeste informatie. Als je besluit beiden in de collectie te houden, dan wel de foto's apart beschrijven. Zeker als er van beide een scan is, waar de metadata bij moet aansluiten. Dat is belangrijk om te onthouden, dat elke scan een metadata-regel in de Excel heeft, anders moet je later toch weer alles bij elkaar gaan zoeken, foto scan en metadata regel.

Verzamelingen scans

V: We willen krantjes gaan scannen, maar dat zijn dan eigenlijk scans die bij elkaar horen. Hoe gaan we daar om met de beschrijvingen en de metadata?

A: In archieftermen gesproken lijken mij de krantjes bij elkaar een *serie* (die moet dus een naam krijgen) waarin elk krantje zelf een *file* is. Binnen die file is elke bladzijde (voorkant en achterkant) een *item*. Afijn, de verschillende niveaus in de hiërarchie van een inventaris blijven maar terugkomen, en daar kom je niet onderuit in het spraakgebruik. Dat heeft niks met een tool (bv ICA-AtoM) te maken, maar heeft alles te maken met de terminologie waarin je een archief. Je kunt weliswaar best krantjes gaan scannen, maar dat je eerst even moet vaststellen hoe je de scans van die krantjes gaat noemen. Dus ik kan me voorstellen bijvoorbeeld dat de

eerste scan (0001) van het eerste krantje uit 1960 (196001) uit de collectie krantjes (KCOL) van het museum (BR-HOLMBRMUS) dan gaat heten: BR-HOLMBRMUS_KCOL_196001_0001.jpg Dit is een suggestie natuurlijk. Omdat nu duidelijk is dat elke bladzijde van een krantje een scan is, en alle bladzijden vrijwel dezelfde metadata zullen hebben, is het dus niet zinvol om met de fotobeschrijving-standaard alle individuele scans van de krantjes te beschrijven, want je krijgt overbodig veel dezelfde data. Metadata van een krantje vastleggen is wel zinvol, en je kunt besluiten wel met de fotobeschrijving-standaard voor elk krantje één regel in te vullen, met in elk geval de algemene metadata die voor alle bladzijden van dat krantje geldt. Het is goed denkbaar dat we alle scans van een krantje uiteindelijk digitaal samenvoegen in een doorzoekbare PDF, zodat je digitaal door het krantje kunt bladeren en zoeken. Dat is het ideale eindresultaat. Dat gaan we met ICA-AtoM niet doen, want daar is het niet voor geschikt.

Volgorde van metadata elementen

V: Moeten we de metadata elementen in de volgorde invoeren zoals in de template is opgezet? En kunnen we nog veranderingen in de elementen aanbrengen?

A: De volgorde van de metadata elementen is op zich niet belangrijk, die kan je altijd omgooien, dus in welke volgorde je wilt invoeren is aan jezelf. Je kunt in de excel template dus kolommen zonder problemen verwisselen. Je kunt ook kolommen weglaten. Ik zou dat zelf niet snel doen, omdat je dan misschien niet meer die metadata gaat bedenken als je er niet meer aan wordt herinnerd in de sheet. Kolommen toevoegen kan ook, maar overleg dat even, zodat we zeker weten dat het ook echt een 'nieuw' element is dat we over het hoofd hebben gezien.

Het invullen van specifieke metadata elementen en hun betekenis

V: Als je datering foto (*Opname jaar, maand, dag*) hebt ingevuld, moet je dan ook de tijdsperiode (*Periode van, tot*) nog aangeven?

A: Als de datering van de foto al exact is dan hoef je geen periode op te nemen, dat is overbodig.

V: Als we *Locaties* invullen, vullen we dan altijd 'Holambra' in bijvoorbeeld, gevolgd door een specifiekere aanduiding? En is het zinvol om een landaanduiding op te nemen?

A: In de beschrijving een land opnemen lijkt me alleen zinnig als het in dit geval van Brazilië afwijkt. Maar ik zou wel consequent een plaatsaanduiding opnemen, dus Holambra, en verder zo exact mogelijk de locatie.

V: Als we *Personen* invullen, dan lopen we tegen het probleem aan van meerdere voornamen die eigenlijk gelden, zoals een doopnaam en een roepnaam. Hoe gaan we hiermee om? Moeten we apart personen en namen gaan vastleggen? En hoe beschrijven we waar de persoon staat op een groepsfoto?

A: Een namenindex aanleggen (op een aparte tab in dezelfde excel met de metadata) is een slim idee om meteen mee te beginnen. In die index neem je op welke schrijfvarianten een naam heeft, zodat die index altijd als aparte lijst of iets dergelijks kan worden gepubliceerd. In excel kun je extra tabbladen maken, in de template heb ik er eentje toegevoegd. Is niet een heel

strikt gestandaardiseerd format overigens.

Ik zou overigens in de rubriek *Personen* alleen persoonsnamen opnemen en geen beschrijving van waar die persoon op de foto precies staat en zo. Die beschrijving zou ik in de rubriek *Beschrijving inhoud* doen, waarbij je de personen qua naam best kan herhalen, het is tenslotte vrije tekst in die beschrijving. Daar kun je iets zetten als: 'v.l.n.r. (van links naar rechts)' of 'staand', 'zittend'.

V: Als we *Personen* invullen, dan kan het zijn dat we de naam van iemand niet weten. Wat doen we dan in de beschrijving? Vraagtekens opnemen?

A: Je zou in de rubriek o.a. kunnen opnemen 'persoon onbekend', zodat in elk geval duidelijk is dat je niet iedereen hebt kunnen identificeren. Vraagtekens zeggen niks en zou ik niet gebruiken.

V: Hoe ga je om met de verdere beschrijving van de foto, als je ook al andere metadata elementen hebt ingevoerd zoals personen en plaatsen hebt ingevuld? Herhaal je de gegevens dan weer in het element *Beschrijving inhoud*?

A: Nee, vermijdt zoveel mogelijk dat dezelfde informatie in verschillende metadata velden terugkomt, dat heeft geen zin.

V: Niet alle foto's van bijvoorbeeld de collectie van Wim Welle zijn door hem zelf gemaakt. Wij willen daarom het gegeven toevoegen *-uit collectie van-*, om aan te geven van wie de foto komt. Is dat zinnig?

A: Als de fotograaf niet bekend is, dan laat *Naam Fotograaf* leeg. Verder stel ik voor het veld toe te voegen: *'Donateur'*, waarin de naam kan worden gezet van degene die de foto aan de collectie heeft afgestaan.

V: Wat is het verschil tussen de elementen *Originele collectie* en *Originele identificatie* en wat zet je daarin?

A: Het kan zijn dat dat niet van toepassing is. Het gaat hier om de naam van de originele fysieke foto-collectie, en een identificatie van elke fysieke foto. Maar het kan zijn dat je nog nooit de collectie en de foto's van unieke namen en identificaties hebt voorzien en dat dat nu eigenlijk pas echt voor de eerste keer gaat gebeuren. Het is in elk geval van belang dat je vanuit de metadata van de scan kunt teruglezen welke fysieke foto erbij hoort en andersom. Dus heb je een foto, dan moet je in de metadata kunnen vinden welke scan erbij hoort en andersom. Afijn, als je 'blanco' begint dan zou ik deze twee elementen 'originele collectie' en 'originele identificatie' gewoon leeglaten, het is niet verplicht ze te vullen. Maar het is dan geen gek idee om op elke fysieke foto de scan-identificatie te schrijven (dus bv: FCOL-Mus01_0003), zodat je dat altijd makkelijk kunt terugvinden in het metadata bestand. Benaming van het origineel en de digitale variant zijn dan feitelijk dezelfde geworden en dat is wel zo eenvoudig.

Voor- en achterkant van een foto scannen en beschrijven

V: Wanneer besluiten we ook de achterkant te scannen, want soms staat er wel wat op, maar niet echt interessant voor de foto-beschrijving.

TULIPANA PROJECT

A: Ik zou alleen de achterkant van de foto scannen als daar inderdaad iets opstaat dat echt iets origineels is en iets toevoegt aan de authenticiteit van de foto (bijvoorbeeld ook zoiets als een label, origineel handschrift of stempels of iets dergelijks is leuk om te laten zien op een scan)

V: Hoe gaan we om met de naamgeving van een foto, waarbij we een voorkant en achterkant scannen omdat dat beide interessant is? Hoe geven we aan dat die scans bij elkaar horen?

A: Het is een goed idee om de scans van de voor- en achterkant een extensie mee te geven, dus bv BR-HOLMBRMUS_FCOL_000001a en 000001b. Indien er geen scan van de achterkant wordt gemaakt dan zou ik geen letter a gebruiken, dat suggereert namelijk teveel dat er ook nog een b is.

V: Nemen we alle metadata van de scan van de voorkant van de foto over in de metadata van de scan van de achterkant van de foto? Dat is dus veel dezelfde data?

A: Metadata van de achterkant van de foto zou ik inderdaad zoveel mogelijk overnemen van de voorkant, je kunt overwegen om in het element '*gerelateerd*' een verwijzing naar de andere scan op te nemen, zodat in de metadata die twee scans naar elkaar wijzen ook.

V: Opschriften op de foto (voor- of achterkant) zijn vaak rommelig en voegen niet veel toe aan de beschrijving van de foto zelf? Moeten we de opschriften letterlijk in de metadata opnemen?

A: Opschriften alleen overnemen als het inhoudelijk iets toevoegt. Het idee is zoveel mogelijk informatie digitaal te krijgen, maar niet zaken dubbel op te nemen, dat heeft geen zin. En of opschriften nou origineel zijn of later aangebracht maakt niet veel uit, als het inhoudelijk iets toevoegt dan zou ik het opnemen.

V: Sommige achterkanten van foto's zijn zo glad dat je er geen identificatie op kunt schrijven met potlood? Wat zouden we kunnen doen?

A: Je zou beschrijfbaar zuurvrije conserverings-tape kunnen gebruiken om op de achterkant de identificatie van de foto te schrijven.

V: Schrijven we op de achterkant van de foto zowel de identificatie van de scan van de voorkant, alsook van de achterkant? Dus twee codes?

A: In geval van meerdere scans per foto, blijft op de foto zelf de identificatie gewoon eenmalig aanwezig, dus bv BR-HOLAMBRMUS_FCOL_000001 (zonder a of b)

Gebruik van trefwoorden

V: Hoe gaan we om met de *Trefwoorden*? Hoeveel kunnen we er invullen, en zit er een hiërarchie in of een bepaalde indeling? Groeit het aantal zinvolle trefwoorden mee naar mate we meer aan het beschrijven zijn?

A: Een trefwoorden register aanleggen (ook op een aparte tab in dezelfde excel met metadata) is een goed idee en deze groeit mee, en zal vast nog wel tot de nodige discussie leiden, en is niet volgens regels op te stellen. Probeer enkelvoudige beschrijvende woorden te gebruiken als trefwoord om het simpel te houden. Dus geen 'agrarische activiteiten' maar gewoon 'landbouw'. En dus inderdaad gedetailleerde woorden en dat mogen er best een aantal zijn per fotobeschrijving. Het is alsof je hardop met een paar woorden aan iemand wilt uitleggen wat er

op die foto staan als belangrijkste, zonder dat die ander de foto voor zich heeft. Kun je ook gewoon op die manier uitproberen trouwens, en ik denk dat je het bepalen van trefwoorden zeker met zijn meerdere personen moet doen, want het is niet zo simpel als het lijkt. Het klopt dat dit een enorme ingewikkelde hiërarchische boom thematische trefwoorden zou kunnen worden. Maar maak daar geen wetenschap van, en misschien moet je je beperkingen opleggen tot een honderdtal trefwoorden of zo, dus ze toch op een vrij algemeen niveau houden. De trefwoorden kunnen later worden gebruikt om mee te selecteren (bv. 'geef alle foto's waar bloemen op staan'). De trefwoorden op de Excel-tab zijn vooral een geheugensteuntje over welke trefwoorden je allemaal hebt gebruikt bij de individuele foto-beschrijvingen en moet vooral wildgroei van trefwoorden voorkomen. Zo moet je het zien, als een overzicht ter raadpleging bij het invoeren van beschrijvingen in de rubriek Trefwoorden.

V: Is het mogelijk om later in een zoekprogramma de termen 'gladiolen' of 'bakkerij' ook onder een trefwoord 'landbouw' of 'detailhandel' te laten vallen, zonder dat je die trefwoorden nu in het schema invult?

A: Met woorden zoeken, die in beginsel zelf niet als trefwoord zijn gebruikt, is wat je kan noemen een combinatie van een thesaurus (termenlijst met betekenis) en een taxonomie (een hiërarchische boomstructuur, een soort van classificatiesysteem van termen in dit geval). Daar kun je wel een hele wetenschap van maken ben ik bang en ik weet niet of je daaraan moet beginnen. Zoiets later toevoegen moet toch wel kunnen denk ik.

V: Hoe scheiden we trefwoorden? Met een komma, of iets anders? En is dat relevant?

A: Als scheidingsteken een puntkomma gebruiken ';'. Dat is duidelijker dan alleen een komma. Eventueel ook later gemakkelijker te gebruiken als we de diverse trefwoorden automatisch uit elkaar zouden willen halen.

Opslag in de cloud

V: Moeten we dropbox gebruiken? Of kan iets anders ook?

A: Dat kan natuurlijk ook elders, en het is eigenlijk zo dat de bekendheid met dropbox aanleiding was voor de aanbeveling. We hebben geen vergelijkend onderzoek gedaan naar alle mogelijke cloudopslagdiensten. Belangrijk is dat je de cloud-opslag met anderen kunt delen, en dat je er niet teveel voor betaald natuurlijk. Hoewel dropbox handig werkt moet je wel degelijk opletten dat als je lokale dropbox folder is gekoppeld met de cloud-dropbox folder (en standaard is dat zo) dat alles wat je lokaal in die folder doet meteen naar de gekoppelde cloud-opslag wordt doorgesluisd. Dat geldt dus ook voor verwijderen en daarmee moet je dus oppassen!

Het vermelden van personen

V: Hoe gaan we om met het vermelden van meisjesnamen in de metadata? [20160916]

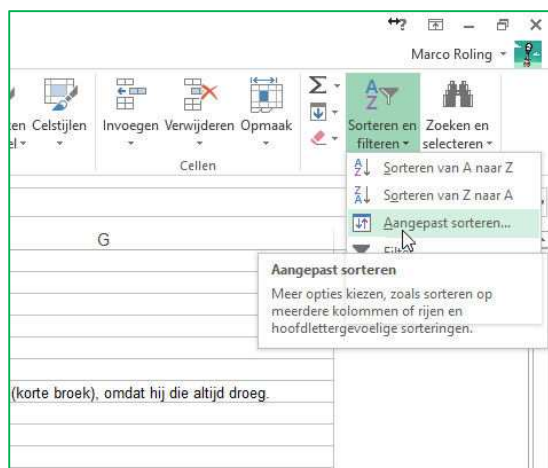
A: Vrouwen kunnen onder hun meisjesnaam, mansnaam of een combinatie voorkomen op de foto's. Het is raadzaam om één naam te kiezen (de keuze is arbitrair) en deze consequent in de metadata van de foto's op te nemen. Een persoonsnamenlijst is op een aparte tab in de Excel metadatasheet opgenomen waarin alle personen kunnen worden geïnventariseerd. In de kolom 'Alternatieve naam' kunnen andere namen waaronder deze persoon bekend is achter elkaar worden genoteerd, elke naam gescheiden door een puntkomma.

De persoonslijst zou gebruikt kunnen worden om een index mee te maken in het collectiepresentatiesysteem en idealiter om dit als zoekfunctie te gebruiken op de collectie.

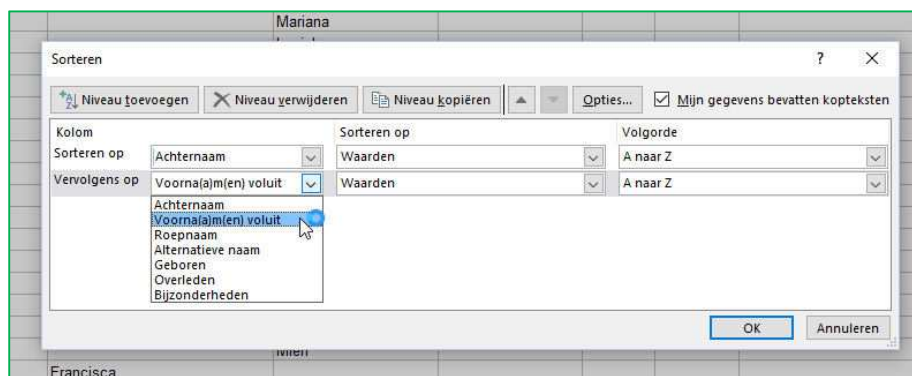
De beschikbare kolommen voor metadata van personen zijn: Achternaam, Voornamen, Roepnaam, Alternatieve naam, Geboren, Overleden, Bijzonderheden.

V: Hoe sorteert je persoonsnamen in de metadata sheet? [20160916]

A: De tab Personen in de Excel metadatasheet kan uitmonden in een lange lijst met namen en sorteren kan ervoor zorgen dat dit nog overzichtelijk blijft. In Excel is daarvoor de functie 'Sorteren en filteren' beschikbaar onder de menutab 'Start'.



Kies voor 'Aangepast sorteren' en het volgende keuzeschermbijbeeld verschijnt:



TULIPANA PROJECT

Hier kunnen eenvoudigweg de sorteerniveaus worden aangegeven, door een kolom te kiezen en alfabetisch te sorteren. Het is dus raadzaam om op twee niveaus te sorteren, namelijk 'Achternaam' en 'Voorna(a)m(en) voluit'. De gehele tab (uitgezonderd de kolomnamen) wordt gesorteerd en dit kan worden herhaald omdat de sortering in de tab onthouden wordt.