



HEEMKUNDIGE KRING  
HET SPEELHOF

# Opbouw van een digitaal archief

**Leo Benneke**

**[hetspeelhof@gmail.com](mailto:hetspeelhof@gmail.com)**

Eerste uitgave.

Beelden en teksten kunnen op twee manieren behandeld worden: tentoongesteld in een museum of opgeslagen in een digitaal archief.

Dit document focust enkel op het opbouwen van een digitaal archief. Dus op het digitaliseren van oude informatie en het bewaren ervan. Het gedigitaliseerd materiaal kan daarna gebruikt worden om afdrukken te maken of om te ontsluiten via registratie in een databank. Het gebruik van het digitaal archief komt in dit document niet aan bod.

## Inhoud

1 Digitaal archief van heemkundige kring Het Speelhof.....	3
1.1 Mappenstructuur.....	3
1.2 Bestandsnamen.....	3
1.3 Digitaliseren.....	4
1.4 Digitaal bronmateriaal.....	4
1.5 Metadata.....	5
1.6 Backup.....	5
1.7 Software tools.....	5
2 Inleiding.....	6
3 Digitalisering: opdelen in blokjes.....	6
3.1 Statische beelden en teksten.....	6
3.2 Digitaliseren van film.....	10
4 Digitale kleuren.....	11
4.1 RGB kleursysteem.....	11
4.2 sRGB.....	11
5 Digitaal formaat.....	12
5.1 Statische beelden en foto's van teksten.....	12
5.2 Video en film.....	14
6 Digitaal bronmateriaal.....	15
6.1 Foto's/afbeeldingen.....	15
6.2 Video/film.....	15
6.3 Tekst.....	15
7 Digitale bestanden bewaren.....	17
7.1 Unieke bestandsnaam.....	17
7.2 Mappen.....	17
7.3 Backup.....	18
7.4 Gebruikte formaten.....	18
8 Speciale categorieën.....	18
8.1 Jaarboeken.....	18
8.2 Tijdschriften.....	19
8.3 Tentoonstellingen.....	19
9 Software.....	19
9.1 XnView.....	19
9.2 Gimp.....	19
9.3 Video bewerking.....	19
9.4 SyncBack.....	19
10 Annex 1: Woordenschat.....	21

## 1 Digitaal archief van heemkundige kring Het Speelhof

De toelichting van de gemaakte keuzes door de kring kan je in volgende hoofdstukken terugvinden. Het archief is geleidelijk aan opgebouwd en niet altijd werden de oude bestanden hernoemd/herschikt naar de 'nieuwe' logica.

### 1.1 Mappenstructuur:

Indeling in mappen is evident. Om mogelijke problemen te vermijden bij later gebruik van het digitaal archief worden er geen spaties in de map namen gebruikt.

Als eerste niveau werd de naam van de bron gekozen (zodat de aanbrenger snel kan gevonden worden ingeval van vragen).

Het tweede niveau geeft de aard van de afbeelding: Gebeurtenissen, Gedenktekens, Kaarten, Locaties, Organisaties, Personen, Voorwerpen

Het derde niveau (indien er veel afbeeldingen zijn op dat niveau):

- Gebeurtenissen: datum gevolgd door korte tekst
- Kaart: jaartal gevolgd door korte tekst
- Locaties: straatnaam (bevat ook gebouwen in die straat)
- Organisaties: naam van de organisatie
- Personen: uitzonderlijk

Een vierde niveau is uitzonderlijk, maar kan natuurlijk

### 1.2 Bestandsnamen

Omdat het bewaren in het digitaal archief en gebruik in een databank op aparte tijdstippen gebeurt, kan het registratienummer niet gebruikt worden als bestandsnaam. De unieke bestandsnaam moet op een andere manier verkregen worden. Ook in bestandsnamen werden geen spaties toegelaten.

- Een lijst met initialen en namen van de bronnen (als in mapnaam) werd opgesteld. Initialen moeten uniek zijn! Alle bestanden aangereikt door die bron beginnen met die initialen.
- Als de bron maar weinig materiaal aanlevert en de bestanden hebben een unieke naam, kan deze naam gebruikt worden als aanvulling (vb. het volgnummer van zijn fotoestel)
- Indien gekend, wordt bij voorkeur het jaartal waarin de foto werd gemaakt gebruikt.
- Voor gebeurtenissen kan een afkorting ervan gebruikt worden (vb. OMD)  
Voor personen hun naam  
Voor locaties de straatnaam
- Meestal is een volgnummer noodzakelijk

Het toekennen van een 'logische naam' is wel niet zonder gevaar. Als later blijkt dat het jaartal verkeerd was of straat(naam) verkeerd, zou je het bestand moeten hernoemen, maar als dat al geregistreerd is in een databank, scheidt dat extra moeilijkheden.

## 1.3 Digitaliseren

Er werd gestreefd om afbeeldingen van ongeveer 8 Mpixel te verkrijgen in het gebruikelijke JPG formaat. Originele beelden met meer pixels worden zo bewaard. Indien er plaatsgebrek zou dreigen, kunnen deze beelden verkleind worden.

De scanner ingesteld om die pixels te verkrijgen. Zo worden postkaarten gescand met 600 dpi. En dia's tegen 3000 dpi.

Plans worden zo ingescand dat alle details nog leesbaar zijn. Dit maakt dat plans soms enorm groot worden (100 Mpixel en meer). Later kan dan voor een tentoonstelling een stukje uitgehaald worden met nog leesbare details.

8mm films: omgezet in AVI/MPEG-4/MP3-formaat 720x576 pixels of gelijkwaardig (\*)

16mm films en groter: omgezet in AVI/H.264/MP3-formaat 1920x1080 of gelijkwaardig (\*)

(\*) hangt af van de mogelijkheden van de omzetter en het beeldformaat van de film (het beeldformaat van de film is het belangrijkste – bijsnijden gebeurt liefst achteraf indien nodig)

## 1.4 Digitaal bronmateriaal

Bronmateriaal wordt soms ook digitaal aangeleverd in vele formaten. Het is belangrijk om het aantal formaten sterk te beperken.

Afbeelding in een ander formaat als JPG: omgezet naar JPG

Affiche in een tekst formaat: kan in PDF of JPG formaat omgezet worden

Tekst van verschillende bladzijden: omgezet naar PDF

Als de tekst later moet bewerkt worden: omgezet naar ODT (en PDF voor ontsluiting)

Video/DVD zonder ondertiteling: omgezet naar AVI/MPEG-4 of H.264/MP3

Video/DVD met ondertiteling: omgezet naar MKV/MPEG-4 of H.264/MP3

Het DVD menu systeem gaat verloren - reproduceren van de originele DVD is dus niet mogelijk  
Bronmateriaal bewerken

Indien een afbeelding wordt gebruikt voor een tentoonstelling, wordt deze afbeelding bijgewerkt: kreuken worden deels weggewerkt, contrast verbeterd om leesbaarheid te verhogen. Donkere foto's verhelderd door gamma correctie. Ontbrekende stukjes deels ingevuld.

Andere afbeeldingen enkel bewerken als ze 'te slecht' zijn.

Filmpjes voorzien van ondertiteling (voorlopig nog geen prioriteit), zwarte balken afsnijden.

## 1.5 Metadata

Een foto van een digitaal fototoestel bevat **EXIF** informatie: informatie die meer uitleg geeft met welke instellingen de foto werd genomen. Soms staat er zelfs GPS info in. Het is belangrijk dat deze data bewaard blijft.

Bijkomend kan de fotograaf (enkel in professionele fototoestellen) IPTC data invullen. Daar kan hij zijn naam en copyright informatie toevoegen. Het IPTC-veld Titel is toegankelijk voor Windows 10 en is dus een goede kandidaat om informatie beschikbaar tijdens het digitaliseren tijdelijk op te slaan, omdat het nog een tijdje kan duren voordat dat de foto wordt geregistreerd.

XnView noemt dit veld **Beschrijving** en Linux noemt dit veld **Omschrijving**.

Je hebt een fotobewerkingprogramma nodig om dit veld aan te passen:

**XnView:** Open foto, ga via **Bewerken – Metadata – Bewerk IPTC data** – tabblad **Beschrijving**. Vul het vak **Beschrijving** in.

**Picasa:** Bewerk foto en vul het tekst vak onder de foto.

## 1.6 Backup

Er werd gekozen voor een externe harde schijf.

## 1.7 Software tools

Voor beeldbewerking: XnView en Gimp

Voor tekstverwerking: Libre Office

Voor Video bewerking: VirtualDub, Subtitle Workshop, Free Video Convertor, VideoPad Video Editor, mkvmerge GUI

Back-up: SyncBack

## 2 Inleiding

In een vroegere vergadering met heemkundigen werd duidelijk dat onder de heemkundigen de kennis en ervaring met de digitale wereld enorm varieert van volslagen leek tot zelfverklaarde expert.

Ook de verzuchtingen en verlangens over die digitale wereld vallen ook in deze categorie, wat maakt dat een betoog over deze materie maar door een beperkt publiek kan gesmaakt worden.

Het digitale verhaal kan ruwweg in twee delen gesplitst worden: het opbouwen van een digitaal archief en het gebruik ervan. Dit document behandelt enkel het opbouwen van een digitaal archief.

Ondanks de grote verschillen in opinies, wil ik mijn benadering documenteren: Het heeft mij moeite gekost om tot deze inzichten te komen. Sterker nog, verdere ervaring zouden mijn inzichten nog kunnen veranderen.

Ik pleit er dus niet voor om dat iedereen deze ervaringen blindelings zou volgen.

## 3 Digitalisering: opdelen in blokjes

### 3.1 Statische beelden en teksten.

#### 3.1.1 Eisen/streefdoel

Dikwijls is het origineel niet in eigen bezit of door de tijd heen fragiel geworden. Daarom zou latere consultatie van het origineel na digitalisering zoveel mogelijk moeten vermeden worden. Dit wil zeggen dat de kleinste details nog leesbaar/herkenbaar moeten zijn in het digitale resultaat.

Digitaliseren wil zeggen een hele foto opdelen in super kleine blokjes (zoals millimeter papier, maar met nog veel kleinere ruitjes). Als vuistregel wordt aangenomen dat een menselijk oog 200 blokjes per duim (2,54 cm) niet meer kan zien zonder hulpmiddelen. Anders gezegd, zorg ervoor dat de digitale beelden 200 dpi (Dots Per Inch) of meer hebben.

Als we tentoonstellingen organiseren, worden de digitale beelden, eventueel met markeringen, opnieuw afgedrukt. Met de stand van de huidige techniek is voor papier het A3-formaat haalbaar. Grotere formaten zijn een veelvoud duurder, terwijl A3 voldoende groot is. Ook voor plans is het formaat A3 dikwijls bruikbaar omdat in een tentoonstelling niet altijd alle details zichtbaar moeten zijn of dat enkel een deel van het plan wordt gebruikt.

Met het formaat A3 in gedachte en 200 dpi, zijn afbeeldingen van 8 Mega pixel een goed streefdoel (dit kan je narekenen: A3 = 16 inch op 12 inch, en  $16 \times 200 \times 12 \times 200 = 8.000.000$  pixels).

## 3.1.2 Digitalisering via een Copy center:

Voor afbeeldingen groter dan A3 geeft een Copy center het meest kwalitatieve resultaat door het gebruik van professionele toestellen. Maak duidelijk dat de kleinste details nog leesbaar moeten zijn en dat het origineel niet mag beschadigd worden. Als het origineel wat te 'fragiel' is, kan het beter gescand of gefotografeerd worden.

Voor grote afbeeldingen (meestal plans) bevatten meestal geen kleine details. Dan is 150 dpi voldoende, maar voor kleinere (vb. postkaarten) wordt best 600 dpi gebruikt (streefdoel 8 Mega Pixel). Een hogere dpi heeft geen zin, want dan komt je de druktechniek tevoorschijn. Zelfs bij 600 dpi moet soms het beeld vervaagd worden.

## 3.1.3 Digitaliseren met een scanner

Dit is meestal de beste methode: Ze is goedkoop en de kwaliteit is goed tot zeer goed.

### **Film, glasplaten, negatieven en dia's**

Je hebt hier een scanner nodig 'met bovenlicht', want de scanner moet het licht opvangen dat door het materiaal gaat. Voor negatieven tot 6 cm zijn betaalbare scanners te vinden. Voor grotere formaten heb je een dure scanner nodig.

Een professioneel gemaakt negatief kan tot 3000 dpi bevatten. Amateur materiaal is veel minder scherp en wordt geschat op 1500 dpi. Natuurlijk gaat de tand des tijds daar nog wat van af doen, maar in welke mate is niet duidelijk.

Als je zeker wil zijn dat je zowat alle details hebt, zou je tegen 3000 dpi kunnen scannen. Digitale beelden kunnen achteraf nog verkleind worden, maar vergroten is sterk af te raden want dan zouden de puntjes zichtbaar kunnen worden of wordt het beeld wazig.

Met 3000 dpi krijg je van een dia een digitaal beeld van ongeveer 12 Mega pixels, met 2400 dpi een 7 Megapixels. Let wel dat het al goede dia's moeten zijn om zoveel detail te geven.

### **Papier, foto's**

Voor foto's gebruik je best 600 dpi (dus 3 keer meer dan met het oog te zien is), voor plans kan je lager gaan, zolang je achteraf de kleinste lettertjes en details nog kan lezen. Achteraf kan je het beeld eventueel verder verkleinen, maar vergroten is totaal uit den boze! Als je een postkaart met 600 dpi scant, krijg je een digitale foto van ongeveer 8 Mega pixels (het streefdoel).

Sommige foto's hebben een geribbeld oppervlak: Je merkt heel fijne lijntjes op. Als je met een nagel over het oppervlag krast. Voel je de ribbeltjes. Soms pikt de scanner die lijntjes op en worden ze zichtbaar in het resultaat. Om dat te vermijden kan je met een hogere resolutie scannen (vb. 1200 dpi) en achteraf de afbeelding verkleinen (kost wel tijd omdat het scannen veel trager wordt).

Grote plans kan je ook inscannen in verschillende delen en met een fotobewerkingsprogramma terug samensmelten. Let wel dat een vervorming aan de randen zichtbaar wordt omdat het papier daar wordt omhooggetrokken wordt door de rand van de scanner. Voor het samensmelten moet die vervorming eraf geknipt worden.

# Opbouw van een Digitaal Archief

---

## Herschalen

Als je de scan nog wil bewerken omdat het origineel in slechte staat is, kan je inscannen met een hogere resolutie (hogere dpi) en na de beeldbewerking verkleinen zolang de kleinste details zichtbaar blijven.

Vergeet dan niet om de dpi ook aan te passen. Vb. als het origineel gescand is met 600 dpi, en na bewerking wordt het beeld 50% verkleind, dan moet de dpi op 300 gezet worden.

Let wel dat de waarde van de dpi zelden gebruikt wordt.

## Kleurherstel

Bij het scannen kan je kiezen voor kleurherstel. Zo krijgen oude dia's met verlopen kleuren bijna hun originele kleuren terug. Let wel dat kleurherstel eist dat er een wit voorwerp in de dia moet voorkomen, zodat de scanner zich daarop kan richten om de juiste kleurbalans te vinden.

Anders gezegd: als er geen wit in het origineel voorkomt, werk je beter zonder kleurherstel, en gebruik je een fotobewerkingsprogramma om de kleur te corrigeren.

## Verscherpen

Digitalisering (omzetten naar vierkante blokjes) heeft een licht vervagend effect. Dit kan gecompenseerd worden door Verscherpen. Lettertjes worden iets beter leesbaar. Let wel: het effect mag amper waarneembaar zijn (er mogen dus geen witte verkleuring naast de zwarte lijnen komen)

## Ontkrassen

Bij duurdere scanners gebruikt bij dia's of negatieven om minder last te hebben van krassen en stof.

## Gamma correctie

Donkere afbeeldingen kunnen wat helderder gemaakt worden via de gamma correctie: De donkerste en helderste delen blijven ongemoeid, de gamma correctie beïnvloedt enkel de tussenliggende waarden. Met een fotobewerkingsprogramma kan deze ook bijgesteld worden en veel fijner. Dus enkel te gebruiken als het origineel sterk onderbelicht is.

### 3.1.4 Fotograferen

Als scannen niet mogelijk of moeilijk is, is fotograferen een mogelijkheid: Fototoestellen hebben wel last van lensvervorming. Niet alleen met het vervormen van hoeken, maar ook, en erger, met veel minder herkenbare lensfouten.

Als er te weinig licht is, krijg je ruis: een egaal oppervlak krijgt spikkeltjes. Gebruik een flash, maar pas op voor weerkaatsing. Dan kun je niet perfect voor het beeld staan (met flash recht voor een beeld achter een glasplaat is geen goed idee). Om dit te corrigeren (parallax vervorming) en om 'vuile' randen af te snijden heb je meestal een fotobewerkingsprogramma nodig.

Rekening houdend met de 200 dpi-regel, heb je een 8 Mega pixel camera nodig voor een A3-foto. Het is aanlokkelijk om dan met een 16 Mega pixel camera een enkele foto van een A2-foto te maken. De nummertjes zijn dan wel juist, maar als je dan later op je computer gaat kijken, merk je, als je volledig inzoomt, rare vervormingen en zijn details soms onleesbaar. De reden hiervoor is dat de lenzen van de meeste (goedkope) fototoestellen tegen hun optische limiet zitten, maar dat voor de fabrikant onbelangrijk is omdat meer megapixels beter verkoopt.

Conclusie: Als je alle details van een groot origineel wil fotograferen, splits het origineel dan (denkbeeldig) op in delen van A3-formaat en smelt die beelden op PC terug samen.



## Opbouw van een Digitaal Archief

---

Foto's van een A3-deel zouden 8 Megapixel moeten worden. Het is nog niet duidelijk of je best het fototoestel op 8 Megapixel zet of dat je de foto's op maximum detail moet zetten en de computer de omvorming moet laten doen.

Controleer zoveel mogelijk de kwaliteit van elke foto: Liefst wordt een statief gebruikt, maar dat is niet altijd mogelijk. Je moet namelijk bij controle de foto zo hard mogelijk inzoomen en zo controleren dat de kleinste details nog zichtbaar zijn. Er zijn twee grote problemen:

- Als je bewogen hebt, zijn de details wazig (soms onleesbaar)
- De kleurechtheid/helderheid is heel wisselvallig. Als je twee (of meer) foto's wil samensmelten merk je altijd kleur/helderheid verschillen

### 3.1.5 Beeldbewerking

Hier wordt de eerste keuze gemaakt: hou je het gedigitaliseerde beeld met al zijn gebreken en dus zo intact mogelijk, of probeer je het beter leesbaar maken?

Wil je plooien, vlekken, scheurtjes wegwerken en zelfs de kleuren wat herstellen of vind je de toestand waarin je het origineel hebt gedigitaliseerd belangrijk?

Natuurlijk heb je voor beeldbewerking wel een computer specialist nodig.

Zo kan randen afsnijden, het beeld wat rechter zetten, enz. al genoeg zijn. Heb je echter stukjes gedigitaliseerd van een groter geheel, dan moeten deze nog bijeengebracht worden met een beeldbewerkingsprogramma zoals Photoshop (Gimp is de gratis tegenhanger voor deze job).

Zeker voor stukjes van een plan merk je dat (goedkope) fototoestellen lensvervorming hebben die je normaal niet opmerkt. Zo merk je, zelfs na het wegwerken van de parallax fouten en vervorming aan de hoeken (vb. bolvorming) dat lijnen van een plan bovenaan en onderaan kloppen, maar ergens daartussen de lijnen van het plan niet meer uitkomen.

Indien mogelijk gebruik je door die lensvervorming best een scanner.

Volgende beeldaanpassingen zijn dikwijls nodig:

- lenscorrectie: het afbuigen van hoeken
- parallax correctie: als je niet perfect in het midden staat van een beeld
- helderheid/contrast correctie: (donkerste delen zwart en de lichtste wit)
- kleurcorrectie: als door de tand des tijds een kleur verzwakt is
- verdoezelen: verschillende puntjes samen nemen om de druktechniek (raster) te verbergen
- clone gereedschap om delen te kopiëren over andere delen om vlekken/plooien weg te werken

Om foto's samen te smelten kan je in Gimp:

- een masker gebruiken om de ene foto geleidelijk transparanter te maken zodat de andere foto geleidelijk aan zichtbaarder wordt om abrupte kleurovergangen te vermijden
- delen van een foto wat verplaatsen zodat ze beter overgaan naar de andere foto. Vooral om lensfouten op te lossen

## 3.2 Digitaliseren van film

Films komen ruwweg in twee formaten:

### Een filmrol (8mm, 16mm, 35mm)

Een filmrol digitaliseren vergt duur/oud/speciaal materiaal. Meestal moet je hiervoor een firma contacteren. Via internet zijn er verschillende aanbieders te vinden.

De beeldverhouding is meestal 4 op 3 (het oude TV-formaat). Voor 8mm film kan je kiezen voor het DVD-formaat, maar liefst voor losse computer filmpjes met DVD beeld grootte (720x576). Het DVD formaat heeft het voordeel dat je de DVD op vele toestellen kan afspelen, maar het wordt dan weer moeilijk als je 1 filmpje er wil uithalen. Als de archivaris voldoende technische kennis heeft, zijn aparte filmpjes aan te raden.

Voor 16 mm film zou je een Blu-Ray disk formaat kunnen vragen of liever weer losse computer filmpjes van ongeveer HD-grootte (1920x1080). De kwaliteit kan hier merkbaar hoger zijn, dus vermijd hiervoor het DVD-formaat: Een Blu-Ray disk kan omgezet worden naar DVD, maar omgekeerd geeft een pover resultaat. Er moet afgewogen worden of de kwaliteit van de film hoog genoeg is om de extra kosten ervan te rechtvaardigen.

Voor 35mm film zou de 4K technologie kunnen gebruikt worden voor kwaliteits-film.

Volgens de 200 dpi regel, zou een beeld van een DVD film op papier afgedrukt een kwalitatief beeld kunnen geven van ongeveer 9 cm op 7 cm. Voor Blu-Ray wordt dat 20cm x 15cm en voor 4K technologie 40cm x 30cm (A3-formaat).

### Een video cassette

De kwaliteit van deze opnamen (VHS, Super VHS, Video-8) is veel lager dan van het DVD-formaat, maar de meeste apparatuur werd voor het DVD formaat voorzien. Sommigen hebben die oude afspeelapparatuur nog thuis staan, maar zelden hebben deze ook een DVD-recorder aangesloten, zodat je meestal ook op zoek moet gaan naar een oplossing via het internet. Ook hier zijn aparte computer filmpjes verkiesbaar, maar vragen meer technische kennis van de archivaris.

Voor losse filmpjes, zie hoofdstuk 5.2 voor het digitaal formaat.

## 4 Digitale kleuren

Na het opdelen van het origineel, moet voor elk vierkantje nog een kleurinformatie gegeven worden.

### 4.1 RGB kleursysteem

De meest gebruikte techniek om kleuren weer te geven komt van de analoge TV's: Het kleurbeeld wordt opgesplitst in Rood – Groen en Blauw (RGB) kanalen. Elke kleur werd heel nauwkeurig vastgelegd (vb. wat exact Rood is). Dit systeem heeft beperkingen: Er zijn blauw en groen tinten die met dit systeem niet kunnen weergegeven worden (neem bvb. maar eens een foto van wolken).

De RGB techniek wordt zowat overal gebruikt: in fotoestellen, in TV's, PC monitors en software programma's en kan omgezet worden voor printers.

Zoals altijd borrelt het wel en zijn er nieuwe systemen die een groter kleurbereik (gamut) hebben. Maar deze technieken zijn nog volop in ontwikkeling en het is niet duidelijk hetwelk er zal overblijven. Daarboven op komt nog dat deze techniek (nu nog) zeer duur is (niet alleen heb je een nieuw fototoestel nodig, maar ook nieuwe software, een aangepaste PC, een nieuw PC scherm en nieuwe printers).

### 4.2 sRGB

Het omzetten van de intensiteit van een kleur naar een digitale waarde kent nu 2 soorten: 8 bit per kleur of 16 bit per kleur. Omdat je 3 kleuren hebt (RGB) heb je dus 24 bit of 48 bit kleuren (8 bit grijs of 16 bit grijs). De 16 bit versie wordt nog weinig gebruikt en is minder gestandaardiseerd.

Het fototoestel, de PC monitor en de programma's gebruiken normaal de 24 bit kleuren. Het gebruik van de hogere bit kleuren vraagt dus wel enige kennis om nuttig te zijn (of duur apparatuur).

24 bit kleuren wil wiskundig zeggen dat er 16 miljoen kleuren kunnen weergegeven worden. Ons oog ziet er maar 4 miljoen, alhoewel ons oog meerdere kleuren kan zien die met het 24 bit systeem niet weer te geven zijn.

RGB met 24 bit kleuren wordt soms ook **sRGB** genoemd: standaard RGB.

Merk op dat de meeste toestellen zich klaarmaken om ook met andere kleurprofielen te kunnen werken. Dit is het terrein van het ICC: Internationaal Color Consortium. Geleidelijk aan zal in digitale bestanden het gebruikte ICC-profiel vermeld worden (nu is dat meestal sRGB), maar dat is toekomstmuziek.

## 5 Digitaal formaat

De blokjes en hun kleuren moeten in een digitaal bestand geordend worden. Er zijn honderden manieren om dit te doen, maar uw PC moet niet alleen nu dat formaat kennen, maar ook de PC van de toekomst. Met andere woorden: het gekozen digitaal formaat is heel belangrijk en we kiezen dus best een heel bekend formaat.

### 5.1 Statische beelden en foto's van teksten

De digitale formaten kunnen ingedeeld worden in twee grote groepen: met of zonder verlies van details. Op het eerste zicht zou je kiezen voor een digitaal formaat zonder verlies van detail, maar het digitaal bestand kan tot 10 keer kleiner worden met een formaat met verlies.

Het resultaat is dat het meest gebruikte formaat in fototoestellen een digitaal formaat met verlies is, en zowat alle toestellen (PC, TV,...) dat formaat kunnen weergeven. De digitale wereld staat nooit stil, het aantal digitale formaten blijft stijgen en worden bestaande formaten regelmatig aangepast.

Het meest bekende formaat is het **JPG**-formaat (Joint Photographic Experts Group). Het is een compressie systeem met verlies dat de digitale bestanden tot 10 keer kleiner maakt dan een systeem zonder verlies, maar met een verwaarloosbaar verlies aan detail. De kans echter dat dit formaat ook nog op toekomstige toestellen kan weergegeven worden lijkt hoger dan andere formaten (lijkt, want wie kent de toekomst). Let wel dat binnen dit formaat uitbreidingen mogelijk zijn die problemen kunnen veroorzaken. Indien mogelijk, probeer deze te vermijden.

JPG heeft een kwaliteitsinstelling. Voor actuele foto's ligt het kwaliteitsniveau op 90 tot 95. Voor oudere foto's wordt meestal 85 gebruikt.

De opslag capaciteit van computer systemen stijgt sterk, maar ook het aantal beelden dat wordt verzameld. Dus het is niet alleen de populariteit van het JPG formaat maar ook de verkleining van de digitale bestanden dat JPG een geschikt formaat maakt.

Voor de puriteinen: JPG is een beeld compressiesysteem met verlies van details. Ook verschijnen minuscule 'artifacts' die zichtbaar worden bij het vergroten. Aangezien de kwaliteit van de oude beelden in ons bezit niet erg hoog ligt (de tand des tijds ...), veronderstellen we dat deze nadelen niet opwegen tegen het voordeel van kleinere bestanden én een heel wijd verspreid systeem.

#### **Tussen formaat PNG**

Als een afbeelding in verschillende stappen moet verwerkt worden, lijkt het JPG formaat iets minder geschikt: Bij sterke contrasten verschijnen er 'artefacten': minuscule rare vervormingen. Bij verdere bewerking zouden deze uitzonderlijk iets meer zichtbaar kunnen worden.

Zwarte lijnen op een wit blad geven zo'n sterke contrasten. Met andere woorden, als plans nog verder moeten afgewerkt worden lijkt JPG niet het beste formaat, een verliesloos formaat is een betere keuze: De bestandsgrootte is dan onbelangrijk. Voor het finale resultaat kan weer JPG gebruikt worden.

Als tussen formaat is PNG een goede keuze: ze is verliesloos en heeft een ingebouwd compressie systeem én is vrij oud (dus ook vrij bekend). Er zijn andere verliesloze formaten (vb. TIFF). Deze kunnen even goed gebruikt worden – het is een kwestie van voorkeur – het is toch maar een tijdelijk

# Opbouw van een Digitaal Archief

formaat.

Voor gewone foto's die verschillende bewerkingen moeten ondergaan kan dit systeem ook gebruikt worden, maar het verschil is hiervoor verwaarloosbaar en de kans is groot dat Metadata (zie volgend hoofdstuk) verloren geraakt.

## 5.1.1 Metadata

Metadata zijn gegevens die bij in het digitaal bestand zitten die meer informatie geven over de inhoud. Voor JPG zijn er ruwweg 2 soorten informatie: **EXIF**: informatie van de camera (belichtingstijd, datum/uur dat de foto werd gemaakt, ...) en **IPTC**: informatie over de foto met o.a. auteursrecht.

Het is belangrijk dat de EXIF informatie niet verloren geraakt. Ook de IPTC data is nuttig. Bij gebruik van een fotobewerkingsprogramma, kan **in de IPTC data uitleg over de foto** bewaard worden totdat de foto wordt geregistreerd.

Het IPTC-veld **Titel** is een goede keus, maar programma's gebruiken verschillende namen ervoor : Picasa en Windows 10 noemen het **Titel**, XnView noemt het **Beschrijving** en in Linux heet het **Omschrijving**.

Bij registratie wordt deze info overzet in de databank. Veranderingen/uitbreidingen van deze data gebeuren enkel in de databank. Het IPTC-veld wordt dan niet meer gebruikt.

### Overzicht van gekende IPTC-data en de programma's die deze kunnen lezen:

1914-18 bevoorrading Vr.jpg	XnView	Picasa	Windows 10	Linux
IPTC beschrijving	<b>Beschrijving</b>	<b>Titel</b>	Titel	Omschrijving
Ikke ni	Auteur beschrijving	Schrijver titel		
Foto vremde	Titel	Koptekst		
Watte?	Speciale instructie	Instructies		
oorlog	<b>Trefwoorden</b>	<b>Tags</b>	Labels	Trefwoorden
Coi's foto's	Toegevoegde categorieen	Extra categorie		
Alleen van mij	Copyright	Copyright	Copyright	Auteursrecht
Verantw. auteur	Auteur/maker	Kunstenaar	Auteurs	Maker
Grote baas	Functie auteur/maker	Naamregel titel		
verantwoor	Verantwoording	Credit		
De coi natuurlijk	Bron	Bron		

## Opbouw van een Digitaal Archief

In onderzoek	Bewerk status	Status		
1234	Job ID	Id sample gegevens		
xnv	Programma	Oorspronkelijk programma	Programmanaam	
1.4	Programma versie	Programma versie		
6	Prioriteit	Urgentie		
b	Object ronde	Objectcyclus		
Diverse naam	Document naam	Titel		
Boechout	Stad	Stad		
Kapelle	Lokatie	Sublocatie		Locatie
Antwe	Provincie	Staat		
Belgi	Land	Land		
032	Landscode	Landcode		
Diverse baas	Functie-ID			
20160804	Aanmaakdatum	Datum waarop de foto is genomen		
2:25	Aanmaaktijd			
20160805	Vrijgave datum	Release datum		
4:45	Uitgifte tijd	Release datum		
thuis	Contact	Contactpersoon		
straatnaam	Content location name	Inhoud, locatie, naam		
home	(inhoud locatie code)	Content locatie van Google code		

XnView kan alle velden aanpassen, Picasa kan enkel de Tags (Trefwoorden) en Titel (Beschrijving) aanpassen, Windows 10 en Linux kan enkel enkele velden laten zien.

Zoals in de tabel hierboven te zien is, zijn er maar 4 parameters die algemeen toegankelijk zijn: De **titel**, **trefwoorden** (of labels), **Auteursrecht** en **Maker**.

## 5.2 Video en film

Voor bewegende beelden heb je een 'container' en een compressie systeem nodig. Als container is AVI heel bekend en als compressie methode MPEG 4. MPEG 4 kent wel verschillende versies, de meest moderne is H.264.

Ondertiteling is mogelijk als er een bestand met dezelfde bestandsnaam is maar van het type srt.

---

# Opbouw van een Digitaal Archief

---

Het DVD formaat gebruikt andere containers (MOV) en compressiemethode MPEG 2, maar coders/decoders zijn auteursrechtelijk beschermd, wat een voordeel kan zijn, maar problemen geeft als je losse filmpjes gebruikt. Het laat wel verschillende geluidskanalen toe en ondertiteling.

Het MKV formaat laat ook verschillende geluidskanalen en verschillende ondertitelingen toe, maar is nog vrij onbekend.

Conclusie: Je zet best alle filmpjes om in hetzelfde formaat, en als er nieuwe computers op de markt komen, controleren of die filmpjes nog wel kunnen afgespeeld worden. Open source of gratis coder/decoders hebben de voorkeur, want die kan je dikwijls downloaden van het internet. Commerciële (betalende) kunnen verdwijnen als de fabrikant dat beslist.

## 6 Digitaal bronmateriaal

### 6.1 Foto's/afbeeldingen

Foto's zijn meestal in het JPG formaat. Indien een ander formaat gebruikt werd, overweeg dan om het formaat om te zetten: Indien het wenselijk is om het origineel formaat te behouden, controleer dan hoe bekend dat formaat is en dat er ook geen dialecten van bestaan!

Maak dan een lijst van de gebruikte formaten en controleer dan in de toekomst of dat formaat ook leesbaar is op nieuwere computers.

Als het aantal foto's heel groot wordt en de opslagruimte beperkt, dan kunnen de foto's dikwijls verkleind worden. De meeste fototoestellen van meer dan 10 Mega pixels geven een onzuiver beeld op 100% uitvergroting. Verklein dan deze foto's tot 8 Mega pixels, voldoende voor een A3-afdruk.

Bij gebruik van fotobewerking, zorg ervoor dat de Metadata niet verloren gaan!

### 6.2 Video/film

Filmpjes worden best omgezet naar eenzelfde formaat. Noteer zowel de video encoding als de geluid encoding.

Voor een analoge opname moet het juiste aantal beelden/seconde gebruikt worden door de mechanische 'shutter': Als een beeld wordt genomen als de shutter juist voor het beeld komt, dan is dat beeld donkerder. Het resultaat is een flikkerend beeld. Anders gezegd, voor analoge opname is beeld per beeld opname belangrijk.

Digitale weergave heeft geen shutter, dus ook geen probleem met een frame rate.

### 6.3 Tekst

Een tekst bestand is een verzameling letters van een bepaald lettertype met afbeeldingen opgemaakt in een digitaal formaat. Ook hier vormen we best het bestand om naar PDF of ODT. Er is dan wel een reële kans dat het omgezette bestand er iets anders uit ziet. Indien gekozen wordt om het bestand in het oorspronkelijke formaat te bewaren, moet weer een lijst bijgehouden worden van de gebruikte digitale tekst formaten en als er nieuwe computers verschijnen, gecontroleerd of dat formaat nog ondersteund wordt.

---

## Opbouw van een Digitaal Archief

---

Als er maar weinig tekst is, op 1 pagina waarvan de opmaak heel belangrijk is (meestal affiches), dan kan de tekst omgezet worden in een foto. Zo worden alle details accuraat vastgelegd, maar latere aanpassingen worden dan wel moeilijker. Mik hier weer naar een afbeelding van 8 Mega pixels.

Als de opmaak belangrijk is, zit het gebruikte lettertype best in het digitaal bestand. Oude tekst formaten, zoals **doc** van Microsoft (Word 97 tot 2003) kunnen dit niet. Als een lettertype wordt gebruikt dat niet aanwezig is in de PC, wordt de tekst niet correct weergegeven. Indien het lettertype gekend is, kan het eerst bij geïnstalleerd worden voordat het wordt omgezet.

Vóór de omzetting worden interessante afbeeldingen best apart bewaard. Mogelijk zitten deze met een hogere resolutie in het origineel.

Als de omzetting niet goed lukt (lay-out te sterk verstoord), dan kan de tekst omgezet worden naar foto's en een PDF gemaakt van die foto's.

Als de opmaak minder belangrijk is kan je eventueel het oudere **doc** document ook nog bewaren in het archief. **Pdf** is een goede keuze, maar als je later de tekst nog (verder) wil aanpassen, dan kies je voor **docx** of **odt**.

### 6.3.1 Docx

Voor het **docx** formaat van Microsoft heb je een jaarlijks te hernieuwen Office licentie nodig voor iedereen die het bestand wil bewerken (voor enkel lezen kan je een gratis programma downloaden). De gratis OpenOffice.org en LibreOffice.org kunnen dit type bestand ook lezen en schrijven, maar er zouden kleine verschillen kunnen optreden.

### 6.3.2 Odt

**Odt** is een open tekstformaat (OpenOffice.org en LibreOffice.org) en lijkt daardoor een betere keuze, alhoewel deze standaard (nog) minder bekend is. Controleer wel dat het lettertype opgenomen wordt in het bestand (via Bestand-Eigenschappen-Lettertype-Inbedding van lettertype).

### 6.3.3 Pdf

**Pdf** lijkt een goede keuze omdat deze standaard meest verspreid is in de commerciële wereld. Als je deze teksten wil echter aanpassen, moet je ook een licentie aanschaffen. De tekst blijft wel bereikbaar want je kan stukken tekst kopiëren.

Kies formaat Archive PDF/A-1a (ISO 19005-1) voor het archief.

Extra: Bij omzetting naar pdf, selecteer op het tab-blad Gebruikersinterface: **Toon documenttitel, Verberg menubalk, Verberg werkbalk**, maar niet Besturingselementen van venster verbergen.

### 6.3.4 Conclusie

OpenOffice.org of LibreOffice.org kunnen zowel het doc formaat als **docx** als **odt** lezen én schrijven maar ook een **pdf** maken. Omdat deze programma's ook nog eens gratis zijn, zijn deze programma's een goede keuze voor een Heemkundige Kring (geen probleem met licenties).

Via een simpele tekstverwerker (zoals kladblok) kunnen nog stukjes tekst uit een **doc** bestand gehaald worden. De tekst in de nieuwere formaten is niet meer leesbaar via zo'n simpele tekstverwerker (voor de veiligheid) en dus zijn de bestanden verloren als het formaat niet meer herkend wordt. Teksten waarvan de opmaak niet belangrijk is mogen daardoor in het doc-formaat



blijven staan.

Voor het archief moeten teksten waarvan de opmaak belangrijk is, worden best omgevormd naar formaat odt, docx of pdf.

Pdf krijgt lichtjes de voorkeur omdat je dit bestand nog 'gemakkelijk' kan doorsturen naar geïnteresseerden. Als het bestand achteraf nog moet aangepast worden, kies je beter voor odt of docx.

Het is een subjectieve keuze of je voor odt of docx kiest. Let wel dat je onder 'Eigenschappen' aanduidt dat het lettertype moet vervat zijn in het bestand. Dit maakt het bestand een 3 Mbyte groter! Er kan een licentie op het lettertype rusten, dus zou in theorie problemen tevoorschijn kunnen komen bij gebruik.

## 7 Digitale bestanden bewaren

### 7.1 Unieke bestandsnaam

Het kan de bedoeling zijn deze digitale bestanden te ontsluiten op een ander systeem. Sommige systemen eisen dat alle bestandsnamen uniek zijn. Omdat je de toekomst niet kent, is het raadzaam te zorgen dat alle bestandsnamen uniek zijn.

Als je de bestanden registreert in een databank, zou je de registratienummer kunnen gebruiken, maar dikwijls gebeurt de digitalisering en de registrering op aparte momenten. Het zou dus eenvoudiger zijn om een andere methode te gebruiken.

De bestandsnaam zou kunnen opgebouwd worden met volgende stukjes:

- de initialen van de persoon die het origineel aanbracht (dus niet noodzakelijk de eigenaar). Als er ooit vragen rond zijn, kan je die persoon voor meer informatie vragen.
- het jaartal wanneer het origineel werd gemaakt (indien gekend) (dus niet het moment van digitalisering)
- het onderwerp van het origineel
- eventueel een volgnummer

Alternatief kan je gewoon een volgnummer nemen per persoon zoals een fototoestel dat doet, want als je je vergist van jaartal of onderwerp heb je een misleidende bestandsnaam. Merk op dat een beeld zonder uitleg meestal niets waard is en dat de uitleg wordt toegevoegd bij registratie in een databank. Als je dus later de bestandsnaam verandert, moet je dat dan ook doen in de databank!

### 7.2 Mappen

Alle bestanden in 1 map geeft een onoverzichtelijk grote map. De bestanden zouden dus best zoveel mogelijk ingedeeld worden in mappen. Ook om unieke bestandsnamen beter in het oog te kunnen houden, kan je een map maken per aanbrenger van een origineel.

In die map per aanbrenger, kan je weer een map maken per soort: gebeurtenissen, gedenktekens, locaties, organisaties, personen, plans, postkaarten, ...

## Opbouw van een Digitaal Archief

---

In de map gebeurtenissen, maak je dan weer een map die begint met de datum van het evenement en een korte naam. De datum beginnend met het jaartal om een tijds gesorteerde mappenlijst te krijgen. Voor de locaties, dan weer de betrokken plaats, ...

### 7.3 Backup

Dit is super belangrijk! Papieren archief verdwijnt zomaar niet, maar een digitaal archief kan zonder waarschuwing ineens verdwijnen vb. als de computer waarop het digitaal archief staat faalt. Er is een reserve kopie vereist. Anders gezegd, er moeten altijd twee digitale archieven zijn, de vorm waarin is onbelangrijk. Zo kunnen 2 computers gebruikt worden, of een externe harde schijf, of ...

Je zou een backup 'in de cloud' kunnen overwegen, maar als het archief wat groter wordt, dan zijn er kosten aan verbonden. Soms kan dit gewenst zijn om meerdere mensen rechtstreekse toegang te geven tot het digitale archief, maar dat vereist een gedegen technische kennis van de gebruikers en kan een fout het hele archief uitvegen.

Daarom is gekozen om het digitaal archief door 1 persoon te laten beheren en een externe harde schijf als backup. Indien nodig kan snel en eenvoudig de verantwoordelijkheid voor het digitale archief doorgegeven worden.

### 7.4 Gebruikte formaten

Bewaar een lijst van gebruikte digitale formaten, met een verwijzing naar enkele bestanden in dat formaat.

Later kan dan een gerichte test op een nieuwe computer uitgevoerd worden om te controleren of de nieuwe computer wel al die formaten ondersteund.

## 8 Speciale categorieën

Tot hier toe werd beschreven hoe om te gaan met losse objecten, met als doel ontsluiten via een databank. Niet alle objecten moeten ontsloten worden en sommigen vormen een geheel.

Diegene die ontsloten worden via het internet, moeten voorzien zijn van een HK-merkteken. Ofwel prominent (links) ofwel rechtsonder discreet (rechts)



### 8.1 Jaarboeken

Deze die we digitaal konden verzamelen moeten nog verwerkt worden.

## 8.2 Tijdschriften

Deze moeten bewaard worden in pdf formaat (zoals andere teksten).

Gedrukte exemplaren kunnen ingescand worden, maar om de tekst over te nemen vraagt dit toch een hoop werk. Deze drempel moet behouden worden als we tijdschriften ontsluiten via het internet. Daarom moet voor het internet de tekst eerst omgezet worden in een foto.

## 8.3 Tentoonstellingen

De bewerkte foto's (dikwijls bijgesneden tot papier formaat) en bindtekst worden apart bijgehouden, maar nooit ontsloten via het internet. De tekst en de beelden worden samengesmolten in een pdf voorzien van een HK-merkteken. Bindtekst wordt eerst omgezet in een foto om de kopie drempel te verhogen.

## 9 Software

Liefst zouden alle betrokkenen dezelfde software gebruiken, maar dat is geen vereiste.

### 9.1 XnView

XnView voor Windows is gratis en handig programma om:

- foto's bij te snijden
- foto's wat te draaien
- kleuren wat bij te stellen
- contrast te regelen
- een hele reeks foto's te hernoemen

### 9.2 Gimp

Gimp is een uitgebreid fotobewerkingsprogramma om:

- vlekken weg te werken
- parallax fouten te corrigeren
- lensfouten te corrigeren

### 9.3 Video bewerking

Video bewerking is een pak complexer, ook de taken zijn dat: Wat wil je doen: De video herschalen, omzetten naar een andere compressie methode, ondertiteling erbij doen, een geluidsspoor erbij zetten,... Omdat voor elk van deze taken gratis programma's ter beschikking zijn en dus ook de nodige kennis verondersteld.

### 9.4 SyncBack

Als je twee archieven gelijk moet houden, dan is een hulpprogramma meer dan wenselijk. Meestal

## Opbouw van een Digitaal Archief

---

werk je op 1 archief en zet je regelmatig de veranderingen op het tweede archief.

SyncBack vergelijkt de tweearchieven en kan enkel de veranderde/nieuwe bestanden overzetten naar het tweede archief. Aangezien 99% onveranderd is, gaat dit dus honderden keren sneller.

## 10 Annex 1: Woordenschat

### **Bit/Byte** bit en byte digitale eenheid

Een bit is de kleinste digitale eenheid en kan maar 2 toestanden waar of niet waar (0 of 1). Zoals met onze cijfers worden er rijen bits gemaakt om grotere getallen te maken. Zo is een rij van 8 bits bekend onder de naam Byte.

Byte is de meest gebruikte digitale maat: Het wordt gebruikt om de grootte van geheugen, harde schijven, ... aan te duiden.

### **Dpi** dots per inch of puntjes per duim (een duim is een Engelse maat voor 2,54 cm)

Digitaliseren wil zeggen alles in vakjes indelen – vergelijkbaar met millimeter papier, maar die ruitjes moeten zo klein zijn dat je ze niet meer ziet. Zo werd bepaald dat een menselijk oog, zonder hulpmiddelen 200 blokjes op 2,54 cm niet meer kan zien (dus nog wel met een vergrootglas).

Je moet er dus op toezien dat de densiteit van de puntjes groot genoeg is. Als er minder dan 200 puntjes per 2,54 cm zijn, wordt de afbeelding onscherp of zie je soms de puntjes.

Rekenvoorbeeld: Een A4-blad is ongeveer 12”x8” (12 duim hoog, 8 duim breed). Het digitale beeld zou dus minstens 2400 puntjes in de hoogte hebben en 1600 puntjes in de breedte of ongeveer 4 miljoen puntjes (4 mega pixels of 4Mp). (Een A3 dus 8 Mp, ...)

### **sRGB** standaard RGB

Digitale kleurweergave met 3 kleuren Rood, Groen en Blauw met per kleur een intensiteit uitgedrukt in een getal van 8 bits, meestal ( $3 * 8 =$ ) 24 bit kleur genoemd.