

Rapport

Digitaal beeldarchief

*Aanbevelingen voor het realiseren van beeldbanken,
aan de hand van 'lessons learnt' en 'best practices'*



Rapport

Digitaal beeldarchief

*Aanbevelingen voor het realiseren van beeldbanken,
aan de hand van 'lessons learnt' en 'best practices'.*

Matthias Vandermaesen, 2005



Met de steun van de
Vlaamse minister van Cultuur,
Jeugd, Sport en Brussel



Publicaties cDAVID - DIGITAAL ERFGOED

1. VANDERMAESEN M. en BAATEN L., *Digitaal geluidsarchief. Krachtlijnen digitalisering: Standaarden, formaten en dragers*, Antwerpen [on line publicatie], 2004, 40 p.
2. VANDERMAESEN M., *Digitaal Beeldarchief. Digitaliseren van polaroidfoto's*, Antwerpen, 2005, 16 p.
3. VANDERMAESEN M., *Digitaal Beeldarchief. Aanbevelingen voor het opzetten van beeldbanken, aan de hand van 'lessons learnt' en 'best practices'*, Antwerpen, 2005, 120 p.

Inhoud

<i>Voorwoord</i>	5
<i>Inleiding</i>	7
<i>> 1 Praktische overwegingen bij het opzetten van een beeldbank</i>	9
1.1 De prioriteiten en doelstellingen	10
1.2 De collectie	12
1.2.1 De doelstellingen van het project	
1.2.2 De waarde van het materiaal	
1.2.3 De fysieke kenmerken en staat	
1.3 De auteursrechten	17
1.3.1 Het openbaar domein	
1.3.2 De wettelijke uitzonderingen	
1.3.3 De auteursrechtelijke bescherming	
1.3.4 Het auteursrecht en foto's	
1.3.5 Conclusie	
1.4 Het doelpubliek	20
1.5 De instelling	22
1.6 De voor- en haalbaarheidsstudie	30
<i>> 2 Het digitaliseringproces</i>	33
2.1 Het belang van kwaliteitseisen	34
2.2 Technische begrippen	37
2.2.1 De werking van scanners en digitale camera's	
2.2.2 Resolutie, bitdiepte en dynamisch bereik	
2.2.3 Andere parameters	
2.3 De bestandsformaten	49
2.3.1 Open versus gesloten formaten	
2.3.2 Compressie	
2.3.3 Een geschikt bestandsformaat kiezen	
2.4 Workflow	54
2.4.1 De voorbereiding van de te digitaliseren stukken	
2.4.2 Het instellen van de apparatuur	

- 2.4.3 Het aanmaken van de archiefkopie
- 2.4.4 Het controleren van de kwaliteit
- 2.4.5 Het optimaliseren van de kopie
- 2.4.6 Het produceren van afgeleide kopieën
- 2.4.7 De nazorg

> 3 Case study: de European Visual Archive beeldbank	77
3.1 Wat is EVA?	77
3.1.1 European Visual Archive - EVA, 1998 - 2001	
3.1.2 European Visual Archive Market Validation project - EVAmp, 2003 - 2004	
3.2 Planning en aanpak	78
3.3 Doelstellingen	79
3.3.1 Het bepalen van de prioriteiten	
3.3.2 Het analyseren en bepalen van het doelpubliek	
3.3.3 Het analyseren van het fotografische materiaal (de collectie), van de auteursrechten en van de tarieven	
3.3.4 Het bepalen van de standaarden voor kwaliteitsvolle digitale beelden	
3.3.5 De workflow	
3.4 De werking van de EVA-beeldbank en -website	102
3.4.1 Welke images of beelden worden op de website geplaatst?	
3.4.2 Welke beschrijvingsgegevens worden overgenomen op de website?	
3.4.3 Hoe verloopt de overdracht van de archieven naar de beeldbank?	
3.4.4 De web front-end en de gebruiker	
Besluit: Lessen geleerd uit het realiseren van de EVA-beeldbank	115

Voorwoord

Het project cDAVID promoot de zorg voor het digitale culturele erfgoed en wil daarom de Vlaamse erfgoedinstellingen zoveel mogelijk instrumenten en aanbevelingen aanbieden.

De lezer zal merken dat de hier geformuleerde aanbevelingen heel uitgebreid zijn. Ze hebben betrekking op het voeren van een grondige voorstudie of het vastleggen van gedetailleerd uitgewerkte werkmethodes of workflow.

Volgens de methodiek die eigen is aan de DAVID-projecten, wordt ook hier aandacht gevraagd voor het werken met standaarden, het aanleggen van documentatie en metadata. Dit alles om de duurzaamheid en betrouwbaarheid van wat archieven en erfgoedinstellingen beheren, te verzekeren. De grote inspanningen en investeringen eigen aan digitaliseringprojecten van archief- en erfgoedmateriaal, moeten bestand zijn tegen de veroudering van hard- en software.

Het EVA- of European Visual Archive project voor een beeldbank met de fotocollecties uit de archieven van Antwerpen en Londen, geeft de mogelijkheid aan de sector om te leren uit ervaringen van derden.

Daarom wordt in dit rapport bij de bespreking van het EVA-project veel aandacht besteed aan de overwegingen die erfgoedinstellingen en gebruikers van beeldbanken maken bij het deelnemen aan een beeldbank.

In het domein van de digitale beeldarchieven is met dit rapport het laatste woord nog niet geschreven. Een aantal aspecten blijven voorlopig nog onderbelicht.

In de toekomst zullen over digitale beeldarchieven zonder twijfel nog meer rapporten en aanbevelingen worden uitgewerkt en gepubliceerd.

Inge Schoups
Stadsarchivaris Antwerpen
2005



Koninklijk Museum voor Schone Kunsten tijdens het Rubensjaar met de zegewagen van Calloo, opgesteld voor het museum, 1977.

Foto Guido Coolens, © Foto Coolens & Deleuil (stadsarchief Antwerpen, FOTO 20846).

Inleiding

Culturele instellingen die ons erfgoed beheren, trekken de laatste jaren meer en meer de digitale kaart. Het digitaal beschikbaar maken van de collectie via nieuwe media zoals het internet, is een populaire oplossing. Er zijn heel wat redenen om een digitaliseringproject te starten. Zo kan moeilijk raadpleegbaar materiaal zoals geluids- of video-opnamen na het digitaliseren via het internet beschikbaar worden gesteld. Heel wat instellingen in binnen- en buitenland hebben via een beeldbank hun belangrijke fotocollecties digitaal opengesteld voor het grote publiek. Zo ook het stadsarchief Antwerpen dat door zijn participatie in het EVA-project bij de eerste projecten in Vlaanderen was dat aan de slag ging met het digitaliseren en digitaal ontsluiten van fotografisch collecties.¹

Ondanks de voordelen is de realisatie van een dergelijk project verre van evident. De technologische veroudering van dragers en digitale formaten bedreigen de leesbaarheid van digitale data op lange termijn. Wat met de technologie van vandaag leesbaar is, is dat over tien jaar misschien niet meer. Het opzetten van een digitale beeldbank is op zich een uitdaging, het operationeel houden van die beeldbank een andere. Digitaliseren, duurzame digitale bewaring en toegankelijkheid eisen net zoveel inspanningen als het onderhouden van analoge collecties. Dit rapport wil de erfgoedbeheerder niet alleen een blik gunnen op wat er allemaal bij digitaliseren komt kijken, maar digitalisering ook in een beter perspectief plaatsen.

In een eerste deel gaan we dieper in op een aantal belangrijke overwegingen die je moet maken wanneer je beslist een beeldbank op te starten. We geven een overzicht van de mogelijke noden en functionele eisen die gebruikers en instellingen aan een digitale beeldbank stellen. We bekijken met welke prioriteiten je rekening moet houden. Het langzame maar onvermijdelijke verval van dragers leidt tot de teloorgang van waardevolle tijdsbeelden. Het digitaliseren is hier het redmiddel bij uitstek. Ten slotte besteden we in het eerste deel aandacht aan items zoals selectie, auteursrechten en kosten.

In het tweede deel brengen we de technische aspecten van het digitaliseringproces in beeld. Het eigenlijke digitaliseren van de collectie is een kritiek moment. De nodige technische expertise is vereist om rekeninghoudend met de doelstellingen van het project, een geschikt digitaal resultaat te produceren. Dit is nodig om de bruikbaarheid van het digitale archief te maximaliseren en de nood om de – vaak fragiele – originelen opnieuw bloot te stellen aan de stress van het digitaliseringproces te minimaliseren.² Het is niet de bedoeling om een volledig technisch overzicht te brengen maar inzicht te verschaffen wat digitaliseren inhoudt.

In een derde deel behandelen we het EVA-project. Het stadsarchief Antwerpen was een belangrijke speler in de uitbouw van het European Visual Archive. We geven hier een overzicht van de realisatie van het project; problemen waarmee men werd geconfronteerd en de praktische oplossingen die hiervoor werden ontwikkeld. De aanbevelingen uit het EVA-project worden in dit rapport gesynthetiseerd. We vergelijken deze met de huidige praktijk en werkwijze in het Antwerpse stadsarchief.

¹ EVA, European Visual Archive: <http://www.eva-eu.org/> (21 maart 2005).

² F. FREY, S. SÜSSTRUNK, Image Quality Issues for the Digitization of Photographic Collections. In Proceedings, IS&T 49th Annual Conference, Minneapolis, MN, 1996. 349 - 353.



Vertrek van koning Albert I en koningin Elisabeth naar Kongo: de koninklijke familie op de eretribune, 5 juni 1928.

Foto E. Bastyns (stadsarchief Antwerpen, PEO 3698).

1 Praktische overwegingen bij het opzetten van een beeldbank

Aan een omvangrijk project zoals het opzetten van een beeldbank gaat heel wat planning vooraf. Dit is de eerste stap in een digitaliseringproject. Vaak gaat men hier snel over omwille van uiteenlopende redenen. Nochtans zijn er heel wat vragen die moeten worden beantwoord en knopen die moeten worden doorgehakt. Een digitaliseringproject beperkt zich immers niet tot het louter “scannen van mijn documenten”. Dit is slechts een onderdeel ervan. Een dergelijk project impliceert ook het efficiënt beheren van personeel, infrastructuur en budget, het plannen van de praktische realisatie van het digitaliseren (*workflow*), de selectie van die stukken die in aanmerking komen, het opstellen van kwaliteitseisen voor de digitale kopieën, het beschrijven van de stukken, het beheren en beschikbaar stellen van de digitale collectie, de voor- en nazorg van de originele stukken, enzovoort.

De voornaamste reden om een goede planning op te stellen is om met succes de beoogde resultaten te behalen. Een goede voorbereiding verhoogt de kans dat het project zal slagen in zijn opzet en zal bijdragen tot een betere werking van de instelling in zijn geheel. Een digitaliseringproject vraagt immers een zekere inzet tijdens de hele operatie, maar ook achteraf. Wanneer hiermee geen rekening wordt gehouden, dan kunnen de gevolgen navenant zijn: een financiële kater, beschadigde stukken, digitale kopieën die niet deugen en zelfs ‘gezichtsverlies’.

Een goede planning heeft onder andere volgende voordelen:

- Een efficiënt gebruik van het personeel en de middelen die je reeds ter beschikking hebt.
- Een beter beeld van het budget. Met name de kosten die moeten worden gemaakt maar ook de mogelijkheden om fondsen te verwerven.
- Een goed idee van de impact die het project zal hebben op de eigen werking of de *core business*. Een belangrijke vraag is: draagt het project bij tot het volbrengen van mijn doelstellingen?
- Het in kaart brengen van de mogelijkheden. Soms moet je halverwege het project de oorspronkelijke doelstellingen laten varen omdat ze onmogelijk zijn uit te voeren. Een zware koerscorrectie betekent extra investeringen, vertragingen, enzovoort.
- Een goede planning laat toe om in de toekomst te kijken. Projecten zijn vaak beperkt in tijd. Eenmaal de digitalisering voltooid is, blijft er een constante nazorg nodig. Zowel voor de systemen die de digitale collectie beschikbaar maken als voor de digitale beelden zelf. Het onderhoud kan in de planning worden ingecalculeerd. Er kan ook worden geanticipeerd op het hergebruik van elementen en beelden in andere projecten.

In dit hoofdstuk bekijken we een aantal praktische afwegingen. Het komt er op aan om vooraf een juiste inschatting te maken van alle factoren die het eindresultaat kunnen beïnvloeden. Onder druk kan je de neiging hebben om hier snel over te gaan, maar voldoende investeren in tijd en middelen tijdens deze eerste voorbereidende planningsfase, rendeert op lange termijn in de kwaliteit van het eindproduct, een soepele en goed geïntegreerde workflow en bovenal de kosten.

Het is de bedoeling om antwoorden te verkrijgen op de volgende vragen:

- Waarom digitaliseren? (1.1)
- Wat gaan we digitaliseren? (1.2 en 1.3)
- Wie willen we met onze beeldbank bereiken? (1.4)
- Wat zijn de kosten en welke fondsen zijn beschikbaar? (1.5)
- Wie moet het werk doen en waar? (1.5)
- Hoe moet het werk verlopen (welke standaarden, welke 'best practices' e.a.)? (1.5)

Op basis van aanbevelingen geformuleerd door gespecialiseerde onderzoekscentra willen we stilstaan bij een aantal praktische overwegingen die iedereen die met een dergelijk project wordt geconfronteerd, in acht zou moeten nemen vooraleer over te gaan tot een feitelijke realisatie.³

1.1 De prioriteiten en doelstellingen

Bij het opstarten van elk project moeten er doelstellingen worden vastgelegd. Wat wil je met het project bereiken? Waarom willen we digitaliseren? 'Iedereen doet het' of 'we hebben er de fondsen voor' zijn geen valabele uitgangspunten. Digitaliseren is immers geen doel op zich maar een middel om de missie van de instelling te ondersteunen. Pas digitaliseren daarom toe wanneer het zich aandient als een goede oplossing om de bestaande situatie te verbeteren.

Vooraleer je met de planning van het project begint is het noodzakelijk om de doelstellingen vast te leggen. Dit is een bepalende factor voor de verdere planning en realisatie. Maak daarom eerst een lijstje met doelstellingen. Gebruik geen vage omschrijvingen maar probeer juist zo concreet mogelijk te zijn. Zo volstaat het bijvoorbeeld niet om te zeggen dat de digitale beelden van "volgende kwaliteit" moeten zijn. Beter is het om op te sommen waarvoor de digitale beelden op korte en op lange termijn moeten dienen.

- ... als digitaal substituuat zodat het origineel duurzaam kan worden bewaard?
- ... in een beeldbank die gericht is op wetenschappelijk onderzoek?

3 NINCH (National Initiative for a Networked Cultural Heritage), The NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials, Verenigde Staten, 2002, <http://www.nyu.edu/its/humanities/ninchguide/index.html> (21 maart 2005).

TASI (Technical Advisory Service for Images), Managing digitisation projects, Bristol, 2005, <http://www.tasi.ac.uk/advice/managing/managing.html> (21 maart 2005).

CLIR (Council on library and information resources), CLIR Reports, Washington DC, 2005, <http://www.clir.org/pubs/reports/index.html> (19 mei 2005).

M.K. SITTS (red.), *Handbook for digital projects. A management tool for preservation and access*, Northeast Document Conservation Center, Massachusetts, 2000, <http://www.nedcc.org/digital/dighome.htm> (18 maart 2005).

- ... in een educatieve, thematische beeldbank?
- ... in een beeldbank waaruit media zoals televisiezenders, kranten, uitgevers en anderen kunnen putten?
- ... in een beeldbank voor het grote publiek die een globale indruk wenst te geven van de collectie?

De twee voornaamste redenen om te digitaliseren zijn enerzijds het *duurzaam bewaren* van de originele stukken en anderzijds het verbeteren en verbreden van de *toegankelijkheid*.

Conservering

Door een digitale kopie ter beschikking te stellen wordt een origineel stuk veel minder blootgesteld aan manipulaties en de risico's die daar aan vasthangen. Daardoor verhoogt de overlevingskans van het stuk.

Digitale toegankelijkheid

Een digitale kopie is veel minder gebonden door fysieke beperkingen eigen aan het analoge origineel. In de digitale wereld kan de informatie veel gemakkelijker ter beschikking worden gesteld van een breed publiek. Digitale stukken kunnen ook op nieuwe manieren worden benaderd.

Hoewel de twee elkaar niet uitsluiten is het belangrijk om de vraag te stellen waar de klemtoon van het project ligt: conservering of een betere toegankelijkheid? Ze veronderstellen elk een andere aanpak naar kwaliteitseisen, workflow, planning, enzovoort toe. Halverwege het project van richting veranderen is moeilijk, impliceert extra investeringen en tijdverlies.

Een andere motivatie om te digitaliseren is *digitale substitutie*. Hierbij wordt het origineel na het digitaliseren vernietigd. De digitale kopie is dan de enige weerslag die nog rest van het origineel. Daarom moet ze van een zeer goede kwaliteit zijn en aan rigoureuze eisen voldoen. Dit is zeer belangrijk aangezien er geen tweede kans is om het stuk te digitaliseren. Deze strategie kan in hoofdzaak worden toegepast wanneer je het origineel onmogelijk kan bewaren. Bijvoorbeeld wegens plaatsgebrek of omdat de originelen sterk beschadigd en niet meer te redden zijn.

Aan de hand van een lijstje met realistische en duidelijk geformuleerde doelstellingen kan je in een voorbereidende fase concrete uitspraken beginnen doen, over onder andere de selectie van de te digitaliseren stukken, maar ook over de te behalen kwaliteitseisen en zelfs over de aanpak. Tijdens de planning van het project worden de doelstellingen verder geconcretiseerd naarmate er meer duidelijkheid bestaat over beschikbare fondsen, middelen, personeel, de termijn waarin het project moet worden gerealiseerd, de hoeveelheid en de aard van het materiaal. Eenmaal er daadwerkelijk aan de slag wordt gegaan mogen de doelstellingen niet meer veranderen.

1.2 De collectie

In eerste instantie lijkt het voor de hand te liggen om de volledige collectie of een zo groot mogelijk deel te digitaliseren. Toch is dit niet noodzakelijk de te verkiezen aanpak.⁴

- Vaak zijn middelen en budget eerder beperkt waardoor een weloverwogen keuze zich opdringt. De investering gaat verder dan louter het digitaliseren. Ook de kwaliteitscontrole, het beschrijven, het voorbereidende werk en de nazorg van zowel de originelen als de digitale bestanden op lange termijn impliceren een niet te onderschatten kost.
- Heel wat erfgoedmateriaal is auteursrechtelijk beschermd waardoor het niet zonder meer mag worden gepubliceerd in een online beeldbank.
- Niet alle stukken in een collectie wekken de interesse van het publiek. Anderzijds kunnen sommige collecties juist gevoelig materiaal bevatten waarvan verspreiding naar een breed publiek toe niet wenselijk is.
- De fysieke staat en de aard van de stukken laten soms niet toe om op een goede manier te digitaliseren met de middelen die voorhanden zijn.

Het is duidelijk dat we niet zomaar letterlijk “alles” kunnen digitaliseren. We moeten een selectie maken van wat hiervoor in aanmerking komt. Daarbij moeten goede selectiecriteria worden opgesteld om problemen zoals extra kosten, juridische geschillen of het niet aanspreken van het doelpubliek te vermijden. Dit is een van de sleutels tot succes.

Het toepassen van selectiecriteria betekent ook het bepalen van prioriteiten: welke stukken moeten eerst worden gedigitaliseerd? Bij conservering ligt de prioriteit bij de stukken die op korte termijn door verval worden bedreigd. Wanneer het accent op het verbeteren van de toegankelijkheid ligt, dan kan je prioriteit geven aan stukken waar veel interesse voor bestaat.

Bij het opstellen van selectiecriteria zijn er een aantal overwegingen in acht te nemen.⁵

- Welke doelstellingen streeft het project na?
 - ~ de conservering van de originelen?
 - ~ het verbeteren van de toegankelijkheid?
- Wat is de inhoudelijke waarde van de stukken?
- Zijn de stukken auteursrechtelijk beschermd? Kan je de toestemming verkrijgen om de stukken te digitaliseren en digitaal ter beschikking te stellen?

4 D. VOGT-O'CONNOR, *Selection of materials for scanning*, in: M.K. SITTS (red.), *handbook for digital projects a management tool for preservation and access*, 2000, p. 45 – 73, <http://www.nedcc.org/digital/dighome.htm> (18 maart 2005).

5 D. VOGT-O'CONNOR, *Selection of materials for scanning*, in: M.K. SITTS (red.), *handbook for digital projects a management tool for preservation and access*, 2000, p. 45 – 73, <http://www.nedcc.org/digital/dighome.htm> (18 maart 2005).

Columbia University Library. Selection criteria for digital imaging, <http://www.columbia.edu/cu/lweb/projects/digital/criteria.html> (23 mei 2005).

D. HAZEN et al., *Selection research collections for digitisation*, Council on Library and Information Resources, Washington DC, 1998, <http://www.clir.org/pubs/abstract/pub74.htm> (23 mei 2005).

- ↪ Wat is de fysieke aard van het te digitaliseren materiaal? Twee elementen zijn hier van belang.
 - ~ De staat van het stuk (beschadigingen, verval, enzovoort)
 - ~ De fysieke kenmerken (formaat, drager, productieprocédé, enzovoort).

Dit zijn slechts een aantal van de overwegingen die je moet maken bij de selectie van het te digitaliseren materiaal. Het opstellen van selectiecriteria veronderstelt dat je ook reeds min of meer een beeld hebt van andere aspecten, zoals bijvoorbeeld het budget dat voorhanden is, de kosten, de beschikbare apparatuur, enzovoort.

1.2.1 De doelstellingen van het project

Een eerste aandachtspunt vormen de doelstellingen van het project. Deze bepalen ten dele de te hanteren selectiecriteria.⁶

- ↪ Selectiecriteria bij conservering.
 - ~ Wat is de fysieke staat van een stuk? In welke mate is het direct bedreigd door verval? Is het nodig om het daarom prioritair te behandelen?
 - ~ In welke mate is een stuk waardevol genoeg om digitalisering en digitale bewaring te rechtvaardigen?
 - ~ In welke mate kan de *authenticiteit* worden gegarandeerd? De digitale kopie kan niet altijd alle essentiële formele en inhoudelijke kenmerken van het stuk bevatten.
 - ~ Garandeert de kwaliteit van de digitale kopieën duurzame digitale bewaring? Opnieuw digitaliseren impliceert kosten en risico's. Probeer dit te vermijden door digitale kopieën te maken die binnen een ruime waaier van toepassingen bruikbaar zijn.
 - ~ Soms kan je een stuk niet langer bewaren. Bijvoorbeeld omwille van plaatsgebrek. Digitale stukken nemen veel minder plaats in. Aangezien er geen tweede kans op digitaliseren bestaat, is het zeer belangrijk om de voorgaande selectiecriteria goed in acht te nemen. Met name het garanderen van de authenticiteit en de duurzame lange termijn bewaring van de digitale kopie.
- ↪ Selectiecriteria bij het verbeteren van de toegankelijkheid.
 - ~ Zijn de analoge stukken (reeds) (voldoende) toegankelijk?
 - ~ Zijn er fysieke kenmerken die de toegankelijkheid van het origineel beperken? (de staat, het type drager, enzovoort)
 - ~ Is de collectie geografisch verspreid (eventueel over verschillende instellingen of departementen)? In een beeldbank kan je de collectie bijvoorbeeld digitaal terug samenbrengen.

⁶ J. GERTZ, *Selection Guidelines for Preservation*. In: *Guidelines for digital imaging. Joint RLG and NPO Conference*, 1999, <http://www.rlg.org/preserv/joint/gertz.html> (18 mei 2005).

IFLA, *Guidelines for digitization projects for collections and holdings in the public domain, particularly those held by libraries and archives*, IFLA, 2002, <http://www.ifla.org/VII/s19/pubs/digit-guide.pdf> (20 mei 2005).

- ~ Is er vraag naar een verbeterde toegankelijkheid van het publiek? Kan het digitaal beschikbaar maken hier een oplossing bieden?
- ~ Kan een collectie integraal worden gedigitaliseerd of bestaat de beeldbank uit een eerder losse verzameling stukken die uit verschillende collecties komen? Ook in de digitale wereld krijgt een stuk meer waarde wanneer ook de context mee kan worden opgenomen.

1.2.2 De waarde van het materiaal

Een tweede aandachtspunt is de waarde van het materiaal. Zijn de stukken zowel inhoudelijk als formeel voldoende aantrekkelijk om te digitaliseren? Het is niet eenvoudig om deze vraag te beantwoorden. Factoren zoals de doelstellingen van het project en de wensen van het doelpubliek spelen een niet onbelangrijke rol. Het antwoord is hoe dan ook subjectief.

Een aantal relevante vragen die toelaten de waarde van een document in te schatten:

- Zal een digitale kopie de interesse van het doelpubliek weten te wekken?
- Is een stuk voldoende representatief voor het onderwerp van de digitale collectie? Zijn andere stukken met hetzelfde onderwerp meer of minder geschikt?
- Is een stuk uniek of historisch waardevol?
- De mate waarin je een stuk goed kan beschrijven, bepaalt eveneens zijn waarde. Iets wat uitgebreid werd beschreven, vind je gemakkelijker terug in de digitale collectie.
- Creëert het digitaal aanbieden een meerwaarde?

Digitale technologie biedt ons nieuwe mogelijkheden om een collectie te bekijken. Een digitaal stuk kent de beperkingen van het analoge origineel niet. Probeer te bepalen in welke mate digitale technologie mogelijk nieuwe inzichten kan verschaffen.

Als voorbeeld geven we een aantal modelvragen voor het opnemen van stukken in een wetenschappelijke beeldbank.

- ~ Digitaliseren en on line ontsluiting maken het mogelijk om een ruimer publiek van wetenschappelijke vorsers te bereiken. Zo bevordert digitale ontsluiting de interdisciplinaire samenwerking. Kan een stuk de interesse wekken van onderzoekers uit andere vakgebieden?
- ~ Digitale ontsluiting biedt de functionaliteit en mogelijkheden die in de analoge wereld niet bestaan en die wetenschappelijke inzichten bevorderen.
- ~ Digitale ontsluiting vergroot de afstand tussen het object en de onderzoeker. Die laatste komt nu immers niet meer rechtstreeks in contact met het originele stuk. Het tactiele element dat soms van belang is in bepaalde onderzoeksvelden, verdwijnt.
- ~ Auteurs- en aanverwante intellectuele rechten kunnen een hinderpaal vormen om de volledige collectie te digitaliseren. Dit belemmert wetenschappelijk onderzoek aangezien dan niet de volledige collectie digitaal ter beschikking wordt gesteld.

~ Het belang van metadata:

- Er moet voldoende informatie over een beeld of de *metadata* aanwezig zijn om degelijk onderzoek mogelijk te maken.
- Metadata identificeren een stuk en zorgen ervoor dat je het kan terugvinden in de collectie. Hoe beter de beschrijving, hoe gemakkelijker een stuk is terug te vinden.
- Metadata wordt meestal bewaard in databanken. Die databanken kunnen on line aan elkaar worden gekoppeld zodat informatie over de grenzen van collecties heen met elkaar kan worden verbonden. Dit noemt men ook wel *interoperabiliteit* of *uitwisselbaarheid*.

Het bepalen van de waarde van een stuk is eerder subjectief. De ervaring van diegene die hiervoor verantwoordelijk is, speelt een doorslaggevende rol. De selectie gebeurt best door iemand die een goede inschatting kan maken van het historische belang, de relevantie en de uniciteit van een beeld enerzijds, en van de meerwaarde van de digitale kopie anderzijds.⁷



Soldaten rond een bierton, 1907 (stadsarchief Antwerpen, OF 1708).

1.2.3 De fysieke kenmerken en staat

Een derde aandachtspunt zijn de *fysieke kenmerken* en de *staat* waarin het materiaal verkeert. Dit heeft een impact op de behandeling van het materiaal voor, tijdens en na het digitaliseren. De hoofdregel is dat je er ten allen tijde op moet letten dat het materiaal zo min mogelijk wordt blootgesteld aan risico's op beschadiging. Dit geldt eveneens bij digitaliseren.

⁷ NINCH, *The NINCH Guide to Good Practice*, 2002, <http://www.nyu.edu/its/humanities/ninchguide/III/> (21 maart 2005).

Volgende selectiecriteria zijn onder andere van toepassing:

- Is het mogelijk om de stukken op een goede en veilige manier te digitaliseren met de apparatuur die voorhanden is?
- Wat zijn de kwaliteitseisen die je aan een digitale kopie stelt? In welke mate bemoeilijkt de staat van het stuk het behalen van die kwaliteitseisen?
- In hoeverre hebben de fysieke kenmerken een impact op de kosten? Een polaroidfoto is gemakkelijker te scannen dan een fragiele landkaart. De kosten zijn navenant. Kunnen die extra kosten worden gerechtvaardigd?

Volgende fysieke kenmerken hebben impact op het digitaliseringproces. Probeer in te schatten welke gevolgen dit zal hebben voor het geheel. Het advies van een expert is daarbij van onmisbare waarde.

- Het type van het object (boeken, manuscripten, glasnegatieven, kranten, ...).
- Het productieprocédé (handgeschreven of gedrukte tekst, lithografie, gravure, ...; in de fotografie: collodium-, cyano- of daguerrotypie, ...).
- De ouderdom van het object.
- De dimensies (bijvoorbeeld de afmetingen van het boek maar ook het gebruikte lettertype; uniformiteit van de grootte in de collectie, ...).
- Welke fysieke drager werd er gebruikt (glas, acetaat, polyester, papier, ...)?
- De lichtgevoeligheid, heel wat objecten zijn gevoelig voor licht. Een te lange blootstelling aan fel licht kan leiden tot beschadigingen en een versneld verval van het object. Er moet worden nagegaan in welke mate het object licht kan verdragen.
- De kleuren. De kleur kan een belangrijke rol spelen in de eigenheid van het object. Er moet worden nagegaan of kleur al dan niet mee moet worden opgenomen tijdens het digitaliseren.
- De toon. In welke mate heeft het object veel lichte en/of donkere gebieden?
- De structuur van de inhoud. Bevat een object zowel illustraties als tekst?
- De structuur van de drager. Is het een ingebonden boek? Een katern uit een krant? Een los document?
- De staat waarin het materiaal verkeert. In welke mate is het document reeds onderhevig aan verval? Is dit reeds in kaart gebracht door een expert? Kan het materiaal na het digitaliseren nog verder worden bewaard? In welke mate is er nood aan een speciale behandeling (schoonmaak, restauratie, ...)?

1.3 De auteursrechten

Een bijzonder aandachtspunt vormen de auteursrechten.⁸ Bij een digitaliseringproject word je bijna onvermijdelijk geconfronteerd met auteursrechten. Dergelijk materiaal kan je niet zonder meer reproduceren en publiceren. Daarvoor is altijd de toestemming van de auteur nodig. Dit negeren is een inbreuk op de auteurswet en kan een zwaar risico vormen voor de betrokken instellingen. Er zijn echter ook uitzonderingen waar toestemming helemaal niet nodig is. Bijvoorbeeld omdat een stuk tot het publieke domein behoort. Het is noodzakelijk om bij de selectie na te gaan of het al dan niet nodig is om de toestemming van een auteur te vragen.

1.3.1 Het openbaar domein

Het openbaar domein vormt een verzameling van werken die niet onder het auteursrecht vallen. Om deze te gebruiken heb je geen toestemming nodig. Er zijn echter een aantal voorwaarden waaraan moet worden voldaan:

- ↪ Een werk is niet oorspronkelijk of is niet voldoende vormelijk. Het is oorspronkelijk wanneer het blijk geeft van een voldoende creatieve inspanning vanwege de auteur. Het is voldoende vormelijk wanneer het voldoende voorstelbaar, geconcretiseerd of gematerialiseerd is.
- ↪ De auteursrechtelijke beschermingstermijn moet verstreken zijn. Elk werk wordt niet eeuwig beschermd. Wanneer deze termijn verstreken is, behoort het werk automatisch tot het openbare domein. In de regel loopt deze periode tot 70 jaar na het overlijden van de auteur of 70 jaar na het moment van het publiekelijk bekendmaken bij een anoniem of pseudoniem werk. Niet alleen foto's, teksten, geluids- en video-opnamen maar ook computerprogramma's worden auteursrechtelijk beschermd gedurende 70 jaar vanaf de bekendmaking of vanaf het overlijden van de auteur.

1.3.2 De wettelijke uitzonderingen

Een werk hoeft niet noodzakelijk tot het publieke domein te behoren om vrij te zijn van auteursrechten. Er zijn een aantal wettelijke uitzonderingen die toelaten zonder toestemming van de auteur en de belanghebbenden, delen van werken publiek te maken.

↪ **Reproductie voor onderwijsdoeleinden, wetenschappelijk onderzoek of archieven.**

Het is mogelijk om werken zonder toestemming van de auteur te reproduceren met het oog op onderwijs, onderzoek of voor de goede bewaring van het cultureel patrimonium. Vaak wordt een digitaliseringproject voor deze doeleinden ontwikkeld. Het ligt dan ook voor de hand te argumenteren dat de inhoud van een beeldbank niet aan het auteursrecht onderhevig is. Deze uitzondering heeft echter enkel betrekking op de reproductie van auteursrechtelijk beschermd materiaal. Audiovisuele beeldbanken gaan echter verder en stellen het materiaal ook beschikbaar voor het publiek. Dit gaat verder dan louter repro-

⁸ A. LEYSEN, *Auteursrechten audiovisuele archieven. Film- en geluidsarchieven*, Stadsarchief Antwerpen, Antwerpen, 2001.

E. MICHIELSEN en A. ROSSENBACKER (red.), *auteursrecht en erfgoed. Handleiding tot het vermijden van uitschuivers*, Culturele Biografie Vlaanderen, Antwerpen, 2004.

ductie. Voor commercieel gebruik of de verspreiding van beschermd materiaal is steeds de toelating vereist.

→ **Citaatrecht en bloemlezing**

Je kan je ook beroepen op het citaatrecht. In dat geval is het mogelijk om een werk aan te halen dat auteursrechtelijk beschermd is. Daardoor wordt het mogelijk om korte tekstfragmenten (een 15 regels) uit beschermde werken over te nemen en zelfs integrale beeldende werken. Een belangrijke beperking is dat het citaatrecht enkel van toepassing is in het kader van een eigen uiteenzetting. Een beeldbank die alleen maar bestaat uit citaten kan dus niet. Een aanhaling als smaakmaker binnen een inhoudelijk kader dat refereert aan kritiek, polemieken, onderwijs of wetenschappelijk onderzoek kan wel.

Veelvuldig gebruik maken van het citaatrecht kan ertoe leiden dat je een werk kan beschouwen als een bloemlezing. Dit geldt ook voor een online beeldbank. In dat geval moet er alsnog worden nagegaan in welke mate de aangehaalde werken auteursrechtelijk beschermd zijn.

Deze wettelijke uitzonderingen laten echter een grote marge open voor twijfelgevallen. Het is moeilijk om vast te stellen wanneer een werk dat zich veelvuldig beroept op het citaatrecht als een bloemlezing moet worden beschouwd. Bovendien is het vermelden van de naam van de auteur of de oorspronkelijke bron verplicht wanneer je gebruik wil maken van deze wettelijke uitzonderingen.

1.3.3 De auteursrechtelijke bescherming

Een werk is auteursrechtelijk beschermd wanneer er een auteur is. Dit is de schepper van het werk. Schrijvers, fotografen, programmeurs of componisten zijn allen auteurs van hun werk. Dit omvat alles wat ze produceren gaande van novellen, gedichten, foto's, muziek tot computerprogramma's. Er zijn twee basisvoorwaarden waaraan een werk moet voldoen om auteursrechtelijke bescherming te genieten.

→ *Het werk moet oorspronkelijk of origineel zijn.*

Een werk wordt als oorspronkelijk of origineel beschouwd wanneer de persoonlijkheid van de auteur erin duidelijk tot uiting komt. Kortom, dat het werk blijkt geeft van een creatieve / intellectuele inspanning van de maker waaraan het zijn individuele karakter ontleent.

→ *Een werk moet een bepaalde vorm bezitten.*

Het gaat om de gematerialiseerde creatie. Een werk moet concreet waarneembaar zijn; een voorstelbaar iets zijn.

De auteur van een werk is diegene die het oorspronkelijke werk heeft gecreëerd en in een oorspronkelijke vorm gegoten. Het gaat er niet om na te gaan wie de auteur is maar om te bewijzen dat iemand de auteur van een werk is. De algemene regel stelt dat naamsvermelding geldt als bewijs van auteurschap. Dit kan op drie manieren:

→ Met een naam. Zowel de echte naam van de auteur als een pseudoniem gelden.

→ Met een letterwoord waarmee de auteur te identificeren is en dat op het werk wordt vermeld.

- Met de naam van de uitgever wanneer het om een anoniem of pseudoniem werk gaat en de oorspronkelijke auteur niet gekend is. In dat geval geldt de uitgever als de auteur van het werk ten aanzien van derden.

Enkel de auteur heeft het recht om het werk te (laten) reproduceren of om het bestaan van het werk aan het publiek mee te delen.

1.3.4 Het auteursrecht en foto's

Voor foto's en afbeeldingen gaat het auteursrecht verder dan de foto zelf. Ook het onderwerp van het beeld kan auteursrechtelijk beschermd zijn. Niet alleen de vraag "wie heeft gefotografeerd en hoe?" maar ook de vraag "wat werd gefotografeerd" is belangrijk. Wat de foto zelf betreft, geldt dat het auteursrecht van toepassing is wanneer er voldaan werd aan de oorspronkelijkheidsvereiste - concreet geïnterpreteerd als vrijheden in interpretatie en opstelling van de fotograaf, de compositie, de uitstraling van het beeld,... - en de vormvereiste (zowel analoog als digitaal). De auteursrechtelijke beschermingstermijn voor foto's is 70 jaar na het overlijden van de auteur of na bekendmaking wanneer het om een anoniem of pseudoniem werk gaat.

Wat betreft het auteursrecht op het onderwerp van een foto, maken we het onderscheid tussen afbeeldingen van menselijke scheppingen en personen. Zo vormt een kunstwerk een menselijke schepping en is het auteursrecht van toepassing waardoor er toestemming aan de auteur van het kunstwerk moet worden gevraagd. Niettegenstaande het begrip 'kunst' voor interpretatie vatbaar is, is het veilig om altijd eerst toestemming te vragen. Een gebruiksvoorwerp is eveneens een menselijke schepping. Dit gaat van gebouwen tot designerobjecten en meubilair. Gebruiksvoorwerpen gelden als uitvinding, model of ontwerp van een auteur en zijn auteursrechtelijk beschermd. Louter functionele vormen of technisch-bouwkundige voorzieningen zoals sleutels, koorden,... vallen doorgaans onder het publieke domein. Ook natuurlijke scheppingen zoals natuurelementen of landschappen, genieten geen auteursrechtelijke bescherming. Tenzij er een duidelijke menselijke invloed is geweest in hun schepping. Voor een foto van een natuurlandschap met een menselijke productie (bruggen, gebouwen,...) geldt het auteursrecht indien de fotograaf de bedoeling heeft om de menselijke productie centraal te plaatsen.

Voor het fotograferen of afbeelden van personen geldt het portretrecht of het recht op afbeelding. Dit betekent dat eenieder het fotograferen of afbeelden van de eigen persoon kan weigeren. Het portretrecht geldt wanneer de afgebeelde persoon duidelijk herkenbaar is en als de afbeelding duurzaam (dit geldt ook voor digitale afbeeldingen) is en kan worden medegedeeld. Het portretrecht geldt tot twintig jaar na het overlijden van de afgebeelde persoon.

Er zijn echter vier bijzondere categorieën:

- Een bekend persoon verleent stilzwijgend de toestemming om te worden afgebeeld. Je kan de persoon dus afbeelden zonder expliciet toestemming te vragen. De afbeelding moet wel betrekking hebben op het beroep, de situatie of de reden waarom hij / zij in het middelpunt van de belangstelling staan. De afbeelding mag ook geen inbreuk vormen op andere rechten zoals het recht op privacy of worden gebruikt voor commerciële of publicitaire doeleinden.
- Voor het afbeelden van een minderjarige is de toestemming van de ouders nodig.
- Je kan eveneens niet zomaar foto's nemen van verdachte of criminele personen die strafrechtelijk worden vervolgd.

- Tenslotte kan je weliswaar foto's nemen van personen op openbare plaatsen zolang de foto anoniem is, er geen winstbejag mee wordt nagestreefd en de afgebeelde personen niet belachelijk worden gemaakt of gekwetst.

1.3.5 Conclusie

Het is duidelijk dat het auteursrecht een niet te onderschatten element is in het opzetten van een beeldbank. De beperkingen en de eis om over de toestemming van de verschillende belanghebbenden te beschikken, maken het allesbehalve gemakkelijk om recente beelden te digitaliseren en in een online beeldbank toegankelijk te maken voor het grote publiek. Bij een omvangrijke collectie foto's wordt het al snel onmogelijk om voor elke foto apart de verschillende belanghebbenden te gaan opsporen en hun toestemming te vragen.

Het valt aan te raden om de auteursrechtelijke problematiek zoveel mogelijk uit de weg te gaan door tijdens de selectie foto's te kiezen die tot het openbare domein behoren en niet meer door het auteursrecht worden beschermd. Concreet betekent dit dat voornamelijk foto's die zeventig jaar en ouder zijn in aanmerking komen voor online publicatie. Dateren van de foto is dus uiterst belangrijk. Vaak staat de naam van de fotograaf en het jaartal wel vermeld op de foto zelf. Indien dit niet zo is, is het beter om op veilig te spelen en ze niet op te nemen in de beeldbank. Ook foto's met mensen kunnen een probleem vormen. Ook hier is het beter om ze niet op te nemen wanneer belanghebbenden (familie of de persoon zelf) niet meer te traceren zijn.

Dit is slechts een summier overzicht van de problematiek rond auteursrechten. Het is aan te raden om bij de start van een digitaliseringproject juridisch advies in te winnen en zo min mogelijk risico's te nemen.

1.4 Het doelpubliek

Veel digitaliseringprojecten hebben als doelstelling hun archieven via elektronische weg open te stellen. Weten welk doelpubliek je probeert te bereiken en wat het doelpubliek zelf wenst is belangrijk om het beoogde resultaat te behalen.

Doorgaans worden gegevens over de raadpleging van de analoge collectie gebruikt om inschattingen over de raadpleging van de digitale collectie te kunnen maken. Dit leidt niet noodzakelijk tot een genuanceerd beeld van het publiek dat uiteindelijk de beeldbank zal raadplegen. Een collectie prenten die moeilijk toegankelijk is in de analoge wereld zal niet zo vaak worden geraadpleegd maar in de digitale wereld kan juist het tegendeel waar zijn.⁹

Het publiek dat van een digitale beeldbank wenst gebruik te maken heeft specifieke noden en wensen. Zo zal de occasionele bezoeker een beeldbank op een andere manier wensen te benaderen dan een wetenschappelijk onderzoeker.

Volgende vragen zijn relevant om een beeld te krijgen van het publiek dat de beeldbank zal raadplegen.¹⁰

⁹ NINCH, *The NINCH Guide to Good Practice*, 2002, <http://www.nyu.edu/its/humanities/ninchguide/XII/> (22 maart 2005).

¹⁰ TASI, *selection procedures*, 2005, <http://www.tasi.ac.uk/advice/creating/selecpro.html#sp2> (22 mei 2005).

→ *Wat is het doelpubliek?*

Demografische data zijn hier interessant: leeftijdscategorieën, interesses, kennisniveau van moderne technologie, studenten, wetenschappers of leken. Het beantwoorden van deze vragen leidt tot het beter identificeren van het publiek dat je wil bereiken. Wees erop voorbereid dat dit een zeer heterogeen karakter kan hebben. De keuze van de te ontsluiten collectie is een belangrijke factor: de aard van de collectie bepaalt ten dele het publiek dat de beeldbank zal raadplegen.

→ *Wat zijn de noden en de wensen?*

Eenmaal het doelpubliek geïdentificeerd, moeten de noden en de wensen worden geïdentificeerd. Zo zal een publiek bestaande uit schoolgaande jeugd geen boodschap hebben aan een systeem met uitgebreide beschrijvingen. Een aantal voorbeelden:

- ~ Zal het publiek reproducties willen bestellen? Waarvoor zullen die dienen: voor publicatie op een website of in een tijdschrift?
- ~ Is het doelpubliek dat je wil bereiken gebaat bij uitgebreide, wetenschappelijke beschrijvingen of volstaat juist de meest essentiële informatie over een digitaal beeld?
- ~ Hoe willen de gebruikers omgaan met je beeldbank? Via een website of op een andere wijze? Hoe moet de toegang tot de informatie dan worden gepresenteerd? Met een sober invulformulier of via een intuïtieve, grafische omgeving?
- ~ Welke afbeeldingen worden het meest geraadpleegd? Welke onderwerpen vindt het doelpubliek interessant? Welke thema's spreken het meest aan?

Op basis van de antwoorden op deze vragen kan je een lijstje met *functionele eisen* opstellen. Dit is een geïntegreerd overzicht van de noden en de eisen van het doelpubliek. Het wordt gebruikt



Havenstaking: lockout van de dokwerkers, september 1907. Foto François SAYS (stadsarchief Antwerpen, OF 1726).

om de oorspronkelijk vooropgestelde doelstellingen aan te passen. Het is goed mogelijk dat het beeld dat het doelpubliek heeft van het eindresultaat, verschilt van dat wat de erfgoedbeheerder voor ogen staat.

Het is nuttig om dit doelpubliek niet alleen passief maar ook actief te betrekken in de ontwikkeling van de beeldbank. Een constante evaluatie van de verschillende aspecten door een representatieve groep van testgebruikers laat toe na te gaan in welke mate de beeldbank beantwoordt aan de eisen en noden van het publiek. Uiteindelijk wil elke beeldbank op een gebruiksvriendelijke manier de digitale collectie ter beschikking stellen het publiek.

Er bestaan verschillende methoden om gebruikersgegevens te verkrijgen. De meest gangbare technieken zijn *enquêtes* en *interviews* onder potentiële gebruikers. In eerste instantie kunnen de leeszaalbezoekers worden ondervraagd. In tweede instantie kan je op zoek gaan naar specifieke doelgroepen. Voor een beeldbank die mikt op een jong publiek kan je veel informatie verzamelen door samen te werken met scholen of andere onderwijsinstellingen. Bij het gebruik van dergelijke technieken moet je er op letten dat je niet louter gaat peilen naar statistische informatie. De achterliggende motivatie speelt evenzeer een rol.

1.5 De instelling

In dit rapport hanteren we de term *instelling* en we geven hieraan een zo ruim mogelijke invulling. Die omvat niet alleen publiekrechtelijke instellingen die erfgoed bewaren en / of beheren zoals musea, bibliotheken of overheidsarchieven maar ook organisaties met een privaatrechtelijke karakter. Naast instellingen die een collectie beheren en een digitaliseringproject opstarten, bestaan er instellingen die alleen maar collecties digitaliseren. Het is de taak van diegene die een collectie effectief beheert om de wenselijkheid van digitaliseren na te gaan. Daartoe moet je onder andere nagaan in welke mate de instelling zelf in staat is om een project op te starten en te onderhouden. Volgende sleutelaspecten spelen hierin een rol. Het is belangrijk om hieraan de nodige aandacht te schenken.

→ *Personeel*

Beschikt de instelling over voldoende personeel om een team van projectmedewerkers samen te stellen? Heeft het voldoende expertise? Is het nodig om (tijdelijk) externe medewerkers aan te werven?

→ *Infrastructuur*

Beschikt de instelling over de nodige plaats en apparatuur om het project uit te voeren? Moet er in nieuwe apparatuur worden geïnvesteerd? Is het beter om bepaalde zaken uit te besteden?

→ *Beschikbare fondsen en budgetmanagement*

Volstaan de financiële middelen om het project goed uit te voeren? Zal het niet meer kosten dan voorzien? Is er een back-up plan? Waar kan je kosten drukken of extra fondsen vandaan halen?

→ *Samenwerking*

Is het mogelijk om ondersteuning te krijgen van professionele partners of andere instellingen die reeds digitaliseringprojecten hebben uitgevoerd? Is het mogelijk om een deel van

het werk uit te besteden aan een externe partner of een ander – beter uitgerust – departement binnen de instelling?

→ *Workflow*

In welke mate is het mogelijk om binnen de instelling een efficiënte workflow te implementeren? Weegt die niet te zwaar door op de dagelijkse werking?

Personeel

Een andere sleutel tot het welslagen van elk digitaliseringproject is een goed personeelsmanagement. Het digitaliseren, beschrijven en het digitaal ontsluiten van de collectie vereisen personeel met de nodige knowhow. Ervaring uit voorgaande projecten is uiteraard welkom. Het is belangrijk dat je de competenties van de medewerkers op een efficiënte manier inzet om niet alleen de productiviteit te maximaliseren maar ook steeds de hoogst mogelijke kwaliteit te halen.

Het samenstellen van een uitgebalanceerd projectteam dat zijn vaardigheden maximaal kan inzetten is niet zo eenvoudig. Het aanwerven van gespecialiseerd en goed opgeleid personeel met de nodige ervaring is een dure investering. Daarom geven we hier een aantal tips.

In eerste instantie moet er worden gekeken naar de vaardigheden van het personeel dat je ter beschikking hebt. Veel van de benodigde expertise kan je dus reeds in eigen huis terugvinden.

Het inzetten van eigen personeel in een digitaliseringproject zou de normale werking van de instelling echter niet in de weg mogen staan. Zeker wanneer digitaliseren slechts een van de nevenactiviteiten is die niet tot de *core business* van de instelling behoort. Je zou kunnen overwegen om voor de duur van het project tijdelijk personeel aan te werven voor gespecialiseerd werk zoals bijvoorbeeld voor het eigenlijke digitaliseren. Een andere mogelijkheid is om een deel van het werk uit te besteden aan een gespecialiseerde firma.

Hoe dan ook is wenselijk dat het projectteam over voldoende expertise beschikt in volgende domeinen:

→ Wetenschappelijk domein

Het opzetten van een beeldbank is interdisciplinair. Historici, informatici, specialisten in conservering, archivariissen of documentalisten, enzovoort kunnen als consulent of adviseur een waardevolle aanwinst zijn voor een beeldbankproject.

→ Juridische ondersteuning

Juridische experts bezitten waardevolle kennis over het portretrecht, auteursrechten, contracten met derden bij outsourcing, enzovoort.

→ Digitale beeldbewerking

Experts met praktische ervaring in domeinen zoals digitale fotografie, scanners, beeldbewerking met softwarepakketten zoals Photoshop.

→ Technische ondersteuning

Netwerk- en systeembeheer, onderhoud van de infrastructuur.

→ Projectcoördinatie.

Infrastructuur

Naast een projectteam met de nodige vaardigheden moet je ook beschikken over de nodige infrastructuur. Dit zijn de werkmiddelen die toelaten om een digitaliseringproject te realiseren. Hieronder verstaan we niet alleen hard- maar ook software. Dit gaat van pc's tot beheerssoftware.

Ook hier is het belangrijk om na te gaan wat reeds voorhanden is, wat eventueel nog moet worden aangeschaft en wat best wordt uitbesteed. Wanneer een instelling over een ICT-departement of -verantwoordelijke beschikt is het aangewezen om deze te betrekken in het project voor de nodige ondersteuning. Het samenstellen van een inventaris met de beschikbare werktuigen is één zaak. Weten of die volstaan om aan de doelstellingen en de eisen van het project te beantwoorden, is een andere zaak.

Een aantal aandachtspunten:

- ↪ De projectmedewerkers beschikken best over de nodige ervaring en scholing om met de hard- en software te kunnen omgaan. Dit is een extra kostenfactor waarmee je rekening moet houden.
- ↪ Het aanschaffen van goedkope, eenvoudig te bedienen apparaten kan dan weer de mogelijkheden beperken om digitale beelden van de juiste kwaliteit te produceren.
- ↪ De nodige technische ondersteuning moet aanwezig zijn: wanneer er een technisch probleem is dan moet daar snel op kunnen worden gereageerd. Wanneer een cruciaal proces in de workflow zoals het digitaliseren, niet of slecht kan worden uitgevoerd door een defecte scanner of een gecrashte computer, dan loopt het digitaliseringproject al snel vertraging op. Als hier niet snel wordt ingegrepen kan de achterstand te groot worden waardoor de collectie misschien slechts gedeeltelijk kan worden gedigitaliseerd.
- ↪ De hard- en software kunnen worden geïntegreerd tot een geheel dat een vlotte workflow toelaat. Digitaliseren met het doel een zo waarheidsgetrouwe kopie te verkrijgen, impliceert dat het aantal stappen tussen de digitaliseringapparatuur (bijvoorbeeld de scanner) en de opslagmedia (cd, magnetische tape, ...), zo kort mogelijk moet zijn.

ICT die je voor een digitaliseringproject hoe dan ook nodig hebt, zijn:

- ↪ Hardware
 - ~ Workstations
Dit zijn de computers waar het projectteam mee aan de slag gaat. De functie van deze machines is zeer uiteenlopend: gaande van beeldbewerking, het digitaliseren (scanners en digitale camera's worden gekoppeld aan een workstation) tot het invoeren van meta-data, controle en toegang tot het digitale depot. Workstations moeten krachtig genoeg zijn om deze taken aan te kunnen. Hedendaagse computers volstaan doorgaans om alle taken uit te voeren.
 - ~ Servers
Dit zijn machines die instaan voor de opslag van digitale beelden, het beschikbaar maken van de digitale beelden op het internet en dergelijke. Servers worden van twee kanten benaderd: enerzijds door het publiek dat de beeldbank on line raadpleegt, anderzijds door het projectteam dat de beeldbank beheert. Het opzetten, onderhouden en operationeel houden van servers is doorgaans een gespecialiseerde en dure aangelegenheid.
 - ~ Digitaliseringapparatuur
Dit zijn scanners en digitale camera's: de apparatuur die nodig is om een analoge afbeelding om te zetten in een digitale vorm.
 - ~ Netwerkinfrastructuur
Dit is de intermediaire infrastructuur die workstations en servers met elkaar laat com-

municeren. Deze bestaat uit netwerkkabels, switches, routers, hubs en wifi-stations. Een grotere organisatie beschikt vaak reeds over een dergelijke infrastructuur zodat extra investeringen hier niet nodig of slechts minimaal zijn.

→ Software

~ Beeldbewerkingsoftware

Hoewel we verder stellen dat de digitale beelden die je wil archiveren slechts minimaal mogen worden gemanipuleerd, is het interessant om over een degelijk beeldbewerking-pakket te beschikken. De bekendste is Adobe Photoshop hoewel er ook alternatieven bestaan zoals Paintshop Pro, XNView of de GIMP. Deze laatste twee zijn vrij verkrijgbaar.¹¹ De standaarden en kwaliteitseisen beschreven in hoofdstuk 3 moeten echter ten allen tijde worden gerespecteerd om een kwalitatief digitaal resultaat te garanderen dat duurzaam kan worden bewaard.

~ Beheerssoftware

Het louter produceren en opslaan van digitale beelden in een *digitaal depot* volstaat niet. De digitale beelden moeten ook nog worden ontsloten. In principe kan je met een eenvoudige databank de digitale collectie beheren op een manier die vergelijkbaar is met de wijze waarop je een analoge fotografische collectie beheert.

Daarnaast bestaat er meer geavanceerde software die men ook wel *Digital Asset Management Systems* noemt. Die vormen een softwareplatform om dit alles binnen één omgeving te kunnen uitvoeren. Het zijn echter professionele softwarepakketten die op een server worden geïnstalleerd.¹²

~ Ondersteunende software

Dit is het geheel van software die ondersteuning biedt en het mogelijk maakt om snel en efficiënt te digitaliseren. Het gaat om besturingssystemen, anti-viruspakketten, office-software, ...

Het belangrijkste is dat de infrastructuur volstaat om digitale beelden van de beoogde kwaliteit te produceren, deze te beheren en ze op een duurzame wijze te bewaren.

Beschikbare fondsen en budgetmanagement

Cruciaal voor het project is het beheer van het budget. Zonder de broodnodige fondsen kan het niet worden gefinancierd. Een uitgebalanceerd en sluitend budgetmanagement staat centraal. Er zijn verschillende redenen om hiermee in de planning rekening te houden:

→ Je kansen op succes verzekeren

Een digitaliseringproject bestaat uit een breed scala aan uitgavenposten gaande van het loon van het projectteam tot het aanschaffen van nieuwe apparatuur. Een verkeerde inschatting van de uitgavenposten kan een zware financiële dobber betekenen. In het slechtste geval is het geld voortijdig op en moet het project misschien zelfs worden afgeblazen.

¹¹ Adobe Photoshop, <http://www.adobe.com/products/photoshop/main.html> (20 mei 2005).

PaintShop Pro, <http://www.corel.com> (20 mei 2005).

GIMP (GNU Image Manipulation Program, <http://www.gimp.org/> (20 mei 2005).

XNView, <http://www.xnview.com/> (20 mei 2005).

¹² Een voorbeeld van een DAMS is het Memorix pakket dat werd ontwikkeld door Pictura Imaginis, <http://www.pictura-dp.nl/> (20 mei 2005).

- Extra fondsen verwerven
Met een uitgebalanceerd financieel plan zullen mogelijke geldschieters sneller over de brug komen met de nodige fondsen.
- Sustainability
Financiële ondersteuning van buitenaf biedt echter geen zekerheden. Vaak beperkt deze financiering zich tot de initiële fase waarbij het project op poten wordt gezet met project-subsidies van overheidswege.

Eenmaal de digitale collectie uitgebouwd en de doelstellingen bereikt, moet er worden nagedacht over het voortbestaan op lange termijn. De mate waarin er een financieel draagvlak bestaat dat het project blijvend ondersteunt noemt men *sustainability*.

Een digitale collectie vraagt een blijvend onderhoud, zowel van de hard- en de software als van de digitale stukken zelf. Door technologische veroudering moeten die op termijn gemigreerd of geconverteerd worden naar nieuwe dragers en formaten.

Wanneer er terug fondsen vrijkomen kan je nieuwe stukken digitaliseren en toevoegen aan de digitale collectie. Of je kan investeren om de mogelijkheden uit te breiden. Of aansluiten bij andere beeldbanken. Op die manier creëer je een meerwaarde en houdt de beeldbank zichzelf in stand.

Centraal in het financiële plan staan de kosten die een instelling denkt te zullen maken. Deze worden omschreven in een *kostenmodel*. Het ontwerpen van een dergelijk model heeft een aantal voordelen.

- Je wordt gedwongen om na te denken over de praktische aanpak: aanschaf van logistieke middelen, samenstellen van een projectteam, workflow, ...
- Op basis van het kostenmodel kan een instelling besluiten het ontwerp – met respect voor de integriteit van de digitale objecten – aan te passen om de financiële haalbaarheid te verbeteren of toch af te zien van het volledige plan.
- Het maken van voorspellingen. Een kostenmodel laat niet alleen toe om een zicht te krijgen op de kosten, maar ook om tot op zekere hoogte op termijn voorspellingen te doen.

Er bestaan tal van modellen die toelaten om de kosten in kaart te brengen. Een interessante aanpak om kosten te berekenen werd ontwikkeld door de Research Libraries Group.¹³ Zij vertrekken vanuit de workflow die wordt gevolgd om de stukken te digitaliseren. Het RLG-kostenmodel onderscheidt hierin 11 stappen. Dit gaat van de selectie tot de nabewerking van de digitale bestanden. Per stap worden de specifieke kosten geïdentificeerd.

De Arts and Humanities Data Service (AHDS) bouwde hierop voort en creëerde een sterk vereenvoudigd model. Ook hier wordt de workflow onderverdeeld in 10 stappen die als een framework dienen waarbinnen de verschillende kosten worden geïdentificeerd.¹⁴

¹³ RLG, RLG worksheet for estimating digital reformatting costs, 1998, <http://www.rlg.org/preserv/RLGWorksheet.pdf> (24 mei 2005).

¹⁴ AHDS, Digitisation. A project planning checklist, 2005, <http://ahds.ac.uk/creating/information-papers/checklist/#appendix> (24 mei 2005).

Voorbeelden van kosten:

- ↪ Selectie
 - ~ Identificatie van de stukken.
 - ~ Inschatten of een stuk in aanmerking komt voor digitaliseren.
- ↪ Digitaliseren
 - ~ Voorbereiding digitaliseren: schoonmaken, uitpakken, enzovoort
 - ~ Vastleggen van de kwaliteitseisen, het instellen van de apparatuur.
 - ~ Het eigenlijke digitaliseren.
 - ~ Kwaliteitscontrole (*quality assurance*).
 - ~ Productie van afgeleide afbeeldingen voor verder gebruik.
 - ~ Post-digitalisering: opnieuw verpakken en terugplaatsen van de originele stukken.
- ↪ Beschrijving
 - ~ Vastleggen werkwijze: naamgeving, ontwikkelen van een datamodel, enzovoort.
 - ~ Het eigenlijke beschrijven.
 - ~ Het aanpassen van bestaande beschrijvingen.
 - ~ Controleren van de beschrijvingen.
- ↪ Beheerskosten
 - ~ Projectmanagement.
 - ~ Het onderhoud van de infrastructuur.
 - ~ Training van het personeel.
 - ~ Garanderen beschikbaarheid digitale collectie.
 - ~ ...

Daarnaast moet je het onderscheid maken tussen de eenmalige en lopende kosten. De eerste spreken voor zich: het aanschaffen van apparatuur is een voorbeeld van een eenmalige kost. Lopende of weerkerende kosten moeten worden gedekt voor de volledige duur van het project. Personeelskosten zijn hiervan een voorbeeld.

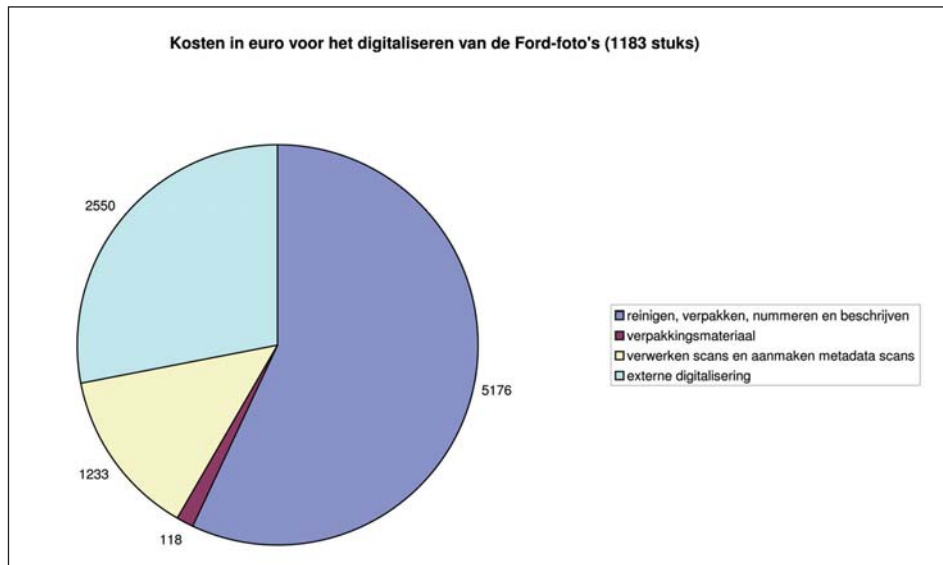
Er moet ook worden opgelet voor zogenaamde verborgen kosten. Die worden vaak niet in een kostenmodel opgenomen. Ze komen echter tot uiting op totaal andere vlakken. Digitaliseren in huis brengt een aantal verborgen kosten met zich mee waarmee soms geen rekening wordt gehouden:

- ↪ Een deel van het personeel dat binnen een project wordt ingezet, heeft binnen de reguliere werking vaak nog andere taken. Het reguliere werk dat omwille van het digitaliseringsproject blijft liggen vormt zo'n verborgen kost.
- ↪ Door gebruik te maken van de infrastructuur die voorhanden is, is deze niet beschikbaar voor de normale werking. Het onderhoud ervan wordt bovendien vaak betaald met reguliere werkingsmiddelen die buiten het projectbudget vallen. Bijvoorbeeld de loonkosten van de systeembeheerder, de kantoorkosten en dergelijke.

Tenslotte moet er een juiste inschatting van de kosten worden gemaakt. Het digitaliseren zelf en het verwerken van de digitale kopieën is niet gratis. Maar het is niet noodzakelijk het duurste aspect van de volledige operatie. Binnen het stadsarchief Antwerpen werd bijvoorbeeld een verzameling van 1 183 foto's gedigitaliseerd. Daarbij werden de verschillende kosten rigoureuus bijgehouden en onderverdeeld in volgende categorieën.

- reinigen, verpakken, nummeren en beschrijven
- aanschaf verpakkingsmateriaal
- verwerken scans en aanmaken metadata
- externe digitalisering.

Daaruit bleek dat de categorie reinigen, verpakken, nummeren en beschrijven op zich reeds 57 % van de totale kosten vertegenwoordigde. Naar het uitbesteden van de eigenlijke digitalisering gaan slechts 28 % van de gemaakte kosten. De verwerking van de scans en het aanmaken de technische metadata nemen met 13,5 % slechts een fractie van de totale kosten in beslag.



Hieruit blijkt duidelijk dat het belangrijk is om te weten wat de precieze impact van elke uitgavenpost is op het plaatje van de totale kosten die tijdens het project worden gemaakt.

Samenwerking

Digitaliseren is geen activiteit die zich tot één instelling, organisatie of project hoeft te beperken. Het is zelfs wenselijk om waar dat mogelijk is samen te werken. Die samenwerking kan verschillende vormen aan nemen: het louter uitwisselen van expertise, een gezamenlijke infrastructuur opbouwen, fondswerving etcetera. Het kan gaan om een intensieve of om een eerder losse samenwerking.

Samenwerking kan ook helpen om fondsen te verkrijgen. Een particuliere sponsor of de overheid zal een samenwerkingverband tussen verschillende instellingen om hun collecties onder één digitaal dak te brengen sneller steunen dan losse projecten.

Het NINCH-onderzoek maakt het onderscheid tussen 4 soorten samenwerkingsverbanden:¹⁵

- Consultancy
Dit is het inbrengen van expertise van buitenaf in een project. Dit kan interessant zijn wanneer je zelf niet over de nodige bekwaamheid beschikt. Je kan ervoor opteren om een extra

¹⁵ NINCH, *The NINCH Guide to Good Practice*, 2002, <http://www.nyu.edu/its/humanities/ninchguide/IX/> (22 maart 2005).

consulent aan te werven of beroep te doen op een extern consultancybureau. Dit is echter een dure oplossing.

↪ Joint projects

Dit is de meest voorkomende vorm van samenwerking. Instellingen en organisaties vormen partnerships om een gemeenschappelijk doel te bereiken: het digitaliseren van hun collecties en het ontsluiten van de data via een gemeenschappelijk platform. Dergelijke verbanden beperken zich niet alleen tot het niveau van de ontsluiting. Ook tijdens het digitaliseren, beschrijven of projectmanagement bestaat er een wisselwerking.

Het EVA-project dat we in het derde deel bespreken is zo'n *joint project* tussen verschillende instellingen. Het stadsarchief Antwerpen en de London Metropolitan Archives zijn de archiefinstellingen die aan dit project deelnamen.¹⁶

↪ Outsourcing

Het uitbesteden van een aantal processen uit de workflow aan professionele, gespecialiseerde bedrijven is een beproefde strategie. Processen zoals digitaliseren, het beschrijven of het online zetten, worden in onderaanneming gegeven.

Hoewel outsourcing een te overwegen optie is, is ze niet risicoloos. De professionele sector beschikt weliswaar over een groot potentieel onder de vorm van infrastructuur en expertise, maar houdt niet altijd rekening met de belangen van de erfgoedinstellingen. Deze bedrijven hebben vaak een achtergrond binnen de kantoorsector (bijvoorbeeld prepress of documentmanagement). Ze digitaliseren niet altijd in open formaten, volgen de standaarden soms niet, manipuleren soms de beelden en dergelijke meer. Sommige bieden kant-en-klare systemen aan die omwille van commerciële belangen niet altijd even compatibel zijn met andere gelijkaardige systemen. Dit kan een hinderpaal zijn.

Om dit te vermijden moet je als instelling duidelijke afspraken maken met het bedrijf over de uitvoering van het proces. Er moet een contract of een offerte worden opgemaakt waarin duidelijk vermeld staat hoe het werk zal verlopen en wat de kwaliteitseisen zijn.

↪ Andere

Alle combinaties van de drie vorige mogelijkheden kunnen ook.

Workflow

De praktische uitwerking van een workflow wordt besproken onder paragraaf 2.4.

¹⁶ <http://www.eva-eu.org> (22 maart 2005).

1.6 De voor- en haalbaarheidsstudie

Zeker bij omvangrijke digitaliseringprojecten is het belangrijk om alle aspecten die een impact hebben op het behalen van de doelstellingen in acht te nemen vooraleer over te gaan tot de effectieve uitvoering van het plan. Het kan interessant zijn om een studie of *assessment* uit te voeren waarin alle nuttige informatie wordt gebundeld.

Deze studie zal dan het uitgangspunt vormen voor de verdere planning en aanpak van het project. Deze voorstudie kan soms tot de conclusie leiden dat het niet is aangewezen om te digitaliseren, bijvoorbeeld omdat de aard van de collectie een efficiënte digitalisering onmogelijk maakt of omdat het project financieel niet haalbaar is.

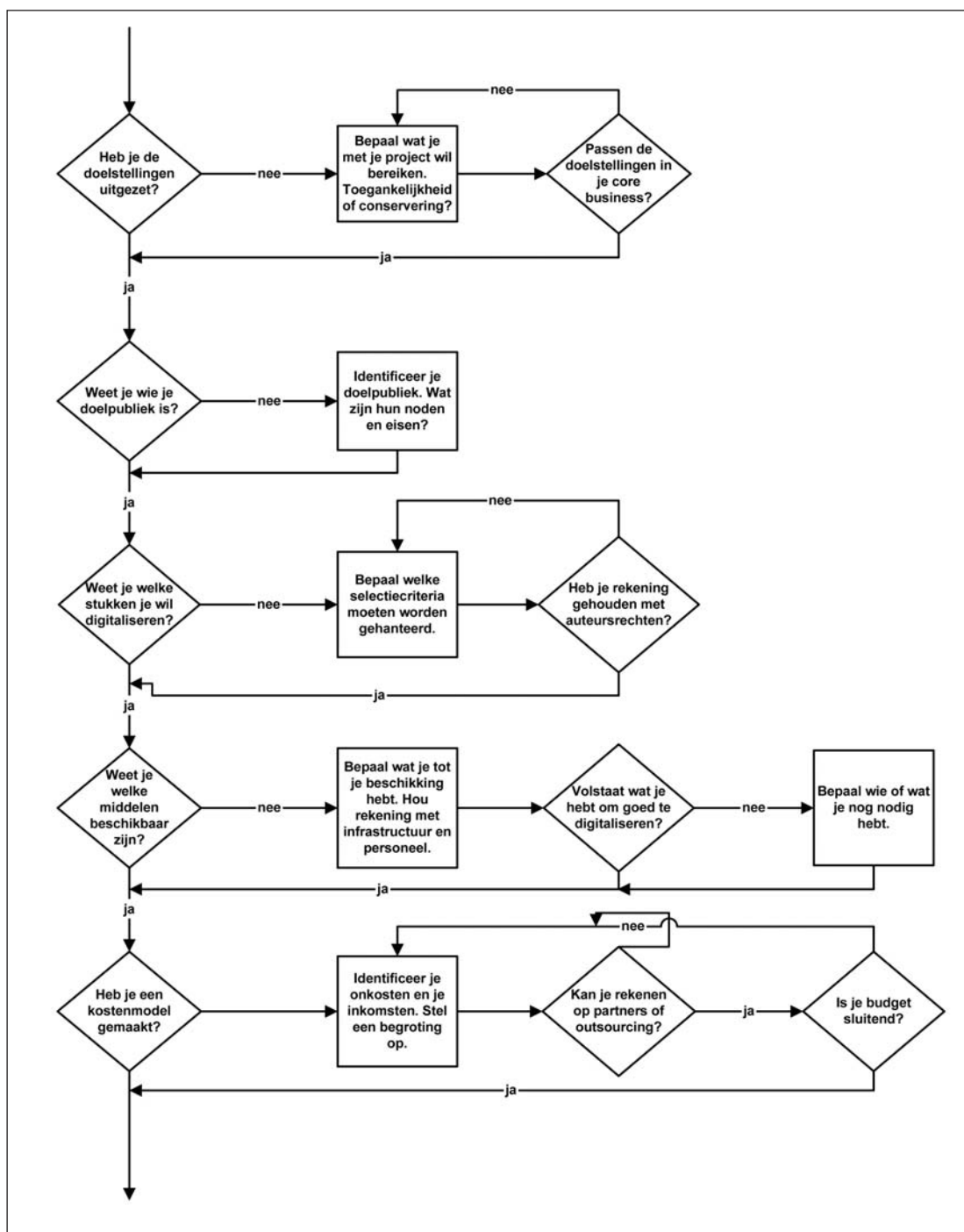
cDAVID aanbeveling: de voor- en haalbaarheidsstudie

Hieronder geven we een voorstel van de verschillende elementen die in een dergelijke voorstudie aan bod kunnen komen.

- ↪ Een abstract of korte en duidelijke omschrijving van het project.
- ↪ De doelstellingen en prioriteiten:
 - ~ wensen en noden van de instelling die een collectie beheert
 - ~ wensen en noden van het doelpubliek.
- ↪ Een beschrijving van de collectie die je wil digitaliseren (hoeveelheid, aard, ...).
- ↪ Een summiere uiteenzetting van de aanpak van het project en de te verwachten resultaten.
- ↪ Een overzicht van de beschikbare en benodigde middelen:
 - ~ personeel
 - ~ infrastructuur
 - ~ noodzakelijke expertise.
- ↪ De beschrijving van de kosten in een kostenmodel of begroting.
- ↪ Conclusie.

Daarnaast is het ook aangewezen om een decision-making matrix op te stellen. Dit is een flowchart of checklist die het projectteam moet doorlopen in de voorbereidende fase. Alle praktische overwegingen die hier aan bod kwamen, worden in deze flowchart opgenomen. Pas wanneer alle stappen in de flowchart werden doorlopen, kan je beginnen met de eigenlijke planning en de realisatie van het project

Hazen, Horrell en Merill-Oldham ontwikkelden een bekende model voor een decision-making matrix. Deze checklist zet de te digitaliseren stukken centraal. Pas wanneer blijkt dat het projectteam de stukken op een goede manier kan digitaliseren, duurzaam digitaal bewaren, digitaal kan ontsluiten en ter beschikking stellen van het publiek, kan van start worden gegaan met de eigenlijke realisatie.¹⁷



Voorbeeld descision making matrix, ontwikkeld door cDAVID, 2005.

17 D. HAZEN, J HORRELL en J. MERRILL-OLDHAM, *Selecting research collections for digitisation*, CLIR, 1998, <http://www.clir.org/pubs/reports/hazen/pub74.html> (21 maart 2005).



Groendienst stad Antwerpen: kwekerij, 1924

(stadsarchief Antwerpen, FOTO 1582).

2 Het digitaliseringproces

Wanneer documenten over een langere periode worden bewaard, dan krijgt de archivaris of erfgoedbeheerder vroeg of laat af te rekenen met het verval van de drager. Er zijn verschillende oorzaken voor dit verval: degradatie van het materiaal, slijtage van de drager, slechte bewaaromstandigheden en andere. In het slechtste geval gaat de informatie onherroepelijk verloren. Fotocollecties zijn hiervan een goed voorbeeld. De degradatie van de fotografische drager leidt er immers toe dat het beeld mee verloren gaat.

Er bestaan verschillende bewaarstrategieën die het verval weliswaar sterk kunnen vertragen maar niet helemaal stoppen. Je kan een kopie maken van de foto. Analoog kopiëren van foto's betekent dat het beeld op een nieuwe analoge drager wordt gezet. Bij deze oplossing worden we al snel geconfronteerd met *generatieverlies*. Fijne details gaan verloren en het contrast en de kleurtinten kunnen verschillen. Daardoor zal een analoge kopie nooit de kwaliteit van het origineel evenaren. Bovendien is ook de analoge kopie onderhevig aan verval zodat je ook deze op termijn moet vervangen.

Een derde mogelijkheid is digitaliseren. Dit is het digitaal kopiëren van een analoge foto of document. De digitale kopie is in wezen een *bitstream* of een reeks van binaire data. Een digitale kopie brengt nieuwe mogelijkheden met zich mee:

- Een digitaal bestand kan je onbeperkt, snel en relatief eenvoudig reproduceren. Het louter kopiëren van de binaire data heeft geen invloed op de kwaliteit van het digitale beeld.
- Qua omvang neemt een digitale kopie veel minder ruimte in dan een traditioneel analoog stuk.

Digitaliseren kent ook heel wat valkuilen. De twee voornaamste problemen die we in dit hoofdstuk zullen behandelen zijn *kwaliteit* en *leesbaarheid*.

Kwaliteit speelt een centrale rol. Het komt er op aan om eisen te stellen aan de kwaliteit van het digitale beeld en de apparatuur waarmee je digitaliseert. De benadering van de kwaliteit bestaat uit zowel een subjectieve als objectieve kant.¹⁸ Enerzijds probeert de archivaris of erfgoedbeheerder naar best vermogen zijn eigen oordeel uit te spreken over de kwaliteit. Belangrijk is dat je daarvoor over de nodige ervaring beschikt. Anderzijds laat de objectieve benadering toe aan de hand van kwantificeerbare parameters het digitale beeld en de apparatuur te toetsen aan eenduidige kwaliteitsnormen.

Bij het bewaren van digitale gegevens vormt de leesbaarheid op lange termijn een belangrijk probleem. Je bent afhankelijk van apparatuur om digitale documenten te raadplegen. ICT evolueert constant waardoor hard- en software zeer snel verouderen en in onbruik geraken. De leesbaarheid van een digitaal document doorheen de tijd is daarom allesbehalve gegarandeerd. In een tweede paragraaf gaan we dieper in op deze problematiek.¹⁹

¹⁸ F. FREY en J. REILLY, *Digital imaging for photographic collections. Foundations for technical standards*, Image Permanence Institute, Rochester Institute of Technology, Rochester, NY, 1999, p. 10, http://www.rit.edu/%7E661www1/sub_pages/digibook.pdf (2 juni 2005).

¹⁹ F. BOUDREZ, *Standaarden voor digitale archiefdocumenten*, stadsarchief Antwerpen, Antwerpen, 2004.

Ten slotte bekijken we de *workflow* van het digitaliseringproces. Welke stappen dien je te nemen en wat zijn de aandachtspunten bij elke stap? Een goed uitgewerkte workflow garandeert dat het digitaliseren efficiënt zal verlopen en het digitale eindresultaat bruikbaar zal zijn.

2.1 Het belang van kwaliteitseisen

Kwaliteit staat centraal tijdens het digitaliseren. Een digitaal stuk moet enerzijds bruikbaar zijn om de doelstellingen van het project te realiseren. Anderzijds moet ze ook een getrouwe weergave van het origineel zijn.

Tijd en geld zijn de belangrijkste redenen om tijdens het digitaliseren oog te hebben voor de kwaliteit. De volledige selectie opnieuw digitaliseren kost extra tijd en geld. Wanneer blijkt dat de digitale stukken niet voldoen is een dergelijke investering vaak niet meer mogelijk. Bovendien is er geen tweede kans wanneer de stukken achteraf worden vernietigd en je enkel de digitale kopie wil bewaren. Door de kwaliteit van een digitaal stuk reeds tijdens het digitaliseren zelf te controleren, kan je op tijd optreden zodat het stuk geen tweede keer onder handen moet worden genomen.

De *doelstellingen* van een digitaliseringproject zijn een tweede reden om oog te hebben voor de kwaliteit van de digitale beelden. De eisen die je stelt aan een digitaal beeld zijn anders wanneer je het beeld enkel op een website plaatst, dan wanneer je het digitale beeld wil gebruiken in een gedrukte publicatie of duurzaam bewaren op lange termijn. Het heeft geen zin om een collectie te scannen met kwaliteitseisen die niet stroken met de uiteindelijke doelstellingen.

Een digitale kopie kan in verschillende vormen verschijnen. Afhankelijk van de doelstellingen en de daaruit volgende kwaliteitseisen.

→ De archiefkopie

Dit is de digitale kopie die we op lange termijn wensen te bewaren. Door de technologische veroudering is dit niet vanzelfsprekend. Wie kan er garanderen dat we over 20 jaar het digitale beeld kunnen bekijken met de technologie die dan voorhanden is? Leesbaarheid vormt hier een speciaal aandachtspunt.

Wanneer het origineel wordt vernietigd is de digitale kopie het enige wat nog rest. Er is geen tweede kans om het stuk te digitaliseren. Naar raadpleging toe, wordt de digitale archiefkopie ook gebruikt om het origineel te vervangen. Dit betekent dat de archiefkopie het origineel zo goed mogelijk moet proberen te benaderen.

De integriteit van de archiefkopie moet worden gegarandeerd. Door technieken zoals bijvoorbeeld compressie toe te passen kan informatie onherroepelijk verloren gaan. Dit moet worden vermeden. Een mogelijke oplossing is het toepassen van de DAVID-bewaarstrategie die de originele bitstream zo lang mogelijk probeert te vrijwaren.²⁰

Naast de archiefkopie is er ook de *veiligheidskopie*. Dit is een exact duplicaat van de archiefkopie. De veiligheidskopie wordt bij voorkeur op een andere locatie bewaard. Ze

²⁰ F. BOUDREZ en H. DEKEYSER, DAVID. *Digitaal archiveren in de praktijk. Handboek.*, Stadsarchief Antwerpen, ICRI, Antwerpen-Leuven, 2003, <http://www.antwerpen.be/david> (31 mei 2005).

dient als back-up wanneer de eigenlijke archiefkopie is beschadigd of dreigt verloren te gaan.

Een archiefkopie wordt ook wel een *digitale master* genoemd.

→ *De afgeleide kopieën*

Hoewel de archiefkopie geschikt is voor bewaring op lange termijn is ze niet in alle situaties bruikbaar. Vaak gaat het om digitale bestanden die een grote opslagcapaciteit vragen. Ze zijn evenmin geschikt om bijvoorbeeld te verspreiden via het internet. Bovendien is het door de auteursrechtelijke bescherming van de beelden vaak niet wenselijk om de archiefkopie vrij ter beschikking te stellen. Afhankelijk van de doelstellingen is het beter om een afgeleide kopie te maken op basis van de archiefkopie.

~ *De raadplegingskopie*

Raadplegingskopieën zijn digitale beelden die snel en gebruiksvriendelijk via het internet kunnen worden verspreid, bijvoorbeeld via een online beeldbank. Dit zijn digitale beelden die door compressie, aanpassen van de resolutie, een voor browsers geschikt bestandsformaat, enzovoort werden aangepast voor online raadpleging.

Auteursrechtelijk beschermd materiaal kan je eventueel watermerken wat het opnieuw gebruiken van het materiaal buiten de beeldbank bemoeilijkt.

Door het toepassen van technieken zoals compressie gaat er wel beeldinformatie verloren. Daarom is een raadplegingskopie zeker niet geschikt voor archivering.

~ *De reproductiekopie*

Dit is de kopie die we aanmaken voor kwalitatieve reproducties voor de meest uiteenlopende doeleinden, bijvoorbeeld drukwerk, gebruik in media zoals televisie of krant, de kunstensector, enzovoort.

In de regel zijn deze niet als dusdanig beschikbaar maar moeten ze worden besteld. Ze worden maar aangemaakt op vraag en volgens de wensen van de opdrachtgever.

Aangezien reproductiekopieën bruikbaar moeten zijn in een specifieke context worden ze vaak gemanipuleerd. Dit maakt ze ongeschikt voor archivering op lange termijn. De reproductiekopie is evenmin geschikt voor vrije (online) verspreiding. Zeker wanneer het beeld auteursrechtelijk beschermd is.

~ *De geoptimaliseerde kopie*

Het origineel kan door verval in zeer slechte staat verkeren. Met moderne technologie is het mogelijk om het beeld digitaal te restaureren. Daartoe wordt het origineel gedigitaliseerd waarna met een beeldbewerkingsprogramma beschadigingen en verval digitaal kunnen worden hersteld.

De digitale, geoptimaliseerde kopie wordt parallel met de archiefkopie bewaard. Toch vervangt ze de archiefkopie niet. Ze kan wel dienst doen als basis voor reproductie- of raadplegingskopieën.

Bij het bewaren van de geoptimaliseerde kopie op lange termijn gelden dezelfde voorwaarden als voor een archiefkopie: maak geen gebruik van technologie zoals compressie of gesloten bestandsformaten die de leesbaarheid en de kwaliteit van het digitale beeld kunnen aantasten.

Maar wanneer is een beeld nu van goede kwaliteit? Doorgaans zal een leek dit in algemene termen zoals "levensecht" of "trouw aan het origineel" proberen te omschrijven. Een dergelijke subjectieve benadering van kwaliteit staat open voor interpretatie. Er is nood aan een geheel van meetbare parameters, referenties en methoden waarmee je op een objectieve manier de kwaliteit van een afbeelding kan gaan bepalen. Zo geeft D'Amato een overzicht van verschillende elementen die kwaliteit van een digitaal beeld kunnen beïnvloeden.²¹

- De staat en de kwaliteit van het origineel.
- De wijze van digitaliseren (scanner, digitale camera of andere) en de instellingen van de apparatuur (resolutie, bitdiepte, kleur, enzovoort).
- Het toepassen van compressie en digitale beeldbewerking om het digitale resultaat bruikbaar te maken voor de doelstellingen van het project. Bijvoorbeeld: het digitaal vergroten of verkleinen van een beeld.
- De hard- en software die worden gebruikt om digitale beelden te bekijken.
- De kwaliteit van apparatuur om digitale beelden te bekijken zoals beeldschermen en printers.
- De omgeving waarin de kwaliteitscontrole verloopt. Bijvoorbeeld: lichtinval op de monitor. De aard van het licht (een gloeilamp verschilt van een neonlamp).

Naast factoren en elementen die de kwaliteit van het beeld direct bepalen, zijn er ook aspecten die onrechtstreeks de bruikbaarheid en de leesbaarheid van een digitaal beeld beïnvloeden.

→ *De metadata*

Metadata is beschrijvende informatie die de inhoud van het beeld verduidelijken. Bijvoorbeeld: de naam van de fotograaf, een beschrijving van het beeld, de aard van de drager, het inventarisnummer, trefwoorden, enzovoort. We onderscheiden drie types: administratieve, technische en inhoudelijke metadata.

Metadata *contextualiseren* het beeld. Ze geven het beeld een plaats in de verzameling maar ze verduidelijken het ook. De mate waarin een beeld gecontextualiseerd is bepaalt ook de *terugvindbaarheid*. Een stuk dat goed beschreven is, zal gemakkelijker terug te vinden zijn dan een stuk dat niet beschreven is. Hetzelfde geldt ook voor digitale stukken.

Ook hiervoor moet er een goede beschrijving voorhanden zijn, willen we ze snel en gemakkelijk kunnen terugvinden in een groeiende digitale collectie.

→ *De keuze van het bestandsformaat*

Een bestandsformaat geeft de digitale output die een scanner of een digitale camera genereert een vorm die herkenbaar is voor een computer. Er bestaan heel wat bestandsformaten maar in een gegeven situatie is niet elk bestandsformaat geschikt.

De keuze van een ongeschikt bestandsformaat voor productie van de archiefkopie kan tot een nadelig informatieverlies leiden.

²¹ D. D'AMATO, *Imaging Systems. The range of factors affecting image quality*, in: Guides to quality in visual resource imaging, Washington, DC, Council on Library and Information Resources, 2000, <http://www.rlg.org/visguides/visguide3.html> (22 maart 2005).

→ *De keuze van de drager*

Er moet een geschikte drager worden gekozen voor de lange termijn bewaring van de archiefkopie. Het heeft weinig zin om te digitaliseren met de beste kwaliteit zonder gebruik te maken van een duurzame drager.

Hoe wordt kwaliteit gegarandeerd? Dit gebeurt op verschillende tijdstippen in het project. Tijdens de planning worden de kwaliteitsnormen vastgelegd. Vervolgens wordt de apparatuur geijkt en ingesteld zodat de digitale output voldoet aan die normen. Ten slotte worden de digitale beelden hieraan getoetst vlak na de productie.

In dit hoofdstuk gaan we op zoek naar bruikbare parameters die invloed hebben op de kwaliteit van een digitale stuk.

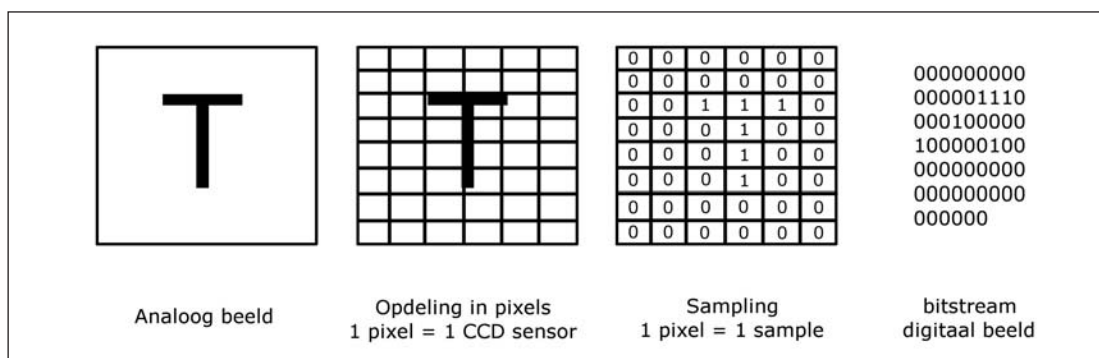
2.2 Technische begrippen

Het digitaliseren van foto's, prenten, tekst, ... is een technisch proces. De spreekwoordelijke druk op de knop die toelaat om op een gebruiksvriendelijke manier documenten te digitaliseren, verbijst dit alles. Ondanks het gebruiksgemak is digitaliseren niet zonder de risico's. Kennis van zaken beperkt de kans dat de doelstellingen niet worden gehaald en het project faalt. Daarom reiken we in dit hoofdstuk een begrippenkader aan waarin we een aantal belangrijke concepten zoals bitdiepte en resolutie verduidelijken.

2.2.1 De werking van scanners en digitale camera's

Fundamenteel zijn er twee soorten apparatuur die je kan aanwenden om een beeld te digitaliseren: de *scanner* en de *digitale camera*. Hoewel deze twee apparaten op het eerste zicht weinig gemeen lijken te hebben, werken ze allebei volgens hetzelfde basisprincipe.

Het hart van deze toestellen wordt gevormd door een chip met lichtgevoelige sensoren. Deze sensoren reageren door een elektrische signaal te geven wanneer ze belicht worden (Foton / Elektron conversie). De sterkte van dit elektrische signaal (Elektron / voltage conversie) wordt gemeten en vervolgens door een analoge / digitale (A/D) convertor omgezet in een digitale waarde. Een digitaal beeld is uiteindelijk niets meer dan een wiskundige matrix opgebouwd met deze digitale waarden. Men noemt deze waarden ook wel *pixels* of *picture elements*.



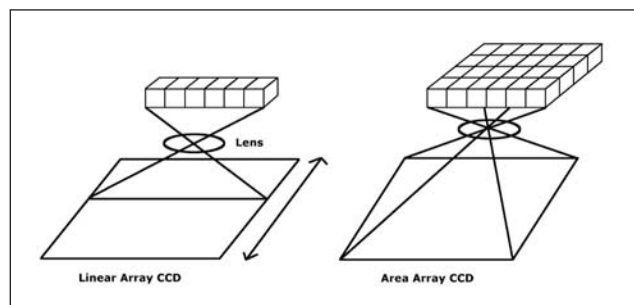
Het samplingproces: van analoog beeld naar digitale bitstream.

De (digitale) waarde gegenereerd door één fotosensor noemt men een *sample*. De output van de apparatuur is niets meer dan reeks of een *bitstream* van samples. Het proces dat de eigenlijke digitale bitstream creëert noemt men ook wel *sampling*. Een computer kan deze bitstream interpreteren en vervolgens het beeld gaan reconstrueren op een computerscherm of via een printerafdruk.²²

Er bestaan verschillende soorten lichtgevoelige chips. De belangrijkste zijn de CCD of *Charged Coupled Device*, de CMOS of *Complementary Metal Oxide Semiconductor* en de Foveon die nog in zijn kinderschoenen staat. Het volstaat te onthouden dat een CCD toelaat om beelden van betere kwaliteit te maken dan een CMOS. Door het prijskaartje vinden we toestellen uitgerust met een CCD-chip eerder in de professionele wereld terug.

Een tweede onderscheid is er tussen *linear array* en *area array* chips. Bij een linear array bestaat het oppervlak van de chip uit een rij sensoren. Een beeld of een document wordt opgedeeld in lijnen. Wanneer de chip een lijn heeft gesampled (omgezet in digitale waarden), wordt de positie van de chip ten opzichte van het te scannen stuk veranderd zodat de volgende lijn kan worden gesampled. Het aantal lijnen dat werd gescand, komt overeen met het aantal rijen pixels waaruit een digitaal beeld bestaat. We noemen dit *progressief scannen*. Het oppervlak van een area array chip bestaat uit een vierkant dat werd opgevuld met lichtsensoren. In tegenstelling tot een linear array chip wordt het beeld niet lijn per lijn afgetast, maar in een keer waargenomen en gesampled.²³

Linear array chips vinden we hoofdzakelijk terug in scanners. Area Array chips worden gebruikt in digitale fototoestellen. Sommige professionele fototoestellen zijn uitgerust met een linear array chip.



Linear array (links), versus area array (rechts).

Een laatste element is de kleur. De lichtsensoren op zich kunnen geen kleuren onderscheiden. Door lichtfilters te gebruiken kan het toestel toch een kleurenbeeld produceren door drie afzonderlijke beelden, respectievelijk gefilterd op rood, blauw en groen licht, met elkaar te combineren.

Een veel toegepaste techniek bij digitale fototoestellen is het gebruik van een *bayer filter*. Daarbij wordt *interpolatie* toegepast. De oorspronkelijke kleur van een pixel wordt dan gereconstrueerd op basis van de waarden van zijn burens. Bij de eerste generaties die deze technologie toepasten

²² We gaan dieper in op het belang van pixels in paragraaf 3.2.1.

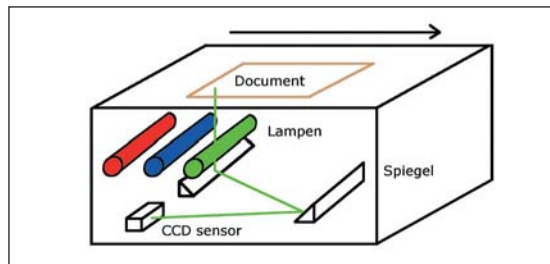
²³ D. D'AMATO en R. KLOPFENSTEIN, *Requirements and options for the digitisation of the illustration collections of the National Museum of Natural History*, 1996, <http://www.nmnh.si.edu/cris/techrpts/imagopts/section4.html> (22 maart 2005).

TASI (Technical Advisory for Images), *TASI Advice Paper. Digital Camera's*, 2003, <http://www.tasi.ac.uk/advice/creating/camera.html> (22 maart 2005).

leidde dat soms tot kleurafwijkingen en *artefacten* of elementen die niet eigen zijn aan het oorspronkelijke beeld.²⁴

Scanners

Het type scanner dat men tegenwoordig aantreft in zowel particuliere als professionele toepassingen is de *flatbed*- of *desktop*-scanner. Deze scanner bestaat uit een transparant oppervlak waarop het te scannen document wordt geplaatst. De scanner zelf bestaat uit verschillende onderdelen: de voornaamste zijn een door een elektromotor aangedreven bewegende arm, één of meerdere lampen, een systeem van lenzen en spiegels en één of meerdere linear array chips.



Werking van de flatbed scanner.

Een eerste type scanner bestaat uit drie verschillende lampen (rood / blauw / groen) die elk afzonderlijk het stuk belichten. Er is slechts één lichtgevoelige sensor en die wordt sequentieel belicht met een rode, groene en blauwe lichtbundel. De drie resulterende kleurbeelden of -kanalen worden met elkaar gecombineerd tot één beeld.

Een andere courante implementatie voorziet drie linear array sensoren en slechts één mobiele lamp. Elke sensor is voorzien van respectievelijk een blauwe, rode of groene kleurenfilter. Ook dit resulteert in drie aparte kleurkanalen die worden gecombineerd tot een kleurenbeeld.

Digitale camera - bookscanners

Een digitale camera is sterk gebaseerd op analoge fotografie. In plaats van een klassieke film bestaat het hart van een digitale camera uit een lichtgevoelige chip. Op het moment dat de sluiterknop wordt ingedrukt, opent het diafragma waardoor licht de camera in kan via de lens en de chip gedurende een kort tijdsinterval wordt belicht. Net zoals bij een analoge camera bepaalt de sluitertijd hoe lang de chip wordt belicht en de grootte van het diafragma hoeveel licht er via de lens binnenkomt.

Net zoals een analoge camera kan een digitale in- en uitzoomen. Een digitaal toestel biedt niet alleen *optische zoom* aan maar ook *digitale zoom*. Wanneer je optisch zoomt wordt de afstand van de lens veranderd ten opzichte van de lichtgevoelige chip. Daardoor verandert de brandpuntsafstand en wordt het beeld vergroot of verkleind. Met digitale zoom kan je nog verder inzoomen op het beeld zonder gebruik te maken van het lenzensysteem. Daarbij past men *interpolatie* toe. Het nadeel is dat er een sterk kwaliteitsverlies optreedt.²⁵

Net zoals bij een scanner is een camera uitgerust met een lichtgevoelige chip die geen onderscheid kan maken tussen kleuren. Ook hier moeten kleurenfilters worden gebruikt om een kleurenopname te kunnen maken. De meest courante is de Bayer-filter.

²⁴ TASI, *digital Camera's*, 2003, <http://www.tasi.ac.uk/advice/creating/camera.html> (22 maart 2005).

M. BROWN, *CCD Bayer masks*, 1999, <http://www.nezumi.demon.co.uk/photo/bayer/bayer.htm> (21 maart 2005).

²⁵ Wikipedia; *digital zoom*, http://en.wikipedia.org/wiki/Digital_zoom (25 mei 2005).

De *bookscanner* of boekenscanner verdient hier een aparte vermelding. De boekenscanner bestaat uit een speciaal ontworpen werkblad waarop het werk wordt geplaatst. Met een vaste gemonteerde camera wordt het werk gedigitaliseerd. De gecontroleerde omgeving laat toe om de resolutie en afmetingen van het scanwerk exact te bepalen, dit in tegenstelling tot bij een losse camera. Een boekenscanner bevat ook een eigen belichtingssysteem zodat het werk onder een optimale belichting kan worden gedigitaliseerd.

Dit apparaat is specifiek ontworpen om op een goede manier ingebonden werken en bijzondere formaten te digitaliseren. Gebonden werken zijn moeilijk te digitaliseren. Het gebruik van een klassieke flatbed-scanner is hiervoor niet geschikt.

cDAVID aanbeveling: wanneer gebruik ik een scanner / digitale camera?

Gebruik een scanner wanneer er geen risico op beschadiging van de originele stukken bestaat.

Voorbeelden van stukken die beter worden gefotografeerd: ingebonden documenten, rollen, fragiele of reeds beschadigde stukken.

Een stuk kan ook te groot zijn voor een gewone scanner. Bijvoorbeeld kaarten of posters. Gebruik in dat geval een digitale camera.

Ingebonden stukken worden bij voorkeur gedigitaliseerd met een boekenscanner.

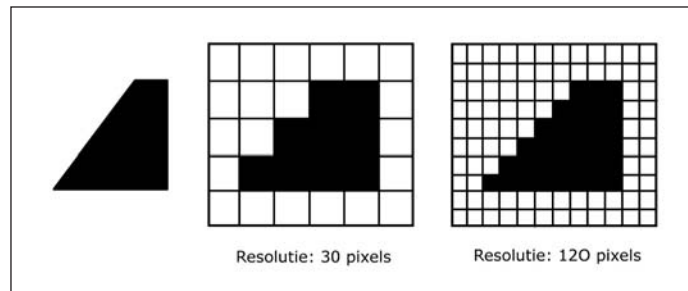
2.2.2 Resolutie, bitdiepte en dynamisch bereik

De drie eenvoudigste parameters om de kwaliteit van een digitaal beeld of een apparaat in te schatten zijn de *resolutie* of het *oplossend vermogen*, de *bitdiepte* en het *dynamische bereik*.

Resolutie

De resolutie of het oplossend vermogen geeft aan in welke mate de apparatuur in staat is om kleine details weer te geven. De resolutie van een digitaal beeld is de mate waarin je de deze details kan onderscheiden. Een beeld met een hoge resolutie zal dus veel kleine details bevatten. Een apparaat met een hoge resolutie kan veel kleinere details waarnemen dan een toestel met een lage resolutie. De maximale resolutie van een digitale camera of een scanner heeft een rechtstreeks verband met het aantal fotosensoren dat een chip bevat. Hoe meer fotosensoren op eenzelfde oppervlakte, hoe groter de resolutie van de scanner of de camera en hoe beter het beeld.

De basiseenheid waaruit een digitaal beeld is opgebouwd, is de *pixel* of picture-element. Elke lichtsensor op een chip neemt slechts een deeltje van het volledige beeld waar. Dit beeldelement noemt men een *pixel*. Per pixel komt één numerieke *pixelwaarde* of *sample* overeen. Dit is de digitale waarde van de door de lichtsensor gemeten lichtintensiteit voor die pixel. Kwaliteit hangt rechtstreeks af van het aantal pixels op eenzelfde oppervlakte. Hoe meer pixels, des te beter de kwaliteit. Onderstaande figuur illustreert dit.



Resolutie: het aantal beeldelementen waaruit een digitaal beeld is opgebouwd.

Bij digitale camera's en scanners kan je de resolutie ook lager instellen. Het digitale beeld zal daarvoor uit een lager aantal pixels bestaan wat een vermindering van de kwaliteit betekent.

De resolutie heeft ook een impact op de lengte van de bitstream die de digitale afbeelding representeert. De bitstream van een digitaal beeld dat uit 120 pixels bestaat zal groter zijn dan van een beeld dat uit slechts 30 pixels bestaat. Dit heeft gevolgen voor de verwerkingstijd van het digitale beeld en de benodigde opslagcapaciteit.²⁶

De maximale resolutie van een apparaat hangt samen met het aantal lichtsensoren dat de chip bevat. Zo kan men strikt genomen geen digitaal beeld maken met een resolutie van 200 x 200 pixels wanneer de maximale resolutie van de apparatuur slechts 100 x 100 pixels is. In de praktijk komt het echter vaak voor dat digitale afbeeldingen toch een grotere resolutie hebben dan de apparatuur waarmee ze gemaakt zijn eigenlijk aankan. Om dit te bereiken wordt er gebruik gemaakt van een *interpolatie*.

Met deze techniek worden er digitale artificiële pixels aan het beeld toegevoegd. De waarden voor deze pixels werden niet werkelijke waargenomen door de lichtsensoren op de chip maar werden afgeleid uit de waarden van de omliggende pixels. Geïnterpoleerde waarden zijn slechts een benadering van de werkelijke waarde. Hoe meer geïnterpoleerde pixels een digitaal beeld bevat, hoe slechter de kwaliteit van het beeld. Dit is een belangrijk element waarop je moet letten bij de aanschaf van apparatuur. Fabrikanten vermelden om marketingredenen graag de resolutie na interpolatie. De reële resolutie van een apparaat is vaak een stuk lager.

Interpolatie wordt ook toegepast wanneer digitale beelden in een beeldbewerkingprogramma zoals Photoshop worden geschaald. Door een beeld na het digitaliseren softwarematig groter of kleiner te maken, gaat er oorspronkelijke informatie verloren en wordt informatie toegevoegd die niet direct werd geproduceerd tijdens het digitaliseren. Bij het aanmaken van een archiefkopie is het af te raden om softwarematig de grootte van een digitaal beeld te manipuleren.

Bitdiepte

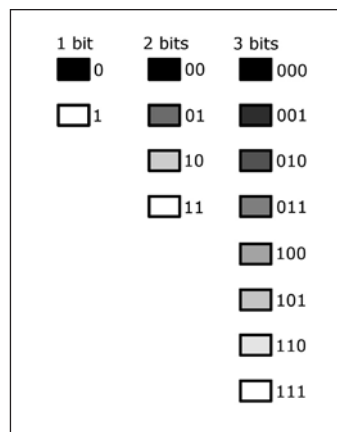
Volgens Frey is een correcte weergave van de originele kleuren zowat het belangrijkste aspect waaraan een digitaal beeld moet voldoen. Dit komt neer op een nauwkeurige reproductie van de originele *tonal range* of het aanwezige spectrum van kleurtonen.²⁷

²⁶ A. KENNEY, O. RIEGER en R. ENTLICH, *Moving theory into practice. Digital imaging tutorial.*, Ithaca, NY, Cornell University Library, 2003, <http://www.library.cornell.edu/preservation/tutorial/intro/intro-06.html> (22 maart 2005)

²⁷ F. FREY, *measuring image quality*, in: *Guides to quality in visual resource imaging*, Washington, DC: Council on Library and Information Resources, 2000, <http://www.rlg.org/visguides/visguide4.html> (22 maart 2005).

Bij het aanmaken van een digitaal beeld, meet elke fotosensor de lichtintensiteit en zet die om in een pixelwaarde. Deze pixelwaarde wordt uitgedrukt door één of een combinatie van meerdere bits. Het aantal gebruikte bits om de waarde binair te kunnen voorstellen noemt men de *bitdiepte*.

Afhankelijk van de bitdiepte kan men tot 16 miljoen unieke kleurtonen voorstellen. Stel dat de bitdiepte van een pixel 1 bit is, dan kan deze pixel slechts 2 waarden aannemen: 1 of 0. De pixel kan daardoor slechts 2 verschillende tonen representeren: zwart of wit. Indien de bitdiepte van een pixel uit 2 bits bestaat, dan kan de pixel 4 verschillende tonen binair voorstellen. Een bitdiepte van 4 bits laat toe om aan 16 unieke tonen een binaire waarde toe te kennen. De meest courant gebruikte bitdiepte voor een digitaal beeld dat bestaat uit *grijswaarden* (*grayscale*) is 8 bits wat het mogelijk maakt dat een pixel 256 verschillende tonen kan voorstellen door een digitale waarde. Het is duidelijk dat het aantal digitaal voorstelbare tonen exponentieel stijgt naar mate men meer bits ter beschikking heeft. $2^1 = 2$ tonen; $2^3 = 8$ tonen; $2^8 = 256$ tonen.



De bitdiepte van een digitaal beeld.

Grijswaarden zijn eigenlijk variaties in lichtintensiteit terwijl de kleurtoon dezelfde blijft. Ondanks de variaties in lichtintensiteit is het digitale beeld dus monochroom.

We hebben reeds vermeld dat een lichtgevoelige chip enkel variaties in lichtintensiteit waarneemt. Om toch meerdere kleuren te onderscheiden wordt er gebruik gemaakt van kleurenfilters. Het gevolg is dat een kleurenbeeld in feite een combinatie is van 3 monochrome beelden namelijk groen, rood en blauw. Elk beeld heeft een bitdiepte van 8 bits (256 verschillende grijswaarden) en representeert 1 kleurkanaal. Een combinatie van de pixelwaarden van deze drie monochrome beelden levert een kleurenbeeld met een bitdiepte van 24 bits. Op deze manier is het toch mogelijk om 16 miljoen kleurentonen weer te geven.

Ook bitdiepte heeft een impact op de grootte van de bitstream. Indien een digitaal beeld van 30 x 30 pixels een bitdiepte heeft van 8 bits (grayscale) per pixel, dan zal de bitstream van dat beeld 720 bits (of 90 bytes) groot zijn. Wanneer de bitdiepte van datzelfde beeld 24 bits (kleur) per pixel was, dan zal de bitstream 2160 bits (of 270 bytes) groot zijn.

Dynamisch bereik

Het *dynamische bereik* of *dynamic range* staat voor het verschil tussen de zwakste en de sterkste lichtintensiteit die een scanner of een camera kan waarnemen. Licht dat op een gekleurd vlak wordt geprojecteerd, wordt deels geabsorbeerd. De rest van het licht wordt gereflecteerd. Een donkere kleur zal minder licht reflecteren dan een lichte kleur. Afhankelijk van de hoeveelheid

gereflecteerd licht waargenomen door de apparatuur, wordt een hogere of lagere digitale waarde toegekend aan de pixel. Tonen die te donker of te licht zijn worden afgerond naar de hoogste of laagste binair voorstelbare waarde (respectievelijk 255 tot 0 voor een bitdiepte van 8 bits). Concreet wil dat zeggen dat details waarvan de toon buiten het dynamische bereik van de camera of de scanner liggen, verloren kunnen gaan omdat ze niet of fout worden weergegeven.

In de context van het digitaliseren van fotomateriaal komt het op het volgende neer: het dynamische bereik van een scanner of een digitale camera moet groter zijn dan dat van de analoge foto's die je wil digitaliseren. Indien dit niet zo is, dan is er informatieverlies.

Op onderstaande foto's wordt het belang van de *dynamic range* gedemonstreerd. Door gebruik te maken van een beeldbewerkingprogramma werd de dynamic range van de rechter foto kleiner gemaakt waardoor ze vooral aan kwaliteit inboet in de donkere tonen. Dat valt op door het verdwijnen van details in donkere gebieden. De bitdiepte van beide foto's werd niet aangepast.



Links het 'originele' en rechts het aangepaste beeld (stadsarchief Antwerpen, OF 10947).

Bitdiepte en dynamisch bereik zijn aan elkaar gerelateerd. Een foto die werd gescand met een toestel met een voldoende groot dynamisch bereik maar een beperkte bitdiepte kan van mindere kwaliteit zijn. Wanneer voor het digitaliseren van diezelfde foto een toestel met een kleiner dynamisch bereik, maar met een grotere bitdiepte werd gebruikt, kan het resultaat kwalitatief juist beter zijn.²⁸ Het komt er dus op neer dat het gebruikte apparaat over een voldoende grote dynamic range moet beschikken om alle tonen van het originele document waar te nemen en dat een voldoende grote bitdiepte wordt gebruikt.

²⁸ A. KENNEY, O. RIEGER en R. ENTlich, *Moving theory into practice. Digital imaging tutorial*, Ithaca, NY, Cornell University Library, 2003, <http://www.library.cornell.edu/preservation/tutorial/intro/intro-05.html> (22 maart 2005).

cDAVID tip: berekenen van het dynamische bereik ²⁹

Dynamisch bereik wordt uitgedrukt in eenheden van lichtintensiteit (D). Het is belangrijk dat het dynamische bereik van de apparatuur gelijk of groter is dan dat van de afbeelding. Zo kan de afbeelding zonder problemen worden gedigitaliseerd. We berekenen het dynamische bereik van de afbeelding en van de apparatuur die we gebruiken om het beeld te digitaliseren.

- Bereken het dynamische bereik van de afbeelding

Bereken de hoogste (D_{min}) en laagste (D_{max}) lichtintensiteit van een afbeelding.

Meet de intensiteit van het licht dat op het stuk valt (I_o) en meet de hoeveelheid licht dat gereflecteerd wordt en door de apparatuur wordt waargenomen (I). Pas volgende formule toe om D te berekenen:

$$D = \log(I_o/I)$$

Neem de lichtste (bijvoorbeeld: een witte jurk) en donkstere gebieden (bijvoorbeeld schaduwen) van een afbeelding en bereken. Het dynamische bereik van de afbeelding bereken je met volgende formule:

$$\text{Dynamisch bereik} = D_{\max} - D_{\min}$$

- Bereken het dynamische bereik van de digitaliseringapparatuur

Dit gebeurt op een gelijkaardige manier. Dat doen we aan de hand van de bitdiepte die de apparatuur aanbiedt. Een bitdiepte van 8 bits levert 255 mogelijke waarden op (2⁸). Het dynamische bereik berekenen we met dezelfde formule. We nemen daarbij de hoogste (255) en laagste (1) waarden.

$$D = \log(255/1)$$

In dit geval is 2.4 het maximale dynamische bereik van een scanner met een bitdiepte van 8 bits.

Opgelet! Digitale beelden in kleur worden voorgesteld door minimale 24 bits. Dit valt echter uiteen in drie aparte kleurkanalen van elk 8 bits. De sensor neemt immers enkel variaties in lichtintensiteit waar. Door filtering kan men kleurenbeelden bekomen. We berekenen in dit geval dus niet het dynamisch bereik voor een bitdiepte van 24 bits maar slechts voor 8 bits. Onderstaande tabel geeft het dynamische bereik voor scanners met een andere bitdiepte:

>>

29 B. ATKINS, Dynamic Range, 2003, <http://www.photo.net/learn/drangle/> (27 juni 2005)

>>

Aantal bits (in kleur)	Dynamische bereik
24 bits (3 x 8 bits)	2.4
30 bits (3 x 10 bits)	3.0
36 bits (3 x 12 bits)	3.6
42 bits (3 x 14 bits)	4.2
48 bits (3 x 16 bits)	4.8

Opgelet! Het gaat hier om maximale waarden. Het digitale resultaat van een foto met een dynamische bereik van bijvoorbeeld 2.0 zal na digitalisering niet opeens het maximale dynamische bereik van de apparatuur aannemen. Omgekeerd zal het digitale resultaat van een afbeelding met een dynamisch bereik van 3.2 met een 24-bits scanner slechts een dynamisch bereik van 2.4 hebben. Daardoor bestaat de kans op informatieverlies.

In welke mate speelt dynamisch bereik een rol?

De Dmax van negatieven is doorgaans niet groter dan 1.5 en zelfs van overbelichte negatieven is de Dmax niet groter dan 2.0.

Het zelfde geldt voor positieven. Foto's op klassiek fotopapier hebben doorgaans een Dmax die niet groter is dan 2.0.

Doorgaans speelt het dynamische bereik dus geen rol van betekenis wanneer we reflectieven willen digitaliseren. Een degelijke scanner volstaat.

Bij het digitaliseren van transparanten zoals film of dia's speelt het dynamische bereik wel degelijk een rol. De Dmax van dia's kan immers gemakkelijk oplopen tot 3.5 afhankelijk van het type film dat werd gebruikt.

2.2.3 Andere parameters

Kleuren

Het correct overnemen van de originele kleuren in een digitale reproductie is een groot probleem. Het moet immers mogelijk zijn om het beeld in zijn originele kleuren digitaal te bewaren voor de toekomst. In realiteit is dit niet eenvoudig te bereiken. Een digitaal beeld dat wordt weergegeven op een computerscherm verschilt van het originele beeld en de kleuren van een afdruk op basis van het digitale beeld is nog anders. Hoe kan dat?

Kleuren kunnen wiskundig worden beschreven op basis van kleurmodellen. Een typisch model beschrijft een kleur als een groep van verschillende kleurcomponenten.³⁰ De kleuren die men bin-

³⁰ D. BOURGIN, *Color space FAQ*, 1994, <http://www.neuro.sfc.keio.ac.jp/~aly/polygon/info/color-space-faq.html>.

nen een kleurmodel kan voorstellen noemt men een kleurenruimte. Veelgebruikte kleurmodellen zijn RGB (Rood / Groen / Blauw) en CMYK (Cyaan / Magenta / Yellow / black). Op basis van dergelijke kleurmodellen worden digitale pixelwaarden weergegeven als kleuren.

Printers zijn gebaseerd op het CMYK-model en beeldschermen op het RGB-model. De conversie tussen RGB en CMYK is niet eenvoudig. Dit verklaart onder andere waarom er een verschil kan zitten tussen de kleur op een drukwerk en op een beeldscherm.

Ook de weergave verschilt van beeldscherm tot beeldscherm. De productiemethodes en de gebruikte materialen verschillen onderling tussen de fabrikanten. Daardoor zijn er verschillen in de kleurenruimte eigen aan de verschillende modellen. Zo valt er een groot kleurverschil te merken tussen de LCD- en de klassieke CRT- beeldschermen. Hetzelfde geldt ook voor scanners en digitale camera's.

Met de introductie van kleurprofielen probeert het ICC³¹ hieraan te verhelpen. Een kleurprofiel bevat de beschrijving van de kleurenruimte van een apparaat. Deze kleurprofielen kunnen worden ingebed in digitale afbeeldingen. De in software geïmplementeerde algoritmen zijn in staat om die kleurprofielen consistent te converteren zodat ze correct worden weergegeven.

De ICC heeft ook een aantal apparaatonafhankelijke kleurprofielen ontwikkeld en gestandaardiseerd. Deze vergemakkelijken de conversie tussen de verschillende apparaatafhankelijke kleurprofielen.

Voor lange termijn archivering is het aan te raden om het digitale beeld op te slaan in een gestandaardiseerd, apparaatonafhankelijk kleurenprofiel. Het bekendste is sRGB IEC 61966-2-1:1999.

*Artefacten*³²

In dit gedeelte geven we een kort overzicht van de voornaamste anomalieën – ook wel artefacten genoemd – die kunnen opduiken tijdens het digitaliseren van een beeld.

– Noise

De constante elektrische spanning, eigen aan elektronica en digitale apparatuur kan bij het digitaliseren storend werken. Deze spanning kan door de sensoren verkeerdelijk worden waargenomen als een stimulus van buitenaf. Wanneer de sensoren effectief een beeld maken, zorgt dit achtergrondsignaal voor variaties in de meting. Hierdoor toont de uiteindelijke waarde van een sample een afwijking. Een pixel zal daardoor een donkerdere of lichtere intensiteit krijgen dan dat ze effectief zou mogen hebben. Dit fenomeen treedt vooral op in egale kleurvlakken (blauwe lucht,...) en kan men herkennen als pixels die een andere intensiteit hebben dan hun burens. Dit fenomeen noemt men *noise*. Men kan dit minimaliseren maar niet volledig uitschakelen. De hoeveelheid noise hangt onder andere af van de gevoeligheid van de sensoren en de bitdiepte. Hoe hoger de gevoeligheid, des te meer kans op noise. Een lagere bitdiepte betekent minder variaties in intensiteit wat maakt dat de afwijking die ontstaat door noise veel groter zal zijn.

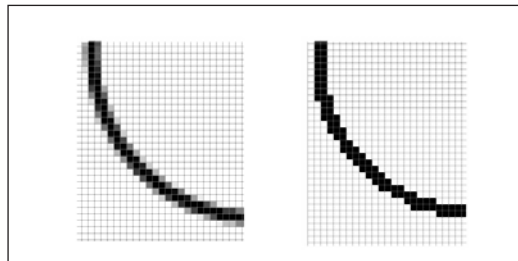
³¹ ICC staat voor International Color Consortium. Dit is een samenwerking van een aantal grote fabrikanten om oplossingen te ontwikkelen voor de correcte weergave van kleuren, <http://www.color.org> (26 mei 2005); <http://www.color.org/faqs.pdf> (26 mei 2005).

³² D. WILLIAMS, *Selecting a scanner*, in: *Guides to quality in visual resource imaging*, Washington, DC: Council on Library and Information Resources, 2000, <http://www.rlg.org/legacy/visguides/visguide2.html#background> (22 maart 2005).

In dit verband moeten we even ingaan op de bitdiepte die je in de specificaties van digitaliseringsapparatuur terugvindt. Scanners die volgens de fabrikant 42 bits kunnen halen, presteren in de praktijk meestal een stuk lager. Dat komt omdat, afhankelijk van de hardware, een aantal bits hoe dan ook onbruikbaar is door noise. Zo kan je met diezelfde scanner een goede kwaliteit bereiken met een reële bitdiepte van bijvoorbeeld 36 bits. De kwaliteit van digitaliseringsapparatuur kan men dus meten door de verhouding tussen het signaal en de hoeveelheid noise of de *signal-to-noise ratio* te onderzoeken.

→ Aliasing

Een tweede probleem is *aliasing*. Dit fenomeen treedt op wanneer een beeld aan een te lage resolutie werd gedigitaliseerd. Een klassiek voorbeeld van aliasing is *moiré*. Een beeld dat zeer veel fijne details in terugkerende patronen heeft, heeft een *hoge spatiale frequentie*. Wanneer je zo'n afbeelding digitaliseert met een te lage resolutie, dan zal het digitale beeld een storend interferentiepatroon vertonen.³³ Zo vormen rasterdrukdocumenten een probleem. Documenten gemaakt in een halve toon druk bestaan namelijk uit een verzameling aparte *dots* in regelmatige patronen die door het oog worden gepercipieerd als één geheel. Wanneer dergelijk drukwerk aan een te lage resolutie wordt gedigitaliseerd dan treedt er moiré op. Een ander voorbeeld van aliasing zijn de zogenaamde *jaggies*. Ronde vormen worden in een te lage resolutie slecht weergegeven: ze worden hoekig en hebben iets weg van *zaagtanden*. Om dit effect tegen te gaan gebruikt men een techniek genaamd *anti-aliasing* die het beeld op lage resoluties softwarematig probeert te verbeteren. De scherpe overgang tussen twee gebieden wordt zachter gemaakt door interpolatie.



Anti-aliasing (links) versus aliasing (rechts).

Meestal volstaat het echter om met degelijke apparatuur aan een hoge resolutie en op een voldoende grote bitdiepte te digitaliseren om deze effecten te vermijden.

~ Stof en krassen

Stof moet worden vermeden. Stof op een document, de digitaliseringsapparatuur en in de omgeving hebben een invloed op de kwaliteit van het document. Stof vormt een storend element in de digitale kopie. Zeker als deze werd geproduceerd met een hoge resolutie. Stof wordt dan zichtbaar.

Soms is een stuk bekrast zonder dat we dit met het blote oog kunnen zien. Door de belichting in een scanner en het instellen van een hoge resolutie kunnen krassen zichtbaar worden en het digitale beeld storen. De huidige apparatuur is uitgerust om krassen te detecteren en te corrigeren in het digitale beeld.

³³ Met een hoge spatiale frequentie bedoelen we dat het aantal wekerende elementen per lengte-eenheid zeer hoog ligt in een patroon. Digitaliseren met een te laag ingestelde resolutie zorgt ervoor dat slechts een deel van de elementen door de sensoren in de scanner wordt gedetecteerd wat een zichtbaar en storend 'patroon' oplevert.

~ *Kleurafwijkingen*

Het gaat hier om storende gekleurde vlakken in het beeld. Dit komt vooral voor bij goedkopere apparatuur. De onderdelen van dergelijke toestellen zijn kwalitatief eerder ondermaats. Dit uit zich in de eerste plaats in kleurafwijkingen.

cDAVID aanbeveling: resolutie, bitdiepte en kleur

Resolutie

De resolutie is het aantal pixels dat een digitaal beeld bevat. Een pixel is het kleinste deelelement van het digitale beeld.

- ↪ Meer pixels betekenen meer beeldinformatie en dus een betere kwaliteit.
- ↪ Meer pixels betekent ook een langere bitstream. Het digitale beeld zal meer opslagcapaciteit vereisen.
- ↪ Pas op voor interpolatie. Een voorbeeld van interpolatie is de digitale zoom op een camera.
- ↪ Wees waakzaam voor zeer hoge resoluties bij de aanschaf van een scanner. Er is een verschil tussen de werkelijke resolutie en de resolutie bereikt door interpolatie. Voor doorsnee scanopdrachten volstaat een scanner met een werkelijke resolutie van 1 200 dpi.
- ↪ De resolutie van een scanner wordt uitgedrukt in dpi of ppi. De resolutie van een camera wordt uitgedrukt in megapixel. De resolutie van de weergave op beeldscherm wordt uitgedrukt in 'pixels x pixels'.

Bitdiepte

De bitdiepte staat voor het aantal bits dat gebruikt wordt om een pixelwaarde voor te stellen. Hoe meer bits er zijn, des te meer informatie kan er worden bewaard.

- ↪ Voor een digitaal beeld in kleur: 24 bits RGB.
- ↪ Voor een digitaal beeld in grijswaarden: 8 bits grayscale.
- ↪ Voor een digitaal beeld in zwart-wit: 1 bits.

Kleuren

- ↪ Converteer het digitale bestand na digitaliseren naar een apparaatonafhankelijk kleurenprofiel: sRGB IEC 61966-2-1:1999 kleurenprofiel.

2.3 De bestandsformaten

De output van digitaliseringapparatuur zoals digitale camera's en scanners is een bitstream. Dit is het digitale beeld in zijn meest elementaire vorm: als een continue reeks enen en nullen. Een computersysteem moet een dergelijke bitstream niet alleen kunnen bewaren maar ook kunnen lezen, interpreteren en ten slotte weergeven als een digitaal beeld. Op zichzelf kan een computer dat niet: ook de bitstream van een geluidsfragment of een tekstdocument zijn een reeks enen en nullen.

Daarom wordt de bitstream in een soort herkenbare container bewaard. Deze container noemt men een *bestandsformaat*. Een bestandsformaat bestaat uit de ruwe bitstream in een vorm die herkenbaar is voor de computer. Dit gebeurt door bijvoorbeeld vooraan een herkenbare hoofding of *header* te plaatsen. Zo kan men het onderscheid maken tussen bijvoorbeeld digitaal geluid en digitaal beeld. Daarnaast bieden bestandsformaten een aantal extra mogelijkheden aan zoals het bewaren van metadata.

Een bestandsformaat wordt beschreven in een *specificatie*. Deze vertelt hoe een bitstream er precies moet uitzien om herkenbaar te zijn. Een voorbeeld hiervan is de specificatie van het JPEG-bestandsformaat die beschikbaar is op het internet.³⁴ Softwareontwikkelaars kunnen op basis van



Restauratie van de zegewagen van "Calloo" voor het Rubensjaar, 1976. Foto Filmgroep stad Antwerpen (stadsarchief Antwerpen, FOTO 405/T).

³⁴ JPEG File Interchange Format version 1.02, 1992, <http://www.w3.org/Graphics/JPEG/jfif3.pdf> (17 maart 2005).

JPEG staat voor Joint Picture Experts Group, <http://www.jpeg.org/> (26 mei 2005).

de specificatie een applicatie schrijven die het formaat van een bitstream herkent en de bitstream kan weergeven als een beeld.

Er bestaan meerdere bestandsformaten, elk met hun typische kenmerken en hun eigen mogelijkheden. Om de bitstream van een digitaal beeld te bewaren is een verantwoorde keuze noodzakelijk. Sommige bestandsformaten zijn bijvoorbeeld totaal ongeschikt voor lange termijn archivering, maar zijn dan weer zeer geschikt voor de distributie via het internet. In deze paragraaf belichten we een aantal aspecten waar je bij de keuze van het bestandsformaat op moet letten. We geven vervolgens een kort overzicht van bestaande bestandsformaten.³⁵

2.3.1 Open versus gesloten formaten

Om een applicatie te kunnen ontwikkelen die een bepaald formaat herkent en kan interpreteren moet de ontwikkelaar de specificatie kennen. De specificatie beschrijft immers het bestandsformaat. Zonder de specificatie kan een ontwikkelaar geen software ontwikkelen die het formaat kan begrijpen. We maken hierbij een onderscheid tussen *open* en *gesloten formaten*.

Van een gesloten bestandsformaat is de specificatie niet vrij beschikbaar. Omwille van commerciële belangen wordt de specificatie vaak niet vrijgegeven. Op die manier wordt het concurrenten moeilijk gemaakt om alternatieve software te schrijven die het bestandsformaat ondersteunt. Een voorbeeld van een gesloten formaat is het Microsoft Word formaat. De specificatie hiervan wordt niet vrijgegeven waardoor het heel moeilijk is voor concurrenten van Microsoft om software te schrijven die het formaat ondersteunt. Bij een open bestandsformaat is de specificatie wel beschikbaar. Iedereen is vrij om de specificatie te implementeren en het formaat te ondersteunen.

Voor het bewaren van digitale beelden op lange termijn is het gebruik van een open formaat is te verkiezen boven een gesloten.

- De fabrikant van een gesloten formaat kan beslissen de ondersteuning te stoppen. Bij gebrek aan alternatieven kan het bestandsformaat op termijn niet meer ingelezen worden.
- Een open bestandsformaat is *uitwisselbaar*. Dat wil zeggen dat men niet gebonden is aan een platform (bijvoorbeeld: het windows-platform). Een open bestandsformaat wordt vaak op verschillende platformen ondersteund. Wanneer er geen ondersteuning (meer) is voor een bepaald platform, dan is men vrij die zelf te ontwikkelen.

We kunnen open en gesloten formaten ook opdelen volgens de mate waarin ze gestandaardiseerd zijn. We maken een onderscheid tussen *officiële* en *de facto standaarden*. Een formaat krijgt de status van officiële standaard wanneer het erkend werd door een standaardisatieorganisatie zoals de ISO of IEEE.³⁶ Formaten hoeven niet noodzakelijk officieel te worden erkend. Door hun wijdverspreid gebruik kunnen we sommige formaten beschouwen als de facto standaarden. Deze laatste groep is veruit de grootste.

Niettegenstaande standaarden in de regel geschikte archiveringsformaten zijn is enige nuancering op zijn plaats. Officiële standaarden bieden niet noodzakelijk meer garanties dan de facto standaarden wat betreft ondersteuning. De effectieve marktevoluties en standaardisatie lopen niet

³⁵ TASI. *Choosing a file format*, 2002, <http://www.tasi.ac.uk/advice/creating/pdf/format.pdf> (22 maart 2005).

³⁶ ISO, International Organization for standardization, <http://www.iso.org> (26 mei 2005).

IEEE, Institute for Electrical and Electronics Engineers, <http://www.ieee.org> (26 mei 2005).

altijd parallel. Of een standaard al dan niet geschikt is om digitale stukken in te bewaren hangt af van de implementatie door softwareontwikkelaars en de toepassing door de gebruikers.³⁷

2.3.2 *Compressie*

Er bestaan twee typen digitale beelden: *raster images* (of bitmapped) en *vector images* (of object oriented). Een raster image of bitmap is een matrix of raster van pixels die in het geheugen worden ingelezen en waarin elke afzonderlijke pixel een unieke locatie en een eigen waarde heeft. Een vector image is in feite een reeks wiskundige instructies die een applicatie interpreteert om een digitaal beeld te genereren. Een raster image is dus op zich een volwaardig digitaal beeld terwijl een vector image in feite een recept is dat je moet volgen om een digitaal beeld te creëren.

Voor digitaliseringdoeleinden is het gebruik van raster imaging beeldtechnieken zoals hierboven beschreven veel geschikter. Digitale beelden gegenereerd door digitale camera's en scanners zijn bitmaps of raster images. Het nadeel van deze techniek is dat de hierin resulterende bitstream al snel zeer groot wordt. De bitmap van een beeld moet al flink wat samples bevatten om een digitaal beeld van goede kwaliteit weer te geven. Het is geen uitzondering dat voor de bewaring van één gedigitaliseerd beeld ettelijke megabytes opslagcapaciteit vereist zijn.³⁸

Op dergelijke grote bestanden kan je compressiealgoritmen toepassen. Deze voeren een aantal wiskundige operaties uit die ervoor zorgen dat de grootte van het digitale bestand wordt gereduceerd. We onderscheiden twee soorten compressie *lossy* en *lossless*. Bij de eerste vorm van compressie is er tijdens het comprimeren onherroepelijk informatieverlies. Lossless compressie probeert dit te vermijden door geen informatie verloren te laten gaan. De bitstream wordt weliswaar ingekort maar de originele bitstream kan je altijd volledig reconstrueren op basis van de comprimeerde bitstream.

Lossless compressie kan de grootte van een bitstream vaak niet genoeg comprimeren om de bitstream werkelijk handelbaar te maken voor bijvoorbeeld distributie via het internet. Lossy compressie biedt wel de mogelijkheden aan om een bitstream met een voldoende grote compressiefactor te reduceren zodat ze gemakkelijk kan worden verspreid. Lossy compressie maakt gebruik van irrelevantie reductie. De bitstream wordt vereenvoudigd door data af te ronden en irrelevante waarden weg te laten. Dit proces is onomkeerbaar waardoor bij decompressie de originele bitstream slechts bij benadering kan worden gereconstrueerd.

Voor het decomprimeren van een digitaal beeld moet de software beschikken over decompressiealgoritmen die het digitale beeld terug kunnen reconstrueren. Naar lange termijn bewaring toe vormt dit een groot probleem: door de technologische vooruitgang geraken verouderde algoritmen in onbruik. Wanneer ze niet meer worden ondersteund door de software dan kan het digitale beeld niet meer worden gelezen. Hoewel het aantrekkelijk is om digitale beelden te comprimeren omwille van plaatsbesparing, is het niet aan te raden om compressie toe te passen wanneer we digitale beelden op lange termijn willen bewaren, bijvoorbeeld bij de archiefkopieën.

³⁷ F. BOUDREZ, *Standaarden voor digitale archiefdocumenten*, Stadsarchief Antwerpen, Antwerpen, 2005, <http://www.antwerpen.be/david/website/teksten/DAVIDbijdragen/Standaarden.pdf>.

J. JOCHMANS en P. STRICKX, *Richtlijnen en aanbevelingen voor het gebruik van open standaarden en/of open specificaties bij de federale overheidsbesturen*, 2004, <http://www.belgium.be/eportal/application?pageid=contentPage&docId=36434> (22 maart 2005).

³⁸ TASI (Technical Advisory for Images), *TASI Advice Paper. File formats and compression*, 2003, <http://www.tasi.ac.uk/advice/creating/pdf/formats.pdf> (22 maart 2005).

Een voorbeeld van een bestandsformaat dat lossy compressie ondersteunt is het JPEG-bestandsformaat. Digitale beelden die werden opgeslagen in dit formaat kunnen tot slechts een fractie van hun oorspronkelijke grootte worden verkleind. Daardoor is het een geschikt formaat voor de online verspreiding van digitaal beeld.³⁹

Een voorbeeld van een lossless bestandsformaat is *PNG* of *Portable Network Graphics*. De compressiealgoritmen die dit formaat ondersteunt kunnen echter niet dezelfde compressieratio's halen als JPEG waardoor het minder geschikt is voor de verspreiding.

Een bestandsformaat kan ook meerdere compressiealgoritmen ondersteunen. Het TIFF-bestandsformaat bijvoorbeeld, kan een bitstream zowel ongecomprimeerd als JPEG-gecomprimeerd opslaan.

cDAVID aanbeveling: wanneer passen we compressie toe?

- Archiveren? Neen! Bewaar het digitale beeld ongecomprimeerd (*uncompressed*) in een bestandsformaat dat dit ondersteunt. Bijvoorbeeld TIFF.
- Verspreiden? Ja! Bewaar het digitale beeld gecompriemerd (*compressed*) in een bestandsformaat dat dit ondersteunt. Bijvoorbeeld JPEG of PNG.

2.3.3 Een geschikt bestandsformaat kiezen

Het kiezen van een geschikt bestandsformaat om de digitale beelden in op te slaan is van cruciaal belang. Heel wat formaten passen immers allerlei bewerkingen toe op de originele bitstream zoals bijvoorbeeld compressie of encryptie. Anderen laten niet toe om essentiële metadata in het digitale bestand in te kapselen. De ruime waaier aan beschikbare bestandsformaten beschouwen we het beste als een gereedschapskist waarvan elk onderdeel zijn eigen finaliteit heeft. Het komt erop neer om het juiste formaat te kiezen voor de doelstellingen die je voor ogen hebt.⁴⁰

De archiefkopie

De archiefkopie is het digitale beeld dat zal worden bewaard op lange termijn. We moeten hier rekening houden bij de keuze van het bestandsformaat. Formaten raken na verloop van tijd in onbruik. Bovendien verhinderen intellectuele rechten het ontwikkelen van alternatieve implementaties. Vooral voor bedrijfseigen of gesloten bestandsformaten die geen breed verspreide toepassing kennen, vormt dit een probleem. Op lange termijn wordt het onmogelijk om bitstreams opgeslagen in een bedrijfseigen bestandsformaat te lezen.

Om dit te vermijden is een *geschikt archiveringsformaat* bij voorkeur een open standaard die ten allen tijde kan worden ondersteund.

³⁹ Eigenlijk moet er een onderscheid worden gemaakt tussen het JPEG-compressiealgoritme en het JFIF-bestandsformaat dat dit algoritme ondersteunt. JFIF staat voor JPEG File Interchange Format. Doorgaans wordt strikt genomen verkeerdelijk naar JFIF verwezen als het JPEG bestandsformaat.

T. LANE, JPEG-FAQ, 1999, <http://www.faqs.org/faqs/jpeg-faq/> (22 maart 2005).

⁴⁰ TASI. *Choosing a file format*, 2002.

Het DAVID project formuleerde een aantal criteria waaraan een formaat moet voldoen wil het beschouwd kunnen worden als een geschikt archiveringsformaat.⁴¹

cDAVID aanbeveling: de archiefkopie

Gebruik het TIFF (Tagged Image File Format) bestandsformaat voor het opslaan van de archiefkopie van een digitaal beeld.

Het TIFF-formaat kent verschillende versies: bij voorkeur slaan we digitale archiefkopieën op als *uncompressed baseline TIFF v6*.⁴²

De afgeleide kopieën

Een digitaal beeld kan op zeer uiteenlopende wijze worden afgeleverd aan de eindgebruiker. We maken het onderscheid tussen:

- ↪ Digitale beelden die via het internet worden verspreid (*raadplegingkopieën*).
- ↪ Digitale beelden die kunnen worden gebruikt voor reproductie in drukwerk (*reproductiekopieën*).
- ↪ Digitale beelden die digitaal werden gemanipuleerd. Bijvoorbeeld digitale restauratie. (*geoptimaliseerde kopieën*).

Dergelijke kopieën noemen we *afgeleide kopieën*. Naar gelang de doelstellingen en de aard worden deze digitale beelden bewaard en verspreid in een geschikt bestandsformaat.

De meest gebruikte vorm van afgeleide kopieën bestaat erin beelden digitaal beschikbaar te stellen via een webinterface. Web-browsers ondersteunen slechts een beperkt aantal formaten onder andere *JPEG* en *PNG*. Onze voorkeur gaat uit naar open formaten die door een ruim scala aan browsers (bijvoorbeeld Mozilla, Internet Explorer, Safari) worden geïmplementeerd. Zowel JPEG als PNG ondersteunen compressie. Dit is nodig aangezien de bandbreedte op het internet eerder beperkt is. Kleinere, gecomprimeerde digitale bestanden zijn dan ook makkelijker te hanteren in een webomgeving.⁴³

De afbeeldingen kunnen ook worden gebruikt in een prepress omgeving. Digitale beelden kunnen op een analoge drager zoals papier worden gereproduceerd. De verschillende druktechnieken impliceren een ruime waaier aan kwaliteitseisen. Je hebt immers een goede digitale kopie nodig om een reproductie met een goede kwaliteit te kunnen maken. Daarom is het belangrijk om met de drukker duidelijke afspraken te maken. Dat geldt ook voor de keuze van het bestandsformaat.

⁴¹ F. BOUDREZ, *DAVID Richtlijn 9: Digitaliseren van analoge archiefdocumenten*.

⁴² F. BOUDREZ, *DAVID Richtlijn 4. Standaarden voor formaten*, <http://www.antwerpen.be/david/website/teksten/Richtlijn4.pdf> (22 maart 2005).

Het TIFF formaat werd erkend als een officiële ISO-standaard: ISO-12639: "Graphic technology -- Prepress digital data exchange -- Tag image file format for image technology", <http://www.iso.org> (27 mei 2005).

⁴³ TASI. *Choosing a file format*, 2002.

Een drukker gebruikt doorgaans het CMYK-kleurenmodel om publicaties te drukken terwijl digitale beelden meestal in RGB werden opgeslagen. Een omzetting is noodzakelijk. Indien je zelf niet de technische kennis bezit is het beter om dit over te laten aan de drukker.⁴⁴

De geoptimaliseerde kopie beschouwen we dan weer als een archiefkopie. Digitale restauratie is arbeidsintensief en duur. Een gerestaureerde digitale kopie moeten we dus als een archiefkopie behandelen willen we deze duurzaam bewaren.

cDAVID aanbeveling: de afgeleide kopieën

Raadplegingkopieën

Gebruik een bestandsformaat dat wordt ondersteund door browsers. Gebruik een bestandsformaat dat compressie toepast.

Bijvoorbeeld: JPEG of PNG

Reproductiekopieën

Beslis over het te gebruiken bestandsformaat in onderlinge afspraak met de drukker of diegene die een digitale reproductie wenst.

Gangbare bestandsformaten: TIFF, PDF en EPS

Geoptimaliseerde kopieën / veiligheidskopie

Behandel een geoptimaliseerde kopie of een veiligheidskopie als een archiefkopie.

Gangbaar bestandsformaat: TIFF

2.4 Workflow

In dit onderdeel brengen we alle concepten die we in dit hoofdstuk hebben behandeld samen in een werkbaar geheel: de workflow. Dit is een verzamelnaam voor de operationele aspecten van het digitaliseren. Wat moet er gebeuren en kadert dit in de doelstellingen van het project? Hoe verloopt het digitaliseren? Wie is er verantwoordelijk voor welke stappen in het proces? Dit zijn vragen die vooraf moeten worden gesteld. Digitaliseren vereist immers een duidelijk overdachte aanpak. Zomaar beginnen digitaliseren kan ertoe leiden dat aan de scope van het project volledig wordt voorbij gegaan. In het slechtste geval betekent dit dat je het digitaliseren van de stukken volledig opnieuw moet doen wat een flinke meerkost en tijdverlies met zich meebrengt. Dergelijke problemen zijn te vermijden door te vertrekken vanuit een duidelijke planning.

We bespreken enkel de algemene concepten die van belang zijn bij het modelleren van de workflow. In de realiteit zijn er heel wat projectspecifieke aspecten die aandacht vereisen. Fragiele stukken vragen een geheel eigen behandeling. Stukken die vuil zijn of aangetast door schimmels vereisen dat er in de planning een schoonmaakbeurt wordt voorzien. Digitaliseren met het oog op louter archiveren vereist niet direct de productie van afgeleide kopieën. Het model dat we

⁴⁴ TASI. *Choosing a file format*, 2002.

hier voorstellen is slechts een mogelijke strategie om digitaliseren aan te pakken. Alles in acht genomen verschilt de situatie van waaruit wordt vertrokken per project en vereist elke project een enigszins op maat ontworpen workflow.

cDAVID aanbeveling: een goed vastgelegde workflow

Bij het ontwikkelen van een workflow dienen we rekening te houden met volgende aandachtspunten.⁴⁵

- *De beste beeldkwaliteit produceren.*
 - ~ Baseer de digitaliseringparameters op de kwaliteit van het analoge origineel en de doelstellingen van het project.
 - ~ Bouw bijkomend een veiligheidsmarge in. Doe dit om te vermijden dat het analoge document meer dan één keer moet worden gemanipuleerd.
- *De 'originale' data als moederkopie archiveren.*
 - ~ Manipuleer bitstreams die direct van de scanner of een digitale camera komen zo min mogelijk.
- *Beleid inzake compressie en beeldmanipulatie bepalen.*
 - ~ Digitale beelden worden gemaakt en gearcheveerd op ware grootte. Dit wil zeggen dat de grootte of de 'schaal' van het digitale beeld noch de resolutie softwarematig mogen worden aangepast.
 - ~ Maak geen gebruik van compressie om digitale beelden te archiveren.
- *Kwaliteitscontroles uitvoeren*
 - ~ Voer regelmatig een kwaliteitscontrole uit.
 - ~ Vergelijk de kwaliteit van nieuwe digitale beelden met oude. Maak gebruik van testkaarten om degradatie van kwaliteit vroegtijdig te kunnen opsporen.
 - ~ Kalibreer de apparatuur op regelmatige tijdstippen zodat de beeldkwaliteit gelijk blijft.
- *Geschikt archiveringsformaat vastleggen.*
 - ~ Gebruik een archiveringsformaat dat geschikt is voor de opslag van digitale moederkopieën.
- *Metadata verwerven.*
 - ~ Een geschikt archiveringsformaat laat toe om metadata in te kapselen in het digitale beeld. Vul de relevante velden aan daar waar nodig (copyright, eigenaar, beschrijving, technische informatie, referentie, ...).
- *Metadata verwerven.*
 - ~ Een geschikt archiveringsformaat laat toe om metadata in te kapselen in het digitale beeld. Vul de relevante velden aan daar waar nodig (copyright, eigenaar, beschrijving, technische informatie, referentie, ...).

>>

⁴⁵ TASI, *generic image workflow*, 2002, http://www.tasi.ac.uk/advice/managing/pdf/workflow_generic.pdf (30 mei 2005).

BOUDREZ F., DAVID Richtlijn 9: *Digitaliseren van analoge archiefdocumenten*, <http://www.antwerpen.be/david/website/teksten/Richtlijn9.pdf> (30 mei 2005).

>>

- *Geschikte drager gebruiken.*
 - ~ Kies een drager die geschikt is voor het opslaan van data op lange termijn. Een mogelijke keuze is cd-r.⁴⁶
- *Keuzen inzake afgeleide kopieën maken*
 - ~ Gebruik de moederkopie enkel voor het archiveren van de data. Maak afgeleide kopieën voor de distributie van de beelden.
 - ~ Hou rekening met de finaliteit van de afgeleide kopieën.
- *Enkel afgeleide beelden van moederkopieën maken.*
 - ~ Maak *nóóit* afgeleide kopieën van digitale bestanden die reeds een afgeleide kopie zijn. Een kopie van een kopie nemen betekent verlies aan kwaliteit.
- *Veiligheidskopieën maken*
 - ~ Maak van elke archiefkopie een exacte kopie die als veiligheidskopie dienst doet. Een veiligheidskopie wordt nooit samen met de archiefkopie op dezelfde locatie bewaard. Ze wordt bewaard op een veilige plaats zoals een kluis. Ze wordt nooit gebruikt tenzij ter vervanging van de archiefkopie indien die laatste niet meer leesbaar is.

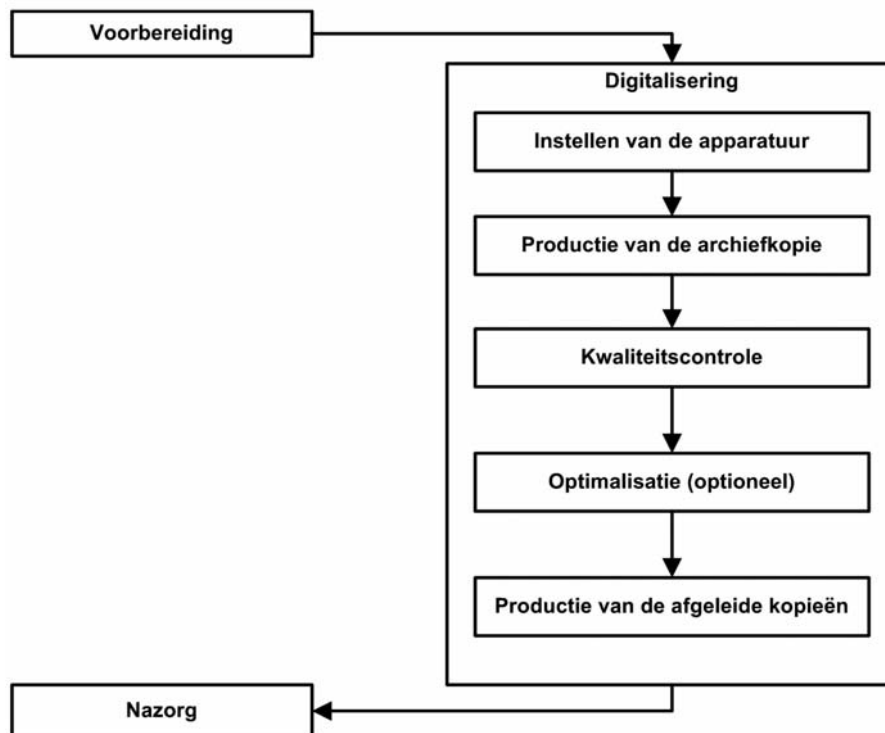
De stappen waaruit de workflow kan bestaan

In het modelleren van de workflow moeten de verschillende taken of stappen in een digitalisering-proces precies worden geïdentificeerd. Dat bestaat uit vijf onderdelen:

- het instellen van de apparatuur
- de productie van de archiefkopie
- de kwaliteitscontrole
- (optioneel) optimalisatie + controle van de geoptimaliseerde kopieën
- productie van afgeleide kopieën + controle van de afgeleide kopieën.

Niettegenstaande het instellen van de apparatuur strikt genomen tot de voorbereiding behoort, beschouwen we het in dit rapport als een wezenlijk onderdeel van de eigenlijke digitalisering. Het ontwerp van de workflow gaat echter verder dan de digitalisering. Er moet rekening worden gehouden met de voorbereiding: de selectie van de stukken, het ophalen uit het archief, het uitpakken, schonen, enzovoort. Maar ook de nazorg na het digitaliseren: het hernummeren, opnieuw verpakken, het terugplaatsen van de stukken in de bewaarplaats, enzovoort.

⁴⁶ BOUDREZ F, DAVID *Richtlijn 2: Duurzame cd's*, <http://www.antwerpen.be/david/website/teksten/Richtlijn2.pdf> (30 mei 2005).



Workflow binnen het digitaliseringsproces.

2.4.1 De voorbereiding van de te digitaliseren stukken

De voorbereiding begint bij het ophalen van de geselecteerde stukken uit de bewaarplaats en loopt tot op het moment waarop het de digitaliseringapparatuur wordt opgestart. De voorbereiding heeft een impact op het resultaat van het digitaliseren. Tijdens de voorbereiding spelen onder andere volgende elementen een belangrijke rol.

- ↪ Het ophalen van de geselecteerde stukken uit de bewaarplaats.
Probeer daarbij de normale werking van de instelling zo min mogelijk te verstoren. Wanneer het een omvangrijke selectie betreft is het aangewezen om niet alle stukken in één keer klaar te maken. Immers, hoe meer stukken, hoe langer de doorlooptijd en dus hoe langer ze niet beschikbaar zijn voor het publiek.
- ↪ De fysieke kenmerken van de drager.
Sommige dragers zoals glasplaten, negatieven of beschadigde stukken vragen speciale aandacht wanneer je deze manipuleert. Tref de nodige maatregelen om verdere beschadiging te vermijden.
- ↪ Verpakking.
Het uitpakken moet met de nodige zorg gebeuren. De fysieke staat van de verpakking kan ervoor zorgen dat dit een tijdrovend werkje is.
- ↪ Klaarmaken van de stukken.
Ze kunnen bevestigd zijn op een drager met nietjes, lijm, plakband of andere middelen.

Overweeg of de afbeeldingen al dan niet worden losgemaakt en afzonderlijk gedigitaliseerd.

In welke mate is de context van belang? Is het wenselijk het stuk en de afbeeldingen te scheiden. Soms is het beter om ineens de context mee te digitaliseren.

- Draag ten allen tijde zorg voor de stukken.
Bewaar stukken in wachtlijn met de nodige zorg zodat ze niet kunnen worden beschadigd.
- Herstel beschadigingen eventueel vooraleer ze te digitaliseren.
Indien nodig of gewenst kan je overwegen de originele stukken te restaureren vooraleer ze te digitaliseren. Dit vereist de nodige expertise.

Een goede voorbereiding verzekert een snelle en efficiënte digitalisering en vormt een garantie voor het welzijn van de stukken tot op het moment van digitaliseren. Het is aangewezen om voor deze fase voldoende tijd te voorzien in de workflow. Het voorbereiden van de stukken gebeurt best door medewerkers die de nodige ervaring bezitten om op een verantwoorde manier met hiermee om te gaan.

2.4.2 Het instellen van de apparatuur

Om tot een resultaat te komen dat op zijn minst *fit-for-purpose* is, moet de hard- en de software juist worden ingesteld. Belangrijke parameters zijn de *resolutie* en de *bitdiepte*. Maar hoe moet dit nu juist gebeuren?

Veel hangt af van de fysieke staat en de aard van de stukken maar ook de doelstellingen spelen een grote rol. We gaan er van uit dat het eindresultaat een archiefkopie is. De kwaliteit van het beeld moet zeer goed zijn om een waarheidsgetrouwe weergave van het origineel te kunnen zijn.

De waarden die we hier voorstellen zijn gebaseerd op de waarden die verschillende grote digitaliseringprojecten hanteren enerzijds, en anderzijds op de aanbevelingen van onderzoek gevoerd aan de Cornell University en de NINCH.⁴⁷ De documenten die worden gedigitaliseerd zijn vaak van zeer uiteenlopend van aard: gedrukte teksten, foto's, prenten, manuscripten, transparanten, enzovoort. In de aanbevelingen wordt veelal een classificatie van documenttypes gemaakt waaraan specifieke waarden en parameters worden gekoppeld.

A. Kenney, O. Rieger en R. Entlich maken het onderscheid tussen vijf grote documenttypes:⁴⁸

- Gedrukte tekst en lijntekeningen
Dit zijn documenten met een monochrome kleurtoon en geen variatie in de kleurintensiteit.

⁴⁷ A. KENNEY, O. RIEGER en R. ENTlich, *Moving theory into practice. Digital imaging tutorial.*, Ithaca, NY, Cornell University Library, 2003, <http://www.library.cornell.edu/preservation/tutorial/contents.html> (22 maart 2005); <http://www.library.cornell.edu/preservation/tutorial/conversion/table3-1.html> (21 maart 2005).

LGC-NLW, *Dros Y Rhiniog. Crossing the threshold. The national library of Wales's Digitisation strategy*, Wales, 2004, http://www.llgc.org.uk/drych/digido_s01.htm (1 juni 2005).

NINCH, *The NINCH Guide to Good Practice*, 2002, <http://www.nyu.edu/its/humanities/ninchguide/VI/> (30 mei 2005).

⁴⁸ A. KENNEY, O. RIEGER en R. ENTlich, *Moving theory into practice. Digital imaging tutorial.*

De grenzen tussen de kleurvlakken zijn scherp afgelijnd. Deze documenten werden veelal machinaal vervaardigd.

→ Manuscripten

Dit zijn teksten waarvan de letters niet altijd even scherp zijn afgelijnd: er zijn variaties in de intensiteit van de inkt. Manuscripten bevatten ook veel specifieke details die eigen zijn aan het handschrift van de opsteller. Manuscripten kunnen bovendien verlucht zijn met afbeeldingen.

→ Halve toon

Dit zijn reproducties van (foto)grafisch materiaal die bestaan uit regelmatige patronen van punten (dots) of lijnen die door het menselijke oog op een afstand worden gezien als één afbeelding.

→ Continue toon

Dit zijn documenten die bestaan uit egale, aaneengesloten kleurvlakken met een variabele intensiteit. Bijvoorbeeld: foto's (zowel zwart-wit als in kleur), tekeningen en schilderijen, enzovoort.

→ Gemengde toon

Dit is een vorm die zowel elementen in halve als continue toon bevat.



De verschillende documenttypes. Van links naar rechts: gedrukte tekst, manuscript, halve en continue toon.

cDAVID aanbeveling: aanbevolen digitaliseringwaarden			
	Resolutie	Bitdiepte	Kleurenprofiel
Gedrukte tekst Geïllustreerd Niet-geïllustreerd (kleur) Niet-geïllustreerd (z/w)	300 dpi 300 dpi 600 dpi	8 bits 24 bits 1 bits	Greyscale sRGB --
Manuscripten Kleur Grijswaarden	300 dpi 300 dpi	24 bits 8 bits	sRGB Greyscale
Halve toon Kleur Grijswaarden	400 dpi 400 dpi (met <i>descreening</i> optie!)	24 bits 8 bits	sRGB Greyscale
Continue toon Kleur Grijswaarden	300 dpi 300 dpi	24 bits 8 bits	sRGB Greyscale
Gemengde toon Kleur Grijswaarden	300 dpi 300 dpi	24 bits 8 bits	sRGB Greyscale

Opmerkingen bij deze tabel:

→ Waarden

Dit zijn de **minimaal aanbevolen waarden** voor de productie van **een archiefkopie**. Afhankelijk van de situatie is het goed mogelijk dat een hogere resolutie een beter resultaat oplevert. Maak een aantal digitale proeven om de optimale resolutie te bepalen.

→ Kleur

De bitdiepte bepaalt of een stuk al dan niet in kleur wordt gedigitaliseerd. Hou rekening met de invloed op de grootte van de bitstream. Een 24 bits beeld zal in omvang drie keer zo groot zijn als een 8 bits beeld.

Digitaliseer de kleur wanneer deze relevant is voor de kwaliteit van het digitale beeld en de doelstellingen van het project. Bijvoorbeeld: indien je enkel de leesbaarheid van een digitale kopie van een manuscript beoogt, dan volstaan 8 bits. Indien ook zaken zoals inkt, textuur van het perkament, enzovoort belangrijk zijn, gebruik dan 24 bits.

→ Digitaliseren van stukken in halve toon

Bij het digitaliseren van halve toon stukken moet erop gelet worden dat de resolutie voldoende hoog is om moiré-effecten te vermijden. Tegenwoordig bezit de meeste apparatuur een (automatische) *descreening*-functie om moiré te vermijden.

Je kan ook de *lines per inch (lpi)* meten van een halve toon druk. Dit is het aantal lijnen bestaande uit afzonderlijke punten waaruit een halve toon druk is opgebouwd. Ten slotte bereken je de dpi met volgende formule:

$$\text{Dpi} = \text{lpi} \times 1,5$$

→ Grote stukken

Een aparte categorie vormen stukken die groter zijn dan het A4-formaat. Grote formaten resulteren in zeer grote digitale archiefkopieën. De National Library of Wales raadt aan om een lagere resolutie te nemen al naargelang het vereiste detail.

Voor kaarten volstaat 200 dpi om leesbare tekst te produceren en 300 dpi om de tekst op een waarheidsgetrouwe manier weer te geven. Een stuk op groot formaat wordt bij voorkeur integraal aan minimaal 300 dpi gedigitaliseerd.⁴⁹

De Cornell Quality Index ⁵⁰

Het onderzoek van de Cornell University resulteerde in een aantal eenvoudige algoritmen waarmee de te behalen resolutie van een digitaal beeld kan worden berekend. De basis van deze formules wordt gevormd door de *quality index*. Dit is een numerieke waardenschaal die door de Cornell onderzoekers werd ontwikkeld om op een eenduidige en objectieve wijze onze waardering van kwaliteit uit te drukken.

Deze methode drukt de relatie uit tussen kwaliteit (QI), de te behalen resolutie (dpi) en de grootte van het kleinste betekenisvolle detail (bijvoorbeeld: de letterhoogte uitgedrukt in millimeter) dat moet worden gedigitaliseerd.

We vermelden deze methode omdat ze de basis vormt van een aantal werktuigen die toelaten om op een snelle en eenvoudige wijze de te behalen resolutie te berekenen. Een voorbeeld is de Image Quality Calculator die ontwikkeld werd door de Digital Services and Development Unit aan de University of Illinois.⁵¹

⁴⁹ LGC-NLW, *Dros Y Rhiniog. Crossing the threshold. The national library of Wales's Digitisation strategy*, Wales, 2004.

⁵⁰ A. KENNEY, O. RIEGER en R. ENTLICH, *Moving theory into practice. Digital imaging tutorial.*, Ithaca, NY, Cornell University Library, 2003, <http://www.library.cornell.edu/preservation/tutorial/conversion/conversion-04.html> (1 juni 2005).

⁵¹ Digital Services and Development Unit, Image Quality Calculator, <http://images.library.uiuc.edu/projects/calculator/> (6 juni 2005).

cDAVID tip: Cornell QI voor gedrukte tekst

Volgens het onderzoek van Cornell University is een QI waarde van 3.0 s ondermaats; 3.6 is eerder slecht, 5.0 is aanvaardbaar en 8.0 is hoge kwaliteit.

Dit zijn de formules die we hanteren om monochroom bitonale (zwart-wit) tekst te digitaliseren:

$$QI = (dpi \times 0,039 \times h)/3$$

$$Dpi = (3 \times QI)/(0,039 \times h)$$

Soms kan een document echter beschadigd zijn of annotaties in kleur bevatten. In dat geval is het noodzakelijk om 8 bits grijswaarden of 24 bits kleur als bitdiepte in te stellen.

De QI formules worden hiervoor aangepast.

$$QI = (dpi \times 0,039 \times h)/2$$

$$Dpi = (2 \times QI)/(0,039 \times h)$$

Heel wat instellingen met expertise in het digitaliseren raden hoe dan ook aan om voor de digitale archiefkopie van een gedrukte tekst een resolutie van 300 dpi of hoger te hanteren.⁵²

cDAVID tip: Cornell QI voor prenten en manuscripten

Voor manuscripten en prenten wordt een andere eenheid gebruikt om de grootte van spatiale details uit te drukken: de spanwijdte (uitgedrukt in millimeter: 'w') van de kleinste details te nemen: afzonderlijke lijnen, strepen, ...

De waarden van de QI werden door Cornell aangepast: 2.0 betekent hoge kwaliteit, 1,5 is goede kwaliteit, 1,0 is twijfelachtig en minder betekent een ondermaatse kwaliteit.

Er wordt een onderscheid gemaakt tussen een QI-formule voor kleuren of grijswaarden en bitonale afbeeldingen.

$$\text{Grijswaarden en kleuren: } dpi = QI/(0,039 \times w)$$

$$\text{Bitonaal: } dpi = (1,5 \times QI)/(0,039 \times w)$$

De resultaten bieden echter geen garantie voor een optimaal resultaat. Zo raden veel instituten aan om een resolutie van 300 dpi en een bitdiepte 8 bits grijswaarden te hanteren wanneer je met manuscripten wordt geconfronteerd.

Ter illustratie: volgende twee tabellen zijn afgeleid op basis van deze formules.

⁵² A. KENNEY, O. RIEGER en R. ENTLICH, *Moving theory into practice. Digital imaging tutorial*.

Table: Representative institutional requirements for conversion, <http://www.library.cornell.edu/preservation/tutorial/conversion/table3-1.html> (21 maart 2005).

Gedrukte tekst / geschatte resolutie voor 1 bit bitdiepte

Grootte van tekst	Detail zichtbaar QI = 3.6	Detail Leesbaar QI = 5	Getrouwe weergave QI = 8
1 mm	300 dpi	400 dpi	600 dpi
2 mm	150 dpi	200 dpi	300 dpi
3 mm	100 dpi	125 dpi	200 dpi
4 mm	70 dpi	100 dpi	150 dpi

Gedrukte tekst / geschatte resolutie voor 8 bit bitdiepte

Grootte van tekst	Detail zichtbaar QI = 3.6	Detail Leesbaar QI = 5	Getrouwe weergave QI = 8
1 mm	150 dpi	200 dpi	300 dpi
2 mm	70 dpi	100 dpi	150 dpi
3 mm	50 dpi	70 dpi	100 dpi
4 mm	35 dpi	50 dpi	80 dpi

Belangrijk: deze methode is slechts een hulpmiddel om kwaliteitseisen terug te koppelen aan de instellingen van de apparatuur.

2.4.3 Het aanmaken van de archiefkopie

Hardware zoals scanners en digitale camera's leveren een digitale output of bitstream. Deze moet reeds alle kwaliteiten hebben van een digitaal bestand dat geschikt is voor bewaring op lange termijn. De centrale regel is dat de bitstream zo min mogelijk mag worden gemanipuleerd tijdens het digitaliseringproces.

Het beeld wordt geïmporteerd in een beeldbewerkingsapplicatie. Hier gebeuren volgende stappen:

- ↪ Het beeld wordt gecontroleerd op kwaliteit (kleurtoon, intensiteit, dynamisch bereik).
- ↪ Het wordt voorzien van de nodige metadata.

- Het wordt opgeslagen in een geschikt archiveringsformaat.⁵³
- Er wordt gecontroleerd of de essentiële metadata correct werden ingekapseld en het gebruikte bestandsformaat conform de specificatie is.
- Het digitale bestand wordt weggeschreven naar een duurzame drager zoals een cd.
- Er wordt gecontroleerd of het bestand correct werd weggeschreven.

Voor de productie van een archiefkopie in een geschikt archiveringsformaat gebruiken we volgende instellingen:⁵⁴

cDAVID aanbeveling: het aanmaken van de archiefkopie.

- *Bestandsformaat*
Baseline TIFF (Tagged Image File Format) versie 6.
- *Compressie*
Geen (uncompressed).
TIFF laat daarnaast compressie via het JPEG- of LZW-algoritme toe. Deze zijn hier echter niet van toepassing. Controleer of ze níet werden gebruikt.
- *Bytevolgorde*
Het TIFF-formaat laat toe de volgorde van de bytes in de bitstream te bepalen. Er zijn twee mogelijkheden. Voor een archiefkopie gebruik je bij voorkeur de "Intel" (of IBM of pc) - volgorde.
- *Metadata inkapselen*
Het TIFF formaat laat toe om in de header (het eerste blok bytes) van het TIFF bestand allerlei extra data over de afbeelding op te slaan. Dit gebeurt door zogenaamde 'tags' of velden in te vullen.

Volgende velden worden best ingevuld:

- Compression (tagnr. 259): 1 (= uncompressed).
- DocumentName (tagnr. 269): bijv. bestandsnaam.
- Make (tagnr. 271): fabrikant van de scanner, camera, ... apparatuur die gebruikt werd om te digitaliseren.
- Model (tagnr. 272): gebruikte scanner.
- Software (tagnr. 305): softwarenaam en versienummer van de software die werd gebruikt.
- DateTime (tagnr. 306): datum van de digitalisering (JJJJ:MM:DD UU:MM:SS).
- Artist (tagnr.315): bijvoorbeeld de naam van de instantie die de collectie beheert of de fotograaf.
- Copyright (tagnr. 33432): houder van het copyright.

Opmerking: een TIFF-document is uitbreidbaar. Je kan zelfgekozen tags of metadatavelden definiëren en die aan het document toevoegen. Dit valt af te raden omdat een aantal applicaties deze zogenaamde *private tags* niet ondersteunt.

⁵³ F. BOUDREZ, DAVID *richtlijn 9: Digitaliseren van analoge archiefdocumenten*. <http://www.antwerpen.be/david/website/teksten/Richtlijn9.pdf> (1 juni 2005)

⁵⁴ F. BOUDREZ, DAVID-*richtlijn 9: Digitaliseren van analoge archiefdocumenten*. (1 juni 2005).

2.4.4 Het controleren van de kwaliteit

Evaluatie van het digitale beeld is van groot belang. Het heeft weinig zin om te digitaliseren zonder te controleren of het digitale beeld een correcte weergave is van het origineel en bruikbaar is om de doelstellingen van het project te behalen.

Bij het controleren van het digitale beeld gaat de aandacht niet alleen uit naar de kwaliteit van het digitale beeld zoals het wordt weergegeven op het beeldscherm of in druk, maar ook naar zaken die een invloed hebben op de leesbaarheid en de bruikbaarheid ervan. Bijvoorbeeld: de kwaliteit van de beschrijving, de keuze van het bestandsformaat of de duurzaamheid van de drager.

De kwaliteitscontrole volgt onmiddellijk op de digitalisering en het importeren van de output in het beeldbewerkingprogramma. Eerst wordt de beeldkwaliteit van de archiefkopie beoordeeld. Voldoet deze, dan wordt het beeld opgeslagen in een geschikt archiverings-formaat: TIFF. Vervolgens worden de verschillende metadata ingekapseld in het digitale beeld. Er volgt een extra controle die nagaat of het gebruikte bestandsformaat voldoet aan de specificatie en of de metadata correct werden opgenomen in het digitale bestand. Ten slotte wordt het digitale bestand weggeschreven naar een duurzame drager zoals cd-r. Ook die wordt aan een kwaliteitscontrole onderworpen vooraleer ze een definitief een plaats krijgt toegewezen in het depot.

Een consequent uitgevoerde controle garandeert de leesbaarheid en de bruikbaarheid van de digitale kopie en ook dat het origineel zo waarheidsgetrouw wordt benaderd. Opnieuw digitaliseren betekent een extra investering in tijd en geld. Bovendien is dit onmogelijk wanneer de originelen na digitalisering werden vernietigd.



Tweede wereldoorlog: piloten schaatsen in het stadspark, december 1944. Foto British Official Foto (stadsarchief Antwerpen, OF 2422).

Controle van de kwaliteit van het digitale beeld

Een kwaliteitsevaluatie kan op twee manieren: op een objectieve manier – aan de hand van concrete parameters – of op een subjectieve manier. Om op een optimale manier een controle uit te voeren worden zowel objectieve als subjectieve methoden toegepast.

De kwaliteitscontrole bestaat uit verschillende onderdelen. We geven een overzicht van de belangrijkste aandachtspunten.

→ *Ontwikkel een consistente aanpak.*

Definieer wanneer een kopie van “aanvaardbare” kwaliteit is en wanneer niet. De finaliteit van de afbeelding speelt hier zeker een rol. Een digitale kopie die goed genoeg is voor reproductie in een publicatie is daarom niet geschikt voor bewaring op lange termijn.

Ook de *intentie* waarmee het beeld werd gedigitaliseerd speelt een rol. Het IPI⁵⁵ onderscheidt in dit verband verschillende intenties. Wil je een exacte weergave van het beeld zoals het er nu uit ziet, de juiste weergave van het originele beeld toen het werd gemaakt (restauratie), een exacte weergave van het originele beeld zoals de fotograaf het bedoelde? Dit houdt in dat je het beeld digitaal manipuleert om storende artefacten eigen aan het originele fotografische procédé te verwijderen.⁵⁶

Ook de keuze van de referentieopname die wordt gehanteerd tijdens de controle speelt een rol. Soms wordt niet het analoge origineel gedigitaliseerd, maar een kopie ervan. Neem je de kopie als basis voor de kwaliteitscontrole? Of grijp je terug naar het analoge origineel?

→ *Baken het terrein af.*

Een doorgedreven kwaliteitscontrole kan soms moeilijk te implementeren zijn als de tijd of de middelen ontbreken. Daarom moet er vooraf worden nagegaan of je letterlijk alle digitale beelden op hun deugdelijkheid zal testen, of dat je slechts een beperkte steekproef (bijvoorbeeld een controle om de 20 digitale beelden).

Er moet ook worden bepaald hoe ver je in de kwaliteitscontrole moet gaan: moet rigoureus aan alle kwaliteitseisen worden voldaan? Of is er een tolerantie tegenover lichte afwijkingen? Hoe ver kan men in die tolerantie gaan?

→ *Wat wordt er gecontroleerd?*

Doorheen de jaren is er heel wat onderzoek gevoerd naar het bepalen van de kwaliteit van digitale beelden. Voor een dergelijk subjectief gegeven is dit geen eenvoudige zaak: er bestaat uiteindelijk een ruime waaier aan modellen en theorieën die toelaten om kwaliteit te beschrijven. Williams reduceert deze tot een vijftal basisparameters: *reproductie van de lichtintensiteit, reproductie van de kleurtoon, resolutie, ‘noise’ en artefacten*.⁵⁷

De kwaliteit van een beeld wordt niet alleen bepaald door de kwaliteit van de digitalisering, maar ook door de mate waarin het beeld correct wordt weergegeven.⁵⁸

⁵⁵ IPI, Image Permanence Institute, Rochester, NY, Rochester Insitute for Technology, <http://www.imagepermanence-institute.org/index.html> (18 maart 2005).

⁵⁶ J.REILLY, F. FREY, *Digital imaging for photographic collections. Foundations for technical standards.*, Image Permanence Institute, Rochester, NY, 1999, http://www.rit.edu/~661www1/sub_pages/8page3g.htm (21 maart 2005).

⁵⁷ Voor een beschrijving van deze elementen: zie paragrafen 2.2.1, 2.2.2 en 2.2.3. D. WILLIAMS, *Image quality metrics*, RLG Diginews 4, nr. 4 (15 augustus 2000), <http://www.rlg.org/preserv/diginews/diginews4-4.html#technical1>.

⁵⁸ A. KENNEY en O. RIEGER, *Using Kodak Photo cd technology for preservation and access. A guide for librarians, archivists and curators.*, Cornell, Cornell University, 1998, <http://www.library.cornell.edu/preservation/kodak/kodak-htm.htm> (2 juni 2005).

- ↪ Monitors verschillen onderling van kwaliteit waardoor de weergave van een gedigitaliseerd beeld grote verschillen kan vertonen.
- ↪ De lichtinval op de monitor speelt ook een rol. In een slecht verlichte ruimte zal een beeld er op een monitor anders uitzien dan in een goed verlichte omgeving. Bovendien wordt het origineel bij voorkeur in een goed verlichte ruimte bekeken terwijl de weergave op een monitor idealiter in een verduisterde omgeving gebeurt.
- ↪ Printers gebruiken een ander kleurenmodel (CMYK) in tegenstelling tot een beeldscherm (RGB). Daardoor kan een afdruk er anders uitzien dan op het scherm. Ook tussen printers onderling zijn er verschillen afhankelijk van de inkt die wordt gebruikt, het procédé waarmee de afbeelding wordt afgedrukt, enzovoort.

Kwaliteitscontrole moet dus gebeuren op apparatuur die op voorhand werd geijkt en in een gecontroleerde omgeving waar storende factoren kunnen worden vermeden. We geven een aantal tips.

***cDAVID aanbeveling: kwaliteitscontrole**⁵⁹*

Ideale controle-omgeving

- ↪ Zorg ervoor dat er geen direct daglicht in de controleruimte is. Verduister de omgeving eventueel (maar niet volledig donker!).
- ↪ Zorg dat slechts de helft van de verlichting aan staat.
- ↪ Richt de monitor naar een wand. Vermijd dat licht van buitenaf (raam, deur) of kunstmatig licht (lampen) de monitor rechtstreeks belichten.
- ↪ Scheid de plaats waar de originelen bekeken worden af van de monitor (dit kan met een simpel wandje).
- ↪ Laat ogen even wennen aan het licht.
- ↪ De lichtbron die het origineel belicht produceert bij voorkeur een fluorescerend daglicht.

Kalibreren van de monitor

- ↪ Gebruik gespecialiseerde hardware en software om de monitor te kalibreren. Indien deze niet voorhanden is bevat bijvoorbeeld: Adobe Photoshop beschikt over een eenvoudige functie om de monitor enigszins te kalibreren.
- ↪ Kalibreer de monitor met volgende instellingen:
 - ↪ Wit punt (kleurtemperatuur) cool white (5000 Kelvin)
 - ~ Bitdiepte: 24 bits
 - ~ Gamma: 2.2
 - ~ Gebruik de gestandaardiseerde Q-60 kleurenkaart van Kodak om de weergave van kleur en lichtintensiteit door de monitor te controleren.⁶⁰

Andere tips

- ↪ Laat de monitor een half uur aan staan zodat het beeld zich stabiliseert.
- ↪ Draag geen felgekleurde kledij.

⁵⁹ A. KENNEY en O. RIEGER, *Using Kodak Photo cd technology for preservation and access.*, 1998, <http://www.library.cornell.edu/preservation/kodak/kodak-htm.htm> (2 juni 2005).

⁶⁰ ISO 12641:1997, *Graphic technology -- Prepress digital data exchange -- Colour targets for input scanner calibration*, <http://www.iso.org> (6 juni 2005).

Eenmaal de omgeving en de hardware goed zijn afgesteld kan je op een goede manier digitale beelden controleren op hun kwaliteit. Daarbij houden we de vijf basisfactoren die de kwaliteit van een digitaal beeld bewaren in het oog.

cDAVID aanbeveling: lichtintensiteit, kleur, resolutie en artefacten

*Lichtintensiteit*⁶¹

→ Wat?

Dit is de mate waarin de weergave van donkere en lichte delen of de lichtintensiteit in een digitaal beeld overeenkomt met een origineel. Aangezien de sensoren van een digitale camera enkel variaties in lichtintensiteit waarnemen is dit een belangrijke parameter voor kwaliteit.

→ Hoe meten?

Gebruik hiervoor een gestandaardiseerde testkaart met neutrale grijswaarden zoals de Kodak Q-60 testkaart. Deze kan voor het eigenlijke digitaliseren worden gebruikt om de scanner te kalibreren.

Een alternatief is de Kodak Q-13 testkaart. Digitaliseer deze samen met het digitale beeld.

Vergelijk visueel het digitale resultaat met de testkaart. Pas indien nodig de instellingen aan van de apparatuur (scanner, digitale camera). Manipuleer het digitale beeld niet door gebruik te maken van de functies, bijvoorbeeld *curves*, in een beeldbewerkingprogramma zoals Adobe Photoshop.

Kleurtoon

→ Wat?

Dit is de mate waarin een kleurtoon correct wordt gedigitaliseerd en weergegeven. Een kleurtoon wordt door de apparatuur gezien en weergegeven als een combinatie van primaire kleuren.

→ Hoe meten?

Gebruik hiervoor een gestandaardiseerde testkaart met kleurtonen. Zowel de Q-60 als de Q-13 en Q-14 van Kodak voorzien testkleuren.

Vergelijk ook het originele stuk met het digitale beeld na het digitaliseren om te bepalen of de kleuren al dan niet overeenkomen.

Resolutie

→ Wat?

Dit is de mate waarin kleine details worden onderscheiden tijdens het digitaliseren en correct worden weergegeven. Hoe hoger de resolutie, hoe meer details het digitale beeld zal bevatten.

>>

⁶¹ D. WILLIAMS, *Selecting a scanner*, in: Guides to quality in visual resource imaging, Washington, DC: Council on Library and Information Resources, 2000, <http://www.rlg.org/legacy/visguides/visguide2.html#background> (22 maart 2005).

>>

- Hoe meten?

Gebruik de RIT Alphanumeric Resolution Test Object kaart om dit te bepalen. Deze wordt afzonderlijk gescand om vast te stellen in welke mate de apparatuur fijne details kan onderscheiden.

Het is ook mogelijk om de resolutie te meten aan de hand van de MTF of Modular Transfer Function.⁶² Met gespecialiseerde software zoals ontwikkelt door MITRE⁶³, kan je bepalen of de apparatuur een voldoende hoge resolutie haalt.

*Artefacten*⁶⁴

Artefacten zijn zichtbaar in het beeld aanwezig en worden als storend ervaren. Vaak vallen ze te vermijden.

- Aliasing en Moiré

De meeste apparatuur heeft een automatische of manuele descreening-optie die moiré uitfiltert. Aliasing kan vermeden worden door met een voldoende hoge resolutie te digitaliseren.

- Krassen en stof

Stof kan het digitale beeld storen. Hou daarom de apparatuur en de werkplek stofvrij. Een goede scanner kan krassen op het originele analoog automatisch filteren.

- Noise en kleurafwijkingen

Noise valt vooral op in egale kleurvlakken als pixels met een afwijkende kleur. Noise is inherent aan de onderdelen van de apparatuur. Degelijke apparatuur kan dit deels onderdrukken. Investeer daarom in goede toestellen.

Ook kleurvariaties vinden ten dele hun oorsprong in de kwaliteit van de gebruikte apparatuur. Bijvoorbeeld kleurkanalen die elkaar slecht overlappen. Dit uit zich in een gekleurde "schaduw" van de elementen in het gedigitaliseerde beeld.

*Andere aandachtspunten*⁶⁵

We vermelden nog een aantal elementen die nefast zijn voor de beeldkwaliteit.

- Detailverlies in lichte of donkere gebieden van het beeld. Dit is een aanduiding dat de dynamic range van de digitaliseringapparatuur niet volstaat.
- Een onscherp beeld.
- Een beeld dat te scherp is. Hierdoor kunnen opvallende halo's ontstaan rond scherp afgelijnde randen.

De mate waarin details worden weergegeven. Schenk aandacht aan de weergave van fijne lijntjes, kleine letters, enzovoort.

⁶² D. WILLIAMS, *What is an MTF... and why should you care?*, RLG Diginews 2, nr. 1 (15 februari 1998), <http://www.rlg.org/preserv/diginews/diginews21.html#technical> (6 juni 2005).

⁶³ MITRE, *MTF Image Quality Evaluation*, 2005, <http://www.mitre.org/tech/mtf/> (6 juni 2005).

⁶⁴ T. SLEEP, *Flaws in film scanners*, Londen, 1999, <http://www.cix.co.uk/~tsphoto/tech/filmscan/flaws.htm> (6 juni 2005).

⁶⁵ NINCH, *The NINCH Guide to Good Practice*, 2002, <http://www.nyu.edu/its/humanities/ninchguide/VIII/> (6 juni 2005).

Het evalueren van de kwaliteit van een beeld is verre van eenvoudig en vraagt enige ervaring en oefening. **Idealiter gebeurt de kwaliteitscontrole en het kalibreren van de apparatuur door iemand met kennis van zaken.**

Controle van de kwaliteit van het digitale bestand

De archiefkopie wordt tijdelijk bewaard op een veilige plaats: een beveiligde folder op een netwerkschijf of lokale harde schijf van de machine waar het digitale bestand werd aangemaakt. Vooraleer alles weg te schrijven naar een duurzame drager, moet er een controle worden uitgevoerd op de ingekapselde metadata en de conformiteit van het gebruikte bestandsformaat met de specificatie.

cDAVID aanbeveling: controle van het digitale bestand

- Controleer of het TIFF-bestandsformaat valid and well formed.
Het TIFF-bestand moet voldoen aan de TIFF-specificatie zoals vastgelegd in de ISO standaard. Het bestandsformaat van een TIFF-bestand kunnen we valideren met JHOVE.⁶⁶
- Controleer of alle metadata aanwezig zijn.
Er moet worden gecontroleerd of de metadata die werden ingevoerd ook daadwerkelijk zijn ingekapseld in het bestand. Dit kan met een programma zoals Leadtools TIFF tag editor.⁶⁷ Controleer of de essentiële metadata zoals de DocumentName en het copyright zijn ingevuld.

Sommige programma's zoals *Adobe Photoshop* hebben de neiging om automatisch private tags toe te voegen aan de kopie. Hoewel dit op termijn geen kwaad kan – de TIFF-specificatie stelt dat private tags de leesbaarheid van de bitstream niet in het gedrang mogen brengen – is het toch beter om dit te vermijden.

Tenslotte wordt de archiefkopie op een duurzame drager gebrand. Als duurzame dragers beschouwen wij optische dragers zoals cd's of magnetische dragers zoals DLT-tapes. Opslag op harde schijven is een dure en risicovolle oplossing. Diskettes zijn dan weer helemaal niet geschikt. Voor de keuze van een geschikte duurzame drager verwijzen we naar de richtlijnen geformuleerd door het DAVID-project.⁶⁸

⁶⁶ JHOVE-JSTOR, Harvard Object Validation Environment, <http://hul.harvard.edu/jhove/> (1 juni 2005).

ISO-12639: "Graphic technology -- Prepress digital data exchange -- Tagged image file format for image technology", <http://www.iso.org> (27 mei 2005).

⁶⁷ Leadtools TIFF tag editor, http://www.leadtools.com/Utilities/TIFF_Editor/TIFF-Editor.htm (1 juni 2005).

⁶⁸ F. BOUDREZ, *DAVID-richtlijn 2. Duurzame cd's*, <http://www.antwerpen.be/david/website/teksten/Richtlijn2.pdf>, (1 juni 2005).

F. BOUDREZ, *DAVID-richtlijn 6. Duurzame magnetische dragers*, <http://www.antwerpen.be/david/website/teksten/Richtlijn6.pdf> (1 juni 2005).

2.4.5 *Het optimaliseren van de kopie*

Er kan overwogen worden om op basis van de archiefkopie een geoptimaliseerde kopie te maken. Het digitaal restaureren houdt in dat je via beeldbewerkingsoftware het origineel beeld terug tracht te herstellen in zijn oorspronkelijke staat. Dergelijke optimalisatie mag echter niet op de archiefkopie gebeuren. De beschadigingen vormen immers een integraal deel van het origineel en hebben als dusdanig een historische waarde.

Optimalisatie veronderstelt dat diegene die het stuk digitaliseert kennis van zaken en de nodige ervaring heeft in beeldbewerking. Optimalisatie houdt onder andere volgende zaken in:

- a. aanpassen van de grootte van het digitale beeld: cropping, orientatie, perspectief, ...
- b. aanpassen van de intensiteit van het beeld: witbalans, levels en curves
- c. schoonmaken: het verwijderen van allerlei mogelijke artefacten en storingen in het digitale beeld (ruis, moiré, vlekken, ...).

cDAVID aanbeveling: de geoptimaliseerde archiefkopie

Net als bij de archiefkopie moet er bij de geoptimaliseerde archiefkopie rekening worden gehouden met bewaring op lange termijn.

Hier gelden net dezelfde eisen als voor het produceren van een archiefkopie met het 'originale' digitale beeld.

We blijven een platformonafhankelijk en open bestandsformaat gebruiken dat als een standaard wordt erkend en de drager die om de geoptimaliseerde kopie te archiveren moet geschikt zijn voor bewaring op lange termijn.

2.4.6 *Het produceren van afgeleide kopieën.*

Na de productie van de archiefkopie volgt die de van *afgeleide kopieën*. Dit zijn beelden die op basis van de archiefkopieën worden gemaakt. In tegenstelling tot deze laatste worden ze gebruikt om het beeld te distribueren onder het brede publiek.

Eerst moet je de afweging maken of je de archiefkopie met zijn tekortkomingen of een geoptimaliseerde kopie gebruikt. Ook dit hangt af van de doelstellingen van het project. Een beeldbank met gerestaureerde beelden zal gebaat zijn bij het gebruik van de geoptimaliseerde archiefkopie als basis van de distributie-exemplaren.

De archiefkopie / geoptimaliseerde kopie is doorgaans een groot bestand van ettelijke megabytes. Het verspreiden van dergelijke bestanden is verre van efficiënt. De opslag op een online server en de benodigde bandbreedte om archiefkopieën snel en gebruiksvriendelijk te raadplegen maakt dit al snel een dure aangelegenheid. Bovendien spelen ook auteursrechten een belangrijke rol: instellingen mogen immers niet zomaar hun foto's ter beschikking stellen. Daarvoor is de toestemming van de belanghebbenden nodig. De hoge kwaliteit van een archiefkopie maakt dat deze gemakkelijk opnieuw kan worden gebruikt door derden bijvoorbeeld in publicaties.

Bij de productie van een afgeleide kopie wordt er gebruik gemaakt van compressietechnieken om een bestand te krijgen dat slechts een fractie van de grootte is van de originele archiefkopie.

cDAVID aanbeveling: de reproductie image

Reproductie-image

Dit zijn digitale distributiekopieën die geoptimaliseerd zijn voor gebruik in publicaties of reproducties. De kwaliteit van deze reproductiekopieën is quasi gelijk aan die van de archiefkopie. Van belang is de keuze van het correcte kleurenprofiel. Printers en drukmachines werken doorgaans met het CMYK-model. Het is dan ook van belang dat tussen diegene die de afbeeldingen aanlevert en de drukker of de uitgever er onderlinge afspraken worden gemaakt wat de keuze van het kleurenprofiel betreft.

- Bestandsformaat: TIFF, EPS of PDF.
- Kleurenprofiel: te bepalen in onderlinge afspraak.

Metadata: niet vereist tenzij uitdrukkelijk gewenst.

cDAVID aanbeveling: de referentie image

Referentie image

Dit zijn digitale beelden die je doorgaans op websites aantreft. Het gaat om relatief kleine afbeeldingen die slechts een paar kilobyte groot zijn. De compressie-ratio die ze hanteren is dan ook eerder groot. Deze afbeeldingen dienen om een indruk te geven van het beeld op een monitor binnen een browser. De lengte van een web image is nooit groter dan 600 pixels.

- Bestandsformaat: JPEG (Joint Picture Experts Group) met voldoende compressie of PNG (Portable Network Graphics).⁶⁹
- Kleurenprofiel: sRGB IEC 61966-2-1.
- Voeg een duidelijk zichtbaar watermerk toe aan het beeld. Het watermerk bevat een copyright notice.

!! Veel beeldbewerkingsprogramma's bieden het toevoegen van metadata aan door gebruik te maken van een IPTC (International Press and Telecommunications Council) wrapper. Deze behoren echter niet tot de JPEG-standaard. Het risico bestaat dat bij een conversie naar een nieuwe versie van het JPEG-bestandsformaat of migratie naar een nieuw bestandsformaat de toegevoegde metadata verloren gaan !!

⁶⁹ W3C, *Graphics on the web*: <http://www.w3.org/Graphics/> (6 juni 2005).

Kleinere bestanden zijn veel gemakkelijker te verplaatsen en kosten minder aan opslag. Het nadeel is dat ze van minder goede kwaliteit zijn dan de archiefkopieën, wat hen ongeschikt maakt voor bewaring op lange termijn. Een distributie-exemplaar kan nooit de archiefkopie vervangen. Het voordeel is dat ze ook niet zo geschikt zijn voor gebruik in publicaties. Door het toevoegen van een watermerk kan men bovendien snel inbreuken op de auteursrechten vast stellen.

cDAVID aanbeveling: de thumbnail image

De thumbnail image

De thumbnail image is een nog kleiner digitaal beeld. Maximaal 100 pixels breed. Thumbnails zijn verkleinde versies van referentie-images. Ze worden gebruikt in de presentatie van zoekresultaten. Wanneer een gebruiker een query lanceert, levert de beeldbank de zoekresultaten af in de vorm van een tabel met thumbnails. Door op een thumbnail te klikken krijgt de gebruiker de referentie-image te zien.

- Bestandsformaat: JPEG (Joint Picture Experts Group) met voldoende compressie of PNG (Portable Network Graphics).⁷⁰
- Kleurenprofiel: sRGB IEC 61966-2-1.

!! De taal waarmee webpagina's worden opgebouwd (XHTML) laat toe om afbeeldingen te verkleinen of te vergroten. Dit is echter maar schijn. De afbeelding wordt weliswaar kleiner getoond maar de hoeveelheid data die moet worden gedownload is nog altijd even groot. Het is verleidelijk om geen aparte thumbnails te maken en via HTML code de referentie-images verkleind weer te geven. Dit valt sterk af te raden. Een scherm met "thumbnails" zal immers veel trager inladen omdat er veel meer data moet worden binnengehaald.

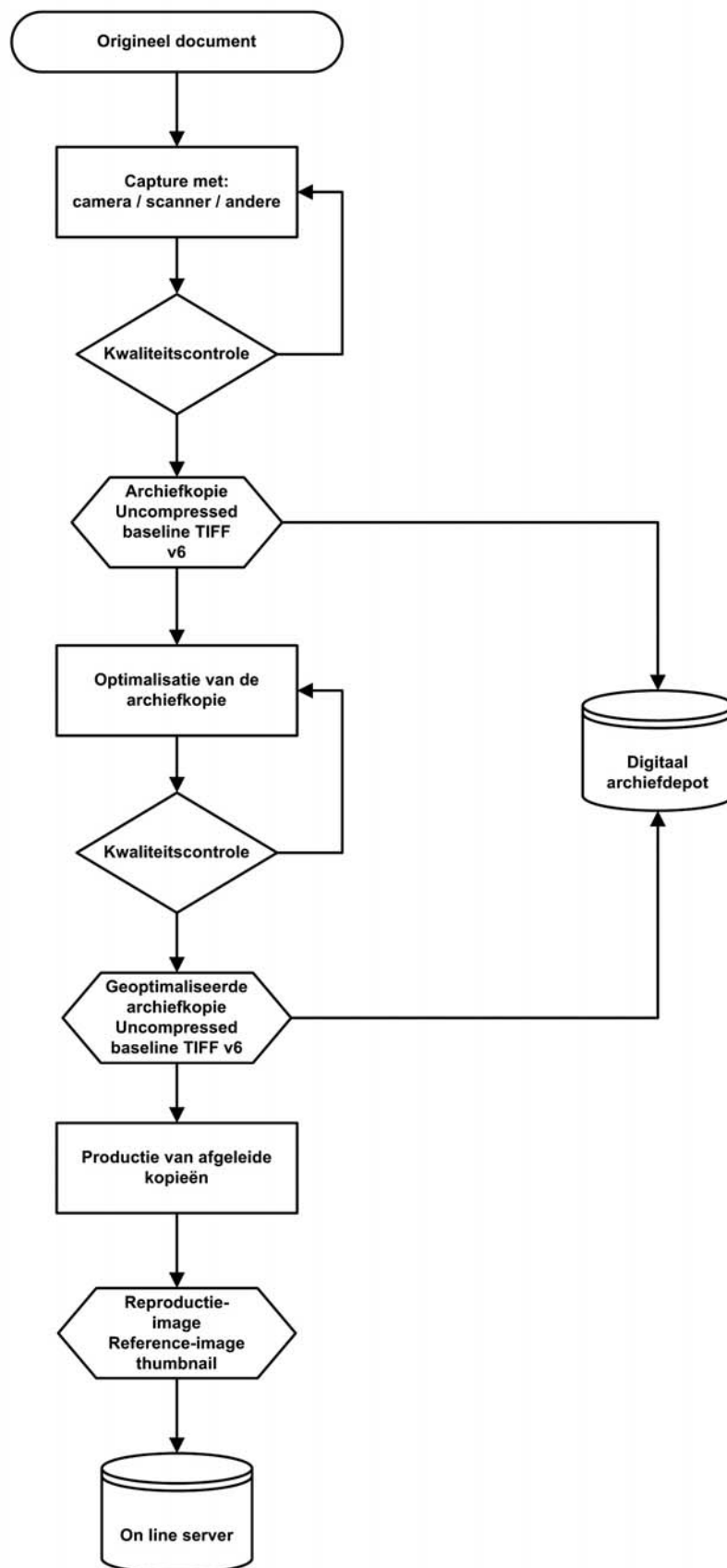
2.4.7 De nazorg

Net zoals bij de voorbereiding is ook de nazorg een moment van belang in de workflow. Afhankelijk van de beslissing van de erfgoedbeheerder worden de stukken klaargemaakt voor vernietiging of het terugplaatsen op de bewaarplaats. In laatste instantie spreekt het voor zich dat de nodige maatregelen worden genomen om het risico op beschadigingen tot een minimum te beperken. Tenslotte is de nazorg het moment bij uitstek om stukken opnieuw te verpakken. Vaak volstaat de oude verpakking al lang niet meer of waren de stukken in het slechtste geval zelfs helemaal niet verpakt. Een duurzame bewaring van de analoge originelen naast de digitale kopieën veronderstelt ook het gebruik van duurzame materialen.

Een digitaliseringproject laat de erfgoedbeheerder toe om een betere inschatting te maken van de waarde van de stukken die hij / zij bewaart en beheert. Indien mogelijk en wenselijk kan je zelfs de volledige bewaring van het originele audiovisuele materiaal herbekijken. Interessant is het werk van Peter Adelstein dat toelaat om snel een inschatting te maken van de bewaring van dit materiaal.⁷¹ Het valt altijd aan te raden om het advies van een expert in te winnen.

⁷⁰ W3C, *Graphics on the web*.

⁷¹ P.Z. ADELSTEIN, *Image Permanence Institute Media Storage Quick Reference*, Image Permanence Institute (IPI), Rochester Institute of Technology, Rochester, <http://www.climatenotebook.org/MSQR/MSQR.pdf> (22 maart 2005).



Model workflow van een digitaliseringsproject.



Stadsfeestzaal, versierd voor het Black en White bal, maart 1948.

Foto A. Lammens (stadsarchief Antwerpen, PEO 3871).

3 Case study: de European Visual Archive beeldbank

3.1 Wat is EVA?

3.1.1 European Visual Archive - EVA, 1998 - 2001

Heel wat erfgoedinstellingen beheren fotocollecties met een grote historische waarde. Het European Visual Archive (EVA) project onderzocht hoe deze collecties voor het publiek met nieuwe technieken konden worden opengesteld. Het doel van het EVA-project was tweeledig. Ten eerste wilde men de fotocollecties van twee Europese archieven, het stadsarchief Antwerpen (SA) en de London Metropolitan Archives (LMA) via het internet toegankelijk maken. Ten tweede wensten de betrokken instellingen expertise rond het digitaliseren van historische fotocollecties verder uitdragen.

Zes Europese partners participeerden aan het EVA-project: Telepolis (nu Digipolis) Antwerpen, de London Metropolitan Archives, het stadsarchief Antwerpen, de European Commission on Preservation and Access (ECPA), Sail Labs, en het Nederlands Instituut voor Wetenschappelijke Informatiediensten (NIWI). Het EVA-project maakt deel uit van het INFO2000-initiatief van de Europese commissie en liep van december 1998 tot februari 2001.⁷²

Met het EVA-project wilden de deelnemers oplossingen aanreiken voor een aantal problemen waarmee men binnen dit kader wordt geconfronteerd. Kwamen onder meer aan bod: auteursrecht, selectie van het te digitaliseren archiefmateriaal, gebruikersonderzoek, geschikte technieken voor digitalisering, beschrijving van gedigitaliseerde foto's, ontwikkeling van digitale informatiesystemen en commercialisering van het aangeboden materiaal.

Bovendien was het systeem zo ontwikkeld dat de drempel voor andere instellingen om mee in het project te stappen voldoende laag lag. Er werd gekozen om zoveel mogelijk aan te sluiten bij de reeds bestaande situatie binnen de deelnemende archieven.

Het EVA-project resulteerde in een werkende website waar 20 000 beelden uit de fotografische collecties van de deelnemende archieven (stadsarchief Antwerpen en London Metropolitan Archives) konden worden opgevraagd, afgeprint of besteld. De ondervraging van de beeldbank kon in het Engels, Nederlands of Duits verlopen.

De website stelde eveneens alle rapporten ter beschikking met de resultaten van het onderzoek dat tijdens het project werd gevoerd.

⁷² R. VAN HORIK, Archives and photographs: the European Visual Archive (EVA) project. In: *Cultivate Interactive*, nr. 3, 2001, <http://www.cultivate-int.org/issue3/eva/> (7 juni 2005).
S.N., Algemene informatie over de EVA- en EVAmp-projecten, <http://www.eva-eu.org/> (7 juni 2005).

3.1.2 European Visual Archive Market Validation project - EVAmP, 2003 - 2004

Het eerste project had duidelijke resultaten opgeleverd. Het ontwikkelde prototype van webinterface was succesvol voor de deelnemende archiefinstellingen. De partners en de EVA-website hadden ook mogelijkheden geboden aan andere erfgoedinstellingen om het EVA-consortium te vervoegen en beeldmateriaal via de EVA-beeldbank aan de bieden aan het Europese publiek.

Omdat geen nieuwe partners of erfgoedbewaarders zich in de periode 2001 - 2002 hadden aangemeld, werd een tweede project opgestart: het EVAmP-project, acroniem voor *EVA Market Validation Project*. Na een bevraging van de erfgoedinstellingen en een analyse van de markt van vooral professionele gebruikers van archieffoto's, moest de bestaande EVA-website kunnen uitgroeien tot een zelfstandig werkende beeldbank waarin meerdere archieven, musea en bibliotheken diensten aanbieden ten aanzien van de gedigitaliseerde beeldverzamelingen in hun bezit.

EVAmP, gesteund door het EU Ten-Telecom programma, liep van februari 2003 tot december 2004 (*eTEN : Information Society Directorate General; Trans-European Telecommunications Networks*).

Drie nieuwe partners, de Koninklijke Bibliotheek van België, het bedrijf PICTURA (Nederland) en het onderzoeksinstituut NDAD University of London Computer Centre, vervoegden de partners van het eerste project, de Londense en Antwerpse archieven LMA en SA, Digipolis Antwerpen en het Nederlands Instituut voor Wetenschappelijke Informatiediensten (NIWI).

De resultaten van dit tweede project werden toegevoegd aan de eerste website en zo ter beschikking gesteld van de sector en het veld.

3.2 Planning en aanpak

Vooraf plannen is noodzakelijk om alle vooropgestelde doelstellingen te kunnen behalen. Zonder een duidelijke strategie bestaat het risico dat projecten zoals EVA of EVAmP het uiteindelijke doel niet of slechts ten dele zouden halen.

Dergelijke projecten worden dan ook opgedeeld in workpackages die uitgebreid worden gedocumenteerd en elk heel duidelijke resultaten moeten opleveren. Standaard worden steeds afzonderlijke workpackages voorzien voor het projectbeheer en voor *dissemination*, verspreiden en publiceren van de onderzoeksresultaten. Het EVA-project streefde er immers naar om de verzamelde expertise naar belanghebbende archiefinstellingen uit te dragen en de gemeenschap van archieven en erfgoedinstellingen optimaal te laten gebruik maken van het gevoerde onderzoek.⁷³

De inhoudelijke onderdelen van EVA waren:

- *Content and data analysis*: analyse van de te digitaliseren collectie en beschikbare beschrijvingen, en van de wensen en noden van de gebruikers.

⁷³ Aan de hand van de resultaten van de verschillende workpackages zijn een aantal Engelstalige rapporten afgeleverd die vrij te verkrijgen zijn op de EVA-website, <http://192.87.107.12/eva/nl/documents.htm> (7 juni 2005).

- *Quality and standards research*: bepalen van de kwaliteit en standaarden van de digitale beelden en de metadata.
- *Property and copyrights agreements*: onderzoek inzake auteursrechterlijke bescherming en mogelijke beperkingen.
- *Input collection*: het digitaliseren en beschrijven van de foto's zelf. Dit moest gebeuren volgens de aanbevelingen van de vorige drie werkpakketten.
- *Working model*: ontwerpen en uitbouwen van de EVA-website.

EVAmp was onderverdeeld in de volgende onderzoeksdoelstellingen:

- Marktonderzoek naar de belangstelling van erfgoedbewaarders om diensten en / of producten ten aanzien van de eigen beeldverzamelingen aan te bieden via een beeldbank.
- Gebruikersonderzoek naar de tevredenheid over de zoekmogelijkheden van de huidige EVA-website.
- Ontwikkelen van een bedrijfsplan zodat zou blijken op welke basis en tegen welke voorwaarden op een kosteffectieve wijze producten en diensten konden worden aangeboden.
- Verfijning van de werkende website, de interface en zoekmogelijkheden en de achterliggende databestanden. Deze verfijning maakt gebruik van het gebruikersonderzoek en het business plan.

3.3 Doelstellingen

Het EVA-project vertrok vanuit volgend standpunt: "de historische fotocollecties in de archieven hebben een groot historisch belang die het louter regionale niveau overstijgt. Toch zijn fotoarchieven vaak de minst toegankelijke bronnen binnen een instelling. Hun complexe ordening, de slechte fysieke staat en de rudimentaire beschrijving verhinderen het maximaliseren van hun waarde. Naar valorisatie toe is een bredere toegankelijkheid en een grotere beschikbaarheid van dit materiaal dus wenselijk."⁷⁴

Concreet wenste men met het EVA-project het volgende te bereiken binnen het domein van de digitale bewaring, het on line ontsluiten en aanbieden van erfgoed.⁷⁵

⁷⁴ S. OSTROW, *Digitizing Historical Pictorial Collections for the Internet*, Washington / Amsterdam, CLIR, 1998.
R. VAN HORIK, Archives and photographs and the European Visual Archive (EVA) project.

⁷⁵ S.N, *The European Visual Archive - EVA -: Access to European Heritage through Digital Preservation of Archival Photographic Collections*, 1998, <http://192.87.107.12/eva/nl/texts/proposal.pdf> (7 juni 2005).

- Standaarden voor het beschrijven van digitale afbeeldingcollecties.

Een gebruiker moet vanuit één punt de verschillende collecties en aangeboden diensten van de deelnemende instellingen kunnen benaderen. Daarvoor is er nood aan een eenduidige standaard voor het beschrijven van digitale objecten. In de praktijk bestaan er echter verschillende modellen en zijn sommige collecties beter ontsloten dan anderen waardoor een transparante benadering van data slechts ten dele mogelijk is.

Het EVA-project probeerde voor dit probleem een oplossing aan te reiken in de vorm van een toolbox voor het ontsluiten en ophalen van fotografische collecties. Centraal staat daarin de implementatie van een standaard voor het beschrijven van digitale afbeeldingen. Een aantal bestaande beschrijvingsstandaarden die daartoe in aanmerking kwamen zijn Dublin Core, EAD, ISAD(G) en ISAAR(CPF).

Dit onderzoek werd opgenomen in workpackage 3 (*Quality and standards research*). Inzake het beschrijven werd één standaard voorgesteld die door alle instellingen gehanteerd kon worden.

- *Het bepalen van de optimale kwaliteitseisen van een digitale afbeelding met het oog op lange termijnbewaring.*

Het digitaliseren van erfgoed biedt nieuwe perspectieven zoals het bewaren van waardevol maar fragiel materiaal. Digitale objecten beperken zich echter niet tot één formaat of één medium. Bij conversie of migratie naar andere formaten moet generatieverlies – een vorm van nadelig informatieverlies dat optreedt bij het omzetten of kopiëren van digitale data – worden vermeden. Tijdens het digitaliseren moet zoveel mogelijk informatie van het origineel digitaal worden opgeslagen om deze op een gepaste manier weer te geven.

De heersende paradox tussen de wens om data op lange termijn te bewaren en de snelle technologische veroudering vormt echter een probleem. Er was nood aan geschikte methoden en *metrics* om digitale foto's van hoge kwaliteit op te leveren die over een langere periode kunnen worden bewaard. Het EVA-project wilde expertise verzamelen. EVA probeerde daartoe een aantal kwaliteitsnormen te definiëren voor het digitaliseren van fotografische bronnen. Men wilde ook een aantal aanbevelingen opstellen voor de keuze en de aanschaf van apparatuur en infrastructuur.

Dit onderzoek werd opgenomen in workpackage 3 (*Quality and standards research*). Inzake de kwaliteit van de digitale afbeeldingen werd een algemene richtlijn afgeleverd betreffende de kwaliteit waaraan de door de deelnemende instellingen aangeleverde afbeeldingen moesten voldoen.

- *De implementatie van moderne technologie om (meta)data te doorzoeken, op te halen uit gedistribueerde databanken en te presenteren.*

Naast het gebruik van gestandaardiseerde tools voor het beschrijven en aanleveren van gedigitaliseerde foto's moest er ook een systeem worden ontwikkeld voor het online aanbieden van de digitale collectie. EVA wilde een prototype van een informatiesysteem aanleveren dat op basis van nieuwe technologie haar doelpubliek ongeacht de plaats of de tijd toegang tot de beeldbank aanbood.

Het systeem moest gebruiksvriendelijk zijn en toelaten om dynamisch doorheen de verschillende digitale collecties te zoeken ongeacht de locatie waar deze worden bewaard of

het moment waarop het systeem wordt benaderd. Een website als interface was de oplossing waarvoor is geopteerd.

Dit onderdeel werd opgenomen in workpackage 6 (*Working model: the EVA system*) en beschrijft de basisprincipes en de functionele eisen van het EVA-systeem. Deze workpackage bevat volgende zeven stappen.

1. Het uitvoeren van een analyse van de verschillende functionele aspecten die geïncorporeerd moesten worden in het systeem. Het ontwikkelen van een functioneel ontwerp op basis van deze analyse.
2. Het ontwerpen van een geschikt databankmodel en de ontwikkeling van een applicatie die toelaat deze databank te beheren.
3. Het ontwikkelen van de verschillende tools die toelaten de data in het EVA-systeem te benaderen.
4. De integratie van technologie die *on-the-fly*, vertaling van de voorkeurstaal van de gebruiker, naar de taal van het EVA-systeem toelaat.
5. Het uitbouwen van de EVA-website. Deze website diende een on line front-end te vormen die toelaat om de verschillende collecties op een transparante en eenduidige manier te benaderen ongeacht plaats en tijd.
6. De ontwikkeling van een interface naar de lokale toegangen toe. Een query of zoekopdracht van een gebruiker zou naar een logische expressie omgezet worden waarmee de gewenste informatie uit de beschikbare lokale toegangen kan worden opgehaald.
7. Het uitbouwen van een operationeel prototype. Het resulterende werkmodel kan men uitdragen naar belanghebbenden die fotocollecties beheren.

→ *Het probleem van de taalverschillen over de grenzen heen.*

Taalverschillen vormen een probleem voor een dergelijk internationaal project. Het EVA-project wilde de fotocollecties van de verschillende participerende archiefinstellingen op internationaal niveau toegankelijk maken.

De bedoeling was dat de beschrijvingen werden gemaakt in de taal van de archiefinstelling die de foto's beheert. Voor het publiek en de gebruikers mocht de taal echter geen barrière zijn, zeker niet binnen de Europese context van dit project: een burger van een ander Europees land zou de mogelijkheid krijgen om in zijn of haar eigen taal te zoeken in de beschrijvingen van de foto's. De projectpartners meenden dat de drempel voor archieven om mee in te stappen hoog zou zijn indien elk archief ook alle gemaakte beschrijvingen zou moeten vertalen.

Het prototype dat het EVA-project bouwde moest de aangeleverde data automatisch presenteren in de taal van de gebruiker. Het systeem moest van een degelijke transparante vertaalinterface voorzien worden.

→ *Coöperatie met commerciële organisaties verbeteren.*

Met de opkomst van informatietechnologie ontstonden er heel wat nieuwe manieren om met data en ontsluiting van data om te gaan. Archiefinstellingen werden bovendien

geconfronteerd met een stijgende vraag om gebruik te maken van deze nieuwe diensten. Vaak moet er daartoe beroep gedaan worden op gespecialiseerde bedrijven omdat men de kennis niet in huis heeft. Het EVA-project wilde deze relatie tussen de publieke en private sector vernieuwen.

→ *Auteursrechten en kosten.*

Auteursrechten en marketing vormen een belangrijk aspect bij het aanbieden van informatie aan het publiek en de professionele mediasector. Het EVA-project wilde antwoorden brengen op vragen omtrent copyright en e-commerce. Men verzag dat de kosten om het publiek en de professionele sector toegang te verlenen tot de beeldbank lager zouden liggen dan voor een traditionele analoge dienstverlening. Het leeuwendeel van de fotocollecties in archiefinstellingen wordt beschermd door het auteursrecht. Het digitaliseren en online aanbieden van deze foto's kan niet zonder de toestemming van eventueel nog levende belanghebbenden.

Workpackage 4 (Property and copyrights agreements) onderzocht de verschillen tussen de deelnemende landen wat betreft de bescherming van het intellectuele eigendom en moest aanbevelingen verstrekken.

3.3.1 *Het bepalen van de prioriteiten*

De eerste opdracht, ook zo vertaald in het eerst aan te pakken werkpakket (*workpackage 2: Content and data analysis*), betrof het uitvoeren van de analyses naar de beelden, data en metadata die zouden worden opgenomen in de EVA-beeldbank. Om dat te bepalen werd een enquête gehouden onder de verschillende potentiële gebruikersgroepen en de archiefinstellingen. Het resultaat is een prioriteitenlijst en een overzicht van de verschillende fotocollecties en data die in aanmerking kwamen voor digitalisering en een identificatie van het doelpubliek.

De verschillende beslissingen en keuzes die moesten worden genomen tijdens de planning en de aanpak van het project werden gevisualiseerd in een decisionmaking matrix. Dit is een schema of flowchart waarin cruciale beslissingen in kaart worden gebracht. Voor het EVA-project worden zeven verschillende stappen onderscheiden: de keuze van de *soort* en het *type* foto dat wordt gedigitaliseerd, de *technische kenmerken* van de digitale kopie, de te gebruiken *beschrijvings- of metadata-schema's*, de *architecturale opbouw* van het EVA-systeem, de aangeboden zoekmogelijkheden en een *e-commerce-aanbod*.

Hierna vindt u de *Decision making matrix* van EVA



7. E-Commerce	Standaard internet	Eigen systeem
6. Zoekmogelijkheden	Alle velden	Onderwerp, datum, fotograaf
5. Systeemarchitectuur	Centraal	Lokaal
4. Metadata/beschrijvingen	Dublin core	ISAD(G)
3. Digital Image	Referentie-image/thumbnail	Master image
2. Type foto	Andere	Zwart-wit foto's
1. Soort foto	Kunstfoto's	Documentaire foto's

3.3.2 Het analyseren en bepalen van het doelpubliek

EVA, 1998 - 2001

Een eerste stap was de analyse van de gegevens die voorhanden waren. Dit resulteerde in een overzicht van de vragen en knelpunten die binnen de archiefwereld voelbaar zijn wat betreft het digitaliseren van foto-opnamen. Een tweede analyse ging na wie het doelpubliek was en in welke onderwerpen dit zich interesseerde. Op basis van dit laatste onderzoek zijn een aantal subject groups of onderwerpen geïdentificeerd. De fotocollecties werden ingedeeld per onderwerp. Ten slotte bleek uit dat de verschillende motieven die binnen de archiefinstellingen leven om over te gaan tot digitalisering, een grote impact op de technische kwaliteitseisen die bij het eigenlijke digitaliseren werden gehanteerd.

Om te achterhalen wie geïnteresseerd is in de historische fotocollecties legden de London Metropolitan Archives (LMA) en het stadsarchief Antwerpen (SA) bij de aanvang van het EVA-project in 1998, een enquête voor aan hun bezoekers. Het was de bedoeling te achterhalen in welke onderwerpen ze geïnteresseerd waren en hun mening te weten te komen over de mogelijkheden om een afdruk van een foto mee naar huis te kunnen nemen. De onderzoeksresultaten en conclusies van deze enquête zijn terug te vinden op de EVA-website.⁷⁶ In het kader van het EVAMP-project werd er in 2003 een nieuwe bevraging onder de EVA-gebruikers ingericht.⁷⁷

In 1998 werd een enquête verstuurd naar iedereen die een reproductie had besteld in dat jaar. Het stadsarchief Antwerpen ontving 32 antwoorden, de London Metropolitan Archives 84. Deze

⁷⁶ EVA CONSORTIUM, *The European Visual Archive - EVA, Analysis of user consultations*, 1999, <http://192.87.107.12/eva/nl/texts/wp23.pdf> (7 juni 2005).

⁷⁷ EVAMP, *European Visual Archives Market Validation Project*, <http://www.EVAMP.org> (7 juni 2005).

EVAMP, *user survey report*, Digipolis, Antwerpen, 2003, [http://www.EVAMP.org/documents/EVAMP_31103_WP2_D2-2_SurveyReport_v1-2\(final\).pdf](http://www.EVAMP.org/documents/EVAMP_31103_WP2_D2-2_SurveyReport_v1-2(final).pdf) (24 maart 2005).

discrepanctie impliceert dat we de conclusies van een vergelijkend onderzoek met de nodige terughoudendheid in beschouwing moeten nemen.

De Antwerpse archiefbezoeker consulteerde het fotoarchief uit persoonlijke interesse. Een significant aantal bezoekers van de London Metropolitan Archives deed dit echter om professionele redenen. De helft van de bezoekers consulteerden het Antwerpse fotoarchief voor de eerste keer terwijl de deelnemers in Londen occasioneel opzoekingen verrichtten in de fotocollecties. Hierbij moeten we onthouden dat bezoekers die om professionele redenen fotoverzamelingen benaderen andere wensen hebben dan diegenen die dit uit louter persoonlijke motivatie doen.

De bezoekers van LMA zijn doorgaans geïnteresseerd in familiegeschiedenis terwijl de bezoekers van het Antwerpse stadsarchief eerder uit algemene interesse de fotocollecties raadplegen. Zowel binnen LMA als Stadsarchief Antwerpen (SA) zoekt men bij voorkeur foto's van gebouwen en straatbeelden. Foto's van mensen en gebeurtenissen komen op de tweede plaats. De oudste beelden genieten de voorkeur. Zo wordt al wat ouder is dan 1920 het vaakst opgevraagd. Het grootste deel van het publiek verkiest eerder op plaats en onderwerp te zoeken dan op datum.

Er werden ook vragen gesteld over het reproduceren van originele foto's. Kwaliteit en prijs zijn de belangrijkste elementen die men in overweging neemt bij het bestellen van reproducties. Een fotografische reproductie krijgt de voorkeur boven een digitale kopie of een print. Bij het bestellen van reproducties via het internet wensen de meeste bezoekers met een bankoverschrijving te betalen. Dit hoeft niet te verbazen aangezien niet elke besteller over een kredietkaart beschikt.

Uit deze enquête blijkt dat in 1998 ongeveer één derde tot de helft van alle bezoekers regelmatig gebruik maakte van het internet. Van die groep was van beide instellingen samen ongeveer 30 tot 40 % bereid om gebruik te maken van een website om de foto's via het internet digitaal te kunnen consulteren. Slechts enkelen gaven de voorkeur aan het traditionele archiefbezoek boven consultatie van een website.

Het tweede onderzoek richtte zich vooral op archiefinstellingen die fotografische collecties beheeren verspreid in Europa. In eerste instantie beschouwen veel instellingen digitaliseren als een aparte problematiek. Op het moment dat het onderzoek werd gevoerd zaten veel instellingen in de planningsfase van hun project. Een aantal waren betrokken bij concrete pilootprojecten en een minderheid was volop bezig met digitalisering op grote schaal. De twee voornaamste redenen om te digitaliseren zijn het verhogen van de toegankelijkheid van de informatie en een betere bewaring van het origineel.

Een concrete consensus over de te hanteren kwaliteitsnormen voor de originele digitale scans bestond niet. Toch was men het wel min of meer eens over de te hanteren formaten en opslagmedia. Veel archiefinstellingen maakten geen gebruik van een standaard beschrijvingsmodel. Slechts zelden werd er een koppeling gemaakt tussen de gedigitaliseerde opname en de metadata. Een ruime waaier aan software was toen reeds voor handen en de instellingen maakten daar als dusdanig gebruik van. Ondanks het grote belang lieten veel instellingen het na om het digitaliseringproces te documenteren.

Een groot deel van de Europese instellingen erkende de mogelijkheden en het belang van digitaliseren. Toch was digitalisering van archief een erg jonge techniek. Een expert die het project opvolgt bleek nodig om de risico's zoveel mogelijk te beperken. Bovendien was er een duidelijke nood om gebruik te maken van algemeen aanvaarde technische normen, een standaard beschrijvingsmodel, gestandaardiseerde software en methoden om alles te documenteren.

Uit het tweede onderzoek kon worden geconcludeerd dat bij aanvang van het EVA-project reeds heel wat instellingen gebruik hadden gemaakt van digitaliseringstechnieken. De concrete toepassing van die technieken verschilde van instelling tot instelling. Het was duidelijk dat het EVA-pro-

ject hier kort op de bal kon spelen door aan de noden en wensen van zowel de bezoekers als de archiefinstellingen tegemoet te komen.⁷⁸

Zes maanden na het afronden van het EVA-project en het lanceren van de website, werd gedurende 6 weken een vragenlijst gepubliceerd op de site (juli - augustus 2001). Van de 1 087 bezoekers vulde slechts 7 % het formulier in zodat de respons te laag werd beschouwd om verder te nemen in het onderzoek. Toch loonden enkele vaststellingen de moeite:

- Twee derde van wie het formulier invulde, bezocht de website niet voor het eerst. Zoekmachines hadden hen niet op de weg gezet.
- Meer dan twee derde bezocht de website omwille van persoonlijke belangstelling. Wanneer de site om professionele redenen werd gebruikt, was wetenschappelijk onderzoek de hoofdreden.
- Ongeveer de helft van de bezoekers vond wat werd gezocht; een derde had geen specifieke zoekopdracht.

EVAmP, 2003 - 2004

Het onderzoek in 2003 gevoerd door EVAmP, focuste vooral op de mate waarin het EVA-project tegemoet kwam aan de noden en de wensen van de eindgebruikers en de deelnemende instellingen. Het gebruikersonderzoek dat in het tweede project werd gehouden resulteerde in veel meer interessante informatie. Er werden verschillende onderzoeken uitgevoerd:

- naar de gebruikers van de EVA-website in de meest algemene zin
- met een beperkte groep naar de wijze waarop het zoeken op de website verloopt
- een vragenlijst aan erfgoedbewaarders.

Ook hier gold het uitgangspunt dat de uitkomst van de ondervraging de nodige input moest leveren voor het verfijnen van de beeldbank en de diensten aangeboden via het internet. Het was de noodzakelijke voorwaarde om de EVA-website nauwer te doen aansluiten bij de noden van de instellingen en de gebruikers. Beide projecten hebben steeds een goede kennis van de gebruikers vooropgesteld.

Er namen 402 gebruikers deel. In tegenstelling tot de enquête in 1998 die enkel onder de leeszaal-bezoekers werd gehouden, werd er een vragenlijst, op de EVA-website gepubliceerd, die vrij kon worden ingevuld.

De groep gebruikers heeft een eerder uiteenlopend karakter qua samenstelling. Het overgrote deel (56 %) zijn gewone bezoekers, gevolgd door mensen met een informaticaopleiding of die actief zijn binnen de IT-sector. Daarna volgen historici, academici, professionelen uit de media sector en tenslotte studenten. 51 % van de deelnemers is vertrouwd met of heeft een basiskennis van on line informatiesystemen. Slechts bij een derde is het tegendeel waar, niettegenstaande ze uiteindelijk toch de gezochte informatie vinden. De EVA-bezoekers beschikken minstens over een

⁷⁸ I. SCHOUPS, Access to Europe's photographic heritage from your desk-top: European Visual Archive (EVA) in INSAR, Supplement IV 2000. Verslagen van DLM-Forum '99: European citizens and electronic information - the memory of the information society (18-19 october 1999, Brussels), <http://192.87.107.12/eva/nl/texts/insar.pdf> (7 juni 2005).

basiskennis informatica of zijn bereid het nodige te doen om toch maar de informatie te kunnen raadplegen.

Het project wou de eindgebruikers vooral ondervragen over de gebruiksvriendelijkheid van de bestaande website en kennis verzamelen over de wijze waarop gebruikers zoekacties formuleren of uitvoeren. Ook was het belangrijk om te vernemen hoe gebruikers websites en beeldbanken vinden en gebruiken in het algemeen. Er werd hen ook een aantal types zoekmodules of beeldbanksystemen voorgelegd.

Een eerste analyse leverde resultaten analoog aan die van de enquête in 2001.

39 % van de gebruikers is hooguit geïnteresseerd in het rechtstreeks downloaden van de afbeeldingen van lage kwaliteit van de website. 26 % wil beschikken over de mogelijkheid om digitale beelden van hoge kwaliteit te kunnen kopen. 20 % wenst enkel informatie over de afbeeldingen. Voor deze groep is het in het bezit krijgen van een digitale of analoge kopie niet aan de orde van de dag. Ten slotte is slechts 10 % van de gebruikers bereid om een analoge reproductie te kopen. 55 % verkiest een goedkope service met een wachttijd van 1 à 2 weken per bestelling tegenover een snelle, gegarandeerde service die duurder is.

Ook naar metadata toe is de mening van de gebruikers vrij uitgesproken. 39 % was hoofdzakelijk geïnteresseerd in beschrijvende informatie over het onderwerp van de beelden. 34 % wilde juist zoveel en zo divers mogelijke informatie. Gebruikers die informatie wensen over de collectie waar toe de foto behoort, vormen de derde grootste groep gevolgd door diegenen die ook informatie over plaats en datering van de foto belangrijk achten. Fototechnische informatie en de identificatie van de fotograaf scoren het laagste. Het grote publiek is eerder geïnteresseerd in een algemene, inhoudelijke beschrijving van het beeld terwijl academici de voorkeur geven aan een volledige beschrijving.

Ten slotte bleek dat de persoonlijke belangstelling de voornaamste motivatie vormt om de beeldbank te raadplegen. Wetenschappelijke onderzoekers vormden de grootste groep die uit professionele belangstelling van de beeldbank gebruik maakt. De helft van de bezoekers verkreeg de gewenste resultaten na een zoekopdracht.

Net zoals tijdens de enquête in 1998, werd er in 2003 gekeken naar de respons van de aan het EVA-project deelnemende instellingen.

In de eerste projecten waren de erfgoedinstellingen en beheerders van fotografische verzamelingen grotere overheids- of gemeentearchieven. Naast deze groep is er in Europa een zeer ruime waaier aan erfgoedinstellingen zoals kleinere archieven, gespecialiseerde bibliotheken, bedrijfsarchieven, bibliotheken en musea, die allemaal fotografische en beeldverzamelingen bewaren en beheren. Zij zijn allemaal potentiële leveranciers van beeldmateriaal voor een beeldbank zoals EVA. Voor onderzoekers is het onderscheid tussen het type instelling dat het beeld aan de oorsprong bewaart, minder relevant.

In het kader van het project werd er gezocht naar de motieven en redenen waarom een instelling gebruik zou willen of kunnen maken van een dienstverlening via een beeldbank. Zijn zij geïnteresseerd in een zo groot mogelijke verkoop van reproducties van documenten uit de eigen verzamelingen of staat de algemene promotie van de instelling voorop? Bij de bevraging werden zowel instellingen betrokken die reeds over gedigitaliseerd materiaal beschikten als zij die nog geen digitale beelden hadden of konden aanleveren.

Zonder uitzondering begrepen de deelnemers het belang van een gemeenschappelijke beschrijvingsstandaard als fundament voor het EVA-project. De mogelijkheid om bestaande beschrijvingen te converteren is echter een *conditio sine qua non*. De mogelijkheden die EVA bood om

digitale moederkopieën veilig te bewaren, wekte interesse onder de deelnemers.

De meerderheid van de deelnemers vond het interessant om een proefperiode met een aantal beperkingen in te lassen vooraleer definitief een deelname aan het EVA-project te bevestigen. Nadien zou een jaarlijkse vergoeding gekoppeld aan het aantal gedigitaliseerde en ingebrachte foto's voor de deelname worden betaald. Een basisvereiste voor deelname aan het project vormt het bestaan van de nodige ondersteuning vanuit EVA en de verzekering van de continuïteit van het project.

Het is duidelijk dat gedurende het project de wensen en de noden van de betrokken gebruikersgroepen en deelnemers deels verschoven zijn. Zo verkozen gebruikers in 1998 de bestelde reproducties in analoge vorm te ontvangen. In 2003 bleken digitale reproducties de voorkeur te krijgen. Het belang van een gemeenschappelijke beschrijvingsstandaard wordt duidelijk onderkend door de deelnemende instellingen maar het vasthouden aan de beschrijvingsmethoden die reeds gebruikt worden blijft een must.

Het volledige rapport van de enquêtes kan via de EVA-website worden geraadpleegd. Hierna worden de belangrijkste conclusies opgesomd omdat ze voor andere beeldbankinitiatieven een aantal interessante overwegingen bevatten. Ook bleek dat een aantal vaststellingen uit het eerste project nog steeds geldig zijn.

- De gebruikers verkiezen kwantiteit; de beeldbankpartners moeten vooral het aantal ter beschikking gestelde beelden verhogen.
- De meertalige zoekmodule is zonder twijfel een meerwaarde.
- Het is beter te focussen op fotografisch materiaal. Gezamenlijk aanbieden van foto's, plannen, tekeningen en teksten in eenzelfde beeldbank vinden de gebruikers niet wenselijk.
- Er bestaat bij het doelpubliek een grote groep gebruikers van het internet. Deze groep is vertrouwd met internet en met zoekmethoden eigen aan internet.
- De eindgebruikers kunnen in twee groepen worden onderverdeeld naargelang persoonlijke of professionele belangstelling. De beeldbank en de instellingen dienen zich tot beide groepen te richten in aanbod en communicatie.
- De partners in EVA en de beeldbank mogen niet uitsluitend overheidsarchieven zijn. Ook andere archieven (universiteits-, particuliere of bijzondere archieven), musea en bibliotheken moeten worden opgenomen.
- Indien EVA nieuwe partners wil aantrekken moet de drempel om in te stappen laag worden gehouden. Dat zou kunnen door bijvoorbeeld een gratis proefperiode te voorzien. Gedurende zes maanden kunnen geïnteresseerde instellingen vertrouwd worden met de werkwijze en de producten beter leren kennen.
- De meeste instellingen die aan de EVA-beeldbank willen deelnemen en foto's op de site ter beschikking stellen van gebruikers en onderzoekers, zijn bereid om daarvoor een jaarlijkse bijdrage te betalen. Bij voorkeur dient deze bijdrage in verhouding te staan tot de grootte van de instelling of van de gedigitaliseerde en ter beschikking gestelde verzameling.
- Een aantal instellingen zouden er de voorkeur aan geven de deelname aan de beeldbank te betalen door een percentage af te staan voor elke via het internet verkochte reproductie.
- Van de voorgestelde modellen om materiaal op te halen bij de erfgoedinstellingen, ging de voorkeur naar de "OAI harvester". Dat model wordt hierna toegelicht.

- Er bestaat belangstelling bij een aantal erfgoedinstellingen om gebruik te kunnen maken van bijkomende diensten wanneer de EVA-groep consultancy en advisering zou kunnen bieden inzake digitalisering en behandeling van de verzamelingen. Aangezien de meeste instellingen deze dienstverlening zeker niet zoeken of wensen, dient het EVA-project met deze twee groepen rekening te houden.
- Heel wat erfgoedinstellingen zijn bereid om het afleveren van reproducties geheel uit te besteden. De meeste zijn zeer terughoudend om zonder meer digitale reproducties van zeer hoge kwaliteit af te leveren.
- Slechts 10 % van de gebruikers is nog geïnteresseerd in afdrucken op papier; een grote meerderheid wenst digitale beelden van een hoge kwaliteit. Hoe deze tegenstrijdige belangen kunnen worden verzoend, moet een uitdaging zijn voor het project en de diensten die men wenst aan te bieden.
- Het beheer van reproductie- en auteursrechten is voor de meeste instellingen een onbekende materie, en vooral een aangelegenheid waarin het EVA-project zekerheid moet kunnen bieden. Zonder duidelijkheid hieromtrent hebben heel wat instellingen te kennen gegeven dat zij aan een beeldbankproject niet zullen deelnemen.

Deze tweede reeks ondervragingen, nu in het kader van het EVA-project, bevestigde de belangstelling van instellingen en gebruikers om fotografische verzamelingen via internet te ontsluiten en aan te bieden of te gebruiken.



Wolhandel 3 Suisses, Meir 51 - 53: de roltoonbanken, 1962 (stadsarchief Antwerpen - OF 10906).

3.3.3 *Het analyseren van het fotografische materiaal (de collectie), van de auteursrechten en van de tarieven*

Deze analyse werd alleen tijdens het eerste EVA-project uitgevoerd. Het EVAMP-project was hoofdzakelijk gericht op marktonderzoek van gebruikers en erfgoedinstellingen; de aanpak van een fotoverzameling in het kader van digitaliseren was niet meer aan de orde.

3.3.3.1 *De collectie*

De fotografische collectie van de London Metropolitan Archives (LMA) bestaat hoofdzakelijk uit zwart-wit foto's. Het grootste deel daarvan zijn zilvergelatine opnamen. Er zijn ook afdrukken op carbon, albumine of collodion. De meest voorkomende formaten zijn 8 x 6, 10 x 8 en 12 x 10 inches. De grootste formaten gaan tot 20 x 16 inches. Niettegenstaande de foto's doorgaans los in dozen worden bewaard, is een flink deel opgekleefd in fotoalbums. De LMA Photographic Collection bestaat uit 2 060 dozen met 329 600 foto's. Deze collectie wordt beheerd door de Corporation of London die eveneens het intellectuele eigendomsrecht van deze collectie in handen heeft.

De collectie van het stadsarchief Antwerpen (SA) bestaat eveneens voor het grootste deel uit zwart-wit foto's. Net zoals in de LMA vormt het overgrote deel zilvergelatine opnamen. Daarnaast zijn er ook afdrukken op carbon, albumine of collodion. De opnamen in het stadsarchief Antwerpen zijn doorgaans niet groter dan 18 x 24 centimeter, hoewel er ook uitzonderingen zijn. Tot in de jaren 1990 werden een aantal foto's op kartonnen platen geplakt. Deze foto's werden aanvankelijk in fiekasten bewaard. Later werden ze los in dozen per onderwerp bewaard. Van deze collectie bestond geen inventaris. Groot formaat foto's werden eveneens op karton opgeplakt en bewaard in portfoliomappen. Vanaf 1994 stapte het stadsarchief van deze aanpak af. Sindsdien worden foto's los bewaard. De collecties worden daarbij ingedeeld volgens afkomst of de naam van diegene die de rechten in handen heeft. Het gaat om ongeveer 100 000 opnamen.

Niettegenstaande de foto's in een ruime waaier aan formaten voorkomen, bleek het merendeel geschikt te zijn om te worden gedigitaliseerd met een 'standaard' flatbed scanner.⁷⁹

3.3.3.2 *De auteursrechten*

Ondanks het feit dat deze collecties in openbare instellingen worden bewaard zijn er toch beperkingen wat betreft het (digitaal) toegankelijk maken van de opnamen. De regelgeving omtrent intellectuele eigendomsrechten bleek een grote hinderpaal te zijn. In de eerste plaats hanteert elk land een eigen wettelijk kader. Op Europees niveau leidt dit tot significante verschillen. Er bestaat geen eenduidige Europese regelgeving. Binnen het EVA-project was het opstellen van eenduidige richtlijnen omtrent copyright eerder problematisch. Uiteindelijk werd er een checklist opgemaakt die de deelnemende instellingen konden gebruiken tijdens het selecteren voor de EVA-beeldbank. De vele juridische onduidelijkheden bij oude foto's hebben geleid tot een eerder conservatieve houding tijdens de selectie. Vaak is immers niet geweten wie de foto heeft genomen, wie er op de foto staat en of de belanghebbenden nog in leven zijn.

– *Privacy:*

Indien er herkenbare, nog levende personen op een foto staan dan mag de foto niet gepubliceerd worden zonder toestemming van de belang- of rechthebbenden.

⁷⁹ EVA CONSORTIUM, *The European Visual Archive - EVA , Inventory of Photographic Materials*, 1999, <http://192.87.107.12/eva/nl/texts/wp22.pdf> (7 juni 2005).

I. SCHOUPS, *Access to Europe's photographic heritage from your desk-top: European Visual Archive (EVA)*, Verslag van DLM-Forum 1999, <http://192.87.107.12/eva/nl/texts/insar.pdf> (7 juni 2005).

→ *Reproducties:*

Indien blijkt dat de foto een reproductie is van een oudere foto, dan mag de foto niet gepubliceerd worden zonder de toestemming van de rechthebbenden van de originele foto.

→ *Bezit van de foto's:*

Indien de archieven (LMA/SA) de foto zelf niet in eigendom hebben maar slechts beheren, dan mag de foto niet gepubliceerd worden zonder toestemming van de belang- of rechthebbenden.

→ *Copyright:*

Dit verschilt van land tot land. Diegene die de foto in bezit heeft, heeft niet noodzakelijk het copyright. Indien er merkbare tekens zijn aangebracht op de foto die kunnen aanwijzen wie het copyright in handen heeft (bijvoorbeeld persagentschappen of de fotograaf zelf) is er toestemming vereist om de foto te mogen reproduceren.

→ *Morele rechten:*

De fotograaf behoudt hoe dan ook de morele rechten over zijn werk.

Vooraf recente foto's vallen onder de bescherming van het auteursrecht. Er moet moeite worden gedaan om alle rechthebbenden op te sporen en hun toestemming te vragen vooraleer tot digitaliseren kan worden overgegaan. Mochten deze onbekend blijven dan raden de EVA-onderzoekers op zijn minst het gebruik aan van disclaimers.⁸⁰

Binnen deze problematiek staat de wettelijke beschermingstermijn van 70 jaar centraal. Nadien worden de foto's niet meer beschermd door het auteursrecht en behoren de ze tot het publieke domein. Omwille van het portretrecht is het noodzakelijk dat de personen op de foto hun toestemming verlenen voor reproductie en distributie.

Binnen het EVA-project was de keuze om oudere foto's te digitaliseren dan ook veeleer tweeledig. Enerzijds omdat het publiek geïnteresseerder was in oudere opnamen en anderzijds omdat men door oudere foto's ter beschikking te stellen de confrontatie met de juridische problematiek vermeed.⁸¹

3.3.3.3 De tarieven

Een digitaliseringproject brengt heel wat kosten met zich mee die openbare instellingen niet zonder meer aan de gebruiker kunnen doorrekenen. Je kan afhankelijk van de situatie afwegen welke kosten wel kunnen worden doorgerekend. Het EVA-project kaderde in een programma dat 'public access' tot de collecties van overheidsarchieven wilde ondersteunen.

Er werd besloten om voor de basisdienstverlening geen kosten aan te rekenen; men kan de beeldbank vrij bevragen en de resultaten bekijken. Wanneer een bezoeker van een digitaal beeld een kopie van hoge kwaliteit wil, dan wordt er een vergoeding aangerekend. Als basis worden de kosten voor het produceren van een dergelijke kopie gebruikt. Bijkomende kosten zijn afhankelijk van bijvoorbeeld het doel waarvoor men de kopie bestelt. EVA besloot dat voor een reproductie die wordt besteld voor professioneel gebruik een hogere vergoeding moet worden betaald dan

⁸⁰ I. SCHOUPS. *Access to Europe's photographic heritage from your desk-top, European Visual Archive (EVA)*, Verslagen van DLM-Forum 1999, <http://192.87.107.12/eva/uk/texts/insar.pdf> (7 juni 2005).

⁸¹ EVA CONSORTIUM, *The European Visual Archive - EVA, Report on Recommendations for the Management of Copyright Issues*, 1999, <http://192.87.107.12/eva/nl/texts/wp42.pdf> (7 juni 2005).

voor een kopie voor persoonlijke of educatieve doeleinden. Ook liggen de kosten voor een bestelling buiten de EU hoger. Tot slot moeten nog de kosten voor vergoedingen voor auteursrechten mee worden aangerekend.

Auteursrechten spelen in het berekenen van de vergoeding en het opvolgen van de bestelling een belangrijke rol. Wanneer er een bestelling geplaatst wordt voor een foto waarvan een archiefinstelling de auteursrechten niet beheert, dan wordt alsnog door die instelling behandeld in naam van de rechthebbenden, indien er een akkoord met die laatste bestaat. Ook dit heeft een impact op de kostenberekening.

Binnen het EVA-project kwam men ook tot de consensus dat er bij grote bestellingen de mogelijkheid moest zijn om korting toe te staan, en dat bestellingen in de nationale munt (voor de eurozone: de euro) moesten worden betaald. Uiteindelijk besloot men dat de inkomsten van de reproducties terug in het EVA-project moesten worden geïnvesteerd. Dit laatste voorstel werd echter niet gerealiseerd.

3.3.4 Het bepalen van de standaarden voor kwaliteitsvolle digitale beelden

3.3.4.1 De kenmerken van digitale images of beelden

Het definiëren van standaarden voor digitale reproductie is niet eenvoudig. De *kwaliteit* staat voorop maar hangt wel af van tal van factoren, onder andere de eisen die de gebruiker stelt. Een occasionele bezoeker van een beeldbank zal zich tevreden stellen met een globale indruk van een foto. Een wetenschappelijk onderzoeker wil daarentegen meestal enkel werken met hoog kwalitatief beeldmateriaal. Het subjectieve karakter van kwaliteitsnormen maakt dat een eenduidige referentie om de kwaliteit van een digitaal beeld te bepalen niet bestaat. Nochtans zijn standaardwaarden voor meetbare parameters nodig, wil men de apparatuur kunnen ijken en kwaliteitscontroles uitvoeren. In de derde workpackage is onderzocht welke kwaliteitsnormen moeten worden gehanteerd om succesvol te kunnen digitaliseren.

Er is een overzicht gemaakt van mogelijke parameters die hiervoor kunnen worden gebruikt en welke procedure moeten worden gevolgd om op min of meer objectieve manier een hoge kwaliteit te bereiken. Het opnameprocédé is het vertrekpunt van het analoge origineel. Sinds het ontstaan van de fotografie zijn er heel wat methoden en technieken ontstaan om foto's te maken en te ontwikkelen. Dit gaat van *daguerreotype* over *cyanotype* tot *zilvergelatine* afdrucken. Al snel werd duidelijk dat binnen het kader van de historische fotografie, het eigenlijke fotografisch procédé van groot belang is. Vaak probeert men dit in reproducties te simuleren. In hoeverre is dit mogelijk binnen een digitaal beeld? In een eerste stap worden de verschillende kenmerken van dit procédé beschreven, vervolgens wordt gecontroleerd in welke mate deze in het digitale beeld terug te vinden zijn.

Het *Department of Preservation and Conservation* aan de *Cornell University* ontwikkelde een methode om fotografische procédés weer te geven in digitale vorm. Volgens de onderzoekers kan men een afdruk op drie niveaus beschrijven. De *structuur* verschaft informatie over het productieprocédé dat gebruikt werd om een opname te maken. Het *detail* beschrijft in welke mate het mogelijk is om het kleinste beduidende onderdeel te onderscheiden van op een korte afstand. De *essentie* beschrijft wat het oog zonder hulp kan waarnemen van op leesafstand. Om een opname te digitaliseren moet het mogelijk zijn om voldoende informatie te verkrijgen op alle niveaus van dit model. Om de structuur van een opname digitaal te beschrijven moet je scannen op een zo

groot mogelijke *resolutie*. De onderzoekers kwam tot de conclusie dat scannen met een resolutie van 400 dpi en een bitdiepte van 8 bits volstaat om detail digitaal te kunnen bewaren.⁸²

Hoewel het onderzoek van Cornell een goede indicatie vormt om tot een standaard te komen, blijft de vraag of een digitale kopie het origineel volledig zal kunnen vervangen tot vandaag onbeantwoord. Wel bestaat er een consensus dat het digitaliseren van foto's, de toegankelijkheid van de collectie vergroot. Informatietechnologie creëert immers nieuwe mogelijkheden om data te ontsluiten en aan te bieden op een gebruiksvriendelijke manier. Omdat na het digitaliseren het origineel niet meer moet worden gemanipuleerd, kan dit ook beter worden beschermd tegen de slijtage en beschadigingen door de raadpleging.

De EVA-onderzoekers concentreerden zich vervolgens op de concrete parameters die binnen EVA konden worden gebruikt om de digitale kwaliteit te waarborgen. De twee meest gebruikte parameters zijn de *resolutie* of het oplossend vermogen en de *bitdiepte*. Toch beveelt het EVA-onderzoek aan om ook andere parameters in beschouwing te nemen bij de aanschaf van een scanner of digitale camera. De performantie van een apparaat wordt vastgesteld op basis van de *modular transfer function*, *scaling capability*, *image noise*, *tonal accuracy*, *colour fidelity* en *task speed*.⁸³ Het bepalen van de performantie van een apparaat of een digitaal beeld kan aan de hand van deze parameters op een min of meer objectieve manier gebeuren. Dezelfde parameters worden ook gebruikt om een apparaat te ijken zodat eenzelfde kwaliteitsniveau ten allen tijde kan worden gegarandeerd.

Daarnaast bestaat er ook een subjectieve wijze waarop men een kwaliteitswaarde aan een digitaal beeld kan toekennen. EVA verwijst in deze context naar de *psychophysic scaling tools* die ontwikkeld werden door het Image Permanence Institute. In dit onderzoek moesten proefpersonen op verschillende wijzen beelden sorteren volgens kwaliteit.⁸⁴

Bij het evalueren van de kwaliteit van een digitaal beeld moet er ook rekening worden gehouden met de *intentie* of de bedoeling waarmee men het beeld digitaliseert. Een onderbelicht origineel kan resulteren in een digitale kopie waarvan men vindt dat de kwaliteit te laag is. Het digitale beeld kan eventueel worden gemanipuleerd om de kwaliteit te verbeteren maar daarmee wordt de relatie met het originele beeld wel ten dele verbroken.

Er worden een onderscheid gemaakt tussen vier verschillende intenties.⁸⁵

- ↪ *Exacte weergave van het bestaande fotografische beeld.*
Het digitale beeld komt overeen met het originele. Het contrast en de helderheid van origineel en de digitale kopie zijn gelijk. Er is geen manipulatie van het digitale beeld.
- ↪ *Weergave van het oorspronkelijke fotografische beeld.*
De originelen kunnen beschadigd zijn door incidenten tijdens hun levensloop (bijvoor-

⁸² A. KENNEY, L. SHARPE (eds), *Illustrated book study. Digital conversion requirements of printed illustrations. Report to the Library of Congress*, 1999, <http://www.library.cornell.edu/preservation/illbk/ibs.htm> (7 juni 2005).

We verwijzen in dit verband naar de cDAVID aanbevelingen omtrent de keuze van resolutie en bitdiepte die in hoofdstuk 2 aan bod komen.

⁸³ EVA CONSORTIUM, *The European Visual Archive - EVA, Reports on standards for image quality*, 1999, <http://192.87.107.12/eva/nl/texts/wp31.pdf> (7 juni 2005).

⁸⁴ J. REILLY, F. FREY, *Recommendations for the evaluation of digital images produced from photographic, microphotographic and various paper formats*. Report to the Library of Congress, National digital Library Project, Rochester, NY, IPI, Rochester Institute of Technology, 1996, <http://memory.loc.gov/ammem/ipirpt.html> (7 juni 2005).

⁸⁵ J. REILLY, F. FREY, *Digital imaging for photographic collections. Foundations for technical standards.*, Image Permanence Institute, Rochester, 1999, http://www.rit.edu/~661www1/sub_pages/8page3g.htm (7 juni 2005).

beeld kreuken, scheuren, waterschade, vergeling, enzovoort) Met beeldbewerkingprogramma's en de nodige technische expertise kan men deze digitaal restaureren.

→ *Weergave van de intentie van de fotograaf.*

De originelen kunnen reeds bij hun ontstaan van een dermate lage kwaliteit zijn door onderbelichting, slechte afdruk of andere. Met beeldbewerkingprogramma's is het mogelijk om de intentie van de fotograaf enigszins te reproduceren.

→ *Exacte weergave van het object dat te zien is op de foto.*

Het is mogelijk om bijvoorbeeld de kleur van een object op een digitaal beeld overeen te laten komen met de kleur van eenzelfde object in een museum. Dit vereist wel een verregaande verfijning in het afstellen van de te gebruiken hard- en software.

Het digitaliseren levert een *bitstream*. Dit is een binaire representatie van het origineel. Een bitstream van een digitaal beeld wordt ook wel een *raster file* of een *bitmap* genoemd omdat het om een matrixvoorstelling gaat van kleurwaarden. De resolutie bepaalt het aantal elementen of *pixels* (picture elements) van deze matrix en de bitdiepte bepaalt hoeveel informatie een element kan bevatten. Op basis van de resolutie en de bitdiepte kan men gemakkelijk de grootte van de bitstream berekenen. Dit is belangrijk om een inschatting te kunnen maken van de benodigde opslagcapaciteit van een server en de kosten die aan een dergelijke investering verbonden zijn. Meer informatie betekent immers dat een grotere opslagcapaciteit nodig is.

De EVA-onderzoekers onderzochten de meest efficiënte manier van dataopslag. Er werd geopteerd voor een systeem van *digital masters*, *reference images* en *thumbnails*. De eis van een waarheidsgetrouwe digitale reproductie resulteert in grote digitale bestanden. Dit zijn de zogenaamde *digital masters* of moederkopieën. Zij worden bewaard in het TIFF of *tagged image file format*. Hun grootte maakt hen ongeschikt voor distributie over het internet. Door de eis om zoveel mogelijk 'oorspronkelijke' informatie te bewaren, is het toepassen van allerlei (verlieslatende) compressietechnieken en andere vormen van manipulatie op deze moederkopieën volledig uit den boze. Om dit probleem op te vangen maakt men kleinere afgeleide afbeeldingen of *reference images* van een moederkopie. Deze worden opgeslagen in het JPEG- of *Joint Photographic Experts Group* formaat. Op deze digitale beelden worden wél compressiealgoritmen en manipulatie (herschalen naar een kleiner formaat) toegepast. De resulterende bitstream is vaak slechts een fractie van de originele bitstream. De kwaliteit is dan ook navenant waardoor deze referentie-images totaal ongeschikt zijn voor lange termijn archivering. Hun minimale grootte maakt ze dan wel weer zeer geschikt voor snelle weergave op een scherm.

Hoewel de kwaliteit van het digitale beeld niet degradeert onderkennen de EVA-onderzoekers wel het probleem van de digitale veroudering. Soft- en hardware verouderen in een hoog tempo. Dit heeft zijn gevolgen voor de leesbaarheid: niemand kan garanderen dat een bestand leesbaar zal zijn met de software van morgen. Er zijn verschillende strategieën om dit probleem te ondervangen. Het EVA-project maakt het onderscheid tussen *emulatie* enerzijds en *conversie* anderzijds. Door oudere software te emuleren op nieuwe systemen kan men bestanden in een ouder formaat openen en lezen. Een andere mogelijkheid is het converteren van een bestand van een oud naar een nieuw formaat.

Verder beveelt het EVA-onderzoek volgende zaken aan om te garanderen dat de bestanden in de toekomst nog kunnen worden gelezen.

1. Documenteer alle methoden, activiteiten en apparatuur die werden gebruikt tijdens de digitalisering.
2. Gebruik open systemen en open standaarden aangezien deze algemeen bekend en vrij te gebruiken zijn.
3. Bewaar de bestanden in een niet-gecomprimeerd formaat. Compressie verhoogt immers de afhankelijkheid van een welbepaald systeem.

4. Gebruik duurzame media zoals cd-rom's of tape om data te bewaren en behandel ze op een zorgzame manier.

Zoals we verder in deze casestudy zullen zien worden de resultaten en aanbevelingen van dit onderdeel geïmplementeerd in de concrete workflow binnen de deelnemende archiefinstellingen.

3.3.4.2 De metadata

Het tweede onderdeel dat in de derde workpackage aan bod kwam, is de beschrijving van de metadata. Dit is minstens zo belangrijk als het scannen van de afbeelding zelf. Het beschrijven van het beeld op een eenduidige wijze in al zijn facetten, is nodig om de fotocollectie digitaal toegankelijk te maken en foto's snel terug te vinden. Het EVA-project definieert metadata als alle data die ontstaat in de marge van de digitalisering, de bewaring en de ontsluiting van een digitaal beeld.⁸⁶

Om metadata te beschrijven bestaan er verschillende metadata-schema's. Deze definiëren de verschillende elementen en toegelaten waarden waarmee men een beeld kan beschrijven. Het rigoreus toepassen van een standaard metadata-schema laat toe om een collectie op een homogene manier te beschrijven en te ontsluiten. De metadata-schema's die voor EVA in aanmerking kwamen, waren: *Dublin Core*, *ISAD(G)/EAD*, *RLG Working group on Preservation Issues of Metadata*. Vooral het Dublin Core metadataschema speelt een voorname rol in de online ontsluiting van digitale bronnen. Om die reden richtte het EVA-rapport zich op dit metadata-schema om de digitale fotocollectie te ontsluiten.

EVA ontwikkelde het volgende beslissingsmodel om een schema te kiezen. Dit is gebaseerd op twee principes. Ten eerste: het evalueren van de bestaande metadata-schema's en standaarden die gebruikt worden binnen archief-, museum- of bibliotheekomgeving; kies het meest relevante schema en pas het aan de eigen noden aan. Ten tweede: consulteer de technische ISO-aanbeveling betreffende de functionele eisen voor een permanente, duurzame bewaring van digitale informatie. Deze aanbeveling staat bekend onder de naam *Reference model of an open archival information system* (OAIS).⁸⁷

3.3.5 De workflow

Voor de eigenlijke digitalisering van de foto's werd een workflow opgezet waarbij rekening werd gehouden met de aanmerkingen geformuleerd in de voorbereidende workpackages (2, 3 en 4)

De concrete aanpak van het digitaliseren kunnen we best beschouwen als een traject bestaande uit verschillende processen dat moet worden afgelegd. Digitaliseren is immers meer dan louter scannen en of digitaal fotograferen van stukken. We kunnen dit traject ruwweg in drie grote stukken opdelen: het selecteren van het beeldmateriaal, het scannen zelf en het maken of invoeren van de beschrijvingen.

⁸⁶ EVA CONSORTIUM, *The European Visual Archive - EVA, Report on Standards for the Description of Photographs*, 1999., <http://192.87.107.12/eva/nl/texts/wp32.pdf> (7 juni 2005).

⁸⁷ L. REICH, D. SAWYER (eds), *Reference model for an Open Archival Information System* (OAIS), CCSDS 650.0-B-1, Blue book, 1999. (ISO 14721:2003).

3.3.5.1 Selectie

Zowel de LMA als het SA gebruikten een *topdown benadering* tijdens de *selectie* van de opnamen die men wilde digitaliseren. Dit is een fundamenteel verschil met andere projecten. Doorgaans gaat men uit van één welbepaalde collectie of archief dat men integraal digitaliseert. In EVA is voor een aparte aanpak gekozen waarbij men wilde dat de inhoud van de digitale beeldbank representatief was voor de integrale fotocollectie bewaard in de depots van de deelnemende instellingen. Daartoe werden een aantal criteria gebruikt om te bepalen of een beeld al dan niet voor digitaliseren in aanmerking kwam.

Kwaliteit en formaat

Niet elke foto komt in aanmerking voor digitaliseren. Een zwaar beschadigde opname is minder geschikt. Enerzijds omdat er een grote investering van tijd en middelen nodig is om dergelijke beelden te digitaliseren, anderzijds omdat het digitale beeld door de beschadiging van het analoge originele hoe dan ook veel minder het verleden zal laten spreken.

Een foto verkeert bij voorkeur in een goede conditie. Het beeld moet helder en duidelijk zijn. In plaats van verschillende types te digitaliseren is er gekozen om zich te beperken tot één welbepaalde soort: zwart-wit afdrukken op papier van éénzelfde grootte. De reden hiervoor is efficiëntie. Voor een partij foto's die uit verschillende soorten en formaten bestaat is het juist moeilijker om eenduidige optimale kwaliteitsstandaarden te hanteren.

De inhoud

Foto's zijn ook geselecteerd op basis van hun inhoud of de *scène*. Daarbij kreeg de historische waarde voorrang op de esthetische. De foto's moesten informatie kunnen verschaffen over het leven in en de geschiedenis van de stad. De esthetische waarde van de foto was echter niet helemaal irrelevant tijdens de selectie: een foto die visueel aantrekkelijk was, kreeg de voorkeur, zolang de historische waarde van beide meer dan voldoende hoog werd geacht. Selectie op basis van de inhoud blijft natuurlijk een subjectief gegeven.

Subject groups

Tijdens het werk in workpackage 2 werd binnen het stadsarchief Antwerpen de fotografische collectie opgedeeld per onderwerp. Achttien onderwerpen werden nog eens onderverdeeld in trefwoorden.

Elke groep wordt apart aangepakt. Foto's worden systematisch uit de collectie gehaald en verwerkt. Ernstige verstoring van de normale dienstverlening werd dankzij deze aanpak vermeden aangezien er telkens slechts een klein gedeelte van de collectie enkel voor een korte periode niet beschikbaar was.

Representatief

Binnen elke groep is er geprobeerd om het onderwerp zoveel mogelijk te verbreden. Zo werd er voor gekozen om binnen bijvoorbeeld het thema handel en industrie, foto's van zoveel mogelijk verschillende beroepsactiviteiten te selecteren en zich niet te beperken tot slechts enkele welbepaalde activiteiten. Via een dergelijke aanpak probeerden we om een zo representatief mogelijk beeld te vormen van elk onderwerp. Het spreekt voor zich dat de keuze van relevante trefwoorden hierin cruciaal is.

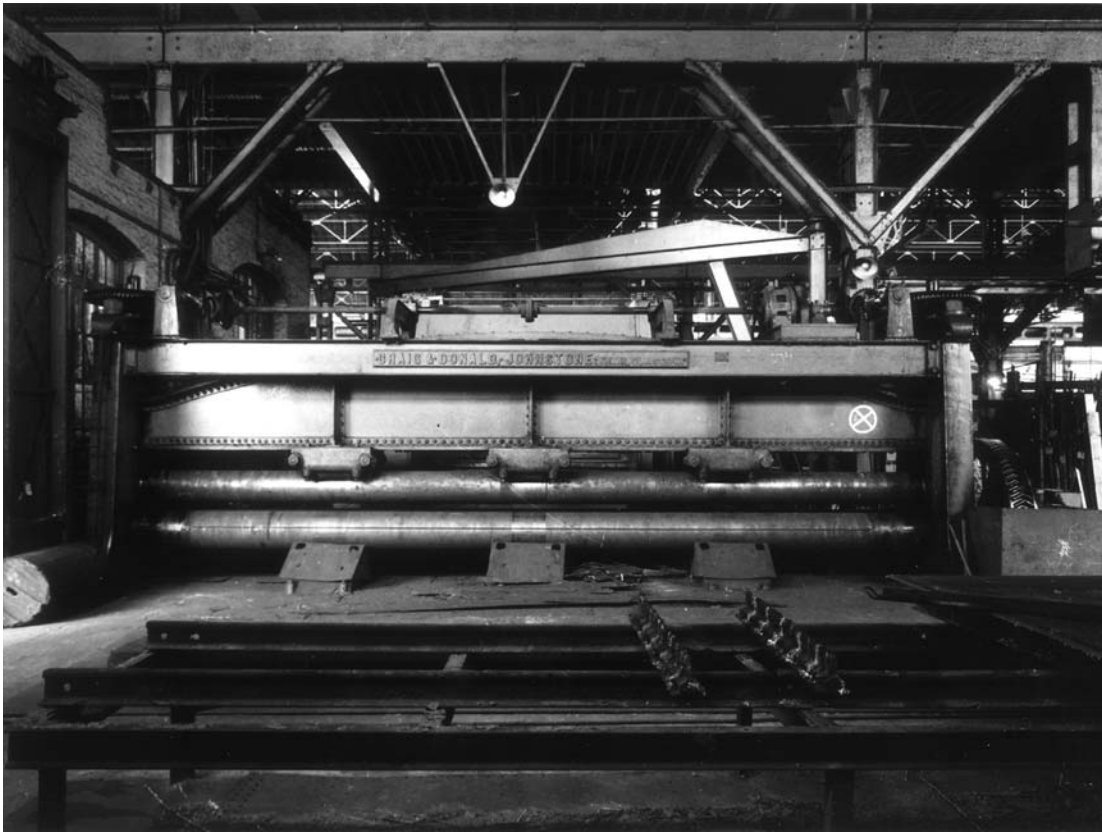
Copyright

Auteursrechten binnen het SA vormen een probleem. Vaak bezit het SA niet alle rechten op een foto. De rechthebbenden moeten hun toestemming verlenen om de foto on line te publiceren. Er is een specifiek administratief proces op poten gezet om de nodige stappen te ondernemen wanneer we niet over alle rechten beschikten. De aandacht is vooral gegaan naar het uitbouwen van een databank met alle namen van fotografen zodat een zorgvuldige identificatie en een correcte vermelding mogelijk was. Pas wanneer aan alle auteursrechtelijke voorwaarden voldaan was, kon de foto opgenomen worden in het digitaliseringstraject.

In tweede instantie speelde ook het portretrecht een rol. Foto's waarop nog levende individuen herkenbaar zijn, kunnen om privacyredenen niet zonder toestemming on line worden gezet. Het opsporen van individuen is in veel gevallen echter onbegonnen werk. Van individuen die herkenbaar zijn op opnamen die vóór 1920 dateren mag men aannemen dat ze overleden zijn waardoor dit bezwaar verval.

Datering

Het moment waarop de originele analoge opname werd gemaakt is eveneens een factor van belang. Uit de eerste bevraging van de gebruikers bleek dat het doelpubliek vooral geïnteresseerd was in foto's die dateren van vóór 1920.



Plooibank met scheepswerfmatten, geleverd door Craig en Donald, Johnston / Glasgow, 1960 (stadsarchief Antwerpen, OF 10859).

3.3.5.2 Scannen

Voor het scannen was er wel een verschil in aanpak tussen het stadsarchief Antwerpen en de London Metropolitan Archives. Dit vond men echter geen onoverkomelijk probleem. We gaan hier dieper in op de modus operandi van het SA.

Binnen het stadsarchief werden in eerste instantie twee scannermodellen gebruikt: de HP Scanjet 6200 en de AGFA Studiostar II. Beiden zijn *flatbed scanners*. De AGFA Studiostar wordt aangestuurd met de bijgeleverde Fotolook software. De HP Scanjet wordt aangestuurd met de bijgeleverde HP Precision Scan software. Voor de postproductie en het aanmaken van de digitale bestanden maakt het stadsarchief gebruik van Adobe Photoshop. Beide scanners zijn aangesloten op een Windows NT netwerkomgeving. De digitale beelden worden opgeslagen op een server of op cd gebrand. De geselecteerde opnamen zijn niet groter dan de dimensies van de scanners. Het eigenlijke scanproces binnen het stadsarchief bestaat uit verschillende stappen

1. Voorbereiden van de originelen.

- ~ Verwijderen van nietjes, paperclips, elastiekjes, enzovoort.
- ~ Sorteren van de afbeeldingen op grootte.

2. Het scannen van de afbeelding.

~ *Scannen van de afbeelding*

De volledige afbeelding wordt gescand. De output van de scanner wordt geïmporteerd in Photoshop. De operator die de apparatuur bedient slaat de output op als een TIFF-bestand.

~ *Image control*

Witte boorden worden digitaal afgeknipt (*cropping*), de digitale foto wordt geroteerd zodat de bovenkant van de foto horizontaal wordt uitgelijnd (*skew*). Dit alles gebeurt in Photoshop. Er wordt ook visueel door de operator vastgesteld of de kleuren van het digitale beeld overeenkomen met de kleuren van het analoge origineel.

Wanneer een digitaal beeld niet overeenkomt met het origineel, dan worden de instellingen aangepast en het stuk opnieuw gedigitaliseerd.

~ *Tussentijdse controle*

De hoeveelheid geselecteerde foto's zou dezelfde moeten zijn als het aantal moederkopieën.

~ *Tijdelijke bewaring*

De digitale beelden worden tijdelijk bewaard in een vooraf bepaalde mappenstructuur. De bestanden worden onder een unieke bestandsnaam tijdelijk opgeslagen op een lokale server binnen het stadsarchief Antwerpen. Deze noemt men de *digitale masters* of *moederkopieën*.

3. Controle van het digitale beeld.

Het digitale beeld moet voldoen aan volgende voorwaarden:

- ~ *Resolutie*: 300 dpi.
- ~ *Bestandsformaat*: TIFF, geen compressie.

- ~ *Skew*: het beeld moest horizontaal uitgelijnd zijn.
- ~ *Cropping*: het digitale beeld moest correct uitgesneden zijn.
- ~ *Naam*: De naam die men toekent aan het digitale bestand moet voldoen aan strikte regels. Het toekennen van de namen gebeurt op een systematische manier.⁸⁸ In casu werd gekozen voor een systeem van een sigel (SA) gevolgd door een uniek volgnummer. Een digitaal beeld draagt dus de naam SAXXXYYY.tif met XXX het volgnummer van de cd-rom en YYY het volgnummer van het bestand op de cd.

4. Definitieve opslag van de digitale beelden.

- ~ De digitale moederkopieën worden op een cd-r gebrand. De ISO-9660 (level 1) norm wordt bij het schrijven gehanteerd.⁸⁹
- ~ Van elke cd wordt een veiligheidskopie gemaakt.
- ~ Van de digitale moederkopie worden afgeleide bestanden (thumbnails, referentie-images, enz.) gemaakt.
- ~ De digitale masters worden van de server waar ze tijdelijk werden bewaard, verwijderd.
- ~ De moeder- en veiligheidskopieën worden op een geschikte wijze opgeslagen (rechttop staand, afgeschermd van licht in een stabiel klimaat). De moederkopieën worden in het archiefdepot van het stadsarchief Antwerpen bewaard. De veiligheidskopieën op een andere locatie.
- ~ De originele analoge opnamen worden niet meer ter beschikking van het leeszaalpubliek gesteld. Dit om een duurzame bewaring van het origineel naast de digitale kopie te waarborgen.
Het schrijven van de cd gebeurt pas wanneer de noodzakelijke metadata (de beschrijving) in het bestand zijn ingekapseld. Het TIFF-formaat laat immers toe om niet alleen beeldinformatie maar ook metadata mee op te slaan.

IJken van de apparatuur

Om een goede kwaliteit te garanderen is het noodzakelijk dat de apparatuur regelmatig wordt *gebenchmarked*. Dat wil zeggen dat men de apparatuur test en eventueel kalibreert zodat de output overeenkomt met de kwaliteitseisen. In de eerste plaats wordt er onderzocht of er nog wel wordt gescand met een *correcte* resolutie. Daartoe wordt de scanner gekalibreerd met een kaart met testpatronen. Ten tweede wordt de bitdiepte gecontroleerd. Er wordt bekeken of de scanner de kleuren van een speciale testkaart correct inleest. Tenslotte wordt gecontroleerd of de *signal-to-noise ratio* binnen de perken blijft. Hoe meer ruis in een digitaal beeld, hoe lager de kwaliteit van het digitale beeld. De ruis moet zoveel mogelijk worden beperkt.

In de praktijk werd de apparatuur binnen het stadsarchief slechts zelden geijkt. De voornaamste reden hiertoe is het ontbreken van de nodige kennis om dit zelf te doen. Het ijken is een precisie-

⁸⁸ ISO 9660:1988 *Information processing -- Volume and file structure of CD-ROM for information interchange*.

⁸⁹ F. BOUDREZ, DAVID *Richtlijn 2. Duurzame cd's*, stadsarchief Antwerpen, Antwerpen, <http://www.antwerpen.be/david/website/teksten/Richtlijn2.pdf> (29 maart 2005).

werk waarvoor de nodige expertise en de juiste hulpmiddelen (testkaarten en dergelijke) moeten aanwezig zijn.

Eindresultaat

Het digitaliseren levert drie eindproducten op: een *digitale master*, een *referentie-image* en een *thumbnail*.

↪ Digitale master

Dit is de moederkopie. Ze vormt een digitaal surrogaat van de originele afbeelding. Ze moet zo zoveel mogelijk details en informatie van het origineel bevatten. Daarom moet ze een zo hoog mogelijke kwaliteit hebben.

Door SA toegepaste parameters: ware grootte 300 dpi, kleuren in 24 bits RGB / zwart-wit in 8 bits grijswaarden, TIFF zonder compressie.

↪ Referentie-image

De EVA-beeldbank wordt beschikbaar gemaakt via het internet. Dat betekent het implementeren van de mogelijkheid om snel beelden via het web aan te leveren bij de gebruiker. De digitale master is hiervoor niet geschikt omwille van de bestandsgrootte. De *referentie-image* kan wel tot dit doel dienen door zijn kleine bestandsgrootte. Dit gaat wel ten koste van de kwaliteit waardoor een referentie-image nooit de functie van de digitale master kan overnemen.

Door SA toegepaste parameters: afmetingen herleid tot kader 400 x 400 pixels, JPEG met gemiddelde kwaliteit, toevoeging *copyright notice*.

↪ Thumbnail

De thumbnail is een afbeelding in een nog kleinere versie. Deze moet een eerste indruk geven van het beeld bij het overzicht van de relevante zoekresultaten. Thumbnails zijn veel te klein om een goed beeld van de afbeelding te krijgen.

Door SA toegepaste parameters: afmetingen herleid tot kader 75 x 75 pixels, JPEG gemiddelde kwaliteit.

3.3.5.3 Invoer van de archiefbeschrijvingen van de foto's

Binnen het stadsarchief Antwerpen zijn de meeste foto's beschreven, gedeeltelijk nog op steekkaarten en grotendeels met summiere beschrijvingen in databanken. Voor de reguliere ontsluiting maakt het stadsarchief gebruik van een documentbeheerssysteem (Documentum-eFloris⁹⁰). Hierin werden ook de MS Access-databanken met de beschrijvingen van de foto's beheerd en onsloten. Voor het beheer van het groot aantal beschrijvingen met de daarbij horende digitale beelden, was er nood aan een gedetailleerder beschrijving. Niet alleen moesten de gegevens over de fotografen en de geportretteerde personen ook worden opgenomen. De deelnemers van het EVA-project wensten de images of digitale beelden zelf te beschrijven en deze beschrijvingen te koppelen aan de originele foto en de daarbij horende data. Tenslotte wilde men een goede en snelle exportmogelijkheid van de beschrijvingen via XML naar de website.

⁹⁰ eFloris, Digitaal stadsarchief Antwerpen, <http://stadsarchief.antwerpen.be> (8 juni 2005).

Daarom werd een invoer- of *input module* ontwikkeld om de metadata op een efficiënte en eenduidige manier te beschrijven: SKOPEO BMS (Beschrijvingsmodule Stadsarchief Antwerpen). Het stadsarchief Antwerpen, Telepolis en NIWI ontwikkelden dit samen.

Het EVA-project maakt een duidelijk onderscheid tussen:

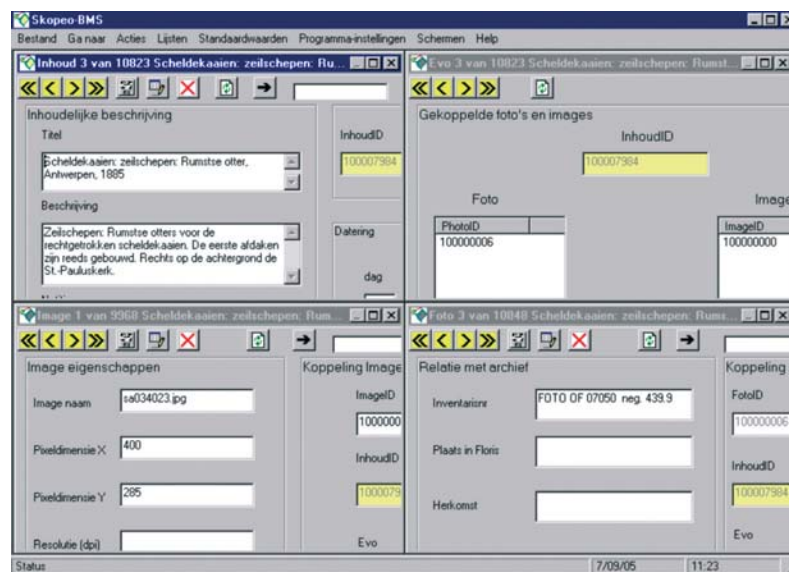
- het analoge origineel als fysiek medium ("foto")
- het digitale beeld als surrogaat voor de originele opname ("image")
- hetgeen te zien is op zowel het analoge stuk als het digitale beeld (ook wel "scène" of "inhoud" genoemd).

Zowel de 'foto' als de 'image' zijn gekoppeld aan de 'inhoud'. Dit maakt het mogelijk om meerdere 'foto's' of 'images' te koppelen aan één inhoud. Omgekeerd kan een image of een foto slechts één inhoud hebben.

Het analoge origineel kan meerdere kopieën of afdrucken hebben, van een verschillende datum of een ander type.

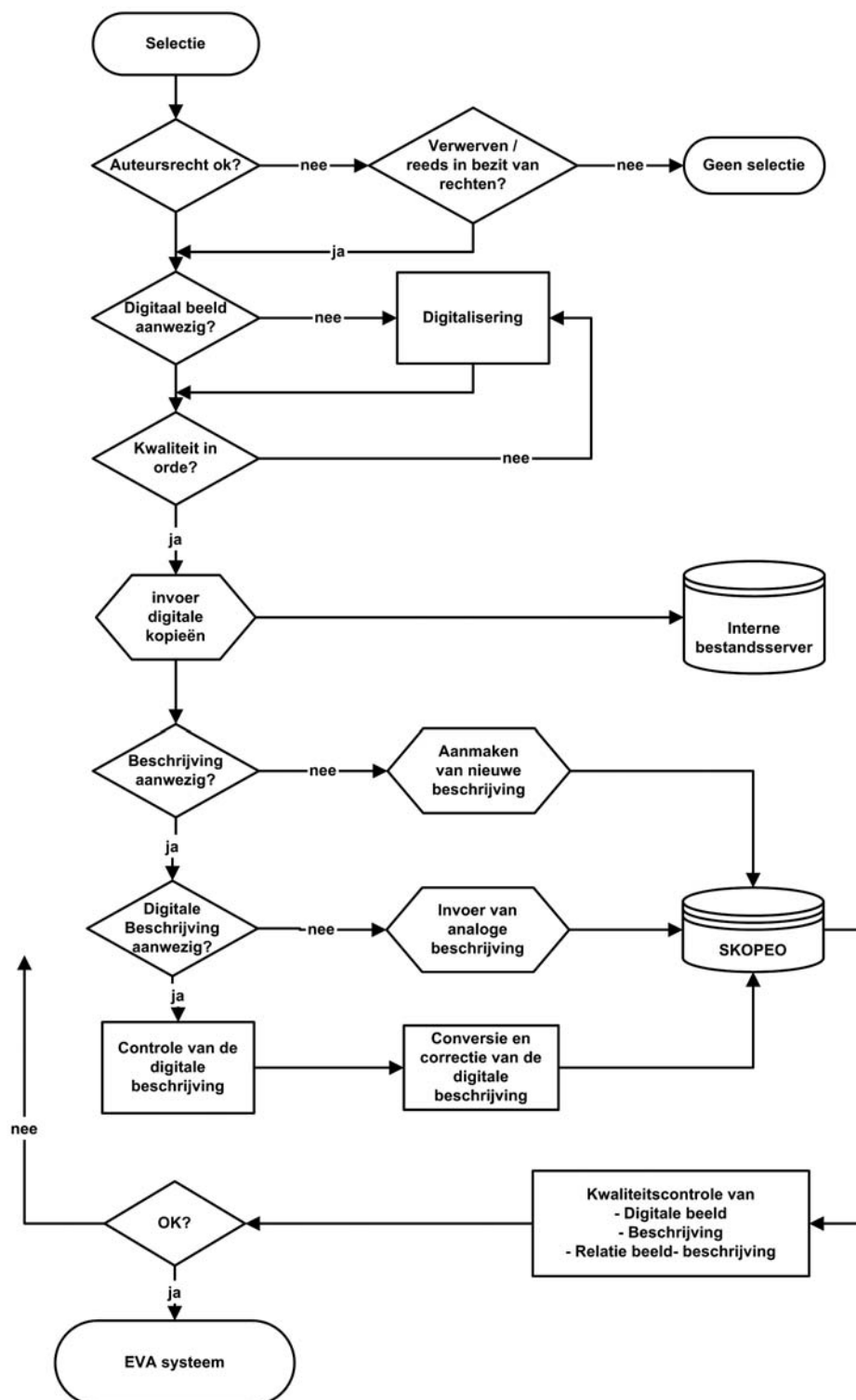
Het andere, tweede belangrijke uitgangspunt van EVA is dat archief- en erfgoedinstellingen bij het beheer van de beschrijvingen zowel het originele als het digitale beeld goed moeten documenteren. Zonder goede documentatie en beschrijving van het digitale beeld, is de duurzaamheid en lange termijn bewaring van het digitale beeld moeilijk en bedreigd. Op termijn moet het digitale beeld immers kunnen worden gemigreerd naar nieuwere formaten of standaarden.

Einde 2003 bleek dat deze invoermodule niet meer voldeed. De opname van alle databanken met reeds beschikbare beschrijvingen bleek geen haalbare oplossing. Bovendien werd er vanuit het NIWI niet langer ondersteuning en verdere ontwikkeling voorzien.



De grafische interface van SKOPEO.

Het stadsarchief Antwerpen ontwikkelt daarom samen met Digipolis een nieuw beeldbanksysteem: AVA of AudioVisueel Archief. Dit is een relatief eenvoudige web-based toepassing met een zoek- en inputmodule. Momenteel is intern reeds een beta-versie operationeel. Het belangrijkste verschil met het EVA-systeem is dat AVA niet enkel gedigitaliseerde foto's en hun beschrijvingen bevat maar alle beschikbare beschrijvingen van fotomateriaal, of het nu is gedigitaliseerd of niet. Voor opzoeken in de fotocollectie van het stadsarchief zal men dus nog slechts op één plaats moeten zoeken en niet langer in tien verschillende databanken.



Visualisering volledige workflow stadsarchief Antwerpen. Na het digitaliseren van (meta)data wordt steeds een grondige kwaliteitscontrole uitgevoerd⁹¹

⁹¹ EVA CONSORTIUM, *The European Visual Archive - EVA, Guidelines on the Methodology of Digitising and Describing Photographs*, 1999, <http://192.87.107.12/eva/nl/texts/wp54.pdf>.

3.4 De werking van de EVA-beeldbank en -website

Het EVA-systeem is een informatiesysteem dat voornamelijk ontworpen is om on line toegang te verlenen tot individuele foto's die deel uit maken van historische fotocollecties in archieven. Via het internet kan een gebruiker toegang verkrijgen tot een catalogus van historische opnamen en deze doorzoeken via beschrijvingsvelden. Om het Europese publiek zo goed mogelijk te bereiken, wordt een vertaalfunctie ingebouwd zodat de ondervrager in een van de ter beschikking gestelde talen kan zoeken. De gebruiker kan vervolgens de digitale beelden bekijken en ze eventueel als afdruk of in digitale vorm bestellen.

De andere pijler van het EVA-project bestaat uit de archieven en erfgoedinstellingen. Zij leveren de beelden en de beschrijvingen aan.

De EVA-beeldbank is een centrale server die de metadata en de digitale afbeeldingen bevat en verbonden is met het internet. Om dit te realiseren moest op volgende vragen een antwoord worden gegeven:

1. Bepalen van de data die ter beschikking zal worden gesteld.
2. Overdracht van de data van de archieven naar de beeldbank.
3. Bepalen van de publieksinterface.

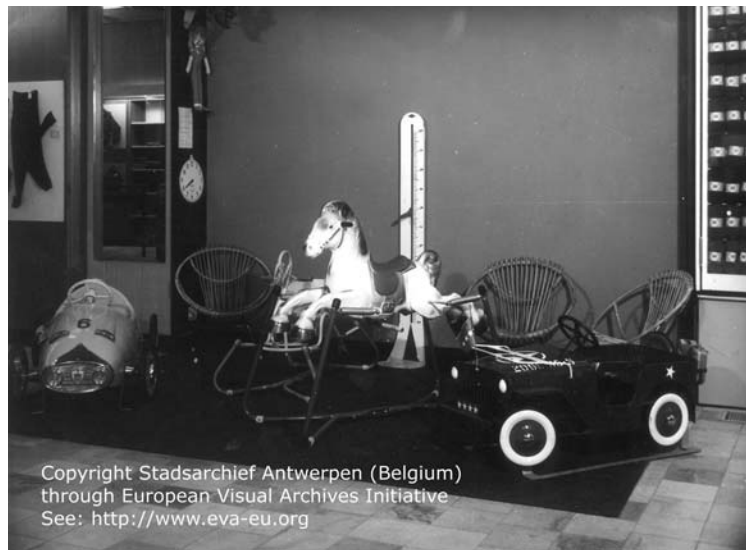
3.4.1 Welke images of beelden worden op de website geplaatst?

De archiefinstelling levert in wezen twee soorten digitale beelden of images af: *thumbnails* en *referentie-images*. De eersten zijn zeer kleine afbeeldingen met een grootte van hooguit een paar kilobyte. Wanneer het systeem een zoekopdracht verwerkt, bestaat het resultaat uit een reeks images die gerepresenteerd worden door deze thumbnails. De minieme grootte van deze bestanden maakt het mogelijk om er meerdere (tot 12) op één webpagina te tonen. Thumbnails zijn doorgaans te klein om een goed beeld te geven van de scène. Daarvoor dienen de referentie-images. Wanneer een bezoeker een beter beeld wil krijgen, dan moet er op de thumbnail worden geklikt om een referentie image te krijgen die een groter en beter beeld geeft.

De referentie-images zijn weliswaar geschikt om op een scherm een voldoende goed beeld te geven van de scène, maar niet geschikt voor drukwerk. Bovendien bezitten de referentie-images met het oog op de auteursrechten een *copyright notice* onder de vorm van een *watermerk*.

Een geschikte digitale kopie van hoge kwaliteit voor drukwerk of een reproductie moet men bestellen via de website. De archiefinstelling levert een hoogwaardige *reproduction image* af. Dit is een kopie van de master image die in het archief wordt bewaard. Deze 'reproduction images' die een exacte kopie van de digitale moederkopie zelf zijn, worden niet op de server geplaatst. Daarvoor zijn de bestanden te groot en het gebruik van deze afbeeldingen kan niet zonder toestemming van of vermelding naar de archiefinstelling aangezien deze de auteursrechten bezit.

Hieronder een voorbeeld van een referentie-image met het watermerk:



Wolhandel 3 Suisses, Meir 51-53: speelhoekje voor kinderen, met copyright notice EVA, 1962
(stadsarchief Antwerpen, OF 2398).

3.4.2 Welke beschrijvingsgegevens worden overgenomen op de website?

De gegevens over de "image", de (originele) foto en de inhoud, en hun onderlinge relaties vormen samen een EVO of *EVA Visual Object*. Het beschrijven van een EVO vereist dat er rekening wordt gehouden met de meest uiteenlopende eigenschappen. Daarom werd er in workpackage 3 uitvoerig onderzoek gevoerd naar de reeds bestaande beschrijvingstandaarden. Op basis van dit onderzoek werd de EVO-dataset samengesteld. Daarbij werden er attributen overgenomen uit EAD, ISAD(G), Dublin Core en ISAAR(CPF).

Het EVA-onderzoek toonde echter aan dat het niet haalbaar was een systeem te ontwikkelen waarin al deze attributen konden worden opgenomen. Voor de ontwikkeling van een webgeoriënteerd systeem baseerde het EVA-project zich uiteindelijk op de reeds bestaande Dublin Core metadata element set (zie 3.4.3). Die kent een brede verspreiding en wordt toegepast als generiek metadata-schema om objecten via het internet ter beschikking te stellen.

Het voor de EVA-website ontwikkelde metadata-schema is dan ook een minimale of vereenvoudigde lijst die bestaat uit een beperkt aantal basiselementen, de zogenaamde 'lite'-versie. Deze EVOLite dataset werd zo ontworpen dat de meest essentiële attributen uit de dataschema's die binnen de beide archieven bestonden en werden gehanteerd, konden worden gemapped naar de EVOLite elementen.

Het EVOLite metadata-schema

Dublin Core	EVA	Beschrijving
--	ID	Unieke code om de EVO te identificeren binnen het EVA-systeem.
Title	Title	Korte beschrijving van de scène.
Creator	Creator	De naam van de fotograaf die de originele analoge opname heeft gemaakt.
Subject / Keyword	Subject	Beschrijvende trefwoorden in de taal van de archiefinstelling waar de opname wordt bewaard.
Description	Description	Vrije beschrijving van het digitale beeld.
Publisher	Archive	Naam van de archiefinstelling die het originele beeld beheert.
Date	Date	Datum van ontstaan van de originele opname. Dit kan ofwel een exacte datum (dag/maand/jaar) zijn of een periode (beginjaar - eindjaar).
Identifier	Location Refimage	Bestandsnaam van de referentie-image en de thumbnail image.
	Laction thumbnail	
Language	Language	De taal gebruikt voor de waarden van de title, description, subject / keyword en coverage.
Relation	Relation	Verwijzing naar het inventarisnummer van de originele opname.
Coverage	Geography	Geografische trefwoorden voor het digitale beeld in de originele taal. (straat, district, stad, land).
Type		Niet gebruikt.
Format		Niet gebruikt.
Contribution		Niet gebruikt.
Source		Niet gebruikt.
Rights		Niet gebruikt. Er wordt een aparte copyright notice gemaakt voor de volledige on line fotocollectie.

Het hanteren van één metadata-set door beide deelnemende instellingen verhoogde de interoperabiliteit tussen de collecties van de instellingen. Bovendien kunnen de oude, lokaal gebruikte beschrijvingstandaarden behouden blijven.

De enige taak voor de archiefinstellingen was het verzorgen van de mapping of concordantie tussen de verschillende datasets zodat er duidelijkheid was welke velden in de eigen databestanden overeenstemden met de data-elementen van de EVA-beeldbank. Dit moest slechts voor 11 attributen worden gedaan.

Zowel Londen als Antwerpen hadden reeds een eigen traditie in het beschrijven van fotocollecties. Hoewel die in grote lijnen overeenkomen worden de accenten toch verschillend gelegd. De afstemming van 11 kernattributen garandeerde de volledige harmonisatie van de archiefbeschrijvingen in de gemeenschappelijke online beeldbank van EVA.

EVAmP, 2003 – 2004

Bij de evaluatie in het tweede project, EVAmP, zijn de uitgangspunten over een beperkte dataset met de kenmerken van de foto's en digitale beelden nog versterkt.

Tijdens het eerste EVA-project was reeds gebleken dat het verkrijgen van een consensus over de gegevenselementen die essentieel zijn om toegang te verschaffen tot de verzamelingen, een zeer langdurig en arbeidsintensief proces is. Om deel te nemen aan de beeldbank, zijn de meeste instellingen niet bereid om extra te investeren in het verkrijgen van een homogene set van metadata en daarvoor de reeds bestaande beschrijvingen aan te passen. De gegevenssets om analoge documenten en digitale beelden en hun inhoud te beschrijven, kunnen heel erg verschillen. Neem het veld 'datum', dit kan door de ene instelling worden ingevuld wordt als een periode met een begin- en eindjaar, en door een andere alleen maar een specifieke datum.

Homogene metadata realiseren met alle betrokken erfgoedinstellingen bleek niet haalbaar en voor de meeste instellingen evenmin wenselijk. Homogene gegevens zijn een belangrijk kwaliteitselement maar zou voor het realiseren van een overkoepelende beeldbank een grote hinderpaal zijn. EVA werkte met de Dublin Core metadata element set (DCMES). Deze erkende ISO-standaard (ISO15836:2003) bestaat uit 15 gegevensvelden. In de voorstellen voor een vernieuwde EVA-website blijft Dublin Core het uitgangspunt maar er zou niet meer gestreefd worden naar een consensus over de inhoud, betekenis en wijze van invullen van alle velden; de concordantie tussen DC en de eigen beschrijvingsvelden van de instelling volstaat. Twee velden moeten steeds worden ingevuld; voor de andere zijn de erfgoedinstellingen vrij. De twee verplichte velden zijn de velden *Description* en *Identifier*. Description moet zoveel mogelijk relevante informatie over de inhoud bevatten omdat dit de basis is voor de zoekopdrachten van de gebruikers. Identifier is het unieke identificatiegegeven van het digitale beeld.

3.4.3 Hoe verloopt de overdracht van de archieven naar de beeldbank?

XML als uitwisselingsformaat⁹²

Het gebruik van XML staat centraal in het EVA-systeem.⁹³ Wanneer beschrijvingen uit verschillende applicaties in eenzelfde XML-formaat kunnen worden omgezet, verloopt de transactie naar een centrale on line server verder eenvoudig en snel. De eigen databestanden van de archief- of erfgoedinstellingen moeten beschrijvingen kunnen exporteren in +XML-bestanden gestructureerd volgens de EVOLite DTD. Deze DTD is de XML-toepassing van het EVOLite metadata-schema.

Op basis van deze DTD maakt het proces op de server in de lokale archieven automatisch XML-documenten aan. Per EVO wordt er een XML-document gemaakt. Hieronder geven we een voorbeeld van zo'n XML-document waarin een EVO wordt beschreven. Bemerkt de tweede regel van het document die de relatie legt naar het digitaal bestand met de EVOLite DTD.

```
<?xml version="1.0" encoding="UTF-8" ?>
<!DOCTYPE EVOLite SYSTEM "EVOLite.dtd">
<EVOLite xml:lang="nl" ID="SAA100008652">
  <title>Twaalfmaandenstraat, ontplofing huis Verpoorten, Antwerpen, 1865</title>
  <description>Twaalfmaandenstraat, ontplofing huis Verpoorten</description>
  <timeperiod>
    <beginyear>1865</beginyear>
    <endyear>1865</endyear>
  </timeperiod>
  <geography>Twaalfmaandenstraat</geography>
  <geography>Antwerpen</geography>
  <subject>Gebeurtenissen</subject>
  <subject>Stadszichten</subject><subject>Rampen</subject>
  <subject>Straten</subject>
  <archive abbrev="SAA">Stadsarchief Antwerpen</archive>
  <location type="refimg">sa011025.jpg</location>
  <location type="thumbnail">sa011025.jpg</location>
  <relation>FOTO OF 07076 neg. 628 b 1</relation>
</EVOLite>
```

Binnen het archief gebeurt de beschrijving van een EVO via de eigen invoermodule of -applicatie. Vervolgens moeten de images en beschrijvingen worden geëxporteerd naar EVOLite zodat ze kunnen worden opgenomen in de EVA-beeldbank. In de lokaal beheerde databank worden alle beschrijvingen verzameld en op gezette tijden worden de nieuwe of bijgewerkte records uit de databank geëxporteerd.

⁹² XML staat voor *Extensible Mark-Up Language*. Een XML document is leesbaar voor zowel mensen als machines. Dit betekent dat een XML document een grote mate van platformafhankelijkheid bezit; er is geen specifieke – platformgebonden – applicatie nodig om het formaat te kunnen lezen (in tegenstelling tot bijvoorbeeld het DOC-formaat van Microsoft). Een XML document bestaat in feite uit een boomstructuur waarvan het kleinste definieerbare deel een element wordt genoemd. De definitie – het opleggen van beperkingen en de syntax – van deze elementen en de boomstructuur wordt gespecificeerd in een DTD of *Document Type Definition*.

W3C, *Extensible Markup Language (XML) 1.0 (Third Edition)*. W3C recommendation 04 february 2004., 2004, <http://www.w3.org/TR/REC-xml/> (7 juni 2005).

⁹³ R. VAN HORIK, Archives and photographs: the European Visual Archive (EVA) project. In: *Cultivate Interactive*, nr. 3 (2001), <http://www.cultivate-int.org/issue3/eva/> (1 april 2005).

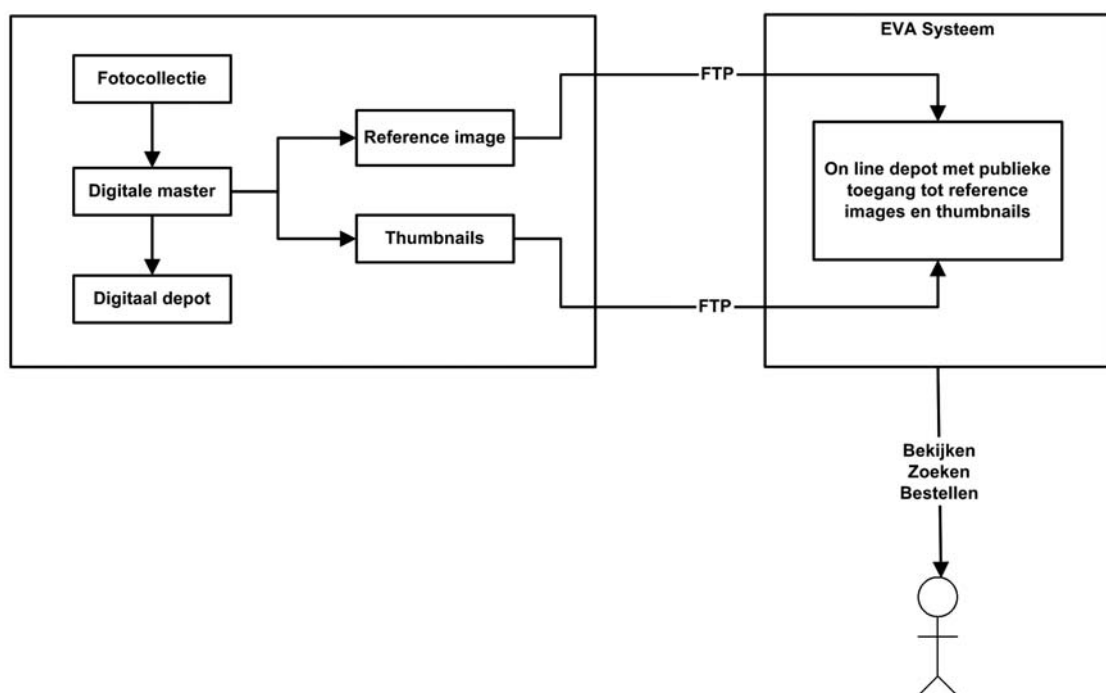
Overdracht van gegevens en digitale afbeeldingen via FTP ⁹⁴

Het uitgangspunt is dat de archief- of erfgoedinstellingen zelf de eigen datasets, databestanden of applicaties (kunnen) blijven gebruiken. Wanneer er een nieuwe digitale beschrijving voorhanden is, moet die bovendien op een eenvoudige en snelle wijze kunnen worden opgenomen in de EVA-beeldbank.

Nadat de datasets conform de EVOLite DTD in XML zijn gegenereerd, worden deze EVOLites met bijhorende referentie-images en thumbnails via FTP naar de EVA-server getransfereerd.

Dit gebeurt in batch; dit wil zeggen dat dit conversieproces op een geautomatiseerde wijze een groot aantal beschrijvingen in de lokale databank wordt uitgevoerd. Op de EVA-server worden de beschrijvingen toegevoegd aan de on line databank. Daartoe worden de gegevens uit de EVOLite XML-bestanden geconverteerd naar het datamodel van die databank.

Het diagram hieronder toont de finaliteit van de verschillende digitale bestanden in het EVA systeem. ⁹⁵



Het EVA-systeem.

⁹⁴ RFC 765, File Transfer Protocol Specification, 1980, <http://www.faqs.org/rfcs/rfc765.html> (7 juni 2005)

⁹⁵ R. VAN HORIK, Archives and photographs: the European Visual Archive (EVA) project. In: *Cultivate Interactive*, nr. 3 (2001), <http://www.cultivate-int.org/issue3/eva/> (7 juni 2005).

EVAmp, 2003 - 2004: harvesting model

Voor EVAmp werd een evaluatie van het systeem uitgevoerd en aanpassingen voorgesteld. Deze voorstellen zijn gebaseerd op de resultaten van de ondervraging van gebruikers en erfgoedinstellingen, en op de veranderde technologie. Vier jaar na het ontwerpen van EVA zijn er heel wat meer beeldbanken op het internet verschenen.

Het tweede EVA-project had vooral een marktonderzoek tot doel en het formuleren van voorstellen om het aansluiten bij EVA aantrekkelijker te maken. Momenteel zijn deze voorstellen (nog) niet gerealiseerd wegens gebrek aan belangstelling van de technische partners van het project. Voor instellingen die belangstelling hebben om beeldbankinitiatieven te ontwikkelen, menen we dat de conclusies van dit onderzoek wel relevant zijn.

Vooraleer aan te sluiten bij een beeldbankinitiatief, bleken erfgoedinstellingen ook volgende overwegingen te maken.

- Het reeds geleverde werk voor het maken van beschrijvingen of het scannen van documenten of foto's, moet kunnen worden opgenomen in de beeldbank. Wanneer aansluiten bij een beeldbankinitiatief zou betekenen dat de beschrijvingen moeten worden bewerkt, aangevuld of opnieuw gemaakt, zullen de meeste instellingen niet deelnemen.
- Veel erfgoedinstellingen werken met eigen applicaties en technologie. Een beeldbankinitiatief dient er rekening mee te houden dat de instellingen met deze eigen technologie zullen blijven werken.
- De meeste erfgoedinstellingen wensen zelf te kunnen beslissen welke gegevensvelden ter beschikking worden gesteld van de beeldbank. Dit kan de helderheid van de beeldbank ernstig verminderen.

Hiervoor werd reeds gemeld dat het verkrijgen van een volledige consensus over alle gegevens, niet werd weerhouden als voorwaarde om deel te nemen aan de beeldbank.

Op basis van deze vaststellingen werd gesteld dat een nieuw beeldbanksysteem op volgende uitgangspunten moet zijn gebaseerd:

1. 'Harvesting' of 'oogsten' als methode om de interoperabiliteit tussen verzamelingen of bestanden te realiseren.
2. Een zeer eenvoudig gestandaardiseerd metadata-formaat om de grote verschillen in beschrijvingsgegevens te overbruggen.
3. Een open source platform voor de ontwikkeling van de beeldbank als noodzakelijke voorwaarde voor de betrouwbaarheid en duurzaamheid van het systeem.

Harvesting

Deze technologie geeft instellingen de mogelijkheid om zelf de volledige controle en beheer over de digitale bestanden te behouden. Zij moeten wel een aantal basisafspraken maken zodat een gezamenlijk initiatief zoals een beeldbank kan worden gerealiseerd. Bovendien wordt deze technologie ondersteund door het *Open Archives Initiative* (OAI) dat het protocol aanlevert om de interoperabiliteit te realiseren: *Open Archives Initiative— Protocol for Metadata Harvesting* (OAI-PMH). OAI-PMH ondersteunt vrijwel alle gegevenssets maar noemt Dublin Core als preferentieel metadata-formaat; wat een troef is voor het EVA-systeem dat gebaseerd is op de Dublin Core metadata element set (DCMES).

Het Open Archives Initiative - Protocol for Metadata Harvesting onderscheidt twee rollen, die van de *data provider* en die van de *service provider*. Een data provider is elke organisatie of instelling die toegang verschaft tot de eigen verzamelingen of gegevensbestanden (bijvoorbeeld het stadsarchief Antwerpen); een service provider is een organisatie die gegevens opvraagt (bijvoorbeeld de EVA-website). Op regelmatige tijdstippen kan een service provider de (bestanden van de) data provider raadplegen en verifiëren op nieuwe gegevens. Die gegevens kunnen dan worden afgeladen naar de gegevensverzameling van de service provider. Het oogsten kan gebeuren bij meerdere instellingen en worden voorgesteld alsof het één grote beeldverzameling is.

Instellingen die op basis hiervan aan een beeldbank deelnemen, moet volgende handelingen stellen.

1. Mogelijk maken dat de beschrijvingen in XML kunnen worden aangeleverd (zoals in het eerste EVA-project).
2. Aanleveren van de digitale beelden in een formaat dat kan worden verwerkt door een internet browser zoals Netscape Navigator of Internet Explorer. De defacto standaard is momenteel het JPEG-formaat.
De instelling bepaalt zelf de specificaties van het JPEG-beeld maar de identificatie van het digitale beeld moet een uniek webadres zijn in de vorm van een URL (Uniform Resource Locator).
3. Beschikken over de software die harvesting mogelijk maakt: de data provider software.
Heel wat data provider software kan werken met courante gegevensbestanden zoals MS Access, MS SQL server of Oracle.

De vernieuwde beeldbank zou volgende voordelen bieden.

- Een gebruiker kan met een enkele zoekopdracht verscheidene verzamelingen bevragen.
- De verzamelingen worden als een geheel voorgesteld zonder dat deze moeten worden samengevoegd in een groot gegevensbestand waarvan het beheer bij een derde partij zou berusten.
- Erfgoedinstellingen kunnen zelf een eigen beeldbank behouden en tegelijkertijd participeren in andere beeldbankinitiatieven om zo een groter publiek te bereiken.
- De inspanning om aan te sluiten bij een beeldbank van dit type vraagt geen bijkomende ICT-ontwikkelingen.
- Het werken met de data provider software maakt het werk om gecorrigeerde beschrijvingen op een eenvoudige en geautomatiseerde wijze toe te voegen aan de beeldbank.
- De complexe handelingen van gegevens uitvoer via FTP zijn niet meer nodig.
- Het synchroniseren van de bestanden wordt een geautomatiseerd proces.
- Een laatste voordeel van deze technologie is het feit dat de bestelmodule waarbij een gebruiker via de EVA-website reproducties van een foto bestelt, hier eveneens mee kan werken.



Modezaak Le Manteau, Meir 80, 1950 (stadsarchief Antwerpen, OF 10870).

3.4.4 De web front-end en de gebruiker

Het EVA-systeem bestaat niet alleen uit een on line databank en een on line depot met digitale afbeeldingen. Er is een *interface* nodig zodat bezoekers op een gebruiksvriendelijke en transparante manier de EVA-beeldbank kunnen bevragen en bekijken. Er bestaan heel wat oplossingen maar veruit de meest voor de hand liggende en geschikte oplossing is via een website. Daarom is er een EVA-website ontwikkeld die toegang bood tot de on line collectie.⁹⁶ Dit is het publieke deel van het EVA-systeem: de *front end*. De databank, de invoermodule en het on line depot noemen we de *back end*.

De EVA-website wil aan de wensen en noden van het publiek voldoen. Daartoe zijn volgende functionele eisen geïdentificeerd bij de eindgebruiker en vervolgens geïmplementeerd.

- ↪ Zoeken van beelden.

De eindgebruiker krijgt de mogelijkheid om op een eenvoudige manier de beeldbank te doorzoeken. Het is daarbij mogelijk om volgende criteria aan te wenden: trefwoord / onderwerp, naam van de fotograaf, periode en geografische locatie. Het gebruik van *Booleaanse* functies (AND, OR) wordt eveneens ondersteund. Het systeem kan een query in de taal van de gebruiker verwerken.

⁹⁶ <http://www.eva-eu.org/nl/>

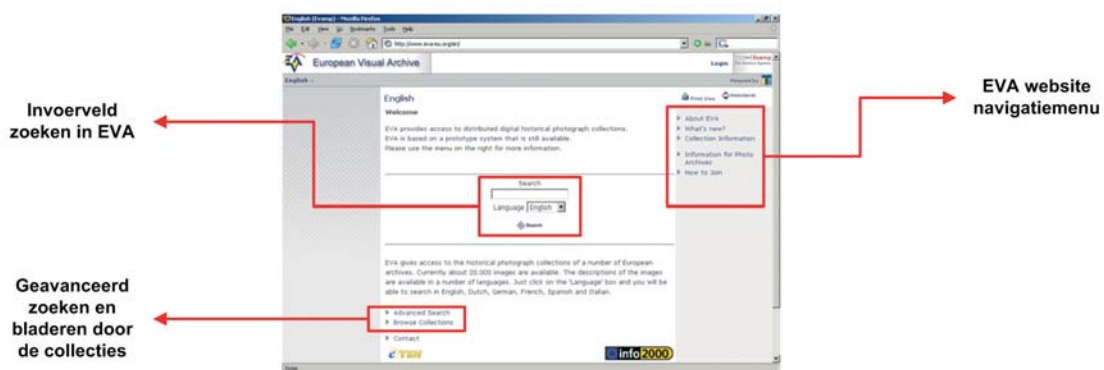
→ Bekijken van beelden.

De eindgebruiker krijgt de mogelijkheid om de collecties te bekijken. Relevante resultaten van een query worden getoond als een galerij met thumbnails. De thumbnails zijn tegelijk *hyperlinks* naar aparte pagina's met een *referentie-image* van betere kwaliteit en de metadata uit de EVA-databank. De gebruiker kan kiezen of enkel de metadata of enkel de afbeelding of beiden worden getoond. De metadata wordt getoond in de originele taal. Tot slot komt er ook een *copyright notice* in beeld.

→ Bestellen van beelden.

Wanneer de eindgebruiker een reproductie wil kopen, dan kan hij via de website een bestelling plaatsen. Het systeem informeert de gebruiker: de bestelling wordt niet bij het 'EVA-project' geplaatst maar wel bij het lokale archief waar de foto wordt bewaard. De klant wordt ook geïnformeerd over de verschillende voorwaarden waaraan de bestelling onderhevig is. Er wordt nog eens een aparte disclaimer in verband met de auteursrechten getoond. Vervolgens moet de gebruiker zijn gegevens via een formulier doorgeven. Een belangrijk veld is de wijze van levering: een reproductie van hoge kwaliteit kan digitaal worden geleverd via e-mail of cd-rom, maar ook analoog als een fotografische afdruk. Bovendien worden de technische mogelijkheden zo benut, dat de gebruikers slechts eenmaal (bij de eerste bestelling) zijn gegevens moet ingeven. Het systeem stuurt vervolgens de bestelling voor verdere afhandeling door naar het archief.

De EVA-website bestaat uit 11 aparte dynamische pagina's of onderdelen die onderling met elkaar verbonden zijn. De eigenlijke inhoud wordt dynamisch gegenereerd op basis van de resultaten van een query. De *homepage* vormt het sluitstuk van de EVA-website. Wanneer een bezoeker de URL van de website intikt dan is dit de eerste pagina die hij te zien krijgt. Het hoofdelement op deze pagina is een zoekformulier dat uit twee velden bestaat; een veld waarin zoektermen (gecombineerd met Booleaanse functies) kunnen worden ingevoerd en een veld om de taal te kiezen waarin de query moet worden verwerkt. Via de link *geavanceerd zoeken* kan men zoeken op meerdere velden (locatie, fotograaf, periode). Tot slot is er een menu voorzien naar informatie over EVA (deelname aan EVA, beknopte bibliografie, ...).



De EVA-website.

*QTE: Query Translation and Expansion*⁹⁷

Een van de doelstellingen van het project was om de fotoarchieven voor een breed publiek in Europa open te stellen. Daarbij stelt zich meteen een onderzoeksvraag over de taal waarin de geïnteresseerde de beeldbank kan ondervragen, de taal van de resultaten van een zoekvraag, de taal van de beschrijvingen die in de lokale archieven of bewaarinstellingen worden gemaakt.

Het eerste uitgangspunt was dat het beschrijven van de foto's in de bewaarinstellingen gebeurt in de eigen, lokale taal van de deelnemende archiefinstellingen: bij EVA was dat dus in het Nederlands of het Engels. Dit is de beste garantie voor de goede kwaliteit van de beschrijvingen. Ook zou het vertalen van de beschrijvingen in de individuele instellingen een te grote werkbelasting betekenen. Tweede uitgangspunt was dat de talen van partners van het EVA-project (Nederlands, Engels, Duits) ook de zoektaalen zijn wanneer een geïnteresseerde de beeldbank bevroegt. Derde uitgangspunt was dat een ondervraging in het Engels, ook de relevante zoekresultaten uit de Nederlandse beschrijvingen uit Antwerpen moest kunnen opnemen om te vermijden dat een bezoeker dezelfde query in verschillende talen zou moeten lanceren om zeker te zijn dat hij of zij alle relevante resultaten in de EVA-beeldbank te zien krijgt.

Voor de vertaling van de beschrijvingen en de zoekvragen werd een specifieke softwarecomponent voor het EVA-systeem ontwikkeld. De grootste handicap bleek het ontbreken van bruikbare, meertalige thesauri. Voor het EVA-systeem werd aan de hand van de bestaande beschrijvingen van de foto's, een trefwoordenlijst met 6 000 termen geëxtraheerd. Deze werden vervolgens vertaald in de andere taal. Naast het EVA-lexicon werd er ook een *expansion list* gemaakt. Deze lijst definieert de bredere relaties tussen de verschillende termen in het lexicon. Zo zijn de termen 'gebouw' en 'kerk' in de expansion list met elkaar gerelateerd: wanneer in een vraag of query 'gebouw' wordt gelanceerd dan zal een deel van de relevante resultaten uit kerken bestaan. Zonder deze expansion list zou het systeem zich beperken tot louter resultaten die expliciet als gebouw worden herkend. De expansion list wordt op eenzelfde manier gebruikt om synoniemen aan elkaar te koppelen.

Het sluitstuk van het vertaalmechanisme is de QTE of *Query Translation and Expansion* softwarecomponent die werd ontwikkeld door het Duitse Saillabs en die tussen het databankbeheersysteem op de EVA-server en de website staat. Wanneer een bezoeker een zoekvraag lanceert, wordt deze eerst verwerkt door de QTE-component. Deze zoekt in het lexicon en de expansion list naar vertalingen, synoniemen of bredere gerelateerde termen. Vervolgens construeert QTE op basis van de gevonden resultaten een nieuwe, machine begrijpbare query (of zoekstring) in SQL.⁹⁸ Deze wordt vervolgens door het databankbeheersysteem van EVA verwerkt waarna de relevante resultaten aan de gebruiker worden getoond via de EVA-website.

De basis van een dergelijke *multilingual search facility* is het lexicon en de expansion list. Veel hangt af van de kwaliteit hiervan. Ze moeten zoveel mogelijk overeenstemmen met de metadata die via de XML-bestanden in de EVA-databank worden ingevoerd. Daarom moet er bij het implementeren van een dergelijke mogelijkheid grote zorg worden besteed aan het correct beschrijven van de beelden enerzijds en het opstellen van het lexicon en de expansion list anderzijds.

⁹⁷ EVA CONSORTIUM, *The European Visual Archive - EVA, Functional and technical design*, 2000, <http://192.87.107.12/eva/uk/texts/wp62.pdf> (8 juni 2005).

⁹⁸ SQL of 'Structured Query Language' is een ANSI/ISO standaard taal voor een relationeel 'database management system' (DBMS). In eenvoudige termen wil dit zeggen dat SQL een gestandaardiseerde taal is die kan worden gebruikt voor taken zoals het bevragen en het aanpassen van informatie in een relationele databank. SQL kan met vrijwel alle moderne relationele databankproducten worden gebruikt, <http://nl.wikipedia.org/wiki/SQL>.

Omdat het bijhouden van het lexicon en de expansion list een manueel en intensief werk is, is dit de achillespees voor de meertalige beeldbank. In het vervolgproject EVamp werd onderzocht of de meertaligheid op een andere wijze kon worden gerealiseerd. Een van de belangrijke conclusies is dat zoveel mogelijk met beschikbare meertalige woordenlijsten moet worden gewerkt. De voorkeur gaat bovendien naar open source oplossingen. Voor het Engels, Frans, Duits, Nederlands, Spaans en Italiaans bleken er publiekelijk beschikbare woordenboeken te bestaan. Door deze woordenboeken te verbinden met *Cross Language Information Retrieval* (CLIR) kunnen meer talen worden ondersteund en is veel minder manueel werk nodig.



Rode tram, 1850 - 1900

Stadsarchief Antwerpen, OF 07122.

Besluit: Lessen geleerd uit het realiseren van de EVA-beeldbank

Uit de EVA-projecten werden een aantal lessen getrokken.

1. Een duidelijke bepaling van het project betekent dat op de hierna volgende vragen een antwoord kan worden gegeven.

- Waarom digitaliseren? Ten dienste van toegankelijkheid, van betere bewaring van de originele documenten, of een omwille van een combinatie van deze twee?
- Welke materialen of documenten zullen worden gedigitaliseerd?
- Welk thema of onderwerp zal worden behandeld? Indien het om meerdere onderwerpen gaat: hoeveel?

Deze antwoorden moeten ook de basis zijn voor het verder uitwerken van het plan van aanpak.

2. Ken uw gebruikers en breng hun belangstelling en wensen in kaart.

- Voor wie is dit project bedoeld? Wat is de doelgroep?
- Wat zijn de verwachtingen of wensen van de gebruikers over de thema's die worden behandeld, het gebruik en de werking van internet, het verkrijgen van reproducties en de wijze van vragen stellen?
- Kunnen nieuwe gebruikers worden aangetrokken, en op welke wijze?
- Wat zijn de noden van verwante sectoren zoals die van multimedia, toerisme of ontspanning?

Deze informatie is van zeer grote waarde en de basis voor het welslagen van het project. Het verdient aanbeveling om het bevragen van de gebruikers niet te beperken tot de aanvang van het project, maar dit te herhalen tijdens het verloop ervan aangezien de wensen van de gebruikers en de technologische mogelijkheden evolueren.

3. Bestudeer en vergelijk lopende projecten.

Erfgoedinstellingen, musea en archieven moeten zo goed mogelijk op de hoogte zijn van lopende gelijkaardige projecten. In het domein van digitalisering werd reeds heel wat gepubliceerd en waardevolle handboeken zijn beschikbaar (bijvoorbeeld van de Research Libraries Group, National Preservation Office, ECPA). Studiedagen en workshops zijn een ideale gelegenheid om van gedachten te wisselen over *best practices*. Advies van collega's is een belangrijke aanvulling bij de informatie die wordt verstrekt door commerciële ondernemingen die digitaliseringsdiensten aanbieden.

4. Selectiecriteria duidelijk vastleggen.

Niet alle foto's of ander erfgoedmateriaal komen in aanmerking om te digitaliseren en de kans is klein dat er voldoende financiële middelen beschikbaar zijn om alles te onder handen te nemen.

De juiste selectie van het materiaal is dan ook een belangrijke succesfactor.

De criteria die worden gehanteerd bij het selecteren van het materiaal moeten heel duidelijk worden vastgelegd (bij voorkeur in een *decision making matrix*) en heel expliciet worden bekendgemaakt aan de gebruikers van de beeldbank. Zo weten gebruikers of de beeldbank relevant is voor het onderzoek dat zij voeren.

5. Het gebruik van standaarden.

Internationaal erkende en aanvaarde standaarden zoals voor het beschrijven of catalogiseren van documenten, foto's of objecten, geven een grote meerwaarde aan het project omdat de uitwisseling en transparantie van het aangeboden materiaal wordt gewaarborgd.

6. Beheer van de auteursrechten.

Van elk gedigitaliseerd object of document moet de auteursrechtelijke informatie gekend zijn.

Bij een digitaliseringsproject is het verduidelijken van de auteursrechtelijke beperkingen één van de allereerste werkzaamheden. Een digitaliseringsproject zou nooit in het gedrang mogen komen omdat niet met de bescherming van deze rechten werd rekening gehouden.

Een duidelijke strategie voor het beheer van deze rechten is dan ook een waardevol instrument en van belang bij het bepalen van de selectiecriteria en de communicatie naar de gebruiker.

7. Gegevensbescherming (data protection).

De wettelijke verplichtingen inzake het beheer van gegevensverzamelingen moeten goed gekend zijn en worden opgevolgd. Dit houdt informatie in over de aard van de gegevens, wie deze waar bewaart en hoe het gebruik door derden van deze gegevensbestanden is geregeld.

8. Bescherming van persoonlijke levenssfeer.

Vooraf tijdens het selecteren van het materiaal en het publiceren (bijvoorbeeld op internet) van de beschrijvingen, dient terdege rekening te worden gehouden met de bescherming van de persoonlijke levenssfeer van levende personen; zoals ook reeds geldt voor de bescherming van de auteursrechten. Algemene principes van mensenrechten gelden ook in het hanteren van een niet-discriminerende taal in de beschrijvingen en in het niet publiceren van foto's die misleidend kunnen zijn.

9. Beleid inzake prijzen of tarieven.

Informatie over prijzen voor reproducties moet op een duidelijke manier worden verstrekt. Het EVA-project heeft ervoor gekozen om de raadpleging van de beeldbank, het opzoeken en bekijken van foto's niet-betalend te houden. Een prijzenpolitiek kan zeer goed een onderscheid maken in groepen gebruikers; zo kunnen voor studenten andere tarieven worden gehanteerd dan voor commercieel gebruik.

10. Aandacht voor de publieke dienstverlening.

Tijdens het digitaliseren, zijn de documenten tijdelijk niet beschikbaar voor raadpleging in de leeszaal of voor andere publiek gebruik. Dit aspect mag niet over het hoofd worden gezien; de niet-beschikbare periode dient zo kort mogelijk te zijn en de gebruikers moeten steeds zorgvuldig worden geïnformeerd.

De mate waarin er een goed uitgewerkte workflow bestaat, maakt het mogelijk om steeds over precieze informatie te beschikken en de gebruikers te kunnen informeren.

11. Veilige manipulatie van het fotografisch materiaal.

In het handboek wordt vanzelfsprekend veel aandacht besteed aan de correctheid van alle technische handelingen van het digitaliseringsproces. Het manipuleren van de kwetsbare foto's vraagt dezelfde zorgvuldigheid en aandacht. Alle medewerkers dienen op de hoogte te zijn van de risico's die originele foto's bedreigen en kunnen beschadigen. Niet alleen het manipuleren van de foto's maar ook de wijze van bewaren tijdens de bewerking en zolang ze uit de oorspronkelijke bewaarplaats verwijderd zijn, moet in goede bewaaromstandigheden gebeuren.

12. Betrouwbaarheid van het scanningmateriaal.

Regelmatige controles van de betrouwbaarheid en goede werking van het scanningmateriaal moeten opgenomen zijn in de workflow en het werkprogramma. Dit kan manueel of geautomatiseerd gebeuren en is een belangrijke voorwaarde voor de integriteit en consistentie van het digitale beeldmateriaal.

13. Bepaling van resolutie- en compressievereisten.

Na het bepalen van de doelstellingen van het project moet de strategie ten aanzien van de te hanteren resolutie en compressie worden vastgelegd. Het documenteren van de gekozen instellingen is een belangrijke vereiste.

14. Een consistente nummering van de digitale beelden.

Vanuit praktische standpunt is een eenvoudige sequentiële nummering aangewezen; deze kan worden gecombineerd met een complexere codering. In ieder geval mag niet uit het oog worden verloren dat het nummer van het digitale beeld een essentieel gegeven is dat steeds moet toelaten het origineel te traceren en dat het digitale beeld doorheen het hele digitaliseringsproces moet kunnen worden gevolgd en gedocumenteerd. Het nummer is tenslotte ook de link met beschrijvingen in de databank en in de beeldbank.

15. Controle van het eigenlijke scannen.

Het controleren van het resultaat van het scannen dient te gebeuren door een andere persoon dan diegene die betrokken is bij het scanningproces. De controle heeft zowel betrekking op de kwaliteit als op de inhoud.

16. Op het scherm steeds de beschrijving bij het beeld.

Wanneer een digitaal beeld op het scherm verschijnt, dient de basisbeschrijving meteen te worden getoond zonder dat de gebruiker daarvoor naar een afzonderlijke pagina moet gaan.

17. Metadata essentieel voor het beheer.

Voor een succesrijk digitaliseringsproces is het essentieel dat het vastleggen van de metadata een routinematige handeling is. Wanneer gegevens niet meteen beschikbaar kunnen zijn voor de gebruikers, moet de erfgoed- of archiefinstelling steeds over de metadata beschikken.

18. Een exit-strategie.

Elk digitaliseringsproject dient rekening te houden met de volgende vragen.

- a. Wat gebeurt er aan het einde van het project?
- b. Waar worden de gegevens bewaard gegenereerd door het project?
- c. Wie is belast met het verder onderhouden van de gegevens?

Kunnen of mogen de gegevens ook in andere projecten, beeldbanken of portaalsites worden aangeboden?

cDAVID - DIGITAAL ERFGOED

Stadsarchief Antwerpen

Venusstraat 11

2000 Antwerpen

Tel. + 32 (0)3 206 94 11

Fax + 32 (0)3 206 94 10

E-mail: cDavid@stad.antwerpen.be

Website: www.cdavid.be

Illustraties: © Stadsarchief Antwerpen en fotografen

Eindredactie: Inge Schoups, Willem Vanneste, Roberte Van Haute

Ontwerp en opmaak: www.gestalte.be

Drukwerk: Roels Printing, Antwerpen

© cDavid, 2005

Met steun van de Vlaamse overheid, de Vlaamse minister voor cultuur, jeugd, sport en Brussel

Met dank aan Expertisecentrum DAVID

De uitgever heeft er naar gestreefd de rechten met betrekking tot de illustraties volgens de wettelijke bepalingen te regelen.

Degenen die desondanks menen zekere rechten te kunnen doen gelden, kunnen zich alsnog tot de uitgever wenden.