DIGITISE MORE, PAY LESS - Optimising the preparation for digitising large collections of images - Case study Photo collection Netherlands Institution of Sound and Vision

Margot Knijn; Netherlands Institution of Sound and Vision; Hilversum, The Netherlands

Abstract

Through the Images for the Future project the Netherlands Institute for Sound and Vision was given the opportunity to digitise large parts of their collections. The sheer size and numbers of material, combined with the necessity of a European Tender procedure, forced Sound and Vision to rethink work processes and workflows.

The photographic department has by now gathered experience in two Tender procedures where the digitisation of more than a million negatives was commissioned.

The general conclusion is that time and money spend on preparing your collection and your workflow with the digitisation process in mind will be earned back easily as the supplier will be able to optimise and automate their work process and more precisely calculate their risk, which will result in a lower price.

This approach can be applied to smaller digitisation project as well, with the same result: Digitise More, Pay Less.

Introduction

In 2006 the audiovisual archives in The Netherlands raised the alarm: if the government would not intervene, the Dutch audiovisual archives would soon no longer have a collection, and the Dutch audiovisual heritage would be lost. Vinegar syndrome was attacking both the film and photographic material, equipment for obsolete video and audio formats was becoming rather rare, and expertise was dying away. Furthermore, in its analogue form, the audiovisual heritage was only available for a lucky few, and offered limited accessibility and search functions.

Collection holders EYE Film Institute Netherlands, the National Archive (NA) and the Netherlands Institute for Sound and Vision (Sound and Vision), combined forces with think tank the Netherlands Knowledgeland Foundation (KL) and the consortium *Images for the Future* was formed (www.imagesforthefuture.org).

Figure 1. *Logo Beelden voor de Toekomst (Images for the Future)*

Arguing that the costs for creating the visual history of the past 100 years must have run into the billions and represented an unprecedented educational, cultural, and economic resource, the consortium received a budget of 154 million Euros from the FES (Fund for the reinforcement of Economic Structure) to restore, preserve, digitise, and make available 137.200 hours of video, 22.510 hours of film, 123.900 hours of audio, and 2.9 million photos over a period of 7 years.
The main goal of the project is realising maximum accessibility of the audiovisual material for the targeted user groups (educational institutions, the general public, and the creative sector).

But before you can make any material digitally accessible, you have to have it digitally mastered. And with such quantities, how do you go about it? The daunting numbers, and the fact that - like in all large archives - detailed information was not available for all material, made the project a challenge. Soon it became obvious that it was necessary to develop a new approach, and we had to think BIG to succeed.

This paper concentrates on the approach for digitising the photographic collection of the Netherlands Institute for Sound and Vision, and how thinking ahead, developing workflows and doing research leads to technical innovation and eventually to lower prices: Digitise More, Pay Less.

The Photographic collection

For six decades, the Netherlands Institute for Sound and Vision and its predecessors (NTS/NOS/NOB, AVAC and NAA) have maintained a collection of Dutch television heritage. From the very beginning the NTS employed photographers to make photographs during all television productions, both in the studio (majority) as well as on location.

Figure 2. *On location, crew and technique in view (1962)*

From the years in which programmes were broadcast live and could not be archived, the photographs are the only remaining evidence. In later times they contribute highly to our knowledge of the productions as they usually show a different angle and perspective than the moving images and often include the audience.

The Photographic Department of the NTS (and later NOS/NOB and FE-NOB-FOTO BV) serviced all the different broadcasters by delivering prints on demand. All broadcasters had contact sheets of the images shot during their productions, while the negatives remained in the archives. Besides photographing actual productions, the photographic department also made photographs during special events (like the Queen visiting the studios, or the annual arrival of St. Nicholas in November), supporting material for productions, stage designs and portraits of production crew and broadcasting staff. In the last 10 years related photographic collections were added to the archive through merging and acquisition.

Figure 3. *The annual arrival of St. Nicholas, a variation on Santa Claus*

The Photographic Collection contains about 2.5 million items, mainly negatives. Within *Images for the Future* Sound and Vision is to digitise 1.2 million negatives.

European Tender Procedures

Spending public money has – understandably - all kinds of constraints. For Sound and Vision it means that all commissions over about 200.000 euro (about 280.000 dollars) need to be done through a European Tender process. This process (originally designed for large building contracts) aims for a transparent procedure where national (or other) preferences should not interfere and where suppliers from Europe and beyond get a fair chance to obtain the contract. It of course also stimulates competition, and therefore increases the chance for a better balance between price and quality. European Tenders also cost money, because they involve lawyers and consultants and generally a lot of work.

Practically a Tender means that the institution needs to write a so-called Descriptive Document: a highly detailed document that describes all needs and wants. The Invitation to Tender is then published and a deadline is set for the offers. The Descriptive Document of the Tender needs to be very precise otherwise you'll run the risk to find out during the set-up phase that the supplier cannot deliver, or you are soon renegotiating the price because some requirement was left in a grey area or overlooked.

Figure 4. *http://ted.europa.eu where most European Tenders are published*

There are two options to weigh the offers made, which has to be chosen for in advance: either Lowest Price, or Most Economically Advantageous. Most Economically Advantageous actually means you can weigh and give points for each of the answers on the listed Requirements, and thus select on perceived quality. This procedure gives you some feeling of choice, although transparency of course stays a priority.
Sound and Vision however decided in an early phase to concentrate on the awarding criterion of Lowest Price. Not just because it makes life simple after opening the offers (no time-consuming grading of the offers), but because it forces the Tender Document to be hyper precise: once a supplier offers the lowest price you have to be 100% sure you have included all your wishes and demands. When successful, a Lowest Price procedure will give the best price for exactly the quality you want.

In 2008 the Consortium partners of Images for the Future started on a joint Tender for photographic material: Sound and Vision together with the National Archive and EYE. The National Archive had made a start with researching technical requirements and EYE and Sound and Vision added their own research, experience and particularities.

Roughly the Tender consisted of the following areas:
- Requirements on the eligibility of the Tenderer (size of company, experience, legal entity etc)
- Technical Requirements: quality of equipment (benchmarking) and scanning methods
- Requirements for metadata, filenames etc.
- Requirements around work process, workflow etc.

Aim was to make a document that contained all quality requirements and at the same time give the tools to measure if the requirements were met, both in the selection phase as well as during production.

So, on the one hand we knew we had to be very detailed and precise in our requirements and instructions. And together with our consortium partners as well as consultants, we were developing instruments for benchmarking and monitoring quality of equipment and quality of digitising. This however did not solve the issue of the sheer numbers. How to tackle this enormous task?

Thinking BIG

When I got involved and started to work for Sound and Vision, I had just come out of a project at the Netherlands Fotomuseum where we digitised 6.000 negatives in one year for the *Memory of the Netherlands*, a Dutch digitising project where more than 80

institutions have brought together digital content. I realised that at this same pace it would take us 200 years - rather than the remaining 5 years - to digitise the required 1.2 million.

First decision was to definitely outsource the digitising and not do it in-house as a form of risk management.
What followed were several months of getting to know the collection and thinking what could be done to make the process more efficient.

The following areas are examples where time and money efficiency was obtained:

- Selection
- Preparation of material
- Re-thinking digitisation methods
- Challenging fixed thoughts
- Do not overask (quality requirements)
- Workflow
- Monitoring procedures

Selection

To tackle the collection it was first divided in several large homogenous chunks, based on material (6 x 6 black and white negatives, 35mm black and white negatives etc.), subcollection (material from own photographic department, acquisitions) and packaging (envelopes, sheets etc). The collection that had been repackaged 10 years ago, was selected to be digitised first. This was a large collection containing 6 x 6 black and white negatives, all packaged in the same way: a strip of two negatives in Melinex sleeves, about 10 sleeves in an envelope. Offering a homogenous lot would allow a more automated work process at the supplier, thus inviting a lower price.
Not all material had to be digitised, but there was no time to contemplate and select every negative at leisure. Therefore the selection was made on packaging level, thus per envelope. Either the whole content of an envelope had to be digitised, or not at all. Not selected enveloped received a stamp with archival ink: DO NOT DIGITISE. Repetition of images within one film was therefore unavoidable, but with 6 x 6 film photographers used to be relatively economic and selection would need to be done by the end user.

The second Tender we made (which was awarded only this last February) we wanted to digitise the 35mm collection: more than 1.5 million in total and therefore selection was inevitable. Particularly because inherent to the 35mm material there was so much repetition that this could not be left to an end user. Too many images would flood the user. But how to select 35mm negatives, how can you see which image is the best? This material was also not repackaged and was still in ordinary acidic postal envelopes. For this dilemma we developed the thought that the supplier should build an on-line selection tool on which all 35mm images had to be displayed, for us to select. But we could not have all the 1.5 million images scanned first. This was solved by repackaging in transparent sheets, so digital contactsheets could be made and used for selection. Therefore our digitising process informed our choice of packaging. These days first choice for repackaging negatives might be paper, but in this case we packaged in polyester.

Preparation of material

The collection has always been numbered with logical archive numbers, starting with the production year and then a sequential number, for instance 65247, is programme 247 of the year 1965. The length of the archive number is variable; for instance programme 10 of 1965 would be 6510. Thinking ahead of the work process both at our end and at the suppliers end these archive numbers were a liability, leading very easily to mistakes, particularly if they are the basis of filenames for the scans. Therefore all material (on envelope or sheet level) was given a barcode. This meant that all checks of material coming in and out, and all the filenames of the scans, could be based on actually read barcodes rather than typed. The barcodes were put in an additional field in our database, so both the old archive number and the new barcode number are available. The dispatchlist of every batch sent to the supplier has these numbers and is the basis for the required XML.
All filenames are derived from the barcode. The filename is the barcode followed by the sequential number of digitising. To still match with the metadata a Requirement Original identifier was developed: in the XML the supplier needs to give the Original negative number as written on the negative as well as the filename of the scan.

Figure 5*. Original Identifier 61320TD VF_0003, Filename FTA001003546_008*

Re-thinking digitisation methods

Thinking of such a large bulk of homogenous material (560.000 6 x 6 negatives) for the first Tender raised the question of the efficiency of the digitisation method itself. If it would be possible to digitise more than one negative at a time, it would speed things up enormously. I started testing with more than one negative under a digital camera and the results were good. With the current CCD sensors we calculated it should be possible to digitise 6 negatives in one go. Of course the Requirements and measurements for sharpness and Illumination uniformity had to be obtained over the whole plane. This multiple scan method was describe in the Descriptive document as an acceptable option and it was left to the suppler to base their offer on either method. The supplier who offered lowest price chose for the multiple scan method and was able to up their production from an average of 800 scans a day to

2500 scans a day. This was of course reflected in the price they could offer.

Challenging fixed thoughts

The Multiple Scan method seemed to have one major drawback: the maximum resolution that could be managed with the available sensors. Each negative would be digitised at a maximum of 900 ppi, which was under the commonly accepted digitisation level of being able to make an A4-size print (slightly bigger than US letter size) at 300 dpi. It should really be 1200 ppi to achieve that. Was the price benefit worth it to have the collection available in a size that was deemed too small? We then embarked on some testing in collaboration with the Netherlands Fotomuseum and a high-end printer of photographic books (Veenman, Rotterdam, The Netherlands). We had images scanned in the range of 600 to 1200 ppi and had them printed on the off-set machines of the printer all at the same size of A4, so some were printed at 300 dpi, others at 200 dpi and other at 100 dpi. The results were amazing. There was no visible difference between the images printed at 300 dpi or 200 dpi. The printers discovered that the cut-off point lies at 170dpi. So our negatives digitised at 900ppi, could easily be used for prints of A4 or bigger, as long as you don't keep to the unchallenged magic number of 300 dpi. In the same test we tested interpolation. In some circles a dirty word but the truth is that some of the methods work very well. This does not mean that the scans in the archive should be interpolated, but it is possible to instruct users of the images a certain way of interpolation if they would need a slightly larger size than is available.

Do not overask

Speaking to the different digitisers it became clear that a common irritation lies in clients that require standards and/or test that actually have no effect or even deliver a worse image. Once all the quality requirements were in place, like Spatial Frequency Response, DNR, Illumination, Sharpening or not etc. we went through it again and asked ourselves, what is this based on? Does this make sense? In the end we want good images of consistent tone and quality.

Workflow

Great advantage was that within the Images for the Future project Sound and Vision had already developed quite a bit of experience. That means that managing large quantities of files, and adapting the database to facilitate the workflow had been done before.
It had been the experience that the quality control after receiving digital files can be complicated. One of the questions was when preparing the first Tender and dealing for the first time with 560.000 negatives, how can we be sure all the negatives in an envelope are digitised? To solve this problem we actually had two people for three months open all the envelopes, and count the negatives. At the same time they checked for oddities. That way accidental colour negatives between the black and white were found, as well as odd sizes like 4.5 by 6 and 6 x 9.

This way we could ensure in the Tender document that we had manually checked all material and taken out possible problems. After adding the numbers to our database we could also add the actual number of negatives per envelope in the dispatch. A new requirement for the XML was developed: number of expected and found negatives had to be specified, which was then checked with the import of the XML. This also gave a tool for the supplier to automatically check the work. Again, more information for the supplier, leading to a better price.

Another decision was made for the delivery of the digital files. Rather than working with hard drives, all files are pushed by FTP into a designated folder. Resulting in less transport, easier and faster for the supplier, easier and faster on the receiving end as well.

Monitoring procedures

In the Descriptive Document of the Tender it was described in detail how Sound and Vision would perform the quality control checks. This made it easy for the supplier to get an insight in the extent of the work, but also made it possible for Sound and Vision to streamline the monitoring process. Calculating random checks and putting the files selected for random checks in a separate folder was automated. Streamlining the monitoring process and quality controls leads to faster feed back, thus picking up problems in an earlier stage as well as signing off on a batch for payment.

Conclusion

1. Writing and rethinking the technical requirements for digitisation lead to better quality of scans and an easier to control work process.
2. Huge detail in the description of the collection and requirements lead to clarity in expectation, and therefore to optimum risk management.
3. Preparing your collection, and thinking creatively with the digitisation process in mind, leads to great cost reduction.

Together, these approaches will leave you with high quality scans for the best price.

Author Biography

Margot Knijn's background is in photographic studies and art history. She has run several international film development institutions for The European MEDIA programme, and worked as a consultant and project manager for the Dutch National Photographic Museum and The Binger Institute. She works currently as a project manager at the Netherlands Institute for Sound and Vision, The Netherlands.

