# ESTABLISHING A DIGITIZATION PROGRAMME FOR NAMIBIA: PROMISES, PITFALLS AND PROGRESS

**Werner Hillebrecht**
National Archives of Namibia
whillebrecht@mec.gov.na

Presented at the 2nd International Conference on African Digital Libraries and Archives (ICADLA-2), University of Witwatersrand, Johannesburg, South Africa, 14th – 18th November, 2011

## Introduction

Digitization has spread world-wide like an infectious disease and no country has been spared. Guidelines, instructions and policies can be found in abundance on the Internet. Nevertheless, it seems that the same mistakes are made again and again. This paper analyses the experiences of the National Archives of Namibia, which has gone through the whole wide array of pitfalls and mistakes. The paper comes to the conclusion that, under African conditions of very limited resources, certain mistakes are unavoidable until the adequate local conditions have been explored and local solutions have been found.

The National Archives of Namibia started with a digitization programme in 1995, at the initiative of a single archivist without much institutional support. Subsequent loss of staff, including the initiator, and lack of continued IT support, ended this first initiative. It was only almost 10 years later that we seriously started again. In the beginning we only wanted to be able to copy private material where the owners wanted the original returned but agreed to have it digitized for access at the archives. We also wanted to be able to supply scanned photographs because it became increasingly difficult to find a local contractor to print from black-and-white negatives, as had been done previously for photograph requests.

## "Quick and dirty" initial programme

Not only donors but also local politicians and administrators want to see quick results. To sustain their interest, it is important to show something impressive without too much effort. If such a result can be shown, one can tackle the more difficult issues at a reasonable pace. In our case, we chose a "quick and dirty" method to digitise our photograph collection; I need to explain the organisation of the collection in some detail to show why it was the easiest and most obvious target.

Firstly, photographs are always visually appealing and serve a wide variety of users; they are a "sexy" digitization object and therefore often a first target of digitization. In our case, two other considerations added to this appeal. Our photograph collection is catalogued on an electronic database, which eliminates the need to create metadata from scratch. This is

in fact not entirely true because the metadata is inconsistent and deficient in many aspects,[1] but the data is sufficient for some basic needs.

Secondly, there was no need to use an elaborate file-naming convention, as the photographs were numbered consecutively with five-digit numbers from one to twenty-something thousand. This greatly reduced the chance of mistakes in data capture, and the unique identifier for each record was already on the database.

Thirdly, our historical photograph collection, which has grown over 50 years, was organised in a three-part arrangement. We hardly have original negatives; most of the collection stems from donated photograph prints from a wide variety of sources, and those originals never were issued to users. Instead, secondary negatives were taken from those prints and "show prints" were printed from those negatives in a standard size and mounted on cardboard, so that users could easily page through them, select what they needed, and order a print from the negative. All three versions – the original, the negative and the show print – carry the same number but are kept in separate series.

The correct method of digitization is of course to use the best available source, the original negative or original print, in high resolution and TIFF format. That is what we are actually doing *now*, at our own pace as resources permit. But we first used the "quick and dirty" solution, scanning in low resolution, compressed format, from the show prints which have significantly reduced quality because they have gone through two additional analogue photographic processes, and sometimes additional wear and tear from handling by users.

We did this for a number of good reasons. It was necessary to achieve fast presentable results. In our very constrained staffing situation we had to use temporary low-cost staff without too much training and supervision. Using the show prints, we did not have to fear damage to the precious and often very delicate originals from this category of staff. The show prints are also in a more or less standard size and in a single sequence, which speeds up the scanning operation, whereas the originals are in a confusing array of series by size, or culled from albums. Lastly, we chose compressed format because of serious constraints on digital storage space at the time, knowing well that the uncompressed high-resolution scans would come later.

This arrangement, against all accepted rules, has served its purpose. The WINISIS[2] database of photographs needed no additional data entry but only a change of display format to link from the existing photograph number to the identical file number on the server. The result, a photographic database that is easily searchable and can deliver images at a single mouse click, serves well to showcase the photographic holdings and to convince

[1] Having grown historically over decades, the indexing and data entry of details is inconsistent; captions are often in Afrikaans or German instead of English; and despite all efforts since independence, some apartheid terminology still persists in captions and even index terms.

[2] The National Archives of Namibia switched all its databases to Unesco's free CDS/ISIS software in 1995, and later started using the Windows version WINISIS of the same software. The main reason is that we found it extremely adaptable to all new demands, including major data upgrading efforts.

policymakers and administrators on various levels that it is worthwhile to support the digitization process.

While this process was still going on, we embarked upon a more ambitious foreign-sponsored project which, we expected, would enable us to gain experience and achieve professional results. In retrospect, I would say it fulfilled these expectations, through both the positive and negative experiences, because however good the existing guidelines and manuals may be, it is necessary to examine them in practice under local conditions. I think it is worthwhile to describe them here in some detail; a number of participants at this conference probably have been involved in this regional project, and their experiences might have been quite different because of local conditions. What I am going to say is not therefore a representative general evaluation of that project (which for this reason should remain unnamed), but describes an experience very specific to local conditions. It should nevertheless be informative.

**The local conditions**

Some of the conditions we are working under should be mentioned here. Namibia is a large country of over 800 000 sq. km with a small population of approximately 2 million. It was decolonized only 21 years ago, emerging from apartheid rule which had reserved educational opportunities for whites and effectively prevented the vast majority of the black population not only from accessing higher education but also qualified technical education, to an even greater extent than in South Africa itself. As a result, there is a very serious shortage of specialized qualified skills as well as a generally low level of education, despite an overall impressive literacy rate. The small size of the population also is an obstacle to wide diversification of local training for skills that are not needed in large numbers, and which therefore have to be obtained abroad. Despite this local shortage of skills, the government implements a very restrictive immigration policy, which often prevents the recruitment of sorely needed specialists.

The library, information science and archives sector is particularly affected by these conditions. A full local library and information studies training programme was established only after independence, and to date no full archival training is available. Another detrimental factor that came into play was uncompetitive salary grading of professionals in this sector, compared with other professionals in the government service. This has led not only to a high staff turnover and brain drain from the governmental information sector, but also to a most unwelcome effect upon the intake of students into this field of specialisation.

In order to counteract low interest in the programme, the relevant university department had lowered the entry requirements, a measure which helped to boost the student numbers but led to the perception that the subject could be used as an easy entrance ticket to the university in order to pursue other studies later on. The low entry requirements also had a negative effect on the average quality of graduates. As a result of all these factors, the library and archives sector in the Namibian government service is chronically understaffed and under qualified, with sometimes more than 50 % of professional posts vacant,

especially in supervisory positions,. These were (and are) the conditions under which we were starting our digitization programme.

**What do we want?**

Before we made a commitment, we had been window-shopping for opportunities to embark upon a larger-scale digitization effort. We had some general idea of what we did not want. We did not want to have any material taken out of the country. We did not want to have foreign specialists coming in and running the programme themselves, doing some token training and leaving us ignorant as before. We wanted to build local capacity. We did not want to sign away our copyright and have our records displayed only on foreign websites over which we had no control. We did want to keep our own preservation copies instead of relying on storage elsewhere. We did not want to have digitization limited to a very narrow project of a single collection that had to be approved by some external board; we wanted to gain wider experience and to be able to select the material according to our own priorities.

These conditions ruled out quite a number of otherwise attractive opportunities. There is a global interest in access to African information resources that elsewhere has been dubbed, quite appropriately, the "digital scramble for Africa" (Limb, 2005), and some of the programmes are indeed rather dubious. The project which looked suitable and which we grasped as our entry ticket to a proper digitization programme did eventually achieve the aim, despite all negative experiences which, in hindsight, have helped us to move forward. This was a project to build a web resource on the history of the liberation struggles in Southern Africa, based and funded in the US but with cooperating national steering committees in all participating countries. An important factor was that this was a regional project, which assisted us to be in contact with developments in neighbouring countries, something that we sorely miss since it ended. Smaller countries have too many constraints to face the task ahead alone. With particular regard for long-term preservation, policy coordination and adequate web presentation, we have to work closely together as a region.

Other attractive features turned out to be a *Fata Morgana*, a mirage. In the beginning, the project had a strong historiographical component, which led to the national and regional committees being dominated by academic historians who perceived this as an opportunity to promote a regional and innovative research agenda. In South Africa especially this element developed rather extensively; in Namibia also we began to thrash out an ambitious programme of incorporating oral history, of filling research gaps, and on the development of educational resources. I cannot speak for the other countries, but in Namibia this programme collapsed because the human resource base was too small overall, and funding for in-depth research was inadequate. As a result, the historians rather lost their initial enthusiasm, and the programme plodded on with more success on the technical side of its implementation.

The project provided for quality scanning equipment and for training by our South African colleagues for our permanent staff (to ensure continuity, although we could not afford to assign anybody permanently to digitization), and for temporary contract staff to do the bulk of the scanning work. While we started with enthusiasm, several problems gradually became

apparent; these revealed that the programme was geared towards the donor's needs with little concern for our local conditions.

**Digital storage issues and resolution**

One problem was the required resolution. The donor requested a standard resolution of 600 dpi, even for printed material, and we got little assistance with the problems this created for us. The donor had unlimited storage space, while we struggled to accommodate the uncompressed preservation files. We were assured that we could always get back copies of their preservation files, but could we trust in that? Eventually the donor agreed to expand our local storage space, and we got a separate RAID array added to our system that would accommodate the scans for a while.

**Processing**

There were other problems. The donor had expensive proprietary software to deal with the automated conversion of the preservation files to the size and functionality needed for their attractive web presentation. This was tailor-made to their specific needs. With the software available to us, we had to do a laborious file-by-file conversion, and when we presented the problem, we just earned a shoulder shrug. We had a growing mountain of uncompressed files with not enough time to reduce them laboriously one by one to a file size suitable for access in the reading room. Eventually some local computer experts made us aware that there were several free shareware products that could do the job perfectly well with batch processing, and the problem was solved.

**Backups**

Of course one must never forget Murphy's Law, that everything that can go wrong will go wrong, at the worst possible moment. We had felt pretty safe with our server with a RAID array and its own uninterruptible power supply. When the space was almost filled to capacity, and before we had reduced all our preservation-quality scans to JPG and taken another backup, disaster struck when a lightning strike destroyed two hard disks at the same time, making the RAID array unrecoverable. Except for those scans already transferred in copy to the partners, a number of scans were lost and had to be done again.

**Metadata**

All of the problems mentioned so far could be rather easily solved with more money and expertise. The creation of metadata proved to be much more problematic, and a serious bottleneck. We were mostly dealing with archival material and with published material of a difficult nature, such as inconsistently numbered periodicals, small brochures and leaflets. We soon found that simple training in the structure of metadata in Dublin Core did not help, because interpreting the sources in a meaningful way required a level of background knowledge that the young people we employed simply did not have. Others who had some background knowledge had atrocious spelling skills and generally inadequate English language skills, so that quality control amounted to rewriting the entire entry. In the end metadata entry rested on the shoulders of one person who could not cope with the amount of work, and this therefore created a serious bottleneck.

**Arrangement**

Closely related to the issue of metadata is the proper arrangement of the material. Our Archives has substantial material about the subject focus of this particular project, which would have been most appropriate for the purposes of the project. Much of the material is however only in the preliminary stages of arrangement, due to the long-term staffing crisis. Digitizing a poorly organised collection that will probably be rearranged in the future is not only not advisable but against all archival rules, because one might in future lose the correlation between the analogue and the digital object. Either the material had to be organised beforehand, or it had to be left out. The differences between the historians, who wanted to open up sources for research, and the archivists, who are concerned about provenance and organic relations between the parts of an archive, became very obvious. We decided not to digitize insufficiently organised archives.

**Copyright**

All the problems mentioned above were essentially temporary issues that could be resolved. I assume it was the issue of intellectual property and permissions that eventually ended this project. The interaction of politics, justified concerns and emotional attitudes created a situation in which it became clear that, for a substantial section of the material that was already digitized or was to be digitized, no rights clearance to put the material on the web would be forthcoming. To put it bluntly, one cannot expect former liberation movements to authorise the use of their materials for a project based in a country that effectively obstructed their liberation for decades, even if the project is spearheaded by scholars who had been in the forefront of support of that struggle. Such subtleties are lost on politicians. This should actually have been clear to all participants from the beginning, and not wishful thinking.

**Results**

As a result there now remains from this particular project an incomplete website. It is still a very useful resource, but substantial digitized material remains hidden from view because it is not allowed to be placed online. Nevertheless, our digitization effort was not in vain, as the material may be used in-house in our local server to prevent use and photocopying of our fragile originals.

One may ask in retrospect what this project did in fact achieve, given all the problems described above. In Namibia, it achieved a great deal. It supplied us with that first-hand practical experience that one cannot gather at training courses, workshops and seminars. It made us realize our limitations and where we have to focus on improving our conditions. And apart from that, we did indeed get a considerable amount of material digitized.

**Excursion: archival web presentation**

For archives, the issue of whether to present "archivalia" (archival materials) on the Web is very controversial. Copyright is one issue but by no means the only one. Giving free access removes a measure of control, context and advice that an archives and archival staff can deliver. Putting a document on the Internet can dissociate it from its immediate context that may be vital for its understanding. For us in Africa, it is also somewhat painful to consider

making our resources freely available to researchers who do not even have to come anymore and spend some money locally to boost our economies. On the other hand, in our vast country we have a responsibility towards our own citizens who currently have to travel over 1000 km to consult their files in Windhoek. We have come up with a plan to make digitized archives available in local copies in archival access centres attached to libraries or records centres in all our 13 regions. This would also circumvent the problem of low bandwidth in rural areas.

An upside is the fascinating opportunity, for example, to load historical photographs on the Internet to solicit further information. However, we are again treading a very fine line here. One of our currently running projects is the digitization of a large photograph collection, the 1940-1970 archives of a local professional photographer. It is a joint project with the Polytechnic of Namibia and an American university, and we are putting these images in reduced quality on the Polytechnic's website. One of the ideas behind this is to invite the identification of places, occasions and people. The Archives purchased the copyright for these particular photographs, so there is no copyright problem as such, but there are other considerations forcing us to be selective. Many of the photographs which we have so far excluded from the programme are family and passport photographs that were made to order. Apart from doubts as to whether these are of interest for public presentation, a host of other legal and ethical questions arises, because when the portrayed persons ordered their photographs, they were usually not intending to make them available them for public consumption. Many of the persons concerned are still alive and while some may be thrilled, others may not be pleased at all.

Technological progress adds new dimensions to the problem. Now that we have electronic face recognition, would these persons want to supply the whole world with a name and quality passport photograph to be recognized automatically in any other picture? Of course many do this voluntarily on Facebook and elsewhere, but the choice should be left to the individual.

**What to digitize?**

Apart from the limitations of the donor-funded programme, there are of course priorities. In our case, we have a very clear priority within our paper-based archives: those records that are most heavily used and therefore endangered by the frequency of consultation and of copy requests. These are readily identified. One such resource is the German record of the German-Namibian war and genocide of 1904-1908, which is in good physical condition but heavily in demand not only for academic research but also by the general public and by the media, as this remains an emotional topic in both countries. Another priority is the collection of records of the South African government office "Native Affairs Ovamboland", which is the key source about Namibia's populous north, and is in appalling physical condition on paper of the poorest quality. Without digitization this resource would disintegrate in a matter of years. Both are large, so they are the sources that we try to smuggle into each programme, to have them finalised.

**How to sustain a programme?**

As the first project slowly petered out, we ran into problems about continuing digitization. The need was obviously there; there was plenty of urgent work as well as trained temporary staff who might not be available for future projects if they remained jobless in the meantime. Up to now we have not succeeded in adding permanent digitization staff to our establishment. So we have been looking for alternative projects to occupy them. The negative aspect of this is of course that project money is tied to certain issues and one cannot create a continuous digitization programme in line with the needs of the institution.

We received temporary funding from a biodiversity project, which resulted in the digitization of Namibian scientific journals: not quite an archival issue, but a worthy undertaking. We received temporary funding from another project relating to the liberation struggle. We struggled with payment modes and productivity issues, experiencing a sharp increase in productivity but also an increase in careless work and the need for quality control when switching to piece-work mode, with a sharp decline when switching to work- time based payment, both of which were difficult to explain to donors. Currently, our efforts are focused on getting permanent staff for this task, and certainly donor-funded programmes could be helpful for further progress.

As it is now, we are caught in a vicious circle: we have too few permanent staff to carry the digitization through, but the formulation of proposals and eventual execution of projects with donor funding also puts a heavy strain on the available human resources. Without adequate core funding and staffing, no adequate digitization can be achieved.

**How to extend a programme?**

Since we started in earnest with the digitization programme, we have been painfully aware that, apart from the two above-mentioned examples, the digitization of paper-based material is not the most pressing task. The most pressing task is to deal with our oral history sound cassettes, with which we are beginning to have problems such as tapes disengaging from hubs, and our videos which are in obsolete formats such as 1-inch tapes and U-matic cassettes. We are again dealing with these on a project basis to gain experience; the first experiences in our cooperation with the Polytechnic of Namibia and Utah Valley University were very positive and we are still waiting for disasters to happen so that we learn from them. We are however determined to build up the local expertise in this area instead of having it done for us.

We are also looking towards other cooperative programmes where local expertise is available and we do not necessarily do the scanning ourselves. The digitization of Namibian government gazettes was financed through a local NGO and carried out by a local firm, with our material because we were the only institution who could supply loose and not bound gazettes for easy automated scanning. The digitization of our maps is done at the local Geological Survey who owns the necessary equipment. This is a very slow process because it requires upgrading of an inadequate database and quite often, in addition, restoration action on the fragile material before it can be scanned. We are looking forward to cooperation with the national broadcaster and the University archives for the sound and video digitization, and have established a working committee with them. We have started a

programme to solicit the electronic printer files of publications from government offices, so that they can be converted into stable electronic versions, stored and made accessible without having to scan a paper copy.

**Perspective**

Let me end with a regional perspective. Although I think that everybody has to build their own capacity, perhaps assisted with training and advice by some centres of excellence such as DISA (who helped us in our first steps), there are a few tasks that may be too large and too complex for smaller countries like Namibia to tackle alone. One might be a backup facility that could be utilized by those who do not trust their local facilities, although some may say that could and should be in the clouds. Another task might be to create a website for marketing our digital resources. While African images, for example, are in worldwide demand, and while their commercial use when requested from European or American sources is fetching substantial fees, we in Africa continue to trade them for "an apple and an egg" (that is, very little) and are constrained by bureaucratic procedures like "treasury-approved price lists" that stay monumentally unchanged throughout inflations and regime changes. A central marketing site with automated billing and image delivery might not solve the problems of African archives, but it would contribute some sorely needed revenue. Maybe it is futile to entertain such dreams, maybe having such facilities beyond the national level would only add another level of bureaucracy to the problem, but at least one should talk about it.

**References**

Limb, Peter. (2005). The digitization of Africa. *Africa today,* vol.52, no.2, p.4-19