WHAT DETERMINES THE QUALITY OF DIGITAL COPIES OF PHOTOGRAPHIC MATERIALS? STRATEGIES TOWARDS DIGITISATION OF PHOTOGRAPHS AT THE NATIONAL DIGITAL ARCHIVES.

Małgorzata Kołtun, Mateusz Bolesta

The National Digital Archives

There are no limits how and by whom a single digital copy of a photographic image can be used. Digitisation is not the last step in the preservation process, it actually opens a variety of new paths for dissemination of photographic materials. However, these paths have great impact on the quality of digital copies. Therefore, what should be taken into consideration while digital copies of photographs are produced? What is the strategy towards digitisation of photographs and what it consists of? As one of three central archives of the state archives network in Poland, the National Digital Archives carry out their statutory activities focused mostly on digitisation and providing access to digital copies of archival records (including photographs), but at the same time the institution brings support to the other state archives within the mentioned areas. Furthermore, on one hand there is the NDA's mission as an archive, on the other hand there are different users' needs that have to be answered. This paper will define, examine and analyse different factors that determine the quality of digital copies produced at the National Digital Archives. It will consider specific role of the institution as well.

As mentioned before, The NDA have to find a balance between being one of the three central archives in Poland and Competency Center on digitisation of archival materials at the same time. And our institution's specific character is founded on this multitude of roles we are given and all the consequences this situation presupposes. The NDA were established as a response to a huge need for introducing advanced technologies to the archival field in Poland and digitisation was among them. Therefore straight from the NDA's point of origin, it was institution's duty to bring support to the other state archives in Poland within the areas mentioned, to assist them in taking the first step onto the digitisation path and to show them what is required when it comes to equipment, tools, financial structures and staff. From more practical point of view, it means we digitise archival materials from the state archives across Poland. Furthermore, we build IT systems and infrastructure aimed to collect and provide access to the information on the collections of all state archives. On the other hand, in 2009, the NDA got entrusted with the performance of the Competency Centre on digitisation of archival materials and it extended institution's mission significantly, mostly with tasks related to

education, consulting and setting up standards and good practices policies for all institutions from the GLAM sector.

Nevertheless, the National Digital Archives belong to the archival network and as one of the state archives have to fulfill their responsibilities. These responsibilities do not differ from one would consider an archive's tasks and they focus on accumulation, protection, storage and provision of archival materials. Where on this list digitisation finds its place? At the NDA we strongly believe digitisation is one of methods for preserving archival records and it is reflected in our policies and strategies and last but least, in our day-to-day work. It is actually one of main tasks. How we preserve in terms of digitisation then?

About as much as our institution's character is specific and unique, we do follow international standards. To be more precise, we produce master copies and access copies and do all our best to ensure these are of the highest quality that is required and on the other hand, possible to obtain. What is crucial at this point, is the role a digital copy (both master and access) plays or, in other words, what is the purpose of a digital copy and what does it have to do in relation to an archive's mission at the National Digital Archives? Needless to say, a master copy is supposed to preserve the analogue original. This refers to those archival materials that show any track of degradation, no matter how heavy the degradation actually is. However, the stronger destruction, the higher priority for digitisation a material is given. One has to keep in mind that it is not only the time that affects the materials' condition resulting in chemical and physical damages, but human factor could be a key issue as well and this refers to the procedures of making archival materials available for public. What does it mean to the materials themselves? To put it simple, for the most popular and desired ones it means being touched by hundreds if not thousands hands a year, including those that belong to the archival staff and also to users. Handling and access could be a major risk to the safety of documents as it is not always possible to supervise people who view the documents in the reading room. Not to mention those not so rare cases, when a particular archival material gets a status "destruct" and every single day counts in an institution's struggle to keep it in good shape. If this battle is lost, a digital copy duties as a replacement for the original.

Second of all, a digital copy of an archival material, no matter if we speak of paper documents or photographs, is supposed to give faster access to actually any kind of information featured in the original analog material. It is the information, not the material itself, that stays in users' filed of focus. How the information was reached before the introduction of digitalization to the state archives in Poland? In, as we call it nowadays, a conventional manner, meaning only in reading rooms. And it involved a list of different activities a user had to undertake. First of all, a visit to an archive was required. Second of all, a registration form with detailed personal information was obligatory to be filled in. Not to mention the call slip. Then came long hours spent in the reading room. Needless to say this was time- and money-consuming. From this perspective digital copies of archival materials open up

access to information and make this access easier if not simply possible, especially if we think of damaged materials that are not supposed to be shown to the viewers. However, it is important for both internal and external archive's users, meaning not only our visitors, but our staff members as well. For the NDA's employees, especially those who are responsible on one hand for arrangement and description of archival materials and on the other for customer service, undisturbed and smooth access to contents of archival documents is essential and this task is assigned to the digital copies.

Speaking of arrangement and description of archival materials, one has to take into consideration where along these processes a digital copy finds its place, especially if we think of photographs. When it comes to public registry records (both as original materials and already microfilmed), our policy states that only these documents are digitised that come together with inventories (and especially digital inventories), while with photographs the procedures are exactly opposite if we speak about archival fonds that are not stock taken and described yet. In this case digitisation happens first as it helps if not enables our archive to organise a photographic collection, for example a collection that has just been purchased by the NDA or donated to the NDA and has never been arranged according to any archival standards. In some cases digitalization would be first aid procedure to maintain the contents which has already become less illegible under the influence of inevitable processes like fading. Moreover, digital copies are those our professionals responsible for description choose over original materials to work with.

Last but not least, high quality digital copies help to notice details which are not detectable by human senses while examining originals, which is crucial for identification purposes. Among our staff we have those professionals who would describe themselves as archivists-detectives and that statement would not be far from the truth as their everyday work bears the mark of very detailed and time consuming investigation within many different areas even though there is no whatsoever crime. On the other hand there are millions of photographs showing unrecognized people, streets, building etc. waiting yet to be named and described. In this process a high quality digital copy lets our detectives see what would stay unseen in the original by giving all these opportunities that are brought by tools and actions like zooming, brightening and contrast enhancement. Sometimes digitalization is often the only way to save and read the vanishing image as graphic software will retrieve the hardly visible image components. Nonetheless, identification is done not only by the NDA employees but we take advantage of crowdsourcing too. But how would it be even possible to gain any profits from contributions from an on-line community without digital copies?

Another issue that determines the quality of digital copies at the NDA is strictly related to our users and their needs. Needless to say, there are many ways one can look at a user of an archive. However, regardless what methodology would be applied, it involves firstly defining groups of users, identifying their expectations and needs, listing their interests, tracking their behaviours and habits in other words – what brings a single user to our archive and what he or she looks for there, taking into consideration

both our Customer Service Office responsible for on-the-spot access and our on-line services szukajwarchiwach.pl (Search the Archives, for public registry records mostly) and audiovis.nac.gov.pl (used to present photographs). Nevertheless, seeing the fact the NDA are one of the state archives in Poland, one of our tasks and part of our mission is to give access to archival collections to public. In other words, it is the whole society we would consider our target group no matter how inaccurate one might find this concept. And within society numerous types of users and preferences can be marked. From other perspective there are two groups of users: internal and external. While the first one covers our own staff members, not only Digitisation Department, the Photographic Collection Department and IT Department, but all our employees actually, the latter refers to all institutions that deal with archiving, digitisation, data storage. Among them state archives received the highest priority on the grounds that our institution brings support to the other state archives within the mentioned areas. Nevertheless, on the other hand, there are GLAM institutions, NGOs, different associations, political parties and churches we have to think of in this case. No matter if we speak of the whole society or the external-internal schema, numerous roles can be assigned to the aforementioned users. Just to mention a producer of archival materials and digital copies of them, keeper of archival documentation and its digital copies and last but not least, simply a consumer of the content of any kind of archival materials. However, this list wouldn't be completed without a digital repository and on-line services administrator.

What do these users need? Society, in the first place, undisturbed and smooth access to the archival materials and their digital copies but it is the content and the information embedded in the files that matter the most. Things changes slightly if we move to the external users of our archive, namely state archives in Poland. With the regard of their statutory tasks and day-to-day work, firstly they have to be provided with professional standards covering the quality of digital copies, metadata and archiving procedures. However, this is directly linked to different tools used to handle the digitisation process, to create metadata, to make digital copies available on-line and to ensure the long-term preservation. In other words, from the perspective of the quality of digital copies determinants, for state archives in Poland is crucial to produce digital copies that would meet the international standards and therefore preserve the original archival materials in the best way possible, but at the same time that would be easy to deal with, accessible via any software available and on-line as well and last but not least, safely stored in a repository allowing both deep archiving using replication for backup purposes and active archiving as well. Need of other institutions, outside the archival network in Poland, are of the same kind. However, what seems to be of the highest priority for them in terms of the NDA's activity as the Competency Centre, is the high level reliability of the repository they are obliged to use to store their digital copies. And finally, when it comes to the internal user, namely all staff members of the NDA, things do not differ that much, however, the NDA being responsible for the szukajwarchiwach.pl (Search the archives) on-line service and Central Digital Repository of the State

Archives, on one hand have to secure the access and archiving processes, and on the other need digital copies that preserve the original archival materials but at the same time are easy to work with taking into consideration everyday projects, szukajwarchiwach.pl architecture, capacity of the Repository and the maintenance cost.

Another factor that determines the quality of digital copies at the NDA is related to the ways scans of photographic materials are used. Needless to say, the more professional purpose behind usage, the higher quality is required from a digital copy. Commercial market (advertising agencies, publishing companies etc.) need high resolution images, without cropping, without whatsoever alternation to the file, produced in accordance with colour management rules, meaning basically raw files, as this fulfills large-format printing standards. But if we speak of amateur or home use, requirements tend to be less exorbitant, involving some cropping, editing, low resolution. In other words, photographic images have to be first of all visually attractive and last but not least, of reasonable file sizes. However, what counts is the file format and its features like adoption and dependency as no one wants to struggle while opening an image with a particular software, using a specific operating system. Following on from that, even though professionals like academics, scholars or researchers may use a digital copy of a photograph for other purposes than these mentioned above, images should keep the same features like those intended for commercial use. High resolution together with no cropping rule allows to see all the details and information encoded in an image. However, if one want to have all the elements clearly visible even in copies of overexposure or underexposed photographs, some post-production editing has to be applied. The question is how far it should go? Nevertheless, an academic or researcher would be interested of not only in obverse, but in the reverse too as it may bare a lot of complementary information. What does it mean to an archive? To put it simple, if needed, digitisation procedure should provide for at least two copies of a photograph. But there is one thing one cannot forget about in this case, no matter if we speak of amateur or professional use – the descriptive metadata without which a digital copy's function is limited only to its visual value.

This brings us to the issue of universality of digital copies. Having in mind aforementioned types of users and their requirements, should an archive take aim in meeting everybody's needs, meaning producing digital copies of the highest quality possible but at the same easily to work with? To go further, is it even possible? Nevertheless, taking into consideration long-term preservation goals, what is within our field of focus, is not only the present variety of users and usage, but future as well. Currently our images are used for different commercial and non-commercial initiatives: as prints to illustrate press articles, for exhibitions, in digital form on Internet, in scientific publications including monographs and academic papers as well as publications from the entertainment area. However, we should keep our minds open and not limit ourselves to what is known and performed nowadays. Moreover, it is important to take into account all legal issues related to reuse of information policies and free access to the archival records as well. In other words, we try to ensure our digital copies of

photographs would be used as extensively as possible and the most obvious way to accomplish this seems to be the one referring to the quality of digital copies and digitisation procedures.

The last aspect to determine the quality of digital copies of photographs at the NDA is related to the nature of work on one hand at the NDA and in all state archives in Poland on the other. All state archives, including the NDA, digitise according to annual plans, meaning we are obliged to produce specific number of scans per year. The number of scans we are able to achieve is linked up with their quality. In what sense? The more advanced and complex techniques and procedures we apply in order to obtain digital copies of high quality, the more time-consuming the digitisation process becomes. Moreover, The National Digital Archives‘ collection consists of 15 million photographs dated from the 1840s to present, while all state archives in Poland hold little over 1 million of photographic archival records. In addition to this, a typical state archive collects and preserves mostly public registry records and if one can find any photographs among them, they would be part of those paper documentation. It results in different digitisation procedures an archive has to apply in order to produce high quality copies. To put it simple, it is easier to deal with an archival fond that consists of photographic documentation only.

About as much as social factor impacts on the quality of digital copies, so does technology. It may be difficult to establish which one is of greater importance, for both are strictly intertwined. Simply put, the kind of appliances we utilise to produce digital copies of photographic documentation enables multipurpose usage of the content by its recipients. However, if one is to consider the issue of scan quality from a technological perspective, one must take into account both boosts and constraints to this attribute of digitisation output. The aim to produce state-of-the-art digital images by cutting-edge equipment, especially in terms of mass production and storage capacity, may come out as slightly out of touch with reality, because often same technological abilities which make up for high quality of scans also curb the efficiency of their processing and archiving. Take ppi resolution for instance, the higher this attribute the better reproduction of details of the original on digital image, but simultaneously the longer scanning process and larger file size. The same rule applies to significant other parameters of a digital copy. Furthermore, even most capable devices will not eliminate certain quality restraints, which are inherently connected to the characteristics of the original source, for example low resolving power of photo-sensitive material, its condition or the flaws committed by the initial author which, by the way, also have to be reflected in the digital image.

In fact, in order to attain the quality of the image, which at the same time would satisfy the needs of consumers, meet our storage capacity and keep up with current standards of digitisation we have to balance between striving for the highest parameters and remaining within the boundaries of spatial availability for saving the data, as well as limited human resources, tight workflow and many other technological variables, which will be discussed further.

Before we proceed to display numerous aspects of digital image quality, we ought to determine

what do we understand as “quality of a digital copy” and what are its constituents. In regards to digitisation in archives there are two main aspects of digital copy quality. First off, digital image has to contain all the information about physical characteristics of a source document, for example its format or the material used in its production. Secondly, it has to capture the content of the original as accurately as it is possible, absent minimal margin of error due to technological limitation. When it comes to photography, all middle tones, colours, contrast and other details must be reproduced on the image.

This introduction barely scratches the surface of the issue. Overall the quality of a digital copy emerges from combined condition and properties of original material, specification and performance of scanning appliances, the extent to which an image is processed during post-production stage, digitisation standards established by The Head Office of the State Archives in Warsaw, both the number and qualifications of scanner operators, storage capacity of Central Digital Repository of the State Archives and last, but not least the ambition to satisfy the needs of even most demanding users, whilst maintaining perfect resemblance between the original and its digital mapping. In this presentation we will describe the criteria and determinants of digital image quality identified during digitisation practice at National Digital Archives, as well as pinpoint certain nuances to this issue that may cause great deal of trouble in the pursuit after an image as perfect as possible.

As strange as it may seem, the quality of digital master copy sometimes begins long before scanning, right at the restoration stage. Many of the photographs, especially on glass medium from 20's and 30's require certain restoration procedure, be it surface cleansing or putting all the broken pieces together. Both degree and quality of restoration processes determine the outcome of digital resemblance of the source object. At National Digital Archives we tend to know what kind of material in terms of its condition we are going to digitise beforehand, that is why we implemented a division of archival documentation on four groups on the basis of its state. Group 0 – nearly perfect condition, fit to scan immediately. Group 1 – Good condition, however requires minor restoration procedures before scanning. This is where much our material is at. Group 2 – bad condition, requires more advanced and extensive restoration labour. Digitisation is put on halt. Group 3 – Very bad condition, the material is on the verge of destruction. Requires long and very complexed restoration. Digitisation is completely out of the question.

First of all when it comes to scanning, there is a number of essential technical parameters, which are eligible for an accurate image quality assessment. Those variables are ppi resolution, colour depth and colour space profile, file format and compression. These parameters are precisely defined within the appendix I to the Ordinance no. 14 of the General Director of State Archives, which constitutes a framework for digitisation practice in all Polish state archives. If we are to embrace the entirety of digital image quality issue, we ought to start with at least brief examination of what this Ordinance has to say about those attributes.

First on the list is ppi resolution. This topic is pretty tricky, because setting this parameter to the highest value does not necessarily do all the job. Of course, higher resolution allows for display of greater amount of detail, however only to a certain extent. If the threshold value of resolution necessary to capture all the details is exceeded, it will only cause the file size to increase without a significant change to the quality, resulting with an unwanted storage space reduction, along with difficult and time consuming reading process.

In terms of digitisation of photographs there are two factors which determine resolution. This parameter should foremost correspond with the format of original documentation so the quantity of pixels would suffice to convey all details. Also we must take into consideration the resolving power of the photo-sensitive material as to avoid exposing too much grain.

At the National Digital Archives the smallest-size material of up to 24 mm is scanned in 3000 ppi, 25-60 mm in 2400 ppi, whereas larger formats such as 61-130 mm and above 130 mm are digitised in 1200 ppi and 600 ppi respectively. Of course these values pertain to master copies only. The idea of ppi resolution increasing accordingly to decreasing format is to have the images preserved in high quality, so that even small pictures would be available for further processing. Another outcome is that the images contain the same amount of pixels on longer edge despite distinguished sizes of source objects. As a result the output is of the same size in MB, yet it is possible to calculate the factual format of an original from each digital copy.

It is absolutely crucial to mark that we only pay attention to scanner's physical optical resolution, that is maximal resolution a device will perform without interpolation, meaning creating artificial pixels on the basis of adjacent ones, which is obviously a bias and does not do much good to the quality of given image.

In 2013 on behalf of the National Digital Archives my superior and co-speaker Małgorzata Kołtun had conducted tests and research aimed at creation of criteria to select a file format to which all state archives in Poland would scan their collections. Those criteria were divided into six categories including procedures and infrastructure (information on the digitisation process, metadata implementation, Central Digital Repository maintenance), digitised material (information on the material eligible to digitisation and manners of its application), feedback from the state archives (especially related to available infrastructure), file saving characteristics (technical information), experience of other institutions in utilizing given file format (*i.e.* Library of Congress of the United States, The National Archives of United Kingdom and Bavarian State Library), economical implications (pertaining to costs of implementation of given file format). During the course of testing stage different file formats, such as PNG, BMP, JPEG and JPEG2000 were analysed and compared in accordance with aforementioned requirements with TIFF 6.0 as to check if it could be replaced by other data record format for master copy production. File formats were tested on two kinds of documentation – a manuscript and a typescript – digitised on Book2net scanner in 300 ppi. Our main concern was the

assessment of quality of an image compared to size of each file format.

Out of this scrutiny came following conclusion. If there is anything about file format that affects the quality of the image it is compression. Files which utilise lossy compression, such as JPEG received smallest size at the expense of compromising or otherwise decreasing the quality of digital image. Had our scans either lossless compression or none their size was larger, such as JPEG2000, PNG and BMP, but the quality of an image was clearly superior.

The National Digital Archives recommended maintaining TIFF 6.0. usage for the production of master copies, simply because our strategic plan for developing Central Digital Repository between 2013 and 2042 already relies on TIFF 6.0. digital output. The NDA also allowed TIFF 6.0. with lossless compression, however only LZW algorithm. As far as the quality and size of TIFF 6.0. image is concerned it ensures high standard of digital copy, while keeping medium size between JPEG2000 and PNG on one side and BMP on the other.

Colour depth is an essential parameter of a digital image. If we choose larger number of bits to save information about colours, the wider colour palette we will achieve on the image, hence its better quality. Generally, 24-bit colour depth, that is 8 bits per each channel of RGB colour space, is sufficient to attain the kind of colour quality that bares resemblance to natural colours, but ever since the majority of our collection are photographs, not to mention very beautiful ones, at National Digital Archives we do not want to settle for less and go for 48-bit colour depth, that is 16 bits per channel of RGB colour space for colour photographs and 16-bit for greyscale.

What are the differences between 8 bits per channel and 16 bits per channel? From a quantitative point of view file size increases along with colour depth. An image scanned at 3000 ppi in 48-bit RGB amounts to approximately 153 MB. This affects heavily the storage capacity in terms of mass digitisation. From a qualitative standpoint instead, if we pick 8-bit colour depth, each of three RGB constituents of RGB will receive 2^8 (because the colour is saved in 2 bytes of memory) or 256 values (0 to 256), whereas with 16-bit colour depth the number of values raises to 2^{16} or 65536 values per each constituent channel. That is 250 times more information about colour. If 256 values falls for each colour channel in RGB colour space in 24-bit colour depth, then we gain the combination of approximately 16,7 million colours.

This dissonance is also visible at post-production stage. If we open an 8-bit image in, say Adobe Photoshop, and apply auto-levels (the histogram will be “stretched”, as a result of adjusting the image to new points of white and black), then again pick levels, we will notice that once black histogram turns somewhat striped thereafter. That is the sign of missing values, which do not apply to any pixels, since 8-bit colour depth only offers 256 values.

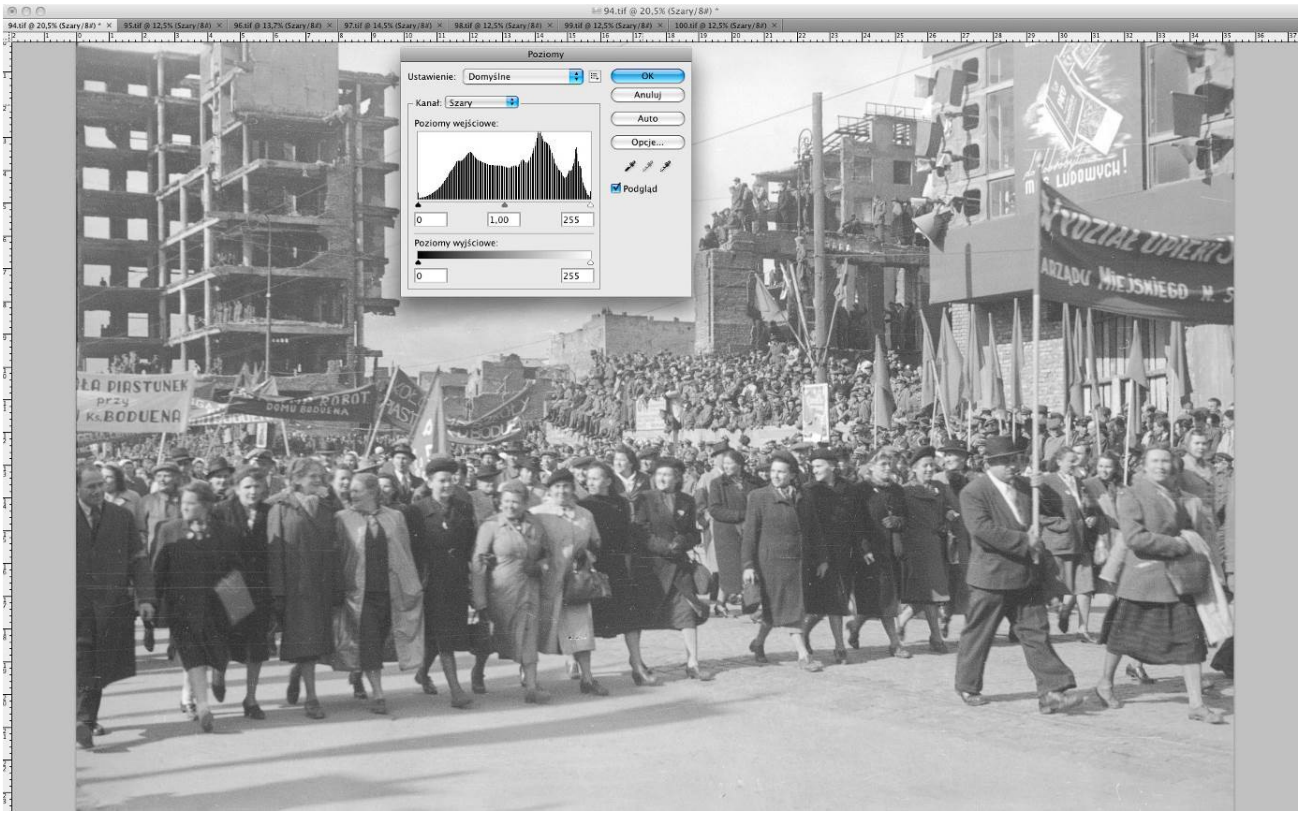

Illustration 1. An image with 8-bit colour depth

The choice of colour depth is crucial to capture entire content of the original, as well as to display it's physical properties, both of which deliver very important information for both the user and archivist. It is however important to stress that efficient processing of 16-bit images needs large computing power and even still takes a lot of time, so if a particular institution has the time to spare and can afford a fast computer, it will not be a problem, however if another organizations deals with mass production of images, daily average quantity of copies and annual work schedule to complete, it may want to consider scanning to inferior 8-bit colour depth, of course at the expense of quality.

An issue which is closely connected to colour depth is colour space. At the National Digital Archives we prefer Gray Gamma 2.2 for grayscale and Adobe RGB 1998 for colour photography. If a digital image is to be used in a book or included to an exhibition, it has to be printed. This is where a wide colour space plays a significant role, for example Adobe RGB 1998 is wider at colour conversion to CMYK model than regular sRGB and still accurately reproduced by printers adding much to its utility.

If user does not consider printing an image, nevertheless it is important to scan with a wide colour space for a better display on the monitor screen, and again Adobe RGB 1998 shows better results than standard sRGB in dimension to viewing images. Henceforth the quality of a digital copy relies on its display and printing possibilities. The wider the colour space, the higher quality, but there is more to

this issue than we already covered.

In fact colour space opens up a vast scope of subjects pertaining to both software and hardware, which are essential for the quality of digital copies and very closely related. This set of issues is known also as colour management.

Another colour space we use is CIE LAB, that helps us to calculate Delta E parameter. At the NDA we measure the values of L*a*b* of three colours only – black, white and grey using Adobe Photoshop and correlate them with their respective reference values. Next we count the Delta E parameter for each of these colours using our Delta E calculator and then we establish the value of mean Delta E. This operation pertains more to textual records rather than photography, because we scan photographic documentation without colour chart. If we did otherwise, we would have to scan every single photography with such chart separately time and time again, whereas we have 15 million pictures and the overall process would take tremendous amount of working hours.

If for example Delta E is between 0 and 1 there is no bias to colour mapping, if its value is between 1 and 2, then only a skilled observer might catch the difference, but if the value of Delta E crosses 3,5-5 threshold there is an apparent discord between original documentation and its digital mapping.

In order to settle the communication between distinguished appliances we constantly operate with ICC profiles. Since ICC profiles are generated during calibration process and profiling it is of major importance that each device be it a scanner, a camera or a monitor had its own profile, to which we pay special attention when it comes to quality. Adding a profile to an input and output device means more control over image colour and lesser possibility to receive a digital copy with apparent colour flaws, that causes time and money losses. For the same reason each digital document must have its own profile, which gives us complete information about colour in this particular file, however in this case such profile is added automatically by the scanner.

Although colour management cannot solve all the problems it adds much to the quality of an image since it harmonizes and unifies colour capture at the scanning stage and colour display on the monitor screen. It also ensures that whatever correction had been made on the input, it is visible on the output.

The topic of colour management leads inevitably to the issue of equipment. There are plenty of scanners and monitors out there, loaded with parameters and armed with marketing cliches, some of which are pretty inaccurate when compared to actual performance, but fortunately we dispose of our Ordinance that gives us precise hints to what are key requirements that digitisation appliances have to meet.

Foremost, since we are still in the subject of colour management, both scanning devices and monitors must have wide gamut, in other words have to be able to display, reproduce or otherwise register a wide range of colours. The wider the gamut, the more colours from a particular colour space

will be visible on the screen and mapped by scanner.

Those same devices must allow for their calibration and profiling. Calibration itself is more or less automatic process, which utilises autonomous control system of the equipment, but may also require external devices, such as colourimeter or spectrophotometer in case of graphic monitors. We prefer spectrophotometers over colourimeters because of their superior precision of colour temperature measurement.

It is also great news if manufacturer provides compatibility with ICM profiles. ICM stands for image colour matching and contains information revolving around differences in colour reproduction and used colour space of a particular device. This data provides foundation for effective colour management.

Thereafter, we have CRI, which stands for Colour Rendering Index. It is estimated that the value of CRI which amounts to 90-100 provides for very good colour mapping, therefore only devices that fit to such range are eligible for usage in digitisation.

Furthermore, the temperature of fluorescent lamps mounted to the scanner should vary between 5000K-6500K and not exceed that range. They should also emit constant light, not flash. This relates more to the safety of archival material, but gains our great concern.

Finally, there are numerous other factors that determine both the quality of the digitising hardware and of digital image, like optic density, depth of field, scanning speed and the ability to record technical metadata, which also contain a lot of important information. Especially when it comes to photography optic density is something to look out for. It is D parameter equal to decimal logarithm of the proportion between incident light and either diffuse or transmitted light, depending on the device. D parameter is very important to the quality of an image. If we scan a very dark slide on a scanner with optic density below 3.6 threshold all darker tones will blend into one black mush and even Photoshop will not help. If we scan the same slide but on a device with, say 4.4 value of optic density, then out of once black stain new details will emerge. So that is the increase of image quality right there! In fact, different threshold values of optical density are proposed for different types of material: 2 for photographic prints and other reflective material, 3 for negatives, 3.6 to 4.1 for slides and other forms of transparent material.

These rules translate into equipment we use to mass produce our master copies. First of all, we exploit three flatbed Kodak/Creo iQsmart^3 flatbed scanners, which are meant specifically to scan photographic documentation. Their maximum resolution amounts to 10000 ppi, optical density (Dmax) to 4.2 and colour depth is 16-bit for grayscale and RGB. These parameters are high and gravely contribute to digital image quality. Another thing that adds to their formidable performance is their universal application, as they are able to scan both transparent and reflective objects regardless of their shape, size and material, be it glass 18x24 cm negative, 35 mm cellulose film or positive image on a paper. They utilise two separate lamps and the diagonal of their glass surface is 13 inches,

allowing us to put as many photographs as it is possible for each scanning session, for example we are able to scan at least 80 frames of 35 mm film without a pause to change the material. They also calibrate automatically and embed colour space profile to each scan, an ability we also factor in considering digital image quality. The drawback of these devices is however scanning speed, as striving for better quality always goes at the expense of time and we maintain a production norm of 50-60 master copies daily.

Largest formats of negatives, namely 13x18 cm and 18x24 cm, are currently undergoing tests on PENTACON SCAN 7000 scanner camera with maximum resolution reaching 400000000 pixels, that is nearly half gigapixel. It has a mounted apochromatic Schneider-Kreuznach 50 mm fixed lens producing very sharp digital images and is attached to Kaiser reprographics column with a light box of approximately 5500K lamp temperature. We chose only the largest formats we have in our collection, because the camera would not scan smaller objects as precisely due to limited possibility of lowering it on the reprographics column.

The scanner camera itself is a more advanced device, allowing greater control over the quality of master copies. First, it has a mounted lens, which sharpness has to be adjusted manually, as it is not equipped with auto-focus. Second, it has an aperture of 6 values or f-stops: 4, 5.6, 8, 11, 16 and 22. Apart from software options of scan processing, these are essential determinants of the image quality in this kind of appliance. If the lens is out of focus, then the image will be blurry. If one mistakes the aperture value by one or two f-stops, again, image will be either underexposed or overexposed. Also, we only use the scanner camera in a photographic darkroom to avoid any unnecessary light sources.

We have also reactivated our multi-format and multi-purpose film scanner Nikon Super CoolScan 9000 ED with 4000 ppi maximum resolution, 4.8 optical density and 16-bit colour depth which is being used to digitise 35 mm frames that appeared covered with Newton's rings on both aforementioned devices.

Although scanner manufacturers provide software enabling image edition at the pre-scan stage, we prefer to process our scans in Adobe Photoshop, which contains better algorithms, be it for levels, curves, auto-contrast or auto-levels. These are the only acceptable options we apply, whilst avoiding retouch and image bias, but the better these functions are programmed, the higher image quality we receive resting assure we did not damage its original content. As far as post-production goes, along with aforementioned operations, we only crop the area of the picture straight to its edges, as it is impossible to do it otherwise with 35 mm films, which are abundant in our collection. The only exception are glass negatives, which are cropped with a small 1 mm black frame around it. This decision came after discussions with archivists working in Photographic Collection Department. We have concluded that cropping an image with such a frame gives more information about the medium. Extremely rarely we do anything else, because any further attempts at image edition would be subject to arguments, as everyone has his own notion of good picture, and we rather give raw images, with

minimal correction, like auto-levels and auto-contrast, to our users, so that each individual can apply his own changes to the image the way he wants it. Often times we simply have to forget whether the picture on the screen “looks nice” and focus only on the accuracy of how it resembles the source object.

A scanner operator's worksite comprises also of graphic monitors for colour management. In our case those are EIZO ColourEdge series monitors CG301W and CG241W, which have a wide colour gamut of 98% reproduction of Adobe RGB colour space. Due to their 16-bit internal processing they produce nearly true black, whereas all grayscale tones are better differentiated from one another in darker areas. Both are identical in terms of specs and are distinguished only in size. Both also allow for calibration, profiling and even post-calibration colour adjustment, which also contributes to the quality so that not only is the colour space chosen properly, but also it is correctly registered and reproduced on the screen.

In the final part of this presentation I would like to tackle the daily basis at National Digital Archives, our workflow, the kind of material we digitise and appliances we use, the difficulties we stumbled upon our practice and discoveries we made during the research part of our work. As you will see all these aspects come together in the framework of our work.

The research we have conducted, especially in most recent time, revolved around one major issue of photography digitisation, that is the accuracy of source document resemblance on the digital copy. While preparing for the completion of our project for digitisation *i.e.* glass negatives from 1914 up to mid-40's which once belonged to upper Silesian Office of Photographic Documentation *Landesbildstelle Oberschlesien*, we were looking for the kind of equipment that would fit to scan 18x24 cm negatives, while matching with its parameters to the devices we employ on our daily basis, that is three Kodak/Creo iQsmart^3 flatbed photography scanners. Our reference model was a scan taken from iQsmart^3 and testing samples were the scans from PENTACON SCAN7000 scanner camera and EPSON Perfection V850 Pro flatbed film scanner. I have brought an excerpt from our tests and as you can see the results might be a little confusing, because in fact all scanners produced nearly entirely different digital images with distinctive contrasts, tones, highlights and shades, as well as sharpness thereby making us ponder which one of them we ought to trust.

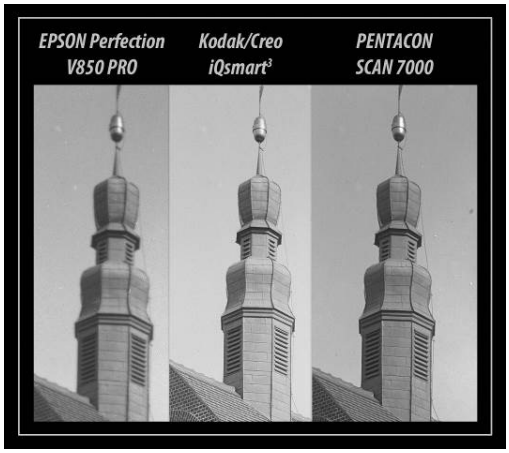

Illustration 2. Comparison between digital copies produced with flatbed scanners and scanner camera.

The same issue occurred when we were searching, under the same project framework, for a method of avoiding Newton's rings, that appeared on 35 mm film negatives belonging to Stefan Rassalski's Photography Archive Fund, a collection of photographs from 1939 to 1948. These films were cut into single frames, which, over the course of time and due to improper storage conditions, bent to a shape of cradle or semicircle so that the structure of the material changed permanently, without any way to unfold it once and for all. So, had we scanned these photographs on our Kodak scanners and PENTACON scanner camera, we received images, large parts of which were covered in Newton's rings. Pressing these frames with special optiwhite glass as well as medical glass further flattened them, but the quantity of Newton's rings decreased by minimum, on the contrary to large discords between samples in dimension especially to sharpness and tones. The results of these tests are displayed on my next graph. All samples were scanned in 600 ppi, 16-bit greyscale.

The quality of these samples is rather poor, primarily of ones made on scanner camera, regardless of additional 150% sharpen mask, with either medical glass or optiwhite glass, which could not flatten the frame completely. IQsmart^3 performed much better with medical glass, maintaining proper sharpness, yet it did not remove Newton's rings. We then incorporated slide mounts with anti-newton glasses and further tested them on PENTACON camera and Kodaks, but in each case images were yet again blurry and their quality did not satisfy us. Also the mounts cut the image on the edges, making us loose important details. The solution came with our once taken out of service film scanner Nikon Super CoolScan 9000 ED which comes equipped with its own anti-newton film holder for 4 frames, so instead of putting each frame to each slide mount, we were able to batch scan them saving a lot of time and effort, whilst eliminating Newton's rings. On the comparison of samples retrieved from Nikon and Kodak we can see that the first one generated an image with bigger contrast, better detail reproduction and better mid-tones than the second. It is also less pale and appears more vivid.

Illustration 3. Comparison between Nikon Super CoolScan 9000 ED film scanner and Kodak/Creo iQsmart3 flatbed scanner.

The conclusion is that there is a variety of methods to deal with Newton's rings, like anti-newton spray or anti-newton slide mounts, and each affects image quality in its own distinctive way. In the very same fashion the specifications, performance, software and components of each individual scanner come together to build different notion of image quality.

Unfortunately, as soon as we dealt with Newton's rings another problem occurred. This time, after making a preview of the image on default settings we noticed a very bizarre image – blank, pale, completely unnatural, somewhat polarized, as if the software could not recognize the characteristics of the negative. This might also be the sign of the loss of silver halides, probably due to poor film development. This effect would be dangerous and calling for immediate restoration in order to save the image. We repeated the procedure but this time scanning negative as positive, in the same fashion as our Kodak scanners operate in Digital Transparency mode which I will discuss next. We then inverted the image with default invert command and applied additional automatic tonal adjustment in Adobe Photoshop only to receive a perfectly normal image. I will display the comparison on the graph.

Illustration 4. Scanning negative as positive and inverting the image with Adobe PhotoShop.

Another thing with Nikon Super CoolScan is that it embeds its own colour space profile NikonGrayG2.2 by default and we cannot change it in its software. The only way to solve this problem is to assign proper colour space profile in Adobe Photoshop.

Kodak iQsmart[3] scanner has yet another useful scanning mode called Digital Transparency. Scanning an object into a DT File allows us to capture entirety of information the scanner is capable of recognizing. Furthermore DT File is in fact TIFF 6.0. file, which is available for further processing in Photoshop or any other graphic software compatible with this data record format. We use Digital Transparency mainly for digitising transparent negative material, which is either in bad condition, is heavily overexposed or otherwise scanned as blank, white image in regular mode. In Digital Transparency mode we simply scan such negative the way human eye sees it, negative as positive, if you will. Thereafter, we invert it in Adobe Photoshop using a special LS Curves preset that inverts the image over the parabola of histogram, instead of default diagonal invert. Such operation reveals a lot of detail, which would remain otherwise hidden in the brightest areas of the image and by that increases the overall quality of the master copy. The preset itself leaves image quite dark after

inversion, so additional automatic tonal adjustment must be applied, but the outcome is simultaneously more attractive from a visual standpoint, with darker tones and sharper contrast.

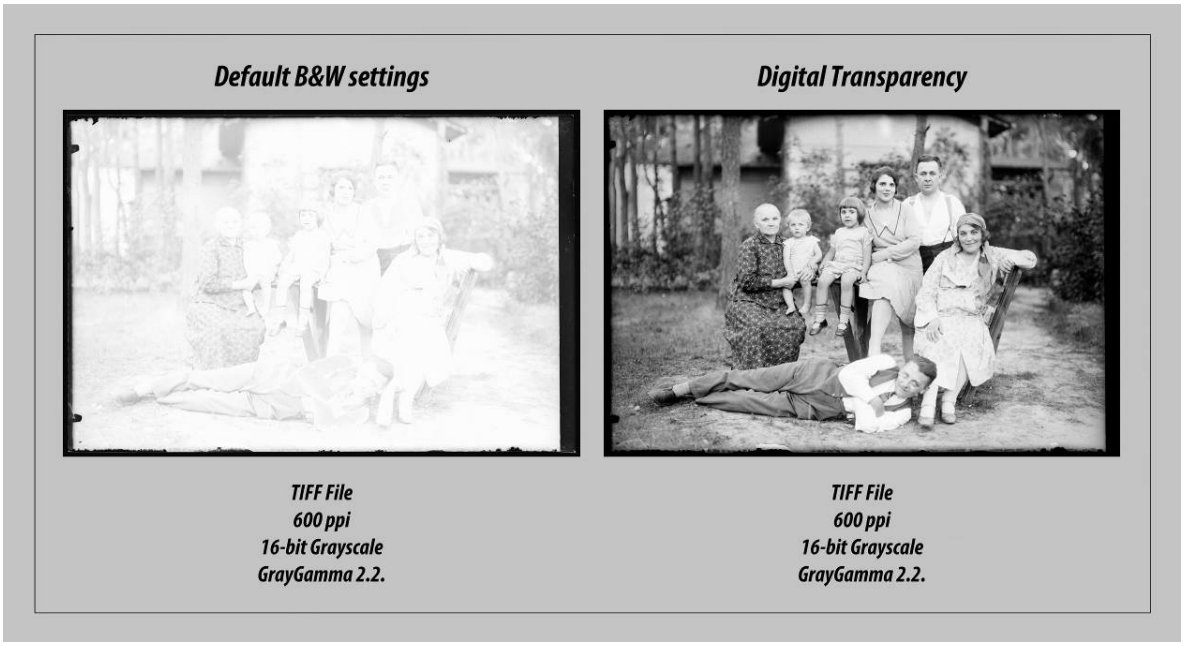

Illustration 5.Default B&W settings vs. Digital Transparency with Kodak iQsmart3.

One more important thing about our Kodak scanners, since we use them as kind of universal devices for all sorts of photographic documentation we store in our archives, are scratches on the glass surface created whilst digitising glass material. These scratches, however small they may seem to human eye, are also scanned and visible on digital images where there are no defects on the original whatsoever. Therefore it is better for us to scan glass negatives only on one scanner, leaving other two safe or scanning glass negatives with scanner camera, but only from a certain format threshold, like in our case 13x18 cm and upwards. Of course, we also change scanner glasses, however, since these scanners are elderly and parts are becoming increasingly rare, this is not done on a frequent basis. Speaking of parts, overall device exploitation affects image quality to a large degree, this applies especially to fluorescent lamps, which burn out after certain time, decreasing the sharpness, contrast, colours and tones, so the image does not appear like the source object at all. The amortization of the equipment is on the other hand quite costly, since there is only one company in Poland, which has an access to spare parts to our scanners. Therefore we have to constantly balance between highest quality possible and expenditures connected with scanner maintenance.

When it comes to scanning output we cannot forget about quality control, which is essential to every digitisation process. It is even more crucial in terms of mass production of master copies, amounting to in our case 37.5 thousand annually. In case of every single digital image from that

number this quality check has to be detailed, meaning each copy has to be compared with source photographic documentation if it is cropped properly and captures all the information visible on the analog original. Furthermore it has to be checked according to technical parameters, especially ppi resolution, colour depth and colour space profile, also we assess proper mapping of colours, tones and contrast on the digital image. Speaking shortly, we basically control copy after copy after copy and if we detect certain flaws, we simply scan such photograph again.

The issue of storage capacity had been mentioned at least a couple of times during this presentation, so it would be suitable to spare some time and shed some light on our experience in this field. Central Digital Repository of the State Archives is managed by the National Digital Archives and is aimed at maintaining the digital version of the entirety of archival material digitised in all state archives in Poland. It was designed to keep deeply high quality master copies of archival collections and to store shallowly access copies which are made available to the outside users. Although Central Digital Repository still operates within that framework, various occurrences make us reach for the master copies time and time again, so it is not a completely deep archive. It is also undergoing regular development in order to meet growing needs of the State Archives, as it is used by entire archival network in Poland. What is constantly being revealed over the course of our work is that the storage space is always on the negative part of our balance sheet. The ever-growing number of high quality master copies seems to outpace the extent of Central Digital Repository development, which is mostly curbed by budgetary restraints. This is yet another case where we have to struggle between high quality of digital data and complications to its production, but this one is of major importance, because in the worst case scenario as soon as the state archives will run out of storage space the whole production is going to stop.

In a brief aftermath, there is a tremendous amount of technological details related to image quality, which we have to have under control. This is why the human factor is most important to keep up with all such nuances and maintain proper quality of digital image. By due diligence of the staff, that comprises of qualifications, meticulousness and most appreciated eagerness to conduct research, as well as to cumulate knowledge we were able to elaborate standards that meet our needs, give direction to other polish state archives and fulfill expectations of our clients.